

基于深度强化学习的自动驾驶车辆 专用道汇入引导

张 健^{1,2,3}, 李青扬^{1,3}, 李 丹², 姜 夏^{1,3}, 雷艳红², 季亚平^{1,3}

(1. 东南大学江苏省城市智能交通重点实验室, 南京 211189; 2. 西藏大学工学院, 拉萨 850013; 3. 东南大学交通学院, 南京 211189)

摘 要:为满足自动驾驶车辆(CAV)与人工驾驶车辆混行过程中安全和效率的需求,自动驾驶车辆专用道应运而生。当高速公路内侧车道设为自动驾驶车辆专用道时,引导自动驾驶车辆从普通车道汇入至专用道的策略研究具有重要的理论意义和实际价值。首先,设计专用道入口并提出车辆控制规则;其次,以使更多自动驾驶车辆换道至专用道为目标,基于深度强化学习,选择换道信号动作引导车辆换道;最后,通过Python语言编译进行数值仿真验证。结果表明:在自动驾驶车辆渗透率、到达专用道自动驾驶车辆比例等不同因素构建的9种场景下,本文算法能够快速收敛;能够有效引导自动驾驶车辆汇入专用道,保证通行效率;相较无信号控制情况,渗透率为20%~40%时,第2车道交通拥堵显著减少;在两段式专用道入口场景下,CAV换道至专用道的比例比单入口场景明显提高。所提出的策略具有较好的适用性,能为工程建设提供参考借鉴。

关键词:交通运输系统工程;汇入引导策略;深度强化学习;高速公路专用道;信号控制;自动驾驶汽车

中图分类号:U491 文献标志码:A 文章编号:1671-5497(2023)09-2508-11

DOI:10.13229/j.cnki.jdxbgxb.20220106

Merging guidance of exclusive lanes for connected and autonomous vehicles based on deep reinforcement learning

ZHANG Jian^{1,2,3}, LI Qing-yang^{1,3}, LI Dan², JIANG Xia^{1,3}, LEI Yan-hong², JI Ya-ping^{1,3}

(1. Jiangsu Key Laboratory of Urban ITS, Southeast University, Nanjing 211189, China; 2. Institute of Technology, Tibet University, Lhasa 850013, China; 3. School of Transportation, Southeast University, Nanjing 211189, China)

Abstract: Exclusive lanes for connected and autonomous vehicles (CAVs) will emerge in order to ensure the safety and efficiency requirements in the process of traffic flow mixed with human-driving vehicles and CAVs. When the inner lane of the expressway is set as the exclusive lane for CAVs, it has important theoretical significance and practical value to study the strategy of guiding CAVs to merge from the

收稿日期:2022-02-02.

基金项目:国家重点研发计划项目(2021YFB1600500);江苏省重点研发计划项目(BE2020013);工信部项目(CEI-EC-2020-ZM02-0100);中国交通建设集团科技研发项目(2019-ZJKJ-ZDZX02-2).

作者简介:张健(1984-),男,教授,博士.研究方向:智能交通运输系统,智慧高速公路,车路协同自动驾驶.

E-mail:zhangjian8seu@163.com

ordinary lane to the exclusive lane. Firstly, the entrance area of exclusive lane was designed and vehicle control rules were proposed. Secondly, with the goal of making more CAVs change lanes to the exclusive lane, the strategy of selecting lane-changing signal actions was proposed based on deep reinforcement learning. Finally, the numerical simulation was carried out with Python language compilation. The results show that the proposed algorithm can converge very quickly under 9 scenarios constructed by different factors, such as the CAV penetration rates and the proportion of CAVs arriving at the exclusive lane; it can effectively guide CAVs to merge into exclusive lanes and ensure traffic efficiency; congestion in the second lane can be significantly reduced compared to the unsignalized control when the penetration rate changes from 20% to 40%; the proportion of CAVs changing to the exclusive lane is significantly higher under the two exclusive lane entrances scenario than that under the one entrance scenario. It shows that the proposed strategy has good applicability and can provide reference for engineering construction.

Key words: engineering of transportation system; merging guidance strategy; deep reinforcement learning; expressway exclusive lane; signal control; connected and autonomous vehicle

0 引言

随着物联网、大数据、人工智能、新一代信息通信等高新技术迅速发展,以自动驾驶为主要特征的新一代智能交通系统成为解决交通问题的新途径。自动驾驶技术发展需要持续迭代更新,势必经历一个漫长的自动驾驶车辆与人工驾驶车辆混合行驶的过渡阶段^[1-3]。结合中国实际,目前实现自动驾驶主要依赖车路协同技术,将技术和成本在车侧和路侧合理分配,以实现车路协同感知、决策、控制一体化^[4-6]。

设置自动驾驶车辆(Connected and autonomous vehicle, CAV)专用道能够为自动驾驶车辆创造更为纯粹的通行环境,进一步减少延误和协同自适应巡航控制系统车辆(Cooperative adaptive cruise control, CACC)退化率、优化安全行车距离,有效提升道路实际通行能力。结合路侧设施,基于CAV专用道的车道管理策略将有效改善混合车流的安全性问题,设置自动驾驶专用车道可以提高通行能力,进一步加快自动驾驶技术应用。高速公路较城市道路场景更简单,因此被认为适合早期的自动驾驶技术发展,更易设置自动驾驶车专用道^[7]。现有CAV专用道的研究包括路网级别和路段级别。路段级别专用道的设置受到CAV渗透率、交通密度、车流的组队强度等因素影响^[8,9],对比相关研究发现^[10-12],高速场景设置一条专用道的条件是CAV渗透率为30%~50%,设置两条专用道的条件是渗透率大于80%。由于CAV专用道还未得到实际应用,所以当前研究均立足于仿真。

研究CAV如何由普通车道汇入专用道,以提高换道安全和效率,保障CAV行驶优先权,具有重要意义和价值,但目前欠缺较为深入的研究,交通信号管控为这一场景提供了可借鉴方案。现有高速公路信号控制主要集中在匝道,针对单匝道控制并已大规模应用的方法包括需求-容量差控制、占有率控制、高速公路入口线性控制等,这些方法基于精确模型,能够解决常发性的拥堵问题,但对于偶发性拥堵预测不准确^[13,14]。一些智能方法(如模糊逻辑控制方法、神经网络、深度强化学习等)可以满足匝道管控场景复杂、非线性以及时变的要求^[15-18]。深度强化学习(Deep Q-learning, DQN)通过与环境交互生成优化控制策略。目前,针对DQN、状态-动作-奖励-状态-动作(State-action-reward-state-action, SARSA)、深度策略梯度(Deep deterministic policy gradient, DDPG)等算法的研究较多^[19-21]。本文提出了基于DQN算法的CAV专用道车辆汇入引导策略。首先,针对CAV专用道入口进行交通安全设计(下文简称交安设计)并制定了车辆控制规则;然后,基于DQN算法提出了换道信号动作选择汇入引导策略,利用信号引导CAV换道至专用道;最后,通过搭建数值仿真平台对本文模型进行仿真实验和分析,以验证汇入引导策略的效果。

1 专用道入口交安设计及车辆控制

1.1 专用道入口的设置

本文研究场景为高速公路单向4车道,行车道沿道路前进方向从左向右依次编号为第1~4

车道,第1车道为CAV专用道,其余为普通车道。专用道仅供CAV行驶,普通车道支持CAV和人工驾驶车辆混行。道路标线设计符合《道路交通标志和标线第3部分:道路交通标线》(GB 5768.3—2009)。专用道与普通车道间为白色实线,表示不可变更车道;普通车道间为白色虚线,表示可变更车道。

为研究自动驾驶车辆换道至专用道的过程,针对专用道入口进行交安设计,如图1所示。借鉴高速公路作业区起点的设置方法,在路侧部署专用道入口距离标志、禁止超车标志等,标志设计

符合《道路交通标志和标线第2部分:道路交通标志》(GB 5768.2—2009)。考虑车路协同场景,部署路侧单元(Road side unit,RSU),RSU含有通信模块、边缘计算模块和车辆状态监测系统,能够实现车路通信、提供主动管控决策、监测一定范围内车辆的位置和速度。RSU的部署需综合考虑布设成本和网络的连通性等因素,在混合交通流场景下,以通信范围为400 m为标准,设置RSU的平均布设间隔为1200 m^[22]。为了保证能够全面监测专用道入口处的车辆状态,在RSU间增设一套电视设备。

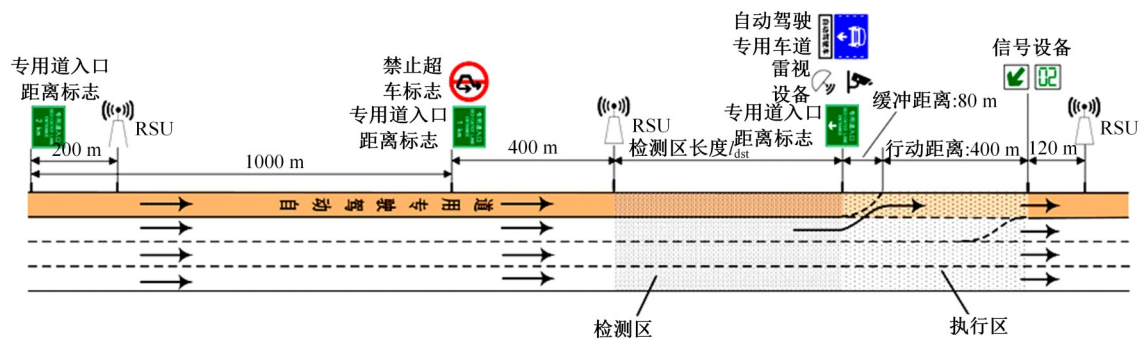


图1 专用道入口示意图

Fig. 1 Schematic diagram of exclusive lane entrance

设计长度为 l_{det} 的检测区和长度为 l_{act} 的执行区。检测区处监测车辆位置和速度信息,以便支持换道信号控制方案的生成。执行区处自动驾驶车辆受信号控制并接受车路(Vehicle to infrastructure, V2I)通信下发的是否可换道指令。在执行区末端设立龙门架,横梁上安装信号灯并显示倒计时。当信号为绿灯时,CAV可换道至内侧专用道;当信号为红灯时,禁止CAV换道至专用道。执行区的长度 l_{act} 为缓冲距离 l_h 与行动距离 l_a 之和,缓冲距离提供给车辆调整其行车状态的空间,行动距离满足车辆换道所需的距离,如下所示^[23]:

$$l_{act} = l_a + l_h = n \frac{v_0}{3.6} t_0 + l_h \quad (1)$$

式中: n 为变换车道的次数,考虑第4车道车辆换道至第1车道,取值为3; v_0 为车辆到达执行区的速度,取值为80 km/h; t_0 为变换一次车道所需的时间,约为6 s^[24]; l_h 取值为80 m^[25]。

1.2 车辆控制规则

车辆控制包括跟驰和换道控制,根据车辆自动驾驶等级不同,人工驾驶车辆的跟驰采用元胞自动机(Cellular automata, CA)模型进行控制。

当CAV前方为人工驾驶车辆,其跟驰行为由自适应巡航控制(Adaptive cruise control, ACC)系统模型控制;当前方为CAV,其跟驰行为由协同式自适应巡航(Collaborative adaptive cruise control, CACC)系统模型控制^[26]。所有车辆换道行为由CA模型控制。

1.2.1 KKW元胞自动机模型

CA模型被广泛应用于交通流仿真,可用于模拟车辆的跟驰和换道行为。模型包括元胞、元胞空间、元胞状态演化规则和领域4个基本组成单元。元胞是最基本的组成要素,道路路段按照时间和空间可划分成长度为等长元胞,长度为 l_{cell} 。元胞属性状态随时间发生变化,变化范围为二进制形式 $\{0, 1\}$,其中0代表元胞无车辆占用,1代表有车辆占用。

三相交通流模型(Kemer-Klenov-Wolf, KKW)作为CA模型中的一种,能较好地反映交通流的时间和空间复杂性,模拟同步流及瓶颈处交通拥挤模式,通过车辆同步间距调整后车速度以达到速度自适应。其更新规则分为确定性和随机性两部分^[27],车辆确定性部分的规则为:

$$\tilde{v}_n(t+1) = \max(0, \min(v_{\max}, v_{\max,n}(t), v_{\exp}(t))) \quad (2)$$

式中: $\tilde{v}_n(t+1)$ 为确定性规则得到的下一时刻速度; v_{\max} 为车辆最大限速; $v_{\max,n}(t)$ 为车辆能达到的最大速度; $v_{\exp}(t)$ 为车辆期望速度。

随机更新部分的规则为:

$$v_n(t+1) = \max(0, \min(\tilde{v}_n(t+1) + a\tau\varphi_n, v_n(t) + a\tau, v_{\max}, v_{\max,n}(t))) \quad (3)$$

式中: $v_n(t+1)$ 为随机更新规则得到的车辆下一时刻的速度; a 为车辆加速度; φ_n 为随机因素; τ 为仿真演化的时间步长。

1.2.2 ACC和CACC模型

本文采取的ACC和CACC模型被证明能模拟车辆各类行为,使车流具有较好的稳定性。根据道路上前、后车辆间的车头时距以及速度差的关系,模型分为巡航、接近、跟驰3类情况,如下所示。

(1)巡航模式。此模式在前、后车辆车头时距大于两倍期望车头时距时被激活,此时后车的加速度计算公式为:

$$a_n(t+1) = k \cdot (v_{\max} - v_n(t)) \quad (4)$$

式中: k 为常数项,取值为 0.4 s^{-1} ; $v_n(t)$ 为目标车辆速度。

(2)跟驰模式。在前、后车辆的距离误差 $e_n(t) < 0.2 \text{ m}$ 或速度误差 $v_{n-1}(t) - v_n(t) < 0.1 \text{ m/s}$ 时激活。此时,若车辆由ACC模型控制,则其加速度计算公式为:

$$a_n(t+1) = k_1 \cdot e_n(t+1) + k_2 \cdot (v_{n-1}(t) - v_n(t)) \quad (5)$$

式中: $k_1 = 0.23 \text{ s}^{-2}$; $k_2 = 0.07 \text{ s}^{-1}$; $e_n(t+1)$ 为前、后车辆下一时刻的距离误差。

若车辆为CACC控制模式,则其速度计算公式为:

$$v_n(t+1) = v_n(t) + k_p \cdot e_n(t) + k_d \cdot \dot{e}_n(t) \quad (6)$$

式中: $k_p = 0.45 \text{ s}^{-2}$; $k_d = 0.25 \text{ s}^{-1}$ 。

(3)接近模式。此模式保证车辆从巡航模式平稳转换至跟驰模式,采取与跟驰模式相似的模型,其中 $k_1 = 0.04 \text{ s}^{-2}$, $k_2 = 0.8 \text{ s}^{-1}$; $k_p = 0.01 \text{ s}^{-2}$, $k_d = 1.6 \text{ s}^{-1}$ 。

1.2.3 换道模型

对于道路上的车辆,在两种情况下会做出换道决策。第1种情况,本车道前车阻碍目标车辆通行,相邻车道前车、相邻车道后车与目标车辆之

间满足最小安全距离时,车辆换道。具体规则如下:

$$d_n(t) < \min(v_n(t) + a, v_{\max}) \quad (7)$$

$$d_{l,n}(t) > d_{\text{safe}} \quad (8)$$

$$d_{l,n,b}(t) > d_{\text{safe}} \quad (9)$$

式中: $d_n(t)$ 为目标车辆在 t 时刻与前车的间距; $d_{l,n}(t)$ 为目标车辆在 t 时刻与相邻车道前车的间距; d_{safe} 为车辆间安全间距; $d_{l,n,b}(t)$ 为目标车辆在 t 时刻与相邻车道后车的间距。

第2种情况,本车道前车不阻碍目标车辆通行,但相邻车道前车与目标车辆间距离大于本车道前车与车辆间距离,即 $d_{l,n}(t) > d_n(t)$,且相邻车道后车与目标车辆之间满足最小安全距离,车辆以一定概率 p_{left} 或 p_{right} 换道,换道概率取决于换道方向。换道轨迹采取等速偏移模型,轨迹为直线。CA模型入口边界条件采用开放性边界条件,出口边界不考虑车辆不能离开的情况。

对于普通车道上的CAV,为保证更多车辆换道至专用道,设置CAV在执行区与检测区不向右变道,并可在满足上述第1、2种情况时向左变道。

2 换道信号动作选择的汇入引导策略

通过深度强化学习可解决连续状态空间下混合交通流的决策问题。智能体(Agent)通过获取检测区交通流状态选择换道信号动作并通过V2I通信发送给CAV,控制结果作为奖励值反馈给智能体用于修正参数。本节分为汇入引导策略深度强化学习模型和策略求解两部分。为简化研究场景,作如下假设:

(1)车与路侧设施、车与车间通信无延迟。

(2)场景中的车辆为性能、外观一致的小汽车。

(3)通行车辆完全按照设定规则行驶,道路无突发事件。

(4)CAV接受换道引导。

2.1 汇入引导策略深度强化学习模型

2.1.1 状态空间

为尽可能多地展示交通流信息,对检测区采取离散交通状态编码(Discrete traffic state encoding, DTSE)^[28],将路网网格化处理,离散的网络间隔分别为车辆的长度 l_{len} 和宽度 l_{wid} 。生成的位置矩阵作为状态空间的信息输入 $\phi(S_t)$,其状态空

间的大小为 $1 \times [l_{\text{det}}/l_{\text{len}}] \times [3.75 \times 4/l_{\text{wid}}]$ 。矩阵中每个元素为刻画交通状态的指标,为强化 CAV 在路网中的重要性,每个元胞取值为 $\{0, 0.5, 1\}$ 。以车辆中心点落在元胞表示对该元胞的占用:0 代表元胞无车辆占用;0.5 代表有人工驾驶车辆占用;1 代表有 CAV 占用,如图 2 所示。

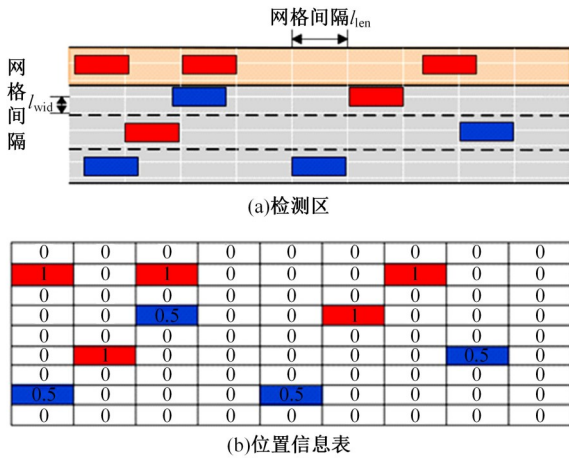


图 2 离散交通状态编码示例
Fig. 2 Schematic of DTSE

2.1.2 动作空间

获取状态空间后,根据 Q 表选取最优动作从而优化换道关键区的开放时间, Q 表的获取如图 3 所示。根据换道关键区的开放与否以及持续时间,将动作空间设置为 $A = \{A_1, A_2\}$, 其中,动作 A_1 表示执行区信号显示为红灯 20 s; 动作 A_2 表示执行区信号显示为绿灯 20 s。Agent 在每个动作实施完成后再选择下一个动作,例如,在某个时刻选择动作 A_2 , 则在执行区绿灯时间 20 s 后再选择下一个动作。

2.1.3 奖励值

奖励函数的设置反映了动作对交通流的影响,为了促进更多的 CAV 换道至专用道,奖励函数设置考虑增加平均绿灯时间内的换道车辆数和降低换道信号开放成本:

$$r = \begin{cases} \varphi, & \text{动作} = A_2 \\ \frac{n_c}{\Delta t_1} - \varphi, & \text{动作} = A_1 \end{cases} \quad (10)$$

式中: φ 为换道信号开放成本,防止绿灯过长时间内无车辆换道; n_c 为绿灯期间成功换道至 CAV 专用道的车辆数; Δt_1 为动作持续时间。

2.1.4 动作选择

在进行动作选择时,可以选择 Q 值更大的动作,这种策略称为利用(Exploitation); 但仅按照 Q

值寻找动作容易陷入局部解,需要在现有状态下进行随机探索,该策略称为探索(Exploration)。为了平衡利用-探索的过程,采用 ϵ -greedy 策略,在每次迭代时产生随机数,若随机数小于 ϵ 则采用随机探索策略,若大于 ϵ 则根据 Q 表选择 Q 值最大的动作。随着迭代次数的增加,神经网络收敛,应该更多按照 Q 表选择动作,因此 ϵ 随迭代次数的增加而减小,如下所示:

$$\epsilon = \epsilon - \frac{\epsilon_{\text{initial}} - \epsilon_{\text{final}}}{10\,000} \quad (11)$$

式中: $\epsilon_{\text{initial}}$ 为初始 ϵ 值; ϵ_{final} 为最终 ϵ 值。

2.2 汇入引导策略求解

策略求解采用 DQN 算法,结合深度学习特征感知能力和强化学习的决策能力,可解决较复杂的问题。Agent 结构采用卷积神经网络(Convolutional neural networks, CNN),包括 2 层卷积层、1 层激励层、1 层池化层和 1 层全连接层,隐藏层激活函数为线性修正单元(Rectified linear units, ReLU)函数,池化层采用最大池化。Agent 算法结构如图 3 所示,输出为 Q 表中的值。

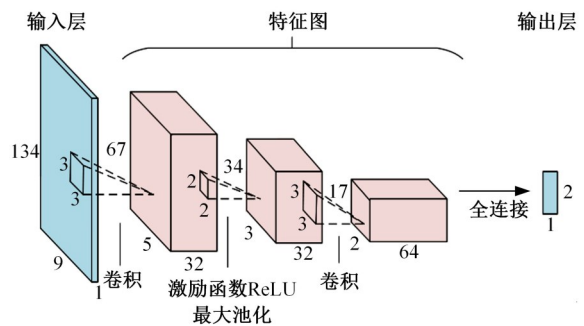


图 3 CNN 算法模型架构

Fig. 3 Architecture of CNN algorithm

在强化学习进行训练时,算法 DQN 以目标 Q 值与实际 Q 值之间的时序差分误差 δ_t 更新神经网络的权重值参数^[28]。智能体训练的算法流程为:首先,初始化样本数为 n_{pool} 的经验回放池 D 和神经网络参数 θ , 将 $\phi(S_t)$ 作为当前值网络的输入,通过 ϵ -greedy 策略选择动作 A_t , 得到下一个交通状态 S_{t+1} 和回报 R_{t+1} , 将 $(S_t, A_t, R_{t+1}, S_{t+1})$ 存储到经验回放池中。在经验回放池中选取 n_{batch} 个样本用于调参,时序差分误差 δ_t 如下所示:

$$\delta_t = y_t^{\text{DQN}} - Q(\phi(S_t), A_t; \theta_t) = R_{t+1} + \gamma \max_A Q(\phi(S_{t+1}), A; \theta') - Q(\phi(S_t), A_t; \theta_t) \quad (12)$$

式中: γ 为折扣率。

进一步采用目标梯度下降法更新神经网络参数 θ 。

经验回放池中不同样本对反向梯度计算的影响不同,其采样时的优先级应有不同,而目前采样则是在经验回放池中均匀采样,忽略了样本的重要程度。因此采取优先级经验回放(Prioritized experience replay, PER)^[29],每个样本采样的概率为 $P(i)=p_i^{\alpha}/\sum_k p_k^{\alpha}$,其中 $p_i=|\delta_i|+\rho_1$, ρ_1 为正数。

优先重放机制引入了偏差,改变了期望值的分布,从而改变了收敛结果,因此引入重要性采样权重修正损失函数 w_j :

$$w_j = \left(\frac{1}{N} \cdot \frac{1}{P(j)} \right)^{\beta} / \max w_i \quad (13)$$

$$\theta \leftarrow \theta + w_j \cdot \delta_j \cdot \nabla_{\theta} Q(\phi(S_j), A_j; \theta) \quad (14)$$

式中: ∇_{θ} 为求出目标函数对策略神经网络参数的梯度,即采用目标梯度下降的方法更新神经网络的参数 θ 。

3 仿真实验及结果分析

3.1 仿真参数设置

为验证本文策略的有效性,基于Python语言编译数值仿真平台,获得模型测试结果。参数设置分为两部分:与车辆控制规则相关的参数设置,如表1所示^[30];与汇入引导策略相关的参数设置,如表2所示^[31]。

表1 车辆控制规则相关参数设置

Table 1 Vehicle control method parameters

参数	l_{cell}	v_{max}	τ	p_{left}	p_{right}	p	p_0	p_{a1}	p_{a2}
取值	1.5	30	1	0.4	0.3	0.04	0.425	0.2	0.052

表2 汇入引导策略相关的参数设置

Table 2 Merging guiding strategy parameters

参数	l_{len}	l_{wid}	$\epsilon_{\text{initial}}$	ϵ_{final}	γ	n_{pool}	n_{batch}	φ
取值	4.5	1.8	0.5	0.01	0.9	5000	64	0.045

考虑行车视距和换道轨迹,检测区长度 l_{det} 设为600 m;执行区总长度 l_{act} 设为480 m。为保证训练数据的一致性,随机生成车流数据统一进行训练,车辆到达率为4000 veh/h,设CAV的市场渗透率为 P ,到达专用道的CAV占CAV总数的比例为 ρ 。根据高速公路4车道设置1条专用道的CAV渗透率条件为30%~50%,过高的渗透率会导致普通车道利用率下降,从而导致专用道交通需求超过其通行能力,因此分别设置 P 为

20%、30%、40%, ρ 为20%、40%、60%的情况进行研究。每轮仿真进行200回合迭代,每回合仿真时间为3000 s。前200 s作为预热,车辆根据1.2节规则运行,不纳入实验结果。

3.2 仿真结果分析

3.2.1 模型有效性验证

为验证本文策略的有效性,在不同CAV渗透率 P 和占比 ρ 条件下对算法模型进行仿真,分析每回合平均奖励值的变化,如图4所示(图中每个点代表每一回合迭代的奖励平均值)。分析可得,在200轮训练过程中各场景平均奖励值变化总体呈现上升趋势,且前期的平均奖励值跳跃明显,后期震荡减少,逐渐趋于平缓,智能体在各场景能有效收敛,采用深度学习DQN算法能较快找到优化的引导策略。奖励值与单位时间内成功换道的CAV数有关,在总体流量相同情况下,CAV渗透率越高,到达专用道CAV占比越低,回合平均奖励值越高。

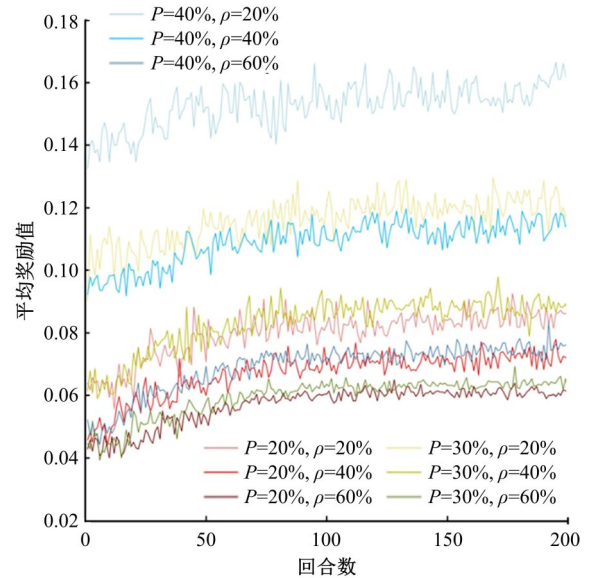


图4 不同场景下每回合平均奖励值变化情况

Fig. 4 Change of episode average reward under different scenarios

3.2.2 交通控制效果分析

研究不同条件对CAV换道至专用道的影响,绘制第200回合第200~3000 s所有车辆的时空轨迹图,如图5所示。图5(a)为到达率为4000 veh/h、 $\rho=40\%$ 、渗透率 P 分别为20%、30%、40%时车辆的时空轨迹。CAV在执行区换道至专用道,因而在图中横坐标约1080 m之后,专用道上增加了大量车辆轨迹,第2车道车辆轨迹明

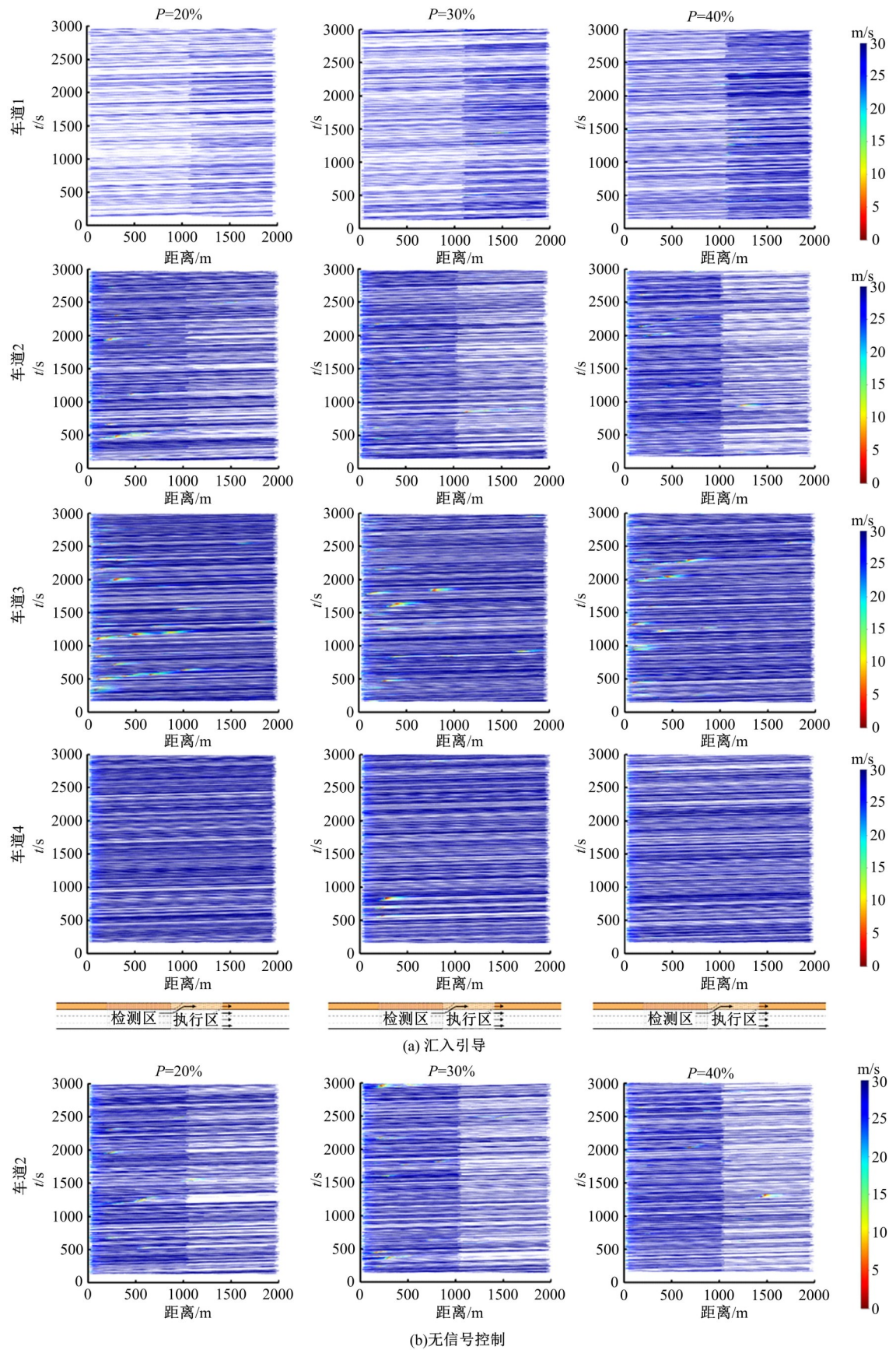


图 5 不同 CAV 渗透率下的车辆时空轨迹图

Fig. 5 Temporal-spatial pattern of traffic speed at various CAV penetration rates

显减少,也呈现了这一换道过程。对于CAV专用道,当CAV渗透率在20%~40%内增大时,车道的占有率增大,车辆汇入导致拥堵增加。针对第2车道,随着渗透率的增大,大量CAV有换道的意愿,但专用道已有较多车辆占用,因此换道间隙的选择更为困难,导致在执行区更易发生拥堵。对于第3、4车道,随着渗透率的增大,拥堵没有发生明显消散,这源于CAV在换道信号引导下有汇入专用道的意愿,导致交通流不稳定,同时CAV车辆间的最小车头时距更小,使拥堵能够较快消散,综述可见拥堵并未有明显变化。因此,针对换道至专用道的场景,CAV高渗透率可能导致交通流产生更多拥堵。图5(b)为无信号控制情况下CAV可自由换道至专用道时,第2车道上车辆的时空轨迹,对比图5(a)中第2车道的车辆轨迹图,采用本文的汇入引导策略可以有效减少第2车道的拥堵,使拥堵快速消散。

图6为每回合仿真结束后,换道至专用道的CAV占到达普通车道CAV的比例,箱型图展现了第181~200回合的统计情况,图中点代表无信号控制情况下换道至专用道的CAV占到达普通车道CAV的比例。由此可知,到达专用道的CAV占CAV总数的比例 ρ 越小,专用道容许更多车辆换道,换道成功的CAV比例越高。对比无信号控制,本文模型换道成功的车辆百分比更低,结合图5(b)可知,本文信号控制下第2车道通过减少部分低效的CAV换道以减少车道拥堵。

表3为不同场景下的道路通行效率,由表3可知,采用本文模型对比无信号控制情况,通行效率无明显变化,因为无信号控制情况下能使更多CAV换道至专用道,进而CAV能够快速行驶,但同时会造成车道更拥堵。总而言之,采用汇入引导策略能够减少交通拥堵、保证行车安全。

3.2.3 两段式CAV专用道入口场景仿真

根据仿真实验,在设置单个CAV专用道入口时,仅有部分CAV在执行区换道至专用道。为提高CAV换道至专用道的数量,考虑增加专用道入口,对两段式CAV专用道入口进行研究。仿真路

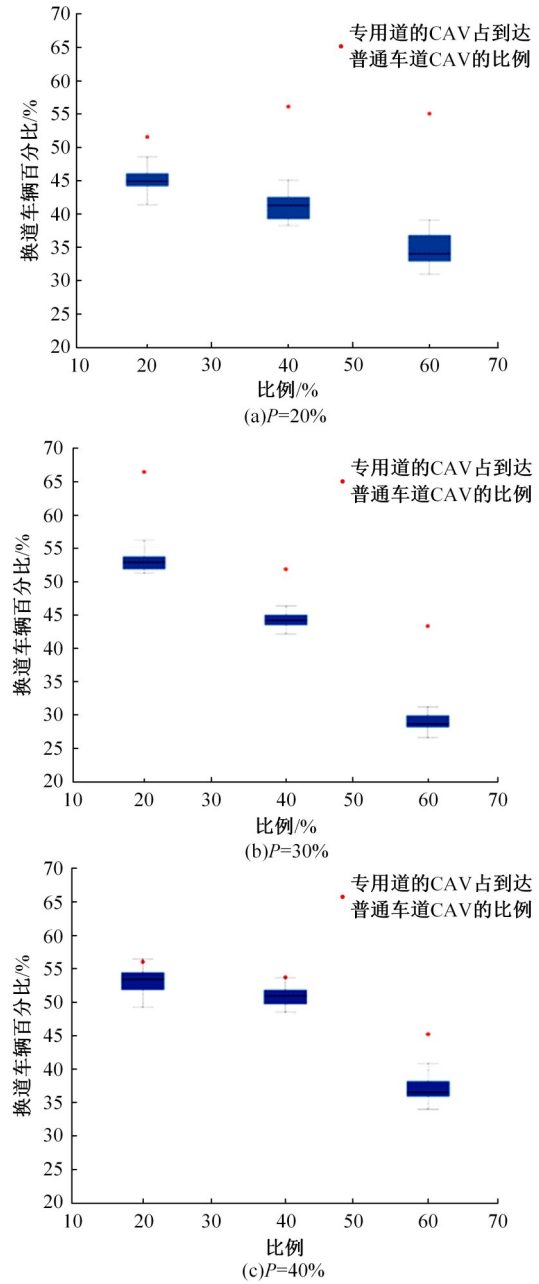


图6 不同场景下换道车辆百分比变化情况

Fig. 6 Change of episode percentage of lane-changing vehicles under different scenarios

段如图7所示,路段长度为3.6 km,共设两个专用道入口,分别称为专用道入口1和专用道入口2,皆由本文所提出的汇入引导控制策略控制。交通需求采取前文到达率为4000 veh/h、 $P=30\%$ 、

表3 不同场景下的通行效率

Table 3 Capacity under different scenarios

veh/h

方法	P=20%			P=30%			P=40%		
	$\rho=20\%$	$\rho=40\%$	$\rho=60\%$	$\rho=20\%$	$\rho=40\%$	$\rho=60\%$	$\rho=20\%$	$\rho=40\%$	$\rho=60\%$
本文模型	3301	3426	3303	3293	3282	3377	3335	3301	3426
无信号控制	3323	3425	3303	2736	3273	3377	3339	3323	3425

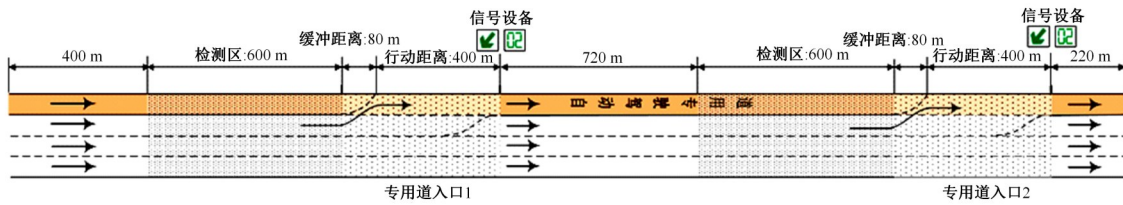


图7 两段式 CAV 专用道入口场景示意图

Fig. 7 Schematic diagram of two CAV exclusive lane entrances

$\rho = 20\%$ 的仿真情况,仿真进行 200 回合迭代,每回合仿真时间为 3000 s,前 200 s 作为路网预热。算法在第 200 回合时已收敛,取第 200 回合的实验情况,得到车辆时空图如图 8 所示。

根据图 8 中车道 1 和车道 2 的车辆时空轨迹,可发现在两段式专用道入口场景下,车道 2 中 CAV 在两处入口附近换道至专用道,换道时在车道 2 的入口附近产生轻微拥堵,但对专用道原行驶的 CAV 无明显影响,CAV 在专用道入口附近的换道行为满足安全和高效的要求。

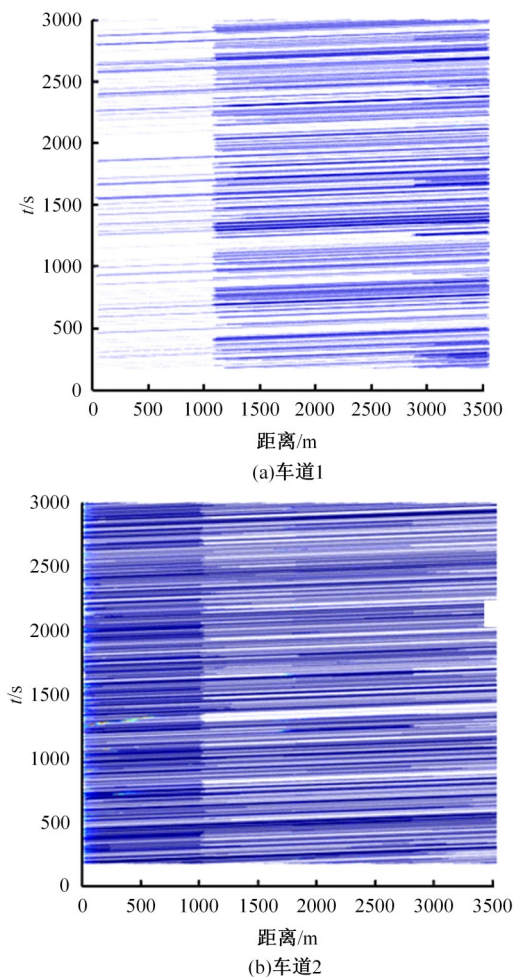


图 8 第 200 回合车辆时空轨迹图

Fig. 8 Temporal-spatial pattern of traffic speed in the 200th episode

在仿真第 200 回合中,普通车道 CAV 在各专用道入口换道至专用道的比例变化如图 9 所示。由图 9 可以看出:专用道入口 1 换道成功数量占到达普通车道 CAV 比例的 53.64%;专用道入口 2 换道成功数量占到达普通车道 CAV 比例的 14.02%,可见通过设置两段式专用道入口,CAV 的换道成功率明显提高,且车辆主要集中在专用道入口 1 处进行换道。多段式专用道入口的设置能够有效提高专用车道的利用率。

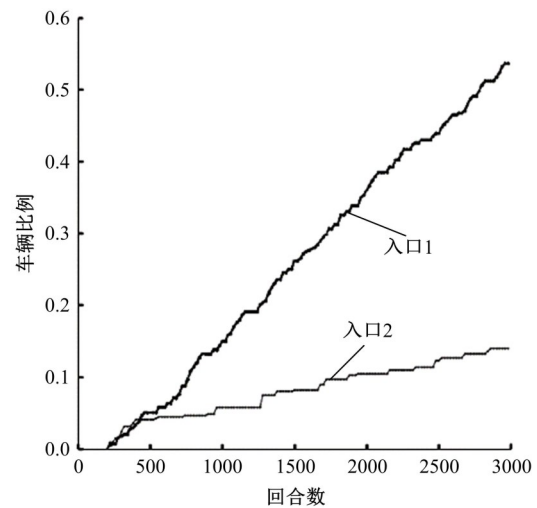


图 9 第 200 回合专用道入口换道成功车辆比例变化情况
Fig. 9 Changes of percentage of vehicles successfully changing lanes at exclusive lane entrances in the 200th episode

4 结束语

针对目前自动驾驶车由普通车道换道至专用道这一过程研究存在的不足,本文基于深度强化学习,以高速公路单向 4 车道场景为例,在行车方向最左侧车道设置 CAV 专用道,进行专用道入口交安设计并提出了车辆控制规则,利用 DQN 算法选择换道信号动作,实现 CAV 的换道汇入引导。最后,基于 Python 语言开发模型数值仿真平台并进行仿真验证,结果表明,采用本文策略能在引导 CAV 换道的情况下有效减少并缓解拥堵,保

证通行效率,同时设置两段式专用道入口的场景可以进一步提高CAV的换道成功率。

参考文献:

- [1] 秦严严,王昊,王炜. 智能网联环境下的混合交通流LWR模型[J]. 中国公路学报, 2018, 31(11): 147-156.
Qin Yan-yan, Wang Hao, Wang Wei. LWR model for mixed traffic flow in connected and autonomous vehicular environments[J]. China Journal of Highway and Transport, 2018, 31(11): 147-156.
- [2] Becker F, Axhausen K W. Literature review on surveys investigating the acceptance of automated vehicles[J]. Transportation, 2017, 44(6): 1293-1306.
- [3] Elliott D, Keen W, Miao L. Recent advances in connected and automated vehicles[J]. Journal of Traffic and Transportation Engineering (English Edition), 2019, 6(2): 109-131.
- [4] 杜豫川,刘成龙,吴荻非,等. 新一代智慧高速公路系统架构设计[J]. 中国公路学报, 2022, 35(4): 203-214.
Du Yu-chuan, Liu Cheng-long, Wu Di-fei, et al. Framework of the new generation of smart highway[J]. China Journal of Highway and Transport, 2022, 35(4): 203-214.
- [5] Shladover S E. Connected and automated vehicle systems: introduction and overview[J]. Journal of Intelligent Transportation Systems, 2017, 22(3): 190-199.
- [6] 冉斌,谭华春,张健,等. 智能网联交通技术发展现状及趋势[J]. 汽车安全与节能学报, 2018, 9(2): 119-130.
Ran Bin, Tan Hua-chun, Zhang Jian, et al. Development status and trend of connected automated vehicle highway system[J]. Journal of Automotive Safety and Energy, 2018, 9(2): 119-130.
- [7] Ma K, Wang H. Influence of exclusive lanes for connected and autonomous vehicles on freeway traffic flow[J]. IEEE Access, 2019, 7: 50168-50178.
- [8] Hua X D, Yu W J, Wang W. Influence of lane policies on freeway traffic mixed with manual and connected and autonomous vehicles[J]. Journal of Advanced Transportation, 2020, 2020: No. 3968625.
- [9] Ghiasi A, Hussain O, Qian Z. A mixed traffic capacity analysis and lane management model for connected automated vehicles: a markov chain method[J]. Transportation Research Part B, 2017, 106: 266-292.
- [10] Chen D J, Ahn S, Chitturi M. Towards vehicle automation: roadway capacity formulation for traffic mixed with regular and automated vehicles[J]. Transportation Research Part B, 2017, 100: 196-221.
- [11] Ye L, Yamamoto T. Impact of dedicated lanes for connected and autonomous vehicle on traffic flow throughput[J]. Physica A, 2018, 512: 588-597.
- [12] Ye L, Yamamoto T. Modeling connected and autonomous vehicles in heterogeneous traffic flow[J]. Physica A, 2018, 490: 269-277.
- [13] Masher D P, Ross D W, Wong P J, et al. Guidelines for design and operation of ramp control systems[R]. Menlo Park: CA United States: Stanford Research Institute, 1975.
- [14] Smaragdis E, Papageogiou M. Series of new local ramp metering strategies[J]. Transportation Research Record, 2003, 1856: 74-86.
- [15] Bellemans T, Schutter B D, Moor B D. Model predictive control for ramp metering of motorway traffic: a case study[J]. Control Engineering Practice, 2006, 14(7): 757-767.
- [16] 陈德望,王飞跃,陈龙. 基于模糊神经网络的城市高速公路入口匝道控制算法[J]. 交通运输工程学报, 2003(2): 100-105.
Chen De-wang, Wang Fei-yue, Chen Long. Freeway ramp control algorithm based on neurofuzzy networks[J]. Journal of Traffic and Transportation Engineering, 2003(2): 100-105.
- [17] 曾筠程,邵敏华,孙立军,等. 基于有向图卷积神经网络的交通预测与拥堵管控[J]. 中国公路学报, 2021, 34(12): 239-248.
Zeng Yun-cheng, Shao Min-hua, Sun Li-jun, et al. Traffic prediction and congestion control based on directed graph convolution neural network[J]. China Journal of Highway and Transport, 2021, 34(12): 239-248.
- [18] Zhang J, Jiang X, Liu Z, et al. A study on autonomous intersection management: planning-based strategy improved by convolutional neural network[J]. KSCE Journal of Civil Engineering, 2021, 25(10): 3995-4004.
- [19] Davarynejad M, Hegyi A, Vrancken J, et al. Motorway ramp-metering control with queuing consideration using Q-learning[C]//2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC), Washington, DC, USA, 2011: 1652-1658.
- [20] 郑思. 面向快速道路远距离瓶颈的深度强化学习交通流控制策略研究[D]. 南京:东南大学交通学院, 2021.
Zheng Si. Research on deep reinforcement learning-

- based active traffic flow control strategies at distant downstream bottlenecks of expressway[D]. Nanjing: School of Transportation, Southeast University, 2021.
- [21] Belletti F, Haziza D, Gomes G. Expert level control of ramp metering based on multi-task deep reinforcement learning[J]. *Transactions on Intelligent Transportation Systems*, 2017, 19(4): 1198-1207.
- [22] 葛家丽. 基于车辆簇的高速公路路侧单元部署研究[D]. 济南: 山东科技大学交通学院, 2020.
- Ge Jia-li. Research on freeway road side unit deployment based on vehicle clusters[D]. Jinan: School of Transportation, Shandong University of Science and Technology, 2020.
- [23] 陈瑜. 高速公路作业区安全分析及交通组织管理方法研究[D]. 哈尔滨: 哈尔滨工业大学交通科学与工程学院, 2006.
- Chen Yu. The safety analysis and organization & management method study of freeway work zone[D]. Harbin: School of Transportation Science & Engineering, Harbin Institute of Technology, 2006.
- [24] 石茂清. 道路交通安全设施设计研究[D]. 成都: 西南交通大学交通运输与物流学院, 2005.
- Shi Mao-qing. Study on design of traffic safety facilities[D]. Chengdu: School of Transportation and Logistics, Southwest Jiaotong University, 2005.
- [25] 孙玲, 张静, 周瀛, 等. 车路协同环境下自动驾驶专用车道入口区域设计[J]. *公路交通科技*, 2020, 37(增刊 1): 122-129.
- Sun Ling, Zhang Jing, Zhou Ying, et al. Design of entrance area of automatic driving special lane in vehicle-infrastructure collaborative environment[J]. *Journal of Highway and Transportation Research and Development*, 2020, 37(Sup. 1): 122-129.
- [26] Milanés V, Shladover S E. Modeling cooperative and autonomous adaptive cruise control dynamic responses using experimental data[J]. *Transportation Research Part C: Emerging Technologies*, 2014, 48: 285-300.
- [27] Xue Y, Wang X, Cen B L, et al. Study on fuel consumption in the Kerner-Klenov-Wolf three-phase cellular automaton traffic flow model[J]. *Nonlinear Dynamics*, 2020, 102(1): 1-10.
- [28] Wang C H, Hwang M C. Value-based deep reinforcement learning for adaptive isolated intersection signal control[J]. *IET Intelligent Transport Systems*, 2018, 12(9): 1005-1010.
- [29] Schaul T, Quan J, Antonoglou I, et al. Prioritized experience replay[C]//*Proceeding of the 4th International Conference on Learning Representations*, San Juan, Puerto Rico, 2016: 322-355.
- [30] 武毅. 基于三相交通流理论的元胞自动机模型研究[D]. 长春: 吉林大学交通学院, 2018.
- Wu Yi. Research on cellular automated traffic flow model for three-phase theory[D]. Changchun: College of Transportation, Jilin University, 2018.
- [31] 舒凌洲, 吴佳, 王晨. 基于深度强化学习的城市交通信号控制算法[J]. *计算机应用*, 2019, 39(5): 1495-1499.
- Shu Ling-zhou, Wu Jia, Wang Chen. Urban traffic signal control based on deep reinforcement learning[J]. *Journal of Computer Applications*, 2019, 39(5): 1495-1499.