



Research  
Safety for Intelligent and Connected Vehicles—Article

# Research on Anthropomorphic Obstacle Avoidance Trajectory Planning for Adaptive Driving Scenarios Based on Inverse Reinforcement Learning Theory



Jian Wu<sup>a</sup>, Yang Yan<sup>a</sup>, Yulong Liu<sup>b</sup>, Yahui Liu<sup>b,\*</sup>

<sup>a</sup>School of Mechanical and Automotive Engineering, Liaocheng University, Liaocheng 252000, China

<sup>b</sup>School of Vehicle and Mobility, Tsinghua University, Beijing 100084, China

## ARTICLE INFO

### Article history:

Received 14 October 2022  
Revised 3 April 2023  
Accepted 30 July 2023  
Available online 31 October 2023

### Keywords:

Obstacle avoidance trajectory planning  
Inverse reinforcement theory  
Anthropomorphic  
Adaptive driving scenarios

## ABSTRACT

The forward design of trajectory planning strategies requires preset trajectory optimization functions, resulting in poor adaptability of the strategy and an inability to accurately generate obstacle avoidance trajectories that conform to real driver behavior habits. In addition, owing to the strong time-varying dynamic characteristics of obstacle avoidance scenarios, it is necessary to design numerous trajectory optimization functions and adjust the corresponding parameters. Therefore, an anthropomorphic obstacle-avoidance trajectory planning strategy for adaptive driving scenarios is proposed. First, numerous expert-demonstrated trajectories are extracted from the HighD natural driving dataset. Subsequently, a trajectory expectation feature-matching algorithm is proposed that uses maximum entropy inverse reinforcement learning theory to learn the extracted expert-demonstrated trajectories and achieve automatic acquisition of the optimization function of the expert-demonstrated trajectory. Furthermore, a mapping model is constructed by combining the key driving scenario information that affects vehicle obstacle avoidance with the weight of the optimization function, and an anthropomorphic obstacle avoidance trajectory planning strategy for adaptive driving scenarios is proposed. Finally, the proposed strategy is verified based on real driving scenarios. The results show that the strategy can adjust the weight distribution of the trajectory optimization function in real time according to the “emergency degree” of obstacle avoidance and the state of the vehicle. Moreover, this strategy can generate anthropomorphic trajectories that are similar to expert-demonstrated trajectories, effectively improving the adaptability and acceptability of trajectories in driving scenarios.

© 2023 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

With the development of information and communications technology (ICT) and artificial intelligence (AI) technologies, intelligence has become an inevitable trend in automobile development [1]. Planning a collision-free trajectory from an initial state to a target state while driving is an important aspect of vehicle intelligence [2].

Trajectory planning, as a link between the decision-making and control execution layers of intelligent driving technology, has been less studied because of the lack of abundant and accurate vehicle

trajectories. Existing trajectory planning techniques can be roughly divided into four types: graph search [3–5], sampling [6–8], curve interpolation [9–11], and numerical optimization [12–14]. Based on the linear time-varying model predictive control method, the vehicle lateral obstacle avoidance trajectory planning problem was defined as a constrained optimal control problem [15]. A real-time solution to the trajectory planning problem was realized based on the spatial-temporal grid method [16]. Based on the potential energy function and vehicle-reachable set, a robust model predictive control method was adopted to generate feasible vehicle trajectories [17]. Based on vehicle-vehicle communication technology, the trajectory-planning problem has been transformed into a constrained-optimization problem to solve the obstacle avoidance trajectory satisfying driving safety and ride comfort [18]. A time-independent polynomial equation was used

\* Corresponding author.

E-mail address: [liyuhui@tsinghua.edu.cn](mailto:liyuhui@tsinghua.edu.cn) (Y. Liu).

to construct a dynamic obstacle-avoidance trajectory planning model [19]. A linear quadratic method of constrained iteration was proposed to solve the trajectory planning problem by transforming it into a nonlinear programming problem [20]. The aforementioned strategies satisfy the vehicle dynamics constraints and effectively address the safety and efficiency of obstacle avoidance trajectories.

However, these strategies mostly require the design of trajectory optimization functions in advance and cannot be extended to complex driving scenarios. The design of the trajectory optimization function plays a key role in trajectory planning [21]. Considering the complexity of driving scenarios, the dynamics of traffic flow, the randomness of driving behavior, and the game between traffic participants, it is a significant challenge for a vehicle decision system to realize autonomous obstacle avoidance trajectory planning and meet safety and comfort requirements. The establishment of a trajectory optimization function requires not only experienced engineers to carefully design each component, but also to formulate its tradeoff strategy. Highly complex and dynamic driving scenarios often require a large number of trajectory optimization functions, which makes their design more tedious and difficult and may fail to obtain the optimal solution [22].

To avoid the numerous complications associated with manual design, scholars have proposed various methods for directly or indirectly recovering planning strategies from expert-demonstrated obstacle avoidance trajectories (EDOATs). By combining imitation learning and optimization, a long time-domain model predictive control method was adopted to obtain expert obstacle avoidance trajectories [23]. A backpropagation neural network model was used to predict obstacle avoidance time to design a constant-speed offset obstacle avoidance trajectory [24]. A hyperbolic-tangent trajectory model was proposed to reconstruct the actual trajectory using a large number of real driver obstacle avoidance scenarios [25]. A trajectory optimization function was designed by constructing the relationship between the obstacle avoidance decisions of real drivers and the trajectory selection probability [26]. Although the aforementioned studies generated effective obstacle avoidance trajectories, obstacle avoidance on structured roads involving surrounding dynamic traffic entities presents a complex challenge. This necessitates a high degree of scenario adaptability for the optimization function of the obstacle avoidance trajectory. The process of establishing a mapping relationship between driving scenarios and obstacle avoidance trajectory optimization functions has not been widely explored.

In addition to security considerations, acceptability is a crucial factor in obstacle avoidance trajectory planning [27,28]. An anthropomorphic trajectory that conforms to the driving habits of drivers can enable vehicles to avoid obstacles more smoothly. This can reduce the tension of drivers and passengers, and improve ride comfort and acceptability. The Gaussian mixture model has been used as a statistical method to model drivers and can effectively describe the differences in drivers' behavioral habits [29]. Remote-learning control modes and real vehicle test data have been used to imitate human driving behavior and develop different styles of driver modeling [30]. To simulate the behavior habits of real drivers, a recurrent neural network-based long short-term memory network architecture was developed, and a trajectory planning strategy based on vision and imitation learning was proposed [31]. By integrating personalized driving habits into obstacle avoidance trajectory planning, a safe driving trajectory that adapts to different driver styles was planned [32]. The state grid method was combined with model predictive control to achieve automatic driving with personalized trajectory characteristics of the driver [33]. A convolutional neural network was used to extract track features from road images collected by cameras, and track planning of intelligent vehicles was realized through classification [34]. A

Gaussian mixture model based on neural network parameterization was proposed to predict the trajectories of vehicles at the entrance of an expressway [35]. A lane-change prediction algorithm adapted to the driver's personalized style was proposed. A trained support-vector-machine decision model was integrated into the model predictive control framework to predict the lane-change behaviors of different drivers [36]. These strategies have ensured the acceptability of obstacle avoidance trajectories and have made important contributions to the development of human-like driving.

Therefore, considering both the acceptability and adaptability of obstacle avoidance trajectory planning strategies, and avoiding the complicated process of manual parameter tuning, this study proposes anthropomorphic obstacle avoidance trajectory (AOAT) planning for adaptive driving scenarios using the HighD natural driving dataset [37]. The main contributions of this study are summarized as follows.

(1) To avoid complicated parameter adjustment of the optimization function, this study applies inverse reinforcement learning theory for the offline learning of EDOATs. A trajectory expectation feature-matching algorithm is proposed to achieve automatic recovery of the optimization function.

(2) To enhance the adaptability of the trajectory optimization function to driving scenarios, this study uses the speed of the ego vehicle and the speed difference between the ego and lead vehicles as the primary driving scenario parameters influencing obstacle avoidance trajectory planning. A mapping relationship between the driving scenario information and the obstacle avoidance trajectory optimization function is established.

The trajectory planning problem is described mathematically in Section 2. The trajectory expectation feature matching algorithm based on maximum entropy inverse reinforcement learning theory is introduced in Section 3. The mapping relationship between the trajectory cost function and key driving scenarios is established and verified in Section 4. The conclusions and future work are presented in Section 5.

## 2. Related work

### 2.1. Problem description

The problem of vehicle obstacle avoidance trajectory planning can be roughly described as follows. After receiving an obstacle avoidance instruction from the decision-making layer, the system preselects a feasible trajectory set  $R = \{r_1, r_2, \dots, r_n\}$  based on control strategies according to the current driving scenarios, where  $r_1, r_2, \dots, r_n$  represent the first, second, ...,  $n$ th feasible trajectories, respectively. Subsequently, the optimal trajectory  $r^*$  is selected to optimize the predesigned optimization function.

$$r^* = \underset{r \in \{r_1, r_2, \dots, r_n\}}{\operatorname{argmin}} \Gamma_{\text{OAF}} \quad (1)$$

where  $r^*$  is the final optimized obstacle avoidance trajectory and  $\Gamma_{\text{OAF}}$  is the trajectory optimization function.

To balance the efficiency, comfort, and safety of the trajectory planning, the optimization function of the obstacle avoidance trajectory is expressed as

$$\Gamma_{\text{OAF}} = \Gamma_{\text{OAF}}(\boldsymbol{\eta}_r, \boldsymbol{\theta}_r) \quad (2)$$

where  $\boldsymbol{\eta}_r$  is a vector composed of trajectory features, and  $\boldsymbol{\theta}_r$  is the weight matrix balancing these trajectory features,  $\boldsymbol{\theta}_r = \{\theta_r^1, \theta_r^2, \dots, \theta_r^n\}$ , where  $\theta_r^1, \theta_r^2, \dots, \theta_r^n$  represents the weight matrices of the trajectory features for the first, second, ...,  $n$ th trajectories, respectively. It should be noted that there is a complex

mapping relationship between feature weights and driving scenarios, which leads to different obstacle avoidance trajectories.

Fig. 1 shows the driving-scenario information of intelligent vehicles for obstacle avoidance on structured highways. In Fig. 1,  $v_x$  represents the longitudinal speed of the vehicle during collision avoidance. In the same driving scenario, different weights of trajectory features result in different obstacle avoidance trajectories planned by the vehicles. If the speed of the ego vehicle is much higher than that of the lead vehicle when it receives an obstacle avoidance command, the red trajectory may be better than the green and blue trajectories, based on the degree of obstacle avoidance urgency. Therefore, the driving scenarios information including the states of the ego vehicle and surrounding vehicles is defined as  $\mathbf{E}$ ,  $\mathbf{E} = \{\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_n\}$ , where  $\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_n$  represents the driving scenarios information for the first, second, ...,  $n$ th trajectories, respectively.

Eq. (2) can be rewritten as

$$\Gamma_{\text{OAF}} = \Gamma_{\text{OAF}}(\boldsymbol{\eta}_r, \boldsymbol{\theta}_r(\boldsymbol{\omega}, \mathbf{E})) \quad (3)$$

where  $\boldsymbol{\omega}$  is the weights parameter of the mapping model from driving scenarios information  $\mathbf{E}$  to trajectory feature weights  $\boldsymbol{\theta}_r$ .

To balance the efficiency, comfort, and safety of trajectory planning, the feature weights of the trajectory optimization function must be manually adjusted by engineers based on their own practical experience. Adjusting the parameters is generally tedious and requires considerable repetition. To solve this problem, this study proposes a new trajectory-planning scheme.

As shown in Fig. 2, the strategy adopted in this study is divided into two stages: offline training and online optimization. In the offline training stage, abundant expert obstacle avoidance trajectories  $\{r_1, r_2, \dots, r_n\}$  and corresponding driving scenarios information  $\{\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_n\}$  are firstly extracted from the HighD natural driving dataset. Then, using the maximum entropy inverse reinforcement learning technology, the feature weights  $\{\theta_r^1, \theta_r^2, \dots, \theta_r^n\}$  of the trajectory optimization function are extracted from the EDOATs. Subsequently, the key driving scenario information that affects the obstacle avoidance trajectory is extracted, and a mapping model of the driving scenarios to the trajectory feature weight is constructed. In this study, the model parameter  $\boldsymbol{\omega}$  is obtained using multivariate nonlinear fitting. In the online optimization stage, the optimization function of the obstacle avoidance trajectory is reconstructed based on the driving scenarios information  $\mathbf{E}$  and the mapping model obtained in the offline training stage.

## 2.2. Natural driving dataset selection and trajectory extraction

As mentioned above, to avoid a series of problems caused by manual adjustment of weights  $\theta_r$ , this research proposes an automatic recovery method of feature weights  $\theta_r$ , based on inverse reinforcement learning theory, which has been widely used as a tool to obtain optimization functions from expert examples. The core idea of inverse reinforcement learning theory is to generate strategies that match expert examples by obtaining the weights of the optimization function. Therefore, the extraction of the obstacle avoid-

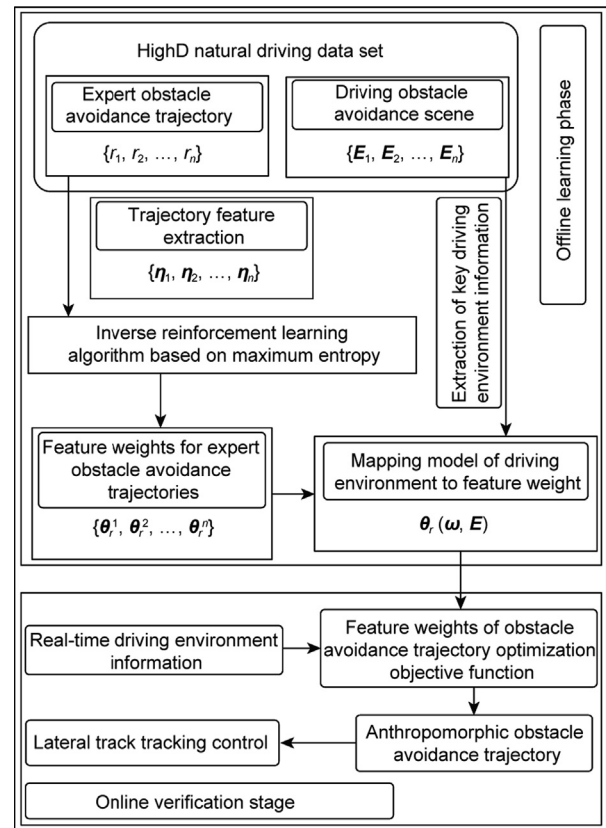


Fig. 2. AOAT planning strategy for adaptive driving scenarios.

ance trajectory of a real driver is necessary to obtain the feature weights of the trajectory optimization function.

In intelligent driving research, natural driving trajectory datasets provide an effective tool for the development and verification of intelligent driving algorithms. Currently, researchers mainly use the next-generation simulation (NGSIM), KITTI, Cityscapes, and HighD datasets. The KITTI and Cityscapes datasets mainly focus on urban road conditions. The trajectories in the NGSIM dataset show poor smoothness of lateral displacement and lateral velocity. This study focuses on obstacle avoidance trajectory planning for autonomous vehicles when driving on structured highways. Through a simple comparison, the HighD dataset was selected to develop and verify the obstacle-avoidance trajectory algorithm. The following is a description of the data used in the study.

Each frame of the HighD dataset provides 25 groups of data centered on the vehicle [37], as shown in Fig. 3. Based on the HighD dataset, obstacle avoidance trajectory planning for vehicles on structured highways was considered. For the main research objective of this study, an excellent obstacle avoidance trajectory in the dataset should first be extracted. In the process of obstacle avoidance trajectory extraction, the following items were considered:

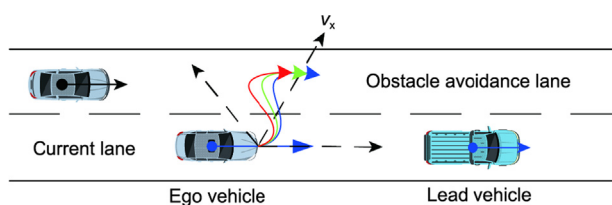


Fig. 1. Obstacle avoidance scenarios in structured highway driving.

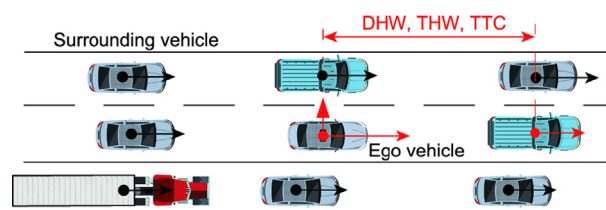


Fig. 3. Information schematic of HighD dataset. DHW: distance-to-headway; THW: time headway; TTC: time-to-collision.

(1) The HighD dataset utilized drones to capture typical road sections. Typical captured sections included ordinary highway and ramp inflow scenarios. This study focuses on the obstacle avoidance behavior of vehicles on ordinary straight roads. Therefore, the free lane-change trajectory data in the barrier-free situation in the front and the ramp inflow trajectory data are discarded in the trajectory extraction process.

(2) To improve the effectiveness of the designed obstacle avoidance trajectory, trajectory data with vehicles around, but not in the captured field of vision were discarded. Trajectory data in which the surrounding vehicles have left the captured field of vision before obstacle avoidance is complete were also discarded.

(3) The moment when the absolute value of vehicle lateral velocity reaches  $0.1 \text{ m}\cdot\text{s}^{-1}$  is considered as the starting moment of obstacle avoidance.

(4) The focus is mainly on high-speed driving obstacle avoidance scenarios, and the data of the obstacle avoidance trajectory when the vehicle speed was less than  $30 \text{ km}\cdot\text{h}^{-1}$  were discarded.

Based on the above four points, 262 groups of EDOAT data were extracted from the HighD dataset. Fig. 4 shows the extracted EDOATs.

To facilitate the design of the trajectory optimization function, the longitudinal velocity of the EDOAT was first analyzed.

The longitudinal velocity at the start time of obstacle avoidance is assumed to be  $v_{x_0}$ , the maximum or minimum longitudinal velocity from the start time to the end of obstacle avoidance is assumed to be  $v_{x_d}$ . The maximum rate of change of the longitudinal velocity during obstacle avoidance is defined in Eq. (4).

$$v_{x_t} = \frac{||v_{x_0}| - |v_{x_d}||}{|v_{x_0}|} \quad (4)$$

where  $v_{x_t}$  reflects the degree of change in longitudinal velocity during obstacle avoidance. Statistical analysis of the EDOATs was performed to obtain a statistical histogram of  $v_{x_t}$ , as shown in Fig. 5.

A total of 96.56% of the vehicles had  $v_{x_t}$  values of less than 10.00%. A total of 75.95% of the vehicles had  $v_{x_t}$  values of less than 5.00%. Therefore, it can be inferred that the degree of longitudinal velocity change is very small during obstacle avoidance. This is closely related to the ride comfort and safety of vehicles. Therefore, it is assumed that the longitudinal velocity of the vehicle was constant during obstacle avoidance.

### 3. Methodology

#### 3.1. Trajectory expectation feature matching algorithm based on inverse reinforcement learning

The trajectory optimization function plays a key role in the design of an obstacle-avoidance trajectory planning strategy. Optimizing a function based on an engineer's manual design is time consuming and laborious. Although the obstacle-avoidance func-

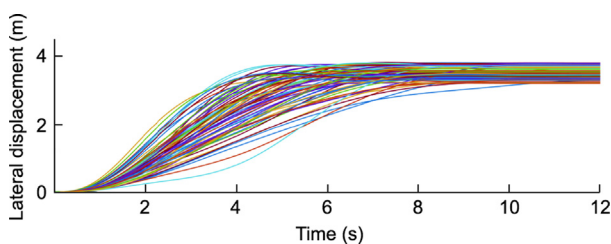


Fig. 4. EDOATs extracted from the HighD dataset.

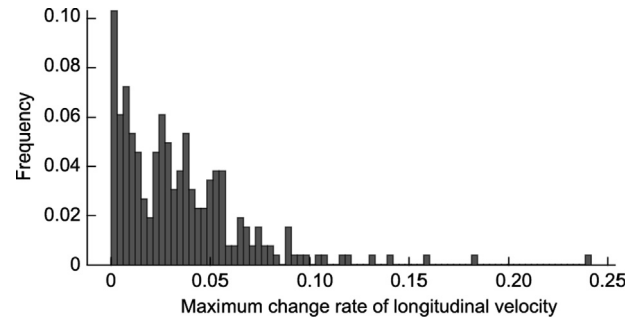


Fig. 5. Maximum change rate of longitudinal velocity during obstacle avoidance.

tion is satisfied, the driving habits of real drivers are ignored, which leads to poor acceptability.

#### 3.1.1. Principle of inverse reinforcement algorithm

The key of inverse reinforcement learning is to find the desired parameter  $\theta_r$  such that the obstacle avoidance trajectory generated by the trajectory optimization function  $\Gamma_{\text{OAF}}$  is similar to the EDOAT. Trajectory  $r$  comprises a continuous sequence of states.

$$r = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_i, \dots, \mathbf{x}_t] \quad (5)$$

where  $\mathbf{x}_i$  is the state vector of the agent at moment  $i$ ,  $i = 1, 2, \dots, t$ . The trajectory optimization function is the mapping of state feature vector  $\boldsymbol{\eta}_r$  to the state feature weights. The state feature vector  $\boldsymbol{\eta}_r$  is defined as the related features that affect the driving state of the vehicle and the surrounding vehicles during obstacle avoidance trajectory planning.

In this study, a linear relationship between optimization function  $\Gamma_{\text{OAF}}$  and state feature vector  $\boldsymbol{\eta}_r$  is preliminarily assumed.

$$\begin{aligned} \Gamma_{\text{OAF}} &= \theta_r^T \\ \text{s.t. } \boldsymbol{\eta}_r &= \sum_{\mathbf{x}_i \in r} \boldsymbol{\eta}_{\mathbf{x}_i} \in \mathbb{R}^n \end{aligned} \quad (6)$$

Based on the above definition, the objective of inverse reinforcement learning can be described as follows. Given a set of expert-demonstrated trajectories and obtaining its state feature weights  $\theta_r$ , the optimization function  $\Gamma_{\text{OAF}}$  parameterized by the feature weights can generate obstacle avoidance trajectories similar to expert-demonstrated trajectories. That is, the probability distribution  $p(r)$  of the trajectory is expected to be obtained such that the expectation of the trajectory features that satisfy the probability distribution model is consistent with the empirical features  $\tilde{\boldsymbol{\eta}}_r$  of the demonstration trajectory.

$$\begin{aligned} E_{p(r)}[\boldsymbol{\eta}_r] &= \tilde{\boldsymbol{\eta}}_r \\ \tilde{\boldsymbol{\eta}}_r &= \left(\frac{1}{m_z}\right)(\boldsymbol{\eta}_1 + \boldsymbol{\eta}_2 + \dots + \boldsymbol{\eta}_{m_z}) \end{aligned} \quad (7)$$

where  $m_z$  is the number of EDOATs. Because the only constraint condition introduced is given by Eq. (7), the maximum entropy principle best reflects existing information. Therefore, the trajectory probability distribution obtained in this study can be expressed as

$$\begin{aligned} p^*(r) &= \underset{p(r)}{\operatorname{argmax}} L(p) = \underset{p(r)}{\operatorname{argmax}} \int_r -p(r) \log p(r) dr \\ \text{s.t. } \int_r p(r) dr &= 1 \\ \tilde{\boldsymbol{\eta}}_r &= \int_r p(r) \boldsymbol{\eta}_r dr \end{aligned} \quad (8)$$

where  $L(p)$  is defined as the distribution entropy of trajectory probability distribution model  $p(r)$ . Based on the Lagrange multiplier method, the Lagrange function is constructed as expressed by Eq. (9).



$$F(p(r), \lambda, \theta_r) = \int_r -p(r) \log p(r) dr + \theta_r^T \left[ \int_r p(r) \eta_r dr - \tilde{\eta}_r \right] + \lambda \left[ \int_r p(r) dr - 1 \right] \quad (9)$$

where  $\lambda$  and  $\theta_r$  are Lagrange multipliers. Thus, Eq. (9) is transformed into the optimization problem of Lagrange function  $F(p(r), \lambda, \theta_r)$ .

$$p^*(r), \lambda^*, \theta_r^* = \operatorname{argmax}_{p(r), \lambda, \theta_r} F(p(r), \lambda, \theta_r) \quad (10)$$

According to the variational method:

$$p(r|\theta_r) = \frac{\exp(-\theta_r^T \eta_r)}{\exp(1 - \lambda)} \quad (11)$$

$$\exp(1 - \lambda) \triangleq Z(\theta_r) = \int_r \exp(-\theta_r^T \eta_r) dr \quad (12)$$

where  $Z(\theta_r)$  is the partition function.

Finally, the trajectory probability distribution model is given by Eq. (13).

$$p(r|\theta_r) = \frac{\exp(-\theta_r^T \eta_r)}{Z(\theta_r)} \quad (13)$$

The distribution entropy of the maximization system under feature matching is equivalent to the likelihood of a maximization expert-demonstrated trajectory under the exponential probability distribution model. Therefore, the distribution parameter  $\theta_r$  of the probability distribution model is given by Eq. (14).

$$\theta_r^* = \operatorname{argmax}_{p(r, \theta_r)} F(p, \theta_r) = \operatorname{argmax}_{p(r, \theta_r)} \sum_{\tilde{r}_i \in \tilde{R}} \log p(\tilde{r}_i|\theta_r) \quad (14)$$

where  $\tilde{R}$  represents the set of EDOATs,  $\tilde{r}_i$  represents the  $i$ th EDOAT.

Eq. (14) does not generally provide an analytical solution. Therefore, this study obtains the value of  $\theta_r$  through numerical iteration. To achieve this, the gradient of the likelihood pair distribution parameter  $\theta_r$  of the expert-demonstrated trajectory must be determined.

$$\log p(\tilde{r}_i|\theta_r) = -\theta_r^T \eta_r - \log Z(\theta_r) \quad (15)$$

The partial derivative with respect to  $\theta_r$  yields Eqs. (16) and (17).

$$\frac{\partial}{\partial \theta_r} \log p(\tilde{r}_i|\theta_r) = -\eta_r + \int_r p(r|\theta_r) \eta_r dr \quad (16)$$

$$\frac{\partial}{\partial \theta_r} \sum_{\tilde{r}_i \in \tilde{R}} \log p(\tilde{r}_i|\theta_r) = -m_z \tilde{\eta}_r + m_z \int_r p(r|\theta_r) \eta_r dr \quad (17)$$

The gradient of the distribution parameter  $\theta_r$  is equal to the difference between the expected trajectory features and empirical features of the expert-demonstrated trajectory under the exponential distribution model, as given by Eq. (16).

When the feature of a trajectory in the exponential distribution model is large, the corresponding weight coefficient of the feature increases, as given by Eq. (17). Combined with Eq. (13), the probability of the probability distribution model choosing this trajectory is “exponentially” reduced. Finally, the trajectory features are reduced to the empirical characteristics of the expert-demonstrated trajectory.

Obtaining the gradient requires the expectations of the trajectory features to be calculated using an exponential probability distribution model. Calculating the expectations of trajectory features using a probabilistic model for high-dimensional continuous space problems is difficult. Similar to Ref. [38], inverse optimal control

theory is adopted in this study. Therefore, the feature of the trajectory with the highest probability is used to approximate the expectation of the trajectory features under the exponential probability distribution model.

$$E_{p(r)}[\eta_r] = \int_r p(r|\theta_r) \eta_r dr \approx \eta \left( \operatorname{argmax}_r p(r|\theta_r) \right) \quad (18)$$

The final gradient,  $G_r$ , is expressed by Eq. (19).

$$G_r = -\tilde{\eta}_r + \eta \left( \operatorname{argmax}_r p(r|\theta_r) \right) \quad (19)$$

Based on the abovementioned inverse reinforcement learning process, the trajectory expectation feature matching algorithm is given as follows.

### 3.1.2. Trajectory expectation feature matching algorithm

In contrast to the global path planning algorithm, the trajectory provided by the local path planning algorithm is generally a continuous function of space coordinates at a certain time in the future.

$$\begin{aligned} x &= x(t), t \in [t_0, t_0 + t_p] \\ y &= y(t), t \in [t_0, t_0 + t_p] \end{aligned} \quad (20)$$

where  $x$  and  $y$  are the abscissa and ordinate of the obstacle avoidance trajectory, respectively, relative to a certain coordinate system;  $t_0$  is the initial obstacle avoidance time;  $t_p$  is the total duration of obstacle avoidance.

According to Fig. 5, the change in the longitudinal velocity during obstacle avoidance can be ignored; thus,  $x$  can be expressed as

$$x = x(t) = x_{t_0} + t_p v_{t_0} \quad (21)$$

where  $v_{t_0}$  represents the longitudinal velocity at the starting time of obstacle avoidance.

The lateral obstacle avoidance trajectory is a continuous function of time; therefore, the obstacle avoidance trajectory optimization problem is a typical infinite-dimensional optimization problem, which must be transformed into a finite-dimensional optimization problem. By introducing the “track support point,” the vehicle’s obstacle avoidance trajectory is represented by a quintic spline curve. Thus, the infinite-dimensional optimization problem is transformed into a finite-dimensional optimization problem.

According to the extracted EDOATs, most vehicles completed obstacle avoidance behavior within 9 s, as shown in Fig. 4. In this study, the starting time of obstacle avoidance was taken as the starting point on the time axis, and seven discrete points were taken at equal intervals of 1.5 s, as shown in Fig. 6. The lateral position corresponding to initial moment  $t_0$  is the lateral position of the vehicle at the initial moment of obstacle avoidance, which is

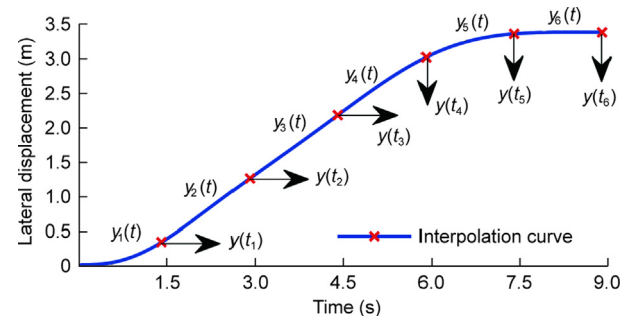


Fig. 6. Quintic spline interpolation curve of lateral obstacle avoidance trajectory based on trajectory support points.

assumed to be known in this study. Finally, the point set  $\pi_{ps}$  composed of six “trajectory support points” is selected as the interpolation point for the spline interpolation of the obstacle avoidance trajectory.

$$\pi_{ps} = [y(t_1), y(t_2), \dots, y(t_6)] \quad (22)$$

The point set  $\pi_{ps}$  divides the vehicle’s lateral obstacle avoidance trajectory into six segments. Based on the principle of quintic spline interpolation, the  $j$ th segment  $y_j(t)$  of the lateral obstacle avoidance trajectory is represented by Eq. (23).

$$\begin{aligned} y_j(t) = & (120h_j)^{-1}(\alpha_j^5 M_{j+1} - \alpha_{j+1}^5 M_j) + (6h_j)^{-1}(\alpha_j^3 T_{j+1} - \alpha_{j+1}^3 T_j) \\ & - \left(\frac{h_j}{36}\right)(\alpha_j^3 M_{j+1} - \alpha_{j+1}^3 M_j) + (h_j)^{-1}(\alpha_j y(t_{j+1}) - \alpha_j y(t_j)) \\ & + \left(\frac{7h_j^3}{360}\right)(\alpha_j M_{j+1} - \alpha_{j+1} M_j) - \left(\frac{h_j}{6}\right)(\alpha_j T_{j+1} - \alpha_{j+1} T_j), \end{aligned} \quad (23)$$

$j = 1, 2, 3, 4, 5, 6$

where  $h_j = t_j - t_{j-1}$ ,  $\alpha_j = y - y(t_j)$ ,  $M_j$  is the fourth derivative of the curve, and  $T_j$  is the second derivative of the curve.

According to the continuity of the first and third derivatives at the interpolation points at both ends of the quintic spline interpolation curve, the following can be obtained:

$$\begin{aligned} \lambda_{j-1} M_{j-1} + \mu_j M_j + \lambda_j M_{j+1} + I_{j-1} T_{j-1} + I_j T_j + I_j T_{j+1} &= 0 \\ p_{j-1} M_{j-1} + q_j M_j + p_j M_{j+1} + \lambda_{j-1} T_{j-1} + \mu_j T_j + \lambda_j T_{j+1} &= g_j \end{aligned} \quad (24)$$

$j = 2, 3, 4, 5$

where

$$\lambda_j = \frac{h_j}{6}, \mu_j = 2(\lambda_{j-1} + \lambda_j), I_j = -(h_j)^{-1}, l_j = 1(I_{j-1} + I_j),$$

$$p_j = -\left(\frac{7h_j^3}{360}\right), q_j = \left(\frac{8p_{j-1} + 8p_j}{7}\right),$$

$$g_j = \left\{ \frac{[y(t_{j+1}) - y(t_j)]}{h_j} \right\} - \left\{ \frac{[y(t_j) - y(t_{j-1})]}{h_{j-1}} \right\},$$

$j = 2, 3, 4, 5$ .

According to Eqs. (23) and (24), there are 12 unknown variables and eight constraint equations. Four of the constraint equations are given by the boundary conditions. The other four constraint equations are obtained by setting the lateral velocity and lateral acceleration of the starting and ending positions of obstacle avoidance to 0. The lateral obstacle avoidance trajectory is expressed as a function of the “trajectory support point” and time, as expressed by Eq. (25).

$$y = y(\pi_{ps}, t) \quad (25)$$

Thus, the optimization function of the obstacle avoidance trajectory is defined as a linear combination of the trajectory feature  $\eta_r$  and feature weight vector  $\theta_r$ . The features of the obstacle avoidance trajectories must be designed in advance. The following three trajectory features are used to describe the trajectory of the vehicle obstacle avoidance:

$$\eta_{r_1} = \frac{\int_t \|y(\pi_{ps}, t) - y_d\|^2 dt}{\int_t \|M_{lpd}\|^2 dt} \quad (26)$$

$$\eta_{r_2} = \frac{\int_t \|\dot{y}(\pi_{ps}, t)\|^2 dt}{\int_t \|M_{lv}\|^2 dt} \quad (27)$$

$$\eta_{r_3} = \frac{\int_t \|\ddot{y}(\pi_{ps}, t)\|^2 dt}{\int_t \|M_{la}\|^2 dt} \quad (28)$$

where  $\eta_{r_1}$ ,  $\eta_{r_2}$ , and  $\eta_{r_3}$  represent the lateral position deviation, lateral velocity, and lateral acceleration characteristics of the trajectory, respectively;  $M_{lpd}$ ,  $M_{lv}$ , and  $M_{la}$  are the maximum lateral position difference, lateral velocity, and lateral acceleration, respectively, during the obstacle avoidance process.  $y_d$  is the abscissa of the vehicle when the expected obstacle avoidance is complete,  $\dot{y}(\pi_{ps}, t)$  and  $\ddot{y}(\pi_{ps}, t)$  respectively represent the first-order and second-order derivatives of  $y(\pi_{ps}, t)$  with respect to time.

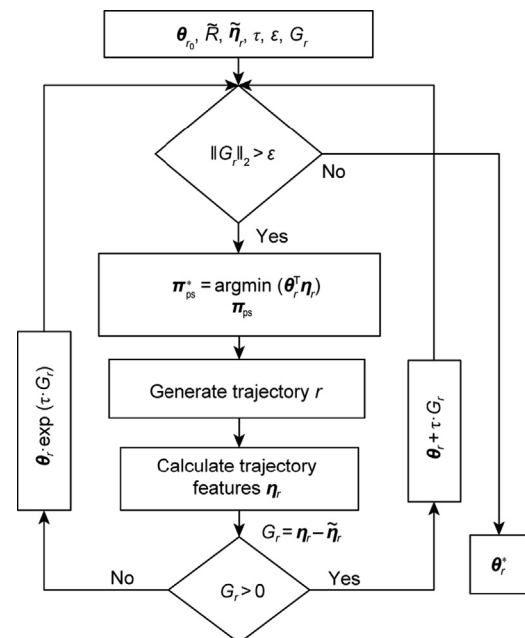
The vehicle obstacle avoidance process can be divided into obstacle avoidance risk assessment, obstacle avoidance target lane selection, and trajectory planning. The first two points are the content of obstacle avoidance behavior decisions, and the focus of this study is vehicle lateral obstacle avoidance trajectory planning. Therefore, it is assumed that the obstacle avoidance instructions and expected obstacle avoidance lanes are known.

The trajectory features in Eq. (26) describe the distance between the lateral positions of the vehicle and target during obstacle avoidance. The trajectory features in Eq. (27) describe the speed at which the vehicle approaches the target position. The trajectory features in Eq. (28) describe the acceleration of the vehicle as it approaches the target position. To a certain extent, this reflects the comfort of the obstacle avoidance trajectory. Combined with the three proposed trajectory features, the optimization function of the obstacle avoidance trajectory can be written as

$$\Gamma_{OAF} = \theta_r^T \eta_r = \theta_{r_1} \eta_{r_1} + \theta_{r_2} \eta_{r_2} + \theta_{r_3} \eta_{r_3} \quad (29)$$

where  $\theta_{r_1}$ ,  $\theta_{r_2}$ , and  $\theta_{r_3}$  are the weight coefficients of  $\eta_{r_1}$ ,  $\eta_{r_2}$ , and  $\eta_{r_3}$ , respectively.

Next, the inverse reinforcement learning algorithm is combined with the obstacle avoidance trajectory based on the quintic spline curve proposed in this study. The trajectory expectation feature matching algorithm based on inverse reinforcement learning is illustrated in Fig. 7.



**Fig. 7.** Trajectory expectation feature matching algorithm based on inverse reinforcement learning.  $\|G_r\|_2$ : the binary norm of the update gradient of the trajectory feature weight;  $\epsilon$ : a pre-given value;  $\tau$ : the coefficient;  $\theta_0$ : the initial value of  $\theta_r$ .

In Fig. 7, when the binary norm  $\|G_r\|_2$  of the update gradient of the trajectory feature weight is less than a pre-given value  $\varepsilon$ , the algorithm terminates. During the learning process, feature weight vector  $\theta_r$  is always positive. When  $G_r$  is negative,  $G_r$  is first multiplied by the coefficient  $\tau$ . Then, we take the exponent of  $G_r \cdot \tau$  and multiply it by  $\theta_r$  to achieve the update.

### 3.2. Obstacle avoidance trajectory planning for adaptive driving scenarios

The following obstacle avoidance behaviors of experienced drivers on a structured expressway are closely related to the surrounding driving scenarios. Based on the above model, this section explores the quantitative model relationship between the optimization function of the obstacle avoidance trajectory and the surrounding driving scenarios. An AOAT planning strategy for adaptive driving scenarios is proposed.

To study the quantitative model of the optimization function and driving scenarios, we define the driving scenarios information  $\mathbf{E}$ , which includes the states of the ego vehicle and surrounding vehicles. Firstly, the mapping relationship between obstacle avoidance environment  $\mathbf{E}$  and weight coefficient  $\theta_r$  of the obstacle avoidance trajectory features is obtained.

$$\begin{aligned} \theta_{r_1} &= \theta_{r_1}(\omega_{r_1}, \mathbf{E}) \\ \theta_{r_2} &= \theta_{r_2}(\omega_{r_2}, \mathbf{E}) \\ \theta_{r_3} &= \theta_{r_3}(\omega_{r_3}, \mathbf{E}) \end{aligned} \quad (30)$$

where  $\omega_r$  is used to represent the mapping relationship between trajectory feature weights and critical driving scenarios,  $\omega_r = [\omega_{r_1}, \omega_{r_2}, \omega_{r_3}]$ .

The mapping model must be interpretable and transferable. Therefore, a polynomial regression model is used to study the mapping relationship between driving scenarios  $\mathbf{E}$  and weight coefficients  $\theta_r$ .

$$\begin{aligned} \theta_{r_1} &= \alpha_0 + \alpha_1 * \mathbf{E}_1 + \alpha_2 * \mathbf{E}_2 + \alpha_3 * \mathbf{E}_1^2 + \alpha_4 * \mathbf{E}_1 * \mathbf{E}_2 \\ \theta_{r_2} &= \beta_0 + \beta_1 * \mathbf{E}_1 + \beta_2 * \mathbf{E}_2 + \beta_3 * \mathbf{E}_1^2 + \beta_4 * \mathbf{E}_1 * \mathbf{E}_2 \\ \theta_{r_3} &= \gamma_0 + \gamma_1 * \mathbf{E}_1 + \gamma_2 * \mathbf{E}_2 + \gamma_3 * \mathbf{E}_1^2 + \gamma_4 * \mathbf{E}_1 * \mathbf{E}_2 \end{aligned} \quad (31)$$

where  $\alpha_0$ – $\alpha_4$ ,  $\beta_0$ – $\beta_4$ , and  $\gamma_0$ – $\gamma_4$  are the fitting parameters of  $\theta_r$ , respectively, and  $\mathbf{E}_1$  and  $\mathbf{E}_2$  denote the key environmental information that affect the weight coefficients,  $\mathbf{E} = [\mathbf{E}_1, \mathbf{E}_2]$ .

## 4. Strategy validation and results analysis

### 4.1. Verification of inverse reinforcement learning algorithm based on feature matching

To verify the effectiveness of the algorithm, an EDOAT (demonstration trajectory 1) was randomly selected from the HighD dataset to verify the algorithm. The initial trajectory feature weight  $\theta_{r_0}$  in the maximum entropy inverse reinforcement learning algorithm was set to [0.1, 10.0, 150.0].

After 76 iterations, the binary norm of update gradient  $G_r$  of the trajectory feature weights is less than the set value  $\varepsilon$ . The optimal feature weight vector  $\theta_r^*$  obtained by inverse reinforcement learning theory is [18.621, 0.479, 1.713].

The blue line represents the lateral position, velocity, and acceleration of demonstration trajectory 1, as shown in Fig. 8. The arrow indicates the change in the direction of the features during the learning process. The final learned obstacle avoidance trajectory has a good degree of feature matching with demonstration trajectory 1, which verifies the effectiveness of the learning algorithm.

To further quantify the learning effect of the algorithm, the change of feature gradient in the learning process is given.

Figs. 9(a)–(c) show the gradient changes in the features given in Eqs. (26) and (28) during the learning process. As the learning continued, the gradients of the three trajectory features gradually approached zero, as shown in Fig. 9. The difference between the features of the demonstration trajectory and the feature expectations of the trajectory decreased.

Next, another expert trajectory (demonstration trajectory 2) was randomly selected from the HighD dataset for algorithm verification. The initial trajectory feature weight  $\theta_{r_0}$  was set to [0.1, 10.0, 150.0].

After 86 iterations, the binary norm of update gradient  $G_r$  of trajectory feature weights is less than the set value  $\varepsilon$ , as shown in Figs. 10 and 11. The optimal feature weight vector  $\theta_r^*$  obtained by inverse reinforcement learning theory is [15.471, 0.452, 1.105].

To explore the influence of the initial track feature weight vector on the final results, we set the initial trajectory feature weights  $\theta_{r_0}$  to [150.0, 1.0, 0.1] for the same demonstration trajectory (demonstration trajectory 2).

After 38 iterations, the binary norm of update gradient  $G_r$  of trajectory feature weights is less than the set value  $\varepsilon$ , as shown in Figs. 12 and 13. The optimal feature weight vector  $\theta_r^*$  obtained using inverse reinforcement learning theory is [51.282, 1.456, 3.655].

It can be seen that under different initial weights of the trajectory features, the final trajectory learned is very similar, but there are significant differences in the numerical weights of the trajectory features. This is because the absolute value of each feature weight coefficient in the trajectory optimization function does not determine the trajectory features, that is, the relative value. The first weighting coefficient  $\theta_{r_1}$  is set to unit 1. Under completely different initial values of the two feature weights, the ratios of  $\theta_{r_2}$  and  $\theta_{r_3}$  to  $\theta_{r_1}$  are [0.0292, 0.0714] and [0.0284, 0.0713], respectively. Under the two different initial values of the feature weights, the relative changes in the feature weight coefficients were less than 3%, as shown in Fig. 14.

To simplify the calculations, the ratio of the feature weights was used to describe the trajectory optimization function.

$$\Gamma_{\text{OAF}} = \theta^T \boldsymbol{\eta}_r = \eta_{r_1} + \left(\frac{\theta_{r_2}}{\theta_{r_1}}\right) \eta_{r_2} + \left(\frac{\theta_{r_3}}{\theta_{r_1}}\right) \eta_{r_3} = \eta_{r_1} + \theta_1 \eta_{r_2} + \theta_2 \eta_{r_3} \quad (32)$$

Based on the HighD dataset, the obstacle-avoidance trajectories of 262 groups of vehicles on structured roads were extracted. The extracted trajectories are learned to obtain the statistical results of feature weight vector  $\theta$ ,  $\theta = \left[1, \left(\frac{\theta_{r_2}}{\theta_{r_1}}\right), \left(\frac{\theta_{r_3}}{\theta_{r_1}}\right)\right]^T$ ,  $\theta_1 = \left(\frac{\theta_{r_2}}{\theta_{r_1}}\right)$ ,  $\theta_2 = \left(\frac{\theta_{r_3}}{\theta_{r_1}}\right)$ .

### 4.2. Validation of the mapping model

To effectively select the key environmental information that affects the vehicle obstacle avoidance trajectory planning behavior, the driving scenario information was analyzed from a practical perspective. Generally speaking, the actual length of the driver's obstacle avoidance time is mainly affected by "obstacle avoidance emergency degree," "obstacle avoidance risk," and the state of the vehicle. When the time headway (THW) between the ego vehicle and lead vehicle is small, the driver is more inclined to complete obstacle avoidance over a shorter time. However, if the speed is high, the driver may be more inclined to complete the obstacle avoidance task over a longer time to ensure the comfort of obstacle avoidance and reduce the tension of the passengers.

Based on practical experience, drivers usually perform obstacle avoidance operations to leave the congested lane and improve driving efficiency.

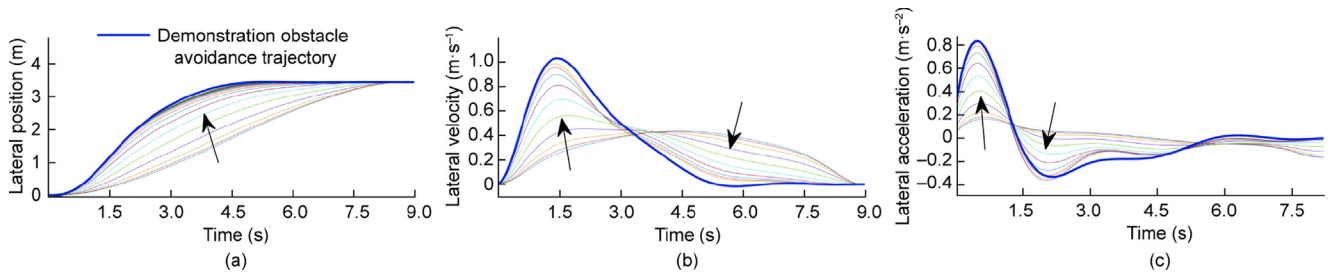


Fig. 8. Feature matching process for demonstration trajectory 1.

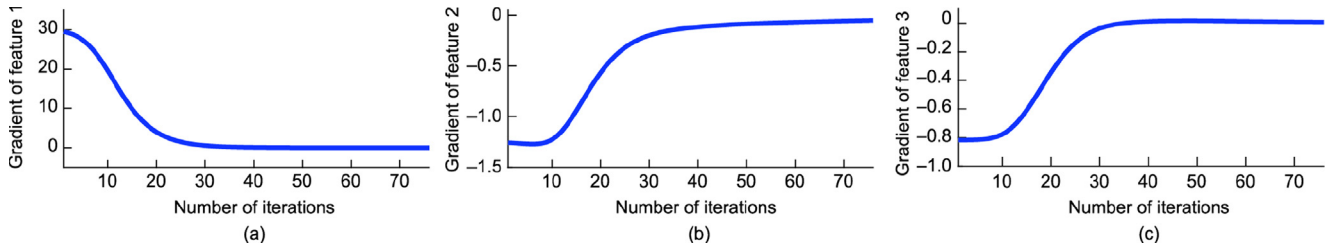


Fig. 9. Gradient change during feature matching for demonstration trajectory 1.

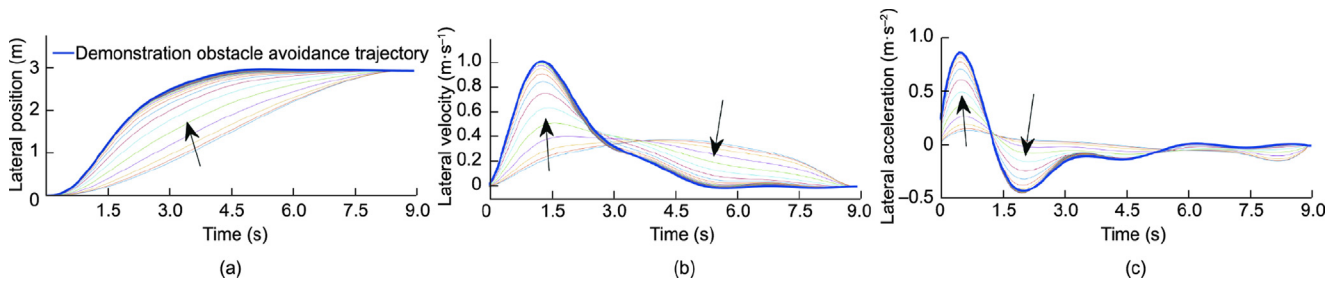


Fig. 10. Feature matching process for demonstration trajectory 2.

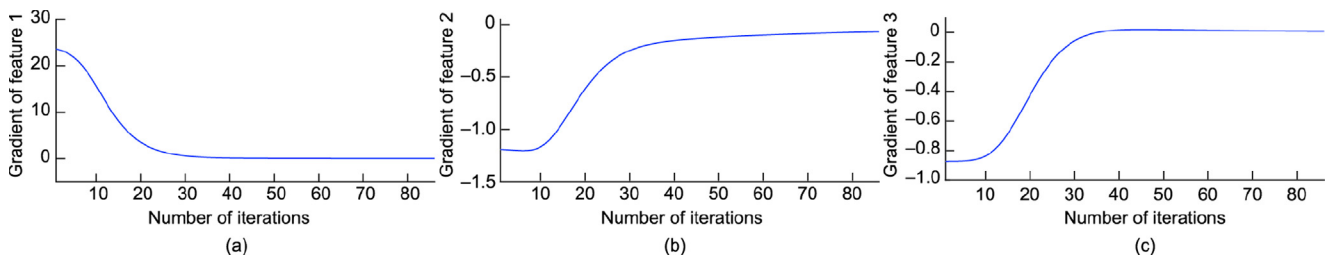


Fig. 11. Gradient change during feature matching for demonstration trajectory 2.

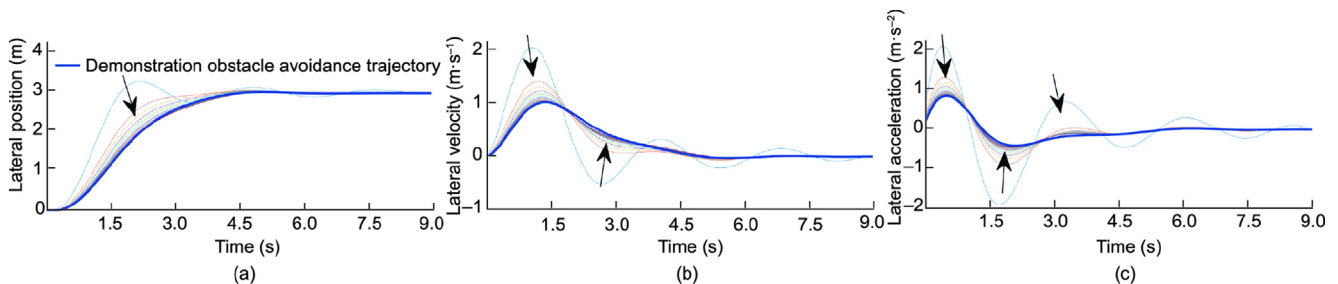


Fig. 12. Feature matching process for demonstration trajectory 2.



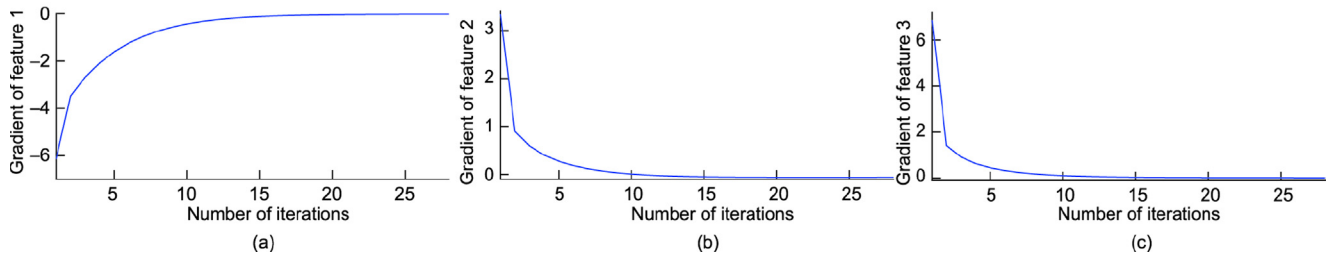


Fig. 13. Gradient change during feature matching for demonstration trajectory 2.

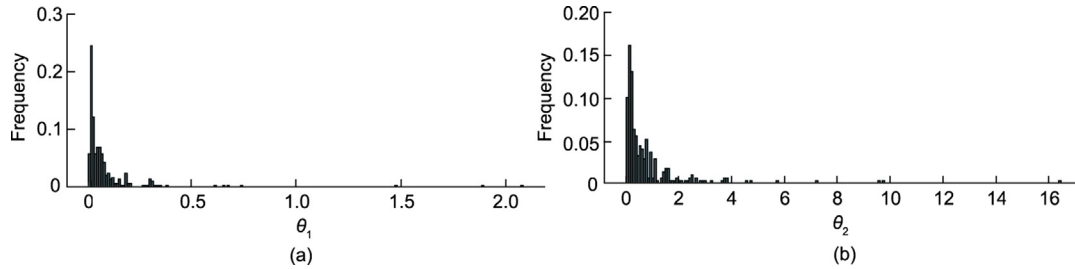


Fig. 14. Learning results of obstacle avoidance trajectory feature weights. (a) Statistical histogram of  $\theta_1$ ; (b) statistical histogram of  $\theta_2$ . For most trajectories,  $\theta_1$  is less than 0.5, and  $\theta_2$  is less than 5.0.

Based on the analysis of 262 groups of obstacle avoidance information extracted from the HighD dataset, it was found that approximately 83.97% of the vehicle speeds were higher than that of the lead vehicle when obstacles were avoided. Approximately 85.50% of the vehicles avoided obstacles with a distance-to-headway (DHW) of less than 100 m; approximately 91.61% of vehicles avoided obstacles with a THW of less than 4 s; and approximately 90.84% of the vehicles avoided obstacles with a time-to-collision (TTC) of less than 50 s.

After the vehicle receives the obstacle avoidance command, the speed difference between the ego and lead vehicles, DHW, and TTC can be used as key driving scenario information to represent the urgency of obstacle avoidance, as shown in Fig. 15.

Considering the above situation, the speed difference between the ego and lead vehicles is regarded as one of the key factors affecting vehicle obstacle avoidance trajectory planning. It is worth noting that the vehicle speed also affects the choice of obstacle avoidance trajectory. When the speed is high, the driver usually chooses a trajectory with a smaller curvature to reduce passenger tension and avoid vehicle instability caused by urgent obstacle avoidance. Therefore, to make the obstacle avoidance trajectory more anthropomorphic and acceptable, the vehicle speed is also used as one of the key factors affecting vehicle trajectory planning.

In summary,  $\Delta v_x$  and  $v_x$  are key scenario factors affecting obstacle avoidance trajectory planning. The relative weight coefficients of each feature in the trajectory optimization function determine

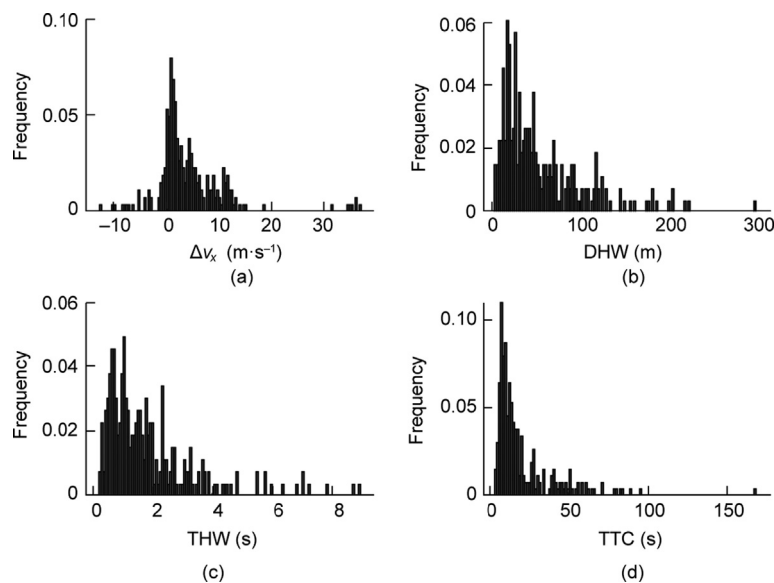


Fig. 15. Statistical histograms of obstacle avoidance trajectory data. (a) Speed difference between the ego vehicle and lead vehicle; (b) DHW; (c) THW; (d) TTC.

the trajectory features. Therefore, the mapping relationship between the key scenario factors  $\mathbf{E}$  and the weight coefficients  $\theta$  of trajectory features is redefined.

$$\begin{aligned} \theta_1 &= \theta_1(\omega_1, \mathbf{E}) \\ \theta_2 &= \theta_2(\omega_2, \mathbf{E}) \end{aligned} \quad (33)$$

Nonlinear fitting technology was adopted for the expert-demonstrated trajectories to obtain the mapping relation from the driving scenario information to the feature weight of the trajectory optimization function:

$$\begin{aligned} \theta_1 &= -0.08414 - 0.005121 * \Delta v_x + 0.0126 * v_x \\ &\quad + 0.0005247 * \Delta v_x^2 - 0.0005662 * \Delta v_x * v_x \\ \theta_2 &= -0.4147 - 0.12 * \Delta v_x + 0.1207 * v_x \\ &\quad + 0.009817 * \Delta v_x^2 - 0.00752 * \Delta v_x * v_x \end{aligned} \quad (34)$$

The fitted curves corresponding to Eq. (34) are shown in Figs. 16(a) and (b), respectively.

Table 1 lists the fitting results of the mapping model between the driving scenarios and weights of the trajectory features. According to Table 1, the estimation model can better explain the mapping relationship between the key information of obstacle avoidance scenarios and feature weights.

As shown in Fig. 16,  $\theta_1$  and  $\theta_2$  are subtractive functions of  $\Delta v_x$ ;  $\theta_1$  and  $\theta_2$  will decrease as the speed difference between the ego vehicle and lead vehicle increases. The greater the vehicle speed difference  $\Delta v_x$ , the higher is the urgency of vehicle obstacle avoidance and collision risk. Therefore, the corresponding obstacle avoidance completion time should be shorter, and the corresponding  $\theta_1$  and  $\theta_2$  should be smaller. Because  $\theta_1$  and  $\theta_2$  are increasing functions of  $v_x$ , the higher the speed of the ego vehicle when avoiding obstacles, the greater are  $\theta_1$  and  $\theta_2$ . This can be explained in terms of obstacle avoidance risk. When the vehicle is moving at high speed, an obstacle avoidance time that is too short will increase the risk of dangerous conditions, such as rollover, thereby increasing the driver's nervousness. Therefore, when the speed is high, the corresponding obstacle avoidance time should be longer, and the corresponding  $\theta_1$  and  $\theta_2$  values should be higher.

At this point, the offline learning phase is completed. The effectiveness of the obstacle avoidance trajectory planning strategy is verified as follows.

### 4.3. Simulation and experimental verification

In this study, 262 groups of real vehicle obstacle avoidance trajectory data points were extracted from the HighD dataset. A total of 220 groups of trajectory data were used as samples to train the mapping relationship between the key information of the driving scenarios and feature weights of the trajectory optimization function. The effectiveness of the proposed trajectory-planning strategy was verified based on the remaining 42 groups of real obstacle-avoidance trajectory data.

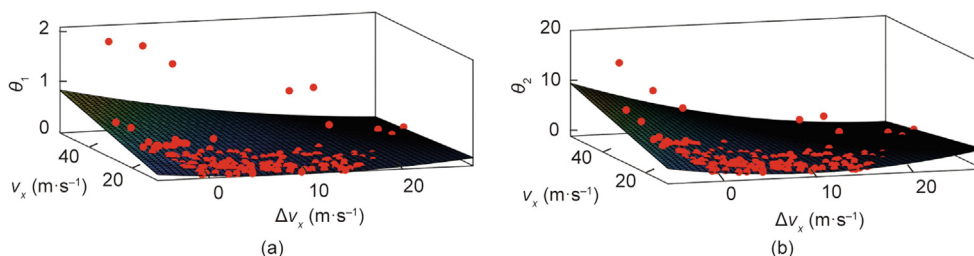


Fig. 16. Mapping relationship between driving scenarios and weights of trajectory features. (a)  $\theta_1$  and driving environment; (b)  $\theta_2$  and driving environment.

Table 1  
Evaluation results of mapping model fitting.

Weight	SSE	$R^2$	RMSE
$\theta_1$	0.1654	0.9853	0.03129
$\theta_2$	0.6493	0.9424	0.06217

SSE: the sum of squares dueto error;  $R^2$ : coefficient of determination; RMSE: root mean squared error.

First, a trajectory optimization function for comparison (Scheme 1) was constructed by taking the average of the learned feature weights from 220 obstacle avoidance trajectories. The trajectory optimization function constructed from the feature weights obtained through the mapping of driving scenario information is shown in Scheme 2. To more intuitively quantify the advantages and disadvantages of the two strategies, the trajectory feature distance vector  $\mathbf{D}_{\eta_r}$  is introduced,  $\mathbf{D}_{\eta_r} = [D_{\eta_1}, D_{\eta_2}, D_{\eta_3}]^T$ .

$$\begin{aligned} D_{\eta_1} &= \|\eta_{r_1} - \tilde{\eta}_{r_1}\|^2 \\ D_{\eta_2} &= \|\eta_{r_2} - \tilde{\eta}_{r_2}\|^2 \\ D_{\eta_3} &= \|\eta_{r_3} - \tilde{\eta}_{r_3}\|^2 \end{aligned} \quad (35)$$

where  $\eta_{r_1} - \eta_{r_3}$  are, respectively, the features of the obstacle avoidance trajectory generated by the trajectory planning algorithm; and  $\tilde{\eta}_{r_1} - \tilde{\eta}_{r_3}$  are, respectively, the features of the EDOAT. A specific description is provided in Eqs. (26)–(28). According to the definition of the trajectory feature distance vector  $\mathbf{D}_{\eta_r}$ , this value describes the distance between the obstacle avoidance trajectory generated by the planning algorithm and features of the expert-demonstrated trajectory. The smaller  $\mathbf{D}_{\eta_r}$  is, the closer the generated obstacle avoidance trajectory is to the expert-demonstrated trajectory.

The proposed algorithm (Scheme 2) and comparison scheme (Scheme 1) were used to generate trajectories for 42 groups of vehicle–obstacle avoidance test scenarios. The average value of  $\mathbf{D}_{\eta_r}$  obtained by the two schemes in 42 groups of test scenarios is taken to obtain the “average” distance between the features of obstacle avoidance trajectories generated by the two schemes and expert-demonstrated, as shown in Fig. 17.

The obstacle avoidance trajectory generated by the two trajectory planning schemes and the expert-demonstrated trajectory had certain trajectory feature differences. This is because neither scheme considered the different driving styles of drivers in the HighD dataset. Compared to Scheme 1, Scheme 2 can significantly improve the similarity between the generated and expert-demonstrated trajectories. Specifically, the average feature difference in the lateral position of the obstacle avoidance trajectory generated in Scheme 2 was only 49.04% of that generated in Scheme 1. The average characteristic difference of the lateral velocity in Scheme 2 was only 42.91% of that in Scheme 1. The average characteristic difference of the lateral acceleration in Scheme 2 was only 55.35% of that of Scheme 1.

To further verify the effectiveness of the proposed scheme, three groups of real obstacle avoidance scenarios were randomly selected

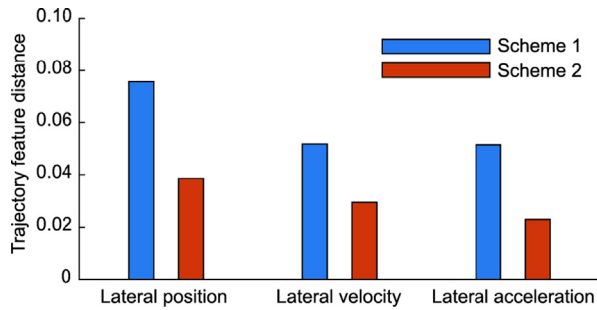


Fig. 17. Comparison of obstacle avoidance trajectory feature distance under two control schemes.

from the 42 groups of test samples. The proposed trajectory planning strategy was used to generate an AOAT, as shown in Fig. 18. The EDOATs in Scenarios 1, 2, and 3 are denoted as EDOAT<sub>1</sub>, EDOAT<sub>2</sub>, and EDOAT<sub>3</sub>, respectively. The AOATs generated based on Scenarios 1, 2, and 3 are denoted as AOAT<sub>1</sub>, AOAT<sub>2</sub>, and AOAT<sub>3</sub>, respectively. The lateral velocity and acceleration of the obstacle avoidance trajectory generated based on real driving scenarios are denoted as LV and LA, respectively. The real lateral velocity and acceleration in this scenario are denoted as LV<sub>T</sub> and LA<sub>T</sub>.

Figs. 18(a)–(f) show the lateral displacement, velocity, and acceleration of the trajectory generated based on Scenarios 1, 2, and 3, respectively. The  $\Delta v_x$  values of the three scenarios are 14.40, 21.30, and 41.76 km·h<sup>-1</sup>, respectively. According to the driving scenario information, Scenario 3 was a more urgent condition than Scenarios 1 and 2. To prevent collision accidents, the vehicle chooses an obstacle avoidance trajectory with high lateral acceleration in Scenario 3.

To verify the real-time performance and acceptability of the strategy, a hardware-in-the-loop (HIL) test platform was built based on the aforementioned obstacle avoidance scenarios, as shown in Fig. 19.

The HIL platform includes a ground resistance moment simulation system, automotive steering assembly, dSPACE MicroLabBox, and national instruments' peripheral component interconnect

extensions for instrumentation national instruments' peripheral component interconnect extensions for instrumentation (NI PXI) real-time system. The HIL experimental process is shown in Fig. 20. An AOAT planning strategy for driving scenario adaptation was run in MicroLabBox to ensure real-time performance. Vehicle and road models in CarSim were embedded in NI PXI real-time systems using NI LabVIEW software. The servomotor simulated the ground resistance moment. These components communicate via a controller area network (CAN) bus. The working process is as follows. After MicroLabBox receives the driving scenario information from NI PXI through the CAN bus, it plans the optimal AOAT  $r^*$  and calculates the optimal control angle  $\delta_f^*$ . The actuator receives control instructions through the CAN bus and causes the corresponding steering action  $\delta_{sw}$  to follow an optimal trajectory. The NI PXI software receives the actual control actions of the platform to complete the closed-loop control.

Figs. 21 and 22 show the trajectory planning and following process of the vehicles for 10 s from the start of obstacle avoidance.

In Scenario 1, the speed difference between the ego and lead vehicles at the beginning of obstacle avoidance was 14.4 km·h<sup>-1</sup>, and the ego vehicle speed was 121.9 km·h<sup>-1</sup>. In Scenario 2, the speed difference between the ego and lead vehicles at the beginning of obstacle avoidance was 21.3 km·h<sup>-1</sup>, and the ego vehicle speed was 118.3 km·h<sup>-1</sup>. In Scenario 3, the speed difference between the ego and lead vehicles at the beginning of obstacle avoidance was 41.76 km·h<sup>-1</sup>, and the ego vehicle speed was 94.50 km·h<sup>-1</sup>. Compared with Scenarios 1 and 2, the speed of the vehicle in Scenario 3 was slightly lower, but the speed difference between the ego and lead vehicles was large. Therefore, the obstacle avoidance urgency in Scenario 3 was relatively high and the obstacle avoidance time was relatively short.

As shown in Fig. 22(a), the maximum lateral position errors in the three scenarios are 0.185, 0.130, and 0.090 m, respectively. By contrast, the tracking accuracy of Scenario 2 was slightly worse than that of Scenario 1, whereas that of Scenario 3 was the worst. The reasons for this phenomenon are as follows. Compared with Scenario 1, the trajectory in Scenario 3 is more urgent, and the system may sacrifice trajectory tracking accuracy appropriately to

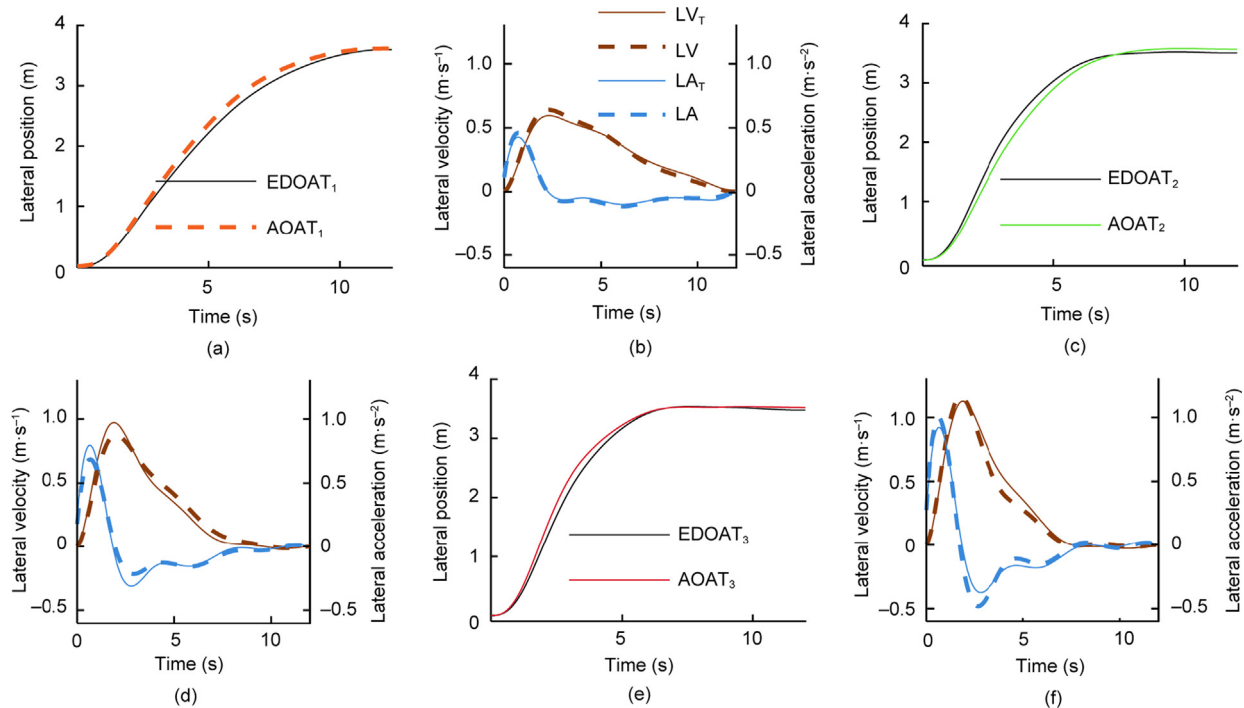


Fig. 18. Simulation results for Scenarios 1, 2, and 3.

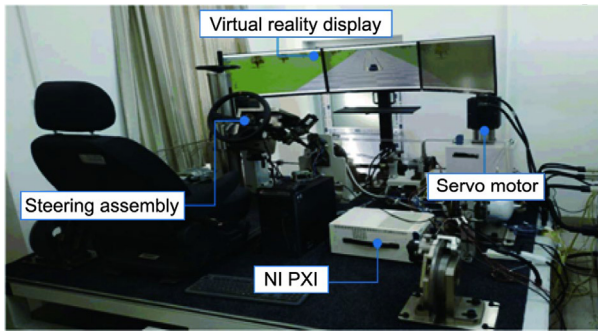


Fig. 19. Intelligent vehicle HIL test platform. NI PXI: national instruments' peripheral component interconnect extensions for instrumentation.

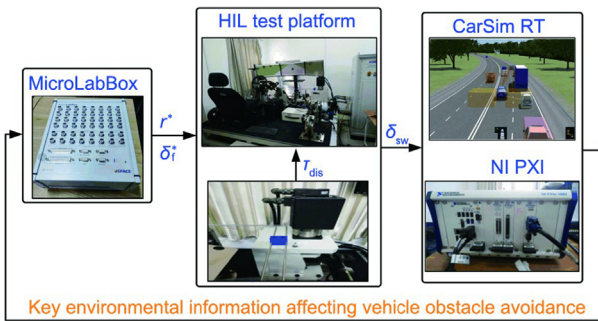


Fig. 20. AOAT planning and following process for driving scenario adaptation.  $\delta_i^*$ : the optimal control angle;  $\tau_{dis}$ : the steering resistance moment to the front wheels;  $\delta_{sw}$ : the actual steering wheel angle of the platform.

ensure vehicle safety. Fig. 22(b) shows the motor control current required to track the three obstacle avoidance trajectories. Through the HIL experiment, it can be observed that the AOAT planned in this study follows well. The data in the HighD dataset are all real traffic scenarios. The simulation and experimental results show that the proposed trajectory planning strategy can automatically adjust its weight coefficients according to the degree of urgency and safety in different driving scenarios. Thus, an AOAT suitable for the current driving scenarios was planned.

### 5. Summary

Trajectory optimization functions are difficult to design and exhibit poor adaptability to various scenarios. Moreover, a single trajectory optimization function cannot generate an obstacle avoidance trajectory that is consistent with driver behavioral habits. Thus, based on maximum entropy inverse reinforcement learning theory, automatic acquisition of the optimization functions of real obstacle avoidance trajectories was realized, and the acceptability of the trajectory optimization function was improved. Then, combined with the key driving scenario information affecting vehicle obstacle avoidance behavior, an AOAT planning strategy for adaptive driving scenarios was proposed. This strategy significantly enhances the adaptability to driving scenarios and improves driving safety, while avoiding obstacles.

In future work, more complex and reliable mapping models will be designed to describe the relationship between driving scenarios and the feature weights of the optimization functions, and the proposed algorithm will be fully validated by real car tests. For inverse

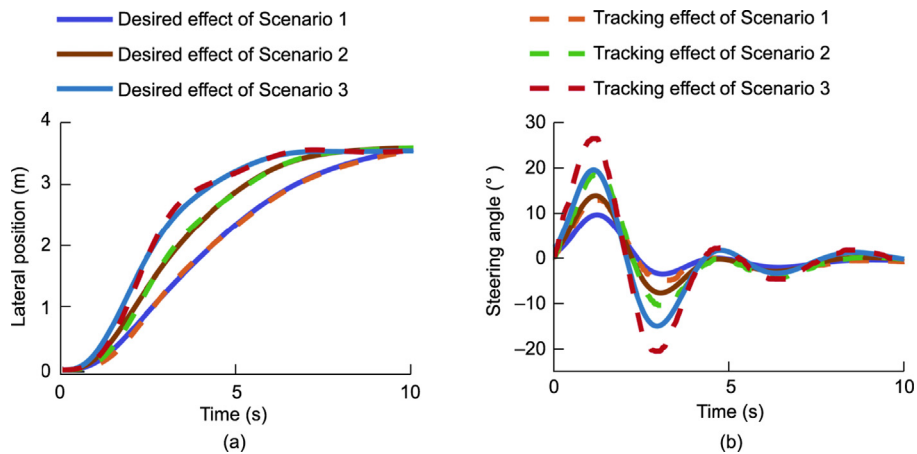


Fig. 21. Tracing effects of Scenarios 1, 2, and 3.

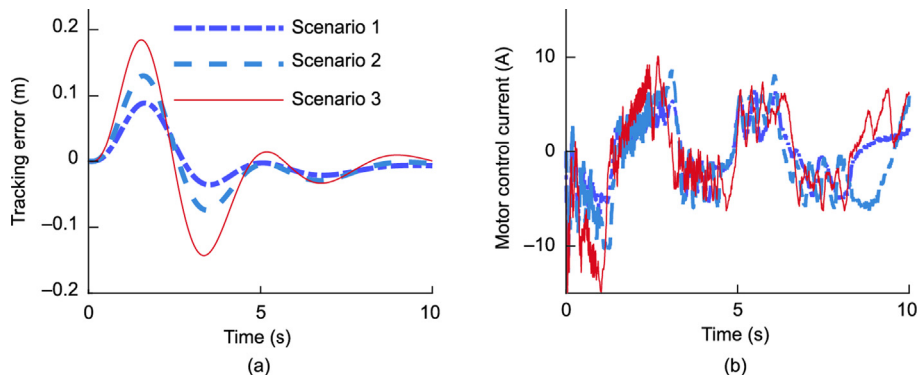


Fig. 22. Tracing effects of Scenarios 1, 2, and 3.



reinforcement learning, obstacle avoidance trajectories of humans with different driving styles were studied in depth. In addition, the vehicle trajectory planning in this study was limited to rectilinear road conditions. Obstacle avoidance trajectory planning for commercial vehicles under typical nonrectilinear road conditions will be an important research direction in the future.

## Acknowledgments

The work was supported by the National Natural Science Foundation of China (51875302).

## Compliance with ethics guidelines

Jian Wu, Yang Yan, Yulong Liu, and Yahui Liu declare that they have no conflict of interest or financial conflicts to disclose.

## References

- [1] Gao C, Wang G, Shi W, Wang Z, Chen Y. Autonomous driving security: state of the art and challenges. *IEEE Internet Things J* 2022;9(10):7572–9.
- [2] Benloucif A, Nguyen AT, Sentouh C, Popieul JC. Cooperative trajectory planning for haptic shared control between driver and automation in highway driving. *IEEE Trans Ind Electron* 2019;66(12):9846–57.
- [3] Dolgov D, Thrun S, Montemerlo M, Diebel J. Practical search techniques in path planning for autonomous driving. *Ann Arbor* 2009;1001(48105):18–80.
- [4] Islam F, Narayanan V, Likhachev M. Dynamic multi-heuristic A. In: Proceedings of IEEE International Conference on Robotics and Automation (ICRA); 2015 May 26–30; Seattle, WA, USA. IEEE; 2015. p. 2376–82.
- [5] Kushleyev A, Likhachev M. Time-bounded lattice for efficient planning in dynamic environments. In: Proceedings of IEEE International Conference on Robotics and Automation (ICRA); 2009 May 12–17; Kobe, Japan. IEEE; 2009. p. 1662–8.
- [6] Arslan O, Berntorp K, Tsiotras P. Sampling-based algorithms for optimal motion planning using closed-loop prediction. In: Proceedings of IEEE International Conference on Robotics and Automation (ICRA); 2017 May 29–Jun 3; Singapore. IEEE; 2017. p. 4991–6.
- [7] LaValle SM, Kuffner JJ. Randomized kinodynamic planning. In: Proceedings of IEEE International Conference on Robotics and Automation (ICRA); 1999 May 10–15; Detroit, MI, USA. IEEE; 1999. p. 473–9.
- [8] Zucker M, Kuffner J, Branicky M. Multipartite RRTs for rapid replanning in dynamic environments. In: Proceedings of IEEE International Conference on Robotics and Automation (ICRA); 2007 Apr 10–14; Rome, Italy. IEEE; 2007. p. 1603–9.
- [9] Berglund T, Brodnik A, Jonsson H, Staffanson M, Soderkvist I. Planning smooth and obstacle-avoiding B-spline paths for autonomous mining vehicles. *IEEE Trans Autom Sci Eng* 2009;7(1):167–72.
- [10] Wu J, Zhang J, Nie B, Liu Y, He X. Adaptive control of PMSM servo system for steering-by-wire system with disturbances observation. *IEEE Trans Transp Electrification* 2022;8(2):2015–28.
- [11] Rastelli JP, Lattarulo R, Nashashibi F. Dynamic trajectory generation using continuous-curvature algorithms for door to door assistance vehicles. In: Proceedings of IEEE Intelligent Vehicles Symposium Proceedings (IV); 2014 Jun 8–11; Dearborn, MI, USA. IEEE; 2014. p. 510–5.
- [12] Gu T, Dolan JM. On-road motion planning for autonomous vehicles. In: Proceedings of International Conference on Intelligent Robotics and Applications (ICIRA); 2012 Oct 3–5; Montreal, QC, Canada; 2012. p. 588–97.
- [13] Lattarulo R, González L, Perez J. Real-time trajectory planning method based on  $n$ -order curve optimization. In: Proceedings of International Conference on System Theory, Control and Computing (ICSTCC); 2020 Oct 8–10; Sinaia, Romania. IEEE; 2020. p. 751–6.
- [14] Lim W, Lee S, Sunwoo M, Jo K. Hybrid trajectory planning for autonomous driving in on-road dynamic scenarios. *IEEE Trans Intell Transp Syst* 2019;22(1):341–55.
- [15] Gutjahr B, Gröll L, Werling M. Lateral vehicle trajectory optimization using constrained linear time-varying MPC. *IEEE Trans Intell Transp Syst* 2016;18(6):1586–95.
- [16] McNaughton M, Urmson C, Dolan JM, Lee JW. Motion planning for autonomous driving with a conformal spatiotemporal lattice. In: Proceedings of IEEE International Conference on Robotics and Automation (ICRA); 2011 May 9–13; Shanghai, China. IEEE; 2011. p. 4889–95.
- [17] Dixit S, Montanaro U, Dianati M, Oxtoby D, Mizutani T, Mouzakitis A, et al. Trajectory planning for autonomous high-speed overtaking in structured environments using robust MPC. *IEEE Trans Intell Transp Syst* 2019;21(6):2310–23.
- [18] Luo Y, Xiang Y, Cao K, Li K. A dynamic automated lane change maneuver based on vehicle-to-vehicle communication. *Transport Res C Emer* 2016;62: 87–102.
- [19] Yang D, Zheng S, Wen C, Jin PJ, Ran B. A dynamic lane-changing trajectory planning model for automated vehicles. *Transport Res C Emer* 2018;95:228–47.
- [20] Chen J, Zhan W, Tomizuka M. Autonomous driving motion planning with constrained iterative LQR. *IEEE Trans Intell Vehicles* 2019;4(2):244–54.
- [21] Liu Y, Liu Y, Ji X, Sun L, Tomizuka M, He X. Learning from demonstration: situation-adaptive lane change trajectory planning for automated highway driving. In: Proceedings of IEEE International Conference on Mechatronics and Automation (ICMA); 2020 Oct 13–16; Beijing, China. IEEE; 2020. p. 376–82.
- [22] Ziegler J, Bender P, Dang T, Stiller C. Trajectory planning for Bertha—a local, continuous method. In: Proceedings of IEEE International Vehicles Symposium Proceedings (IVSP); 2014 Jun 8–11; Dearborn, MI, USA. IEEE; 2014. p. 450–7.
- [23] Sun L, Peng C, Zhan W, Tomizuka M. A fast integrated planning and control framework for autonomous driving via imitation learning. In: Dynamic Systems and Control Conference (DSCC); 2018 Sep 30–Oct 3; Atlanta, GA, USA; 2018.
- [24] Wang Y, Pan D, Liu Z, Feng R. Study on lane change trajectory planning considering of driver characteristics. *SAE Technical Paper* 2018;2018-01-1627.
- [25] Zhou B, Wang Y, Yu G, Wu X. A lane-change trajectory model from drivers' vision view. *Transport Res C* 2017;85:609–27.
- [26] He X, Xu D, Zhao H, Moze M, Aioun F, Guillemard F. A human-like trajectory planning method by learning from naturalistic driving data. In: Proceedings of IEEE Intelligent Vehicles Symposium (IV); 2018 Jun 26–30; Changshu, China. IEEE; 2018. p. 339–46.
- [27] Wu J, Kong Q, Yang K, Liu Y, Cao D, Li Z. Research on the steering torque control for intelligent vehicles co-driving with the penalty factor of human-machine intervention. *IEEE Trans Syst Man Cybern* 2023;53(1):59–70.
- [28] Nguyen AT, Rath J, Guerra TM, Palhares R, Zhang H. Robust set-invariance based fuzzy output tracking control for vehicle autonomous driving under uncertain lateral forces and steering constraints. *IEEE Trans Intell Transp Syst* 2020;22(9):5849–60.
- [29] Miyajima C, Nishiwaki Y, Ozawa K, Wakita T, Itou K, Takeda K, et al. Driver modeling based on driving behavior and its evaluation in driver identification. *Proc IEEE* 2007;95(2):427–37.
- [30] Xu L, Hu J, Jiang H, Meng W. Establishing style-oriented driver models by imitating human driving behaviors. *IEEE Trans Intell Transp Syst* 2015;16(5):2522–30.
- [31] Cai P, Sun Y, Chen Y, Liu M. Vision-based trajectory planning via imitation learning for autonomous vehicles. In: Proceedings of IEEE Intelligent Transportation Systems Conference (ITSC); 2019 Oct 27–30; Auckland, New Zealand. IEEE; 2019. p. 2736–42.
- [32] Li H, Wu C, Chu D, Lu L, Cheng K. Combined trajectory planning and tracking for autonomous vehicle considering driving styles. *IEEE Access* 2021;9:9453–63.
- [33] Zhang C, Chu D, Lyu N, Wu C. Trajectory planning and tracking for autonomous vehicle considering human driver personality. In: Proceedings of Conference on Vehicle Control and Intelligence (CVCI); 2019 Sep 21–22; Hefei, China; 2019. p. 1–6.
- [34] Wu P, Cao Y, He Y, Li D. Vision-based robot path planning with deep learning. In: Proceedings of International Conference on Computer Vision Systems (ICVS); 2017 Jul 10–13; Shenzhen, China; 2017.
- [35] Lenz D, Diehl F, Le M, Knoll A. Deep neural networks for Markovian interactive scene prediction in highway scenarios. In: Proceedings of IEEE Intelligent Vehicles Symposium (IV); 2017 Jun 11–14; Los Angeles, CA, USA. IEEE; 2017. p. 685–92.
- [36] Vallon C, Ercan Z, Carvalho A, Borrelli F. A machine learning approach for personalized autonomous lane change initiation and control. In: Proceedings of IEEE Intelligent Vehicles Symposium (IV); 2017 Jun 11–14; Los Angeles, CA, USA. IEEE; 2017. p. 1590–5.
- [37] Krajewski R, Bock J, Kloeker L, Eckstein L. The HighD dataset: a drone dataset of naturalistic vehicle trajectories on German highways for validation of highly automated driving systems. In: Proceedings of International Conference on Intelligent Transportation Systems (ITSC); 2018 Nov 4–7; Maui, HI, USA. IEEE; 2018. p. 2118–25.
- [38] Kuderer M, Gulati S, Burgard W. Learning driving styles for autonomous vehicles from demonstration. In: Proceedings of IEEE International Conference on Robotics and Automation (ICRA); 2015 May 26–30; Seattle, WA, USA. IEEE; 2015. p. 2641–6.