

DOI:10.11918/202403056

基于 DQN 的快速路合流区 CAV 换道决策方法

程国柱¹, 王文志¹, 陈永胜¹, 徐 亮²

(1. 东北林业大学 土木与交通学院, 哈尔滨 150040; 2. 长春工程学院 土木工程学院, 长春 130012)

摘 要: 为解决快速路合流区的交通拥堵和安全问题, 保障智能网联车辆 (connected and automated vehicle, CAV) 在快速路合流区高效、安全、舒适、稳定地行驶, 采用深度强化学习方法之 DQN (deep q-network) 算法, 综合考虑车辆行驶安全性、高效性与舒适性等因素, 建立了用于神经网络训练的奖励函数模型, 提出了快速路合流区 CAV 换道决策方法。应用开源自动驾驶仿真场景 highway-env, 搭建快速路合流区仿真环境, 对其主线路段及匝道进行仿真试验。仿真试验结果表明: 相比于智能驾驶人模型 (intelligent driver model, IDM) 和 highway-env 中换道决策方法, 采用所提出的快速路合流区 CAV 换道决策方法时, CAV 能够快速达到 22.22 m/s 行驶的稳态, 同时避免频繁换道及加减速行为, 并优化了车头时距, 很大程度上提高了交通流运行效率和乘车舒适性。研究成果在智能交通领域中, 可为智能网联环境下城市快速路合流区的车辆通行提供了一种新方法, 为未来智能网联车辆的换道提供了决策方法。

关键词: 智能交通; CAV 换道决策方法; DQN; 智能网联车辆; 深度强化学习; 快速路合流区

中图分类号: U461

文献标志码: A

文章编号: 0367-6234(2025)03-0098-12

Decision method of CAV lane change in expressway merging area based on DQN

CHENG Guozhu¹, WANG Wenzhi¹, CHEN Yongsheng¹, XU Liang²

(1. School of Civil Engineering and Transportation, Northeast Forestry University, Harbin 150040, China;

2. School of Civil Engineering, Changchun Institute of Technology, Changchun 130012, China)

Abstract: In order to tackle traffic congestion and safety issues in expressway merging areas and to ensure efficient, safe, comfortable, and stable travel of connected and automated vehicles (CAVs) in these areas, this study employs the DQN (deep q-network) algorithm from deep reinforcement learning. By considering factors such as vehicle safety, efficiency, and comfort, a reward function model for neural network training has been established, and a CAV lane-change decision-making method for merging areas has been proposed. Using the open-source highway-env simulation scenario, a simulation environment for expressway merging areas has been set up, and experiments have been conducted on the mainline and ramps. The results of the simulation experiments show that compared to the intelligent driver model (IDM) and the lane-change decision-making method in highway-env, the proposed CAV lane-change decision-making method enables CAVs to quickly reach a stable driving state at a speed of 22.22 m/s. It also reduces frequent lane changes and acceleration/deceleration behaviors, and optimizes the time-headway between vehicles. This significantly improves the efficiency of traffic flow and ride comfort. The research findings provide a new method for vehicle traffic management in urban expressway merging areas under intelligent networked conditions. They also offer a decision-making approach for lane changes in future connected and automated vehicles.

Keywords: intelligent transportation; CAV lane change decision method; DQN; intelligent connected vehicles; deep reinforcement learning; expressway merging area

随着机动车拥有量的不断提升, 城市交通拥堵日益严重, 为了解决这一问题, 人们以高速公路的建设理念在城市中打造立体交通, 从而建设了城市快速路^[1]。但城市机动车保有量的进一步增长造成了城市快速路流量的急剧增加, 主线车流与匝道车

流形成的交通冲突问题, 使合流区经常发生交通拥堵状况, 交通流运行风险随之提高^[2]。智能网联车辆的发展对于解决城市快速路合流区交通安全、拥堵问题有着巨大的影响。且在 CAV 的发展阶段中, 深度强化学习 (deep reinforcement learning, DRL) 领

收稿日期: 2024-03-26; 录用日期: 2024-04-28; 网络首发日期: 2025-02-18

网络首发地址: <https://link.cnki.net/urlid/23.1235.T.20250217.1656.002>

基金项目: 中央高校基本科研业务费专项资金 (2572023CT21); 国家自然科学基金 (52378433)

作者简介: 程国柱 (1977—), 男, 教授, 博士生导师

通信作者: 程国柱, guozhucheng@126.com

域的进步极大推动了自动驾驶决策技术的发展^[3]。因此,本文将利用深度强化学习方法建立快速路合流区CAV换道决策方法,为进一步解决快速路合流区的交通拥堵问题提供新的方法和思路。

彭浩楠等^[4]利用逆强化学习方法,实现神经网络对最优轨迹规划的学习。朱冰等^[5]通过实际测量驾驶数据的方式,针对车辆纵向行为对神经网络进行训练。高振海等^[6]采用实际驾驶数据训练逆强化学习算法模仿驾驶员的行为。以上研究都显示了机器学习和逆强化学习在自动驾驶车辆轨迹规划和行为模仿中的应用潜力。

李伟东等^[7]提出结合决策层和控制层的双层强化学习模型训练神经网络以提高算法的计算效率。刘卫国等^[8]通过分布式训练的方法提升系统的计算速度。Wang等^[9]通过在多个驾驶演示数据集上并行异步执行多个监督学习过程提高强化学习训练阶段的初始性能和收敛速度。宋晓琳等^[10]结合模仿学习与强化学习构建智能车辆换道策略提升决策速度。以上研究都体现了在智能车辆领域,如何通过不同的方法和技术提高强化学习算法的效率和性能,提高其在实际道路环境中的适应性和安全性。

张志勇等^[11]改进双延迟深度确定性策略梯度算法提高Q值的估计精度,以提高并道决策的准确性。杨顺等^[12]通过对多种特征数据进行卷积的方式训练神经网络,提高自动驾驶车辆运行速度和偏移控制效果。张雪峰等^[13]提出双决斗深度Q网络强化学习模型对神经网络进行训练来进行智能网联车辆换道决策。谷梦路等^[14]提出一种结合随机森林算法和深度Q网络算法的ICV汇入控制模型,提高ICV在与高速公路主线上的传统人工驾驶车辆交互时的汇入效率和成功率。姚福星等^[15]提出一种基于混合专家模型的智能网联汽车换道决策方法,实现连续控制、高安全、高可靠的自适应高速换道。以上研究都表明了深度强化学习算法的改进和应用对于提升自动驾驶和智能网联车辆的决策能力具有重要作用。通过精确的Q值估计、高效的特征

提取和稳定的决策机制,可以显著提高车辆的自主驾驶性能。

张羽翔等^[16]采用强化学习方法学习决策参数,在奖励函数中考虑了安全性、当前驾驶员的意愿和平均人类驾驶风格。这项研究表明,强化学习可以为自动驾驶车辆提供一种有效的决策学习框架,尤其是在需要综合考虑多种因素时。通过设计的奖励函数,可以在不牺牲安全性的前提下,使自动驾驶车辆的行为更加符合人类的期望和习惯。

综上所述,在智能网联车辆的研究领域,尤其是针对城市快速路合流区的行驶策略,尽管已有一些研究运用深度强化学习方法^[17-19],但主要聚焦于车辆行驶效率及安全性等方面,少有研究考虑驾驶过程中的稳定性及乘客舒适性等因素。

本文采用了DQN算法^[20],并对其奖励函数进行了创新性的改进。在设计奖励函数时,本文不仅关注于提升车辆的期望速度,以确保在城市快速路上的高效通行,同时也将行驶安全性和乘客乘车舒适性纳入考量。通过这种多维度的奖励设计,本文旨在实现一个综合性的优化目标:在确保车辆快速通过路段的同时,兼顾行车安全和乘客的良好乘车体验。

1 CAV换道决策方法

1.1 换道决策场景及研究对象

本文研究场景为城市快速路合流区,行驶车辆包括目标CAV和环境车辆,如图1所示。CAV能与道路传感通信控制设施(即路侧单元,road side unit,RSU)网联通信,并受其控制,而环境车辆则由路侧单元通过传感器雷达等监测其行驶信息。在此场景下,当目标CAV位于匝道时,需结合车辆的行驶安全、速度和加速度,考虑并入主线的最佳时机;当目标CAV位于主线时,会受到匝道车辆并入行为和其他主线车辆行驶状态的影响。因此,本文将基于马尔科夫性,设置深度强化学习的奖励函数,以保障目标CAV在合流区中不同类型道路上行驶时的安全、高速和稳定。

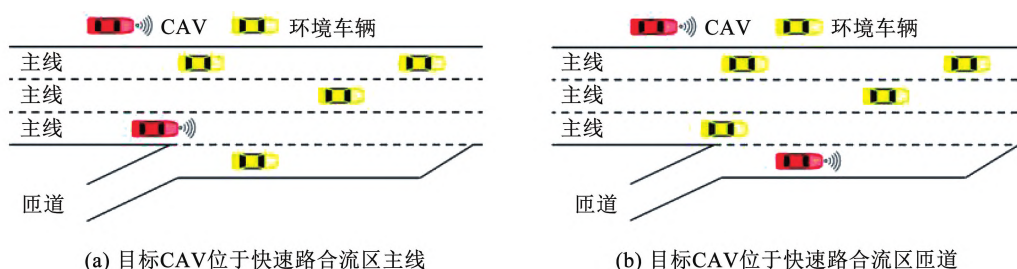


图1 快速路合流区

Fig. 1 Merging area of expressway

1.2 换道决策思路与流程

本文针对 CAV 换道决策过程主要考虑目标 CAV 位于快速路合流区主线和匝道两个场景。目标车辆(主车)首先判断处于哪种行驶环境。若为快速路合流区主线,则进一步判断前方车辆是否处于影响自身保持原车道行驶的范围,此影响范围通过比较目标车辆以期望速度在期望超车总时间 $18\text{ s}^{[21]}$ 内的行驶距离是否大于前车在此段时间的行驶距离,若小于则前车处于影响范围内,反之前车则不处于影响范围内。若前车未在影响范围内,目标车辆继续保持原车道行驶;若在影响范围内,目标车辆继续判断前方车辆的速度是否大于自身速度,若前车速度大于主车速度,则目标车辆保持原车道行驶;反之,目标车辆进入换道判断阶段。首先由最小安全距离计算模型^[22]判断换道是否安全,若不安

全,则放弃换道,减速跟随前方车辆;若安全,进一步通过对比奖励值决定换道方向,最终完成换道。以此方式往复循环,直至主车到达目的地。若为快速路合流区匝道,本文只考虑合流区入口。目标车辆到达合流区起点时开始判断换道是否安全,若不安全,则继续保持原车道行驶,此情况下,如果目标车辆到达合流区终点,则以“先到先服务”的原则进行强制换道,即在主线后车与合流区终点的间距满足最小安全距离、可以安全实现减速或停车让行的情况下,目标车辆忽略速度、乘车舒适度等因素采取换道行为并入主线,避免造成匝道的排队拥堵情况;若安全,则通过对比奖励值判断最佳的换道时机,完成换道,进入快速路合流区主线。换道决策流程如图2所示。

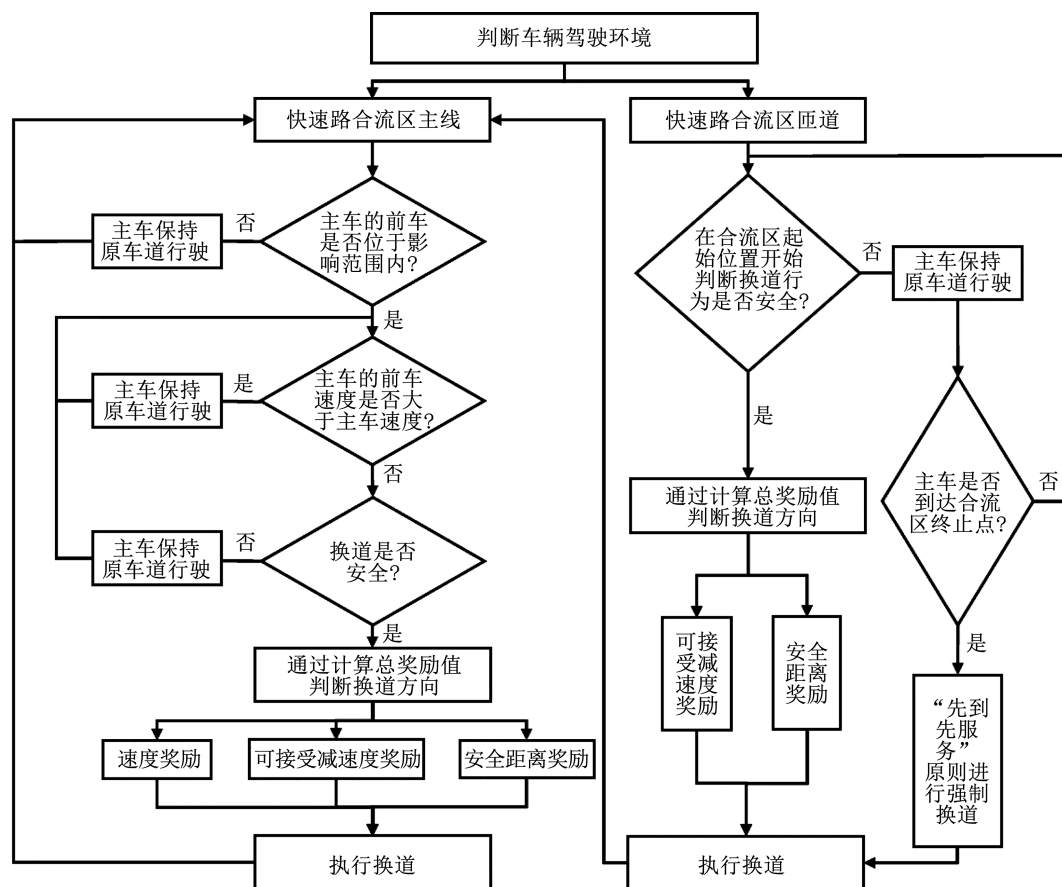


图2 快速路合流区智能网联车辆换道决策流程

Fig. 2 Lane change decision-making process of intelligent networked vehicles in expressway merging area

2 基于马尔科夫过程的奖励函数

2.1 奖励函数构建

马尔科夫决策过程(Markov decision processes, MDPs)^[23-24]作为强化学习的数学基础,通常由状态空间、动作空间、状态转移矩阵、奖励函数以及折扣因子等组成。本文通过马尔科夫决策过程建立奖励

函数,并通过训练神经网络确定折扣因子。最终利用 DQN 算法评估最优动作价值,即 CAV 在快速路合流区行驶的最优决策。

在马尔科夫过程的基础上增加奖励函数 R 和衰减系数 γ , 表示为 (S, P, R, γ) 。 R 是一个奖励函数, S 状态下的奖励是某一时刻(t)处在状态 s 的下一个时刻($t+1$)能获得的奖励期望,具体形式为

$$R_s = E[R_{t+1} | S_t = s] \quad (2)$$

基于此,从 t 时刻开始所得到的折扣回报总和为

$$G_t = R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \cdots = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \quad (3)$$

式中折扣因子 γ 代表了对未来奖励的折扣。 γ 越大,代表更重视未来奖励;相反, γ 越小,则更重视近期奖励。

在城市快速路场景中,智能网联车辆(目标车辆)在不同的道路环境,即状态 s_i 下将选择不同的驾驶行为 a_i ,每个驾驶行为得到相应的奖励或惩罚 r_i ,使得目标车辆进入下一个状态 s_{i+1} ,以此循环往复得到最优的驾驶决策,流程如图3所示。其中,状态 $s_i \in S$,行为 $a_i \in A$,奖励 $r_i \in R$ 。

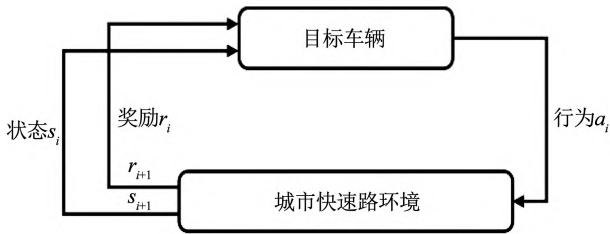


图3 智能网联车辆强化学习流程

Fig. 3 Reinforcement learning process for connected and automated vehicles

2.1.1 快速路合流区主线的奖励函数设计

当目标车辆位于快速路合流区的主线时,其在行驶过程中的主要需求为更短的通行时间、更舒适的乘坐体验和更安全的驾驶行为。因此,在奖励值的计算中,本文主要考虑以下4个方面的奖励。

1) 速度奖励 R_v 。当目标车道前车在影响主车的距离范围内时,主车速度 v_{MV} 与目标车道前车的速度 v_{FV} 之比。考虑目标车辆可能在换道期间与前车发生追尾事故,通过最小安全距离模型计算换道安全距离,若不满足换道安全条件,则速度奖励的折扣因子为0,目标车辆将放弃追求更高的车速,保持安全条件下的跟驰行为:

$$R_v = \frac{v_{MV}}{v_{FV}} \quad (4)$$

2) 可接受减速度奖励(针对前车) R_{b1} 。当目标车道前车在影响主车的距离范围内时,主车可接受减速度 b_{MV} 与目标车道前车减速度 b_{FV} 之比:

$$R_{b1} = \frac{b_{MV}}{b_{FV}} \quad (5)$$

3) 可接受减速度奖励(针对后车) R_{b2} 。当目标车道后车在影响主车的距离范围内时,主车可接受减速度 b_{MV} 与目标车道后车减速度 b_{LV} 之比:

$$R_{b2} = \frac{b_{MV}}{b_{LV}} \quad (6)$$

4) 安全距离奖励 R_d 。换道临界间隙 s_{CG} 与行驶安全距离 s_{safe} 之比:

$$R_d = \frac{s_{CG}}{s_{safe}} \quad (7)$$

R_v 值越高,代表车辆的换道行为将会获得更高的行驶速度; R_{b1} 和 R_{b2} 值越高,代表车辆的换道行为对于乘客来说更加舒适; R_d 值越高,代表车辆的换道行为更安全。综上,奖励函数从速度、安全以及舒适性三方面综合考虑车辆换道行为产生的影响。

R_{b1} 和 R_{b2} 的计算方法是利用SITRAS模型^[25]判断是否可以换道,分为两个条件:1) 主车跟随目标车道前车的减(加)速度是否大于主车可接受的减速度;2) 目标车道后车跟随主车时的减速度是否大于目标车道后车可接受的减速度。减速度计算公式为

$$b_n = [2 - (D - x_n(t)) / (10v_n)] \cdot b_{LC} \cdot \theta \quad (8)$$

式中: b_n 为车辆 n 在 t 时刻可接受的减速度, D 为预定的转向或障碍物的位置, x_n 为车辆 n 在 t 时刻的位置, v_n 为车辆 n 的期望速度, b_{LC} 为车辆愿意接受的减速度的平均值, θ 为司机的风险系数。

目标车辆行驶决策的奖励具体形式为

$$R_t = \lambda_v R_v + \lambda_{b1} R_{b1} + \lambda_{b2} R_{b2} + \lambda_d R_d \quad (9)$$

式中:折扣因子 λ 代表了对不同奖励的折扣,用以平衡各奖励之间的相互影响。其中 λ_v 为速度奖励折扣因子, λ_{b1} 为可接受减速度奖励(针对前车)折扣因子, λ_{b2} 为可接受减速度奖励(针对后车)折扣因子, λ_d 为安全距离奖励折扣因子。

为保障交通流安全,便于神经网络进行计算,设置偏置项 w_1 作为车辆发生碰撞的惩罚;设置偏置项 w_2 作为车辆频繁换道的惩罚来保障乘客乘车的舒适性;设置偏置项 w_3 作为车辆出界的惩罚。则奖惩机制下的神经网络计算函数具体形式为

$$Q_{\text{value}}(v, b, s, w_1, w_2, w_3) = \lambda_v R_v + \lambda_{b1} R_{b1} + \lambda_{b2} R_{b2} + \lambda_d R_d + w_1 + w_2 + w_3 \quad (10)$$

2.1.2 快速路合流区匝道的奖励函数设计

当目标车辆处于快速路匝道时,车辆在到达换道终止点前必须完成换道,否则匝道将产生排队拥堵现象。因此,在快速路合流区环境中,匝道上的车辆换道时无需考虑主车速度与目标车道前车速度之间的关系,则 t 时刻所得到的奖励具体形式为

$$R_t = \lambda_{b1} R_{b1} + \lambda_{b2} R_{b2} + \lambda_d R_d \quad (11)$$

因为匝道处车辆行驶入主路仅需进行一次换道,所以无需考虑车辆的频繁换道行为,所以对于在合流区的匝道车辆,仅需设置偏置项 w_1 、 w_3 ,所以此环境下神经网络计算函数具体形式为

$$Q_{\text{value}}(b, s, w_1, w_3) = \lambda_{b1} R_{b1} + \lambda_{b2} R_{b2} + \lambda_d R_d + w_1 + w_3 \quad (12)$$

2.2 奖励函数参数设置

深度强化学习方法之深度 Q 网络(DQN)结合 Q 学习和深度神经网络,利用训练后的神经网络确定目标体在某一状态下执行动作的期望奖励。DQN 算法的神经网络结构需要根据问题的性质进行设计,即需要通过调整学习率等参数来提高算法的学习效率。

神经网络是一种模仿生物神经网络的结构和功能的数学模型,在本文用于对目标车辆的行为进行识别、预测和执行。参考文献[6-7],设置神经网络架构包含两个隐藏层,每个隐藏层有 256 个神经元,通过设置学习率控制权重更新的步长,设置经验回放缓冲区存储智能体的经验,设置折扣因子用于未来奖励的当前价值,设置 tensor board 用来监控训练过程。具体参数设定值见表 1。

表 1 神经网络参数设置

Tab. 1 Neural network parameter setting

参数名称	代码名称	参数标定值
学习率	learning_rate	0.000 5
缓存区大小	buffer_size	15 000
学习前预热步数	learning_starts	200
样本数量	batch_size	32
网络折扣因子	gamma	0.8
更新网络步数	train_freq	1
更新网络次数	gradient_steps	1
训练过程控制参数	verbose	1
网络更新频率	target_update_interval	50

折扣因子用于平衡神经网络中的各项奖励,确保模型在学习时不会只关注某一特定行为,形成适用多种情况的策略。奖励值的偏高或偏低都会导致目标车辆偏重某一方面而忽略其他,从而造成目标车辆运行过程中奖励值突变,即受到惩罚。本文设置不同数值的 5 组折扣因子,对比目标车辆运行期间其奖励值的变化情况如图 4 所示。

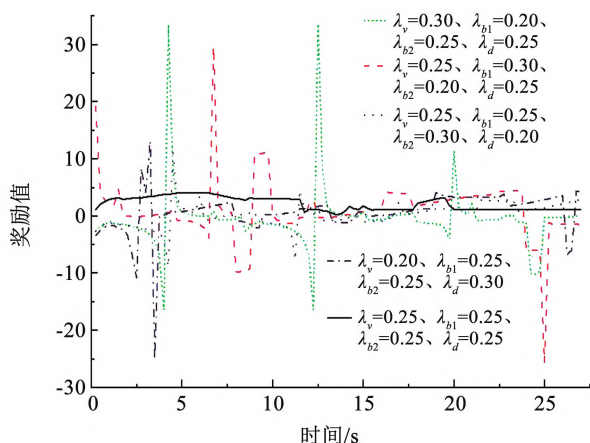


图 4 奖励值变化曲线

Fig. 4 Reward value change curve

可知,当各折扣因子均为 0.25 时,目标车辆在运行过程中所获的奖励值变化最为平稳,即在此条件下目标车辆可以平等地考虑每个驾驶行为,此时各项奖励均有效。因此确定各项奖励折扣因子 λ_i 取值见表 2。

表 2 折扣因子设置

Tab. 2 Discount factor setting

CAV 位置	λ_v	λ_{b1}	λ_{b2}	λ_d
CAV 位于主线	0.25	0.25	0.25	0.25
CAV 位于匝道	0	0.25	0.25	0.25

偏置项有助于神经网络模型更快收敛,提高计算效率,减少过拟合。目标车辆的表现受多种因素影响,需要通过训练和仿真测试来调整偏置项值,以明确任务目标,避免低效学习,并应对环境复杂性。因此,需要综合考量不同组 w_i 取值下完成有限的训练次数后目标车辆的平均碰撞发生率、平均车速和平均换道次数确定偏置项值。其中,碰撞发生率设为 1×10^6 组随机仿真试验中目标车辆发生碰撞的概率。各组偏置项取值见表 3。其中,3 组事故的发生率分别为 24%、16%、13%,CAV 位于主线时的平均车速分别为 21.39、22.21、15.07 m/s,CAV 位于匝道时的平均车速分别为 17.48、18.32、15.51 m/s,CAV 位于主线时的平均换道次数分别为 3.1、2.6、1.2,CAV 位于匝道时的平均换道次数均为 3。可见,组 2 的事故发生率较组 1 低 0.08%,且在平均车速更高的同时换道次数更少,因此组 2 设置值优于组 1;组 3 的事故发生率虽然较组 2 低 0.03%,但因过度考虑车辆运行的稳定性而忽略了运行效率,因此大幅度降低了车速,且事故发生率相差不大,所以组 2 设置值优于组 3。最终选取组 2 设置值作为奖励函数的偏置项设置值。

表 3 偏置项设置

Tab. 3 Offset item setting

CAV 位置	组别	w_1	w_2	w_3
CAV 位于主线	组 1	-5	-1	-15
	组 2	-10	-1	-20
	组 3	-15	-1	-25
CAV 位于匝道	组 1	-25	0	-15
	组 2	-30	0	-20
	组 3	-35	0	-25

3 CAV 换道决策方法验证

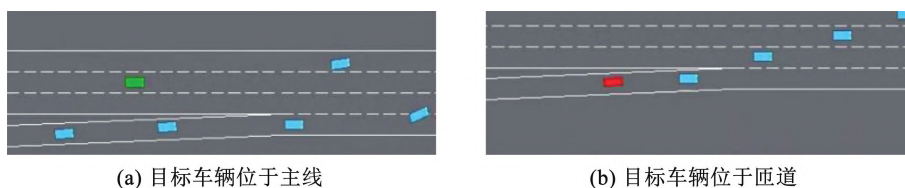
3.1 仿真环境搭建

为了验证本文 CAV 换道决策的可行性,依托

Python 搭建仿真试验环境,考虑城市快速路的建设规范,以及车辆行驶时的不同环境,搭建快速路合流区。参考《城市快速路设计规程》设置3条主线车道和1条匝道,合流区长度为80 m,单车道宽度为3.5 m,主线限速为80 km/h,匝道限速为40 km/h,试验车道总长度为510 m,试验部分匝道长度为150 m。

在上述设置条件下,依托 Python 搭建的城市快速路合流区仿真试验环境能够模拟不同的驾驶场景,其中目标 CAV 可能位于合流区的主线或匝道,

如图5所示。图5(a)中,目标CAV位于合流区主线,图5(b)中,目标CAV位于合流区匝道的情况。除目标车辆外,环境中其他车辆均采用IDM模型^[26]进行控制。此外,为简化研究,参考文献[27]对车辆特征以及车辆跟驰模型的参数进行标定,车辆长度为5 m,车辆宽度为1.8 m,最大加速度为 1.4 m/s^2 ,最大减速度为 3 m/s^2 ,舒适减速度为 2 m/s^2 ,车辆间最小安全距离为2 m,司机风险系数为1。



(a) 目标车辆位于主线

(b) 目标车辆位于匝道

图5 快速路合流区仿真环境

Fig.5 Simulation environment of expressway merging area

3.2 试验方案

在开源自动驾驶仿真场景 highway-env 中,原DQN算法配置的奖励函数主要侧重于3个关键目标:提升目标CAV的速度,尽量保持在最右侧车道行驶,并避免与其他车辆发生碰撞(简称VRL策略)。而在改进的DQN算法中,奖励函数进一步考虑了提供舒适的减速度,并减少目标CAV的频繁换道和加减速行为,同时确保目标CAV与其他车辆之间保持安全的行驶距离(简称VBSD策略)。在这个仿真环境中,其他车辆采用智能驾驶模型进行模拟。

针对目标CAV在合流区的不同位置,设置了不同的起始速度范围:当目标CAV位于合流区主线时,其起始速度设置在50~70 km/h之间;当目标CAV位于合流区匝道时,其起始速度设置在20~40 km/h之间。此外,设定仿真环境中其他车辆的数量为10辆,且控制仿真环境的空间占有率在10.78%。整个仿真的总时长设定为50 s,在这段时间内持续记录仿真运行的数据。分别在以下4种场景训练神经网络模型:VRL策略下,目标车辆位于快速路合流区主线和匝道的场景;VBSD策略下,目标车辆位于快速路合流区主线和匝道的场景。每种场景迭代 1×10^8 个回合,每个回合起始于目标CAV进入试验路段起点,终止于目标CAV到达试验路段终点,完成所有迭代后将获得适应不同策略的神经网络模型。

在模型训练完成后,进行仿真模拟。在奖励函

数设计中设置了速度奖励、可接受减速度奖励、安全距离奖励以控制CAV更高速且平稳地行驶,因此在仿真试验中收集目标车辆和环境车辆的行驶速度、加速度、与前车的车头时距数据,以及仿真过程中的换道次数。通过这些数据,可以对不同策略下的神经网络模型进行对比分析,从而评估VRL和VBSD策略在实际仿真环境中的表现效果,验证本文所设计奖励函数的可行性。

3.3 主线仿真结果分析

3.3.1 车速和加速度对比分析

为避免数据集偏差,神经网络训练完成后分别进行了4组速度和加速度的随机场景仿真试验。通过不同的交通环境和条件来检验模型的泛化和性能一致性,确保结论的可靠性。试验结果如图6和图7所示。

对试验数据进行分析,在采用VBSD策略的情况下,目标车辆在4种不同的场景中能够在仿真开始后的11、13、11、11 s时达到22.22 m/s的速度,并在此速度水平上保持稳定行驶。与此同时,加速度在这些场景中均保持在0值附近的小幅波动。相比之下,VRL策略下的目标车辆在仿真开始后的46、27、27、32 s时速度达到22.22 m/s,但在此期间存在超速行为,表明神经网络模型对车辆行驶的控制出现了问题。在VRL策略中,尽管目标车辆的加速度在某些时段保持稳定,但经常出现急剧增加或减少的情况,这可能导致乘客体验不佳。

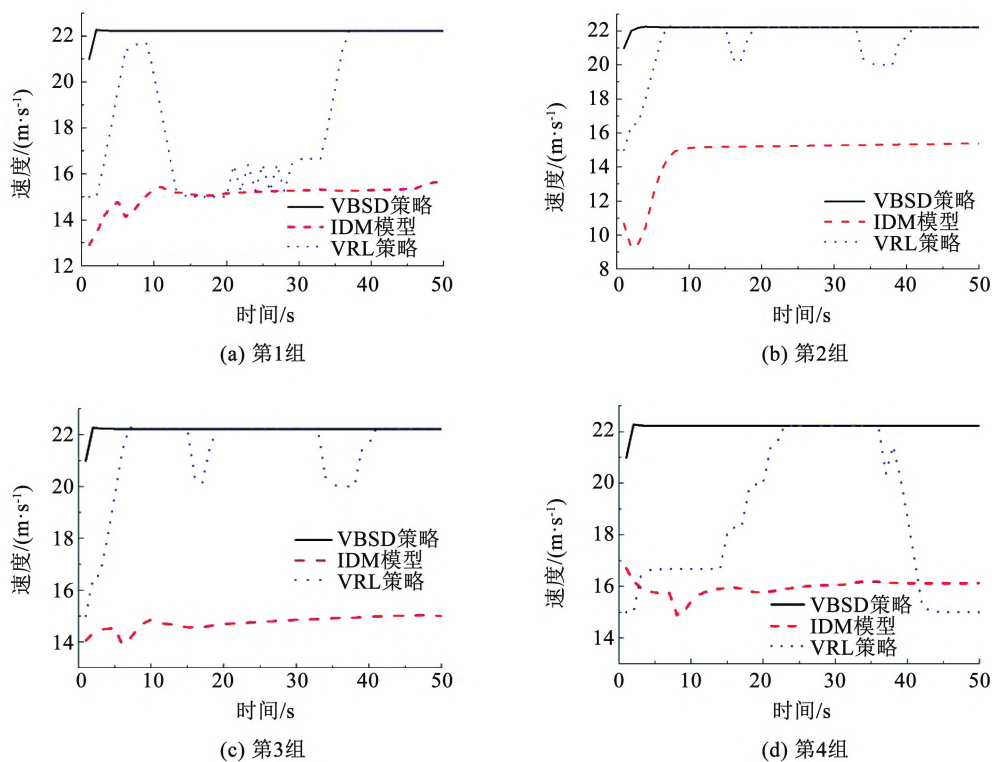


图 6 合流区主线车辆速度对比

Fig. 6 Comparison of the speed of the main line vehicles in the merging area

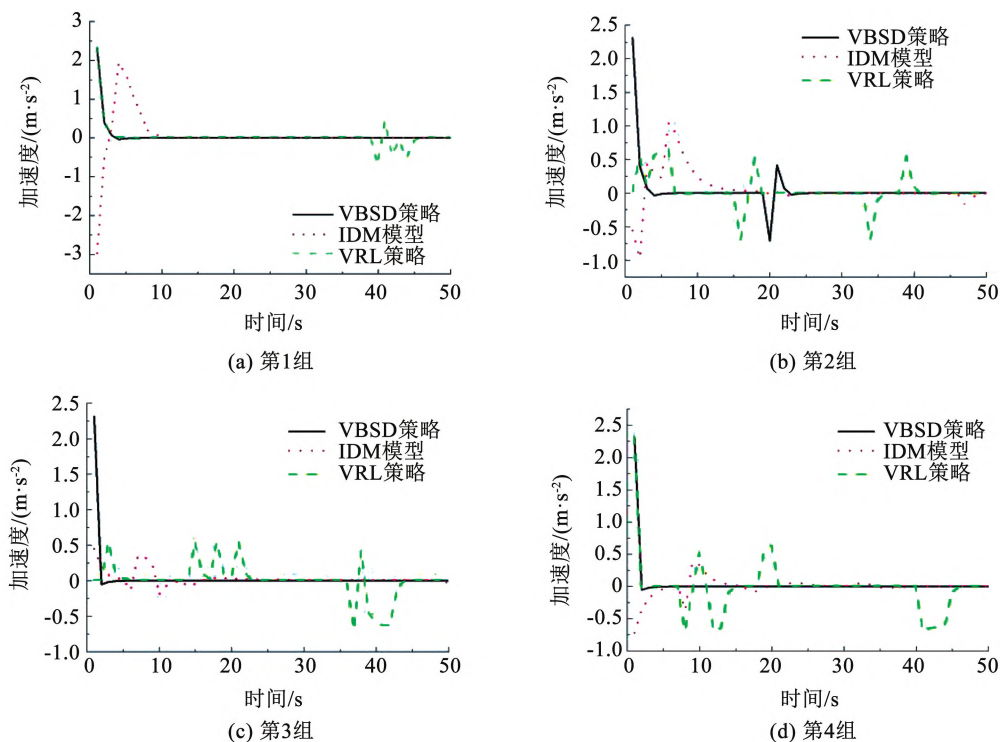


图 7 合流区主线车辆加速度对比

Fig. 7 Comparison of the acceleration of the main line vehicles in the merging area

VBSD 策略在与 VRL 策略和 IDM 模型比较时,具有以下优势:与 VRL 和 IDM 相比,VBSD 能控制车辆以更高速度行驶;VRL 策略下车辆速度和加速度波动较大,可能影响乘客舒适度;VBSD 策略能维持车辆期望速度,并保持速度和加速度稳定,这种稳

定性在提升行驶效率的同时,也为乘客提供了更舒适的乘车体验。因此,VBSD 在提升行驶效率和乘客舒适度方面都优于 VRL 和 IDM。

3.3.2 换道次数对比分析

换道次数是影响乘车舒适性的一个重要因素之一,

频繁的换道会导致乘客感到不适。为了提高乘车舒适性,CAV需要在确保安全的前提下,尽量避免不必要的换道行为。本文进行了5组随机仿真试验,统计每次仿真试验中车辆行驶过程的换道次数,并

定义连续换道次数为目标CAV在行驶中连续执行了两次及以上换道行为的次数,无效换道次数为目标CAV在前方无车辆影响其行驶的情况下执行换道行为的次数,见表4。

表4 CAV主线行驶过程换道次数统计

Tab.4 Statistics on lane changes during the running of the CAV main line

仿真次序	换道次数		连续换道次数		无效换道次数	
	VBSD策略	VRL策略	VBSD策略	VRL策略	VBSD策略	VRL策略
1	4	9	0	3	0	1
2	8	16	1	5	3	3
3	4	8	0	1	1	2
4	6	14	0	4	1	8
5	6	14	1	3	0	7

根据表4所获数据进行系统分析,在换道次数方面,VBSD策略控制下的CAV平均换道次数为5.6次,VRL策略控制下的CAV平均换道次数为12.2次,VBSD策略相较VRL策略降低了54.10%;在连续换道次数方面,VBSD策略控制下的CAV平均连续换道次数为0.4次,VRL策略控制下的CAV平均连续换道次数为3.2次,VBSD策略相较VRL策略降低了87.50%;在无效换道次数方面,VBSD策略控制下的CAV平均无效换道次数为1.0次,VRL策略控制下的CAV平均无效换道次数为

4.2次,VBSD策略相较VRL策略降低了76.19%。综上所述,VBSD策略在以下方面优于VRL策略:能更有效地减少频繁和无效的换道行为;提高了行车安全性和乘客乘车舒适度,可为乘客提供更佳的乘车体验。

3.3.3 车头时距分析

分别采用VRL策略与VBSD策略进行两组随机的仿真模拟试验,仿真结果如图8所示,图8(a)、8(b)分别为目标车辆位于快速路合流区主线的两种不同随机场景。

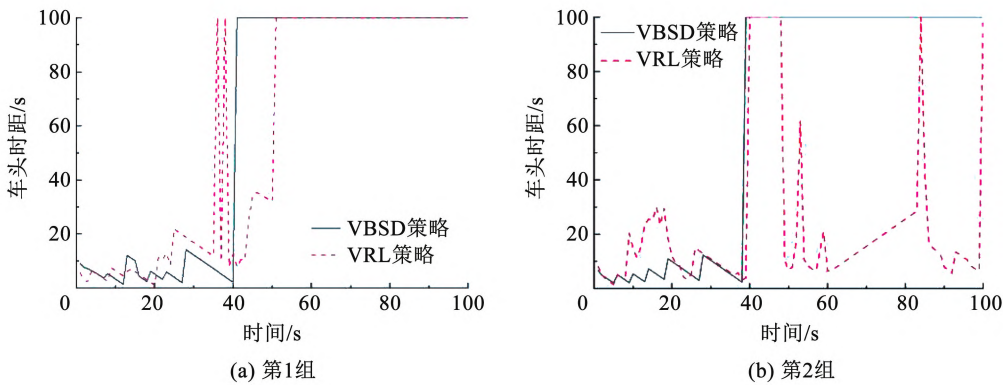


图8 合流区主线车头时距对比

Fig.8 Comparison of the time distance between the front of the main line in the confluence area

根据试验数据,VBSD策略中目标车辆与前车车头时距在两种场景中分别于仿真开始41s和39s后进入持续平稳状态;VRL策略中此车头时距分别于仿真开始51s和40s后进入平稳状态,且随后仍有车头时距的大幅波动。综上,VBSD策略相较于VRL策略有以下优势:能更长时间维持车辆无影响行驶,优化了行驶环境;减少换道行为,提升了乘客舒适性;解决了VRL策略中车头时距突增突降导致

的行驶不稳定问题;通过改进,显著提高了神经网络计算和车辆行驶的稳定性。

3.4 匝道仿真结果分析

3.4.1 车速和加速度对比分析

同主线车速和加速度仿真试验分析,对目标车辆位于匝道的情况分别进行4组随机场景速度仿真试验和4组随机场景加速度仿真试验,试验结果如图9、10所示。

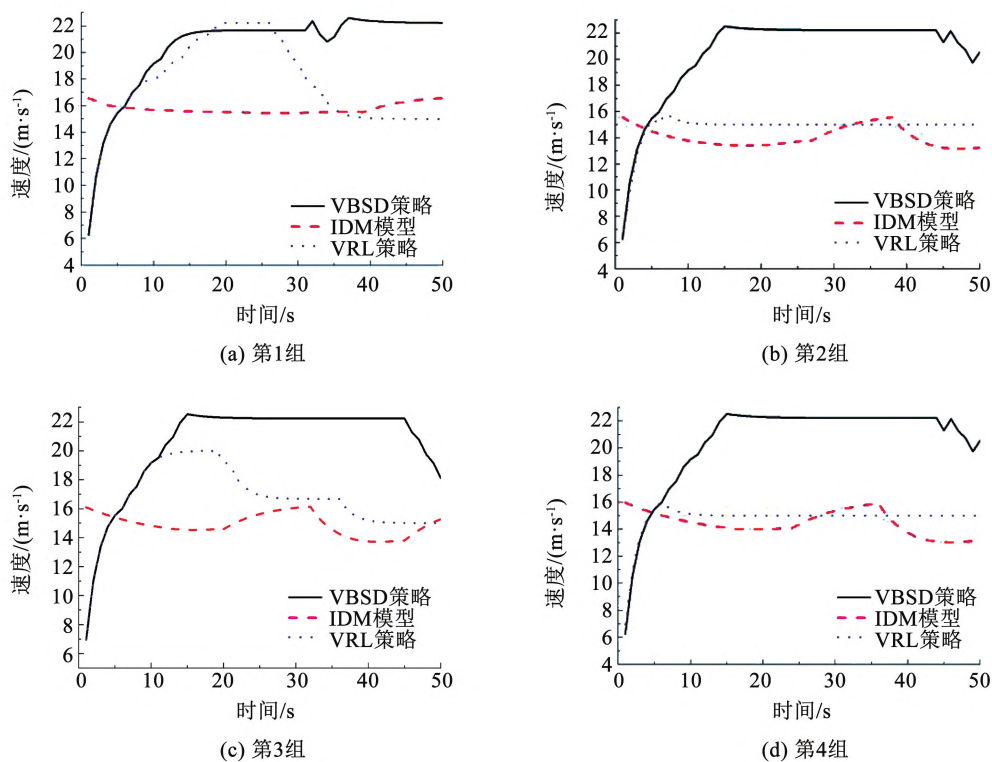


图 9 合流区匝道车辆速度对比

Fig. 9 Comparison of vehicle speeds on ramps in merging areas

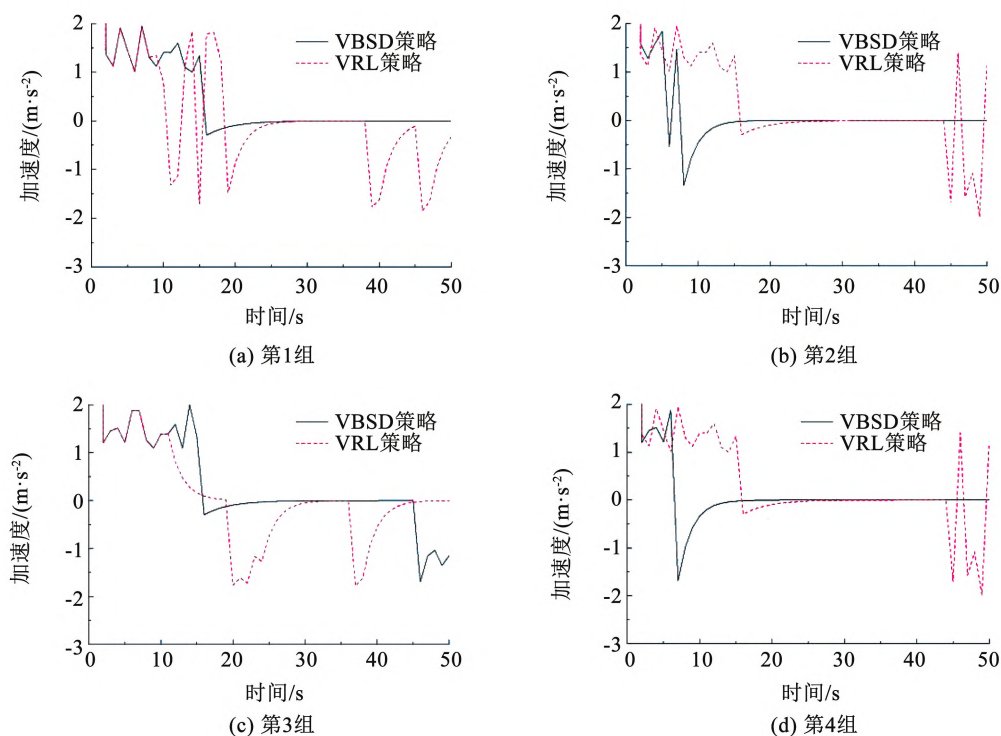


图 10 合流区匝道车辆加速度对比

Fig. 10 Comparison of vehicle acceleration at the ramp in the merging area

对试验数据进行分析,在采用 VBSD 策略的情况下,目标车辆在 4 种不同的场景中均能在仿真开始后的 36、15 s 时达到大约 22 m/s 的最高速度,并在此速度下保持稳定行驶。在这一策略控制下,加

速度基本维持在 0 值附近,仅出现小幅波动。相比之下,VRL 策略下的目标车辆在稳定运行时的最高速度仅达到 15 m/s,且在车辆有机会高速运行时,速度和加速度都会发生较为剧烈的波动。在相同场

景的对比中,VBSD策略在以下方面优于VRL策略和IDM跟驰模型:能更有效地控制车辆长时间高速行驶;在匝道行驶阶段,能更好地控制加速度变化,减少乘客不适;驶入主线后,能维持车辆稳定高速行驶,提升乘车体验。

3.4.2 换道次数对比分析

为比较VBSD策略和VRL策略在CAV从匝道并入主线时的稳定性和舒适度,记录仿真试验中的换道次数、连续换道次数和无效换道次数,来量化评估两种策略的影响。换道次数统计见表5。

表5 CAV由匝道并入主线换道次数统计

Tab.5 Statistics on the number of CAV lane changes from ramp to main line

仿真次序	换道次数		连续换道次数		无效换道次数	
	VBSD策略	VRL策略	VBSD策略	VRL策略	VBSD策略	VRL策略
1	3	3	0	1	0	2
2	3	3	0	1	1	2
3	1	3	0	1	0	2
4	3	3	0	1	1	2
5	3	3	0	1	0	2

根据表5所获数据进行统计分析,在换道次数方面,VBSD策略控制下的CAV在5次仿真试验中平均换道次数为2.6次,VRL策略控制下的CAV在5次仿真试验中平均换道次数为3次,VBSD策略相较VRL策略降低了13.33%;在连续换道次数方面,VBSD策略控制下的CAV在5次仿真试验中不存在连续换道行为,VRL策略控制下的CAV在5次仿真试验中平均连续换道次数为1次,VBSD策略相较VRL策略降低了100%;在无效换道次数方面,VBSD策略控制下的CAV在5次仿真试验中平均无效换道次数为0.4次,VRL策略控制下的CAV在5次仿真试验中平均无效换道次数为两次,VBSD策

略相较VRL策略降低了80%。

综上所述,VBSD策略在CAV匝道并入主线时比VRL策略更高效和稳定,主要优势如下:提升了行车安全性和乘客舒适度;减少了并入行为对主线交通流的负面影响;通过优化并入过程,减少了交通拥堵,提高了道路网络运行效率;能确保乘客享受到平稳且安全的乘车体验。

3.4.3 车头时距分析

分别采用VRL策略与VBSD策略进行两组随机的仿真模拟试验,仿真结果如图11所示,图11(a)、11(b)分别为目标车辆位于快速路合流区匝道的两种不同随机的场景。

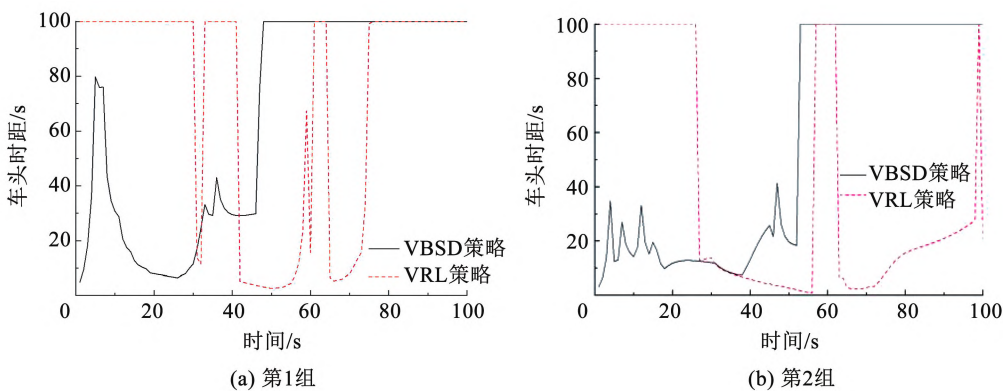


图11 合流区匝道车头时距对比

Fig.11 Comparison of the time distance between the front of the ramp in the merging area

由此可知,VBSD策略在匝道行驶时与VRL策略相比,有以下优势:能保持更小的车头时距,避免交通流速率降低;完成换道后,目标车辆能快速进入稳定行驶状态,保持前方无影响车辆。

4 结 论

本文针对智能网联车辆在城市快速路合流区的驾驶行为进行了深入研究,主要考虑CAV在合流区主线和匝道两种不同环境下的行驶状态。结合马尔科夫

过程对深度强化学习中的 DQN 算法进行创新性改进,提出了 VBSD 换道决策方法,该策略不仅考虑通行效率,还兼顾了行驶的稳定性与乘客的舒适性。利用 Python 搭建了仿真试验平台验证所提出的 VBSD 策略,得出如下研究结论:

1) 引入马尔科夫过程理论,对 DQN 算法的奖励函数进行了改进设计,实现了算法的优化。

2) 构建快速路合流区的仿真环境,使得单一智能体能够在该环境中进行有效的仿真测试。

3) VBSD 策略显著提升了 CAV 在快速路合流区的行驶表现。在主线上,使车辆高速且稳定行驶,减少无效换道,提升乘车体验;在匝道上,使车辆保持较短车头时距,为解决拥堵提供了理论支持;并入主线时,使车辆执行高效换道;并入后,使车辆进入高速稳定状态,解决了状态衔接问题。

本文的研究成果为智能网联车辆在城市快速路合流区的通行提供了新的方法和思路,但仍有待进一步探索。未来的研究将更深入地分析目标 CAV 对整体交通流的影响,并考虑多智能体构成的复杂交通流。

参考文献

- [1] 郝媛, 孙立军, 徐天东, 等. 城市快速路交通拥挤分析及拥挤阈值的确定[J]. 同济大学学报(自然科学版), 2008, 36(5): 609
HAO Yuan, SUN Lijun, XU Tiandong, et al. Analysis of traffic congestion on urban expressway and congestion threshold identification[J]. Journal of Tongji University (Natural Science), 2008, 36(5): 609
- [2] 程国柱, 王婉琦, 王连震, 等. 车辆换道行为风险评价方法研究进展[J]. 哈尔滨工业大学学报, 2023, 55(3): 139
CHENG Guozhu, WANG Wanqi, WANG Lianzhen, et al. Research progress on risk assessment of vehicle lane-changing behavior[J]. Journal of Harbin Institute of Technology, 2023, 55(3): 139
- [3] 金立生, 韩广德, 谢宪毅, 等. 基于强化学习的自动驾驶决策研究综述[J]. 汽车工程, 2023, 45(4): 527
JIN Lisheng, HAN Guangde, XIE Xianyi, et al. Review of autonomous driving decision-making research based on reinforcement learning[J]. Automotive Engineering, 2023, 45(4): 527
- [4] 彭浩楠, 唐明环, 查奇文, 等. 智能网联汽车基于逆强化学习的轨迹规划优化机制研究[J]. 北京理工大学学报, 2023, 43(8): 820
PENG Haonan, TANG Minghuan, ZHA Qiwen, et al. Research on inverse reinforcement learning-based trajectory planning optimization mechanism for autonomous connected vehicles[J]. Transactions of Beijing Institute of Technology, 2023, 43(8): 820
- [5] 朱冰, 蒋渊德, 赵健, 等. 基于深度强化学习的车辆跟驰控制[J]. 中国公路学报, 2019, 32(6): 53
ZHU Bing, JIANG Yuande, ZHAO Jian, et al. A car-following control algorithm based on deep reinforcement learning[J]. China Journal of Highway and Transport, 2019, 32(6): 53
- [6] 高振海, 闫相同, 高菲. 基于逆向强化学习的纵向自动驾驶决策方法[J]. 汽车工程, 2022, 44(7): 969
GAO Zhenhai, YAN Xiangtong, GAO Fei. A decision-making method for longitudinal autonomous driving based on inverse reinforcement learning[J]. Automotive Engineering, 2022, 44(7): 969
- [7] 李伟东, 马草原, 史浩, 等. 基于分层强化学习的自动驾驶决策控制算法[J/OL]. 吉林大学学报(工学版). (2023-12-19). <https://doi.org/10.13229/j.cnki.jdxbgxb.20230891>
LI Weidong, MA Caoyuan, SHI Hao, et al. An automatic driving decision control algorithm based on hierarchical reinforcement learning [J/OL]. Journal of Jilin University (Engineering and Technology Edition). (2023-12-19). <https://doi.org/10.13229/j.cnki.jdxbgxb.20230891>
- [8] 刘卫国, 项志宇, 刘伟平, 等. 基于分布式强化学习的车辆控制算法研究[J]. 汽车工程, 2023, 45(9): 1637
LIU Weigu, XIANG Zhiyu, LIU Weiping, et al. Research on vehicle control algorithm based on distributed reinforcement learning [J]. Automotive Engineering, 2023, 45(9): 1637
- [9] WANG Yunpeng, ZHENG Kunxian, TIAN Daxin, et al. Pre-training with asynchronous supervised learning for reinforcement learning based autonomous driving [J]. Frontiers of Information Technology Electronic Engineering, 2021, 22(5): 673
- [10] 宋晓琳, 盛鑫, 曹昊天, 等. 基于模仿学习和强化学习的智能车辆换道行为决策[J]. 汽车工程, 2021, 43(1): 59
SONG Xiaolin, SHENG Xin, CAO Haotian, et al. Lane-change behavior decision-making of intelligent vehicle based on imitation learning and reinforcement learning [J]. Automotive Engineering, 2021, 43(1): 59
- [11] 张志勇, 黄大洋, 黄彩霞, 等. TD3 算法改进与自动驾驶汽车并道策略学习[J]. 机械工程学报, 2023, 59(8): 224
ZHANG Zhiyong, HUANG Dayang, HUANG Caixia, et al. TD3 algorithm improving and lane-merging strategy learning for autonomous vehicles [J]. Journal of Mechanical Engineering, 2023, 59(8): 224
- [12] 杨顺, 蒋渊德, 吴坚, 等. 基于多类型传感数据的自动驾驶深度强化学习方法[J]. 吉林大学学报(工学版), 2019, 49(4): 1026
YANG Shun, JIANG Yuande, WU Jian, et al. Autonomous driving policy learning based on deep reinforcement learning and multi-type sensor data [J]. Journal of Jilin University (Engineering and Technology Edition), 2019, 49(4): 1026
- [13] 张雪峰, 王照乙. 基于双决斗深度 Q 网络的自动换道决策模型[J]. 东北大学学报(自然科学版), 2023, 44(10): 1369
ZHANG Xuefeng, WANG Zhaoyi. Automatic lane change decision model based on dueling double deep Q-network [J]. Journal of Northeastern University (Natural Science), 2023, 44(10): 1369
- [14] 谷梦路, 葛振振, 王畅, 等. 智能网联车辆加速车道类人化汇入控制研究[J]. 中国公路学报, 2024, 37(3): 134
GU Menglu, GE Zhenzhen, WANG Chang, et al. Human like merging control of intelligent connected vehicles on the acceleration lane[J]. China Journal of Highway and Transport, 2024, 37(3): 134
- [15] 姚福星, 孙超, 兰云港, 等. 基于混合专家模型的智能网联汽车换道决策方法[J]. 汽车工程, 2024, 46(5): 882
YAO Fuxing, SUN Chao, LAN Yungang, et al. A lane change decision method for intelligent connected vehicles based on mixture of expert model[J]. Automotive Engineering, 2024, 46(5): 882
- [16] 张羽翔, 何钢磊, 李鑫, 等. 基于参数描述的换道场景自动驾驶精确决策学习[J]. 同济大学学报(自然科学版), 2021,

- 49(增刊1): 132
- ZHANG Yuxiang, HE Ganglei, LI Xin, et al. Precise decision-making learning for automated vehicles in lane-change scenario based on parameter description[J]. Journal of Tongji University (Natural Science), 2021, 49(Sup. 1): 132
- [17] MIRCHEVSKA B, PEK C, WERLING M, et al. High-level decision making for safe and reasonable autonomous lanechanging using reinforcement learning [C]//The 21st International Conference on Intelligent Transportation Systems (ITSC). Maui: IEEE, 2018: 2156
- [18] LI G F, YANG Y F, LI S, et al. Decision making of autonomous vehicles in lane change scenarios: deep reinforcement learning approaches with risk awareness[J]. Transportation Research Part C: Emerging Technologies, 2022, 134: 103
- [19] WANG S, FUJII H, YOSHIMURA S. Generating merging strategies for connected autonomous vehicles based on spatiotemporal information extraction module and deep reinforcement learning[J]. Physica A: Statistical Mechanics and Its Applications, 2022, 607: 128
- [20] 罗鹏, 黄珍, 秦易晋, 等. 基于DQN的车辆驾驶行为决策方法[J]. 交通信息与安全, 2020, 38(5): 67
- LUO Peng, HUANG Zhen, QIN Yijin, et al. A method of vehicle driving behavior decision based on DQN algorithm[J]. Journal of Transport Information and Safety, 2020, 38(5): 67
- [21] 戢晓峰, 徐迎豪, 郝京京, 等. 双车道公路超车行为风险演化时间协变量建模[J/OL]. 西南交通大学学报. (2024-03-26). <https://kns.cnki.net/webvpn.nsf/edu.cn/kcms/detail/51.1277.U.20240322.1638.002.html>
- JI Xiaofeng, XU Yinghao, HAO Jingjing, et al. Time covariate modeling of overtaking behavior risk evolution on two-lane highway [J/OL]. Journal of Southwest Jiaotong University. (2024-03-26). <https://kns.cnki.net/webvpn.nsf/edu.cn/kcms/detail/51.1277.U.20240322.1638.002.html>
- [22] 王荣本, 游峰, 崔高健, 等. 车辆安全换道分析[J]. 吉林大学学报(工学版), 2005, 35(2): 179
- WANG Rongben, YOU Feng, CUI Gaojian, et al. Analysis on lane changing safety of vehicle [J]. Journal of Jilin University (Engineering and Technology Edition), 2005, 35(2): 179
- [23] KOMOROWSKI T, PESZAT S, SZAREK T. On ergodicity of some Markov processes[J]. Annals of Probability, 2010, 38(4): 1401
- [24] ARNS M, BUCHHOLZ P, PANCHENKO A. On the numerical analysis of inhomogeneous continuous-time Markov chains [J]. Inform Journal on Computing, 2010, 22(3): 416
- [25] 关羽, 穆岩, 杨小宝. 微观交通仿真中换道模型的研究综述[J]. 交通与计算机, 2008, 26(5): 66
- GUAN Yu, MU Yan, YANG Xiaobao. Review of lane-changing model in microscopic traffic simulation research[J]. Computer and Communications, 2008, 26(5): 66
- [26] KESTING A, TREIBER M, HELBIG D. Enhanced intelligent driver model to access the impact of driving strategies on traffic capacity[J]. Philosophical Transactions of the Royal Society A, 2010, 368(1928): 4585
- [27] TREIBER M, HENNECKE A, HELBIG D. Congested traffic states in empirical observations and microscopic simulations[J]. Physical Review E, 2000, 62(2): 1805
- (上接第97页)
- [18] DIENG L, MARCHAND P, GOMES F, et al. Use of UHPFRC overlay to reduce stresses in orthotropic steel decks[J]. Journal of Constructional Steel Research, 2013, 89: 30
- [19] 王占飞, 程浩波, 程志彬, 等. 桥面铺装对正交异性钢桥面板疲劳性能的影响[J]. 沈阳建筑大学学报(自然科学版), 2018, 34(2): 257
- WANG Zhanfei, CHENG Haobo, CHENG Zhibin, et al. Influence of bridge deck pavement on the fatigue performance of orthogonally opposite-sex steel bridge [J]. Journal of Shenyang Jianzhu University(Natural Science Edition), 2018, 34(2): 257
- [20] 罗军, 邵旭东, 曹君辉, 等. 钢-超高性能混凝土组合板开裂荷载正交试验及计算方法[J]. 浙江大学学报(工学版), 2020, 54(5): 909
- LUO Jun, SHAO Xudong, CAO Junhui, et al. Orthogonal test and calculation method of cracking load of steel-ultra-high performance concrete composite specimen [J]. Journal of Zhejiang University (Engineering Science), 2020, 54(5): 909
- [21] 叶华文, 段智超, 刘吉林, 等. 正交异性钢-混组合桥面的轮载扩散效应[J]. 吉林大学学报(工学版), 2022, 52(8): 1808
- YE Huawen, DUAN Zhichao, LIU Jilin, et al. Wheel-load diffusion effect on orthotropic steel-concrete composite bridge deck [J]. Journal of Jilin University (Engineering and Technology Edition), 2022, 52(8): 1808
- [22] 黄义, 何芳社. 弹性地基上的梁、板、壳[M]. 北京: 科学出版社, 2005
- HUANG Yi, HE Fangshe. Beams, plates, and shells on elastic foundations[M]. Beijing: Science Press, 2005
- [23] 中交公路规划设计院有限公司. 公路钢结构桥梁设计规范: JTG D64—2015[S]. 北京: 人民交通出版社, 2015
- China Communications Highway Planning and Design Institute Co., Ltd.. Code for design of highway steel structure bridges: JTG D64—2015[S]. Beijing: China Communications Press, 2015
- [24] YE Huawen, WANG Wenchao, HUANG Ao, et al. Reinforcing effect of CFRP bar on concrete splitting behavior of headed stud shear connectors [J]. Steel and Composite Structures, 2023, 48(2): 131
- [25] 赵明华, 张玲, 刘敦平. 考虑摩阻效应的弹性地基梁幂级数解[J]. 铁道学报, 2010, 32(6): 72
- ZHAO Minghua, ZHANG Lin, LIU Dunping. Power-series analytic solutions for the finite beam on elastic foundation with consideration of interface resistance[J]. Journal of the China Railway Society, 2010, 32(6): 72