OVERVIEW PAPER

Behavior signal processing for vehicle applications

CHIYOMI MIYAJIMA[1], PONGTEP ANGKITITRAKUL[1] AND KAZUYA TAKEDA[1]

Within the past decade, analyzing and modeling human behavior by processing large amounts of collected data has become an active research field in the area of human–machine interaction. The research community is striving to find principled ways to explain and represent important behavioral characteristics of humans, with the goal of developing more efficient and more effective cooperative interactions between humans, machines, and environment. This paper provides a summary of the progress we have achieved to date in our study, which has focused specifically on interactions between driver, vehicle, and driving environment. First, we describe the method of data collection used to develop our on-the-road driving data corpus. We then provide an overview of the data-driven, signal processing approaches we used to analyze and model driver behavior for a wide range of practical vehicle applications. Next, we perform experimental validation by observing the actual driving behavior of groups of real drivers. In particular, the vehicle applications of our research include driver identification, behavior prediction related to car following and lane changing, detection of emotional frustration, and improving driving safety through driver coaching. We hope this paper will provide some insight to researchers with an 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 [1–3]. In regards to driving vehicles, as in driver–vehicle–environment interaction (Fig. 1), human errors contribute to more than 90% of fatal traffic accidents [4]. Understanding driver behavior can be useful in preventing traffic collisions [5], as well as enhancing the effectiveness of interactions between drivers, vehicles, and the environment. The study of driver behavior is a very challenging task due to its stochastic nature and the high degree of inter- and intra-driver variability. To cope with these issues, over recent decades data-centric approaches have gained much attention in the research community [2, 6–8]. Large amounts of data are being used to approximate parameters of models in order to optimize the performance of systems. However, at present, there are no standard methodologies for processing this data and using it to represent human behavior. In this article, we focus on understanding human behavior from a signal processing perspective, and on developing methodologies to analyze and model the meaningful behavioral information that is extracted.

The first step towards analyzing and modeling driver behavior is to collect a reasonable amount of realistic, multi-modal observations. Here, observations, in the form of driving signals, represent behavioral variables as a time series which possesses particular dimensions of behavioral characteristics. With recent advances in sensing and computing technologies, it is now practical to acquire large amounts of real-world driving signals using an instrumented vehicle, and to store all of this data. We took extra care in designing and developing our instrumented vehicle so that we could collect a broad range of driving signals which would represent relevant information regarding the driver, the vehicle, and the driving environment. Therefore, this paper will begin with a description of the data collection process and the driving data corpus, which is one of the largest on-the-road driving corpora in existence, with data collected from more than 550 participants at the time this article was written [9]. Subsequently, we will discuss behavior signal processing and modeling methods with practical vehicle applications.

Driving behavior involves multiple layers of information regarding a driver, ranging from short-term characteristics such as a driver's mental, physical, and cognitive states, to longer-term characteristics such as goals, personality, and driver identity. Hence, modeling driver behavior allows us to detect or predict behaviors of interest during vehicle operation, as well as to assess and improve driver behavior

[1]Graduate School of Information Science, Department of Media Science, Nagoya University, Nagoya, Japan

Corresponding author: Pongtep Angkititrakul
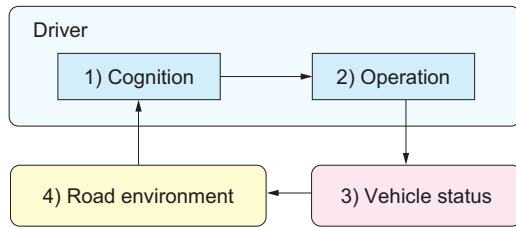Email: pongtep@g.sp.m.is.nagoya-u.ac.jp

Fig. 1. Recursive relationship between driver, vehicle, and environment.

Fig. 2. Instrumented vehicle.

after vehicle operation. In this paper, we demonstrate that driver-behavior models obtained from driving signals can be used to identify drivers, predict driving behavior, detect driver frustration, and assess recorded driving behavior.

In the course of our research, we have developed driver-behavior models based on a probabilistic Gaussian mixture model (GMM) framework. We first applied our GMM-based driver model to capture relationships among the related parameters of car-following behavior. We showed that GMM-based car-following models representing patterns of pedal operation in the cepstral 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 vehicle operation behavior, in terms of pedal-pressure, 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 an individual's driving style. The experimental results showed that the framework could achieve a prediction performance of 19 dB signal-to-deviation ratio (SDR). In order to model vehicle trajectory during lane changes, we developed a hidden Markov model (HMM)-based model to capture the dynamic movement of vehicles. In conjunction with a proposed hazard map of surrounding vehicles, we were able to generate predicted vehicle trajectories during lane changes under given traffic conditions with a prediction error of 17.6 m.

In addition, by employing a Bayesian network (BN), driver frustration could be detected at a true-positive (TP) rate of 80% with only a 9% false-positive (FP) rate. Finally, using a system that automatically detects hazardous situations, we developed a web-based interface which drivers can use to locate and review each hazardous situation which occurred during their own recorded driving sessions. The system also provides drivers with feedback on each risky driving behavior that was detected, and coaches drivers on how to appropriately respond to such situations in a safer manner. Experimental evaluation showed that driving behavior could be improved significantly when drivers used the proposed system. In this paper, we will describe in detail the signal processing approaches we employed for collecting, analyzing, modeling, and assessing human behavior signals, and demonstrate the advantages of using these techniques in vehicle applications.

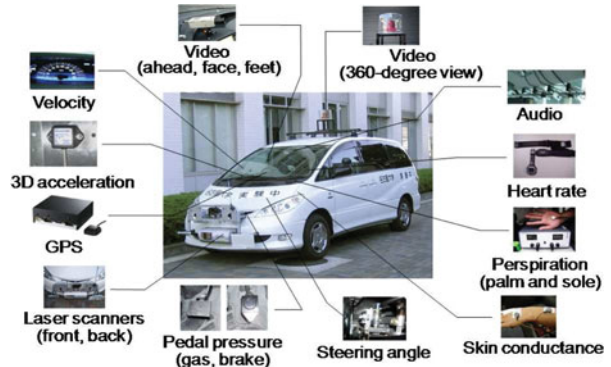This paper is organized as follows. We first introduce our driving signal corpus and data collection methods in the next section. In Section III, we describe our first application of driver modeling for driver identification. Then, in Section IV, driver behavior prediction is discussed for both car-following and lane-changing tasks. Section V describes our analysis of driver frustration using a combination of speech and pedal actuation signals. An application involving the use of recorded driving data for driver education is discussed in Section VI. Finally, we summarize and discuss future work in Section VII.

II. DRIVING CORPUS

A data collection vehicle was designed for synchronously recording audio and other multimedia data as drivers operated the vehicle on public roadways [9, 10]. Various sensors were mounted on a Toyota Hybrid Estima with a 2360 cc displacement engine and an automatic transmission, as shown in Fig. 2. Table 1 summarizes all the driving signals recorded by the system. Participants drove the instrumented vehicle on city streets and expressways in the city of Nagoya, Japan. The data collection vehicle, route, and equipment were the same for all drivers, and all the drivers were trained and treated in the same manner. During the experiment, drivers also performed carefully designed secondary tasks similar to activities likely to occur during everyday driving. The secondary tasks were carried out in the same order at very similar locations, so that data from different drivers could be analyzed and compared. Throughout each data collection session, an experimenter monitored the experiment from the rear seat.

A) Collection protocol

In order to develop a technique for quantifying the cognitive state/stress level of drivers, driving data are recorded under various driving conditions with four different secondary tasks. Detailed descriptions of the tasks and examples of the spoken sentences are follow as:

- 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", "Dennys".
- Navigation dialog task: Drivers are guided to an unfamiliar location by a navigator via a cell phone with a

Table 1. Summary 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 of 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 leading 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	6 × 15 fps	Images	Driving environment
GPS	Navicom GPS-M1zz	1	1 Hz	Standard	GPS information

hands-free headset. Drivers do not have maps, and only the navigator knows the route to the destination.
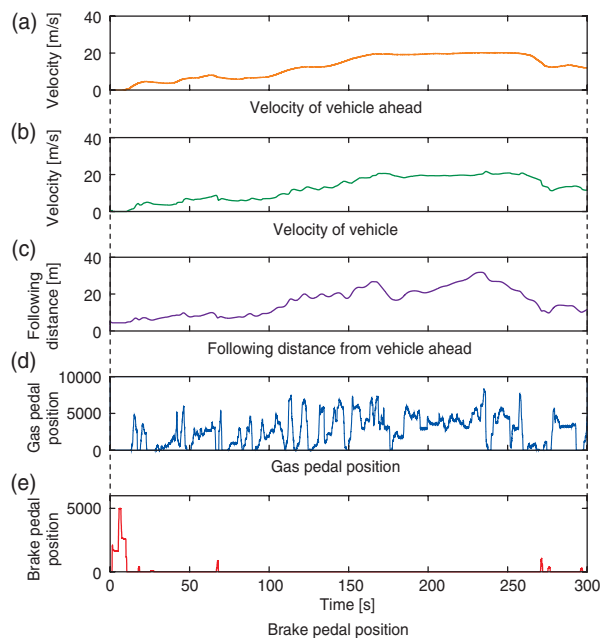
- Alphanumeric reading task: Drivers repeat random four-character strings consisting of the letters a–z and digits 0–9, e.g., "UKZ5", "IHD3", "BJB8". The original four-character strings are supplied through an earphone.
- Music retrieval task: Drivers retrieve and play music using a spoken dialog interface. Music can be retrieved by the artist's name or the song title, e.g., "Beatles", "Yesterday".

Drivers start from Nagoya University and return after about 70 min of driving. Driving data are recorded under the above four task conditions on city roads, and under two task conditions on an expressway. Driving data without any tasks are recorded as a reference before, between, and after the tasks. Figure 3 shows some samples of driving signals.

After completing the route, the participant is asked to assess his or her subjective level of frustration while driving, by simultaneously viewing the front-view camera and facial videos while listening to the corresponding audio recording. The assessment is done on a custom designed computer interface, on which the driver continuously indicates the intensity of their frustration by sliding a bar along a scale from 0 to 30, i.e., from "no frustration" to "extremely frustrated". The interface output is a continuous signal, and the level of frustration is recorded every 0.1 s.

B) Data annotation

An effective annotation of the collected multimedia information is crucial for providing a more meaningful description of the situations drivers experience. In our study, we proposed a data annotation protocol that covers most of the factors that might affect drivers and influence their subjective feedback responses. The annotation labels are comprised of six major groups: driver's affective state (level of irritation), driver actions (e.g., facial expression), driver's secondary task, driving environment (e.g., type of

Fig. 3. Examples of driving behavior signals.

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 fields. Further details of the annotation protocol and a more detailed description can be found in [11]. Having introduced our driving corpus, in the following sections we will discuss signal processing techniques and data-driven approaches for various vehicle applications.

III. DRIVER IDENTIFICATION

Driving behaviors differ among drivers. They differ in how they hit the gas and brake pedals, in the way they turn the

steering wheel, and in how much distance they keep when following a vehicle. Consequently, intelligent transportation systems (ITS) applications are expected to be personalized for different drivers according to individual driving styles. One way to achieve this is to assist each driver by controlling a vehicle based on a driver model representing the typical driving patterns of the target driver. Driver models for individual drivers or for subgroups of drivers classified based on their driving styles would be trained either in offline or online mode, and an ITS application would choose a driver model appropriate for assisting the target driver by identifying the driver or finding the model that suits his/her driving styles.

The objective of this section is to demonstrate the ability to identify a driver from an observed pattern of pedal operation. As a result of individual characteristics (e.g., personality traits, driving styles), pedal operation patterns while driving differ among drivers. Figure 4 shows examples of gas pedal operation signals of 150 sec in duration collected in a driving simulator from two drivers, recorded while they were following the same leading vehicle. Pedal operation patterns are consistent for each driver, but differ between the two drivers. We modeled the differences in gas and brake pedal operation patterns with GMMs [12] using the following two kinds of features: (1) Raw pedal operation signals and (2) Spectral features of pedal operation signals [13].

GMM is a statistical model widely used in pattern recognition, including speech and speaker recognition [12]. It is defined as a mixture of multivariate Gaussian functions, and the probability of D-dimensional observation vector $\boldsymbol{o}$ for GMM λ is obtained as follows:

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

where M is the number of the Gaussian components and $\mathcal{N}_i(\boldsymbol{o})$ is the D-variate Gaussian distribution of the ith
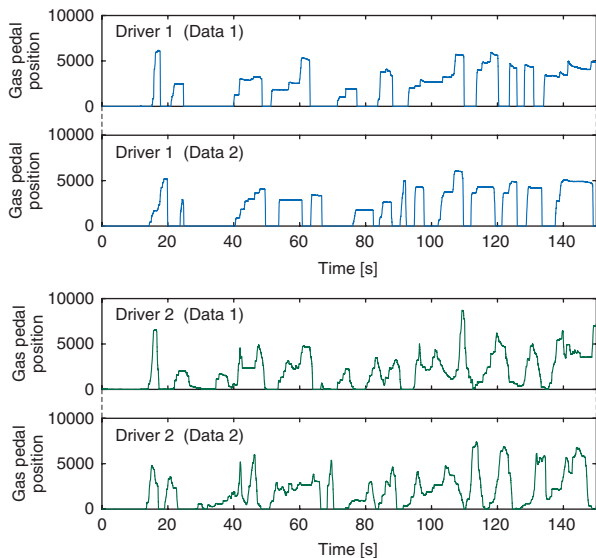

Fig. 4. Examples of gas pedal operation patterns for two drivers (Top: driver 1, Bottom: driver 2) following the same leading vehicle.
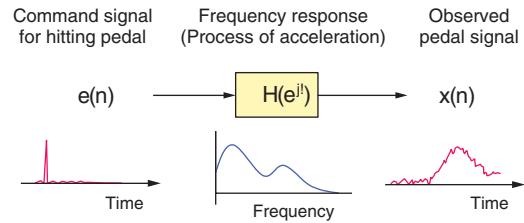

Fig. 5. General modeling of a driving signal.

component defined by mean vector $\boldsymbol{\mu}_i$ and covariance $\boldsymbol{\Sigma}_i$:

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

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

A) Spectral features of pedal operation signals

As shown in Fig. 5, in driver modeling we assume that the command signal for hitting a pedal $e(n)$ is filtered with driver model $H(e^{j\omega})$, represented as the spectral envelope, and that the output of the system is observed as pedal signal $x(n)$. In other words, a command signal is generated when the driver decides to apply pressure to the gas pedal, and $H(e^{j\omega})$ represents the process of acceleration. This can be described in the frequency domain as follows:

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

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 $H(e^{j\omega})$.

A cepstrum is a widely used spectral feature for speech and speaker recognition [14], defined as the inverse Fourier transform of the log power spectrum of the signal. Cepstral analysis allows us to smooth the structure 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. In addition, 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
In a way similar to research done on speech and speaker recognition, we found that the dynamic features of driving signals contain large amounts of information about driving behavior. Dynamic features are defined as the following linear regression coefficients:

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

where $o(t)$ 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 from preliminary experiments that the regression window is $2K = 800$ ms for both raw pedal signals and cepstral coefficients. If $o(t)$ is a D-dimensional vector, D dynamic coefficients are obtained from the static coefficients, combined into a $2D$-dimensional feature vector, and modeled with GMMs.

B) Driver identification experiments

1) EXPERIMENTAL CONDITIONS

Driving data from 276 drivers, collected on city roads in the data collection vehicle, were used, excluding data collected while the vehicle was not moving. Three minutes of driving signals were used for GMM training and another 3 min 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 driving than during 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. For driver identification, the unknown driver was identified as driver $\hat{k}$ who gave the maximum weighted GMM log-likelihood over the gas and brake pedal signals:

$$\hat{k} = \arg\max_k \{\gamma \log P(\boldsymbol{G} \mid \lambda_{G,k})$$
$$+ (1-\gamma) \log P(\boldsymbol{B} \mid \lambda_{B,k})\}, \quad 0 \le \gamma \le 1, \quad (5)$$

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 kth driver models of the gas and brake pedals, respectively. γ 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. 6. The identification performance was rather low when using raw driving signals: the best identification rate for raw signals was 47.5% with $\gamma = 0.80$. By applying cepstral analysis, however, the identification rate increased to 76.8% with $\gamma = 0.76$. We thus conclude that cepstral features capture individual variations in driving behavior better than raw driving signals and achieve better performance in driver identification.

IV. DRIVER BEHAVIOR PREDICTION

Driver behavior models can be employed to predict future vehicle operation patterns given the available observations at the time of prediction. Here, two driving tasks are considered: car following [15, 16] and lane changing [17].

A) Car following

Car following characterizes longitudinal behavior of a driver while following behind another vehicle [18]. In this study, we focus on the way the behavior of the driver of the
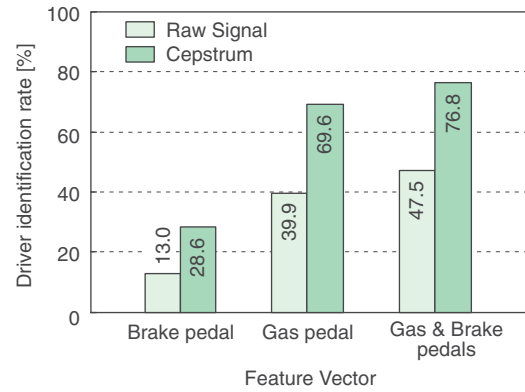

Fig. 6. Comparison of identification rates using raw pedal signals and cepstral coefficients.

following vehicle is affected by the driving environment and by the states of his or her own vehicle. There are several factors that affect 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 ability and reaction time of the follower. Figure 7 shows a basic diagram of car following and the 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 model represents patterns of pedal operation corresponding to the observed velocity and following distance. The underlying premise of this modeling framework is that a driver determines gas and brake pedal operation in response to the stimulus of vehicle velocity and following distance. Consequently, such relationship can be modeled by the joint distribution of all the correlated parameters. Figure 8 illustrates a car-following trajectory (gray dashed line) on different 2-D parameter spaces, overlaid with the contour of corresponding two-mixture GMM distribution.

1) FEATURE EXTRACTION AND MODEL REPRESENTATION

In our framework to model a pedal pattern, an observed feature vector at time t, $\mathbf{x}_t$, consists of vehicle velocity, following distance, and pedal pattern (G_t) with their first- (Δ) and second-order (Δ^2) time derivatives as

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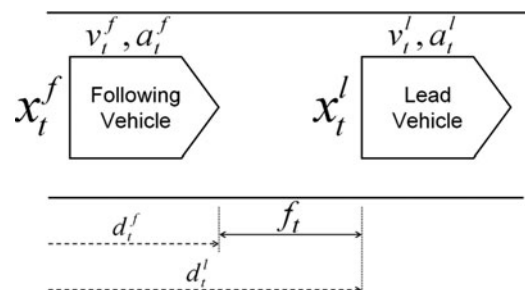$$(6)$$
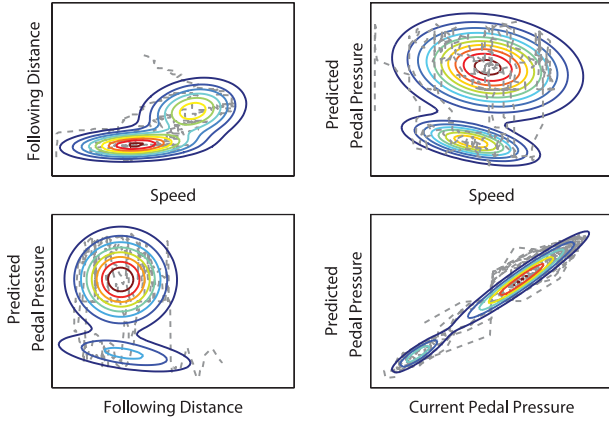

Fig. 7. Car following with corresponding parameters.

Fig. 8. A car-following trajectory (gray dashed line) on different two-dimensional parameter spaces, overlaid with the contour of corresponding two-mixture GMM distribution.

where, in this modeling, the $\Delta(\cdot)$ operator is defined as

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

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

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

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 kth mixture expressed as

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

2) PEDAL PATTERN PREDICTION
The predicted gas pedal pattern $\hat{G}_{t+1}$ is computed using the weighted predictions resulting from all the mixture components of the GMM, as

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

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 kth mixture component which is given by

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

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

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

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

3) MODEL ADAPTATION
We applied Bayesian or MAP adaptation to re-estimate the model parameters individually, by shifting the original statistic (i.e., mean vectors) toward the new adaptation data [12]. The *universal* or *background* 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's capability, we took a further step of adapting the parameters of the universal driver models, as described in the following two scenarios.

- Driver adaptation: The goal of driver adaptation is to adapt the model parameters to better represent a given individual's 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 are used to adapt the universal model, and subsequently, the on-line adapted driver model is used to represent driving behavior for the rest of each particular car-following event. The objective of on-line adaptation is to capture the overall unique car-following characteristics of a particular event (e.g., driver and environment) that deviate from the average characteristics of the universal models.

4) EXPERIMENTAL EVALUATION
Evaluation is performed using approximately 300 min 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 s are considered. The prediction was performed on every sample (i.e., every 0.1 s). Figure 9 compares the prediction performance of the universal, driver-adapted, and 30 s-on-line-adapted (using 30 s of driving data) driver models with 4, 8, 16, and 32 mixtures, in terms of SDR.

The measurement SDR is defined as follows:

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

where T is the length of a signal, $G(t)$ is the actually observed signal, and $\hat{G}(t)$ is the predicted signal. From the results, we can see that the driver-adapted models showed the best performance (approximately 19 dB SDR or 11.22% normalized errors with 32-mixture GMMs).

B) Lane changing

In this section, we consider driving behavior related to control of vehicle trajectory during lane changes. Since lane change activity consists of multiple states (i.e., examining the safety of traffic environments, assessing the positions
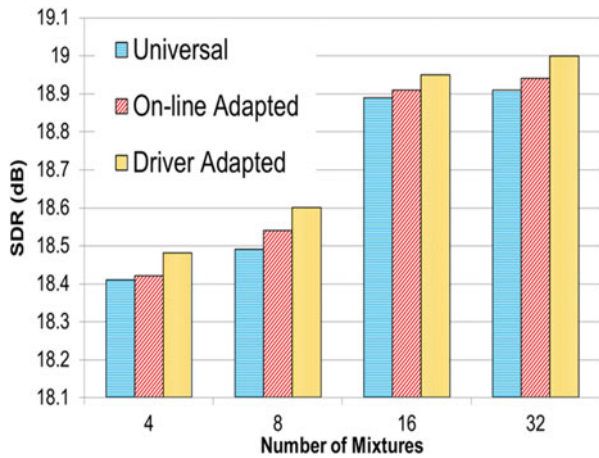
Fig. 9. Comparison of pedal prediction performance for car-following task using different driver models.

Fig. 10. Lane-change trajectory and geometric positions of surrounding vehicles.

of other vehicles, moving into the next lane, and adjusting driving speed to traffic flow) [19], a single dynamic system cannot model vehicle trajectory. In addition, the boundaries between states cannot be observed from the vehicle's trajectory.

To study lane-changing behavior, a set of vehicle movement observations were made using a driving simulator. Relative longitudinal and lateral distances from a vehicle's position when starting a lane change, $x_i[n]$, $y_i[n]$, and the velocity of the vehicle, $\dot{x}_i[n]$, $\dot{y}_i[n]$, were recorded every 160 ms. Here, $i = 1, 2, 3$ are the indexes for the locations of surrounding vehicles (Fig. 10), and $\{x_0[n], y_0[n]\}$ represents the position of the driver's own vehicle. The duration of lane-change activity, $n = 1, 2, \ldots, N$, starts when V_0 (the drivers own vehicle) and V_2 are at the same longitudinal position and ends when V_0's lateral position reaches the local minimum as shown in Fig. 10.

1) MODELING TRAJECTORY USING A HMM
We used a three-state HMM [20] 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:

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

Here, the Δ operator is defined as in equation (4). 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 total duration of the lane change activity. When a driver performs 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
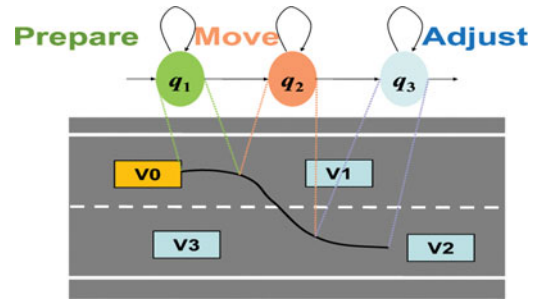
uniform re-sampling. Once a set of state durations is determined, we apply either the maximum likelihood HMM signal synthesis algorithm (ML method) [21] or the sampling algorithm [22] to generate the most probable trajectory. Simply repeating this process will produce a set of probable vehicle trajectories which characterize a driver's 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 leading vehicle than to those of vehicles to the side.) 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 a 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^{-1}) as a metric of the cognitive distance. Then we calculate the hazard map function M_i for surrounding vehicle V_i as follows:

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

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 $r_i^t R_i^{-1} r_i$.

Each hazard map M_i can be regarded as *a posteriori* probability of being in the safe driving condition within the range of distances $Pr(\text{safe}|r_i)$, when the likelihood is given as an 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 $M = \sum_{i=1,2,3} \lambda_i M_i$. 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 duration of the lane change, we can compare the possible trajectories. Then the optimal trajectory that has the lowest value is selected from among 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 with moderate traffic. The velocity of the vehicles in the passing lane ranged from 82.8 to 127.4 km/h and the distance between successive vehicles in the passing lane ranged between 85 and 315 m. The drivers were instructed to pass the lead vehicle when they were able to, once during each trial. The trained hazard maps M for the two drivers shown in Fig. 11 depicts differences in sensitivity to surrounding vehicles.

We generated possible lane-change trajectories for the vehicles over a 20-s period using the two above-mentioned methods, and then selected the optimal trajectory. Figure 12 illustrates sample trajectories generated using the sampling method and the corresponding optimal trajectory, and compares them with the actual trajectory. For quantitative evaluation, we calculated the difference between the predicted and actual trajectories based on dynamic time warping (DTW) [23], using the normalized square difference as a local distance, and measured it in terms of signal-to-deviation ratio (SDR). Figure 13 (top) shows average SDRs of the best trajectory hypothesis and all trajectory hypotheses (mean), using the ML method (left) and the sampling method (right). The sampling method was better at generating vehicle trajectories similar to the actual driver trajectories than the ML method. Figure 13 (bottom) 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 actual lane-change duration. 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, as a result of predicting vehicle trajectories over a distance of about 600 m (i.e., over a 20-s time period).
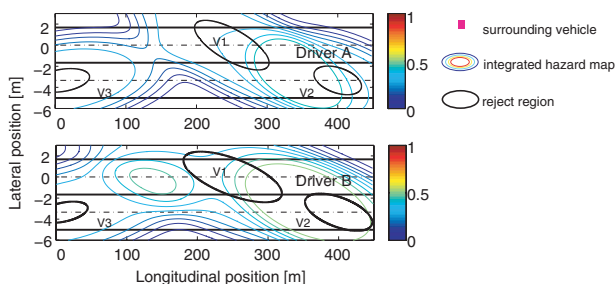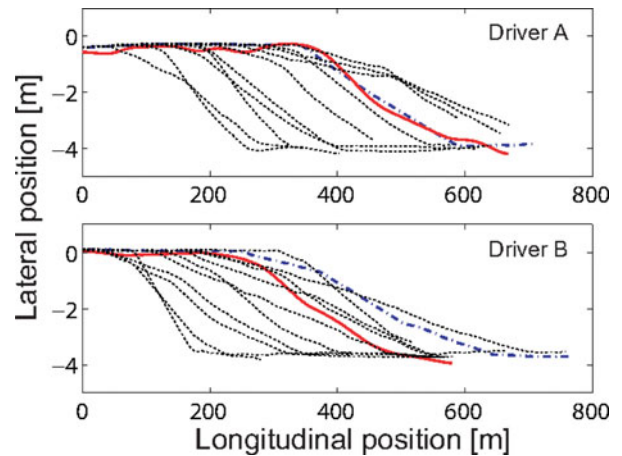

Fig. 12. Examples of generated trajectories (black dotted lines) and optimal trajectory (blue dashed line) using sampling method, compared with actual trajectory (red solid line).
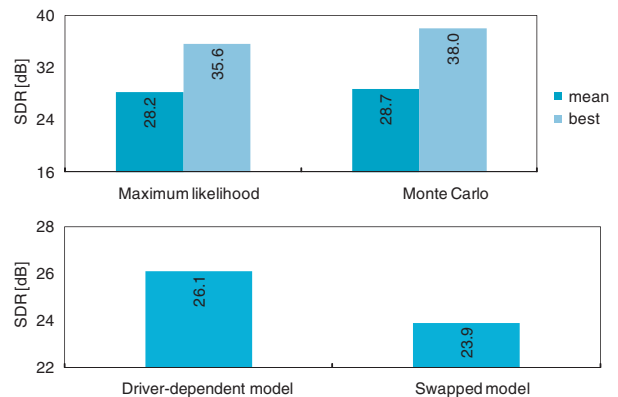

Fig. 13. Average SDRs of lane-change trajectories. Top: the best and mean trajectories using ML method (left) versus sampling method (right). Bottom: using a driver's own model (left) versus using the other driver's model (right).

V. DETECTION OF DRIVER FRUSTRATION

In this section, we propose a method that integrates features of a different nature, in order to detect driver frustration. 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 expressions, or actions [24].

A) Analysis

A method for combining all of the different features and annotation results in an efficient language was needed, and a BN [25] was the natural choice to deal with such a task. One of the important characteristics of a BN is the ability to infer the state of an unobserved variable, given the state of the observed ones. In our case, we wanted to infer a participant's frustration given the driving environment, speech recognition errors (i.e., communication environment), and the participant's responses measured through the physiological state, overall facial expression, and pedal actuation.
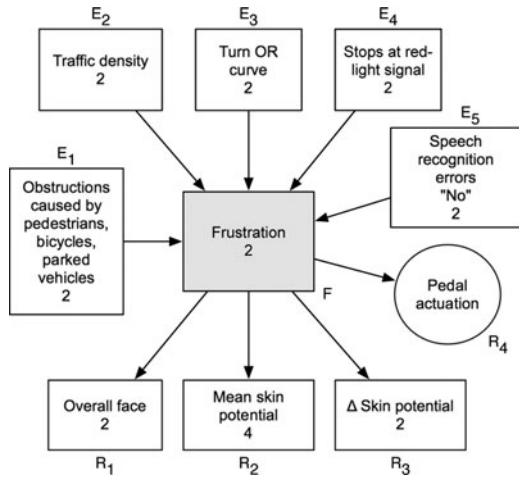

Fig. 11. Hazard maps for two drivers when surrounding vehicles were in the same positions.

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 the node can assume. Labels outside nodes identify random variable type.

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 (i.e., traffic density, stops at red-lights, obstructions, turns or curves, and speech recognition errors) may also result in driver frustration; (2) a frustrated driver is likely to exhibit changes in his or her facial expression, physiological state, and gas- and brake-pedal actuation behavior. In Fig. 14, the 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 (e.g., "2" for yes/no binary states, "4" for four levels of arousal). Random variables were identified by a label outside each node: "F" (frustration), "E" (environment), and "R" (responses).

In addition to the graph structure, it is also 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 of a child node taking on each of its different values for each combination of the values of its parent nodes. On the other hand, if the variable is continuous, the CPD is assumed to have a linear-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

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

The denominator was calculated by summing (marginalizing) out F. In addition, in this study we set a uniform Dirichlet prior [26] to every discrete node in the network. This was done in order to avoid over-fitted results due to the ML approach used for calculating the CPTs. Without a priori, 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 signals, skin potential, and other binary signals (environmental factors and speech recognition errors). 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. The results served as the network inputs. The value of each binary label at the current time step was directly entered into the network without further processing. Frame shift was kept fixed at 0.5 s. For two consecutive frames, the value of current traffic density continues to remain in effect, for example, on future skin potential and pedal actuation signals, in order to account for delayed physiological and behavioral reactions. In addition, frustration was estimated for every frame (i.e., we did not pre-select segments where we were certain of the presence or absence of frustration, while ignoring ambiguous regions.)

B) Experimental evaluation

Within the data used in our experiments, 129 scenes of frustration (segments with an original value above 0) were found. On average, participants became frustrated 6.5 times while driving in our experiment. The mean strength of frustration scenes was 10.5, and the mean duration was 11.8 s. Figure 15 shows the estimation results for all drivers concatenated side by side: actual frustration detected for all 20 participants (top); the posterior probability of the frustration node calculated using the *entire* network (center); and the quantized posterior probability using a threshold of 0.5 (bottom). The quantized probability for each driver was further median-filtered to remove unwanted spikes. The overall results show that the model achieved a TP rate of
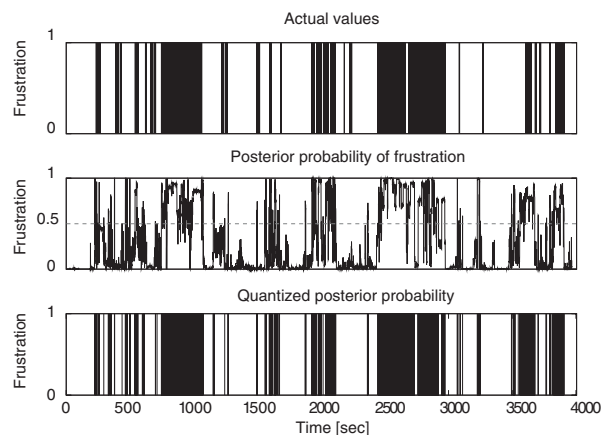

Fig. 15. Results for individual drivers (arranged side by side) calculated using the entire network. Comparison between actual frustration detected for all drivers (top), posterior probability of the frustration node (center), and its quantized version using a threshold of 0.5 (bottom).

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

80% with a FP rate of 9% (i.e., the system correctly estimate 80% of the frustration and, when drivers were not frustrated, mistakenly detected frustration 9% of the time).

VI. DRIVER COACHING

We also developed a next-generation event data recorder (EDR) by employing driver-behavior modeling. It is capable of detecting a wide range of potentially hazardous situations that would not be captured by conventional EDRs [27], and we have shown that it can be used to improve driving safety by making drivers aware of their unsafe behavior. Our automated diagnosis and self-review system was developed on a server computer as a web application for easy access via networks from PCs or smart phones [28, 29]. The system automatically detects nine types of potentially hazardous situations from the driver's own recorded driving data. These include:

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

The current version will display up to five of the most hazardous scenes for each hazard type by automatically gauging the hazard level, using the magnitude of the difference from pre-defined thresholds (for hazard types 1–7), or by using 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 detailed information

on hazardous situations detected on a given day, represented by balloon icons on an actual driving map. Each balloon represents a unique hazardous situation, with different colored balloons corresponding to different types of hazards, as shown in Fig. 16. The system also provides statistics on all the hazardous situations the driver encountered, using the archived data recorded for that driver, and displays it on a pie chart using the number of occurrences for each type of hazard. Therefore, the system could be used to identify a tendency toward risky driving behavior, or other personality traits possessed by an individual driver.

After clicking on a balloon on the map of the driving route, the corresponding video and driving signals are displayed, along with explanations of the hazardous behavior detected at that location and instructions on how the user can improve the safety of their driving. 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 hazardous situation, based on a potential driving danger analysis manual [31] (the manual is based on traffic psychology and Japanese driving rules). In general, the system will inform the user why a particular detected driving behavior is considered unsafe in a given situation, and then coach the user by suggesting safe driving behavior for that situation, in order to improve their driving skills. Figure 17 shows an example of the interface diagnosing a hazardous situation at an intersection. The system notifies the user that he or she did not stop at the stop sign, and crossed the intersection at a speed of 17 km/h. The system then suggests that, in this situation, the driver should stop completely at the stop sign and confirm that it is safe to cross the intersection before proceeding.

A) Experimental evaluation

In order to validate the effectiveness of our system in reducing the number of detected hazardous situations (e.g., to improve driving behavior), we recruited 33 drivers, including 6 expert drivers, to participate in our experiment. The subjects were asked to drive the instrumented vehicle three
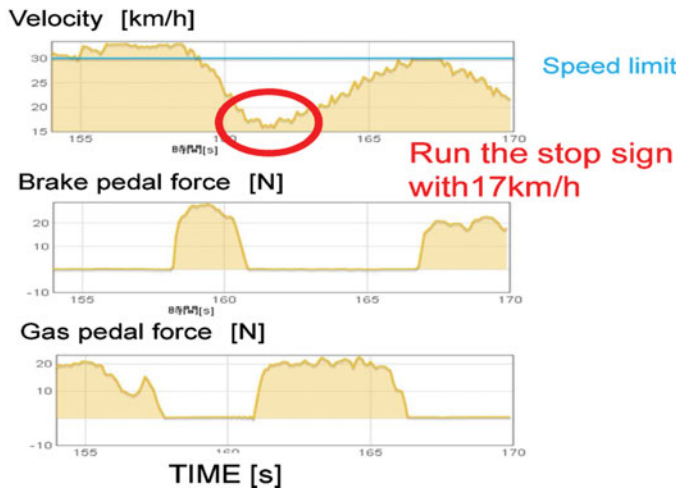
---

[1]Here, two GMM-based driver-behavior models were used, one to represent safe driving behavior, and the other to represent risky driving behavior. Risky driving behavior could be determined by performing a hypothesis test of the observations against a pre-defined threshold [30]

Fig. 17. An interface diagnosing a hazardous situation at an intersection.

times on three different days, following the same route, which takes approximately 90 min to complete. We used data from the second and the third sessions for our analysis, because we allowed the subjects to get familiar with the vehicle during the first session. After the second session, 27 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. Figure 18 compares the number of detected hazardous situations. We can see that the number of detected hazardous scenes for the non-expert drivers decreased by more than 50% after using the system, while there was no significant change for the drivers who did not use the system.

VII. SUMMARY AND FUTURE WORK

We have presented some examples of human-behavior signal processing and related modeling approaches, with a focus on the interaction between driver, vehicle, and environment. Utilizing multi-modal driving signals (e.g., brake/gas pedal pressure, steering-wheel angle, distance between vehicles, vehicle velocity, vehicle acceleration, etc.),
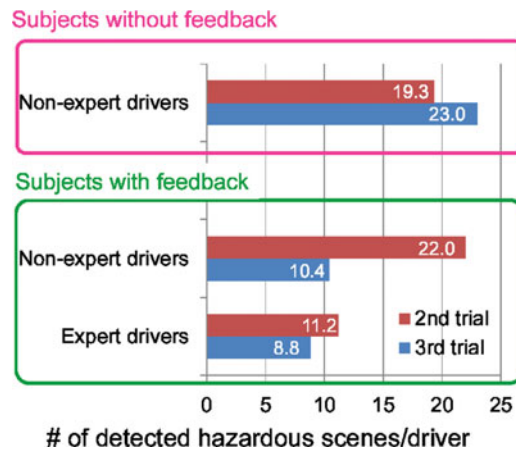

Fig. 18. Number of detected hazardous scenes for non-expert drivers who did not receive feedback (top), and for non-expert and expert drivers before and after using the system (bottom).

we were able to capture meaningful characteristics of driver behavior, and driver models were then employed to detect, predict and assess driving behavior. Experimental evaluations using real-world driving data have shown promising outcomes with a wide range of vehicle applications, such as recognizing driver identity, predicting driver behavior during maneuvers (i.e., car following and lane changing),

detecting mental states of drivers (i.e., frustration), and assessing driving behavior for driver coaching. Our future work will focus on identifying individual driving characteristics, as well as examining variations in driving behavior between drivers, by comparing drivers from different countries, for example.

ACKNOWLEDGEMENTS

REFERENCES

[1] Levin, E.; Pieraccini, R.; Eckert, W.: A stochastic model of human-machine interaction for learning dialog strategies. IEEE Trans. Speech Audio Process, **8**(1), (2000), 11–23.

[2] Pentland, A.; Liu, A.: Modeling and prediction of human behavior. Neural Comput., **11**, (1999), 229–242.

[3] Wickens, C.D. Engineering Psychology and Human Performance, *HarperCollins Publishers*, New York, 1992.

[4] Lum, H.; Reagan, J.A.; Interactive highway safety design model: accident-predictive module. Public Roads Mag., **59**(2), (1995). http://www.fhwa.dot.gov/publications/publicroads/95winter/p95wi14.cfm

[5] Wouters, P.I.J.; Bos, J.M.J.: Traffic accident reduction by monitoring driver behaviour with in-car data recorders, *Accid. Anal. Prev.*, **32**(5), (2000), 643–650.

[6] McCall, J.C.; Trivedi, M.M.: Driver behavior and situation aware brake assistance for intelligent vehicles. IEEE Trans. Intell. Transp. Syst., **95**(2), (2007), 374–387.

[7] Oliver, N.; Pentland, N.P.: Driver behavior recognition and prediction in a SmartCar, in *Proc. SPIE Aerosense, Enhanced and Synthetic Vision*, vol. 4023, April 2000, 2280–2290.

[8] Salvucci, D.D.; Boer, E.P.; Liu, A.: Toward an integrated model of driver behavior in a cognitive architecture, *Transp. Res. Rec.*, 2001, 9–16.

[9] Takeda, K.; Hansen, J.H.L.; Boyraz, P.; Malta, L.; Miyajima, C.; Abut, H.: An international large-scale vehicle corpora of driver behavior on the road, *IEEE ITS*, **12**(4), (2011), 1609–1623.

[10] Miyajima, C.; Kusakawa, T.; Kitaoka, N.; Itou, K.; Takeda, K.: On-going data collection of driver behavior signals, in *DSPINCARS*, 2007.

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

[12] Reynolds, D.A.; Quatieri, T.F.; Dunn, R.B.: Speaker verification using adapted gaussian mixture models. Digit. Signal Process., **10**(1), (2000), 19–41.

[13] Miyajima, C. et al., Driver modeling based on driving behavior and its evaluation in driver identification. Proc. IEEE, **95**(2), (2007), 427–437.

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

[15] Nishiwaki, Y.; Miyajima, C.; Kitaoka, H.; Itou, K.; Takeda, K.: Generation of pedal operation patterns of individual drivers in car-following for personalized cruise control, in *IEEE-IV*, Taiwan, 2007, 823–827.

[16] Angkititrakul, P.; Miyajima, C.; Takeda, K.: Modeling and adaptation of stochastic driver-behavior model with application to Car-following, *IEEE IV,* Baden-Baden, Germany, June 2011, 814–819.

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

[18] Brackstone, M.; McDonald, M.: Car-following: a historical review. Transportation Research Part F, **2**(4), (1999), 181–196.

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

[20] Rabiner, L.R.: A tutorial on hidden markov models and selected applications in speech recognition. Proc. IEEE, **77**(2), (1989), 257–286.

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

[22] Rubinstein R.Y. and Kroese, D.P.: Simulation and the Monte Carlo Method, 2nd ed., *John Wiley & Sons*, New York, 2007.

[23] Sakoe, H.; Chiba, S.: Dynamic programming algorithm optimization for spoken word recognition, IEEE Trans. Acoust., Speech, Signal Process., **26**(1), (1978), 43–49.

[24] Malta, L.; Miyajima, C.; Kitaoka, N.; Takeda, K.: Analysis of real-world driver frustration. IEEE ITS, **12**(1), (2011), 109–118.

[25] Murphy, K.P.: Dynamic Bayesian networks: representation, inference, and Learning, PhD Dissertation, Univ. of California at Berkeley, 2002.

[26] Ferguson, T.S.: A Bayesian analysis of some nonparametric problems. *Ann. Stat.*, **1**(2), (1973), 209–230.

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

[28] Takeda, K.: et al. Improving driving behavior by allowing drivers to browse their own recorded driving data, in *IEEE-ITSC*, Washington, October 2011.

[29] Takeda, K.: Self-coaching system based on recorded driving data: learning from one's experiences, *IEEE ITS*, **13**(4), (2012), 1821–1831.

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

[31] Renge, K.: Manual of Potential Danger Analysis (in Japanese), Internal Document, 2009.

Chiyomi Miyajima received the B.E., M.E., and Dr.Eng. degrees, all in Computer Science from Nagoya Institute of Technology, Nagoya, Japan, in 1996, 1998, and 2001, respectively. From 2001 to 2003, she was a Research Associate with the Department of Computer Science, Nagoya Institute of Technology. She is currently an Assistant Professor with the Graduate School of Information Science, Nagoya University. Her research interests include human behavior signal processing and speech processing.

Pongtep Angkititrakul received the B.Eng. degree in Electrical Engineering from Chulalongkorn University, Bangkok, Thailand, in 1996 and the M.S. and Ph.D. degrees in Electrical Engineering from the University of Colorado, Boulder, in

1999 and 2004, respectively. From 2007 to 2010, he was a Visiting Researcher with Toyota Central R&D Laboratories, Aichi, Japan. He is currently a Research Associate with the Graduate School of Information Science, Nagoya University, Nagoya, Japan. His research interests are human behavior signal processing and speech signal processing.

Kazuya Takeda received his B.E., M.E. in Electrical Engineering and Doctor of Engineering degrees from Nagoya University, Nagoya Japan, in 1983, 1985, and 1994, respectively. From 1986 to 1989 he was with the Advanced Telecommunication Research Laboratories, Japan. His main research interest at ATR was corpus-based speech synthesis. From 1989 to 1995, he was at KDD Laboratories, Japan. He led a research project on a voice-activated telephone extension system at the Laboratories. From 1995 to 2003, he was an Associate Professor of the School of Engineering at Nagoya University, where he started a project to collect human behavior signals while driving, and to apply signal processing technologies for understanding the signals. Since 2003, he has been a Professor at the Department of Media Science, Graduate School of Information Science, Nagoya University. His current research interests are media signal processing including; spatial audio, robust speech recognition, and human behavior modeling.