Awareness on Present and Future Trajectory of Vehicle Using Multiple Hypotheses in the Mixed Traffic of Intersection

Yunhyoung Hwang and Seibum B. Choi, *Member, IEEE*

***Abstract*— In the transition period, autonomous vehicles are mixed with unconnected traffic occupants, such as non-autonomous vehicles and pedestrians, resulting in a major hurdle toward autonomy in urban areas, especially at intersections. In this context, the cooperative-intelligent transportation system (C-ITS) affords a promising solution to achieve a breakthrough with its omniscient sensors network and computing capability. From the perspective of a C-ITS-based service, the trajectory of non-autonomous vehicle is a critical uncertainty that resides at the intersection. Therefore, this paper proposes a unique interactive framework, which is installed in the edge server of C-ITS and can estimate the present trajectories and predict the future trajectories of the non-autonomous vehicles at intersections. The proposed framework was based on multiple hypotheses of possible maneuvers that formed the confined prior set to reduce the high uncertainties posed by the complicated environment of the urban intersection. The resulting all-in-one framework provided a stable long-term trajectory prediction with intrinsic maneuver classification and improved tracking in an integrated way by incorporating the interactions between the multiple hypotheses. This situation awareness can assist autonomous vehicles to drive safely and defensively. The proposed framework was verified using a dataset collected at a real urban intersection.**

***Index Terms*— Autonomous vehicle, C-ITS, edge computing, intelligent transportation system, intersection, maneuver classification, situation awareness, trajectory prediction.**

I. Introduction

SAFETY is the key challenge for achieving the wide spread of autonomous driving. Among the various operational domains, the road intersection is the most challenging situation, because the traffic is mixed with non-autonomous vehicles and pedestrians, the traffic flows are crossed, and the line of sight of on-board sensors of autonomous vehicles is often obstructed by buildings or other traffic participants.

Autonomous driving at the intersections therefore inevitably requires the assist of cooperative-intelligent transportation system (C-ITS), which is promising given its omniscient sensor system and computing capability. In the advanced C-ITS, as demonstrated in [1], [2], multiple road-side sensors are utilized to monitor the mixed traffic at the intersection. This C-ITS with the sensor network is currently evolving toward autonomous intersections without traffic lights and thus realize fully autonomous traffic [3].

In a fully autonomous traffic, every autonomous vehicle would be connected with each other and to the supporting infrastructure. However, the problem arises in the transition period, where non-autonomous vehicles are mixed in the traffic, and several of them are "not connected." This is one of most critical uncertainties for both the autonomous vehicles and C-ITS. It has been forecasted [4] that the penetration rate of autonomous vehicle of level 4 and 5 would be below 40% in 2035 even from an optimistic outlook. Although the rate for the connected vehicle would be higher, the replacement of existing vehicles will take time.

From the perspective of C-ITS for the urban intersection, certain challenges with the non-autonomous vehicles remain. First, a typical uncertainty that needs awareness is "Where will they go?" The trajectory of the vehicles should be predicted for the path planning of nearby autonomous vehicles. In particular, it might be important to know which lane they will move toward, rather than pursue the exact prediction of future trajectories, which is intractable owing to the high level of uncertainty in the intersection environment. This criterion was the main concern of this study. Moreover, an abnormal lane change should be recognized even during the turning maneuver. In contrast to the existing advanced driver-assistance system (ADAS), such as the autonomous emergency braking (AEB), the false negative rate would be an important factor for the defensive driving assist from the infrastructure, because the auto-recognition of an abnormal lane change would be very challenging for the autonomous vehicle. Second, for the multi-sensors fusion by the server, their states should be tracked as accurately as possible. Moreover, the quality of tracking is closely related to the performance of the trajectory prediction.

In the literature, trajectory prediction can be roughly classified into motion-based and context-based approaches. In the motion-based approach, the future trajectory is predicted from the current and past motion states.

In [5], the motion states were tracked based on the multiple motion models that are associated with interactive multiple model (IMM) structure [6], and the trajectory prediction at the

Manuscript received 12 January 2021; revised 30 November 2021 and 24 February 2022; accepted 6 April 2022. Date of publication 3 May 2022; date of current version 11 October 2022. This work was supported in part by the Korea Ministry of Science, ICT and Future Planning under Grant R-20160302-003057; and in part by the Brain Korea 21 (BK 21+) Program through the National Research Foundation of Korea (NRF) by the Ministry of Education of Korea. The Associate Editor for this article was J. W. Choi. *(Corresponding author: Seibum B. Choi.)*

Yunhyoung Hwang is with the Mechanical Engineering Department, Jeonbuk National University, Jeonju 54896, South Korea (e-mail: yhh@jbnu.ac.kr).

Seibum B. Choi is with the Mechanical Engineering Department, Korea Advanced Institute of Science and Technology, Daejeon 34141, South Korea (e-mail: sbchoi@kaist.ac.kr).

Digital Object Identifier 10.1109/TITS.2022.3169030

curve was corrected with GIS information. In [7], the motion hypothesis was associated with a map-based trajectory by treating the states as pseudo-measurements. Similarly, in [8], a most probable path was prepared by clustering for the prediction, and the predicted motion was utilized for the iterative path planning of the ego vehicle. Reference [9] estimated the motion states of the ego vehicle with double Kalman filters and information from the in-vehicle network, and the estimated states were associated with the statistically obtained trajectory for motion prediction.

In another study [10], the authors used a motion history to predict the future trajectories at the intersection, which was based on the variational Gaussian mixture model (VGMM) framework. In a subsequent work [11], they proposed a hierarchical mixture of experts (HME) framework to approximate the conditional probability density function to lower the dependency on the input data. Another study [12], which was motivated from [13], proposed the particle filter-based approach that the likelihood is determined from the similarity test with the quaternion-based rotationally invariant longest common subsequence (QRLCS) metric, where the particles are sampled implicitly in the trajectories database. In addition, motion prediction in [14] used a mixture of experts (MOE) that combined the predictions from GMM-based expert that is trained with deep neural network (DNN) and additional odometry-based expert.

In case of motion-based prediction that usually associates the motion states to a certain path, there could be sharp and sudden switching between possible trajectories, making the prediction less flexible unless prepared densely as in [10] and [12]. It is the distribution of the possible trajectories (i.e., the prior distribution) that makes trajectory prediction difficult.

In the context-based approach, the prediction is usually made in the machine learning frameworks, because the context is inferred or considered as a statistical input.

References [15] and [16] predicted the trajectory from the inference with the context information, including the intent of the driver, using Bayesian Network. In [16], the authors used the intelligent driver model (IDM) [17] to model the longitudinal behavior of the vehicle. Furthermore, in [18], the results of motion- and maneuver-based prediction models were mixed in the Frenet frame. The prediction was more weighted to the motion-based model for a short prediction horizon. The trajectory was predicted even before the entrance of the intersection in [19]. This was possible because the context information, including the layout of the intersection and kinematic variables, was utilized. In [20], the hidden Markov model (HMM) was used for maneuver recognition from which the trajectory is predicted in the VGMM framework proposed in [10]. In [21], the authors demonstrated that large-scale data can improve the performance of the maneuver classification as well as trajectory prediction. In addition, a unique integrated approach that combined the maneuver classification and trajectory prediction in the GMM framework was proposed in [21], as an alternative to the MOE approach. Currently, studies based on recurrent neural networks (RNNs) are achieving substantial improvements in the field of trajectory prediction. In [22], the authors introduced the long-short term memory (LSTM) based framework for trajectory prediction in highway situations. In another study [23], the authors combined the attention mechanism to an LSTM-based framework to learn the interaction between vehicles. Reference [24] proposed the mixture density network (MDN) coupled framework featuring innovative multi-modal trajectory prediction for urban roundabouts.

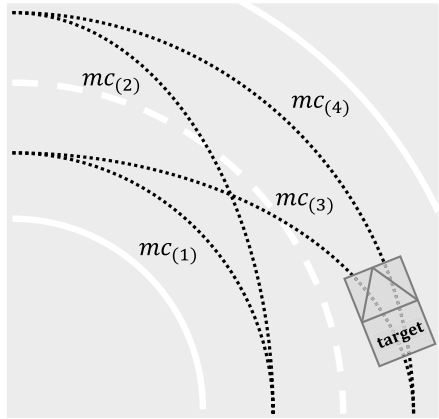

Fig. 1. Exemplary illustration of the maneuver components for the left turn, which is the case of $N = 4$ for two lanes of entry and exit. The $mc_{(i)}$ denotes the i-th maneuver component.

However, regardless of the background theory, previous works have mainly focused on trajectory prediction or maneuver classification in highway situations. However, in case of intersections, the states of vehicle change more dynamically compared to those in the highway scenario because of the added degrees of freedom; thus, the uncertainty is amplified with the layout of the roads. A simple example is the change of lane during turning. Therefore, the long-term prediction would be increasingly uncertain as the complexity of the environment increases. Thus, performance must be secured for urban intersections. Moreover, the existing decision system consists of separate tracking, maneuver classification, and trajectory prediction modules, thus complicating the structure of the system.

This study focuses on situation awareness in urban intersections, including tracking the present and predicting the long-term future, as well as classifying abnormal lane changes. These tasks are accomplished with a confined set of prior knowledge, which is utilized in the proposed framework to reduce the high levels of uncertainty at the intersection, thus making the prediction in the complicated environment more certain. Although recent relevant studies are focusing on using increasingly more data to overcome the uncertainty, it is expected that these basic attributes of the road could prevent excessive reliance on data.

The proposed framework is based on the IMM structure [6], which is widely used for target tracking and prediction [5], [25]–[28]. Unlike the popular approaches that use multiple hypotheses on various types of motion models [5], [27], [28], the proposed framework uses the multiple hypotheses on possible maneuvers coupled with a confined prior set of maneuver patterns—referred to as *maneuver component* and depicted in Fig. 1. This is a unique approach that has not been addressed

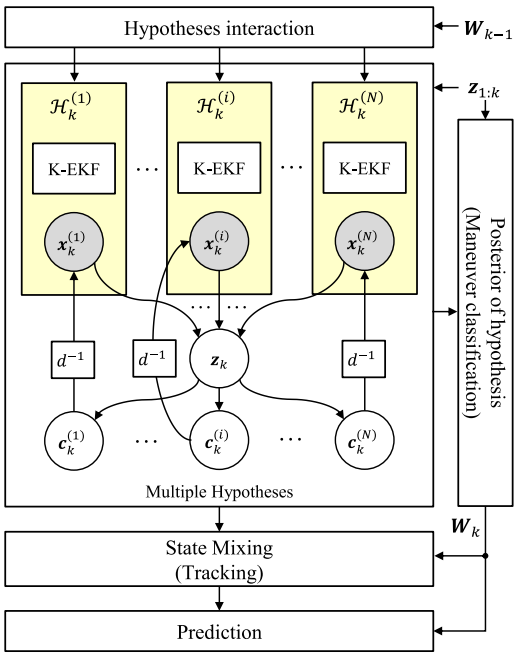

Fig. 2. Architecture of the proposed IMM-based framework. The yellow box represents each hypothesis, and gray and white boxes represent the hidden states and observations, respectively. $\boldsymbol{W}_k$ denotes the set of the posterior of multiple hypotheses; and $i \in \{1, \ldots, N\}$ denotes the index of the hypothesis, where N is the maximum value, and the k denotes the index of current time step.

in literatures, and the basic concept was motivated from the assumption that the trajectory of the vehicle can be represented as a mixture of these multiple maneuver components. Thus, the predictions can be flexible, because numerous maneuvers can be made from the mixture. The prior knowledge is defined from these few maneuver components and the transitions between them.

The main contribution of this study is as follows:

- A unique interactive framework based on IMM structure is proposed for long-term trajectory prediction at urban intersections, in which the multiple hypotheses are set for possible maneuvers, rather than motion models as conventional tracking problems. Further, experiment results specific to the urban intersection are presented, which was not much introduced in the literature. Because the prior knowledge of patterned maneuvers at a specific intersection is utilized in the proposed framework, it is expected that these basic attributes of the road could reduce excessive reliance on data.
- Moreover, the proposed framework inherently makes maneuver classification in the course of interaction, which can be utilized to detect abnormal lane changes and thus can be a practical feature for safe autonomous driving at intersections.
- Furthermore, the quality of tracking can be enhanced in a dynamic situation such as at an intersection because prior knowledge of the maneuver components serves as a feed-forward input in the context of tracking. This enhanced tracking will contribute to the improvement of the stability of the predictions. Thus, the architecture for situation awareness is simplified with the resulting all-in-one framework. In contrast to existing decision systems where each task is separately operated and thus separate learning is required, maneuver classification, tracking, and prediction are all integrated into one software. Moreover, all these operations are executed with only the black-box information that can be observed from outside the target vehicle.

The ultimate goal of this contribution was to help C-ITS or the edge-server assist safe and defensive autonomous driving in the mixed traffic of the intersection with awareness on the behavior of surrounding non-autonomous vehicles. This study specifically focuses on the left-turn case at the intersection because of the frequent conflicts occurring herein. Evidently, the proposed framework can be applied to any kind of maneuver at the intersection.

The remaining part of this article is organized as follows. Section II introduces the proposed interactive framework based on IMM structure. The concept of *maneuver component* is explained and its formulation is introduced. Section III defines the multiple hypotheses for the IMM structure, which are coupled with the maneuver component and kinematic extended Kalman filter (K-EKF) design. Subsequently, Section IV introduces the IMM structure in which multiple hypotheses interact with each other and the posterior of each maneuver component is obtained from Bayesian inference. In Section V, the quality of tracking, trajectory prediction, and the maneuver classification are synthetically evaluated with the experiment results, which is followed by a description of the conclusions drawn.

II. Proposed Framework

A. Architecture of the Framework

The proposed framework was motivated by the representation of the arbitrary trajectory as a mixture of multiple maneuver components; thus, it was implemented based on the IMM structure [6] with multiple hypotheses on the maneuver components. In addition, the maneuver classification, tracking, and trajectory prediction were conducted by virtue of the interactions between these hypotheses in an all-inclusive manner. As presented in Fig. 2, all the tasks of the architecture are conducted in the following steps:

1) Step I: Multiple maneuver components are modeled as depicted in Fig. 1.
2) Step II: A maneuver hypothesis is defined with a filter design for each maneuver component $mc_{(i)}$, that is, the i-th maneuver hypothesis for the i-th maneuver component is defined.
3) Step III: Multiple hypotheses interact with each other in the IMM structure with the posterior probabilities of each hypothesis obtained at the previous time step. Following the interaction, each hypothesis generates a-priori and a-posteriori state estimates with a property $\boldsymbol{c}_{k-1}^{(i)}$ of the coupled maneuver component $mc_{(i)}$, which is a function of observation $\boldsymbol{z}_k$ and characterizes each hypothesis.
4) Step IV (Maneuver classification): An intrinsic maneuver-classification is conducted when evaluating the posterior probability of each hypothesis using the

observations and the a-priori state estimates, which can be defined as follows:

$$p(\mathcal{H}_k^{(i)} \mid \boldsymbol{z}_{1:k}) \tag{1}$$

where $\mathcal{H}_k^{(i)}$ denotes the maneuver hypothesis of the current time step on $mc_{(i)}$, and $\boldsymbol{z}_{1:k}$ denotes the observation of the current time step and history.

5) Step V (Tracking): The a-posteriori state estimates from each hypothesis are mixed with the posterior probabilities of each hypothesis in (1) obtained at the current time step, thus generating a mixed track.
6) Step VI (Prediction): The future states are predicted from the posterior probabilities of each hypothesis for the prediction horizon, implying that the future is predicted from the current combination of maneuver hypotheses.

B. Definition of Maneuver Component

The maneuver component is a representative maneuver or a path at the intersection and forms the basis for the maneuver classification, tracking, and trajectory prediction; it can be defined statistically from the dataset or simply from the GIS information. In this study, the maneuver component is defined from the dataset and is modeled with a quartic Bézier curve as follows ($M = 4$):

$$\mathbf{P}(t) = \Sigma_{i=0}^{M} b_M^{(i)}(t)\mathbf{P}_{(i)}, \{\mathbf{P}_{(i)} | i = 1, \ldots, M+1\}, \tag{2}$$

where $\mathbf{P}(t)$ denotes the point on the curve at the parameter t and $0 \leq t \leq 1$, and $\mathbf{P}_{(i)}$ denotes a control point that controls the shape of curve. The binomial coefficient $b_M^{(i)}$ is defined as follows:

$$b_M^{(i)}(t) = \binom{M}{i}(1-t)^{M-i}t^i, \binom{M}{i}\frac{M!}{i!(M-i)!}. \tag{3}$$

Bézier curve is widely used for planning the path for autonomous driving. Thus, the maneuver component defined by the Bézier curve is advantageous and useful for planning the paths of autonomous vehicles that pass through the intersection. Moreover, this would be beneficial, especially when the lane marking inside the intersection is absent or not realistic. In addition, only a few control points are delivered to the autonomous vehicles for their safe autonomous driving at the intersection. Furthermore, the Bézier curve provides simple control with the parameter t and features convenient tools to acquire a heading or curvature for a certain point on it.

In this study, the methodology suggested in [29] is utilized to model the maneuver component based on the Bézier curve. The multiple control points are tuned to determine the best fit to the trajectories in the dataset. Fig. 3 illustrates an exemplary configuration of five control points for modeling the maneuver component based on the quartic Bézier curve. The control points $\mathbf{P}_0/\mathbf{P}_1$ and $\mathbf{P}_3/\mathbf{P}_4$ are aligned with the roads, and $\mathbf{P}_2$ is placed at the intersecting point of $\overline{\mathbf{P}_0\mathbf{P}_1}$ and $\overline{\mathbf{P}_3\mathbf{P}_4}$. Finally, the distance between the control points, d_1 and d_2, are tuned by adjusting the control points $\mathbf{P}_1$ and $\mathbf{P}_3$ from the dataset. The constructed maneuver components for the left turn at the real intersection are represented subsequent sections detailing the experiment results.

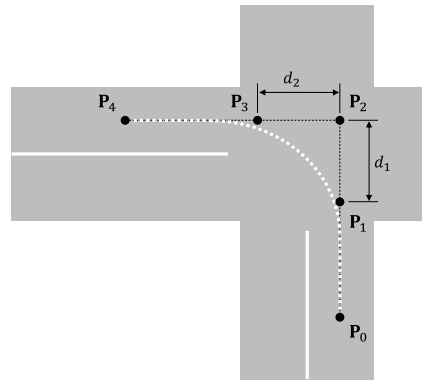

Fig. 3. Exemplary illustration of constructing a maneuver component for the left turn based on the Bézier curve. There are five control points, because it is quartic. A white dotted line represents the modeled maneuver component.

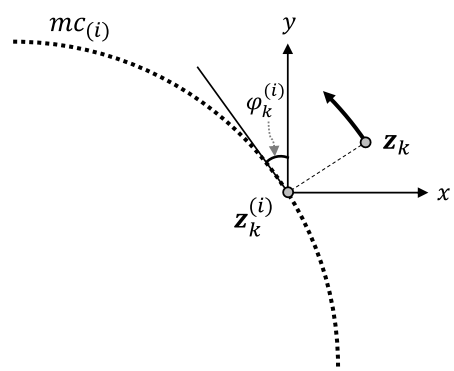

Fig. 4. Definition of maneuver hypothesis and the property $\boldsymbol{c}_k^{(i)}$. The bold arrow denotes the i-th hypothesis $\mathcal{H}_k^{(i)}$ from which the states evolve with the property of $\boldsymbol{c}_k^{(i)}$ on the $mc_{(i)}$.

III. Maneuver Hypothesis

A. Definition

A maneuver hypothesis defines that the target vehicle moves with the property of a specific maneuver component. In the proposed framework, each i-th hypothesis $\mathcal{H}_k^{(i)}$ generates the a-posteriori state estimate as follows:

$$p(\boldsymbol{x}_k^{(i)} | \boldsymbol{z}_{1:k}, \mathcal{H}_k^{(i)}), \tag{4}$$

where $\boldsymbol{x}_k^{(i)}$ is the state estimate from the i-th hypothesis, and is implemented in the following form:

$$p(\boldsymbol{x}_k^{(i)} | \boldsymbol{z}_{1:k}, \boldsymbol{c}_k^{(i)}), \;\; \boldsymbol{c}_k^{(i)} = (\varphi_k^{(i)}, \kappa_k^{(i)}). \tag{5}$$

As shown in (5), the property $\boldsymbol{c}_k^{(i)}$ is composed of the heading angle φ and curvature κ. As depicted in Fig. 4, $\boldsymbol{c}_k^{(i)}$ is defined at the projected observation $\boldsymbol{z}_k^{(i)}$ on the $mc_{(i)}$. In addition, the maneuver hypothesis is modeled with an EKF based on the kinematic motion model, as will be discussed later.

B. Basic Consideration

The sensors—such as the radar [1], lidar, and camera [2]—of the C-ITS are installed in fixed locations that guarantee

an unobstructed sight for perception of the mixed traffic. As multiple sensors of complementary types and field of views are fused with each other, it was verified that the position of the target vehicle could be observed with fairly high accuracy [30]. Moreover, only the black-box information is available, which can be observed from outside the vehicle.

Therefore, the observation for the filter design comprises only the position, and the kinematic model assumes a point mass for the motion model, because the targets for the situation awareness in this study are the non-autonomous vehicles without connectivity. In addition, property $\boldsymbol{c}_k^{(i)}$, which is a function of observation z_k, becomes an input to the model that characterizes each hypothesis. Definition of maneuver hypothesis and the property $\boldsymbol{c}_k^{(i)}$. The bold arrow denotes the i-th hypothesis $\mathcal{H}_k^{(i)}$ from which the states evolve with the property of $\boldsymbol{c}_k^{(i)}$ on the $mc_{(i)}$.

C. Motion Model

The motion model of the hypothesis is based on the constant-turn-rate-and-velocity (CTRV) kinematic model for simplicity. Evidently, the motion model can be extended to higher orders such as the constant acceleration models. There are two types of CTRV model that have been widely used in the literature. First, the CTRV I model is based on the calculus of trigonometric functions [28], [31], [32] and provides more accurate results, but it is disadvantageous for straight driving cases where the denominator becomes zero. Therefore, a special treatment is required to resolve this problem for the configuration of multiple models. Second, the CTRV II model can provide relatively less accurate results because it is based on geometric approximation [33], [34]. However, the CTRV II model poses no risk of indeterminate results as compared to the CTRV I model. Moreover, the error from the geometric approximation would be insignificant for the motion at the intersection.

Thus, the CTRV II model is selected as the basis of the kinematic model in this study, i.e., the motion model for each hypothesis is designed based on the CTRV II model. The property $\boldsymbol{c}_k^{(i)}$ from each maneuver component is arranged as the input that characterizes each hypothesis $\mathcal{H}_k^{(i)}$ as discussed in the basic consideration. The states of the model for $\mathcal{H}_k^{(i)}$ are defined as follows:

$$\boldsymbol{x}_{k-1}^{(i)} = [x_{k-1}^{(i)}\ y_{k-1}^{(i)}\ v_{k-1}^{(i)}]^T = [x_1^{(i)}\ x_2^{(i)}\ x_3^{(i)}]^T, \tag{6}$$

where x and y denote the positions for each coordinate axis, and v denotes the velocity. The input is defined from $\boldsymbol{c}_k^{(i)}$ as follows:

$$\boldsymbol{u}_k^{(i)} = \boldsymbol{c}_{k-1}^{(i)} = [\varphi_{k-1}^{(i)}\ \kappa_{k-1}^{(i)}]^T, \tag{7}$$

Thereafter, the state equation for the motion model is arranged as follows:

$$\boldsymbol{x}_k^{(i)} = \boldsymbol{x}_{k-1}^{(i)} + \Delta T \begin{bmatrix} -x_3^{(i)}\sin(u_{1,k}^{(i)} + \frac{1}{2}u_{2,k}^{(i)}x_3^{(i)}\Delta T) \\ x_3^{(i)}\cos(u_{1,k}^{(i)} + \frac{1}{2}u_{2,k}^{(i)}x_3^{(i)}\Delta T) \\ 0 \end{bmatrix}. \tag{8}$$

Furthermore, (8) can be represented as the following non-linear equation for the prediction model of the filter:

$$\boldsymbol{x}_k^{(i)} = f(\boldsymbol{x}_{k-1}^{(i)}, \boldsymbol{u}_k^{(i)}, \boldsymbol{w}_u, \boldsymbol{w}_a), \tag{9}$$

where $\boldsymbol{w}_u$ and $\boldsymbol{w}_a$ are the additive noises for the input and the $x_3^{(i)}$ state, respectively. These noises are assumed to be uncorrelated and Gaussian as follows:

$$\boldsymbol{w}_u \sim \mathcal{N}(0, \boldsymbol{Q}_u),\ \ \boldsymbol{Q}_u = \begin{bmatrix} \sigma_\psi^2 & 0 \\ 0 & \sigma_\kappa^2 \end{bmatrix}, \tag{10a}$$

$$\boldsymbol{w}_a \sim \mathcal{N}(0, \boldsymbol{Q}_a),\ \ \boldsymbol{Q}_a = \sigma_a^2 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{10b}$$

and

$$\boldsymbol{v} \sim \mathcal{N}(0, \boldsymbol{R}),\ \ \boldsymbol{R} = \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix} \tag{10c}$$

where $\boldsymbol{v}$ is the observation noise. Although the observation noise can affect the uncertainty in the curvature input, the stated effect would be less if the path around the projected observation $z_k^{(i)}$ is assumed to be locally circular.

D. Filter Design

The filter is designed in the EKF framework because of the nonlinearity of the motion model as represented in (9). Consequently, the state–space model in (8) is linearized with Jacobian transformations as follows:

$$\begin{aligned} \boldsymbol{F}_k^{(i)} &= \frac{\partial f}{\partial \boldsymbol{x}_{k-1}}|_{\hat{\boldsymbol{x}}_{k-1|k-1}^{(i)}} \\ &= \begin{bmatrix} 1 & 0 & -\Delta T sin\hat{\phi}_k^{(i)} - \frac{1}{2}u_{2,k}^{(i)}x_3^{(i)}\Delta T^2 cos\hat{\phi}_k^{(i)} \\ 0 & 1 & \Delta T cos\hat{\phi}_k^{(i)} - \frac{1}{2}u_{2,k}^{(i)}x_3^{(i)}\Delta T^2 sin\hat{\phi}_k^{(i)} \\ 0 & 0 & 1 \end{bmatrix}, \end{aligned} \tag{11}$$

$$\begin{aligned} \boldsymbol{B}_k^{(i)} &= \frac{\partial f}{\partial \boldsymbol{u}_k}|_{\hat{\boldsymbol{x}}_{k-1|k-1}^{(i)}} \\ &= \Delta T \begin{bmatrix} -x_3^{(i)} cos\hat{\phi}_k^{(i)} & -\frac{1}{2}(x_3^{(i)})^2\Delta T cos\hat{\phi}_k^{(i)} \\ -x_3^{(i)} sin\hat{\phi}_k^{(i)} & -\frac{1}{2}(x_3^{(i)})^2\Delta T sin\hat{\phi}_k^{(i)} \\ 0 & 0 \end{bmatrix}, \end{aligned} \tag{12}$$

where $\boldsymbol{F}_k^{(i)}$ and $\boldsymbol{B}_k^{(i)}$ denote the system and input matrices, respectively, and

$$\hat{\phi}_k^{(i)} = u_{1,k}^{(i)} + \frac{1}{2}u_{2,k}^{(i)}x_3^{(i)}\Delta T. \tag{13}$$

The a-priori state estimate $\hat{\boldsymbol{x}}_{k|k-1}^{(i)}$ is obtained from following prediction model:

$$\hat{\boldsymbol{x}}_{k|k-1}^{(i)} = f(\hat{\boldsymbol{x}}_{k-1|k-1}^{(i)}, \boldsymbol{u}_k^{(i)}), \tag{14}$$

where $\hat{\boldsymbol{x}}_{k-1|k-1}^{(i)}$ denotes a-posteriori state estimate at the previous time step. Therefore, the a-priori covariance matrix from (11)–(13) can be derived as

$$\boldsymbol{P}_{k|k-1}^{(i)} = \boldsymbol{F}_k^{(i)}\boldsymbol{P}_{k-1|k-1}^{(i)}(\boldsymbol{F}_k^{(i)})^T + \boldsymbol{Q}_a + \boldsymbol{B}_k^{(i)}\boldsymbol{Q}_u(\boldsymbol{B}_k^{(i)})^T. \tag{15}$$

The last term in (15) is induced from the uncertainties in the input as represented in (10).

For the calculation of the a-posteriori state estimate for (4), a residual for each hypothesis is defined as

$$\tilde{z}_k^{(i)} = z_k^{(i)} - \boldsymbol{H}\hat{\boldsymbol{x}}_{k|k-1}^{(i)}, \tag{16}$$

where $\tilde{z}_k^{(i)}$ denotes the residual of i-th hypothesis at current time step, and the observation matrix $\boldsymbol{H}$ is defined for the position as

$$\boldsymbol{H} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}. \tag{17}$$

Therefore, the a-posteriori state estimate $\hat{\boldsymbol{x}}_{k|k}^{(i)}$ becomes

$$\hat{\boldsymbol{x}}_{k|k}^{(i)} = \hat{\boldsymbol{x}}_{k|k-1}^{(i)} + \boldsymbol{K}_k^{(i)}\tilde{z}_k^{(i)}, \tag{18}$$

where the Kalman gain $\boldsymbol{K}_k^{(i)}$ is

$$\boldsymbol{K}_k^{(i)} = \boldsymbol{P}_{k|k-1}^{(i)}\boldsymbol{H}^T(\boldsymbol{S}_k^{(i)})^{-1}. \tag{19}$$

In (19), $\boldsymbol{S}_k^{(i)}$ is the residual covariance matrix.

$$\boldsymbol{S}_k^{(i)} = \boldsymbol{R} + \boldsymbol{H}\boldsymbol{P}_{k|k-1}^{(i)}\boldsymbol{H}^T \tag{20}$$

This residual covariance is used as the basis for maneuver classification in this study.

IV. IMM Structure

A. Hypotheses Interaction

The interactions between the hypotheses are vital to the IMM structure [6], as these interactions yield smooth transitions between multiple possible maneuvers in the proposed framework. If it is assumed that the trajectory of the vehicle has multiple internal maneuver modes, the posterior probability of state estimate can be represented as

$$p(\boldsymbol{x}_k|z_{1:k}) = \Sigma_{i=1}^N p(\boldsymbol{x}_k|z_{1:k}, \mathcal{H}_k^{(i)})p(\mathcal{H}_k^{(i)}|z_{1:k}). \tag{21}$$

This representation implements the basic idea of this study described in Section II-A, where N denotes the number of hypotheses and $p(\mathcal{H}_k^{(i)}|z_{1:k})$ denotes the posterior probability of each hypothesis that will be obtained subsequently. Moreover, the term $p(\boldsymbol{x}_k|z_{1:k}, \mathcal{H}_k^{(i)})$ corresponds to the posterior probability of the state estimate for each hypothesis in (4), where it can be represented using Bayes' theorem as follows:

$$p(\boldsymbol{x}_k|z_k, z_{1:k-1}, \mathcal{H}_k^{(i)}) = \frac{p(z_k|\boldsymbol{x}_k, \mathcal{H}_k^{(i)})p(\boldsymbol{x}_k|z_{1:k-1}, \mathcal{H}_k^{(i)})}{p(z_k|z_{1:k-1}, \mathcal{H}_k^{(i)})}, \tag{22}$$

The prediction term in this Bayes filter can be represented as the interactions with other hypotheses, incorporating the inherent transitions through the Markov process as depicted in Fig. 5:

$$\begin{aligned} &p(\boldsymbol{x}_k|z_{1:k-1}, \mathcal{H}_k^{(i)}) \\ &\quad = \Sigma_{j=1}^N p(\boldsymbol{x}_k|z_{1:k-1}, \mathcal{H}_k^{(i)}, \mathcal{H}_{k-1}^{(j)})p(\mathcal{H}_{k-1}^{(j)}|\mathcal{H}_k^{(i)}, z_{1:k-1}), \end{aligned} \tag{23}$$

In (23), $p(\mathcal{H}_{k-1}^{(j)}|\mathcal{H}_k^{(i)}, z_{1:k-1})$ is the probability of reverse transition between hypotheses for the transition $(j \rightarrow i)$,

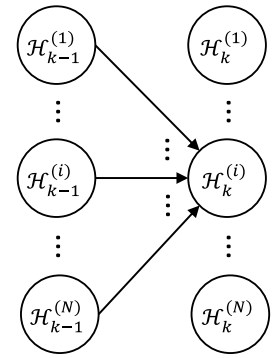

Fig. 5. Markov process during the interaction between multiple maneuver hypotheses.

which is denoted as P_{ji}^{-1}. In the IMM structure, the prediction in (23) is approximated as

$$\begin{aligned} &p(\boldsymbol{x}_k|z_{1:k-1},\ \mathcal{H}_k^{(i)}) \\ &\quad \sim N\Big(\boldsymbol{x}_k;\ E[\boldsymbol{x}_k|\mathcal{H}_k^{(i)},\ \Sigma_{j=1}^N \hat{\boldsymbol{x}}_{k-1|k-1}^{(j)} P_{ji}^{-1}],\ cov(\cdot)\Big). \end{aligned} \tag{24}$$

The right-hand side of (24) implies that the state prediction should be made from the mixture of previous a-posteriori state estimates from each hypothesis. The reverse transition probability P_{ji}^{-1} can be translated into the transition probability P_{ji} using Bayes' theorem again.

$$P_{ji}^{-1} = \frac{P_{ji}\,p(\mathcal{H}_{k-1}^{(j)}|z_{1:k-1})}{\Sigma_{j=1}^N P_{ji}\,p(\mathcal{H}_{k-1}^{(j)}|z_{1:k-1})}, \tag{25}$$

where $P_{ji} = p(\mathcal{H}_k^{(i)}|\mathcal{H}_{k-1}^{(j)}, z_{1:k-1})$ and $p(\mathcal{H}_{k-1}^{(j)}|z_{1:k-1})$ is the a-posteriori of each hypothesis at previous time steps. Further details on the derivation of the IMM framework can be obtained from [6] and [27]. Thus, the interactions between hypotheses yield the prediction in (14) at each filter as

$$\hat{\boldsymbol{x}}_{k|k-1}^{(i)} = f(\hat{\boldsymbol{x}}_{k-1|k-1}^{(i+)}, \boldsymbol{u}_k^{(i)}), \tag{26}$$

where $\hat{\boldsymbol{x}}_{k-1|k-1}^{(i+)}$ is the interacted a-posteriori state estimate from (24), described as follows:

$$\begin{aligned} \hat{\boldsymbol{x}}_{k-1|k-1}^{(i+)} &= \Sigma_{j=1}^N \hat{x}_{k-1|k-1}^{(j)} P_{ji}^{-1} \\ \boldsymbol{P}_{k-1|k-1}^{(i+)} &= \Sigma_{j=1}^N P_{ji}^{-1}\Big[\boldsymbol{P}_{k-1|k-1}^{(j)} + (\hat{\boldsymbol{x}}_{k-1|k-1}^{(j)} \\ &\qquad - \hat{\boldsymbol{x}}_{k-1|k-1}^{(i+)})(\hat{\boldsymbol{x}}_{k-1|k-1}^{(j)} - \hat{\boldsymbol{x}}_{k-1|k-1}^{(i+)})^T\Big]. \end{aligned} \tag{27}$$

$\boldsymbol{P}_{k-1|k-1}^{(i+)}$ is the interacted a-posteriori covariance [6].

B. Posterior of Hypothesis (Maneuver Classification)

The posterior probability of each hypothesis is obtained from the Bayesian inference, and this process can be interpreted as a maneuver classification. Thus, using Bayes' theorem,

$$p(\mathcal{H}_k^{(i)}|z_{1:k}) = \frac{p(z_k|\mathcal{H}_k^{(i)})p(\mathcal{H}_k^{(i)}|z_{1:k-1})}{\Sigma_{i=1}^N p(z_k|\mathcal{H}_k^{(i)})p(\mathcal{H}_k^{(i)}|z_{1:k-1})}. \tag{28}$$

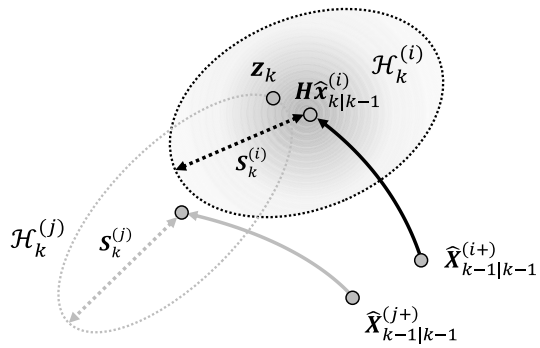

Fig. 6. Illustration of the process determining the marginal likelihood for each hypothesis in (29). The dotted curve represents the ellipse of residual covariance. In this example, the likelihood of $\mathcal{H}_{(i)}$ is greater than that of $\mathcal{H}_{(j)}$, i.e., the motion of the vehicle is more expected to follow hypothesis $\mathcal{H}_{(i)}$.

In (28), $p(\mathcal{H}^{(i)}_k|z_{1:k-1})$ indicates the prior probability that is calculated from prior knowledge. The key to calculate the posterior probability in (28) is to derive the marginal likelihood $P(z_k|\mathcal{H}^{(i)}_k)$.

Claim: The marginal likelihood for each hypothesis in (28) can be defined as

$$p(z_k|\mathcal{H}^{(i)}_k) \sim \mathcal{N}(\boldsymbol{H}\hat{\boldsymbol{x}}^{(i)}_{k|k-1},\ \boldsymbol{S}^{(i)}_k). \tag{29}$$

Proof: The likelihood of observation for each hypothesis is

$$p(z_k|\boldsymbol{x}^{(i)}_k,\ \mathcal{H}^{(i)}_k). \tag{30}$$

The marginal likelihood can be calculated by marginalizing out the states $\boldsymbol{x}^{(i)}_k$ of each hypothesis. From the total probability theorem,

$$p(z_k|\mathcal{H}^{(i)}_k) = \int p(z_k|\boldsymbol{x}^{(i)}_k, \mathcal{H}^{(i)}_k)p(\boldsymbol{x}^{(i)}_k|z_{1:k-1}, \mathcal{H}^{(i)}_k)d\boldsymbol{x}^{(i)}_k, \tag{31}$$

which can be expressed as

$$\begin{aligned} &p(z_k|\mathcal{H}^{(i)}_k) \\ &= \int p(z_k|\boldsymbol{x}^{(i)}_k, \mathcal{H}^{(i)}_k)p(\boldsymbol{x}^{(i)}_k|z_{1:k-1}, \mathcal{H}^{(i)}_k)d\boldsymbol{x}^{(i)}_k \\ &= \int \mathcal{N}(z_k;\ \boldsymbol{H}\boldsymbol{x}^{(i)}_k, \boldsymbol{R})\mathcal{N}(\boldsymbol{x}^{(i)}_k;\ \hat{\boldsymbol{x}}^{(i)}_{k|k-1}, P^{(i)}_{k|k-1})d\boldsymbol{x}^{(i)}_k. \end{aligned} \tag{32}$$

Therefore, the marginalization of normal distribution results in

$$\begin{aligned} p(z_k|\mathcal{H}^{(i)}_k) &\sim \mathcal{N}(z_k;\ \boldsymbol{H}\hat{\boldsymbol{x}}^{(i)}_{k|k-1}, \boldsymbol{R} + \boldsymbol{H}\boldsymbol{P}^{(i)}_{k|k-1}\boldsymbol{H}^T) \\ &\sim \mathcal{N}(z_k;\ \boldsymbol{H}\hat{\boldsymbol{x}}^{(i)}_{k|k-1}, \boldsymbol{S}^{(i)}_k). \end{aligned} \tag{33}$$

■

Furthermore, Fig. 6 illustrates the process of calculating the marginal likelihood from the current observation z_k with the distribution in (29).

For calculating the prior probability in (28), the prior knowledge includes the maneuver components that can be obtained from the dataset as discussed in Section II, and the following Markov transition matrix (MTM) that defines the transition probabilities between maneuver hypotheses.

$$\boldsymbol{\Lambda} \in \boldsymbol{R}^{N\times N},\ \boldsymbol{\Lambda}(j,\ i) = P_{ji}. \tag{34}$$

Although the MTM can be time-variant, it is considered time-invariant in this study to ensure simplicity. For instance, the transition probability would be reduced as the vehicle approaches the exit of the intersection. Thereafter, the prior probability in (28) can be calculated as follows:

$$\begin{aligned} &p(\mathcal{H}^{(i)}_k|z_{1:k-1}) \\ &= \Sigma^N_{j=1} p(\mathcal{H}^{(i)}_k|\mathcal{H}^{(j)}_{k-1}, z_{1:k-1})p(\mathcal{H}^{(j)}_{k-1}|z_{1:k-1}) \\ &= \mathrm{col}_i(\boldsymbol{\Lambda})^T\boldsymbol{M}_{k-1}, \end{aligned} \tag{35}$$

where $\mathrm{col}_i(\boldsymbol{\Lambda})$ denotes the i-th column vector of matrix $\boldsymbol{\Lambda}$. In (35), $\boldsymbol{M}_{k-1}$ denotes the vector of the posterior probabilities of each hypothesis at the previous time step as

$$\boldsymbol{M}_{k-1} = [p(\mathcal{H}^{(1)}_{k-1}|z_{1:k-1})\ \cdots\ p(\mathcal{H}^{(N)}_{k-1}|z_{1:k-1})]^T. \tag{36}$$

Moreover, the initial posterior probabilities of each hypothesis are defined as the following categorical distribution:

$$\begin{aligned} &p(\mathcal{H}^{(i)}_o|\boldsymbol{P}_o) = P^{(i)}_o, \\ &\qquad \boldsymbol{P}_o = \{P^{(i)}_o|\Sigma_i P^{(i)}_o = 1,\ i = 1,\ \cdots,\ N\}, \end{aligned} \tag{37}$$

where $P^{(i)}_o$ denotes the initial posterior probability of each hypothesis.

C. State Mixing (Tracking)

The a-posteriori state estimates from each hypothesis are mixed with the posterior probabilities of each hypothesis.

$$\hat{\boldsymbol{X}}^M_{k|k} = \Sigma^N_{i=1} p(\mathcal{H}^{(i)}_k|z_{1:k})\hat{\boldsymbol{X}}^{(i)}_{k|k} \tag{38}$$

In the context of tracking, all the maneuver components and posterior probabilities of each hypothesis act as feed-forward elements. Thus, $\hat{\boldsymbol{X}}^M_{k|k}$ in (38) is the mixed tracking result generated with this feed-forward information. The a-posteriori covariance of mixture $\boldsymbol{P}^M_{k|k}$ is calculated as follows [6]:

$$\begin{aligned} \boldsymbol{P}^M_{k|k} = \Sigma^N_{i=1} p(\mathcal{H}^{(i)}_k|z_{1:k})\Big[&\boldsymbol{P}^{(i)}_{k|k} + (\hat{\boldsymbol{X}}^{(i)}_{k|k} \\ &- \hat{\boldsymbol{X}}^M_{k|k})(\hat{\boldsymbol{X}}^{(i)}_{k|k} - \hat{\boldsymbol{X}}^M_{k|k})^T\Big]. \end{aligned} \tag{39}$$

D. Prediction

The proposed framework obtains predictions based on the posterior probabilities of each hypothesis calculated at the current time step. The future trajectory of the target vehicle is assumed from the current combination of hypotheses, and multiple prediction tracks are generated within the prediction horizon. This approach is useful for delivering the trajectory information in the form of waypoints. The most intuitive way will be to use the mixture in the prediction horizon as follows:

$$\hat{\boldsymbol{x}}^M_m = \Sigma^N_{i=1} p(\mathcal{H}^{(i)}_k|z_{1:k}) f(\hat{\boldsymbol{x}}^{(i)}_{m-1}, \boldsymbol{u}^{(i)}_m = \boldsymbol{c}^{(i)}_{m-1}). \tag{40}$$

$\hat{\boldsymbol{x}}^{(i)}_{m-1}$ in (40) denotes the prediction from each hypothesis at the previous prediction step, where the states in the prediction

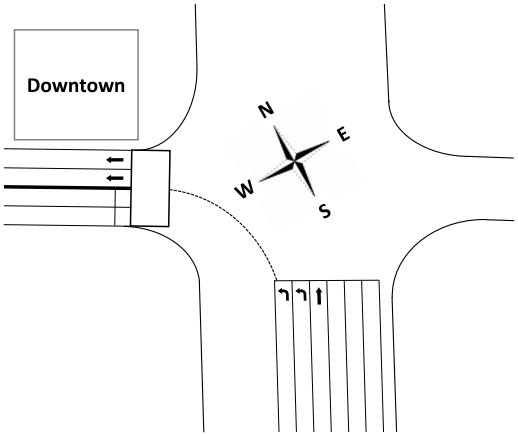

Fig. 7. Test site: Bang-I Station intersection. The data were collected for the left turn from the southwest to the northwest direction.

domain are indexed and represented with m. The prediction is initiated from the currently mixed track in (38). Therefore, for the first prediction step of $m = 1$,

$$\hat{\boldsymbol{x}}_{m-1}^{(i)}|_{m=1} = \hat{\boldsymbol{X}}_{k|k}^{M}. \tag{41}$$

$\hat{\boldsymbol{x}}_m^M$ in (40) denotes the track predicted at the m-th prediction step. As the observation is not available for $m \geq 2$, the predicted track is utilized as a pseudo-observation. Thus, the projected observation $\boldsymbol{z}_m^{(i)}$ for each hypothesis is obtained from $\hat{\boldsymbol{x}}^{M_m}$. The velocity during the prediction is assumed to be constant. The computation cost can be further reduced by using the mixture of properties to approximate the prediction in (40) instead of calculating the prediction for every hypothesis, as follows:

$$\hat{\boldsymbol{x}}_m^M = f(\hat{\boldsymbol{x}}_{m-1}^M,\ \boldsymbol{u}_m^M = \boldsymbol{c}_{m-1}^M), \tag{42}$$

where the mixture of property $\boldsymbol{c}_{m-1}^M$ is

$$\boldsymbol{c}_{m-1}^M = \Sigma_{i=1}^N p(\mathcal{H}_k^{(i)}|\boldsymbol{z}_{1:k})\boldsymbol{c}_{m-1}^{(i)}. \tag{43}$$

V. Experiment Results

A. Test Site

The proposed framework was verified with the case of left turn at the intersection. The dataset was collected at two only-left-turn lanes on the 4-way intersection of 'Bang-I Station' (37°30′31.01″N, 127°7′33.79″E, Seoul, South Korea) from southwest to northwest, as depicted in Fig. 7. The local Cartesian coordinate was set at the origin located on the bottom-left side of Fig. 7 from which the y-axis was aligned to the inlet lane.

In addition, Table I lists four classes of representative maneuvers observed during the left turn at the test lanes along with their frequency, where a certain amount of abnormal lane changes was observed. Apparently, the higher frequency of abnormal lane changes from 1st to 2nd lane indicated an effect of the downtown—located downstream—on the turning behavior of the vehicles.

TABLE I
Representative Maneuvers During Left Turn

Entry	Exit	
	1st	2nd
1st	80.3 %	19.7 %
2nd	17.8 %	82.2 %

For example, 1st (Entry) to 2nd (Exit) represents the lane change from 1st lane to 2nd lane during the left turn. The cases of 1st to 2nd and 2nd to 1st are classified as abnormal lane changes.

TABLE II
Dataset Composition

Class	Maneuver	Count	Portion
M1	1st to 1st	40	28.0 %
M2	1st to 2nd	33	23.1 %
M3	2nd to 1st	27	18.9 %
M4	2nd to 2nd	43	30.0 %

The trajectory data were collected for 3 months using an GNSS aided internal navigation system (OxTS RT3003 DGPS-RTK) with which the test vehicle was equipped. The data were collected at 100 Hz and in a forced manner, ignoring any possible bias, because naturalistic data collection requires an extensive period of time to indicate an unspecified majority and was beyond the scope of this study. The current experiment was aimed at verifying the ability of distinguishing between various representative maneuvers and highlighting the interaction between them. In this experiment, five drivers were asked to make the left turns for representing the four classes of maneuvers:

– Normal left turn (1st to 1st lane, 2nd to 2nd lane)
– Abnormal left turn (1st to 2nd lane, 2nd to 1st lane)

In total, 175 left turns were executed, and 143 cases were valid among them. The composition of dataset is represented in Table II.

B. Maneuver Component

The maneuver components were defined from four representative maneuvers in Table I ($N = 4$) and learned from the dataset in Table II. As discussed in Section II-B, a total five control points $\{\mathbf{P}_{(i)}|i = 1, \ldots, 5\}$ of the quartic Bézier curve were tuned to minimize the following root-mean-square error (RMSE):

$$RMSE\ [m] = \sqrt{\frac{\Sigma_{l=1}^{L_1} e\left(z_l, \mathbf{P}(t_l)\right)^2}{L_1}} \tag{44}$$

where $\mathbf{P}(t_l)$ is the closest point on the Bézier curve to the observed l-th track z_l of each sample data, and $e(a, b)$ denotes the Euclidean distance between a and b; L_1 denotes the number of tracks in the sample data. As depicted in Fig. 3, the distances d_1 and d_2 were iteratively tuned to minimize the RMSE in (44). The dataset of M2 maneuvers collected at the test site are depicted in Fig. 8, and Table III lists the five control points learned from the dataset for each maneuver component in the local coordinate system.

Fig. 8. Dataset of M2 maneuvers collected at the test site. The thin curves represent the collected samples, and the bold red curve denotes the learned maneuver component $mc_{(2)}$.

TABLE III
LEARNED CONTROL POINTS FOR MANEUVER COMPONENT

CP	$mc_{(1)}$		$mc_{(2)}$		$mc_{(3)}$		$mc_{(4)}$	
	x $[m]$	y $[m]$	x $[m]$	y $[m]$	x $[m]$	y $[m]$	x $[m]$	y $[m]$
$\mathbf{P}_0$	83.17	37.76	83.17	37.76	86.08	38.14	86.08	38.14
$\mathbf{P}_1$	83.45	55.25	83.55	56.47	86.56	57.13	86.68	60.37
$\mathbf{P}_2$	83.75	74.25	83.95	77.47	86.98	74.13	87.12	77.36
$\mathbf{P}_3$	64.76	75.06	62.96	78.24	70.00	74.85	70.13	77.98
$\mathbf{P}_4$	50.78	75.66	50.97	78.68	50.78	75.66	50.97	78.68

CP: control point. All values are in local Cartesian coordinate.

C. Parameters

The sample time ΔT_k was set to 0.1 s for maneuver classification and tracking. A short sample time would decrease the resolution during calculation of likelihood and make the system prone to noises in the observation, whereas a relatively long sample time would decrease the resolution of the prediction. Similarly, the sample time ΔT_m was set to 0.2 s for the prediction in this experiment. Moreover, the parameters for noises in (10) were set as per Table IV, and the MTM in (34), which was considered time-invariant in this study, was tuned as follows:

$$\mathbf{\Lambda}_1 = \begin{bmatrix} 0.8 & 0.05 & 0.05 & 0.1 \\ 0.05 & 0.75 & 0.15 & 0.05 \\ 0.05 & 0.15 & 0.75 & 0.05 \\ 0.1 & 0.05 & 0.05 & 0.8 \end{bmatrix} \tag{45}$$

In addition, the initial a-posteriori of each hypothesis in (37) was uniformly set because the hypotheses were not distinguishable at the initial stage of observation in the configuration of this experiment.

$$P_0 = \{0.25\ 0.25\ 0.25\ 0.25\} \tag{46}$$

D. Tracking Results

The tracking results with the mixture in Section IV-C are presented here. The results from two frameworks—the conventional CTRV II model-based EKF framework without multiple hypotheses and the proposed IMM-based framework—were comparatively analyzed. Consequently, the efficacy of using multiple hypotheses was verified with the experiment results from the proposed framework. Fig. 9 presents the tracking results for the M2 sample, where Fig. 9(b) depicts the calculation process of marginal likelihood from (29), as an implementation of Fig. 6. At the instant of capturing the Fig. 9(b), the likelihood for $\mathcal{H}_k^{(2)}$ and $\mathcal{H}_k^{(4)}$ had the highest value as depicted in Fig. 9(c). As represented in Fig. 9(a), the generated mixture track accurately followed the observation.

TABLE IV
NOISE PARAMETERS

Parameter	Unit	Value
σ_x	[m]	0.30
σ_y	[m]	0.30
σ_ψ	[rad]	0.2
σ_κ	$[\mathrm{m}^{-1}]$	2
σ_a	$[\mathrm{ms}^{-2}]$	1

TABLE V
AVERAGED RMSE VALUES FOR TRACKING OF POSITION AND VELOCITY

RMSE	Model	M1	M2	M3	M4
Position (norm.)	CTRV II (a)	0.488	0.613	0.889	0.485
	CTRV II-IMM (b)	0.457	0.415	0.469	0.448
Velocity [km/h]	CTRV II (a)	1.401	0.890	1.198	1.164
	CTRV II-IMM (b)	1.411	0.910	1.193	1.178

CTRV II-IMM denotes the proposed IMM based approach.

Table V represents the overall tracking results for all the 143 samples, where the normalized RMSE was used for evaluating the position tracking as follows:

$$RMSE(norm.) = (\sqrt{\sigma_x^2 + \sigma_y^2})^{-1} \sqrt{\frac{\Sigma_{k=1}^{L_2} e(z_k, \boldsymbol{H}\hat{\boldsymbol{x}}_{k|k}^M)^2}{L_2}}, \tag{47}$$

where z_k denotes the actual observation at each time step, and L_2 is the number of tracks. Although there were no remarkable differences between the two models during the mild curvature change for maneuver classes of M1 and M4, the proposed IMM-based framework showed significant improvements during the highly dynamic curvature change for maneuver classes M2 and M3. This improvement was attributed to the prior set of maneuver components, which acted as a feed-forward element in the context of tracking, as discussed in Section IV-C. The larger value of normalized tracking error could possibly result in a poorer classification performance.

E. Trajectory Prediction Results

The accuracy of the long-term trajectory prediction was determined using the mean-absolute-error (MAE). For evaluation in terms of travel direction, the Euclidean distance between the predicted track $\boldsymbol{H}\hat{\boldsymbol{x}}_m^M$ in (42) and the closest point

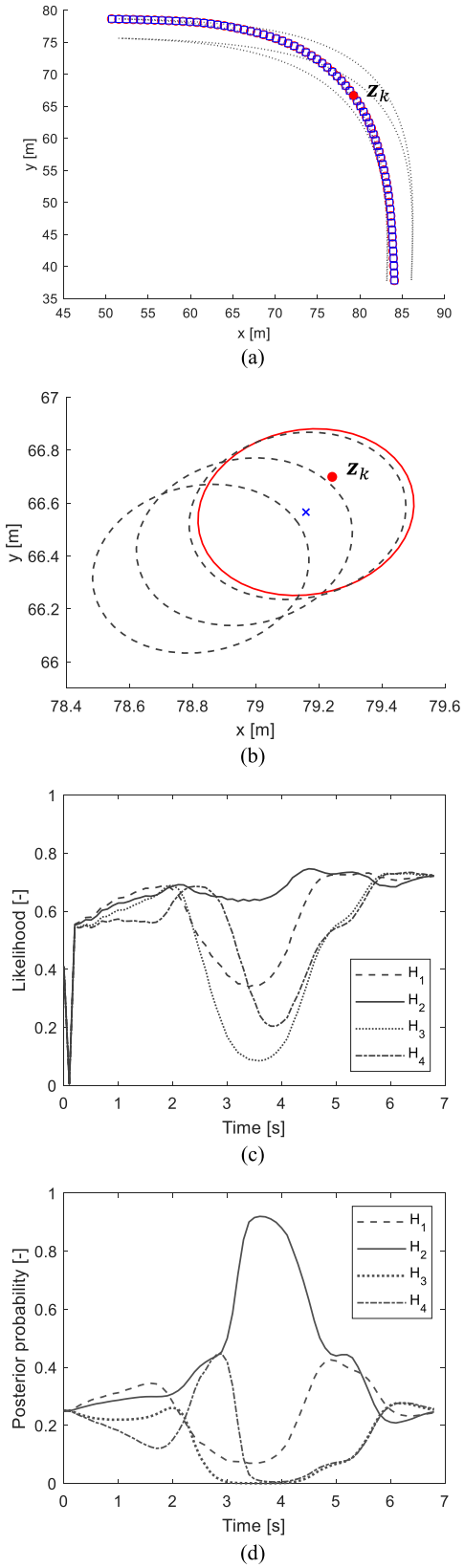

Fig. 9. Tracking results for the M2 sample. (a) Tracking results of all observations. Red circle: observations, z_k: observation at 3 s, blue square: mixture track, (b) Calculation of marginal likelihood at 3 s. The red and dashed curves represent the ellipse of $\mathcal{H}_k^{(2)}$ and the ellipses of other $\mathcal{H}_k^{(i)}$, respectively; z_k: observation, blue cross: a-priori position of $\mathcal{H}_k^{(2)}$; (c) Calculated marginal likelihood, (d) Calculated a-posteriori of each hypothesis.

TABLE VI

MAE VALUES FOR TRAJECTORY PREDICTION AT t_h

Framework	M1 $[m]$	M2 $[m]$	M3 $[m]$	M4 $[m]$	Total $[m]$
Proposed	**1.117**	**0.916**	**0.828**	**0.722**	**0.897**
LSTM encoder decoder [35]	1.745	1.229	1.054	0.862	1.230
RNN-FF [24]	2.173	1.506	1.497	2.339	1.942
GMM [10]	2.482	1.792	2.151	1.949	2.098

The LSTM encoder decoder features the input data of xy-position, heading, and velocity and was trained from xy-position. The input sequence contains the past trajectory of 3 s, and the hidden size was set to 100. In this configuration, a total of 88,573 snippets were generated. The parameters of other models were set the same as in the literatures, and data were split into training (70%) and test (30%) sets randomly.

to $\boldsymbol{H}\hat{\boldsymbol{x}}_m^M$ on the line fitted for observations was evaluated for each prediction, and the errors were averaged for all samples. This can be regarded as a metric that measures the error from lateral displacement. The prediction horizon was set to $t_h = 5\ s$. If the predicted instant exceeded the instant of final observation, the horizon was limited to the instant of final observation at the exit of the intersection. The MAE values were calculated for the predictions at t_h, which were predicted at $t_o + 3\ s$; t_o denotes the first instant of observation.

Table VI presents the results of trajectory prediction for all the samples. This result of trajectory prediction, which is dedicated to maneuvers at the intersection, is a unique part of our study. For the comparative analysis, we conducted the experiments using three other machine-learning-based approaches to which our dataset is applicable, including the LSTM-based one that is currently popular in the field of trajectory prediction. As shown in Table VI, the proposed framework outperformed the other machine-learning-based approaches for all the types of maneuvers in terms of the prediction accuracy. This might be owing to the high level of uncertainty in the maneuvers at intersection, and it was verified that the prior knowledge of maneuver patterns can be an effective solution to counteracting the uncertainty at a specific intersection. Overall, the accuracy for the M1 and M2 samples was relatively worse because two maneuvers are not distinctive of each other at the early stage of the turn. Notably, a certain level of prediction error existed for the M1 and M4 samples that followed normal lane driving without lane change, because the turning itself creates a larger uncertainty as compared to the cases of highway driving. Because the dataset was not large-scale, it might be improved for other machine-learning-based models with a lot more data for a specific location, but it would be very costly because there are a lot of urban intersections of different dimensions and layouts. In other words, the proposed framework is able to provide accurate and stable trajectory predictions with a relatively small amount of data and an easy initialization process.

Moreover, Fig. 10 represents the sequential prediction results for several samples, where the accuracy increased with time owing to the clearer distinction between the maneuver components. Similarly, the samples of M3 class exhibited better results than those of M2. As shown in Fig. 10, the prediction was stable during the entire horizon and indicated one of the benefits and strengths of the proposed framework.

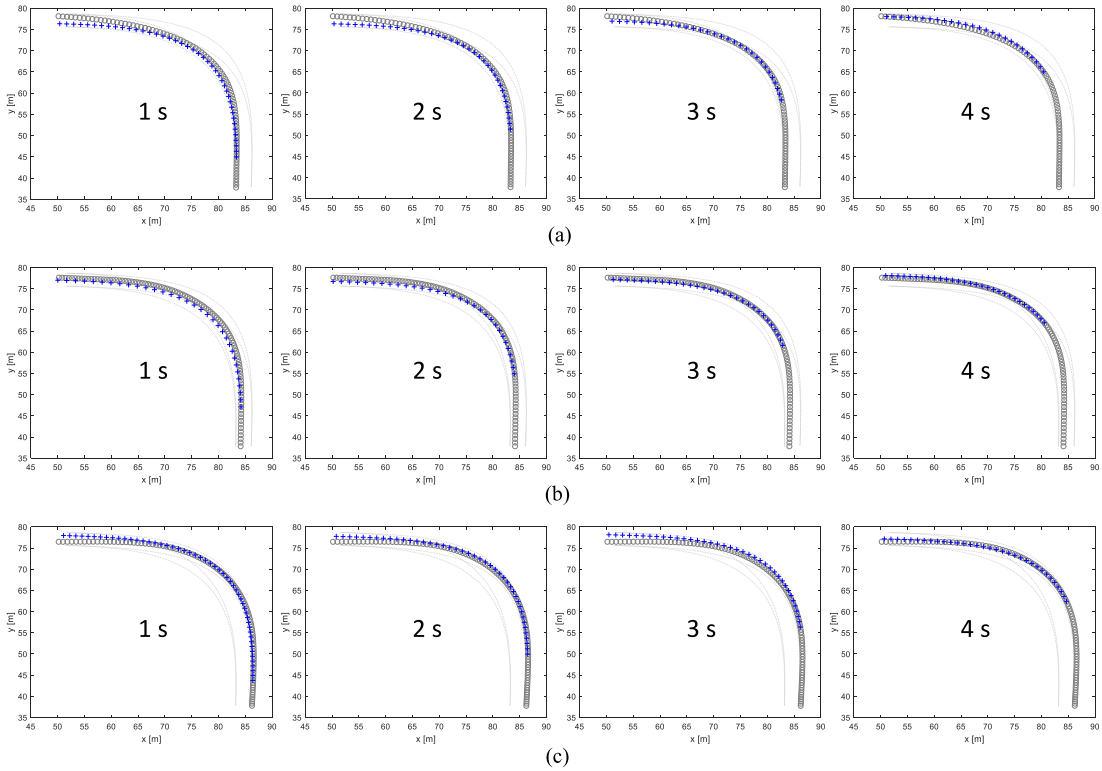

Fig. 10. Results of trajectory prediction per time step until 4 s from the start of the observation. (a), (b): M2 samples, (c): M3 sample. The blue plus sign, circle, and dotted curves represent the predicted tracks, actual observations, and maneuver components, respectively.

In addition, a sensitivity analysis was conducted to verify the stability of trajectory prediction by varying the values of noise parameters and MTM, which are the representative parameters of the proposed model. A previous study [27] demonstrated that the tracking performance is not sensitive to the variation of MTM in the IMM-based approach. To verify such a characteristic of the IMM-based framework, the accuracy of the trajectory prediction was investigated for the following three additional variants of MTM, in which the value of each element was arbitrarily set.

$$\mathbf{\Lambda}_2 = \begin{bmatrix} 0.6 & 0.2 & 0.1 & 0.1 \\ 0.2 & 0.6 & 0.1 & 0.1 \\ 0.1 & 0.1 & 0.6 & 0.2 \\ 0.1 & 0.1 & 0.2 & 0.6 \end{bmatrix}$$
$$\mathbf{\Lambda}_3 = \begin{bmatrix} 0.7 & 0.3 & 0 & 0 \\ 0.3 & 0.69 & 0.01 & 0 \\ 0 & 0.01 & 0.69 & 0.3 \\ 0 & 0 & 0.3 & 0.7 \end{bmatrix}$$
$$\mathbf{\Lambda}_4 = \begin{bmatrix} 0.9 & 0.1 & 0 & 0 \\ 0.1 & 0.89 & 0.01 & 0 \\ 0 & 0.01 & 0.89 & 0.1 \\ 0 & 0 & 0.1 & 0.9 \end{bmatrix} \tag{48}$$

As shown in Table VII, the prediction accuracy for M1 and M4 was enhanced for larger values of the diagonal elements in MTM, whereas the accuracy for M2 and M3 was enhanced for larger values of the transition elements in MTM. However, there was no significant difference in accuracy between the variants of MTM, which corroborates the discussion in the literature [27]. In the case of noise parameters, we conducted additional experiments with various noise levels at the observed position for each coordinate axis. Because the relevant parameters were set in accordance with the actual observation noise levels, we applied non-trivial additional noises of various levels to the observations in the test samples. During these additional tests, MTM was set to $\mathbf{\Lambda}_1$. Although the overall accuracy became slightly worse, it has been shown that the proposed framework is fairly robust to additional observation noises, as shown in Table VIII. This would be

TABLE VII
MAE VALUES FOR TRAJECTORY PREDICTION AT t_h FOR VARIOUS MTM

MTM	M1 $[m]$	M2 $[m]$	M3 $[m]$	M4 $[m]$	Total $[m]$
$\mathbf{\Lambda}_1$	1.117	**0.916**	**0.828**	0.722	**0.897**
$\mathbf{\Lambda}_2$	1.112	0.967	0.851	0.733	0.915
$\mathbf{\Lambda}_3$	1.214	1.004	0.843	0.780	0.965
$\mathbf{\Lambda}_4$	**1.002**	0.993	1.057	**0.675**	0.912

TABLE VIII
MAE VALUES FOR TRAJECTORY PREDICTION AT t_h FOR VARIOUS NOISE LEVELS IN OBSERVATION

$\sigma_{a,x} = \sigma_{a,y}$ $[m]$	M1 $[m]$	M2 $[m]$	M3 $[m]$	M4 $[m]$	Total $[m]$
0	**1.117**	0.916	**0.828**	**0.722**	**0.897**
0.1	1.274	**0.911**	0.863	1.033	1.040
0.2	1.284	0.943	0.838	1.086	1.062

$\sigma_{a,x}$ and $\sigma_{a,y}$ denote the standard deviations of additional noise applied to the observation of the x and y positions, respectively.

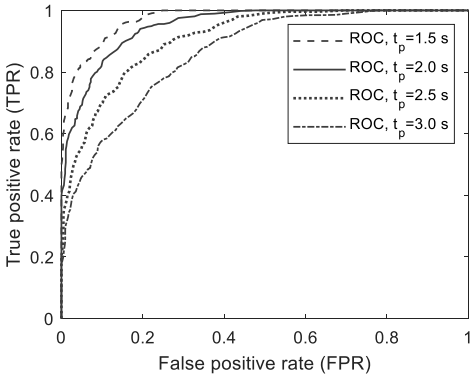

Fig. 11. The ROC curves of maneuver classification for various t_p.

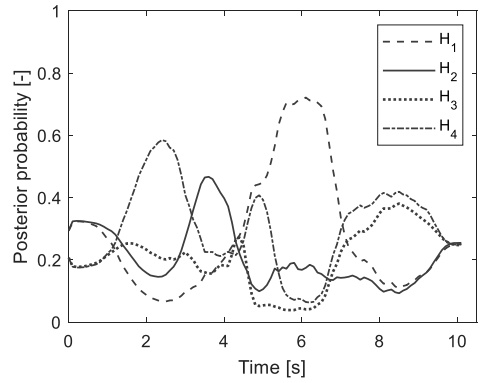

Fig. 12. Posterior probability of hypothesis for the sample in Fig. 11(b).

TABLE IX
AUC FOR MANEUVER CLASSIFICATION

t_p	AUC
1.5 s	0.973
2.0 s	0.951
2.5 s	0.910
3.0 s	0.862

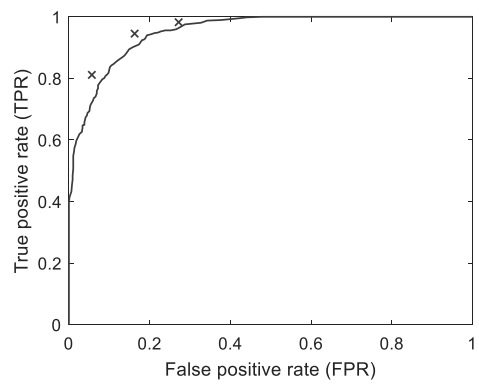

Fig. 13. Maneuver classification by the trajectory prediction for t_p = 2 s. The solid curve denotes the ROC curve, and the three crosses denote the three operating points of classification by the trajectory prediction.

the benefit from the filtering structure. It is interesting that the accuracy for M2 even improved with additional noises, and it seems that the enlarged noise parameter worked in a positive way for the trajectory prediction of the M2 samples.

F. Maneuver Classification Results

The main objective of maneuver classification is to detect the abnormal lane change by the target vehicle, which is crucial for the path planning of the autonomous vehicle. The proposed IMM-based framework can fulfill this objective with the posterior probability of the hypothesis described in (28) for $\mathcal{H}_k^{(2)}$ and $\mathcal{H}_k^{(3)}$ corresponding to the M2 and M3 cases, respectively. The quality of classification was evaluated with the receiver operating characteristic (ROC) curve, which is a popular measure for evaluating classification problems. The approach stated in [36] was used to meet the criteria of lane change and labeling. The sample was labeled as *positive* in case the lane change actually occurred after t_p, and the ROC was plotted for the time span of $[t_{lc} - t_p \ t_{lc}]$, where t_{lc} denotes the instant of lane change. Moreover, numerous intersections including the test site do not have lane lines following the turning. Thus, the t_{lc} was set to the instant at which the observation crossed the virtual lane line set at a point that laterally deviated d_{lc} from the maneuver component $mc_{(1)}$ or $mc_{(4)}$. For example, the M2 sample was labeled as *positive* if the observation crossed a point that deviated d_{lc} from $mc_{(4)}$ after t_p. The d_{lc} was set to 1.5 m in the analysis.

The ROC curves for various values of $t_p[s] = [1.5\ 2.0\ 2.5\ 3.0]$ are charted in Fig. 11, and the values of area under curve (AUC) are listed in Table IX. The results were remarkable because they were comparable to the results for the cases of highway driving as summarized in [21]. The AUC dropped below 0.9 when $t_p = 3.0$ s, because the lane-changing sub-maneuver was not clearly distinguished from the main maneuver of turning at the intersection; this creates the main difficulty for the prediction at intersection. As stated in the Introduction and observed from Fig. 11, a certain amount of false-positive rate is inevitable when focusing on the reduction of the false negative rate for the safety of autonomous vehicles. This negotiation could be regarded as defensive driving by the machine.

The lane change during turning can occur with lower components of $mc_{(2)}$ or $mc_{(3)}$. The M2 sample presented in Fig. 10(b) indicates this case. As presented in Fig. 12, the posterior probability of $\mathcal{H}_k^{(2)}$ was not prominent throughout the observations for this sample. In such cases, the maneuvers can be deterministically classified using the trajectory prediction presented in the previous section. The trajectory is considered a *positive* case when it is predicted to cross the virtual lane line. The classification results from the trajectory prediction at various operating points for $t_p = 2$ s are presented in Fig. 13, which show that the quality of classification can be improved with a deterministic approach. For example, the margin in the false negative rate increases by 9.2% for the false positive rate of about 5.6%.

VI. Conclusion

The maneuver classification, tracking, and trajectory prediction could be integrated in an all-in-one interactive framework with a confined prior set of maneuver components. This confined set of prior knowledge is vital to reduce the large uncertainties posed at the intersection, which makes the future less uncertain. The proposed framework could make smooth and stable predictions over time by virtue of the interactions between multiple hypotheses. In addition, the quality of the maneuver classification is guaranteed to improve in the inherent probabilistic framework, which can be further deterministically elaborated with the trajectory prediction. Moreover, the improved quality of tracking would enhance the quality of classification and prediction. This awareness achieved on the present and future trajectories will potentially enable a stable support using C-ITS for the planning of autonomous vehicles in urban areas.

The limitation of our work is the computational cost when the number of maneuver components increases because the proposed framework uses multiple Kalman filters for the interactions between multiple hypotheses. The computational complexity is multiplied by the number of hypotheses.

Because other types of models including neural-network model can be integrated into the proposed interactive framework if they are combined with appropriate hypotheses and probabilistic treatments for interacting mechanism, we will further explore the possibilities of computational optimization with various configurations of interactive framework in the future works. We will also address improved harmonization between prior knowledge and data, as well as the interaction of other traffic participants especially in traffic incidents.

References

[1] J. Misener *et al.*, "Cooperative intersection collision avoidance system (CICAS): Signalized left turn assist and traffic signal adaptation," California PATH, Berkeley, CA, USA, Res. Rep. UCB-ITS-PRR-2010-20, 2010.

[2] M. Goldhammer, E. Strigel, D. Meissner, U. Brunsmann, K. Doll, and K. Dietmayer, "Cooperative multi sensor network for traffic safety applications at intersections," in *Proc. 15th IEEE ITSC*, Anchorage, AK, USA, Sep. 2012, pp. 1178–1183.

[3] K. Dresner and P. Stone, "A multiagent approach to autonomous intersection management," *J. Artif. Intell. Res.*, vol. 31, pp. 591–656, Mar. 2008.

[4] S. Trommer *et al.*, "Autonomous driving—The impact of vehicle automation on mobility behaviour," IFMO, BMW Group, Munich, Germany, Res. Project Rep., Dec. 2016. [Online]. Available: http://www.ifmo.de/publications.html?t=45

[5] C. Barrios and Y. Motai, "Improving estimation of vehicle's trajectory using the latest global positioning system with Kalman filtering," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 12, pp. 3747–3755, Dec. 2011.

[6] H. A. P. Blom and Y. Bar-Shalom, "The interacting multiple model algorithm for systems with Markovian switching coefficients," *IEEE Trans. Autom. Control*, vol. AC-33, no. 8, pp. 780–783, Aug. 1988.

[7] D. Petrich, T. Dang, D. Kasper, G. Breuel, and C. Stiller, "Map-based long term motion prediction for vehicles in traffic environments," in *Proc. 16th IEEE ITSC*, The Hague, The Netherlands, Oct. 2013, pp. 2166–2172.

[8] F. Large, D. Vasquez, T. Fraichard, and C. Laugier, "Avoiding cars and pedestrians using velocity obstacles and motion prediction," in *Proc. IEEE IV*, Parma, Italy, Jun. 2004, pp. 375–379.

[9] P. Wang, J. Wang, C.-Y. Chan, and S. Fang, "Trajectory prediction for turning vehicles at intersections by fusing vehicle dynamics and driver's future input estimation," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2602, pp. 68–77, Dec. 2016.

[10] J. Wiest, M. Höffken, U. Kreßel, and K. Dietmayer, "Probabilistic trajectory prediction with Gaussian mixture models," in *Proc. IEEE IV*, Alcala de Henares, Spain, Jun. 2012, pp. 141–146.

[11] J. Wiest, F. Kunz, U. Kreßel, and K. Dietmayer, "Incorporating categorical information for enhanced probabilistic trajectory prediction," in *12th ICMLA*, vol. 1, Miami, FL, USA, 2013, pp. 402–407.

[12] C. Hermes, C. Wohler, K. Schenk, and F. Kummert, "Long-term vehicle motion prediction," in *Proc. IEEE IV*, Xi'an, China, Jun. 2009, pp. 652–657.

[13] H. Sidenbladh, M. J. Black, and L. Sigal, "Implicit probabilistic models of human motion for synthesis and tracking," in *Proc. ECCV*. London, U.K.: Springer-Verlag, 2002, pp. 784–800.

[14] X. Huang, S. G. McGill, B. C. Williams, L. Fletcher, and G. Rosman, "Uncertainty-aware driver trajectory prediction at urban intersections," in *Proc. ICRA*, Montreal, QC, Canada, May 2019, pp. 9718–9724.

[15] S. Lefèvre, C. Laugier, and J. Ibañez-Guzmán, "Risk assessment at road intersections: Comparing intention and expectation," in *Proc. IEEE IV*, Alcala de Henares, Spain, Jun. 2012, pp. 165–171.

[16] M. Liebner, F. Klanner, M. Baumann, C. Ruhhammer, and C. Stiller, "Velocity-based driver intent inference at urban intersections in the presence of preceding vehicles," *IEEE Intell. Transp. Syst. Mag.*, vol. 5, no. 2, pp. 10–21, Summer 2013.

[17] M. Treiber, A. Hennecke, and D. Helbing, "Congested traffic states in empirical observations and microscopic simulations," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 62, no. 2, pp. 1805–1824, 2000.

[18] A. Houenou, P. Bonnifait, V. Cherfaoui, and W. Yao, "Vehicle trajectory prediction based on motion model and maneuver recognition," in *Proc. IEEE/RSJ IROS*, Tokyo, Japan, Nov. 2013, pp. 4363–4369.

[19] S. Klingelschmitt and J. Eggert, "Using context information and probabilistic classification for making extended long-term trajectory predictions," in *Proc. IEEE 18th Int. Conf. Intell. Transp. Syst.*, Sep. 2015, pp. 705–711.

[20] N. Deo, A. Rangesh, and M. M. Trivedi, "How would surround vehicles move? A unified framework for maneuver classification and motion prediction," *IEEE Trans. Intell. Vehicles*, vol. 3, no. 2, pp. 129–140, Jun. 2018.

[21] F. Wirthmuller, J. Schlechtriemen, J. Hipp, and M. Reichert, "Teaching vehicles to anticipate: A systematic study on probabilistic behavior prediction using large data sets," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 11, pp. 7129–7144, Nov. 2021.

[22] F. Altche and A. de La Fortelle, "An LSTM network for highway trajectory prediction," in *Proc. IEEE ITSC*, Yokohama, Japan, Oct. 2017, pp. 353–359.

[23] K. Messaoud, I. Yahiaoui, A. Verroust-Blondet, and F. Nashashibi, "Attention based vehicle trajectory prediction," *IEEE Trans. Intell. Vehicles*, vol. 6, no. 1, pp. 175–185, Mar. 2021.

[24] A. Zyner, S. Worrall, and E. Nebot, "Naturalistic driver intention and path prediction using recurrent neural networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 4, pp. 1584–1594, Apr. 2020.

[25] U. Orguner, T. B. Schon, and F. Gustafsson, "Improved target tracking with road network information," in *Proc. IEEE Aerosp. Conf.*, Big Sky, MT, USA, Mar. 2009, pp. 1–11.

[26] F. Caron, M. Davy, E. Duflos, and P. Vanheeghe, "Particle filtering for multisensor data fusion with switching observation models: Application to land vehicle positioning," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2703–2719, Jun. 2007.

[27] Y. Bar-Shalom, K. C. Chang, and H. A. P. Blom, "Tracking a maneuvering target using input estimation versus the interacting multiple model algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 25, no. 2, pp. 296–300, Mar. 1989.

[28] A. F. Genovese, "The interacting multiple model algorithm for accurate state estimation of maneuvering targets," *Johns Hopkins APL Tech. Dig.*, vol. 22, no. 4, pp. 614–623, 2001.

[29] M. Reed, "Intersection kinematics: A pilot study of driver turning behavior with application to pedestrian obscuration by A-pillars," UMTRI, Ann Arbor, MI, USA, Tech. Rep. UMTRI-2008-54, 2008.

[30] D. Meissner, S. Reuter, E. Strigel, and K. Dietmayer, "Intersection-based road user tracking using a classifying multiple-model PHD filter," *IEEE Intell. Transp. Syst. Mag.*, vol. 6, no. 2, pp. 21–33, Summer 2014.

[31] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*. Norwood, MA, USA: Artech House, 1999, pp. 200–213.

[32] R. Schubert, E. Richter, and G. Wanielik, "Comparison and evaluation of advanced motion models for vehicle tracking," in *Proc. 11th ICIF*, 2008, pp. 1–6.

[33] C. M. Wang, "Location estimation and uncertainty analysis for mobile robots," in *Proc. IEEE ICRA*, Philadelphia, PA, USA, Apr. 1988, pp. 1231–1235.
[34] T. D. Larsen, K. L. Hansen, N. A. Andersen, and O. Ravn, "Design of Kalman filters for mobile robots; evaluation of the kinematic and odometric approach," in *Proc. IEEE CCTA*, vol. 2, Kohala Coast, HI, USA, Aug. 1999, pp. 1021–1026.
[35] K. Cho *et al.*, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," 2014, *arXiv:1406.1078*.
[36] J. Schlechtriemen, A. Wedel, J. Hillenbrand, G. Breuel, and K.-D. Kuhnert, "A lane change detection approach using feature ranking with maximized predictive power," in *Proc. IEEE IV*, Dearborn, MI, USA, Jun. 2014, pp. 108–114.

Yunhyoung Hwang received the B.S. degree in electrical engineering from Hanyang University, Seoul, South Korea, the M.S. degree from the Graduate School of Automotive Technology, KAIST, Daejeon, South Korea, and the Ph.D. degree in mechanical engineering from KAIST in 2021. From 2008 to 2021, he was with the AI-Bigdata Research Centre, Korea Automotive Technology Institute. He is currently an Assistant Professor with the Mechanical Engineering Department, Jeonbuk National University. His current research interests include decision and control for autonomous vehicles.

Seibum B. Choi (Member, IEEE) received the B.S. degree in mechanical engineering from Seoul National University, Seoul, South Korea, the M.S. degree in mechanical engineering from KAIST, Daejeon, South Korea, and the Ph.D. degree in control from the University of California, Berkeley, CA, USA, in 1993. From 1993 to 1997, he was involved in the development of automated vehicle control systems at the Institute of Transportation Studies, University of California. In 2006, he was with TRW, Livonia, MI, USA, where he was involved in the development of advanced vehicle control systems. Since 2006, he has been with the Faculty of the Mechanical Engineering Department, KAIST. His current research interests include fuel-saving technology, vehicle dynamics and control, and active safety systems. He is a member of the American Society of Mechanical Engineers, the Society of Automotive Engineers, and the Korean Society of Automotive Engineers.