2018 IEEE International Conference on Vehicular Electronics and Safety (ICVES)
September 12-14, 2018, Madrid, Spain

Intention-aware Decision Making in Urban Lane Change Scenario for Autonomous Driving*

Weilong Song
China North Vehicle Research Institute
Beijing, China
songweilong8896@gmail.com

Bo Su
China North Vehicle Research Institute
Beijing, China
bosu@noveri.com.cn

Guangming Xiong
School of Mechanical Engineering
Beijing Institute of Technology
Beijing, China
xiongguangming@bit.edu.cn

Shengfei Li
China North Vehicle Research Institute
Beijing, China
lishengfeizdh@126.com

***Abstract*— Autonomous vehicles need to face human-driving vehicles with their uncertain intentions in dynamic urban environment. Thus it leads to a challenging decision-making problem. In this paper, we focus on solving this problem in lane driving situation including performing lane changing or lane keeping maneuvers. A general POMDP model is formulated to represent autonomous driving decision-making process, and several approximations are applied to reduce the complexity of solving POMDP model. Firstly, we proposed a maneuver-based decomposition method to represent the possible candidate policies using path and velocity profiles in policy generation process. Secondly, a deterministic machine learning model is built to recognize human-driven vehicles' driving intentions. Then, a situation prediction model is proposed to calculate the possible future actions of other vehicles considering cooperative driving behaviors. Finally, we build a multi-objective reward function to evaluation each policy. In addition, we test our methods in realistic simulation software. The experimental results show that our algorithm could perform lane keeping or lane changing maneuvers successfully.**

Keywords—Autonomous vehicle, decision making, POMDP, situation prediction

I. Introduction

During the past decades, autonomous and automated driving technology has been developed rapidly. In DARPA Urban Challenge [1], several autonomous driving technologies including perception, behavior generation and motion planning were tested. In 2011, Google released its autonomous driving platforms. Over 10,000 miles of autonomous driving for each vehicle was completed under various traffic conditions [2]. Audi tested automated vehicle (i.e., A7 piloted driving concept vehicle) from Stanford to the Consumer Electronics Show 2015 in Las Vegas for 550 miles. Tesla developed Autopilot V7 system for restricted automated driving in Model S vehicle type. With these significant progresses, autonomous/automated vehicles have shown their potential abilities to reduce the number of traffic accidents, remit driving fatigue and solve the problem of traffic congestions.

One key problem in autonomous driving domain is behavioral decision making. To generate a high-level motion decision (e.g. change to the left lane), autonomous vehicles need to interact with surrounding vehicles, consider and predict possible future motion of traffic elements. The interaction between autonomous vehicles can be solved by data sharing with the development of intelligent transportation system. However, human-driven vehicles will still be predominance in a short time. Autonomous vehicles need to understand human drivers' uncertain driving intentions and choose proper actions to behave cooperatively.

In this paper, we focus on behavioral decision making in lane driving scenario. The typical lane driving scenario is show in Fig.1, the autonomous vehicle needs to select lateral driving strategies and longitudinal driving behaviors (e.g., acceleration) to interact with other vehicles.

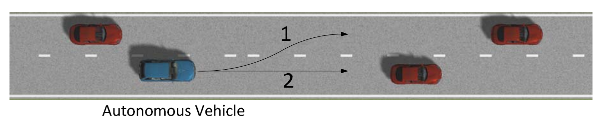

Fig. 1. An example of lane driving scenario. Autonomous vehicle has two potential strategies, path 1 representing lane change and path 2 corresponding to lane keeping. Autonomous vehicle should decide which decision is better considering the future evolution of the situation.

With these requirements, we proposed an intention-aware decision making framework. Firstly, we use *Partial Observed Decision Making Process* (POMDP) to build a general model considering other vehicles' uncertain intentions and several approximations are applied to reduce the complexity. Secondly, the policy generation algorithm is proposed to produce candidate strategies and a deterministic *Hidden Markov Model* (HMM) with *Gaussian Mixture Model* (GMM) is built to recognize other vehicle's lateral and longitudinal motion intentions. Thirdly, the interactive situation prediction model is built based on other vehicles' driving intentions and cooperative driving behaviors. Then, reward functions are designed to evaluate each strategy. In addition, we test our methods in realistic simulation software named Prescan [3]. The main contributions of this paper are:

(1) Formulating a decision making framework considering motion intentions of other vehicles in lane change scenario.

(2) Designing a decomposed policy generation algorithm considering relations between decision making and motion planning layer.

(3) Building an interactive situation prediction model considering cooperative driving behaviors.

*Research supported by the National Natural Science Foundation of China (No. 91748211). (Corresponding author: Bo Su.)

The structure of this paper is as follows. Section II reviews the related work in this research area. Section III models general decision-making process in POMDP and make suitable approximations, while the detail model for the approximation is proposed in Section IV. The experimental result is analyzed in Section V and Section VI is conclusion and future work.

II. Related Work

The most common approach is to make actions corresponding to specific situation based on manually designed rules. In DARPA Urban Challenge, the winner "Boss" used a human reasoning framework for behavior extraction based on state machine [4]. Team "Junior", ranking second in DUC, used a hierarchical state machines framework with manually defined 13 states [5]. However, this method does not deeply understand the traffic situation and the rules are manually pre-defined. Besides, surrounding vehicles' future reactions to host vehicle's actions are often ignored. As a result, the driving behavior of this method may be dangerous and not comfort for the human passengers [6].

Another approach is based on planning technics which is able to consider the evolution of future scenario. Bahram et al. proposed a predict-based reactive decision-making approach [7]. A Bayesian model is used to predict the lateral driving maneuvers of surrounding vehicles and a searching mechanism is proposed to calculate the optimal policy. However, their predict horizon is not enough for generating sequential actions such as overtaking. Wei et al. proposed a decision-making model considering autonomous vehicle social behavior in highway entrance ramp [8]. They consider the human-driven vehicles' future reactions, which is based on the yield/not-yield intentions at the first prediction step. However, their approach is focused on single lane behavior and the lane change behavior is not considered. Lenz et al. [9] used Monte Carlo Tree Search (MCTS) algorithm to perform tactical cooperative planning in autonomous highway driving scenario. However, the driving intentions of other vehicles were not considered.

POMDPs provide a mathematical framework for solving the decision making problem. Liu et.al [11] used POMDP to model autonomous driving decision making process considering situation uncertainties. Bai et al. [12] proposed an intention-aware POMDP model for autonomous driving to interact with pedestrians. Ulbrich et al. designed a POMDP-based decision making algorithm in lane change scenario [13]. Eight states are manually designed in state space to reduce the complexity. Cunningham et al. proposed a multi-policy decision-making framework [14]. With the available policy sets, a multi-vehicle simulation mechanism is built to select the best high-level policy for autonomous vehicle to execute. Hubmann et al. [15] focused on decision-making problem in dynamic and uncertain environments. POMDP are used to model the uncertainties of intentions and interactions of human-driven vehicles. However, the action space is based on a preplanned path in intersection scenario, while we focus on lane change scenario that both lateral offsets and longitudinal accelerations should be considered.

Overall, the autonomous driving decision making is still a challenging problem.

III. Problem Statement

In this section, POMDP is used to model general decision-making process for autonomous driving and several assumptions are applied to get approximate decision-making model.

A. Modeling autonomous driving decision-making process in POMDP

A POMDP model can be formulized as a tuple $\{\mathcal{S},\mathcal{A},T,\mathcal{Z},O,R,\gamma\}$. $\mathcal{S}$ is state space, $\mathcal{A}$ denotes the action space and $\mathcal{Z}$ is the observation space. The conditional function $T(s',a,s)=\Pr(s'\mid s,a)$ represents state transition probabilities to state $s'\in\mathcal{S}$ with an action $a\in\mathcal{A}$ in the state $s\in\mathcal{S}$. The observation function O models the probability of observing $z\in\mathcal{Z}$, when the state is $s'\in\mathcal{S}$ and an action $a\in\mathcal{A}$ is taken. The reward function $R(s,a)$ returns an immediate reward and $\gamma\in[0,1]$ is the discount factor.

The state space $\mathcal{S}$ contains the vehicle pose $[x,y,\theta]$ and velocity v . For the human-driven vehicles, the lateral and longitudinal intention $[I_{lat},I_{lon}]$ also needs to be contained. However, the road context knowledge is static reference information from map so that it is not added to the state space.

The joint state $s\in\mathcal{S}$ could be denoted as $s=[s_{host},s_1,s_2,...,s_N]^T$, where s_{host} is the state of host vehicle, $s_i,i\in\{1,2,3,...,N\}$ is the state of other vehicles and N is the number of other vehicles involved. Let define metric state $s_m=[x,y,\theta,v]^T$, including the position, heading angle and velocity. Thus the state of host vehicle can be defined as $s_{host}=s_{m,host}$, while the human-driven vehicle state s_i is $s_i=[s_{m,i},I_{lat,i},I_{lon,i}]^T$. With the advanced perception system, we assume that the metric state s_m could be observed. However, the intentions cannot be directly observed, so it is treated as partially observable variables.

The action space $\mathcal{A}$ is defined as a discrete set $\mathcal{A}=[LK,LC]$, because tactical decision-making system of typical autonomous vehicle is used to execute suitable discrete maneuvers. The action LK means keep in current lane, and LC means changing to the target lane. In addition, we introduce a hidden action named lane change prepare (LCP). LCP is a special action, in which autonomous vehicle follows the trajectories in current lane for a short time before performing LC. With LCP action, we can double check the situation for LC and reduce the influence of occasionally perception error to improve the consistency of strategies.

Similar to the joint state space, the observation space $\mathcal{Z}$ is denoted as $z=[z_{host},z_1,z_2,...,z_N]^T$, where z_{host} and z_i are the host and human-driven vehicles' observations, respectively. Each vehicle's observations should consist of its position, heading angle and speed to properly update surrounding vehicle's motion intentions.

In state transition process, state transition probability $\Pr(s' \mid s, a)$ can be formulized by equation (1).

$$\Pr(s' \mid s, a) = \Pr(s'_{host} \mid s_{host}, a_{host}) \prod_{i=1}^{N} \Pr(s'_i \mid s_i) \tag{1}$$

$\Pr(s'_{host} \mid s_{host}, a_{host})$ represents state update of ego vehicle, which can be represent by equation (2) given the time interval $t_{interval}$, the longitudinal acceleration a_{lon} and the differential of heading angle $\Delta\theta$. With path-velocity decomposition strategy, $\Delta\theta$ and a_{lon} can be calculated given reference path f_{traj} and candidate velocity profiles, respectively.

$$\begin{cases} x' = x + (v + a_{lon} t_{interval} / 2) t_{interval} \cos(\theta + \Delta\theta) \\ y' = y + (v + a_{lon} t_{interval} / 2) t_{interval} \sin(\theta + \Delta\theta) \\ \theta' = \theta + \Delta\theta \\ v' = v + a_{lon} t_{interval} \end{cases} \tag{2}$$

According to total probability formula, this probability $\Pr(s'_i \mid s_i)$ can be factorized as a sum in whole action space.

$$\Pr(s'_i \mid s_i) = \sum_{a_i} \Pr(s'_i \mid s_i, a_i) \Pr(a_i \mid s_i) \tag{3}$$

With the assumption that the intention I_i is updated in the observation space, I'_i is not changed in the state transition model. $\Pr(s'_i \mid s_i)$ can be reformulated as:

$$\Pr(s'_i \mid s_i) = \sum_{a_i} \Pr(s'_{i,m} \mid s_{i,m}, a_i) \Pr(a_i \mid s_{i,m}, I_i) \tag{4}$$

We should only consider the probability of other vehicle's future actions given states and intentions. Then $\Pr(s'_{i,m} \mid s_{i,m}, a_i)$ can be updated by equation (2).

The observation model O is built to simulate the measurement process. With conditional independent assumption of other vehicles' measurements, the observation model can be calculated as:

$$\Pr(z \mid a, s') = \Pr(z_{host} \mid s'_{host}) \prod_{i=1}^{N} \Pr(z_i \mid s'_i) \tag{5}$$

The host vehicle's observation function can be modeled as Gaussian distribution.

$$\Pr(z_{host} \mid s'_{host}) \sim \mathcal{N}(z_{host} \mid s'_{m,host}, \Sigma_{z_{host}}) \tag{6}$$

Other vehicles' observation probability $\Pr(z_i \mid s'_i)$ are reformulated by metric observation $z_{i,m}$ and intention observation z_{I_i}.

$$\Pr(z_i \mid s'_i) = \Pr(z_{i,m}, \mathrm{z}_{I_i} \mid s'_i) = \Pr(z_{i,m} \mid s'_i) \Pr(\mathrm{z}_{I_i} \mid s'_i) \tag{7}$$

Where the probability of metric observation $\Pr(z_{i,m} \mid s'_i)$ can be modeled as a Gaussian distribution.

$$\Pr(z_{i,m} \mid s'_i) \sim \mathcal{N}(z_{i,m} \mid s'_{i,\mathrm{m}}, \Sigma_{z_{i,m}}) \tag{8}$$

In addition, intention prediction probabilities $\Pr(\mathrm{z}_{I_i} \mid s'_i)$ can be calculated by a machine learning model which could be used in the intention update process.

The reward function R is modeled as a weighted sum format, considering and balancing several factors including safety, efficiency and so on.

$$R(s, a) = \sum_i \mu_i R_i(s, a) \tag{9}$$

B. Approximations for general POMDP model

Solving POMDP is always intractable. Thus we make several approximations to reduce the complexity of the general model. Four approximations are proposed in this process.

(1) The lateral behaviors should be consistent which means that we should avoid actions switching from LK to LC frequently. Besides, the predict horizon in decision-making process is normally less than 10s, thus the number of lateral actions is limited. With these backgrounds, we design limited candidate policy sets and policy evaluation process can be taken to select the best action in current state.

(2) We treat perception error as constant values instead of probabilistic distributions because the error is always small with advanced perception technics. In other words, it means that we only consider the intention update in observation model and the metric states are the same as the state space. The constant observation error is ignored in this process and used as for safety check in policy evaluation.

(3) We build deterministic intention prediction model to represent $\Pr(z_{i,m} \mid s'_i)$. Thus the observation model is deterministic in each state.

(4) The situation prediction model is proposed to calculate $\Pr(a_i \mid s_{i,m}, I_i)$ in state transition process. Cooperative driving behaviors are considered in this model. With this model, the most proper actions of other vehicles are calculated. Using situation prediction model and deterministic intention prediction model in (3), the belief update process is formulized as a deterministic transition dynamic.

Algorithm 1 Approximate decision-making process

Input :
Predict horizon H, time step Δt, discount factor γ
Current *metric* states : $s^0_{m,host}, s^0_{m,j}$
$\Pi \leftarrow$ GeneratePolicyset()
$I^0_j \leftarrow$ IntentionPredict($s^0_{m,host}, s^0_{m,j}$)
foreach policy $\pi_k \in \Pi$, **do**
  for $i = 1$ **to** $H / \Delta t$, **do**
    $a^{i-1}_{host} \leftarrow \pi^{i-1}_k$
    $s^i_{host} \leftarrow$ UpdateState($s^{i-1}_{host}, a^{i-1}_{host}$)
    for $j = 1$ **to** N, **do**
      $a^{i-1}_{m,j} \leftarrow$ SituationPredict($s^i_{host}, s^{i-1}_{host}, s^{i-1}_{m,j}, I^{i-1}_j$)
      $s^i_{m,j} \leftarrow$ UpdateState($s^{i-1}_{m,j}, a^{i-1}_{m,j}$)
      $I^i_j \leftarrow$ IntentionPredict($s^i_{m,host}, s^i_{m,j}$)
    end
    $R^i_k \leftarrow$ CalculateReward($s^i_{host}, s^i_{m,j}$)
  end
  $R^0_k \leftarrow$ CalculateReward($s^0_{host}, s^0_{m,j}$)
  $V^{\pi_k} \leftarrow \sum_{i=0}^{H/\Delta t} \gamma^i R^i_k$
end
$\pi^* \leftarrow \underset{\pi_k}{\mathrm{argmax}}\,(V^{\pi_k})$
return π^*

With these approximations, the POMDP model can be solved by a policy search framework with deterministic transition dynamics. Thus the best policy can be selected by policy evaluation process (**Algorithm 1**) to maximize the value function with a predict time H and time step Δt .

$$\pi^* = \underset{\pi}{\text{argmax}} \left(\sum_{i=0}^{H/\Delta t} \gamma^i R(s_i, \pi(s_i)) \right) \tag{10}$$

IV. MODELING APPROXIMATE DECISION MAKING PROCESS

In this Section, we propose the detail model for the approximations. Policy generation model is used to generate the candidate policies. HMM-based intention prediction model is proposed to recognize other vehicles' driving behavior and situation prediction model is built to calculate the possible actions for other vehicles in state transition process. In addition, the reward functions are designed for the evaluation process.

A. Policy generation

In policy generation model, we aim to get the possible policy set Π where each policy $\pi_k = [a_k^0, a_k^1, \ldots a_k^N]$ consists of actions a_k^i from action space $\mathcal{A}$, where N is equal to $H / \Delta t$. As the tactical actions *LK* and *LC* cannot be sampled to calculate the state transitions directly, we use path-velocity decomposition approach to represent each action. Each tactical action is corresponding to a reference path with several possible expected velocity profiles along with the path. Because LCP is a forced hidden action before LC, it is not considered in policy generation process.

Thus we introduce the lateral policy $\pi_{lat,k} \in \Pi_{lat}$ where $\pi_{lat,k}$ is a path corresponding to the policy π_k . Besides, the longitudinal policy sets $\pi_{lon,k}^i = [a_{lon,k}^{i,1}, a_{lon,k}^{i,2}, \ldots a_{lon,k}^{i,M}]$ are introduced to indicate the velocity profiles. The policy set $\pi_{lon,k}^i$ means the possible accelerations corresponding to the action a_k^i , where i and k are the index value.

The predict horizon H is normally less than 10s and the strategies should be consistent, so we only consider no more than three sequential actions. As is shown in Fig.2, we define motion templates for the complete maneuvers as a binary Tree with combining the same actions in three stages. Parameters $(\Delta t_1, \Delta t_2, \Delta t_3)$ with $t_{horizon} = \Delta t_1 + \Delta t_2 + \Delta t_3$ are used to divide the time for each stage. The actions do not changed in each stage due to the consistent assumption.

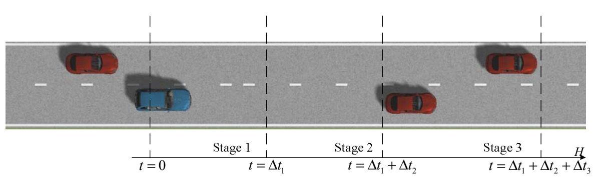

Fig. 2. Stage-based policy design.

Due to the consistent properties for the corresponding path of LK actions in the adjacent stage, such LK behaviors are combined in this process. Besides, to show the driving aggressive level, LC actions can be perform in one or more stages. The LC action in one stage (e.g. stage 1) is considered as an aggressive maneuver, while the normal LC action is defined as performing in adjacent two stages (e.g. stage 1 and stage 2) and the conservative LC action is performed in the whole prediction time.

The remaining problem for policy generation model is to design the possible path and corresponding acceleration profiles for each action and the candidate policies are the combination of such paths with acceleration profiles.

We use fifth-order polynomial trajectory [16] for each maneuver with the time Δt .The end position of the trajectory is determined by the relative distance $\Delta d = v_0 \Delta t$ in the lane direction and the middle of the target lane for the lateral direction.

The longitudinal policies for each maneuver are also divided into three time segments with the same parameters $(\Delta t_1, \Delta t_2, \Delta t_3)$. It means that keeping constant acceleration/deceleration actions in the first two segments and performing constant velocity in the third segment. To guarantee comfort, the acceleration value is limited to the range from -4 m/s^2 to 2 m/s^2 and such action is discrete into a multiple of $[-0.5, 0.5, 0]$. Thus the action space is represented by a discretizing acceleration set.

B. GMM-HMM based lateral and longitudinal intention prediction model

Based on the previous work [17], we use GMM-HMM to build the intention prediction model.

A HMM could be completely defined by hidden states *N*, a set of *M* observable symbols per state, and the probability tuples $\boldsymbol{\lambda} = (\pi, A, C, \mu, \Sigma)$. In the training process, Baum-Welch method [18] are used to estimate parameters for each motion intention *I*. Once the parameters corresponding to each motion intention have been trained, recognition process can be performed. Forward algorithm is used to compute the conditional probability $P(O \mid \boldsymbol{\lambda}_i)$ and obtain the behavior.

$$I = \arg\max_i P(O \mid \boldsymbol{\lambda}_i) \tag{12}$$

The motion intention I of other vehicles considered in this paper is divided into two aspects, lateral intention $I_{lat} \in \{I_{LCL}, I_{LK}, I_{LCR}\}$ (i.e. lane change to the left lane, keep in current lane and lane change to the right lane) and longitudinal intention $I_{lon} \in \{I_{Yield}, I_{NYield}\}$.

In order to satisfy the demand for autonomous driving, we select easy observed features as the model inputs. Besides, we only use features from other vehicles themselves to build the prediction model. Four features are examined in feature selection process including distance to the lane marker d_{lane} (defining vehicles are left of the lane marker as positive value), velocity v , relative heading angle θ_{lane} (clockwise as positive value) and yaw velocity δ .

However, the longitudinal intention reflects a cooperative behavior only occurred when interacting with other vehicles.

Based on the behavior properties $B_i = \{x_i, y_i, v_i, a_i\}$ of each vehicle, the training feature is defined as $\{\Delta D, \Delta v, \Delta a\}$, where ΔD is the relative longitudinal distance, Δv is the relative speed and Δa is relative acceleration.

Using such features, the corresponding HMMs can be trained, including $\boldsymbol{\lambda}_{LCL}$, $\boldsymbol{\lambda}_{LK}$, $\boldsymbol{\lambda}_{LCR}$ for lateral intention and $\boldsymbol{\lambda}_{Yield}$, $\boldsymbol{\lambda}_{NYield}$ for longitudinal intention.

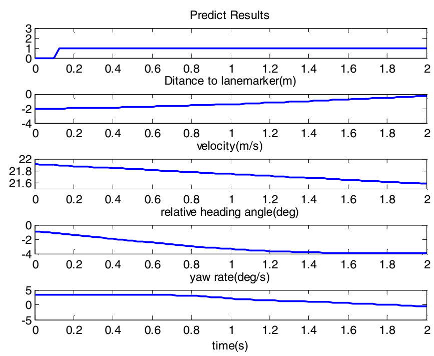

Fig. 3. Prediction result of lane change maneuver. 1 represent lane change to the left.

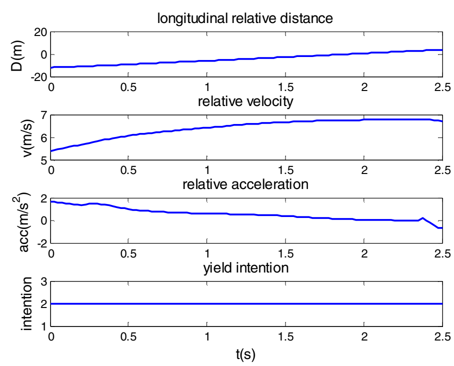

Fig. 4. Yield intention prediction result. 2 represent not yield and 1 represent yield to autonomous vehicle.

C. Situation prediction model based on cooperative driving

The situation prediction model aims to get the state of other vehicle's future actions given current states and intentions. With this process, a deterministic model is built to represent $\Pr(a_i \mid s_{i,m}, I_i)$ in state transition process. The lateral and longitudinal actions are calculated in this model. Similar to the policy generation model, the reference path is used to represent lateral behavior and the acceleration profiles are used to represent longitudinal behavior. As the reference path can be obtained by the start and end position, we only need to determine the end position to get the corresponding path.

In each predict time step, the region of interest (ROI) map is built to get the potential interaction region. The map contains the current front (CF), adjacent front (AF) and adjacent rear (AR) vehicles. The current rear (CR) does not directly affect autonomous vehicle's driving behavior so it is not in ROI.

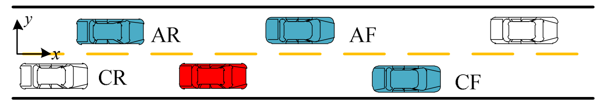

Fig. 5. A example of vehicles in ROI. The red vehicle is autonomous vehicle. The blue vehicles are the interaction vehicles in ROI region, while the white vehicle is the non-interaction vehicle.

The front vehicle of each lane in perception area is considered as keeping current speed. Then we use intelligent driver model [10] to calculate the acceleration for vehicles outside of ROI region and CF vehicle, we assume such vehicles performing LK behavior.

$$\dot{v}_\alpha = a^{(\alpha)}[1 - (\mathrm{v}_\alpha / v_0^{(\alpha)})^\delta - (s_{gap}(v_\alpha, \Delta v_\alpha)/\mathrm{s}_\alpha)^2] \tag{13}$$

Where v_α is the velocity of vehicle α. Δv_α and s_α are the velocity difference and the gap to the front vehicle, respectively. The desired gap $s^*(v_\alpha, \Delta v_\alpha)$ is as follows:

$$s_{gap}(v, \Delta v) = \mathrm{s}_0^{(\alpha)} + T^\alpha v + v\Delta v / (2\sqrt{a^{(\alpha)} b^{(\alpha)}}) \tag{14}$$

Where $v_0^{(\alpha)}$, $a^{(\alpha)}$, $b^{(\alpha)}$, $\mathrm{s}_0^{(\alpha)}$, T^α and δ are the parameters.

Then we focus on calculating the actions of vehicle AR and AF. These two vehicles are directly interacting with autonomous vehicle. If host vehicle merges into the target lane or vehicle AF merges into current lane, the passing sequence should be considered. In order to cooperative driving, the lateral intention of AF and longitudinal intention of AR are considered in this interaction process.

For autonomous vehicle, with each given policy, the future state of autonomous vehicle can be calculated by equation (2). For cooperative driving demand, the future state of vehicle AF and AR should be both no-collision and comfort.

For vehicle AR, if the longitudinal intention is yield, we can compute its action using following equations:

$$v_{IV} t_{merge} = d_1 \tag{15}$$

$$v_{AR} t_{merge} + \frac{1}{2} a_{expect} t_{merge}^2 = d_2 \tag{16}$$

$$d_1 - d_2 = d_{MSS} \tag{17}$$

$$a_{expect,1} = \frac{2v_{IV}^2 (d_1 - d_1 v_{AR} / v_{IV} - d_{MSS})}{d_1^2} \tag{18}$$

where d_{MSS} is the minimum safety distance.

We use the constraints $a_{\min,comfort} \le a_{expect} \le a_{\max,comfort}$ to ensure comfort. Although $a_{\min,comfort}$ may not avoid collision, it means that such policies cannot be selected and the policies which don't affect the AR vehicle will be selected. The actions can be calculated in equation (19).

$$a_{expect} = \begin{cases} a_{\min,comfort}, & a_{expect,1} \le a_{\min,comfort} \\ a_{\max,comfort}, & a_{expect,1} \ge a_{\max,comfort} \\ a_{expect,1}, & a_{\min,comfort} < a_{expect} < a_{\max,comfort} \end{cases} \tag{19}$$

Similarly, if the longitudinal intention of AR is not yield, the expected action is based on equation (20).

$$a_{expect}=\begin{cases}a_{\max,comfort}, & a_{expect,1}\ge a_{\max,comfort}\\ \dfrac{2v_{IV}^2(d_1+d_{MSS}-d_1v_{NR}/v_{IV})}{d_1^2}, & a_{expect}<a_{\max,comfort}\end{cases} \tag{20}$$

For vehicle AF, we first recognize its lateral behavior. If the result is LK, we calculate it based on equations (13)(14). If the intention is LC, the end point is selected as the middle point of the its AF and AN vehicles and fifth-order polynomial profile is generated to represent its future actions.

D. ROI-based reward function design.

The total reward is calculated by vehicles only in ROI to ensure consistency. A weight-sum reward model is built to computer the optimal policy.

$$\begin{aligned}R_{total}&=\mu_1R_{safety}+\mu_2R_{comfort}+\mu_3R_{time}+\mu_4R_{velocity}\\&+\mu_5R_{LK}+\mu_6R_{economy}\end{aligned} \tag{21}$$

The safety reward is calculated by the state of autonomous vehicle and the metric state of ROI vehicles in the same lane. If the host vehicle is on crossing the lane, both vehicles in two lanes are used in reward calculation process.

Time to collision (TTC) and time inter vehicle time (TIV) are used as the indicator for safety reward, which can be calculated as follows:

$$t_{TTC,i}=\frac{x_i-x_{IV}}{v_{IV}-v_i} \tag{22}$$

$$t_{TIV,i}=\begin{cases}(x_i-x_{IV})/v_{IV}, & if\ x_i>x_{IV}\\(x_{IV}-x_i)/v_i, & otherwise\end{cases} \tag{23}$$

Where x_{IV} and v_{IV} are the *x*-axis position and longitudinal velocity of autonomous vehicle and x_i and v_i is from vehicle *i*. the safety reward $R_{TTC,i}$ and $R_{TIV,i}$ to vehicle *i* can be represented by a normalization process.

$$R_{TTC/TIV,i}=\begin{cases}0, & t_{TTC/TIV,i}<t_{TTC/TIV,\min}\\1, & t_{TTC/TIV,i}>t_{TTC/TIV,\max}\\ \dfrac{t_{TTC/TIV,i}-t_{TTC/TIV,\min}}{t_{TTC/TIV,\max}-t_{TTC/TIV,\min}}, & otherwise\end{cases} \tag{24}$$

Then the weighted safety reward $R_{safe,i}$ and the final reward R_{safe} are proposed in the following equations with parameters μ_{TTC} and μ_{TIV}. The final reward R_{safe} is calculated by minimizing rewards of related vehicles.

$$R_{safe,i}=\mu_{TTC}R_{TTC,i}+\mu_{TIV}R_{TIV,i} \tag{26}$$

$$R_{safe}=\min_i R_{\mathrm{safe},i} \tag{27}$$

In special, if the TTC value is infinite due to the same speed, R_{TTC} is defined to 1. Meanwhile, if the relative distance Δd is less than minimum safe distance d_{safe}, the reward will be set to negative infinite.

$$R_{safe}=-\infty, \qquad \Delta d<d_{safe} \tag{28}$$

Comfort reward $R_{comfort}$ is based on jerk reward R_{jerk} and acceleration reward $R_{acceleration}$.

$$R_{comfort}=\mu_{jerk}R_{jerk}+\mu_{acceration}R_{acceleration} \tag{29}$$

Time reward R_{time} is computed by the velocity of autonomous vehicle and the maximum speed.

$$R_{time}=v_{IV}/v_{\max} \tag{30}$$

Velocity reward $R_{velocity}$ is used to make the action approaching the speed limit.

$$R_{velocity}=\begin{cases}1, \text{if } & |v_{policy}-v_{expect}|<\Delta v_{\min}\\0, \text{if } & |v_{policy}-v_{expect}|>\Delta v_{\max}\\1-\dfrac{|v_{policy}-v_{expect}|-\Delta v_{\min}}{\Delta v_{\max}-\Delta v_{\min}}, & otherwise\end{cases} \tag{31}$$

The gap reward R_{LK} is an indicator to maintain a suitable gap to the lead vehicle.

$$R_{LK}=\begin{cases}1, \text{ if } & |d_{current}-d_{expect}|<\Delta d_{\min}\\0, \text{ if } & |d_{current}-d_{expect}|>\Delta d_{\max}\\1-\dfrac{|d_{current}-d_{expect}|-\Delta d_{\min}}{(\Delta d_{\max}-\Delta d_{\min})}, & otherwise\end{cases} \tag{32}$$

The economy reward $R_{economy}$ is represented by fuel consumption rate $F(y)$ in the following equation.

$$R_{economy}=\begin{cases}1, & F(y)<\mathrm{E}_{\min}\\0, & F(y)>\mathrm{E}_{\max}\\1-(F(y)-\mathrm{E}_{\min})/(\mathrm{E}_{\max}-\mathrm{E}_{\min}), & otherwise\end{cases} \tag{33}$$

Where $F(y)=e^{a+ey+fy^2+gy^3}$, *y* is velocity.*a*, *e*, *f*, *g* is constant value. we select $a = -0.67944$; $e = 0.029665$; $f = -0.00028$; $g = 0.00000149$ according to the reference [19].

V. Experiment Result

In this section, we evaluate our model through Prescan software, a realistic simulation tool for autonomous driving and connected vehicles [3]. Using this software, the testing scenarios can be built and vehicles can be added with dynamic model and real world effects such as errors on localization or delays in communications. Gaussian noise is applied in the perception system to simulate observation error. Two scenarios are evaluated in this section including dealing with vehicle AF suddenly merging into current lane and active lane change situation.

A. *Facing sudden merging vehicle*

The visualizing result is shown in Fig. 6. When facing sudden merging vehicle, autonomous vehicle could recognize its LC intention and decelerate to avoid potential conflict (Fig. 6(b)(c)(d)). Then, autonomous vehicle changes to the left lane to overtake the slow vehicle for more frequent driving (Fig. 6(e)-(h)).

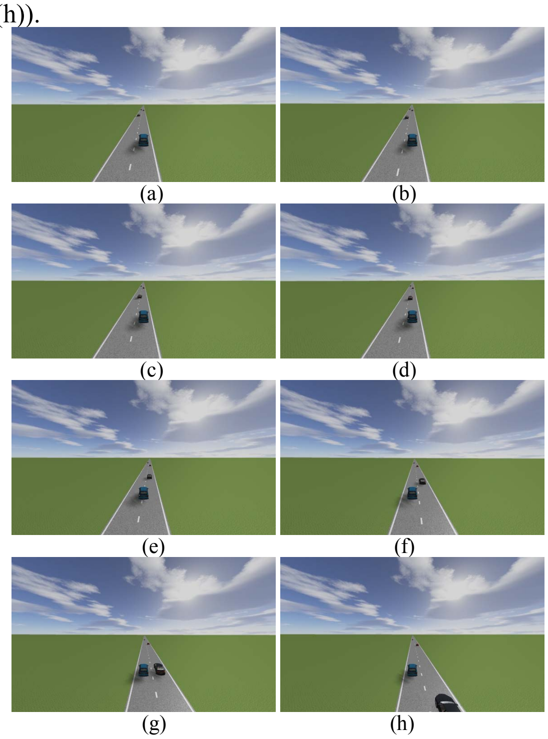

Fig. 6. Visualizing result in sudden merging scenario. Blue vehicle is the host vehicle.

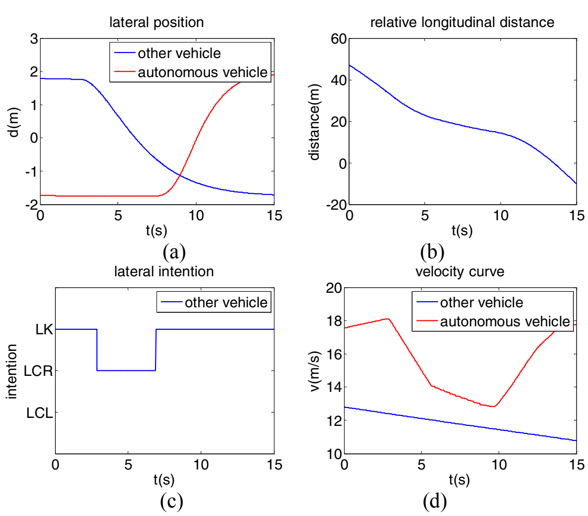

Fig. 7. Simulation result of handling suddenly merging scenario

The experiment result is shown in Fig. 7. Fig. 7(a) is the relative lateral position. Fig. 7(b) is relative longitudinal position, while Fig. 7(c) shows driving intention of vehicle AF and velocity curves are described in Fig. 7(d). It is dangerous in this situation because the AF vehicle merges into current lane in a low speed. At the time of 3 seconds, autonomous vehicle recognizes the LC behavior of other vehicle and then decelerates to avoid potential collision. It proves that our decision-making model has capability to deal with such dangerous situation.

B. *Performing active lane change maneuver*

The visualizing result in active lane change scenario is shown in figure 8. In Fig. 8(a), autonomous vehicle is driven in right lane. Due to facing slow vehicle, autonomous vehicle is prepared to change to the left lane with deceleration in Fig. 8(b). Then autonomous vehicle performs LC action to the left lane in Fig. 8(c) and 8(d). Then autonomous vehicle turns back to the right lane to complete a full overtaking behavior (8(e)-(g)).

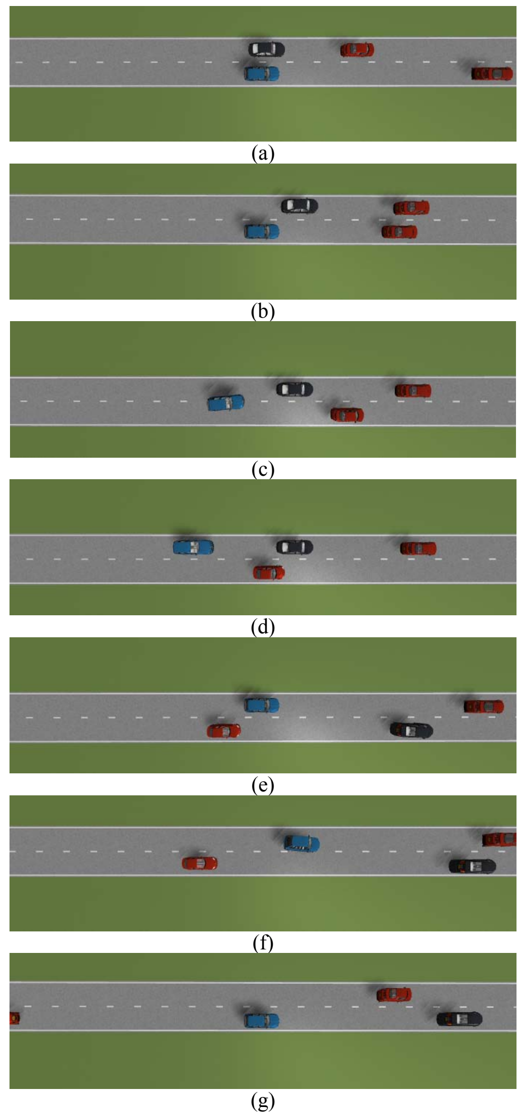

Fig. 8. Visualizing result in active lane change scenario. Blue vehicle is the host vehicle.

The simulation data is shown in Fig. 9. The speed of autonomous vehicle is shown in Fig. 9(a), faster speed is achieved after overtaking behavior which is useful to improve time efficiency. The lateral position of autonomous vehicle is shown in Fig 9(b) and the detail rewards for each action is shown in Fig. 9(c)(d). Based on the detailed rewards, the total reward is described in fig. 9(e). Finally, actions in this process are shown in Fig. 9(f).

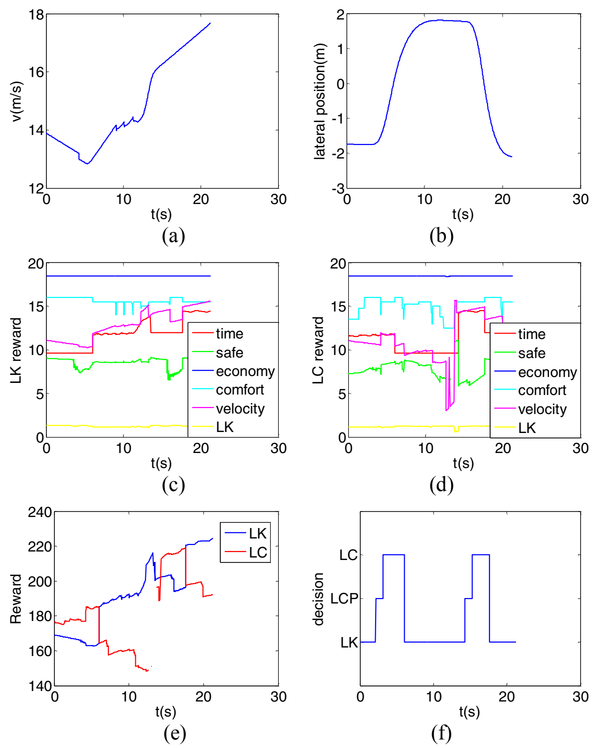

Fig. 9. The state variables in active lane change process

Overall, the simulation results in these two scenarios show that autonomous vehicle with our model can interact with other vehicles in lane driving scenario successfully and effectively.

VI. CONCLUSION

In this paper, we have proposed an intention-aware decision making framework which is useful in both autonomous vehicle and driving assistant systems. Four approximations have been used to reduce the complexity of solving general POMDP model. Stage-based policy generation model has been built to obtain the possible candidate actions. Other vehicle's lateral and longitudinal motion intentions have been modeled by GMM-HMM. Besides, situation prediction model has been built to predict the future actions of other vehicles and reward functions have been designed to evaluate each strategy. In addition, we have tested our methods in realistic simulation software. The results have shown that our model is effective in urban lane change scenario.

In the near future, we will design a more precious situation prediction model using machine learning technics. Besides, deep reinforcement learning algorithm could be tried to get more human-like driving behaviors. Furthermore, we will focus on improving our model and applying it in a real world autonomous vehicle platform.

REFERENCES

[1] C. Urmson, J. Anhalt, D. Bagnell, C. Baker, R. Bittner, M. Clark, *et al.*, "Autonomous driving in urban environments: Boss and the urban challenge," *Journal of Field Robotics,* vol. 25, pp. 425-466, 2008.

[2] J. Markoff, "Google Cars Drive Themselves, in Traffic," *New York Times,* vol. 9, 2010.

[3] M. Tideman and M. van Noort, "A simulation tool suite for developing connected vehicle systems," in *Intelligent Vehicles Symposium (IV), 2013 IEEE*, 2013, pp. 713-718.

[4] C. Urmson, J. Anhalt, D. Bagnell, C. Baker, R. Bittner, M. Clark, et al., "Autonomous driving in urban environments: Boss and the urban challenge," Journal of Field Robotics, vol. 25, pp. 425-466, 2008.

[5] M. Montemerlo, J. Becker, S. Bhat, H. Dahlkamp, D. Dolgov, S. Ettinger, et al., "Junior: The stanford entry in the urban challenge," Journal of field Robotics, vol. 25, pp. 569-597, 2008.

[6] L. Fletcher, S. Teller, E. Olson, D. Moore, Y. Kuwata, J. How, et al., "The MIT–Cornell collision and why it happened," Journal of Field Robotics, vol. 25, pp. 775-807, 2008.

[7] M. Bahram, A. Wolf, M. Aeberhard, and D. Wollherr, "A prediction-based reactive driving strategy for highly automated driving function on freeways," in Intelligent Vehicles Symposium Proceedings, 2014 IEEE, 2014, pp. 400-406.

[8] J. Wei, J. M. Dolan, and B. Litkouhi, "Autonomous vehicle social behavior for highway entrance ramp management," in Intelligent Vehicles Symposium (IV), 2013 IEEE, 2013, pp. 201-207.

[9] Lenz, David, Tobias Kessler, and Alois Knoll. "Tactical cooperative planning for autonomous highway driving using Monte-Carlo Tree Search." *Intelligent Vehicles Symposium (IV), 2016 IEEE*. IEEE, 2016.

[10] Treiber, Martin, Ansgar Hennecke, and Dirk Helbing. "Congested traffic states in empirical observations and microscopic simulations." *Physical review E* 62.2 (2000): 1805.

[11] Liu, Wei, et al. "Situation-aware decision making for autonomous driving on urban road using online POMDP." *Intelligent Vehicles Symposium (IV), 2015 IEEE*. IEEE, 2015.

[12] H. Bai, S. Cai, N. Ye, D. Hsu, and W. S. Lee, "Intention-aware online POMDP planning for autonomous driving in a crowd," in Robotics and Automation (ICRA), 2015 IEEE International Conference on, 2015, pp. 454-460.

[13] S. Ulbrich and M. Maurer, "Probabilistic online POMDP decision making for lane changes in fully automated driving," in Intelligent Transportation Systems-(ITSC), 2013 16th International IEEE Conference on, 2013, pp. 2063-2067.

[14] A. G. Cunningham, E. Galceran, R. M. Eustice, and E. Olson, "MPDM: Multipolicy decision-making in dynamic, uncertain environments for autonomous driving," in Proc. IEEE Int. Conf. Robot. and Automation, Seattle, WA, USA, 2015.

[15] Hubmann, Constantin, et al. "Automated Driving in Uncertain Environments: Planning With Interaction and Uncertain Maneuver Prediction." *IEEE Transactions on Intelligent Vehicles* 3.1 (2018): 5-17.

[16] Piazzi A, Bianco C G L. Quintic G 2-splines for trajectory planning of autonomous vehicles. IEEE Intelligent Vehicles Symposium. 2000: 198-203.

[17] Song, Weilong, Guangming Xiong, and Huiyan Chen. "Intention-aware autonomous driving decision-making in an uncontrolled intersection." *Mathematical Problems in Engineering* 2016 (2016).

[18] L. E. Baum and T. Petrie, "Statistical inference for probabilistic functions of finite state Markov chains," *The annals of mathematical statistics,* pp. 1554-1563, 1966.

[19] Wu, Changxu, Guozhen Zhao, and Bo Ou. "A fuel economy optimization system with applications in vehicles with human drivers and autonomous vehicles." Transportation Research Part D: Transport and Environment 16.7(2011):515-524.