Human-like motion planning model for driving in signalized intersections

Yanlei Gu a,⁎, Yoriyoshi Hashimoto b, Li-Ta Hsu a, Miho Iryo-Asano a, Shunsuke Kamijo a

a Institute of Industrial Science, The University of Tokyo, Japan
b Graduate School of Information Science and Technology, The University of Tokyo, Japan

Abstract

Highly automated and fully autonomous vehicles are much more likely to be accepted if they react in the same way as human drivers do, especially in a hybrid traffic situation, which allows autonomous vehicles and human-driven vehicles to share the same road. This paper proposes a human-like motion planning model to represent how human drivers assess environments and operate vehicles in signalized intersections. The developed model consists of a pedestrian intention detection model, gap detection model, and vehicle control model. These three submodels are individually responsible for situation assessment, decision making, and action, and also depend on each other in the process of motion planning. In addition, these submodels are constructed and learned on the basis of human drivers’ data collected from real traffic environments. To verify the effectiveness of the proposed motion planning model, we compared the proposed model with actual human driver and pedestrian data. The experimental results showed that our proposed model and actual human driver behaviors are highly similar with respect to gap acceptance in intersections.

© 2016 Production and hosting by Elsevier Ltd. on behalf of International Association of Traffic and Safety Sciences. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Recent developments in advanced driver assistance systems and autonomous robots seem to suggest that cars will be able to drive without human intervention in the near future. Thus, autonomous vehicles will join human drivers on the road soon. Currently, research studies on autonomous vehicles focus on their safety aspects to reduce accidents. These studies have adopted various sensors, such as LIDAR, radar, and vision, to perceive the surrounding environment and avoid collision with other vehicles and pedestrians. There is another critical issue in a hybrid traffic situation. Humans, including pedestrians and drivers, should not be affected by autonomous vehicles. In other words, the behavior of an autonomous vehicle is supposed to be similar to that of a human-driven vehicle, to avoid confusing pedestrians and other drivers in decision making. The accident reports on Google’s driverless car also suggested that robot cars might actually be too cautious and careful. Google is actually working to correct this cautiousness and make its cars drive more similarly to humans to reduce the number of accidents [1]. This paper proposes a human-like motion planning model that can control vehicles like humans do.

Vehicle motion models can be divided into three levels with an increasing degree of abstraction: physics-based motion models, maneuver-based motion models, and interaction-aware motion models [2]. The physics-based motion models explain the vehicle motion by velocity, acceleration, mass of the vehicle, road surface friction coefficient, and the laws of physics. This type of models can be used for predicting the evolution of the state of the vehicle [3,4], but is limited to short-term (less than 1 s) motion prediction [2]. The maneuver-based motion models represent vehicles as independent maneuvering entities and could provide long-term predictions of driver intentions. Campbell et al. and Amsalu et al. proposed to use the continuous vehicle dynamics for long-term (less than 1 s) motion prediction [2]. The maneuver-based motion models represent vehicles as independent maneuvering entities and could provide long-term predictions of driver intentions. Campbell et al. and Amsalu et al. proposed to use the continuous vehicle dynamics for long-term (less than 1 s) motion prediction [2]. The maneuver-based motion models represent vehicles as independent maneuvering entities and could provide long-term predictions of driver intentions. Campbell et al. and Amsalu et al. proposed to use the continuous vehicle dynamics for long-term (less than 1 s) motion prediction [2].
situation derived from the local situational context [7,8]. Platho et al. proposed to decompose the complex situations into smaller and more manageable parts to recognize and understand the driving situations [9]. Hulsen et al. suggested that driving behavior is greatly influenced by four aspects: traffic rules, assessment of allowed actions, expected behaviors, and impacts of traffic participants on each other [10,11]. They introduced an ontology to model traffic situations at complex intersections and enabled reasoning about traffic rules for involved vehicles. Obviously, the influence of contextual information, such as traffic rules, road structure, and actions of other road users, should be considered in the motion planning model.

To model and represent human-like motion planning, we need to understand how the contextual information affects a driver’s action. The influence can be modeled and analyzed on the basis of data collected from real traffic environments. In particular, most research studies on autonomous vehicles focus on right- or left-turning vehicles at intersections and discuss how vehicles pass the intersections in the case of sharing the road with other road users. The driving maneuver, in which a turning vehicle passes the intersection, is called gap acceptance. The basic idea of gap acceptance is to estimate the time difference between two consecutive pedestrians and vehicles [12]. Ragland et al. analyzed the distribution of accepted and rejected gaps in the left turn across path/opposite direction scenarios and proposed to characterize gap acceptance by a logistic model [13]. Zohdy also proposed to determine the critical gaps using a logit function [14]. Schroeder et al. explored factors associated with driver-yielding behavior at unsignalized pedestrian crossings and developed predictive models by using logistic regression [15]. Rather than at common intersections, Salamati et al. aimed to identify the contributing factors affecting the likelihood of a driver yielding to pedestrians at two-lane roundabouts [16]. Alhajyaseen et al. [17] and Wolfemann et al. [18] explained the stochastic speed profiles and the stochastic path models of free-flowing left- and right-turning vehicles from the aspect of intersection layout. Moreover, Alhajyaseen et al. [19,20] analyzed the vehicle gap acceptance behaviors against pedestrians and further proposed an integrated model. The integrated model represented the variations in the maneuvers of left-turners (left-hand traffic) at signalized intersections, and the proposed model dynamically considered the vehicle reaction to intersection geometry and crossing pedestrians [21]. Those research studies focused on analyzing how contextual information affects the driver’s behavior.

Recently, researchers applied motion models to control vehicles. Kye et al. presented intention-aware automated driving at unsignalized intersections. The intention-aware decision-making problem is modeled as a partially observable Markov decision process [22]. As for collision avoidance, Kohler et al. proposed to recognize the pedestrians standing at the curb and intending to cross the street despite an approaching car. The proposed active pedestrian protection system can perform an autonomous lane-keeping evasive maneuver in urban traffic scenarios to avoid braking [23]. Keller et al. and Braeuchle et al. proposed an active pedestrian safety system that combines sensing, situation analysis, decision making, and vehicle control. The proposed system can decide whether it will perform automatic braking or evasive steering and reliably execute this maneuver at relatively high vehicle speed [24]. Moreover, Pongsathorn and Akagi et al. proposed to reduce collisions at potentially hazardous areas by suggesting an appropriate speed, which is learned from actual driving data of expert drivers [25,26].

This paper focuses on the scenario at an intersection, one of the most challenging traffic scenarios, and proposes a human-like motion planning model for left-turning vehicles. Fig. 1 illustrates a traffic scenario wherein a vehicle turns and passes an intersection while there are pedestrians walking on or to the crosswalk. In this case, the driver will wait for an appropriate moment and then cross the intersection by iteratively assessing pedestrian situations, making decisions, and adjusting actions. The proposed model represents the whole driving process, as shown in Fig. 2. The proposed model consists of three submodels: pedestrian intention detection model, gap detection model, and vehicle control model. These three submodels are separately responsible for situation assessment, decision making, and action. They also depend on each other in the proposed motion planning model.

In addition, the construction of the motion planning model was conducted on the basis of the analysis of actual human driver data. To obtain a credible model, we collected real data at an intersection in Tokyo City. In the verification of the effectiveness of the proposed idea, the model was implemented as a virtual driver, which allows for comparison with the behavior of human drivers. The contribution of this paper is the development of a human-like motion planning model by integrating a pedestrian intention detection model, gap detection model, and vehicle control model. This paper presents the proposed model and its performance in Sections 2 and 3, respectively. Finally, this paper will be concluded in Section 4.

2. Motion planning model

As shown in Fig. 2, the proposed motion planning model includes different submodels. This section explains the construction of each submodel and describes the relationships between the submodels as well. Before the explanations, we clarify the assumptions for the developed models.

a. The vehicle trajectory has been determined before the vehicle turns. It means that the proposed model controls the vehicle position along the longitudinal direction rather than changing the trajectory [27]. This assumption is consistent with the common actions of human drivers at intersections.

b. The road structure, traffic signal phase, and elapsed time of the phase are assumed to be known, which can be transmitted from a vehicle-to-infrastructure system [28].
2.1. The pedestrian intention detection model

Pedestrian behaviors are affected by the surrounding traffic situation. The intensive research studies on traffic engineering have suggested that pedestrian behaviors are potentially related to the signal phase, the intersection layout, the vehicles, and even to other pedestrians at signalized intersections [29–31]. Regarding the behavioral flow of pedestrians: assessment, decision making, and physical movement, as a stochastic process, our previous works constructed a probabilistic model of pedestrian behavior using a DBN [32,33]. The developed model takes into account not only pedestrian physical states but also contextual information, and integrates the relationship between them. It is important to note that our pedestrian behavior model can recognize a pedestrian’s crossing or waiting intention before he or she enters the crosswalk area or stops at a road side. The detailed description of the pedestrian behavior model has been published in our previous work [32, 33]. This section describes only the conception of the model.

Fig. 3 illustrates the pedestrian behavior model graphically, which is represented by a DBN. A DBN is a Bayesian network that relates variables to each other over adjacent time steps. Nodes in a DBN, which correspond to the variables in rectangles or ellipses in Fig. 3, represent the temporal process and its possible states. The arcs, which are indicated by solid or dash lines, represent the local or transitional dependencies among variables in a DBN. The construction of a DBN consists of building a network structure and learning the parameters for describing the “arcs.” In this research, parameters in the arcs were learned from actual pedestrian data collected at intersections.

Table 1 summarizes the variables, which are shown in the pedestrian behavior model. As illustrated in Fig. 3, the proposed model first assesses the situation, which is indicated by \( C_t \). In this research, the contextual information \( (C) \) includes the traffic signal phase \( (S_{t1}) \), the elapsed time of the phase \( (S_{e1}) \), the vehicle conditions in the surrounding environment \( (V_t) \), the road side relative to the left-turning vehicle \( (S_{d2}) \), the group situation \( (G_t) \), and the length of the crosswalk \( (C_l) \). According to the graph, the contextual information \( (C_t) \) pedestrian positions \( (P_{t1}−1) \), and decision \( (D_{t1}−1) \) at the last time step jointly affect the decision of crossing or waiting at \( D_t \) at the current time step. In fact, the connection between \( P_{t1}−1 \) and \( D_t \) is represented by \( L^2_t \), which is the distance to the destination. In the case of a pedestrian preparing to enter a crosswalk \( (W=before) \), \( L^2_t \) is the distance to the entrance edge of the crosswalk \( (L^{ent}) \). For a pedestrian in the crosswalk area \( (W=on) \), \( L^2_t \) is the distance to the exit edge of the crosswalk \( (L^{exit}) \). This configuration was inspired by actual pedestrian behaviors.

Next, with the contextual information \( (C_t) \), crossing/waiting decision \( (D_t) \), pedestrian position \( (P_{t1}−1) \), and motion type \( (M_{t1}−1) \), the model estimates the probability of motion transition, e.g., the possibility that a pedestrian will change from walking to running. The reason we estimate the motion transition is that pedestrians show different distributions of speed at different motion types. After the estimation of \( D_t \) and \( M_t \), the model predicts the pedestrian speed, moving direction, and pedestrian position. Furthermore, the proposed model uses the observation of the pedestrian position \( (Z_t) \) to update the probability of the predicted pedestrian position.

To consecutively estimate the pedestrian state including the decision, motion type, and dynamics, we employ a sample-based method, the particle filter (PF) algorithm, as inference. The general PF algorithm has three steps: sampling, importance sampling, and resampling. In the sampling step, which corresponds to the prediction, each particle moves in its state space \( [D_t,M_t,D_t1,S_{p1},P_t] \) according to its previous state and the proposed graphical model. The probability of the predicted state is evaluated by the contextual information. In the importance sampling step, which corresponds to the correction, the importance weight of each particle is updated on the basis of the observation of the pedestrian position. In the resampling step, particles are reproduced/discarded so

Please cite this article as: Y. Gu, et al., Human-like motion planning model for driving in signalized intersections, IATSS Research (2016), http://dx.doi.org/10.1016/j.iatssr.2016.11.002
that the number of distributed particles will be proportional to the importance weight distribution. Afterwards, all particles are assigned the same importance weights before going to the next epoch. With this mechanism, the pedestrian behavior model can estimate the crossing intention, motion type, moving direction, speed, and position.

2.2. The gap detection model

After obtaining the crossing/waiting intention, position, and speed of pedestrians, the proposed motion planning model estimates which gap is appropriate to pass on the basis of the gap detection model. Generally, a human driver decides whether he or she can pass a gap between two pedestrians when the first pedestrian enters the conflict area of the vehicle trajectory, which was empirically setup as a 2.5-m width in this research. This value was obtained by analyzing real human-driving data. Fig. 4 shows the statistics result on the real human-driving data. The blue line and red line indicate the distributions of the pedestrian distance to the vehicle trajectory at the moment of the drivers accelerating the vehicles, in the case of hard yield and soft yield, respectively. The negative value along the horizontal axis means that the pedestrian has crossed the vehicle trajectory. From this figure, we can see that 85% of the drivers did not start accelerating until the pedestrian had approached the trajectory of about 2.5 m/1.0 m in the case of a hard yield/soft yield, respectively. Moreover, we can see that, in the hard-yield cases, the drivers could accelerate the vehicles earlier than in the soft-yield cases. Thus, our proposed system determined that the conflict area had a 2.5-meter width, and the proposed gap detection model estimated the probability of gap acceptance on the basis of the situation at this moment (the first pedestrian just entered the conflict area).

Fig. 5 illustrates the configurations of two pedestrians and a vehicle at this moment. The dash line is the vehicle trajectory, and the blue rectangle is the conflict area. Pedestrian 1 just arrives to the conflict area at time t. At this moment, the distance from pedestrian 2 to the vehicle trajectory is \( D_{p2,t} \), and the speed of pedestrian 2 is \( V_{p2,t} \). In addition, the distance from the vehicle to the conflict point is \( D_{v,t} \), and the vehicle speed is \( V_{v,t} \). This paper proposes to model the gap acceptance behavior using these four parameters. The probability of gap acceptance \( L(x) \) is formulated as follows:

\[
L(x) = \frac{\exp(\alpha + \beta x)}{1 + \exp(\alpha + \beta x)}
\]  

(1)

\[
x = [D_{p2,t}, V_{p2,t}, D_{v,t}, V_{v,t}]
\]  

(2)

where \( x \) is a vector of explanatory variables; \( x \) consists of “pedestrian 2” distance \( D_{p2,t} \), and speed \( V_{p2,t} \), and vehicle distance \( D_{v,t} \), and speed \( V_{v,t} \). \( \alpha \) and \( \beta \) are a constant and the coefficient for the explanatory variables, respectively. The values of \( \alpha \) and \( \beta \) are learned from actual human driver data using maximum likelihood estimation. The parameters will be visualized in Section 3.3. This model can be used to decide whether to accept or reject the gap according to the output of Eq. (1).
2.3. The vehicle control model

This driving process can be basically divided into the in-flow and out-flow stages, which correspond to the deceleration before entering a crosswalk and to the acceleration while passing the crosswalk, respectively. Fig. 6 shows the speed of some actual vehicles in the intersection area. The red lines are the speed of the vehicles when there is no pedestrian. In this case, passing an intersection is called free flow. The blue and green lines are the speed of the vehicles when pedestrians are present. Obviously, the minimum speed in free flow is higher than that in other cases. In addition, some vehicles stop during the period of passing an intersection, such as the blue lines in Fig. 6. This type of passing is called a hard yield. In contrast, the green lines do not reach zero during the period of passing. This case is called a soft yield. Theoretically, soft yield can be considered as a preparation for passing the intersection because the vehicle does not stop and can pass the intersection at a shorter time compared to hard yield. This paper proposes to choose either a hard yield or a soft yield for passing on the basis of the gap detection model.

In most of the cases, vehicles approach an intersection with deceleration and pass it with acceleration. Wolfermann et al. suggested that the speed curves split according to the acceleration and deceleration, and can be approximated by cubic functions [18]. In this case, the rate of change in acceleration, i.e., jerk, is represented by a linear function. Thus, with the initial jerk, slope of jerk, acceleration, and speed, the jerk, acceleration, speed, and distance at time can be determined by Eqs. (3) to (6), respectively.

\[
\begin{align*}
J_i &= kt + J_0 \\
A_i &= k \frac{t^2}{2} + J_0 t + A_0
\end{align*}
\]

where the slope of the jerk describes the change rate of the jerk. Generally, a big jerk makes passengers uncomfortable owing to the high dynamics of the inertial force. In this model, if \(J_0, k, A_0, v_0\), are fixed, the speed profile is also determined. It is important to note that our proposed control model does not follow one constant profile. It dynamically chooses the profiles according to the pedestrian conditions. For example, with the expected speed and acceleration, the required \(J_0, k\) can be adjusted using Eqs. (3) to (5).

In a real traffic situation, when there are many pedestrians at an intersection, it is difficult to find a gap to pass the intersection. In this case, drivers usually stop in front of the crosswalks. With the vehicle position, speed, and acceleration at current time, the vehicle is expected to stop at the stop point. With this assumption, \(J_0, k\) can be determined. Alhajyaseen et al. [20] used this profile (stopping profile in the paper) before accepting a gap. Our paper also adopted this idea for generating a hard-yield profile.

However, some drivers adopt a soft-yield profile for passing the intersection. This paper proposed to use the gap detection model to find potential gaps and determine which profile (hard- or soft-yield profile) should be chosen. To apply the gap detection model, we need to predict potential gaps and determine which profile is suitable. In the paper, the pedestrian states and vehicle state, to make the first pedestrian satisfy the requirement in the gap detection model. At the predicted moment, the pedestrian should just arrive near the edge of the conflict area, as shown in Fig. 5. The predicted states of pedestrian 2 and the vehicle are used for evaluating the acceptance probability of the potential gap using Eq. (1).

If the potential gap is determined, the system moves to the next function, selecting a hard or soft yield. Fig. 7 visualizes the idea of the profile selection. Suppose that the positions of pedestrian 1 and pedestrian 2 are \(P_{1,t}\) and \(P_{2,t}\) at time \(t\) and their speeds are \(V_{1,t}\) and \(V_{2,t}\), respectively. At moment \(t\), the vehicle position and velocity are assumed as \(V_{v,t}\) and \(V_{v,t}\), respectively. Pedestrian 1 needs time \(\Delta t\) to arrive to the far side of the conflict area:

\[
\Delta t = \frac{D_{1-t-f}}{|V_{v,t}|}
\]

where the speed is determined by the speed profile \(V_{v,t}\) used at time \(t\). \(V_{v,t}\) is the hard-yield profile. The proposed vehicle control model determines the type of profile on the basis of the following equations:

\[
\begin{align*}
P_{v,t+\Delta t} &= P_{v,t} + \int_{t}^{t+\Delta t} V_v(s) \, ds \\
V_{v,t+\Delta t} &= V_v(t + \Delta t)
\end{align*}
\]

where the speed is determined by the speed profile \(V_v(s)\) used at time \(t\). \(V_v(s)\) is the hard-yield profile. The proposed vehicle control model determines the type of profile on the basis of the following equations:

\[
P_{v,t+\Delta t} = \begin{cases} 
V_v(t + \Delta t), & \text{if } V_v(t + \Delta t) > 0 \\
0, & \text{if } V_v(t + \Delta t) = 0
\end{cases}
\]

Please cite this article as: Y. Gu, et al., Human-like motion planning model for driving in signalized intersections, IATSS Research (2016), http://dx.doi.org/10.1016/j.iatsr.2016.11.002
After the vehicle arrives to the stop point, the system evaluates the situation for clearing. In the out-flow profile, the initial jerk $j_0$ and slope $k$ are constant values and empirically determined. Moreover, in the out-flow profile, the system first judges whether the vehicle can...
pass the gap with the determined out-flow profile at every time step. The determination is conducted by maintaining a maximum margin to the pedestrians. If the time is appropriate, the system changes to the out-flow process and passes the crosswalk.

3. Experiments

3.1. Experiment setup and data collection

To learn and verify the human-like motion planning model, we collected actual data from one intersection in Tokyo City. The road structure of the intersection is illustrated in Fig. 8, which is a snapshot from Google Earth. We installed cameras at a high floor of a building, which is located around the intersection. Fig. 9 shows the image captured from the camera. The captured video had 10 fps (frame per second) and an 842 × 480 pixel resolution. The captured video was manually calibrated to remove the perspective effect. Because of the occlusion caused by trees and the limited view field of the pedestrian

![Fig. 9. Image captured from a camera installed at a high floor of a building around the intersection. The red points are the labeled pedestrian trajectories, and the green points are the labeled vehicle trajectories.](image)

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Summary of the collected data from the experiment intersection.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>West crosswalk</td>
</tr>
<tr>
<td>Signal cycles</td>
<td>Total</td>
</tr>
<tr>
<td>Pedestrians</td>
<td>Near side</td>
</tr>
<tr>
<td></td>
<td>Far side</td>
</tr>
<tr>
<td></td>
<td>Alone</td>
</tr>
<tr>
<td></td>
<td>Group</td>
</tr>
<tr>
<td></td>
<td>Crossing</td>
</tr>
<tr>
<td></td>
<td>Waiting</td>
</tr>
<tr>
<td></td>
<td>Total</td>
</tr>
<tr>
<td>Vehicles</td>
<td>Left turning</td>
</tr>
</tbody>
</table>

![Fig. 10. Crossing decision probability at the onset of PFG.](image)

Please cite this article as: Y. Gu, et al., Human-like motion planning model for driving in signalized intersections, IATSS Research (2016), http://dx.doi.org/10.1016/j.iatssr.2016.11.002
Table 3
Logit regression results of the pedestrian intention recognition model at the onset of PFG.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficient</th>
<th>Z-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>5.5302</td>
<td>5.902</td>
</tr>
<tr>
<td>Crosswalk length (m)</td>
<td>−0.0968</td>
<td>−2.601</td>
</tr>
<tr>
<td>Pedestrian group condition ∈(0.1)</td>
<td>−2.2165</td>
<td>−3.975</td>
</tr>
<tr>
<td>Vehicle condition ∈(0.1)</td>
<td>−0.9314</td>
<td>−1.887</td>
</tr>
<tr>
<td>Distance to crosswalk entrance (m)</td>
<td>−0.2593</td>
<td>−5.485</td>
</tr>
<tr>
<td>Number of observations</td>
<td>166</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.3321</td>
<td></td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>−71.745</td>
<td></td>
</tr>
</tbody>
</table>

walking areas, only the “West crosswalk” and “South crosswalk” were considered in this research. The lengths of these crosswalks are 23 m and 10 m, respectively. In this intersection, the cycle time of the traffic signal is fixed. Therefore, we can easily label the signal phase in the whole video by deciding the start time of the first cycle.

The pedestrian and vehicle positions were labeled from the image. An example of the labeling is shown in Fig. 9. The red points and green points correspond to the pedestrians and vehicles, respectively. The sequential position of one vehicle was used as the vehicle trajectory. Because we could not determine the true pedestrian intention at each time step, the decision values were labeled as “waiting” if the pedestrian did not cross. In addition, we applied a Kalman filter to each trajectory and regarded it as the ground truth trajectory. Table 2 shows the statistics of the collected data from the pedestrians and vehicles. Totally, we had 921 pedestrians and 148 left-turning vehicles.

3.2. Evaluation of the pedestrian behavior model

In the evaluation of pedestrian behaviors, we used a fourfold cross-validation to divide the dataset into training and test sequences. The parameters in the proposed model were determined by applying the maximum likelihood estimation in the training sub-datasets. Fig. 10 visualizes the learned relationship between contextual information and crossing probability at the onset of the pedestrian flashing green (PFG) time. Actually, the probability of crossing decision is represented by a logistic function. The variables of the logistic function include the distance from the pedestrian to the crosswalk, vehicle situation, group situation, and crosswalk length. The vertical direction in Fig. 10 is the probability of the crossing decision of pedestrians, and the horizontal direction is the distance from the pedestrian to the edge of the crosswalk. The positive value means that the pedestrian did not enter the area of the crosswalk. It can be seen clearly that the crossing probability, which is the value of the logistic function, increased as the pedestrian was approaching the crosswalk. The different color lines correspond to different contextual conditions. The green line means the crossing probability in the case where a lone pedestrian was crossing a 23-m-length crosswalk with a vehicle waiting. By comparing the green line and the black line, we can see that, if a vehicle appeared at the intersection, the crossing probability would decrease. It means that pedestrians sometimes gave up crossing because of the vehicles. In addition, if the crosswalk was shorter, pedestrians would have more intention to cross during the PFG time, which is indicated by the red line. Moreover, the blue line indicates that pedestrians in group had a lower probability of crossing compared to lone pedestrians (black line).

Table 3 shows the coefficients of the variables obtained from the training. All the four parameters had negative effects on the crossing intention of pedestrians. Moreover, the magnitude of the Z-values indicated the pedestrian group condition and the pedestrian distance to the crosswalk entrance; these two variables are more significant compared to the vehicle condition and the crosswalk length in our model.

In the pedestrian behavior model, the observation was the pedestrian position. We did not directly use the manually labeled position as the observation. The accurate observation could not show the noise tolerance feature of the proposed model. To verify the reliability of the system, we added different levels of the noise to the labeled pedestrian position. The noise was assumed to be normal probability distributions with variances of 0.1, 0.4, and 1.0 m. The DBN model used the noise position as the observation. The decision recognition accuracy is shown in Fig. 11. The left image of Fig. 11 shows the recognition accuracy at different positions relative to the crosswalk. In the smallest position noise case ($σ = 0.1$ m), the proposed model could achieve a 90% recognition rate. In the high-noise conditions, the system still maintained a recognition rate higher than 80%. The right image in Fig. 11 shows the recognition accuracy at different times after the onset of PFG. With the increase in time, the recognition accuracy increased. On average, the system could recognize the pedestrian decision in 83% of the cases at a noise level of 0.4 m.

3.3. Evaluation of the human-like motion planning model

In this paper, we proposed the gap detection model. The probability of the gap acceptance is affected by four parameters: longitudinal

![Fig. 12. Visualization of the gap acceptance model.](image-url)

Fig. 11. Recognition rate of the crossing/waiting decision with respect to the distance to the crosswalk (left) and to the time from the onset of PFG (right). The different colors indicate the noise level in the position observation.

Please cite this article as: Y. Gu, et al., Human-like motion planning model for driving in signalized intersections, IATSS Research (2016), http://dx.doi.org/10.1016/j.iatssr.2016.11.002
distance from the vehicle to the conflict point, vehicle speed, distance from the pedestrian to the conflict point, and pedestrian speed. The coefficients of the four parameters, which are generated by learning from actual human driver data, indicate how the four parameters affect the probability. Fig. 12 illustrates the relationship between the gap acceptance probability and the parameters. The gap acceptance probability is the value of the logistic function in Eq. (1). The vertical direction in Fig. 12 is the probability of the gap acceptance, and the horizontal direction is the distance from the pedestrian to the conflict point. It can be seen clearly that the probability increased with the increase in distance from the pedestrian to the conflict point. The different color lines correspond to different situations represented by pedestrian speed, vehicle distance, and vehicle speed. The green line means the acceptance probability when a pedestrian was moving at a speed of 1.5 m/s and the vehicle was moving at a speed of 1 m/s at a distance of 15 m. By comparing the green line and the black line, we can see that, if a vehicle moved closer to the conflict point, the acceptance probability would become higher. Moreover, if the vehicle had a faster speed, it would be easy to pass the gaps between pedestrians, which can be concluded by comparing the red line and the black line. It also proves that the soft yield was more effective than the hard yield.

Table 4 shows the coefficients of the variables in the gap acceptance model, which is denoted by Eq. (1). We can see that the pedestrian distance and the vehicle speed were positively affected by the probability of the gap acceptance, whereas the pedestrian speed and the vehicle longitudinal distance had a negative effect on the model. The magnitude of the Z-values indicates that the variables were significant in the gap acceptance model.

Moreover, we evaluated the distance between pedestrians and a vehicle’s path when the vehicle arrives to the conflict point. The comparison between human drivers and our proposed model is visualized in Fig. 13. The blue line corresponds to the human drivers, and the red line is estimated from our model. The average difference between two lines was approx. 0.3 m. In addition, the red line is located at the right side of the blue line. We can conclude that our proposed model is similar to human drivers and even safer than human drivers. Moreover, the blue line indicates that more than half of the drivers maintained a distance of 3.5 m for a safe margin.

Finally, we compared our proposed model with real human driver behavior to demonstrate how human-like our model is in gap detection. The second row of Table 5 shows the similarity in gap acceptance between our proposed model and human drivers. Our model and human drivers chose the same gap in the 40 cases of 48 total sequences. However, there were 7 cases wherein our model was delayed and 1 case where it was ahead. In fact, the developed model could be considered as equivalent to the average behavior of human drivers. In addition, we also compared our model with a conventional model, which uses a hard-yield profile and fixed comfort margins [34]. The margin was set to 3.5 m, which was suggested by the human driver data in Fig. 13. The third row of Table 5 shows the similarity in gap acceptance for the conventional model. “Ahead” means that the system accepted and passed the previous gaps, which came earlier than the gap selected by the human driver. The result demonstrates that the conventional model had four more delays compared with our proposed method.

4. Conclusions and future work

This paper proposed a human-like motion planning model to represent how human drivers operate vehicles in a signalized intersection. The developed model can assess pedestrians’ crossing intention, find the appropriate gap to pass, and optimize the vehicle control profile. The performance of the system was mainly evaluated on the basis of the comparison with actual human pedestrian and driver data. The proposed motion planning model achieved an 83% recognition rate for
Fig. 14. An example of gap acceptance for the demonstration of the delay case of the conventional model.

Fig. 15. An example of gap acceptance related to waiting pedestrians.

Please cite this article as: Y. Gu, et al., Human-like motion planning model for driving in signalized intersections, IATSS Research (2016), http://dx.doi.org/10.1016/j.iatssr.2016.11.002
pedestrian intention. Moreover, the model finally selected the same gap as human drivers did in 83% of the cases.

In this paper, the developed model is represented as a left-turning vehicle as an example because it is one of the most complicated cases in the traffic configuration of Japan. However, the driving behavior in other situations and countries could be explained by the proposed methodology as well. This paper describes the motivation from the viewpoint of autonomous vehicle technology, but the proposed model can also be used for other applications in traffic engineering, such as in the simulation and analysis of the traffic effectiveness of intersections. On the basis of the analysis of the experimental result, we found that a portion of the driver’s data were not appropriate as a representative of driving maneuver. The further improvement for our model will be the mining of safe driving maneuver from human driver data. In this way, autonomous vehicles will work just like safe human drivers do. In addition, whether pedestrians recognize the vehicles or not is an important factor for decision making. Combining face detection and decision making would be a good research topic for further works.

Acknowledgment

This research is permitted by the Compliance Committee of the university.

References