International Journal for Housing Science and Its Applications
Publish August 6, 2025. Volume 46, Issue 3 Pages 7566-7576

https://doi.org/10.70517/ijhsa463649

A multi-vehicle steering trajectory tracking method for tunnel scenarios based on improved residual networks

Zhixian Li[1,*], Nianfeng Shi[1,2], Guoqiang Wang[1,2] and Liguo Zhao[1]
[1] School of Computer and Information Engineering, Luoyang Institute of Science and Technology, Luoyang, Henan, 471023, China
[2] Henan Key Laboratory of Green Building Materials Manufacturing and Intelligent Equipment, Luoyang Institute of Science and Technology, Luoyang, Henan, 471023, China
Corresponding authors, (e-mail, lzx@lit.edu.cn).

Abstract Accurate tracking of vehicle steering trajectories is crucial to the safety of traveling in tunnels. In this paper, a multi-vehicle steering trajectory tracking method based on improved residual network is proposed for tunnel scenarios, which combines the attention mechanism and model predictive control technology to realize accurate tracking. Aiming at the problem that the traditional twin network tracking algorithm is not satisfactory enough in tunnel scenarios, the ECA channel attention mechanism is introduced to improve the structure of the residual network and enhance the feature extraction capability; the feature fusion module is designed to effectively integrate different levels of feature information; and the model predictive controller based on the spatial deviation model is constructed to realize accurate tracking. The experimental results show that in the simple occlusion scenario, the algorithm in this paper improves the tracking accuracy MOTA by 3.6% to 83.42% compared with SiamCAR algorithm, and the tracking precision MOTP improves by 3% to 88.19%, and the number of identity switching is reduced to 5 times; in the complex traffic scenario, the tracking accuracy improves by 2.4% to 78.77%, and the tracking precision improves by 4.2% to 85.69%. The active steering experiment based on data recharge verifies the effectiveness of the control method, and the system is able to adjust the trajectory deviation to ensure the smooth driving of the vehicle. The method can realize accurate tracking of multi-vehicle steering trajectories in tunnel scenarios and improve driving safety.

Index Terms Vehicle trajectory tracking, Attention mechanism, Residual network, Model predictive control, Spatial deviation model, Twin network

I. Introduction

With the improvement of people's daily life standard demand, automobile industry technology is also developing rapidly [1]. Among them, automatic driving technology is one of the main core technologies researched in recent years, and the ability of vehicles to travel under various extreme working conditions has been widely concerned [2]-[4].

In recent years, self-driving vehicles, as a new subject area of technology and cutting-edge applications, can reduce the occurrence of traffic accidents and greatly improve the efficiency of transportation travel [5], [6]. Autonomous driving technology is the development of vehicles as an intelligent mobile transportation system to the future and the development of road intelligence of the general trend [7], [8]. And the realization of vehicle driving intelligent system is to realize the most effective method to improve the driving safety of driving road, improve the efficiency of passenger transport travel is crucial way, but to fully realize the commercialization of self-driving cars, there are many other practical technical difficulties need to be solved [9]-[12]. In fact, there are various types of steering interferences in the driving process of self-driving vehicles, and these steering interferences will have a certain impact on the steering system of the vehicle, which in turn will affect the trajectory tracking accuracy of the vehicle [13]-[15]. Especially in the tunnel scenario, due to the weak GPS signal in the tunnel, and the existence of other obstacles and insufficient light lead to the vehicle in the process of driving by the steering impact interference, the vehicle's steering system is affected, which will affect the vehicle's ability to drive stably to a greater extent, so the study of tunnel scenarios of multi-vehicle steering trajectory tracking is of great significance [16]-[19].

Vehicle trajectory tracking is a key technology in the intelligent transportation system, which is of great significance in ensuring road traffic safety. Especially in special scenes such as tunnels, traditional tracking algorithms are often difficult to cope with complex environmental changes due to factors such as insufficient light, dense vehicles, and restricted view angles. In recent years, deep learning technology has made significant progress in the field of computer vision, providing new ideas for solving the vehicle trajectory tracking problem. Among them, the target tracking algorithm based on deep residual network shows good performance, which is able to effectively extract image features and realize the precise positioning of the target. However, in practical applications, the traditional

residual network still faces the problems of insufficient feature extraction, easy to lose targets, and frequent identity switching in tunnel scenarios. This is mainly due to the drastic changes in lighting conditions in tunnel environments, the frequent occurrence of vehicle occlusion, and the high similarity of vehicle appearance features, which increases the tracking difficulty. Meanwhile, the steering trajectory tracking of vehicles in tunnels requires higher accuracy to ensure driving safety. Although existing twin network tracking algorithms perform well in general scenarios, for deep networks with a large number of channels, not all feature channels are valuable for the tracking task, which leads to inefficient feature extraction. In addition, traditional tracking control methods are often difficult to balance tracking accuracy and control smoothness when dealing with vehicle steering trajectories, which affects the driving experience and safety. Therefore, how to design a high-precision multi-vehicle steering trajectory tracking method for tunnel scenarios has become a key issue in current research.

In order to solve the above problems, this study proposes a multi-vehicle steering trajectory tracking method based on improved residual network for tunnel scenarios. Firstly, the traditional residual network structure is improved by introducing the ECA channel attention mechanism, which enhances the network's ability to extract key features and enables the model to adaptively adjust the weights of different feature channels. Second, the feature fusion module is designed to effectively integrate the feature information at different levels by adopting a two-branch structure, so as to improve the model's adaptive ability to target changes. Then, a spatial-based deviation model is established to convert the vehicle motion to be analyzed under the Freightliner coordinate system, which reduces the dimensionality of the vehicle model and better understands the motion relationship of different objects on the road. Finally, a model prediction controller is designed based on this model to achieve accurate tracking of the vehicle steering trajectory by optimizing the objective function. In order to verify the effectiveness of the method, the appearance feature extraction network is trained on the vehicle re-identification dataset, and comparative experiments are carried out in different scenarios, and the reliability of the control method is further verified by the double-shifted line data backfeeding experiment.

II. Multi-vehicle steering trajectory tracking methods

II. A. Deep residual network model

Resnet consists of multiple convolutional layers, batch normalization layers and pooling layers.Resnet50 has 50 convolutional layers with an initial layer of ordinary convolutional structure which is responsible for initial feature extraction of the input image. Layer1 contains three Residual Blocks, each of which internally consists of multiple convolutional layers and jump connections used to capture low-level patterns of input features. layer2 contains 4 Residual Blocks, which gradually extract higher level features by deepening layer by layer. layer3 contains 6 Residual Blocks that continue to deepen the network and capture more abstract and complex features. layer4 contains 3 Residual Blocks to further increase the depth of the network, allowing the network to learn more complex image features. With the addition of a 5-category output layer after the last fully connected layer, the network ends up outputting 5-category predictions.

II. B. Attention mechanisms

II. B. 1) SE Attention Mechanism

SENet learns the importance of each channel, enabling the network to better understand the relationship between different channels and adaptively adjust the weights of each channel. Specifically, the Squeeze operation reduces redundant information between channels, while the Exciation operation reinforces useful information, thus improving the performance and generalization of the model. In this way, the neural network is able to focus its attention on specific channels and ignore those that do not play much of a role in the classification effort. It is just like a human being will only notice his own target in his own field of vision and will not pay attention to other targets.

In SE attention mechanism, the input feature maps are first pooled in pooling operation and then fed into the fully connected layer (FC), which is then fed into the next FC layer learning using the ReLU function, and then outputted through the Sigmoid function denoted as S. The formula is denoted as,

$$z = \frac{1}{W \times H} \sum_{i=1}^{W} \sum_{j=1}^{H} u_c(i, j) \tag{1}$$

$$S = \sigma(g(z, W)) = \sigma(W_2 \delta(W_1 z)) \tag{2}$$

where (2) represents the activation function ReLU, and σ is the activation function Sigmoid. $W_1 \in R^{\frac{c}{r} \times C}$, $W_2 \in R^{\frac{c}{r} \times C}$ denote the weight matrices of the two fully connected layers, respectively. weight matrices.R denotes the number

of hidden nodes in the intermediate layer. z is the output in Eq. (1), (W,H) is the dimension of the image, and c is the number of image channels.

II. B. 2) CBAM Attention Mechanism

CBAM consists of two parts, except for the addition of the maximum pooling parallel operation in the SENet module. The channel attention module generates weight coefficients by fusing the feature maps after average and maximum pooling, and then multiplies the weight parameters by the input features to enhance the representation of useful features.

If the initial feature input is $F \in R^{H\times W\times C}$ and the number of network channels is C, the output features F_{avg}^{c} and F_{max}^{c} are obtained through the maximum pooling layer and the global average pooling layer, and on this basis, the dimensionality reduction, ReLU activation and dimensionality enhancement are carried out on the feature F_{avg}^{c}, F_{max}^{c}. Two new feature maps are obtained. The two graphs are added and activated with the Sigmoid function, as shown in Eq. (3).

$$\begin{aligned} M_c(F) &= \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \\ &= \sigma(W_1(W_0(F_{avg}^{c})) + W_1(W_0(F_{\max}^{c}))) \end{aligned} \tag{3}$$

where $w_0 \in R^{\frac{c}{r}\times C}$, represents the process of dimensionality reduction; $w_1 \in R^{C\times\frac{c}{r}}$, represents the process of dimensionality enhancement, and r stands for the multiplicity of the dimensionality reduction or enhancement; σ represents the Sigmoid activation operation. After obtaining the weight information $M_c(F)$, it is multiplied by the input feature F to obtain the output feature $F1$ of the channel dimension.

After being processed by the spatial attention module, $F1$ undergoes two pooling operations and merges to 2, and then undergoes a convolution operation to compute the weighting coefficients using the Sigmoid function, which is used to generate the output of the spatial attention module; finally, the weighting coefficients $M_c(F1)$ are multiplied by the features $F1$ supplied to the module to obtain the new feature map $F2$. The specific process is expressed in Equation (4).

$$\begin{aligned} M_c(F1) &= \sigma(f^{7\times7}([AvgPool(F1); MaxPool(F1)])) \\ &= \sigma(f^{7\times7}([F_{avg}^{s}; F_{\max}^{s}])) \end{aligned} \tag{4}$$

Here, $f^{7\times7}$ denotes a two-dimensional convolution operation with a convolution kernel size of 7 × 7, and σ denotes the Sigmoid activation function.

II. B. 3) ECA Attention Mechanism

ECANet has been shown that channel feature dimensionality reduction can affect how the attention mechanism learns [20]. Therefore, ECANet avoids multi-channel dimensionality reduction by using one-dimensional convolution in multi-channel dimensionality interactions, which helps to reduce complexity.

Assuming the input features $F \in R^{H\times W\times C}$, the input features will be compressed through the pooling layer to get $F_{avg} \in R^{1\times1\times C}$; secondly, one-dimensional convolution is applied to the interactions between each channel and k some of the neighboring channels. Here k represents the coverage of the cross channels and the size of the 1D convolution core, the value of k can be customized, if $k=3$, then we can obtain $F_{conv} \in R^{1\times1\times C}$; secondly, the Sigmoid function is used to F_{conv} for activation operation to obtain the weight coefficient $M_c(F)$, denoted as $M_c(F) \in R^{1\times1\times C}$, with the formula (5),

$$M_c(F) = \sigma(f^{k}(AvgPool(F))) = \sigma(f^{k}(F_{avg}^{c})) \tag{5}$$

Here, f^{k} denotes the convolution operation with convolution kernel size k, F_{avg}^{c} denotes the average pooling operation, and σ denotes the Sigmoid activation function.

Equation (6) describes the feature extraction process of a one-dimensional convolutional neural network. The weight coefficients $M_c(F)$ are multiplied with the input features F element by element to obtain the new output features $F1$.

$$F1 = M_c(F) \times F \tag{6}$$

II. C. Tracking algorithm based on attentional residual networks

II. C. 1) General framework of the algorithm

The current twin network based tracking algorithm uses the picture pair matching idea, if two pictures have their features close together, then the Siamese network thinks that the gap between them is very small, and the two pictures are very close together.The Siamese network extracts the features of the tracked target from the given labeled box, and the feature extraction network it adopts is generally a deeper residual network, and as the deepening of the feature extraction network, the network's The number of channels also increases, but not all feature channels play a positive role in the tracking task, and some feature channels are worthless for tracking, so the tracking model needs to learn different channels biasedly, and the channel attention mechanism can help the twin network learn the template features better and enhance the discriminative ability of the model.

In order to adapt the algorithm to the tracking scenarios under different challenges, this paper is based on the twin network structure, firstly, the channel attention mechanism is introduced in the feature extraction module to improve the quality of the extracted features, secondly, the attention fusion is performed on the multi-layer features, which fuses the spatial information of the lower layers with the semantic information of the higher layers, making the model adaptable to the changes of the target.In this section, based on the SiamCAR algorithm, the proposed tracking algorithm with channel attention feature fusion mechanism.

II. C. 2) Attention residual networks

Since visual single-target tracking algorithms all learn the target to be tracked in the first frame, the quality of the features in the first frame is crucial for later tracking, inspired by the attention mechanism of the ECA channel, in this paper, we introduce this attention mechanism into the residual unit, which enables the Siamese network to assign weights to different feature channels and suppress the features that cause interference to the tracking through the weighting factor The structure of the attentional residual unit is shown in Fig. 1. The structure of the attention residual unit is shown in Fig. 1.

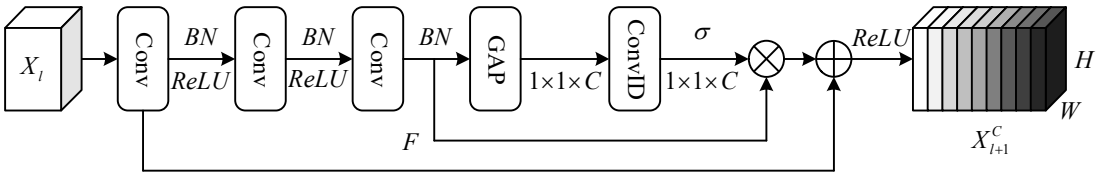


Figure 1, Attention residue unit

II. C. 3) Feature Fusion Module

The traditional feature fusion method performs addition operation for the features of different layers, this fusion scheme does not increase the model parameters, but the features between different layers are more different, and the addition operation may lead to the loss of some of the features, inspired by the literature, the fusion scheme based on the attention mechanism [21] is used to fuse the features of different layers, in which the attention fusion module is shown in Fig. 2, for the residual network The outputs of the last three stages $AR_j \in \mathbb{R}^{H \times W \times C_j}, j = 3,4,5$ are selected for attention fusion. In this case, the fusion module is divided into an upper and lower double-branch structure, where the upper branch is used to extract the global channel attentional weights W_g and the next branch is used to extract the local channel attentional weights W_t. First the $AR_j \in \mathbb{R}^{H} \times W \times C_j AR_{j-1} \in \mathbb{R}^{H \times W \times C_{j-1}}$ element-by-element for the features output from the residual module are Summing (taking into account that the channels and sizes of features at different layers are inconsistent, the most primitive outputs of the last three stages of the residual attention network will be converted to a tensor of the same size and number of channels by convolution operation before that) yields the initial integration feature $A \in \mathbb{R}^{H \times W \times C}$, which then goes through the upper and lower double branches respectively to get the localized attention mechanism feature $A_l \in \mathbb{R}^{H \times W \times C}$ and the feature $A_g \in \mathbb{R}^{1 \times 1 \times C}$ of the global attention mechanism, and the two are subjected to element-by-element addition.

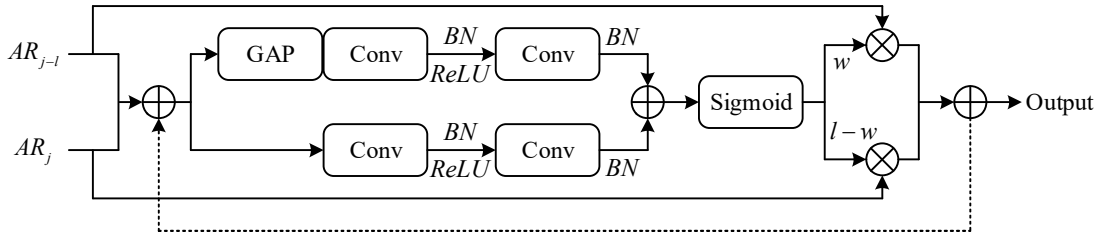


Figure 2, Feature Fusion Module based on attention

The fusion process can be expressed as Eq,

$$A = AR_{j-1} \oplus AR_j \tag{7}$$

$$W_g = Conv2D(ReLU(\beta(Conv2D(A)))) \tag{8}$$

$$W_l = GAP(Conv2D(ReLU(\beta(Conv2D(A))))) \tag{9}$$

$$W = \sigma(W_g + W_l) \tag{10}$$

$$Output = AR_{j-1} \otimes W + AR_j \otimes (1 - W) \tag{11}$$

where $\otimes$ is the element-by-element multiplication operation, $\oplus$ is the element-by-element addition operation, σ is the Sigmoid activation function, β denotes the batch regularization process, GAP is the Global Average Pooling Layer, and $Cov2D$ is the two-dimensional convolution operation.

II. D.Experimental results and analysis

II. D. 1) Experimental data set

In this section, the Veri-wild dataset is used as a vehicle re-identification dataset for training the appearance feature extraction network of vehicles.

The Veri-wild dataset is a classic dataset in the field of vehicle re-identification, which is widely used in the field of intelligent transportation and traffic safety. The dataset consists of 24 folders, which contain 416314 image data and 40671 vehicle types.

In the Veri-wild dataset, the dataset is too large and the number of images under each folder (vehicle ID) is uneven, in order to make the vehicle target features extracted by the network more complete, a program is written to screen the dataset. The first 16 compression terms of the dataset are selected to screen each vehicle ID with the number of vehicles between 40-60, and then 4 images are randomly selected from under each vehicle ID as the test set. The appearance images of 570 vehicles are obtained after screening, constituting the training set train_570 and the test set test_570.

II. D. 2) Evaluation indicators

In this paper, the MOTChallenge evaluation criteria are selected to evaluate the R-DeepSORT tracking performance, and the specific performance evaluation indexes are as follows.

(1) MOTA

MOTA is a measure of the tracking accuracy of the tracking algorithm in the continuous tracking task, which indicates the proportion occupied by wrong matches, the higher MOTA value indicates the less misdetection, omission and IDSwitch phenomenon, and the better tracking performance, and the calculation formula is shown in (12).

$$MOTA = 1 - \frac{\sum_t (IDSW(t) + f_p(t) + f_n(t))}{\sum_t GT(t)} \tag{12}$$

where, $GT(t)$ represents the number of detection frames that really exist in the t th frame image; $IDSW(t)$ indicates the number of times an identity switching occurs in the t th frame image target; $f_p(t)$ represents the

number of times the image is misdetected in the t th frame image; and $f_n(t)$ the number of missed detections in the t th frame image.

(2) MOTP

MOTP is a measure of the accuracy of the detector of the tracking algorithm in the continuous tracking task, expressed in the position error between the tracking frame and the real frame, more concerned about the performance of the detector, the calculation formula is shown in (13).

$$MOTP = \frac{\sum_{t,i} d_t^i}{\sum_t c_t} \tag{13}$$

where, c_t denotes the number of successful matches between the tracking frame and the real frame in the t-th frame image; d_t^i denotes the intersection and concatenation ratio of the detection frame between the i-th tracking frame and the corresponding real frame in the t-th frame image.

(3) IDSwitch

IDSwitch represents the number of times an ID switch occurs in the target to be tracked, and the smaller the value indicates the better performance of the tracker when occlusion occurs.

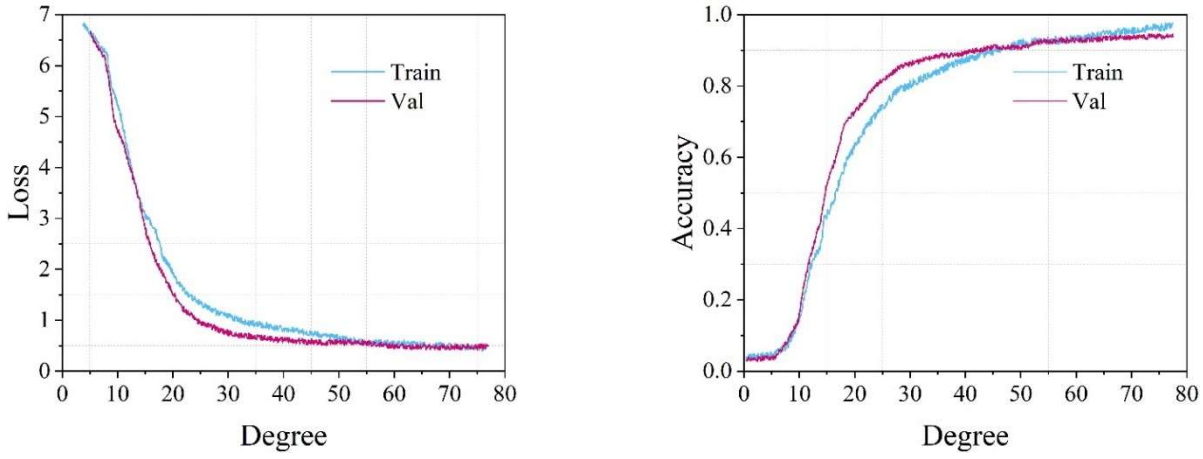
(4) Detection speed

The detection speed is evaluated using the frame rate FPS, which represents the tracking rate of the tracker.

II. D. 3) Experimental results and analysis

(1) Vehicle re-recognition experiment

This experiment focuses on training the appearance feature extraction network on the constructed vehicle appearance feature dataset, and the optimized appearance feature extraction network is trained 80 times by using a fixed-step learning rate decay strategy, and the loss function and accuracy curve of the model training process are shown in Figure 3. It can be seen that when the model is trained to the 40th epoch, the vehicle feature appearance extraction network of the twin network tracking model based on the attention mechanism tends to converge, no underfitting or overfitting phenomenon occurs, and the accuracy rate on the validation set is 98.8%, which indicates that the vehicle appearance feature network of the twin network tracking model based on the attention mechanism has a higher ability to extract vehicle features.



(a) Loss function curve (b) Accuracy curve

Figure 3, Training loss curve and accuracy curve

(2) Road vehicle tracking experiment

This section of the experiment is to prove the effectiveness of the improved algorithm proposed in this paper, in the experimental process SiamCAR algorithm for comparison, the experimental process will be for the occlusion scene, complex traffic scene to verify the model tracking performance.

Scene 1, Simple Occlusion Scene

In order to verify the tracking performance of the vehicle tracking algorithm in the occlusion scene, the training dataset MVI_40991 file in the UA-DETRAC public dataset is selected for the tracking test, and the duration of the surveillance video is 93s, with an image resolution of 960×540 and a frame rate of 26PS.

The evaluation results of the algorithm in the occlusion scene are shown in Table 1. As can be seen from the table, the algorithm proposed in this paper compared to the SiamCAR algorithm tracking accuracy increased by 3.6%, tracking precision increased by 3%, the number of identity switching reduced by 7 times, proving the effectiveness of the twin network tracking model based on the attention mechanism proposed in this paper, but due to the deepening of the network structure brought about by the increase in the amount of computation of the model, the detection rate is reduced, but still meet the real-time requirements of vehicle tracking.

Table 1, The algorithm evaluation results under the occlusion scenario

Tracking algorithm	MOTA/%	MOTP/%	IDS	FPS
SiamCAR	79.85	85.17	14	36
This article tracking algorithm	83.42	88.19	5	29

Scene 2, Complex Traffic Scene

In order to verify the tracking performance of the vehicle tracking algorithm in complex traffic scenes, traffic scenes with fuzzy scenes and high traffic flow are selected from the UA-DETRAC public dataset for testing. The MVI_20035 file of the training set is selected, the total duration of the surveillance video is 42s, the image resolution is 960×540, and the frame rate is 26FPS.

The algorithm evaluation results of different tracking algorithms are shown in Table 2. From the experimental results, it can be seen that the twin network tracking model proposed in this paper based on the attention mechanism tracking accuracy MOTA improves the accuracy by 2.4% precision, MOTP improves by 4.2%, and the number of identity switching is reduced by 13 times. However, the speed of the tracking algorithm is reduced, which is due to the fact that the tracker needs to match and track all the vehicles on the video screen, and the complexity of the algorithm is increased, which has a certain impact on the tracking performance.

Table 2, Algorithm evaluation results in complex traffic scenarios

Tracking algorithm	MOTA/%	MOTP/%	IDS	FPS
SiamCAR	76.41	81.54	30	31
This article tracking algorithm	78.77	85.69	1	24

III. Multi-vehicle steering trajectory tracking prediction controller

III. A. Space-based bias modeling

Converting the vehicle motion into a space-based deviation model requires decomposing the vehicle position information from the right-angle coordinate system to the Freightliner coordinate system. Given a smooth reference trajectory, according to the position of the vehicle, it is projected onto the reference trajectory, and the motion state of the current vehicle in the map coordinate system is decomposed based on the projection point to obtain the motion state along the reference trajectory. The advantage of doing motion analysis in the Freightliner coordinate system is to do motion decomposition with the help of the reference trajectory, which reduces the dimensionality of the vehicle model, ignores the influence of the shape of the road in the map coordinate system, and is more helpful to understand the motion relationship of different objects on the road. The correspondence of the vehicle position information in the Cartesian coordinate system and the Freightliner coordinate system is shown in Fig. 4. The variable e_φ is defined to denote the deviation of the vehicle heading angle φ from the heading angle φ_s that corresponds to the heading angle on the reference trajectory, and e_y denotes the vehicle's transverse deviation with respect to the reference trajectory. Then, the variable s (the projected displacement of the vehicle with respect to the reference trajectory) is introduced to establish e_φ and e_y from a function of time t into an equation of state with respect to space s.

The kinematic equation is derived from the geometric relationship in Fig. 4 as,

$$v_s = (\rho_s - e_y)\dot{\varphi}_s = v\cos e_\varphi \tag{14}$$

$$\dot{s} = \rho_s \dot{\varphi}_s = \frac{\rho_s}{\rho_s - e_y} v\cos e_\varphi \tag{15}$$

where, v_s is the projected vehicle speed along the reference trajectory; ρ_s is the radius of curvature of the reference trajectory; φ_s is the heading angle of the reference trajectory; and $\dot{s}$ is the speed of the vehicle traveling along the reference trajectory.

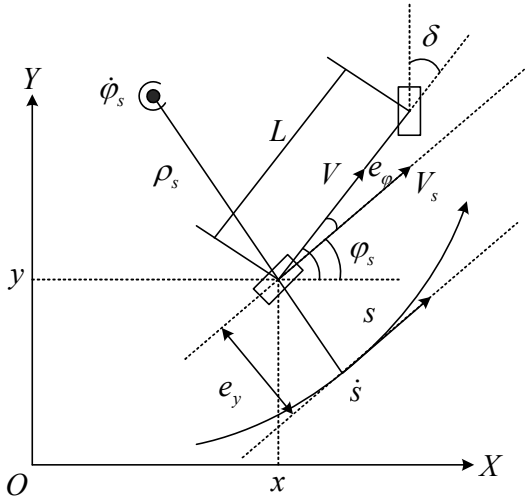


Figure 4, Vehicle model in the Freiner coordinate system

The derivatives of the lateral displacement deviation and heading angle deviation with respect to time are,

$$\dot{e}_y = v\sin(e_\varphi) \tag{16}$$

$$\dot{e}_\varphi = \dot{\varphi} - \dot{\varphi}_s \tag{17}$$

Assuming $\dot{s} \neq 0$, the derivative with respect to space can be replaced as a function of the derivative with respect to time, i.e., $\frac{d(\cdot)}{ds} = \frac{d(\cdot)}{dt}\frac{1}{s}$. The curvature of the vehicle's motion κ is expressed in terms of the vehicle's front wheel steering angle δ as $\kappa = \frac{\tan\delta}{L}$. Thus, the space-based deviation model expression is,

$$e_y' = \frac{e_y}{\dot{s}} = \frac{\rho_s - e_y}{\rho_s}\tan(e_\varphi) \tag{18}$$

$$e_\varphi' = \frac{e_\varphi}{\dot{s}} = \frac{\rho_s - e_y}{\rho_s\cos(e_\varphi)}\kappa - \frac{1}{\rho_s} \tag{19}$$

The space-based incomplete vehicle deviation model is defined as,

$$\xi' = f(\xi, u) \tag{20}$$

Where, the state quantity is $\xi = \begin{bmatrix} e_y & e_\varphi \end{bmatrix}^T$; the input quantity is $u = \kappa$.

III. B. Model Predictive Controller

III. B. 1) Linearization and discretization

Assume that any point (ξ_{ref}, u_{ref}) on the reference trajectory satisfies the above kinematic equations. Linearize at the point (ξ_{ref}, u_{ref}) using the 1st order Taylor approximation of the reference trajectory and discretize using Euler's formula. Setting the discrete walk as $\Delta s = \frac{\rho_s^k \cos e_\varphi^k}{\rho_s^k - e_y^k} vT$, the spatial deviation model for linear discretization is,

$$\xi(k+1|k) = A_{k,c}(k)\xi(k|k) + B_{k,c}(k)u(k|k) + d_{k,c} \tag{21}$$

$$A_{k,c}=\begin{bmatrix} 1-\frac{\tan e_{\varphi}^{ref}}{\rho_s}\Delta s & \frac{\rho_s-e_y^{ref}}{\rho_s}\cdot\frac{\Delta s}{\cos^2 e_{\varphi}^{ref}} \\ -\frac{\kappa}{\rho_s\cos e_{\varphi}^{ref}}\Delta s & 1+\frac{\rho_s-e_y^{ref}}{\rho_s}\cdot\frac{\sin e_{\varphi}^{ref}}{\cos^2 e_{\varphi}^{ref}}\cdot\kappa\cdot\Delta s \end{bmatrix}\in R^{2\times 2} \tag{22}$$

$$B_{k,c}=\begin{bmatrix} 0 \\ \frac{\rho_s-e_y^{ref}}{\rho_s}\cdot\frac{\Delta s}{cose_y^{ref}} \end{bmatrix}\in R^{2\times 1} \tag{23}$$

Let the prediction time domain of the system in the trajectory tracking layer be N. From Eq. (23), we can find the respective outputs of the system state quantities at each moment in the prediction time domain as,

$$e_y(k)=\psi(k)\xi(k\mid k)+\Pi(k)u(k)+O(k)D_{k,c} \tag{24}$$

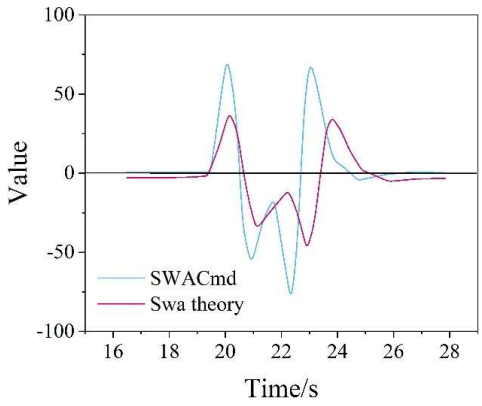
III. B. 2) Spatial model-based model predictive control algorithm

In order to realize safe, accurate and smooth trajectory tracking of multi-vehicles in tunnels, it is necessary to design a suitable objective function. According to the control object and control goal, the selected objective function should be able to optimize the vehicle lateral displacement deviation and steering curvature, so as to minimize the deviation of the vehicle in the driving process, and at the same time to avoid the wear and tear of the steering system and tires caused by the high steering frequency. Therefore, the objective function of the selected SMPC algorithm consists of two parts, which ensure the smoothness of the vehicle during driving and the accuracy of trajectory tracking, respectively.

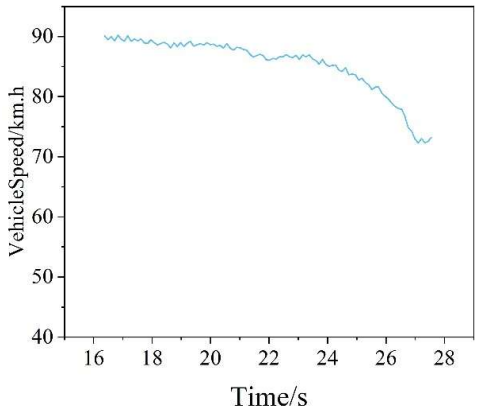
III. C. Double-shifted line data recharge experiments

This subsection takes the results of the 85km/h double shift line experiment as an example to demonstrate the control scheme and vehicle position information acquisition method described in this paper. Since the simulated vehicle model parameters have a certain error in the tire nonlinearity region and the experimental vehicle parameters, the steering wheel corner command is adjusted in time and frequency domains according to the situation during the experiment. Eventually, the amplitude of the cornering command of the real-vehicle experiment described in this subsection is 0.56 times of the theoretical value, and the frequency is 0.7 times of the theoretical value. Meanwhile, during the experiment, it was found that there was a difference of about -2.4deg between the angular zero position of the experimental vehicle's EPS and the zero position of the steering wheel, so that the steering wheel corner command value and the feedback value at the beginning and end of the experiment did not coincide with the y = 0deg.

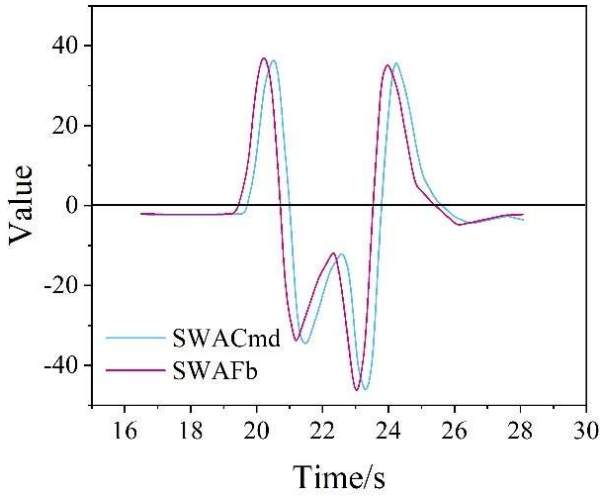
Figure 5 shows the variation of each kinematic parameter of the vehicle with time during the real vehicle experiment. From the simulation results, it can be seen that the active steering method based on data backfeeding adopted in this section can enable the vehicle to successfully complete the double-shift line experimental maneuver, which effectively verifies the reliability and effectiveness of the control method proposed in this paper.



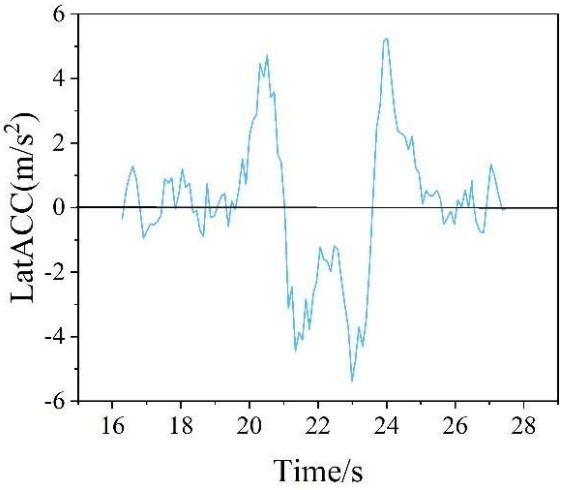
(a) Value contrast



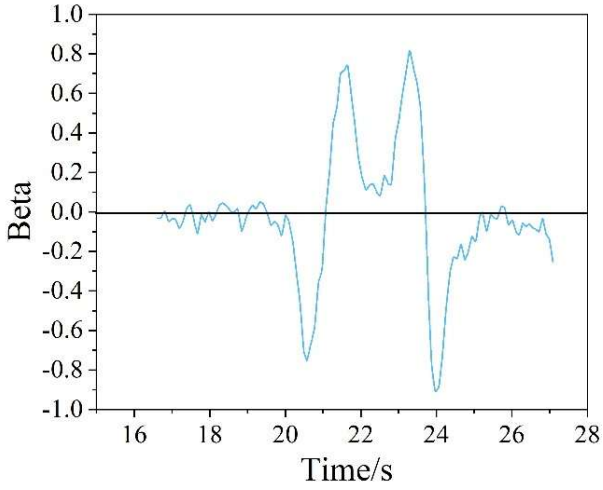
(b) Speed of a motor vehicle



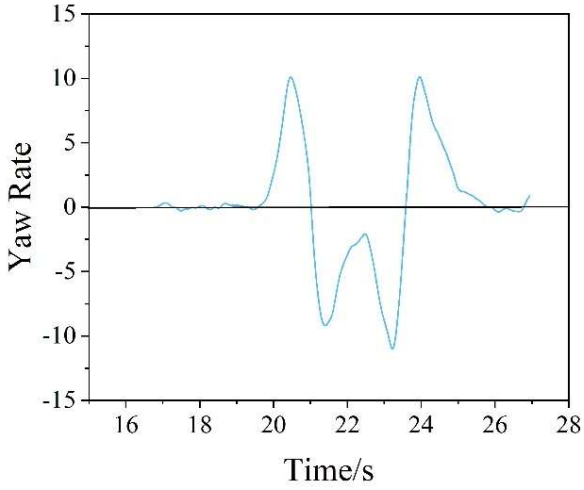
(c) Turn the corner command value and feedback value



(d) Lateral acceleration



(e) Centroid side bias Angle



(f) Angular velocity

Figure 5, Curve of vehicle parameters with time

IV. Conclusion

Through the in-depth study of the multi-vehicle steering trajectory tracking problem in tunnel scenarios, the vehicle tracking performance is improved by improving the residual network structure and introducing the attention mechanism. The experimental results demonstrate that the twin network tracking model based on the ECA channel attention mechanism performs well in feature extraction and achieves 98.8% accuracy on the validation set. Comparison experiments show that in simple occlusion scenarios, the tracking accuracy MOTA of this improved algorithm reaches 83.42%, which is 3.6% higher than that of the SiamCAR algorithm; the tracking accuracy MOTP reaches 88.19%, which is 3% higher; and the number of identity switching times is reduced from 14 times to 5 times. In the complex traffic scene test, the MOTA value of the improved algorithm reaches 78.77%, which is 2.4% higher than the SiamCAR algorithm; the MOTP value reaches 85.69%, which is 4.2% higher; and the number of identity switching times is reduced to 1, which is substantially better than the 30 times of the SiamCAR algorithm.

At the control level, the model predictive controller based on the spatial deviation model realizes the accurate tracking of the vehicle steering trajectory. The results of the double-shift line experiment show that the active steering method with data recharge enables the vehicle to successfully complete the test maneuver, which verifies the effectiveness of the control method. Although the computational complexity of the improved algorithm increases, resulting in a slight decrease in tracking speed, it still maintains a processing speed of 29 FPS in simple occlusion scenarios and reaches 24 FPS in complex traffic scenarios, which is able to meet the real-time tracking requirements.

The comprehensive results show that the improved residual network combined with the attention mechanism can effectively improve the accuracy and stability of multi-vehicle steering trajectory tracking in tunnel scenarios, which provides reliable technical support for the application of intelligent transportation systems in complex environments.

Funding

This work was sponsored in part by Science and Technology Research Project of Henan Province (232300420157).

References

[1] Klink, G., Mathur, M., Kidambi, R., & Sen, K. (2014). Contribution of the automobile industry to technology and value creation. Auto Tech Review, 3(7), 18-23.
[2] Yurtsever, E., Lambert, J., Carballo, A., & Takeda, K. (2020). A survey of autonomous driving, Common practices and emerging technologies. IEEE access, 8, 58443-58469.
[3] Barabas, I., Todoruţ, A., Cordoş, N., & Molea, A. (2017, October). Current challenges in autonomous driving. In IOP conference series, materials science and engineering (Vol. 252, No. 1, p. 012096). IOP Publishing.
[4] Brenner, W., & Herrmann, A. (2018). An overview of technology, benefits and impact of automated and autonomous driving on the automotive industry. Digital marketplaces unleashed, 427-442.
[5] Hansson, S. O., Belin, M. Å., & Lundgren, B. (2021). Self-driving vehicles—an ethical overview. Philosophy & Technology, 34(4), 1383-1408.
[6] Zmud, J., Sener, I. N., & Wagner, J. (2016). Self-driving vehicles, Determinants of adoption and conditions of usage. Transportation Research Record, 2565(1), 57-64.
[7] Liu, L., Lu, S., Zhong, R., Wu, B., Yao, Y., Zhang, Q., & Shi, W. (2020). Computing systems for autonomous driving, State of the art and challenges. IEEE Internet of Things Journal, 8(8), 6469-6486.
[8] Behere, S., & Törngren, M. (2015, May). A functional architecture for autonomous driving. In Proceedings of the first international workshop on automotive software architecture (pp. 3-10).
[9] Gao, C., Wang, G., Shi, W., Wang, Z., & Chen, Y. (2021). Autonomous driving security, State of the art and challenges. IEEE Internet of Things Journal, 9(10), 7572-7595.
[10] Hu, L., Zhou, X., Zhang, X., Wang, F., Li, Q., & Wu, W. (2021). A review on key challenges in intelligent vehicles, Safety and driver-oriented features. IET Intelligent Transport Systems, 15(9), 1093-1105.
[11] Inoue, H., Raksincharoensak, P., & Inoue, S. (2017). Intelligent driving system for safer automobiles. Journal of Information Processing, 25, 32-43.
[12] Martinez, C. M., Heucke, M., Wang, F. Y., Gao, B., & Cao, D. (2017). Driving style recognition for intelligent vehicle control and advanced driver assistance, A survey. IEEE Transactions on Intelligent Transportation Systems, 19(3), 666-676.
[13] Saleem, H., Riaz, F., Mostarda, L., Niazi, M. A., Rafiq, A., & Saeed, S. (2021). Steering angle prediction techniques for autonomous ground vehicles, a review. IEEE Access, 9, 78567-78585.
[14] Li, Z., Liu, Y., Shin, K. G., Li, J., Guo, F., & Liu, J. (2019). Design and adaptation of multi-interference steering. IEEE Transactions on Wireless Communications, 18(7), 3329-3346.
[15] Lin, C. F., Juang, J. C., & Li, K. R. (2014). Active collision avoidance system for steering control of autonomous vehicles. IET Intelligent Transport Systems, 8(6), 550-557.
[16] Bao, C., Feng, J., Wu, J., Liu, S., Xu, G., & Xu, H. (2020). Model predictive control of steering torque in shared driving of autonomous vehicles. Science Progress, 103(3), 0036850420950138.
[17] Li, Y., Cai, Y., Sun, X., Wang, H., Jia, Y., He, Y., ... & Chao, Y. (2024). Trajectory tracking of four-wheel driving and steering autonomous vehicle under extreme obstacle avoidance condition. Vehicle system dynamics, 62(3), 601-622.
[18] Liu, J., Guo, H., Song, L., Dai, Q., & Chen, H. (2020). Driver-automation shared steering control for highly automated vehicles. Science China Information Sciences, 63, 1-16.
[19] Leng, B., Jiang, Y., Yu, Y., Xiong, L., & Yu, Z. (2021). Distributed drive electric autonomous vehicle steering angle control based on active disturbance rejection control. Proceedings of the Institution of Mechanical Engineers, Part D, Journal of Automobile Engineering, 235(1), 128-142.
[20] Helong Yu,Xianhe Cheng,Ziqing Li,Qi Cai & Chunguang Bi. (2022). Disease Recognition of Apple Leaf Using Lightweight Multi-Scale Network with ECANet. Computer Modeling in Engineering & Sciences,132(3),711-738.
[21] Zhengda Zhou,Yeming Dai & Mingming Leng. (2025). A photovoltaic power forecasting framework based on Attention mechanism and parallel prediction architecture. Applied Energy,391,125869-125869.