2018 IEEE Intelligent Vehicles Symposium (IV)
Changshu, Suzhou, China, June 26-30, 2018

Toward Human-like Behavior Generation in Urban Environment based on Markov Decision Process with Hybrid Potential Maps

Chunzhao Guo, Kiyosumi Kidono, Ryuta Terashima and Yoshiko Kojima

Abstract— **It is crucial for automated vehicles to interact with the surrounding vehicles properly in urban environment that can make both passengers and surrounding drivers feel safe and at ease. In this paper, we propose a human-like behavior generation approach, which can make safe as well as efficient decisions and generate an appropriate path with the corresponding speed profile to interact with the surrounding vehicles. Firstly, a cumulative cost map is constructed by integrating the hybrid potential maps of the surrounding vehicles according to their categories. Subsequently, three candidate paths, i.e., a route-following path, an in-lane circumventing path and a collision-avoidance path, are extracted to form a Markov decision process (MDP) model. The optimal decision to cope with the current situation, e.g., when and how to perform a passing maneuver with parked vehicles in the host lane, are finally made in the MDP model such that the actual path and target speed profile will be generated. Particularly, three driving modes, i.e., *Leisure Mode*, *Normal Mode* and *Efficiency Mode*, are provided to the driver/passengers to adjust the decision making strategy such that the generated human-like automated driving commands will make them not only feel safe by the collision-free trajectories but also feel at ease by meeting their preferences and needs. Experimental results in various typical but challenging urban traffic scenes have substantiated the effectiveness of the proposed system.**

I. INTRODUCTION

Next generation cars will be highly or fully automated on regular roads. This could greatly improve the driving comfort and, more importantly, the road safety, thereby potentially saving over a million lives each year. Since nearly half of fatal crashes feature a lane or roadway departure, and most of the crashes feature more than one vehicle [1], various automated driving systems have been developed and implemented in the real world in the last decade [2], e.g., LKA (lane keep assist), ACC (adaptive cruise control), etc. These functions usually work quite well in highways, where the surrounding vehicles travel in the same direction with a relatively steady speed. Whereas, they may have problems in urban environment with daily traffic. For instance, a parked vehicle blocks the host lane while an oncoming vehicle is approaching in the neighboring lane. In this situation, the ego vehicle has to decide when and how to perform the passing maneuver. Furthermore, such a maneuver has to be not only safe but also efficient and practical, like a good human driver, so as to make the passengers or surrounding drivers feel safe and at ease.

C. Guo, K. Kidono, R. Terashima and Y. Kojima are with Toyota Central R&D Labs., Inc., 41-1, Yokomichi, Nagakute, Aichi 4801192, Japan. E-mail: {czguo, kidono, ryuta, yoshiko-k}@mosk.tytlabs.co.jp

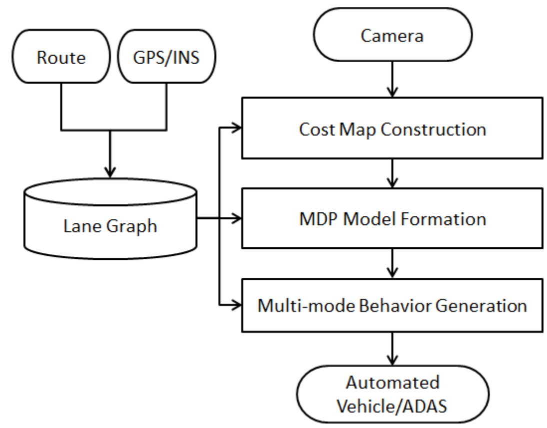

Fig. 1: Flow diagram of the proposed approach.

In this paper, we focus on the automated driving functions along a predefined route while interacting with the surrounding vehicle properly in urban environment with daily traffic. Such a system is supposed to not only avoid collisions with other objects, but also act like a good human driver. Human-like performance of automated driving systems is not just an option to make the system better, but a necessary requirement in the future mixed urban traffic. On the one hand, there will be many vehicles in the surroundings which are driven by human drivers. The human-like performance will not confuse those drivers, which is crucial to achieve harmony in the mixed traffic. On the other hand, there will be human passengers in the automated vehicles. The human-like performance will not only need to make them feel safe with collision-free trajectories but also make them feel at ease with efficient and reliable behaviors. Fig. 1 shows the flow diagram of the proposed approach. Given the route and driving lane information together with the current vehicle position, a cost map of the road is firstly constructed based on the environment perception result and lane graph information. Subsequently, a number of candidate paths are extracted from the cost map to form a Markov decision process (MDP) [3] model for the current circumstance. Finally, the optimized behavior will be generated in the MDP model according to the preferred driving mode/strategy of the passenger such that the actual path and target speed profile will be obtained.

II. RELATED WORK

Developing a behavior generation system that is sophisticated enough to operate in urban environment with daily traffic is difficult due to the road complexity. In the past decades, various safe and smooth path planning systems have been presented. Gu *et. al.* used dynamic programming

978-1-5386-4451-5/18/$31.00 ©2018 IEEE

to generate trajectories for automated driving [4]. Karaman *et. al.* employed Rapidly-exploring Random Trees (RRT) to find the best feasible path, which accounts to almost-sure convergence to global optimal solutions [5]. Guo *et. al.* generated the local trajectories by online learning from the leader vehicle [6]. However, the proper behavior of the ego vehicle to interact with the surrounding vehicles is not only determined by a or multiple good paths to go, but also involved with the decision on when and how to go due to the dynamic environment. To make decisions for an automated vehicle, various machine learning-based techniques have been adopted, such as Gaussian mixed models [7], Gaussian process regression [8], Inverse reinforcement learning [9], etc. Alternatively, the probabilistic planning formalisms have been well used for decision making, such as MDP and partially observable MDP, which can handle the uncertainty of the dynamic environment well. For example, Wei *et. al.* presented a two-step decision making that finds suitable velocity profiles by evaluating a set of candidate actions [10]. Cunningham *et. al.* used multipolicy decision making process to select the best policy for vehicle control [11]. The proposed system belongs to this line of systems. However, unlike the previous MDP-based methods which generate the candidate actions/policies based on manual tunning, the proposed approach combines machine learning techniques with the MDP-based reasoning in terms of MDP model formation and optimal behavior generation.

In summary, the proposed approach is distinguished from the previous ones in the following aspects, which are the contributions of this paper:

1) We propose a hybrid potential map, which include a trajectory-induction potential and a risk-prevention potential, accounting for the apparent obstacles and the underlying risks of the environment, respectively.

2) We propose a MDP model for the decision making of an automated driving system, which is formed based on a number of candidate paths, including a route-following path, an in-lane circumventing path and a collision-avoidance path.

3) We propose three driving modes for the automated driving system that balance between safety margin and time efficiency with respect to decision strategy. Such a scheme is able to handle the obscure or dilemma situations, e.g., following or overtaking a slow preceding vehicle, thereby meeting the needs of different passengers and different tasks.

III. PROPOSED SYSTEM

To interact with the surrounding vehicles properly, the proposed system firstly detect the on-road vehicles based on a deep neural network (DNN) and track them using extended Kalman filter (EKF). Subsequently, the detected vehicles are classified into six categories by Bayesian network (BN), i.e., *leader vehicle*, *parked vehicle*, *tail-end vehicle*, *exiting vehicle*, *merging vehicle* and *other vehicle* [12]. Meanwhile, the free road space is detected based on stereo vision[13], which will be used to ensure the ego vehicle will not hit other obstacles, such as curbs, pedestrians, cyclists, etc.

A. Cost Map Construction

To generate an appropriate behavior for the ego vehicle, a cumulative cost map must be constructed that accounts for different aspects of the road environment and the driving task. An artificial potential field offers an effective and elegant representation of the cost map. Typically, the route is assigned with low potentials and obstacles are assigned with high potentials. Nevertheless, in this work, we generate a hybrid potential map, which includes trajectory-induction potentials and risk-prevention potentials, for each detected vehicle according to its category.

1) Trajectory-induction potential map: The trajectory-induction potential accounts for the apparent objects of the environment and is learned from the real driving data. Such a potential encodes the information of the path and speed profile of human driving in the same situation. More details can be found in [12].

2) Risk-prevention potential map: One of the key differences between good drivers and poor drivers lies in the ability of predicting the underlying risks of the road environment. Such a foresight could allow early actions to be taken to prevent the ego vehicle from getting involved in the inevitable collision situations. There are two ways to improve such an ability. One is from the education, such as in the driving schools or on some driving safety related websites [14]. The other is from the experiences of driving in various traffic situations. In this work, we try to mimic the same mechanism to provide the risk prediction ability to an automated vehicle. On the one hand, multiple expert human drivers were asked to elaborate all of the major underlying risks for each of the vehicle categories with respect to what risks to expect, where to look out and how to do to prevent collisions [15]. On the other hand, the naturalistic driving data is examined to extract and compare the crash and near-crash patterns related to each of the vehicle categories. In this way, the risk-prevention potential can be learned, which accounts for the underlying risks in the environment. Fig. 2 shows the underlying risks by each vehicle category. Such a potential is used to adjust the velocity in advance to prevent the ego vehicle getting involved in a inevitable collision situation, e.g., running into a darting-out pedestrian at high speed.

More formally, the risk-prevention potential map consists of multiple speed potentials, corresponding to each of the underlying risks. At first, the potential collision points, as shown in Fig. 2, are estimated, which is the point of intersection of ego vehicle's local path and the distribution of the underlying risks. For each potential collision point, the spatial distribution of the maximum target speed is subsequently computed, which can ensure the driving safety, i.e., absolute enough braking distance, as well as driving comfort, i.e., relatively gentle deceleration once the emergency brake is necessary. More specifically, the total stopping distance of a full braking is a function of vehicle's current speed, i.e., $D_{\text{total}} = f(v)$, which can be computed as [16]

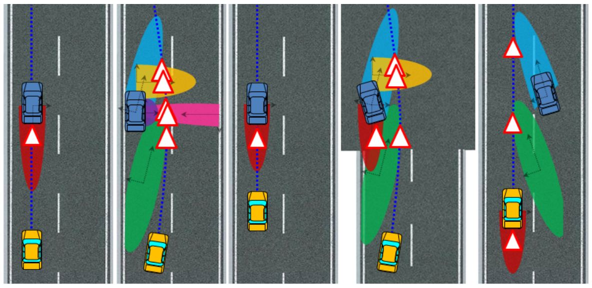

Fig. 2: Proposed risk-prevention potential maps of surrounding vehicles. Red: rear end collision due to sudden braking of the vehicle. Yellow: objects darting from behind the vehicle. Light blue: sudden start or sudden change in directions of the vehicle. Purple: sudden door opening of the vehicle. Green: bikes entering roadways to circumvent the vehicle. Magenta: pedestrians crossing the road to get in the vehicle. Blue dots: local trajectory of ego vehicle. Red triangles: potential collision points.

$$D_{\text{total}} = vt_{p-r} + \frac{v^2}{2\mu g} \tag{1}$$

where t_{p-r} is the perception-reaction time. In the proposed system, it would be the computational time of the environmental perception module. μ is coefficient of friction between the tires and the road surface. g is the gravity of Earth. Therefore, for a location which is d away from the collision point, the target speed, which is then used to construct the risk-prevention potential field, can be computed as

$$s_{\text{target}} = \gamma_v f^{-1}(d - d_{\text{safety}}) \tag{2}$$

where d_{safety} is a predefined safety margin. $f^{-1}(d - d_{\text{safety}})$ computes the maximum speed at the location, which can be derived from (1), allowing the vehicle to stop d_{safety} away from the collision point in a full braking. γ_v is a discount factor to account for early deceleration action, which will allow the ego vehicle to stop without using a full brake so as to enhance the driving comfort and make the passengers feel at ease.

B. MDP Model Formation

When we human drivers see a vehicle on the road, we do not consider how far we should keep from it, but what we should do to handle it. According to the instructions of expert human drivers and the naturalistic driving data, the following behaviors are considered as appropriate ones with respect to the vehicle categories.

- *Leader vehicle*: The ego vehicle should follow it while avoiding physical collisions with other objects.
- *Parked vehicle*: The ego vehicle should circumvent it smoothly while avoiding physical collisions with other objects.
- *Tail-end vehicle*: The ego vehicle should perform the "stop-and-go" behavior behind it while avoiding physical collisions with other objects.
- *Exiting vehicle*: The ego vehicle should keep driving behind it until it leaves the host lane. If it stops, e.g., in front of the crosswalk, the ego vehicle should circumvent it if and only if there is a feasible circumventing path in the vicinity of the host lane while avoiding physical collisions with other objects.
- *Merging vehicle*: The ego vehicle should give room for its merging action if and only if the deceleration magnitude is under a predefined threshold ν. Otherwise, the ego vehicle should drive normally and react to the merging vehicle's next movement while avoiding physical collisions with other objects.
- *Other vehicle*: The ego vehicle should drive normally in the host lane while avoiding physical collisions with it and other objects.

In summary, to interact with the leader vehicle, the tail-end vehicle, the merging vehicle and the other vehicle, the ego vehicle needs to follow the predefined route in the host lane and adjust its speed to keep a safe distance with them. To interact with the parked vehicle and the exiting vehicle, the ego vehicle may have to go beyond the boundaries of the host lane to pass them. In this case, it is crucial for the ego vehicle to decide when and how to perform the passing maneuver. Here, in this paper, we propose a MDP model for the above decision making.

MDP is a discrete time stochastic control process, which provides a mathematical framework for modeling decision making in dynamic environment. It is characterized by a set of 5 quantities, i.e., $(\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma)$. In this work, they are defined as follows,

1) $\mathcal{S}$: This is a set of states of the road environment. Conventional methods usually define it based on a grid map-based representation of the road. However, we human drivers do not make decisions of the passing maneuver over the whole road space, but switch between a number of candidate paths, corresponding to different behaviors. Therefore, in the proposed approach, we generate three candidate paths, i.e., a route-following (RF) path, an in-lane circumventing (IC) path and a collision-avoidance (CA) path. As shown in Fig. 3, the states in $\mathcal{S}$ are set as locations equally distributed on the RF and IC candidate paths with an additional state on the CA path. The RF path is aimed to keep the ego vehicle staying on the center of the host lane, or make it go back to the host lane smoothly, which is generated based on the potential map of the host lane. The IC path is aimed to circumvent the parked vehicle smoothly. It is generated based on the trajectory-induction potential map of the parked vehicle from the ego vehicle's current location to

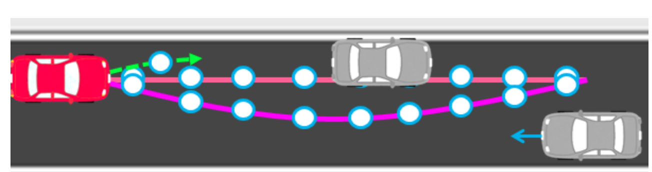

Fig. 3: States (dots) of the proposed MDP model. Pink: the RF path. Magenta: the IC path. Green: the CA path.

the endpoint of the trajectory-induction potential. The CA path is aimed to move the ego vehicle to a safer place, which is usually on the roadside away from the oncoming vehicle, to avoid immediate collisions. The generation of these candidate paths is based on Covariant Hamiltonian Optimization for Motion Planning (CHOMP) [17], which is a fast local path planner and always generates a smooth collision-free path.

2) $\mathcal{A}$: This is a set of actions, which are defined as $\{a_{\text{go\_forward}}, a_{\text{go\_change}}, a_{\text{stop}}, a_{\text{go\_avoid}}\}$. $a_{\text{go\_forward}}$ will lead the ego vehicle to go forward on the current path. $a_{\text{go\_change}}$ will lead the ego vehicle to change to the other candidate path from the current path between RF and IC paths. a_{stop} will lead the ego vehicle to stop on the current path. $a_{\text{go\_avoid}}$ will lead the ego vehicle to avoid immediate collisions by moving to a safer place if the oncoming vehicle is getting too close and may hit on the stopped ego vehicle. Note that the ego vehicle will not stay on the CA candidate path. The path on which the ego vehicle was located before the $a_{\text{go\_avoid}}$ action will still be set as its current path at the following time step.

Moreover, in the MDP, the best action in $\mathcal{A}$ will be selected with the discrete states in a fixed interval. The ego vehicle's motion in the intervals is determined based on the vehicle's dynamic model with the linear speed profile [18]. Therefore, $a_{\text{go\_forward}}$ will accelerate the ego vehicle until the target speed. a_{stop} will decelerate the ego vehicle until it finally stops. The target speed is computed in the hybrid potential map, which takes a few factors into account, such as the speed limit, the path's curvature, the constraints of lateral and longitudinal accelerations and the minimum break distance, etc.

3) $\mathcal{P}$: This is a transition probability $P(s'|s,a)$ that action a in state s will lead to state s'. It is determined by the vehicle's dynamic model with the speed and heading angle of the ego vehicle with respect to the action taken by the automated driving system.

4) $\mathcal{R}$: This is the immediate reward $R(r|s,a)$ received after transitioning from state s to state s' due to action a. Since the reward should account not only for safety, but also for time efficiency and driving comfort, in this work, it is computed as

$$R = \mu_1 r_{\text{safety}} + \mu_2 r_{\text{smoothness}} + \mu_3 r_{\text{efficiency}} \tag{3}$$

where r_{safety} is the safety term. For collisions, r_{safety} will be assigned with a large negative constant, e.g., -100, for penalty. For non-collisions, it is computed based on the cost map. $r_{\text{smoothness}}$ is the smoothness term, which measures the driving comfort. For $a_{\text{go\_forward}}$ and a_{stop} actions, it will be set as 0, since the smoothness of the RF and IC candidate paths have already been ensured during the path extraction step. For the other actions, a negative constant will be assigned for penalty. $r_{\text{efficiency}}$ is the efficiency term. It can be computed based on the spatial progress towards to the goal.

5) γ: This is the discount factor, ranging in value from 0 to 1. It represents the difference in importance between future rewards and present rewards.

C. Multi-mode Behavior Generation

The objective of this work is to generate human-like behaviors that are not only collision-free but also able to make the passengers in the automated vehicle feel at ease during the interaction with the surrounding vehicles. There are two requirements to make the passengers feel at ease. One is that the ego vehicle's behavior should coincide with the passengers' prediction and expectation. If the automated driving system always make the same reliable decisions as the passengers', the passengers will trust the system, thereby feeling at ease during the vehicle driving. The other is that the ego vehicle's behavior should meet the passengers' needs of the current driving task. The automated driving system should make decisions with higher time efficiency for an urgent task, e.g., taking an ill person to a hospital, than for a leisure task, e.g., taking families to a park.

In this work, we cover these two requirements as follows. On the one hand, we proposed a hybrid potential map, as introduced in Subsection III.A. These potentials are learned from naturalistic driving data of the same situations with the daily urban traffic such that the human driving skills and manners can be encoded. Therefore, the behavior generated based on the proposed approach can be expected to resemble the performance of a good human driver, which will coincide with the passengers' predictions and expectations [12]. On the other hand, we design three driving modes to realize different decision strategies for different passengers and different tasks. These modes are *Leisure Mode*, *Normal Mode* and *Efficiency Mode*, which are similar as the three driving modes of a Toyota Prius [19] to meet the requirements for different drivers and different tasks. This can be achieved by three sets of parameters in the MDP model, or simply changing the discount factor γ. The *Leisure Mode* will have a greater γ, thereby resulting in a more conservative decision. The *Efficiency Mode* will have a smaller γ, thereby resulting in a relatively more aggressive decision. The default values of γ for the three modes are learned from real driving data in the corresponding driving style or strategy. Moreover, the users can also adjust them via a human-machine-interface (HMI) according to their preferences. Note that all of the three modes are defined in the safety prerequisite, as shown in the experimental results in the following section.

With the parameters in the selected mode, the MDP can be solved based on the dynamic programming technique [20]. More specifically, the value function $v(s)$ is defined as

$$v(s) = \mathbb{E}[R(r|s,a) + \gamma \sum P(s'|s,a)v(s')] \tag{4}$$

The optimal value function $v^*(s)$ is given by

$$v^*(s) = \max_a \mathbb{E}[R(r|s,a) + \gamma \sum P(s'|s,a)v(s')] \tag{5}$$

Identifying the optimal values $v^*(s)$ will led to determining the optimal policy $\pi^*(s)$ by

$$\pi^*(s) = \arg\max_a \big(R(r|s,a) + \gamma \sum_{s' \in S} P(s'|s,a)v(s')\big) \tag{6}$$

Following this optimal policy will lead to computing the optimal behavior, i.e., a local trajectory with corresponding speed profile, for the ego vehicle. Note that the behaviors of the surrounding vehicles are derived from the tracking results with constant motion assumption.

Furthermore, these three modes are also used to determine the discount factor γ_v in the risk-prevention potential, as introduced in Subsection III.A. Generally, the determination of γ_v is quite difficult. A smaller γ_v, i.e., a larger advanced deceleration, could make the ego vehicle safer, however, they may also upset the passengers in the ego vehicle since the chances of the underlying risks turning into real dangers are actually quite low statistically in the daily driving. Whereas, with the provided driving mode, the γ_v could be set as the appropriate values that coincide with the passengers' preference.

It should be noted that although we have proposed three types of candidate paths, it is not necessary to always generate all of them in the behavior generation process. As mentioned previously, the appropriate behaviors are determined with respect to the vehicle categories in the current traffic circumstance. In the cases with the leader vehicle, the tail-end vehicle or the merging vehicle, the ego vehicle only needs to generate the RF path with a safe speed to keep itself driving in the center of the host lane. In the cases with the parked vehicle or the exiting vehicle with other vehicles, e.g., oncoming vehicles in the neighboring lane, the ego vehicle will need to generate both of the RF path and the IC path. The CA path will not be generated unless an other vehicle is getting close towards the stopped ego vehicle. Besides, the IC path is generated only based on the trajectory-induction potential map of the parked vehicle, thereby, it should be an ideal trajectory to cope with the parked vehicle without the influence of the oncoming vehicle. The final behavior of the ego vehicle will be generated by the optimal policy in the proposed MDF model. Such a continuous replanning over a longer time horizon while only executing over a shorter horizon, like the model-predictive control techniques, will enable the ego vehicle to react to changing vehicle behavior safely and efficiently while maintaining a smooth human-like trajectory.

IV. Experimental Results

In the experiments, the proposed system has been tested in a wide variety of typical but challenging situations. The results have substantiated the effectiveness and robustness of the proposed approach.

Fig. 4 shows the comparative example results in a curved road with two parked vehicles. The conventional potential field-based method makes the ego vehicle pass the parked vehicles one by one, whereas, the proposed system generates a behavior to pass them together with a safe and smooth trajectory. From the quantitative evaluations of the lateral displacement we can see that the proposed system's performance resembles the human driving result well. Since the generated local path is very close to the road centerline, the ego vehicle may encounter dangers during the passing maneuver if an oncoming vehicle approaches. In order to evaluate the robustness of the system, more challenging tests have been conducted in simulations.

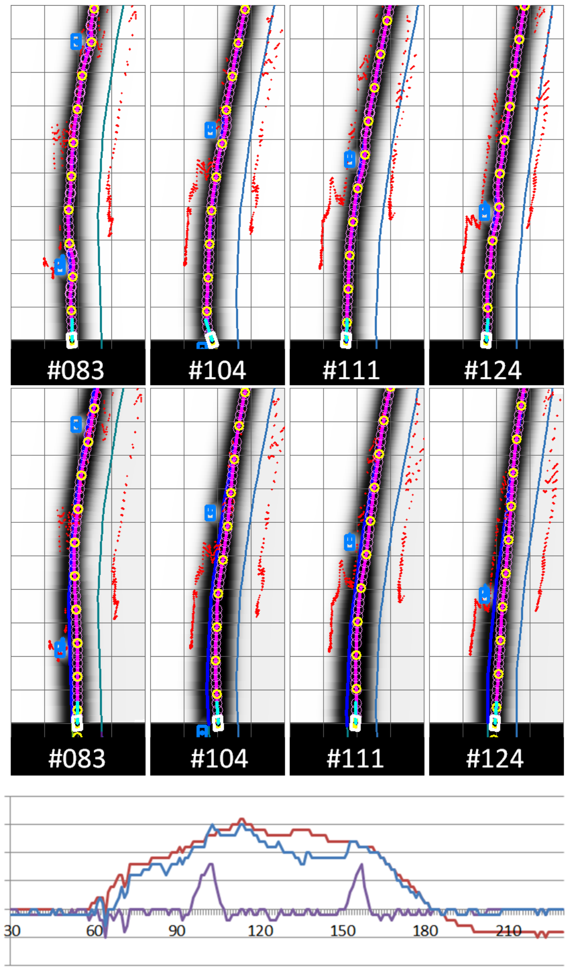

Fig. 4: Example results on a curved road with two parked vehicles. Upper: results by a conventional potential field-based method. Middle: results by the proposed approach. Lower: comparison of the lateral displacements. Purple: trajectory by the conventional potential field-based method. Light blue: trajectory by the proposed approach. Red: trajectory by a human driver.

Fig. 5 shows the example sequential results when the ego vehicle needs to circumvent a parked vehicle with an oncoming vehicle in a double-lane road. The ego vehicle starts the passing maneuver as soon as it enters the trajectory-induction range of the parked vehicle (Fig. 5(a)). When the oncoming vehicle is approaching, the ego vehicle decelerates until stop to wait for the oncoming vehicle (Fig. 5(b)). It restarts the passing maneuver when the oncoming vehicle is nearly passing by, and finally passes the parked vehicle and returns to the host lane (Fig. 5(c)).

Fig. 6 compares the proposed system's performances with the oncoming vehicle at different lateral distances from the host lane. In the case of the oncoming vehicle driving $4.3m$ away from the host lane, it has no influence on the ego vehicle, and the ego vehicle circumvents the parked vehicle

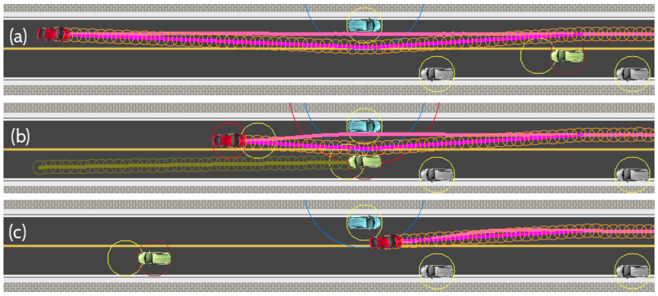

Fig. 5: Example sequential results with an oncoming vehicle. Pink: the RF candidate path. Magenta: the IC candidate path. Continuous circles: the selected path. Note that the oncoming vehicle is always going forward, whose lateral position is controlled by a collision-free path planner.

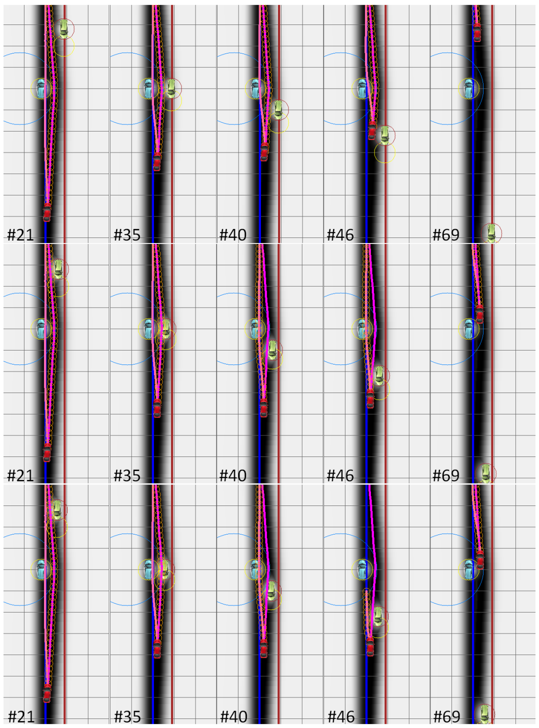

Fig. 6: Example sequential results with the oncoming vehicle at different lateral positions away from the host lane. Upper: $4.3m$ away. Middle: $3.0m$ away. Lower: $2.7m$ away.

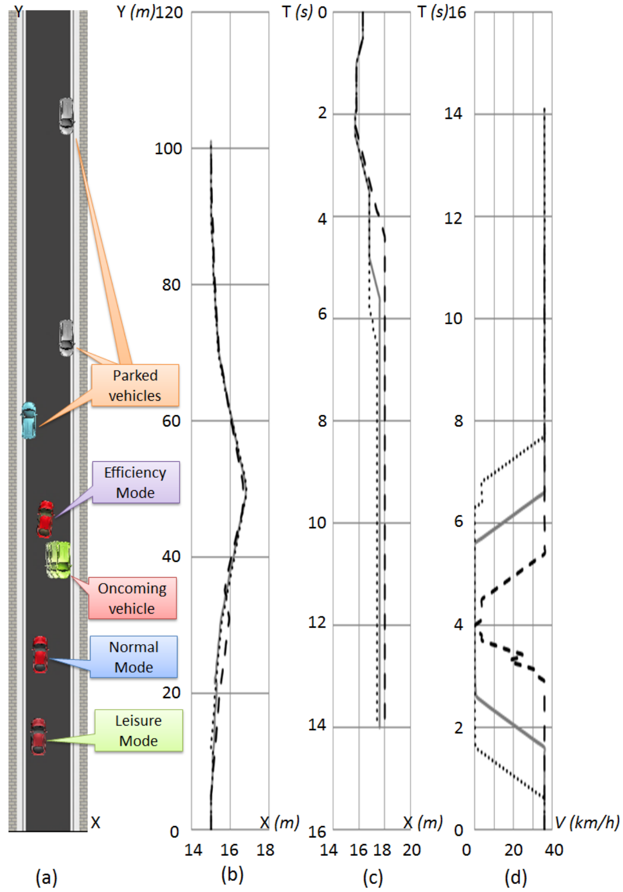

Fig. 7: Quantitative evaluations of the three driving modes. (a) Vehicle positions of different modes at the same moment. (b) Trajectories of the ego vehicle. Note that the scales of horizontal and vertical axes are different. (c) Lateral position of the oncoming vehicle. (d) Speed of the ego vehicle. Solid lines: *Normal Mode*. Dot lines: *Leisure Mode*. Dash lines: *Efficiency Mode*.

smoothly. In the case of the oncoming vehicle driving $3.0m$ away from the host lane, the ego vehicle decelerates until stop during the IC behavior since the oncoming vehicle is getting close. It then starts the RF behavior to go back to the host lane a little bit to give room to the oncoming vehicle. When the oncoming vehicle is nearly passing by, the ego vehicle re-starts the IC behavior to pass the parked vehicle. In the case of the oncoming vehicle driving $2.7m$ away from the host lane, the ego vehicle decelerates until stop during the IC behavior, and then starts the RF behavior to go back to the host lane to give room to the oncoming vehicle. As the oncoming vehicle is getting closer and closer towards the ego vehicle, the CA behavior is triggered and the ego vehicle moves to the left side of the host lane to reduce the risk of collision with the oncoming vehicle. When the oncoming vehicle is nearly passing by, the ego vehicle re-starts the IC behavior again to pass the parked vehicle.

Moreover, Fig. 7 provides the evaluations of the three driving modes when the ego vehicle is circumventing a parked vehicle with an oncoming vehicle in a narrow one-lane road. As we can see from the results, in the *Leisure Mode*, the ego vehicle stops early during the IC behavior, and then continues the IC behavior as the oncoming vehicle is nearly passing by. Whereas, in the *Efficiency Mode*, the ego vehicle starts the IC behavior early and starts to decelerate late. When the oncoming vehicle is getting close, it does not stop in the IC path but selects the RF behavior to go back to the host lane a little bit, and then stop to wait for the oncoming vehicle. Meanwhile, for the oncoming vehicle on the other hand, it does not want to hit the ego vehicle, either. Therefore, the oncoming vehicle, whose lateral position is controlled by a collision-free path planner, naturally moves to the left side of its lane to give more room to the ego vehicle. In practice, the driver of the oncoming vehicle can also understand the intention of the ego vehicle according to the ego vehicle's IC behavior. In this case, the ego vehicle in

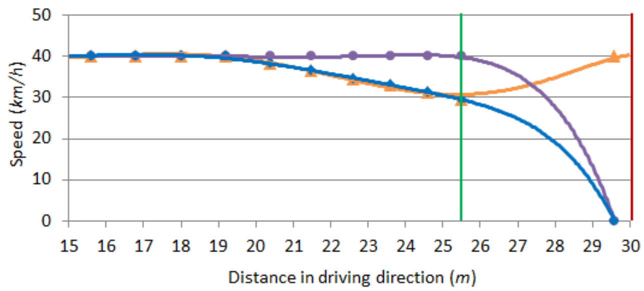

Fig. 8: Evaluation of speed profiles in the situation of an object may dart out from behind a parked vehicle. Orange triangle and blue diamond marks: target speeds of the risk-prevention potential. Purple dots: target speeds without early speed adjustment. Red line: the position of the potential collision point. Green line: the position from where the area behind the parked vehicle enters the camera's field of view. Blue and orange curves: the speed profiles with the use of risk-prevention potential in the cases with and without a darting-out object, respectively. Purple curve: the speed profile without the use of risk-prevention potential in the case with a darting-out object.

the *Efficiency Mode* can pass the parked vehicle efficiently thanks to the interactions between the two vehicles.

Furthermore, Fig. 8 gives the example results of the advanced speed adjustment by the proposed risk-prevention potential. From the results we can see that the ego vehicle is driving at $40km/h$, which is a target speed computed in the trajectory-induction potentials. As the ego vehicle is approaching to the potential collision point near the parked vehicle, the target speed, given by the risk-prevention potential, comes into effect, which reduces the ego vehicle's speed in advance so that it can be assured that the ego vehicle will be able to stop before the potential collision point once an object darts from behind the parked vehicle. Once the ego vehicle reaches the location where the driver's sight or the camera's field of view is not obstructed by the parked vehicle, the underlying risk of darting-out object will no longer exist. It will either simply disappear if no objects are detected behind the parked vehicle, or turn into an actual danger if it is found that an object is darting out. In the former case, the ego vehicle will accelerate again to the previous target speed, as indicated by the orange line in Fig. 8. In the latter case, the ego vehicle will perform an emergency brake, as indicated by the blue line in Fig. 8. From the results we can see that the deceleration of the emergency brake with the use of the proposed risk-prevention potential is relatively gentle, thanks to the advanced speed adjustment. Whereas, the one without using the proposed potential is quite hard, as indicated by the purple line in Fig. 8. In this case, the ego vehicle may have a large chance to cause a collision, since the braking distance at $40km/h$ is about $9m$ in general. Even it can avoid the collision, the large deceleration magnitude will make the passengers in the ego vehicle uncomfortable.

V. Conclusions

In this paper, we proposed a human-like behavior generation approach based on a Markov decision process with hybrid potential maps, in which, the automated driving system moves by selecting the optimal candidate path continuously. Such a mechanism can allow the ego vehicle to react to changing vehicle behavior safely and efficiently. Moreover, three different driving modes have been proposed to meet the requirements of different passengers and different tasks. Such a system will not only make passengers feel safe with collision-free performance but also make them feel at ease with efficient and reliable behaviors. Experimental results in various typical but challenging urban traffic scenes have substantiated the effectiveness of the proposed system.

Future work will be focused on the tests with real vehicles and the evaluations of the human-like performance in terms of feelings of safety and at ease with different drivers and passengers.

References

[1] National Highway Traffic Safety Administration, Traffic safety facts, 2011.

[2] http://en.wikipedia.org/wiki/Advanced_driver_assistance_systems.

[3] R. Bellman, A Markovian decision process, Journal of Mathematics and Mechanics, pp. 679−684, 1957.

[4] T. Gu, and J. Dolan, On-road motion planning for autonomous vehicles, International Conference on Intelligent Robotics and Applications, pp. 588−597, 2012.

[5] S. Karaman, and E. Frazzoli, Incremental sampling-based algorithms for optimal motion planning, Robotics Science and Systems VI, 104, 2010.

[6] C. Guo, *et al*, Learning-based trajectory generation for intelligent vehicles in urban environment, Proc. IEEE IV Symp., pp. 1236−1241, 2016.

[7] F. Havlak, and M. Campbell, Discrete and continuous, probabilistic anticipation for autonomous robots in urban environments, Transactions on Robotics, vol. 30, pp. 461474, 2014.

[8] Q. Tran, and J. Firl, Modelling of traffic situations at urban intersections with probabilistic non-parametric regression, Proc. IEEE IV Symp., pp. 334339, 2013.

[9] S. Lee, and S. Seo, A learning-based framework for handling dilemmas in urban automated driving, Proc. Int. Conf. on Rob. and Aut., pp. 1436−1442, 2017.

[10] J. Wei, J. Snider, T. Gu, J. Dolan, and B. Litkouhi, a behavioral planning framework for autonomous driving, in 2014 IEEE Intelligent Vehicles Symposium, pp. 458−464, 2014.

[11] A. Cunningham, E. Galceran, R. Eusice, and E. Olson, MPDM: Multipolicy decision-making in dynamic, uncertain environments for autonomous driving, 2015 IEEE International Conference on Robotics and Automation, pp.1670−1677, 2015.

[12] C. Guo, *et al*, Toward human-like lane following behavior in urban environment with a learning-based behavior-induction potential map, Proc. Int. Conf. on Rob. and Aut., pp. 1409−1416, 2017.

[13] C. Guo, *et al*, A multimodal ADAS system for unmarked urban scenarios based on road context understanding, IEEE Trans. Intell Transp. Syst., 16(4), pp. 1690−1704, 2015.

[14] http://www.jaf.or.jp/e/risk_prediction/index.htm.

[15] C. Guo, *et al*, Human-like behavior generation for intelligent vehicles in urban environment based on a hybrid potential map, Proc. IEEE IV Symp., pp. 197−203, 2017.

[16] https://en.wikipedia.org/wiki/Braking_distance.

[17] M. Zucker, *et al*, CHOMP: Covariant Hamiltonian optimization for motion planning, The International Journal of Robotics Research, 32(9-10), pp. 1164−1193, 2013.

[18] D. Ferguson, *et al*, Motion planning in urban environments, Journal of Field Robotics 25.1112, pp. 939−960, 2008.

[19] http://toyota.jp/prius/

[20] R. Howard, Dynamic programming and markov processes, 1960.