TOWARDS PREDICTION OF DRIVING BEHAVIOR VIA BASIC PATTERN DISCOVERY WITH BP-AR-HMM

Ryunosuke Hamada† *Takatomi Kubo*† *Kazushi Ikeda*† *Zujie Zhang*† *Tomohiro Shibata*†
Takashi Bando‡ *Masumi Egawa*‡

† Nara Institute of Science and Technology, Japan ‡ DENSO CORPORATION, Japan

ABSTRACT

Prediction of driving behaviors is important problem in developing the next-generation driving support system. In order to take account of diverse driving situations, it is necessary to deal with multiple time series data considering commonalities and differences among them. In this paper we utilize the beta process autoregressive hidden Markov model (BP-AR-HMM) that can model multiple time series considering common and different features among them using the beta process as a prior distribution. We apply the BP-AR-HMM to actual driving behavior data to estimate VAR process parameters that represent the driving behaviors, and with the estimated parameters we predict the driving behaviors of unknown test data. The results suggest that it is possible to identify the dynamical behaviors of driving operations using BP-AR-HMM, and to predict driving behaviors in actual environment.

***Index Terms*—** driving behavior prediction, Bayesian nonparametric approach, beta process autoregressive hidden Markov model, beta process

1. INTRODUCTION

Even nowadays approximately 690,000 per year of traffic accidents still occur, although the number of accidents develops a trend to decrease in Japan [1], so it is imperative to strive to prevent accidents furthermore. Practically, some researchers have developed enthusiastically the indices of the risk of collision and the automatic emergency brake system for automotive vehicle, to suppress the number of traffic accidents [2, 3]. And recently researchers turn to think about the estimation of driving scenes and the prediction of behaviors of drivers to realize novel driving support systems, in order to support drivers in diverse environment [4–9], not just to prevent collisions. If we can estimate driving scenes or driver's behaviors, it is possible to utilize the driving support system like the collision preventing system according to present driving scene, which is effective to prevent accidents beforehand.

When we intend to model driver's behaviors in order to estimate driving scenes or predict driving behaviors, hidden Markov model (HMM) [3, 5] that treats time series data, or its extension such as autoregressive hidden Markov model (AR-HMM) [6–8] are often utilized. Ikeda et al. modeled time series data of brake pressure, and proposed the training data selection method of the adaptive brake alerting system using likelihood for the model [5]. Other researchers used the AR-HMM whose output vector under a state is subject to its own vector-autoregressive (VAR) process [6–8]. Although it is possible to model time series dataset that have the same set of states, transition probabilities and output processes jointly using HMM or AR-HMM, a dataset that doesn't satisfy the assumption must be modeled separately. In practice it is possible to occur that it exhibits a specific behavior in a certain time interval, so it is not easy to judge whether we can model a set of time series data jointly or not. And if we model time series dataset including a time series that exhibits its own behaviors, we may fail to discover such behaviors. In order to model driving behaviors under diverse driving scenes, it is necessary to utilize the novel method that can solve the problem to deal with common or different features across multiple time series data.

Fox et al. proposed novel efficient modeling method, beta process autoregressive hidden Markov model (BP-AR-HMM) that utilizes beta process prior and enables to model multiple time series data considering common or different features across a set of data [10]. They identified not only the common behaviors across multiple time series, but also the specific behavior exhibited in a specific time series using multiple time series of motion capture data.

In this paper, we will apply BP-AR-HMM to the actual driving time series dataset and verify whether it can discover the dynamical behaviors of driving operations, and predict the driving behaviors in actual environment.

2. DRIVING BEHAVIOR MODELING

To model whole driving operation time series dataset, we utilized BP-AR-HMM that is an extension of HMM and AR-HMM. This section describes an outline of BP-AR-HMM with reference to HMM and its extension.

2.1. HMM and its model extension

To model time series data, hidden Markov model (HMM) and autoregressive hidden Markov model (AR-HMM) are widely utilized. In HMM each time point of time series has its latent state, and the latent state generates observable variables to model time series. And each latent state is subject to the Markov process, so its transition to a succeeding state is controlled by the transition matrix that describes the probabilities of transition from a state to all probable states. In AR-HMM that is an extension of HMM, observable variables are subject to the identical VAR process as it belongs to the identical state. According to this property we can expect that AR-HMM will give more promising result than HMM, when we apply it to data that exhibit its dynamical behavior contiguously. If we adopt either HMM or AR-HMM, however, it is necessary to determine the number of states using the cross-validation or according to the information criterion.

Fox et al. proposed an extensional model that can determine the number of states according to training data, sticky hierarchical Dirichlet process hidden Markov model (sHDP-HMM) [11]. The sHDP-HMM is a kind of methods referred to as Bayesian nonparametric approaches that are developed actively by researchers recently. The methodology of Bayesian nonparametrics is one of the method of Bayesian statistics, attempting to learn the model complexity automatically according to training data [12, 13]. Taniguchi et al. utilized sHDP-HMM to model driving behaviors, and succeeded to segment driving time series [9]. In contrast, we utilized BP-AR-HMM that is an extension of AR-HMM as a Bayesian nonparametric approach. BP-AR-HMM can take into account either common or different features across multiple time series data, modeling them jointly with the prior probability distributions generated by beta process [10].

2.2. BP-AR-HMM

Fox et al. proposed BP-AR-HMM as a Bayesian nonparametric approach that can model multiple related time series data taking into account commonalities and differences among them. Each state has its dynamical behavior, and each dynamical behavior is represented by a specific VAR process. As is for sHDP-HMM, it allow the number of states to be countably infinite in theory, and the number is determined according to the intrinsic complexity of a training dataset. Transition from a state to its succeeding state is subject to the Markov process as well as AR-HMM, but transition probabilities are determined for each time series respectively. Fig. 1 shows the graphical model of BP-AR-HMM.

Authors assume that there exits N time series data and they share common dynamical behaviors $\theta_1, \theta_2, \ldots$. Binary indicator variable $\mathbf{f}_i = [f_{i1}, f_{i2}, \ldots]$ represents which dynamical behaviors time series i exhibits. When time series i exhibits dynamical behavior k , it is represented as $f_{ik} = 1$, and

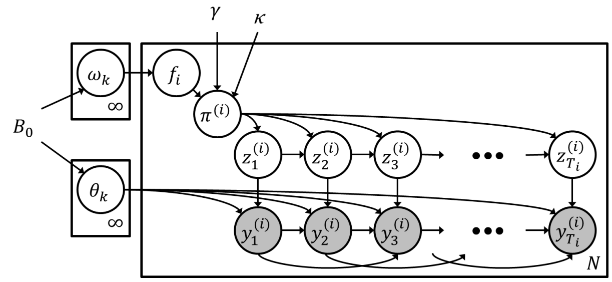


Fig. 1. Graphical model of BP-AR-HMM.

f_{ik} can be defined by Bernoulli process and represented as:

$$f_{ik}|\omega_k \sim Bernoulli\left(\omega_k\right) \tag{1}$$

where mass ω_k is a mass of an atom in a draw B that is generated by beta process conjugate to Bernoulli process, which is represented by base measure B_0, ω_k and θ_k:

$$B|B_0 \sim BP\left(1, B_0\right) \tag{2}$$

$$B = \sum_{k=1}^{\infty} \omega_k \delta_{\theta_k} \tag{3}$$

where δ represents Kronecker's delta. Beta process is conjugate to Bernoulli process, and marginalizing it along B results to gain predictive distributions known as Indian buffet process (IBP) [14]. In time series i , transition from a state to its succeeding state is subject to Dirichlet distribution:

$$\pi_j^{(i)}|\mathbf{f}_i, \gamma, \kappa \sim Dir\left([\gamma, \ldots \gamma, \gamma+\kappa, \gamma, \ldots] \otimes \mathbf{f}_i\right) \tag{4}$$

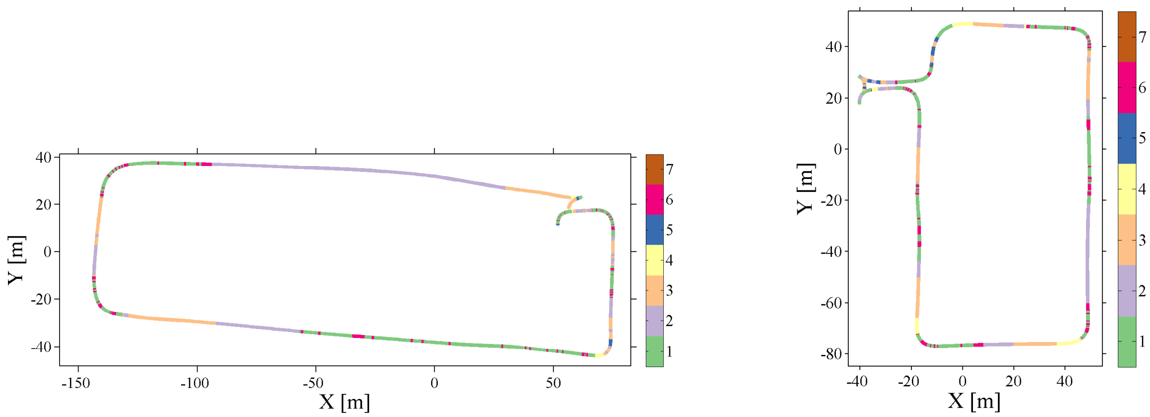
where $\otimes$ denotes the element-wise vector product, and κ is a hyperparameter that adds additional mass to self-transition probability. Let $\boldsymbol{y}_t^{(i)}$ denote observable variable of time series i at time t , and $z_t^{(i)}$ latent state. If we assume each dynamical behavior is r-order VAR process, the relation between a state and a corresponding observation can be formulated as follow:

$$z_t^{(i)} \sim \pi^{(i)}_{z_{t-1}^{(i)}} \tag{5}$$

$$\boldsymbol{y}_t^{(i)} = \sum_{m=1}^{r} \mathbf{A}_{m, z_t^{(i)}} \boldsymbol{y}_{t-m}^{(i)} + \boldsymbol{e}_t^{(i)}\left(z_t^{(i)}\right) \tag{6}$$

$$\boldsymbol{e}_t^{(i)}\left(k\right) \sim \mathcal{N}\left(\mathbf{0}, \boldsymbol{\Sigma}_k\right), \tag{7}$$

where dynamical behavior θ_k consists of $\theta_k = \{\mathbf{A}_k, \boldsymbol{\Sigma}_k\}$, and autoregressive coefficient matrix $\mathbf{A}_k = [\mathbf{A}_{1k}, \mathbf{A}_{2k}, ..., \mathbf{A}_{rk}]$. They applied matrix-normal inverse-Wishart distribution (MNIW) [15] to $\{\mathbf{A}_k, \boldsymbol{\Sigma}_k\}$ as prior distribution. The MNIW is consists of a matrix-normal distribution $\mathcal{MN}\left(\mathbf{A}_k; \mathbf{M}, \boldsymbol{\Sigma}_k, \mathbf{K}\right)$ given $\boldsymbol{\Sigma}_k$, and inverse-Wishart distribution $\mathcal{IW}\left(\mathbf{S}_0, n_0\right)$, where $\mathbf{M}, \boldsymbol{\Sigma}_k, \mathbf{K}^{-1}$ denote mean matrix, covariance matrices



(a) Course 1 (b) Course 2

Fig. 2. Course 1 and 2 with estimated state sequences using BP-AR-HMM. The subject was instructed to drive the car (a) clockwise on course 1, and (b) counterclockwise on course 2. These figures show the estimated state sequences of first lap among five laps respectively.

for column and row of $\mathbf{A}_k$, and $n_0, \mathbf{S}_0$ denote degree of freedom and scale matrix in inverse-Wishart distribution.

In this study, we assumed that the order of VAR process r is 1, and assigned parameters of matrix-normal inverse-Wishart distribution $\mathbf{M} = \mathbf{0}, \mathbf{K} = 0.1\mathbf{I}_d, n_0 = d + 2$, and $\mathbf{S}_0$ product of 0.75 and covariance matrix of observed data. We estimated the other parameters with Markov chain Monte Carlo (MCMC) sampling method, utilizing sum-product algorithm [16] and reversible jump MCMC [17]. In addition, Fox et al. applied gamma-distributed priors to γ, κ and enabled to estimate these parameters through MCMC sampling [10]. The code of BP-AR-HMM developed by Fox is available on [18].

2.3. Measurement of driving behavior data

In this study, we measured data in a real road environment. A subject was a 35 year-old male who drove on a daily basis, and we instructed him to drive our experimental car along the two courses (Fig. 2), and to make a stop after every lap he went around the course. The total number of laps are five for each course respectively. During the experiment there are other cars than ours around and people occasionally walked across the road. We attached sensors to the experimental car, so we could measure gas pedal opening rate, brake pressure and steering angle of the car. We measured these three driving operations with sampling rate 10Hz, and concatenated them into the observation column vector $\boldsymbol{y}_t^{(i)}$. We have already confirmed the correspondence between the estimated state sequences obtained from applying BP-AR-HMM to our time series data and the locations of the car on the courses, which is consistent across laps (Fig. 2) [19]. Fig. 2 (a) and (b) show the state sequences drawn on the course 1 and 2 respectively.

3. RESULT

We first applied BP-AR-HMM to training time series data that are consist of four laps for each course, sum up to eight time series. And we obtained four estimated state sequences and transition matrices of states for each course, as well as seven VAR process parameters $\theta_k = \{\mathbf{A}_k, \mathbf{\Sigma}_k\}\,(k = 1, 2, \ldots, 7)$. Fig. 3 shows the relationship between the state sequence and three driving operations during the first lap of course 2 (Fig. 2 (b)). Each color in Fig. 3 corresponds to that in Fig. 2. Now we focus our attentions on the locations just facing left corners of the course 2, which reveal the reproducible representation of state sequence across laps. The driving state revealed state 4 (cream-color) followed by state 7 (brown) at the location just before turning left, lower left, lower right, and upper left corner of the course 2. The reason why upper right corner did not reveal such state sequence pattern, which is probably because the subject stopped the car in front of the upper right corner (Fig. 3, from 78 to 84 second) and did not stop in front of the other corners. Actually driving operations of the former and the latter differed from each other.

The VAR process coefficient matrices of state 4 and 7, $\mathbf{A}_4$ and $\mathbf{A}_7$, are:

$$\mathbf{A}_4 = \begin{bmatrix} 0.1545 & -0.0082 & 0.0001 \\ 0.0061 & 0.9975 & -0.0001 \\ -0.6746 & 17.2974 & 0.9764 \end{bmatrix} \tag{8}$$

$$\mathbf{A}_7 = \begin{bmatrix} 1.0031 & 0.0272 & 0.0004 \\ -0.0694 & 0.5159 & -0.0001 \\ 12.5130 & 0.7974 & 1.0626 \end{bmatrix}. \tag{9}$$

Each element of the column vector $\boldsymbol{y}_t^{(i)}$ represents gas pedal opening, brake pressure and steering angle in order. When the car was in front and beyond the corner turning left ex-

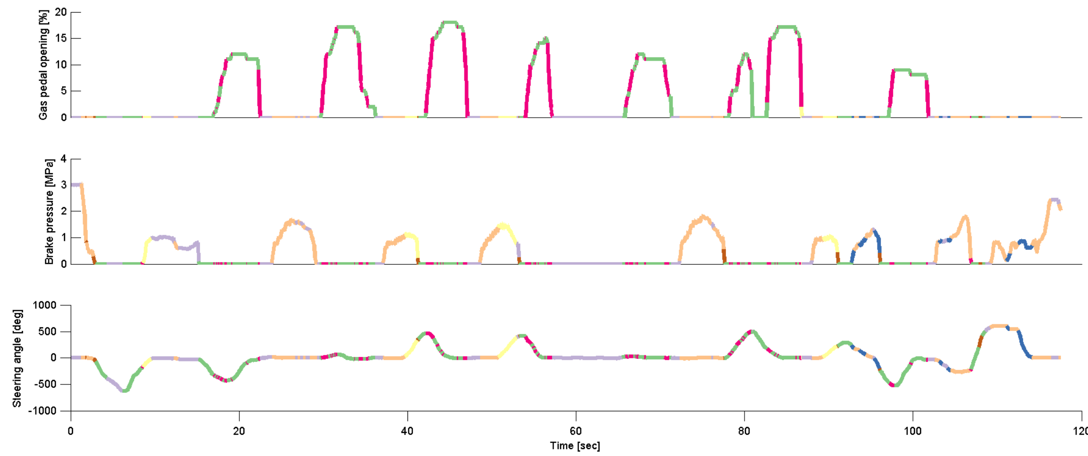


Fig. 3. Driving operations time series of first lap on the course 2. The four corners of turning left at the location of lower left, lower right, upper right and upper left correspond to at the time of 38, 51, 77 and 88 second respectively.

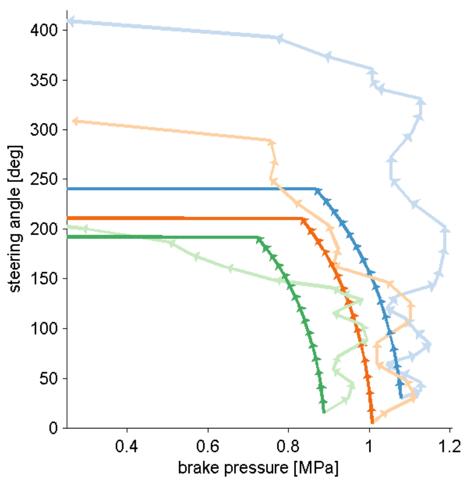


Fig. 4. Prediction of driving operations. *light-colored arrows*: actual observations of driving operations, *deep-colored arrows*: predicted driving operations. Orange, blue and green arrows are those on the course 2 at lower left, lower right and upper left corners of turning left, respectively.

cept upper right, gas pedal opening always kept 0%. And just when the state got to be fourth state, brake pressure took positive value and steering angle was 0 degree. As a result of these facts, when $\boldsymbol{y}_t^{(i)}$ was subject to VAR process of state 4 brake pressure attenuated gradually and steering angle increased progressively (around 50 second in Fig. 3).

Next we show the result of predictions of driving operations of test data, fifth lap of course 2 (Fig. 4). Light-colored and deep-colored arrows show actual observations and predicted driving operations respectively. Each arrow represents the change of driving operations during the interval of 0.1 second. Predicted driving operations almost trace the trajectory of actual observations, except the inherent fluctuation. We could predict the sudden decrease of brake pressure before turning left, although the predicted decrease occurred earlier than the actual observation on a deep-colored blue arrow. Ratios of the variance of prediction error to that throughout time series are 0, 0.164 and 0.0438 for each operation. As a consequence, we concluded that driving operations can be predicted by dynamical behaviors estimated with BP-AR-HMM.

4. DISCUSSION

In this paper, we applied BP-AR-HMM to multiple time series of driving operation data considering the common and different features among multiple data, in order to model driving behaviors. We successively tested whether the driving operation sequence of novel time series can be predicted with estimated dynamical behaviors by BP-AR-HMM, and confirmed that predicted driving operations have the same trajectories as actual data, while we did not use any information of course configuration.

In this experiment we measured the driving operation data of only one subject on two courses, and the number of time series applied to BP-AR-HMM was eight. Since we place this study preliminary, it is necessary to model the driving behaviors using long-time and much more driving data in order to reflect much more diverse driving situations. Our future work includes inspecting (i) change of the number of estimated states, (ii) relationship between estimated dynamical behaviors and actual driving operations, and (iii) difference of driving operation characteristics across multiple subjects. These inspections might give us profound knowledge in developing the novel adaptive driving support system.

5. REFERENCES

[1] Traffic Bureau of Japanese National Police Agency, "Occurences of traffic accidents in 2011," `http://www.e-stat.go.jp/SG1/estat/List.do?lid=000001086731`, September accessed on Sep. 16, 2012, *in Japanese*.

[2] S. Kitajima, Y. Marumo, T. Hiraoka, and M. Itoh, "Comparison of evaluation indices for estimating driver's risk perception of rear-end collision," *JARI Research Journal*, vol. 30, pp. 495–498, 2008.

[3] H. Mima, K. Ikeda, T. Shibata, N. Fukaya, K. Hitomi, and T. Bando, "Estimation of driving phase by modeling brake pressure signals," in *Proc. of International Conference on Neural Information Processing*, 2009.

[4] T. Kumagai and M. Akamatsu, "Prediction of human driving behavior using dynamic Bayesian networks," *IEICE Transaction on Information and Systems*, vol. 89-2, pp. 857–860, 2006.

[5] K. Ikeda, H. Mima, Y. Inoue, T. Shibata, N. Fukaya, K. Hitomi, and T. Bando, "An adaptive real-end collision warning system for drivers that estimates driving phase and selects training data," *ISCIE Journal 'Systems, Control and Information'*, vol. 24-8, pp. 193–199, 2011, *in Japanese*.

[6] A. Utsugi, S. Kato, and M. Akamatsu, "Modeling of driving behavior based on long distant driving behavior datasbase," in *Proc. of Conference on 'Systems, and Information'*. SICE, 2005, pp. 124–127, *in Japanese*.

[7] Y. Kishimoto and K. Oguri, "A modeling method for predicting driving behavior concerning with driver's past movements," in *Proc. of IEEE International Conference on Vehicular Electronics and Safety*. IEEE, 2008, pp. 132–136.

[8] K. Abe, H. Miyatake, and K. Oguri, "Driving behavior prediction based with model switching by heart rate variability," *IEICE Transaction A on Fundamentals of Electronics, Communications and Computer Sciences*, vol. 91-1, pp. 78–86, 2008, *in Japanese*.

[9] T. Taniguchi, S. Nagasaka, K. Hitomi, N.P. Chandrasiri, and T. Bando, "Semiotic prediction of driving behavior using unsupervised double articulation analyzer," in *Proc. of 2012 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2012, pp. 849–854.

[10] E.B. Fox, E.B. Sudderth, M.I. Jordan, and A.S. Willsky, "Sharing features among dynamical systems with beta processes," in *Proc. of Advances in Neural Information Processing Systems*. NIPS, 2009, vol. 22, pp. 549–557.

[11] E.B. Fox, E.B. Sudderth, M.I. Jordan, and A.S. Willsky, "An HDP-HMM for systems with state persistence," in *Proc. of International Conference on Machine Learning*. ICML, 2008, pp. 312–319.

[12] N. Hjort, C. Holmes, P. Müller, and S. Walker, *Bayesian nonparametrics*, Cambridge University Press, 2010.

[13] P. Orbanz, "Tutorials on Bayesian nonparametrics," `http://stat.columbia.edu/~porbanz/talks/npb-tutorial.html`, accessed on Sep. 14, 2012.

[14] T.L. Griffiths and Z. Ghahramani, "Infinite latent feature models and the Indian buffet process," in *Proc of Advances in Neural Information Processing Systems*. NIPS, 2005, vol. 18, pp. 475–482.

[15] M. West and J. Harrison, *Bayesian forecasting and dynamic models*, Springer Verlag, 1997.

[16] F.R. Kschischang, B.J. Frey, and H.A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Transactions on Information Theory*, vol. 47-2, pp. 498–519, 2001.

[17] P.J. Green, "Reversible jump Markov chain Monte Carlo computation and Bayesian model determination," *Biometrika*, vol. 85-4, pp. 711–732, 1995.

[18] E.B. Fox, "Homepage of Emily B. Fox," `http://stat.wharton.upenn.edu/~ebfox/software.html`, accessed on Aug. 30, 2012.

[19] T. Kubo, K. Ikeda, Z. Zhang, T. Shibata, T. Bando, and M. Egawa, "Estimation of Driving Situation from Multiple Time Series Data with Beta Process Autoregressive Hidden Markov Model," in *Symposium of Systems and Information*. SICE, 2012, *in Japanese*.