Investigation of Driving-Behavior Modeling for Recognition of a Driving Situation

Junki Ema*, Longbiao Wang*, Atsuhiko Kai* and Toshihiko Itoh†
* Shizuoka University, Hamamatsu 432-8561 Shizuoka Japan
† Hokkaido University, Sapporo, Hokkaido 060-8628, Japan

***Abstract*—We investigate a method for recognizing driving situations on the basis of driving signals for application to a safe human interface of an in-vehicle information system. A hidden Markov model (HMM)-based pattern recognition method is used to model and recognize the driving situation, which is classified as one of seven categories. We attempt to find the optimum HMM configuration to improve the performance of driving situation recognition. We also analyze factors that degrade the recognition performance and discuss a solution. CIAIR in-vehicle corpus was used to evaluate the HMM-based recognition method. Driving situation categories were recognized using five driving signals. The accuracy of recognition was 78.3% without indicator information and 84.1% with indicator information.**

I. Introduction

Recently, an increasing number of studies on in-vehicle signal processing have been reported [1], [2]. One of the main applications of in-vehicle signal processing is a safe in-vehicle information system, especially for a speech-based car navigation system. Lee et al. reported that a problem with a more natural spoken dialogue interface as part of a car-navigation system is distraction whereby machine operation and voice conversation influence driving of the vehicle [3]. One possible solution to this problem is that a spoken dialogue system changes its dialogue rhythm according to the driving situation, which should improve safety when using a speech-based car-navigation system. The prediction or recognition of driving situations is an important issue in such a case. Oliver and Pentland presented a driver behavior recognition and prediction method based on dynamic graphical models, hidden Markov models (HMMs) and coupled HMMs [4] trained using video and vehicle signals [5]. The computational cost was high due to the complexity of the model architecture and algorithm and the large amount of signal data including video and vehicle signals. Nishiwaki et al. proposed a stochastic framework for modeling driving behavior during lane changing where driver habitual and cognitive characteristics are modeled by an HMM and geometrical probability function [6]. However, only the scenario of lane changing was discussed. One of the authors reported the results of an experiment on the prediction of driving actions (that is driving situations) [7]. The prediction system employed the HMM-based pattern recognition method only for driving signals and did not use positional and video information. The best driving action (driving situation) prediction accuracy was 63.2% using 10 seconds of driving signal for the prediction. In this paper, we extend the HMM-based method to recognition of a driving situation. Driving behaviors are classified into seven situations (stopping, turning right, turning left, changing lanes to the right, changing lanes to the left, avoiding obstacles and other actions, which differ from the six categories previously used [7]). The driving behaviors are modeled by HMMs by training with driving signals from many drivers. The driving-behavior model with highest likelihood is selected as the result of driving-situation recognition. We aim to find the optimum HMM configuration (that is, the number of states and number of mixtures etc.) to improve driving-situation recognition. We also analyze factors that degrade driving-situation recognition and discuss solutions.

The remainder of this paper is organized as follows. Section II describes our driving-situation recognition task. In Section III, a technique to model and recognize the driving situation is presented. Section IV describes the experimental results of driving-situation recognition. Finally, Section V summarizes the paper and describes the direction of future work.

II. Driving-Situation Recognition Task

A. Driving Signal Database

In this paper, the Center for Integrated Acoustic Information Research (CIAIR) database is used to evaluate our method. Multi-modal information including audio and video as well as information on vehicle operation and position was collected as part of a CIAIR project. Detailed information on CIAIR corpus can be found in [8]. In this paper, only the five driving signals given in Table I are used: forces on the gas and brake pedals, engine speed, car velocity and steering angle. The five signals are synchronized with the multiple modality driving data. They were sampled at 1 kHz, and each sample was encoded in 16 bits.

B. Labeling of Driving Situations

It is considered that driving situations can be classified into many categories according to the combination of traffic situations and driver behavior. However, we considered that

TABLE I
Driving signal data

Driving signal	Range
Gas pedal pressure (accel.)	0–50 kgf / cm^2
Brake pedal pressure (brake)	0–50 kgf / cm^2
Steering angle	–1800° to 1800°
Engine speed	0–8000 rpm
Car velocity	0–120 km/h

Proceedings of the Second APSIPA Annual Summit and Conference, pages 161–164,
Biopolis, Singapore, 14-17 December 2010.

TABLE II
CATEGORIES OF DRIVING SITUATIONS

Driving situations	Label
Avoiding obstacles	A
Changing lanes to the left	CL
Changing lanes to the right	CR
Turning left	L
Turning right	R
Stopping	S

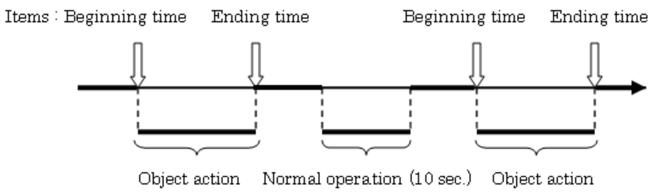

Fig. 1. Selected rough-driving situations and normal operation

very detailed classification of the driving situation is not needed in changing the dialogue rhythm according to the driving situation to avoid distraction that would influence the driving of the vehicle. Therefore, in this study, we aim to classify driving situations into categories that require great control of the steering wheel, accelerator and brake. Table II shows the driving situations and their corresponding labels. Here, these six plus one (others) categories are the object of modeling and recognition. The labeling was carried out by one person using animation data.

III. TECHNIQUE FOR MODELING AND RECOGNITION OF A DRIVING BEHAVIOR

A. Outline of the Recognition Technique

In this paper, we focus on recognizing driving situations using the features of the gas pedal pressure, brake pedal pressure, steering angle, engine speed and car velocity. We adopt an HMM because the driving signals are patterns of sequential data and HMM works well in modeling categories of such data patterns. The modeling of driving behavior and the training data required are described in Section III-B, and the method of recognizing driving behavior is described in Section III-C.

In this study, the driving-behavior segment is assumed to be already known, and thus, processing of an unknown driving-behavior segment is future work.

B. Construction of the Driving-Behavior Model

Each driving-behavior segment used as training data or test data is determined by the time information of the beginning and end of the corresponding driving behavior. Some real driving situations differ from the rough-driving situations presented in Table II. The other driving behavior is referred to as the normal driving situation (N) in this study. Therefore, it is also necessary to model the normal driving situation (normal operation). The training and evaluation data of normal operation are selected at random for 10 seconds, which is the average duration for the other six driving situations. Fig. 1 shows the selected rough driving situations and normal operation. There is no overlap between the segment of normal operation of 10 seconds duration and the segment of other rough driving situations indicated in Table II, and the segment of normal operation is between the end of the previous situation and the beginning of the next situation.

The driving signals sampled at 1 kHz are downsampled to 10 Hz since we consider that a sampling interval of about 100 ms is empirically adequate to capture features in recognizing a driving situation. A vector of 10 feature dimensions (five driving signals plus their first-order derivatives) is used for training and testing. A left-to-right HMM is adopted to model each driving behavior.

C. Recognition of Driving Situations

In this paper, the driving behaviors are recognized employing a two-stage method.

It is obvious that the stopping behavior should be recognized by the feature of car velocity alone. Thus, in the first stage, it is determined whether evaluation data indicate the stopping behavior or another driving behavior. The judgment of stopping behavior is determined by a linear classifier with a threshold θ. Precision, recall and the F-measure are employed as evaluation measures. Precision is defined as the number of recognized true driving situations divided by the total number of elements labeled as belonging to the category of a positive driving situation, and recall is defined as the number of recognized true driving situations divided by the total number of elements that actually belong to the positive driving-situation category. The F-measure is the harmonic mean of precision and recall. In this paper, the threshold θ is set to 0.7 km/h because the best F-measure (97.6%) is obtained for development data at this speed . That is, if the average car velocity is lower than the threshold, the segment is judged as a stop situation. The performance of the distinction between stopping and other driving behaviors of the evaluation data is performed. The F-measure of the recognition of the stopping behavior is 99.0%.

In the second stage, driving situations other than stopping are recognized as follows. Given a segment of an unknown driving situation, the likelihood of each driving-behavior model is calculated, and the driving-behavior model with the highest likelihood is determined as the recognition result. The recognition technique is illustrated in Fig. 2.

IV. EXPERIMENTS

A. Experimental Setup

As mentioned in Section III-A, it is assumed that a segment of each driving situation is already known and has been used as a recognition unit. Driving signals of 70, 10 and 30 people (about 18 minutes per person on average) were used as training data, development data and evaluation data, respectively. Table III shows the statistics of driving-behavior data.

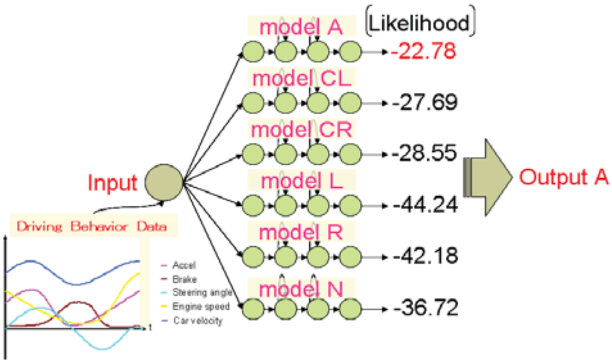

Fig. 2. Illustration of HMM-based recognition of the driving situation

TABLE III
BREAKDOWN OF DRIVING SIGNALS (UNIT: # OF SEGMENTS)

	A	CL	CR	L	R	S	N
Training set (70 people)	504	64	20	112	9	545	768
Development set (10 people)	61	8	3	16	2	80	108
Evaluation set (30 people)	236	27	11	45	5	258	344

B. Experimental Results

We aim to determine the optimum HMM configuration (number of emitting states and number of mixtures) to obtain the best driving-situation recognition using all possible combinations of HMM configurations. However, this is impractical because of the computational costs. We verify three combinations of HMM configurations using the development set to improve the driving-behavior recognition. The three methods for optimizing HMM configurations are as follows.

method 1: The number of states and the number of mixtures of each driving behavior are the same, and the HMM configuration with highest recall rate is selected.

method 2: The HMM configuration with highest likelihood in the training stage is selected for each category model.

method 3: The initial HMM configuration is chosen employing *method 1*, and it is adjusted empirically.

In *method 1*, the maximum number of states is seven and the number of mixtures is four because the training data were insufficient beyond these numbers of parameters. The results of *method 1* are presented in Table IV. The best driving-behavior result was obtained when HMMs with five states and four mixtures were used.

In *method 2*, the HMM configuration of each driving-behavior model is decided by the likelihood in the training stage. We trained each driving-behavior model using 2–64 mixtures, but training of the model may fail because the size of the training dataset is limited. In training the HMMs, the number of EM iterations was set to 20 for each mixture, and the number of mixtures was increased as 2, 4, 6, 8, 16, 32 and 64. The best configurations are decided considering the change in the likelihood curve of the model and setting the same number of configurations for the right-turn behavior and left-turn behavior and the same number for changing lanes to the right and changing lanes to the left. Table V shows the HMM configuration for each driving-behavior category.

TABLE IV
RECALL OF DRIVING-SITUATION RECOGNITION OF THE DEVELOPMENT SET EMPLOYING *method 1*

# of states	# of mixtures	Recall (%)
1	2	57.0
3	2	71.7
5	2	72.0
7	2	71.0
1	4	65.0
3	4	71.7
5	**4**	**78.3**
7	4	73.1

TABLE V
HMM CONFIGURATIONS OF *method 2*

Category	# of states	# of mixtures
A	13	16
CL	3	16
CR	3	16
L	7	4
R	7	4
N	7	32

In *method 3*, the initial configuration was chosen employing *method 1* and then adjusted empirically. We consider that the number of states for category A (avoiding obstacles) should be increased because category A includes many precise driving operations. Moreover, the number of mixtures for category N (normal operation) should be increased because category N includes various driving behaviors that differ from the other six driving behaviors listed in Table II. Thus, the HMM configurations of *method 3* were adjusted as in Table VI. Driving-behavior recognition results of the development set obtained using *method 3* are presented in Table VII. The best performance was achieved in the case of adjusting the HMM configuration of only category N. In the remainder of this paper, *method 3* refers to the HMM configurations given in Table VI-(b).

Table VIII presents the driving-situation results of the test set obtained using *method 1*, *method 2* and *method 3*. The average driving-behavior recognition results obtained using *method 1*, *method 2* and *method 3* were 74.0%, 75.1% and 78.3%, respectively. *Method 3* had the best performance (significant with $p < .05$). However, the recognition performance of CL and CR (lane changes) is very low, and they are often falsely recognized as A (avoidance of obstacles) or lane changes in the opposite direction. As they essentially seem to be confusable, we thought that additional indicator information should be

TABLE VII
PERFORMANCE OF THE DRIVING-SITUATION RECOGNITION OF THE DEVELOPMENT SET EMPLOYING *method 3*

Parameter configuration	Recall (%)
Only A & N has changed	78.7
Only N has changed	**79.0**
Only A has changed	66.4

TABLE VI

COMPARISON OF HMM CONFIGURATIONS (*method 3*)

(a) Adjusting configuration of categories A & N

Category	# of states	# of mixtures
A	13	16
CL	5	4
CR	5	4
L	5	4
R	5	4
N	5	32

(b) Adjusting configuration of category N

Category	# of states	# of mixtures
A	5	4
CL	5	4
CR	5	4
L	5	4
R	5	4
N	5	32

(c) Adjusting configuration of category A

Category	# of states	# of mixtures
A	13	16
CL	5	4
CR	5	4
L	5	4
R	5	4
N	5	4

TABLE VIII

DRIVING-SITUATION RECOGNITION RESULTS FOR THE TEST SET (%)

	method 1			*method 2*			*method 3*		
Category	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure
A	64.3	58.5	61.3	61.3	60.0	60.6	46.4	76.8	57.8
CL	23.1	16.2	19.1	15.4	13.8	14.6	3.9	5.6	4.6
CR	9.1	5.6	7.0	0.0	0.0	0.0	9.1	14.3	11.1
L	97.8	83.0	89.8	97.8	84.6	90.7	97.8	86.3	91.7
R	100	66.7	80.0	100	66.7	80.0	100	66.7	80.0
S	99.6	98.5	99.0	99.6	98.5	99.0	99.6	98.5	99.0
N	64.6	76.0	69.9	70.5	72.4	71.5	90.3	70.2	79.0
ALL	74.0			75.1			78.3		

TABLE IX

DRIVING-SITUATION RECOGNITION RESULTS FOR THE TEST SET USING ADDITIONAL INDICATOR INFORMATION (%)

	method 1			*method 2*			*method 3*		
Category	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure
A	71.1	62.6	66.5	65.5	64.4	65.0	50.6	83.2	63.0
CL	92.3	100.0	96.0	88.5	95.8	92.0	92.3	100.0	96.0
CR	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
L	100.0	95.7	97.8	97.8	93.6	95.7	100.0	95.7	97.8
R	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
S	99.6	98.5	99.0	99.6	98.5	99.0	99.6	98.5	99.0
N	70.5	77.9	74.0	74.9	75.8	75.4	92.9	73.1	81.8
ALL	81.0			81.0			84.1		

useful. The effect of indicator information on driving situations recognition is discussed in the next section.

C. Effect of Additional Indicator Information

Indicator information was not collected for the CIAIR database; therefore, it is unrealistic to use indicator information in this experiment. However, indicator information can be collected from the signal line of a car. In this section, we discuss the effect of indicator information on driving-situation recognition if it is available. In this work, we assume that an indicator is not used when avoiding obstacles and in normal driving, the left indicator is used when turning left and changing lanes to the left, and the right indicator is used when turning right and changing lanes to the right. Fig. 3 compares the driving-situation recognition performance with and without indicator information, and shows an average improvement of about $5 \sim 6\%$. The detailed results for each driving behavior are presented in Table IX. When indicator information was given, the accuracy of CL and CR (lane changes) improved remarkably, and the improvement in A (avoidance of obstacles) and N (normal operation) was less than that in CL and CR.

V. CONCLUSION AND FUTURE WORK

In this paper, we investigated a method of recognizing driving situations on the basis of driving signals for application to a safe human interface as part of an in-vehicle information system. Driving situations were categorized using five driving signals. We verified three combinations of HMM configurations using a development set to improve driving-behavior recognition. As a result, when HMM configurations of each driving-behavior model for best performance was investigated, the method that the number of states and number of mixtures of each driving-behavior are the same, and it is adjusted empirically (*method 3*) was achieved best performance with 78.3% ($p < .05$). Moreover, when indicator information was provided, the driving-behavior recognition of CL and CR (lane changes) improved remarkably, and an average improvement of about $5 \sim 6\%$ was achieved. In future work, we will attempt to evaluate recognition when the driving-behavior segment is unknown.

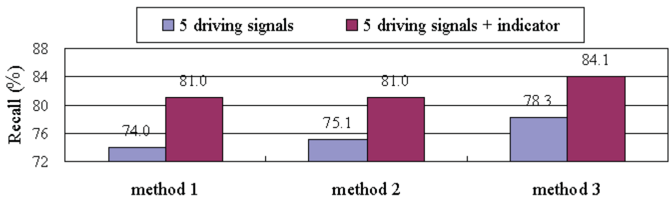

Fig. 3. Performance of driving-situation recognition when indicator information is included

REFERENCES

[1] K. Takeda, H. Erdogan, J. .H. L. Hanson and H. Abut, *In-Vehicle Corpus and Signal Processing for Driver Behavior*, Springer, 2008.

[2] L. Malta, C. Miyajima and K. Takeda, "A Study of Driver Behavior Under Potential Threats in Vehicle Traffic", IEEE Trans. on Intelligent Transport System, Vol. 10, No. 2, pp. 201–210, 2009.

[3] J. D. Lee, T. L. Brown, B. Caven, S. Haake and K. Schmidt, "Does a speech-based interface for an in-vehicle computer distract drivers?", Proc. World Congress on Intelligent Transport System, 2000.

[4] N. Oliver, *Towards Perceptual Intelligence: Statistical Modeling of Human Individual and Interactive Behaviors,* PhD thesis, MIT, Media Lab, Cambridge, MA, 2000.

[5] N. Oliver and A. P. Pentland, "Driver behavior recognition and prediction in a SmartCar", Proc. SPIE Aerosense 2000, Enhanced and Synthetic Vision, Apr., 2000.

[6] Y. Nishiwaki, C. Miyajima, N. Kitaoka and K. Takeda, "Stochastic Modeling of Vehicle Trajectory During Lane-changing", Proc. ICASSP 2009, pp. 1377–1380, 2009.

[7] T. Itoh, S. Yamada, K. Yamamoto and K. Araki, "Prediction of Driving Actions using Driving Signals", Proc. Biennial on DSP for in-Vehicle and Mobile Systems, PaperID : TM-6 206, 2007.

[8] N.Kawaguchi, S.Matsubasa, K.Takeda, and F. Itakura, "Multimedia data collection of in-car speech communication", Proc. EUROSPEECH 2001, pp.2027–2030, 2001.