

DOI:10.11918/202410035

# 考虑 CAV 专用道汇入需求的高速公路 交织区集成控制策略

席殊,陈旭梅,李培坤,马嘉欣

(北京交通大学 交通运输学院,北京 100044)

**摘要:**为提升高速公路通行效率,优化网联自动驾驶车辆(connected and autonomous vehicle,CAV)专用道设置条件下的交织区时空资源配置,以保证 CAV 高效安全汇入主线专用道,提出了一种基于深度强化学习的交织区集成控制策略。以主线三车道高速公路为研究对象,并设置内侧车道为 CAV 专用道,设计了充分考虑 CAV 专用道汇入需求同时兼顾主线通行效率和匝道排队长度的多目标奖励函数,利用深度确定性策略梯度(deep deterministic policy gradient,DDPG)算法实现包括入口匝道信号控制、主线车道级可变限速以及 CAV 车队间隙调整的集成控制策略,最后使用 SUMO 和 Python 搭建高速公路交织区仿真场景,验证所提集成控制策略的有效性。研究表明:CAV 渗透率为 30% 时,在低、中、高不同水平交通需求场景下,对比无控制情况集成控制策略可使 CAV 汇入专用道的纵向位置有所提前,CAV 汇入专用道成功率分别增加了 19.34%、22.86%、25.55%;此外,车辆平均行程时间也分别降低了 5.42%、17.41%、20.65%。所提出的交织区集成控制策略效果显著,不仅实现了 CAV 汇入专用道的有效引导,还提升了主线的通行效率及运行安全,为改善 CAV 专用道设置条件下高速公路交织区交通运行状况提供了理论依据和技术参考。

**关键词:**智能交通;交通管控;深度强化学习;高速公路交织区;CAV 专用道;混合交通流

中图分类号:U491.2

文献标志码:A

文章编号:0367-6234(2025)11-0001-11

## An integrated control strategy for freeway weaving area considering CAV dedicated lanemerging demand

XI Shu, CHEN Xumei, LI Peikun, MA Jiaxin

(School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China)

**Abstract:** To improve freeway operational efficiency and optimize spatiotemporal resource allocation in weaving area under the setting of dedicated lane for connected and autonomous vehicles (CAVs), a deep reinforcement learning-based integrated control strategy is proposed. This strategy aims to ensure the efficient and safe merging of CAVs into the dedicated lane. The research focuses on a three-lane freeway configuration, with the innermost lane designated for CAVs. A multi-objective reward function is developed to address the dedicated lane merging demand of CAVs while simultaneously considering the efficiency of mainline traffic and the length of ramp queue. The deep deterministic policy gradient (DDPG) algorithm is employed to implement the integrated control strategy, which encompasses entrance ramp signal control, variable speed limit at the lane level, and adjustment to the gaps between CAV platoon. A simulation environment for the freeway weaving area is designed using SUMO and Python to assess the effectiveness of the proposed control strategy. The results demonstrate that, with a CAV penetration rate of 30%, the integrated control strategy advances the longitudinal positioning of CAVs entering the dedicated lane, merging success rate increases of 19.34%, 22.86%, and 25.55% under low, medium, and high traffic demand scenarios, respectively. Additionally, average vehicle travel time is reduced by 5.42%, 17.41%, and 20.65% under the same conditions. The proposed integrated control strategy for weaving area demonstrates significant effectiveness by not only achieving effective guidance for CAV merging dedicated lane but also enhancing the traffic efficiency and operational safety of the mainline, providing a theoretical basis and technical reference for optimizing the traffic operation in weaving areas of freeway under CAV dedicated lane conditions.

**Keywords:** intelligent transportation; traffic control; deep reinforcement learning; freeway weaving area; CAV dedicated lane; mixed traffic flow

收稿日期:2024-10-16;录用日期:2024-11-15;网络首发日期:2025-08-01

网络首发地址:https://link.cnki.net/urlid/23.1235.T.20250801.1148.004

基金项目:国家自然科学基金(72271020);中央高校基本科研业务费专项资金(2024JBZX025)

作者简介:席殊(1998—),男,硕士研究生;陈旭梅(1974—),女,教授,博士生导师

通信作者:陈旭梅, xmchen@bjtu.edu.cn

随着自动驾驶技术的快速发展,网联自动驾驶车辆逐渐成为未来交通系统的重要组成部分。车路协同作为自动驾驶领域的前沿概念和发展趋势,其通过先进的无线通信和新一代互联网技术,实现车辆与车辆、车辆与行人、车辆与道路之间的动态实时信息交互,成为交通运输领域新质生产力的重要体现。车路协同技术为高速公路带来了新的发展机遇,而设置高速公路 CAV 专用道则是车路协同技术的典型应用场景,旨在为 CAV 提供协同感知、协同决策、协同控制服务。

设置 CAV 专用道可充分发挥 CAV 对提升交通流稳定性的技术优势<sup>[1-3]</sup>,目前在学术研究上多通过计算机微观仿真方法评估 CAV 专用道设置效果<sup>[4-5]</sup>,多数研究结果表明当 CAV 达到一定的市场渗透率后,设置 CAV 专用道有利于提升道路通行能力并降低通行延误<sup>[6-8]</sup>。然而在此技术背景下,由于 CAV 汇入专用道将会加剧主线车流交织,造成高速公路交织区车流运行周期性拥堵,现有高速公路路端或车端的单一控制策略难以解决专用道设置条件下交织区的广域时空资源配置问题,如何通过更全面且更有针对性的控制手段使 CAV 高效安全地汇入专用道是当前 CAV 专用道落地使用面临的现实问题与挑战。

由于深度强化学习 (deep reinforcement learning, DRL) 算法在解决高速公路交通流及车辆汇入控制场景非线性、时变性强、序列决策等问题时具有显著优势,现已成为交通管控的有效研究方法,而已有研究基于 DRL 的单一控制策略则主要集中在信号控制和可变限速方面。例如:Han 等<sup>[9]</sup>提出了一种基于物理信息的深度强化学习匝道信号控制策略,该策略以高速公路通过量为奖励函数,使用历史数据和交通流模型生成的数据相结合来训练强化学习模型,结果表明该策略能显著降低车辆行程总时间。张健等<sup>[10]</sup>利用深度 Q 网络 (deep Q-network, DQN) 算法实现换道信号控制,以 CAV 进入专用道数量为奖励函数,结果表明该算法能有效增加 CAV 换道至专用道的比例。Lu 等<sup>[11]</sup>提出了一种基于双延迟深度确定性策略梯度 (twin delayed deep deterministic policy gradient, TD3) 算法的车道级可变限速控制方法,以瓶颈区车速和碰撞风险作为奖励函数,结果表明该方法能够在降低碰撞风险的同时提高交通效率。韩磊等<sup>[12]</sup>提出一种基于改进竞争双深度 Q 网络可变限速控制策略,以最小化车辆行程时间作为奖励函数,结果表明能够有效提高混合交通流环境下瓶颈区的通行效率。此外,也有一些学者针对智能网联环境下 CAV 汇入主线的车端控

制算法开展了相关研究<sup>[13-14]</sup>。

综上所述,现有相关文献研究缺乏同时从路端和车端考虑的集成控制算法,在智能网联环境下如何充分发挥车路协同信息共享优势保证 CAV 有充足的时空条件顺利汇入专用道行驶,同时尽可能降低 CAV 换道对主线运行效率的影响是当前 CAV 专用道应用亟待解决的问题;此外也缺乏考虑 CAV 专用道汇入需求的多目标奖励函数设计方法,现有研究奖励函数主要集中考虑平均车速、行程时间等交通流特征,难以有效表征本文研究场景下 CAV 汇入专用道成功率和交织区运行效率的复杂关系。

鉴于此,本文提出一种基于深度强化学习的车路协同集成控制策略。面向智能网联环境设置 CAV 专用道仿真场景,设计了集成路端匝道动态信号控制和车道级可变限速、车端 CAV 车队间隙动态调整的一体化控制策略,并运用 DDPG 深度强化学习算法实现,并通过构建优先经验回放机制提高算法响应效率,考虑 CAV 专用道汇入需求提出渐进式换道意图奖励函数,并融合 CAV 换道安全约束条件以保障 CAV 高效安全汇入专用道,同时设计了兼顾优化主线、匝道运行负荷的多目标奖励函数以指导最优策略的选择。

## 1 问题描述与转化

### 1.1 问题描述

高速公路交织区因上下匝道车辆流线交织易形成通行瓶颈,智能网联环境下 CAV 专用道的引入将使交织区车辆流线更为复杂,但 CAV 可通过 V2X (vehicle to everything) 技术实现 CAV 与 CAV、CAV 与道路基础设施的信息共享,这一优势为解决高速公路 CAV 专用道设置后 CAV 的高效汇入问题提供了新的解决思路,基于此,本文提出了面向智能网联环境的高速公路交织区车路协同集成控制模型。

所提出的模型架构如图 1 所示。本文研究场景为高速公路交织区路段,包含三车道的主线路段、单车道的入口匝道以及单车道的出口匝道,行驶方向最内侧设置为 CAV 专用道。交织区上游方向为车道级可变限速区域,旨在通过动态调整交织区入口车道速度为匝道车辆汇入主线提供充足的时间条件,同时起到提前控制专用道内 CAV 队列速度作用,避免后续车辆因前方 CAV 汇入汇出专用道而出现的车辆紧急制动行为。入口匝道设有动态信号控制器,目的是调节匝道汇入主线的交通流量<sup>[15-16]</sup>;交织区下游路段为 CAV 专用道汇入引导区域,通过专用道内 CAV 队列动态间隙调整为匝道 CAV 提供充足的换道空间。该系统通过路段检测器实时检测

状态信息,交通控制单元(traffic control unit,TCU)基于交通状态信息,随后将集成控制指令发送至路侧单元(road side unit,RSU),CAV 通过 RSU 接收控制中心传来指令并实施,包括车道限速值、红绿灯时长、CAV 车队间隙调整,人工驾驶车辆(human

driven vehicle,HDV)则通过观察可变信息标志(variable message sign,VMS)发布的限速信息、入口匝道红绿灯状态以及 CAV 运动状态变化调整自身驾驶行为。

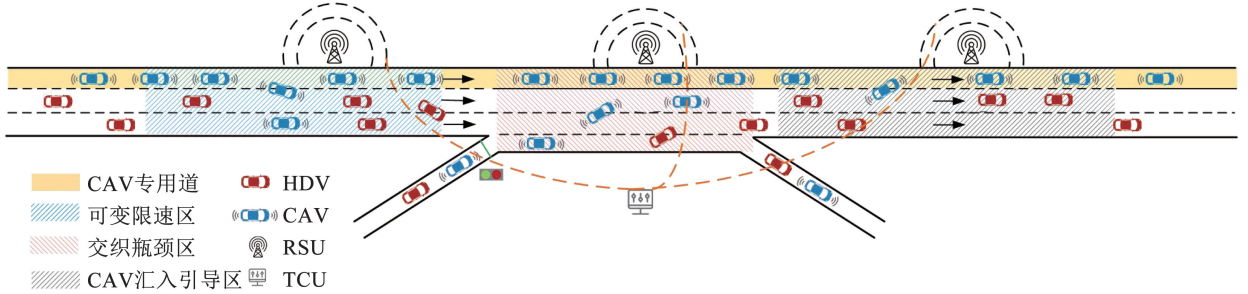


图 1 CAV 专用道设置条件下交织区集成控制系统结构

Fig. 1 Structure of integrated control system for weaving area under CAV dedicated lane

## 1.2 问题转化

本文提出的交织区车路协同集成控制问题可以等效为马尔可夫决策过程(Markov decision process,MDP),适合使用深度强化学习方法完成控制过程,其组成要素包括状态空间  $S$ 、动作空间  $A$ 、奖励函数  $R$ 、折扣因子  $\gamma$ 。就本文研究来说,状态空间应包含交织区道路状况信息,动作空间则是指车道级限速值、匝道红绿灯时长以及 CAV 车队间隙,智能体在与环境反复交互试错学习的过程中,在当前状态  $s_t$  根据策略  $\pi$  采取某个动作,之后外部环境将反馈给智能体一个数值奖励  $r_{t+1}$ ,同时返回一个新的状态  $s_{t+1}$ ,智能体重复迭代以上过程,直至累计折扣奖励最大化。各要素的详细分析如下。

### 1.2.1 状态空间

本文研究的交织区集成控制问题可将状态空间分为路端环境变量和车端环境变量,路端环境变量表示路侧实时交通流状况,包括车道占有率、车道车流平均速度、车道限速值、绿灯时长信息,同时将上一周期内车道限速值、绿灯时长控制信息作为历史特征输入控制算法有利于智能体感知状态做出最优策略;车端环境变量来自于 CAV 状态信息,包括 CAV 速度、加速度、纵向位置等信息,二者共同组成本文研究的状态空间。因此定义状态空间:

$$S = \{S_{road}, S_{CAV}\} \quad (1)$$

$$S_{road} = \{\rho_{up}, \rho_{weave}, \rho_{down}, \rho_{ramp}, V_{flow}, V_{limit}, T_{green}\} \quad (2)$$

$$S_{CAV} = \{v_i, a_i, x_i, \Delta v_{pi}, d_{pi}, \Delta v_{fi}, d_{fi}\} \quad (3)$$

式中: $S_{road}$  为路端环境变量集合, $S_{CAV}$  为车端环境变量集合, $\rho_{up}$  为交织区上游各车道占有率集合, $\rho_{weave}$  为交织区各车道占有率集合, $\rho_{ramp}$  为交织区入口匝道占有率, $V_{flow}$  为交织区各车道车流平均速度, $V_{limit}$

为上个控制周期内可变限速区各车道限速值集合, $T_{green}$  为上个控制周期内入口匝道绿灯控制时间; $v_i$ 、 $a_i$ 、 $x_i$  分别为 CAV 专用道相邻车道内 CAV 的速度、加速度、纵向位置, $\Delta v_{pi}$ 、 $d_{pi}$  分别表示 CAV 与专用道内前车的速度差和间距, $\Delta v_{fi}$ 、 $d_{fi}$  分别表示 CAV 与专用道内后车的速度差和间距。

### 1.2.2 动作空间

上述车路协同控制问题形成了较大的离散动作集合,为降低动作空间维度,本文定义在时刻  $t$  生成一个联合动作空间  $a_t$ , $a_t$  中第 1 个变量  $a_{limit}$  表示车道级限速值动作空间,第 2 个变量  $a_{green}$  用来定义入口匝道绿灯控制时间,第 3 个变量  $a_{gap}$  表示 CAV 队列间隙动作空间。为简化问题并提高算法可实现性,CAV 各行驶阶段换道动作均由换道模型生成,不受深度强化学习算法控制。

动作空间  $a_t$  定义为

$$a_t = [a_{limit}, a_{green}, a_{gap}] \quad (4)$$

车道级可变限速具体计算方法为

$$v_i = v_0 + \Delta v \cdot \text{int}(a_{limit}) \quad (5)$$

式中: $v_0$  表示主线车道级可变限速初始值,取值为 90 km/h; $\Delta v$  表示主线车道级可变限速变化值,根据工程实际情况,限速值一般为 10 的整数倍,因此  $\Delta v$  取 10 km/h; $\text{int}(a_{limit})$  表示将  $a_{limit}$  数值向下取整, $a_{limit}$  取值为  $[-3, 3]$ 。

匝道绿灯时长计算方法为

$$t_{green} = t_0 + \Delta t \cdot \text{int}(a_{green}) \quad (6)$$

式中: $t_0$  表示匝道初始绿灯时长,取值为 30 s; $\Delta t$  表示匝道绿灯时长变化值,取值为 5 s; $a_{green}$  取值为  $[-3, 3]$ 。

考虑到专用道内 CAV 车队运行中车辆动力特性较为复杂,本文使用 SUMO 插件 Simpla<sup>[17]</sup> 控制

CAV 车队的纵向驾驶行为,通过 openGap 效用函数实现 CAV 车队的通信协同,进而完成 CAV 专用道内车队的动态间隙调整,其调整机理如图 2 所示,位于 2 号车道内的 CAV 发出换道至专用道请求,由

RSU 传递至专用道内 CAV,为其匹配距离其空间位置最近的专用道 CAV,此时专用道内 CAV 通过改变队列中与前车 CAV 前后间距,调整至合适间隙后实现 2 号车道内 CAV 顺利切换进入 3 号专用车道。

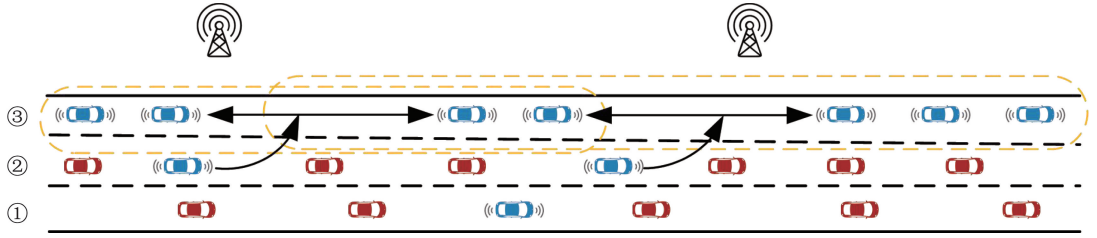


图 2 专用道内 CAV 车队间隙动态调整策略

Fig. 2 Dynamic adjustment strategy for CAV platoon gap in dedicated lane

调整 CAV 专用道内车队动态间隙的计算方法为

$$G = G_0 + \Delta G \cdot \text{int}(a_{\text{gap}}) \quad (7)$$

式中: $G_0$  表示初始车队间隙,取值为 15 m; $\Delta G$  表示车队间隙变化值,取值为 3 m; $a_{\text{gap}}$  取值为  $[-4, 4]$  [18-19]。

### 1.2.3 奖励函数

深度强化学习奖励函数的设计直接影响算法的响应速度和表现性能,本文要解决的主要问题是保障匝道 CAV 驶入主线专用道在时空上的优先性,同时尽可能减小对匝道和主线交织区运行的不利影响。因此本文将奖励函数定义为 CAV 换道意图奖励、匝道及交织区运行奖励、CAV 专用道通行效率奖励。

1) CAV 换道意图奖励。为保证 CAV 顺利由匝道驶入专用道,本文提出一种“渐进式”换道意图奖励满足 CAV 专用道汇入需求,以非线性数值函数形

式评价每个时间步长内交织区时空条件对 CAV 换道的影响,定义为

$$r_{1c0} = \exp[-(x_0 - L)^2 / (100L)] \quad (8)$$

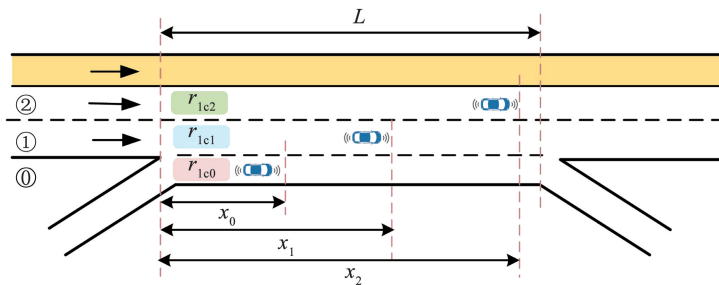
$$r_{1c1} = 1 - \exp[-(x_1 - L)^2 / (70L)] \quad (9)$$

$$r_{1c2} = 1 - \exp[-(x_2 - L)^2 / (45L)] \quad (10)$$

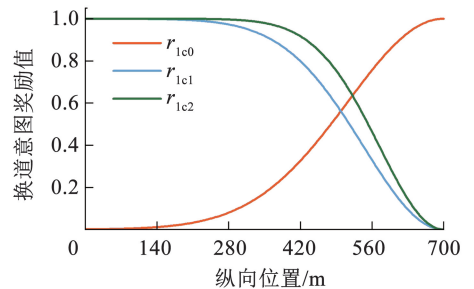
$$R_{1c} = -w_0 r_{1c0} + w_1 r_{1c1} + w_2 r_{1c2} \quad (11)$$

式中: $r_{1c0}$ 、 $r_{1c1}$ 、 $r_{1c2}$  分别表示 0、1、2 号车道的换道意图奖励; $x_0$ 、 $x_1$ 、 $x_2$  分别表示 0、1、2 号车道上 CAV 距离交织区起点的距离; $L$  表示交织区长度; $w_0$ 、 $w_1$ 、 $w_2$  表示各换道意图奖励的权重系数,其取值经多次训练测试微调确定; $R_{1c}$  表示各车道换道意图奖励总和。

如图 3 所示,其中图 3(b) 由式 (8) ~ (10) 绘制,当 CAV 位于 0 号车道时,给予智能体负奖励,位于 1 号车道时给予正奖励,位于 2 号车道时给予相对更大的正奖励,即 CAV 横向距离专用道越近,奖励值越大,以此激励 CAV 靠近专用道的换道意图。



(a) 不同车道 CAV 换道意图定义



(b) 换道意图奖励值变化曲线

图 3 换道意图奖励定义

Fig. 3 Definition of lane-changing intention reward

此外,为避免换道意图奖励函数使 CAV 忽略低奖励,发生由 0 号车道连续换道至专用车道的高风险且低舒适的驾驶行为,本文参考文献 [20],设计了 CAV 换道决策安全约束算法(程序如算法 1 所示),以此在最大程度上规避同一辆 CAV 在较短时间间隔发生连续换道以及同一车道前后 CAV 的换

道决策位置过于接近等危险换道行为。

算法 1: CAV 换道决策安全检测机制

Input: 目标 CAV 车辆  $v_i$  上一次换道执行时刻  $t_1$ , 当前换道决策时刻  $t_2$ ;  
 目标 CAV 通信范围内 CAV 车辆 ID 列表:  $\text{CAVid} = \{sv_1, sv_2, sv_3, \dots, sv_i\}$ ;  
 换道决策编码  $\text{LC}_{\text{decision}}(i) = \{0, 1\}$ ,  $i \in \text{CAVid}$ , 0 表示保持当前车道, 1 表示执行换道决策;

车辆  $i$  的当前车道  $\text{Lane}_{\text{current}}(i) = \{0, 1, 2\}, i \in \text{CAVid}$ ;

车辆  $i$  的目标车道  $\text{Lane}_{\text{target}}(i) = \{0, 1, 2\}, i \in \text{CAVid}$ ;

车辆的纵向换道决策位置  $x_{\text{lon}}(i)$ ;

同一 CAV 紧邻换道决策时刻时间间隔最小值  $\Delta t_{\text{lc}}$ ;

同一车道前后相邻 CAV 换道纵向位置最小值  $\Delta s_{\text{lc}}$

```

1: 初始化  $\text{LC}_{\text{decision}}(i) = 0$ 
2: If  $|t_2 - t_1| \leq \Delta t_{\text{lc}}$  then
3:    $\text{LC}_{\text{decision}}(i) = 0$ 
4: Else
5:   If  $|x_{\text{lon}}(v_i) - x_{\text{lon}}(sv_i)| \leq \Delta s_{\text{lc}}$  and  $\text{Lane}_{\text{current}}(v_i) = \text{Lane}_{\text{current}}(sv_i)$  and  $\text{Lane}_{\text{target}}(v_i) = \text{Lane}_{\text{target}}(sv_i)$ 
6:      $\text{LC}_{\text{decision}}(i) = 0$ 
7:   Else
8:      $\text{LC}_{\text{decision}}(i) = 1$ 
9:   End if
10: End if
11: Return  $\text{LC}_{\text{decision}}(i)$ 

```

2) 匝道及交织区运行奖励。CAV 驶入主线后可能会引起主线交通流的震荡, 同时也要防止入口匝道排队长度过长导致延误增加和效率损失。因此, 以含交织区平均速度和匝道车辆排队长度变量的函数表示运行情况奖励, 从而衡量车道级可变限速和匝道信号控制动作的可行性, 计算方法如下所示:

$$R_v = \frac{1}{n} \sum_{i=1}^n (v_i / v_{\text{max}}) \quad (12)$$

$$R_q = -1 / (1 + e^{\alpha(\beta - n_q) + \delta}) \quad (13)$$

$$R_f = w_v R_v + w_q R_q \quad (14)$$

式中:  $R_v$  表示交织区运行速度奖励;  $n$  表示交织区的车辆数量;  $v_i$  表示交织区内第  $i$  辆车在某一时间步长的速度;  $v_{\text{max}}$  表示交织区路段最大限速;  $R_q$  表示匝道排队奖励函数;  $n_q$  表示入口匝道排队车辆数;  $\alpha$ 、 $\beta$ 、 $\delta$  为系数<sup>[21]</sup>;  $R_f$  表示交织区运行奖励总和;  $w_v$ 、 $w_q$  表示奖励权重系数<sup>[22]</sup>。

3) CAV 专用道通行效率奖励。为评价集成控制动作对 CAV 专用道通行效率的影响, 通过计算 CAV 专用道的流量变化表征, 计算方法为

$$R_{\text{DL}} = (F_{\text{out}}^{\text{CAV}} - F_{\text{in}}^{\text{CAV}}) / (F_{\text{out}}^{\text{Edge}} - F_{\text{in}}^{\text{Edge}}) \quad (15)$$

式中:  $R_{\text{DL}}$  表示 CAV 专用道通行效率奖励函数;  $F_{\text{in}}^{\text{CAV}}$ 、 $F_{\text{out}}^{\text{CAV}}$  分别表示在 CAV 专用道汇入引导区起终点单位时间内的 CAV 流入量和流出量;  $F_{\text{in}}^{\text{Edge}}$ 、 $F_{\text{out}}^{\text{Edge}}$  分别表示 CAV 专用道汇入引导区所在路段起终点单位时间内的所有车辆的流入量和流出量。

4) 多目标奖励函数。本文研究问题涉及多目标决策, 可通过构建合成奖励函数将多目标问题转换为单目标问题, 实现不同目标之间的权衡优化, 多目标奖励函数为

$$R = R_{\text{lc}} + R_f + R_{\text{DL}} \quad (16)$$

## 2 集成控制算法设计

### 2.1 DDPG 算法

采用上述的动作空间设计方法能将离散动作值转化为连续动作值, 有效降低动作空间维度, 而 DDPG 算法融合了深度神经网络和确定性策略梯度方法优势, 能够处理高维状态空间和连续动作空间问题, 故本文使用 DDPG 算法实现控制策略。DDPG 算法直接输出智能体动作, 主要包括 Actor 模块、Critic 模块、经验池模块以及对偶网络结构。其目标  $Q$  值通过从目标网络中复制参数进行更新, 从经验回放池中随机抽取样本进行训练, 以最小化  $Q$  值网络的均方误差为训练目标, 从而不断接近目标  $Q$  值。

其损失函数和策略梯度的公式分别表达如下:

$$L(\theta^Q) = E_{(s,a,r,s')} [(r + \gamma Q'(s', \mu'(s') | \theta^Q) - Q(s, a | \theta^Q))^2] \quad (17)$$

$$\nabla_{\theta} J(\theta^{\mu}) \approx E_s [\nabla_{\theta} \mu(s | \theta^{\mu}) \nabla_a Q(s, a | \theta^{\mu}) |_{a=\mu(s)}] \quad (18)$$

DDPG 算法将  $Q$  值网络的均方误差最小化实现更新 Critic 网络, 使用策略梯度方法更新 Actor 网络, 达到在连续动作空间高效率学习的目的。

### 2.2 优先经验回放机制

一般的经验回放机制常采用均匀抽样, 从回放缓冲池中随机抽取样本进行训练, 但这种情况下重要样本的训练价值得不到充分利用, 因此本文算法采用优先经验回放机制训练智能体, 以提高算法的训练速度和响应效率, 保证智能体能够充分利用经验回放池中先前经验, 这对本文提出的交织区集成控制智能体在经验数据训练上更为有利。

经验样本抽样概率定义为

$$P(i) = p^e(i) / \sum_i p^e(i) \quad (19)$$

式中:  $P(i)$  表示样本  $i$  的抽样概率,  $p^e(i)$  表示样本  $i$  的优先级,  $e$  表示优先级权重。

使用优先经验回放机制会使 DDPG 的网络输出产生偏差, 对此使用重要性采样权重进行修正:

$$\omega_i = \left[ \frac{1}{N} \cdot \frac{1}{P(i)} \right]^{\lambda} \quad (20)$$

式中:  $\omega_i$  表示经验样本  $i$  的重要性采样权重;  $N$  为经验回放缓冲池中的经验样本数量;  $\lambda$  表示超参数, 其取值为  $[0, 1]$ 。

### 2.3 算法框架

基于 DDPG 的高速公路交织区集成控制策略流程如图 4 所示。其主要分为 SUMO 仿真交互环境和 DDPG 控制中心, 在第  $i$  个控制周期, 从 SUMO 获取状态空间变量  $s_t$ , 并通过 RSU 传递至控制中心, 经

Actor 网络输出动作  $a_t$  后由 TCU 将集成控制指令传输至路端和车端发布实施,随后进入下一状态空间  $s_{t+1}$ ,经验被聚合为  $\langle s_t, a_t, r_t, s_{t+1} \rangle$  四元数组存入

经验回放池,并发送到 Critic 网络进行训练,上述过程不断循环迭代,直至算法收敛。其详细算法流程如算法 2 所示。

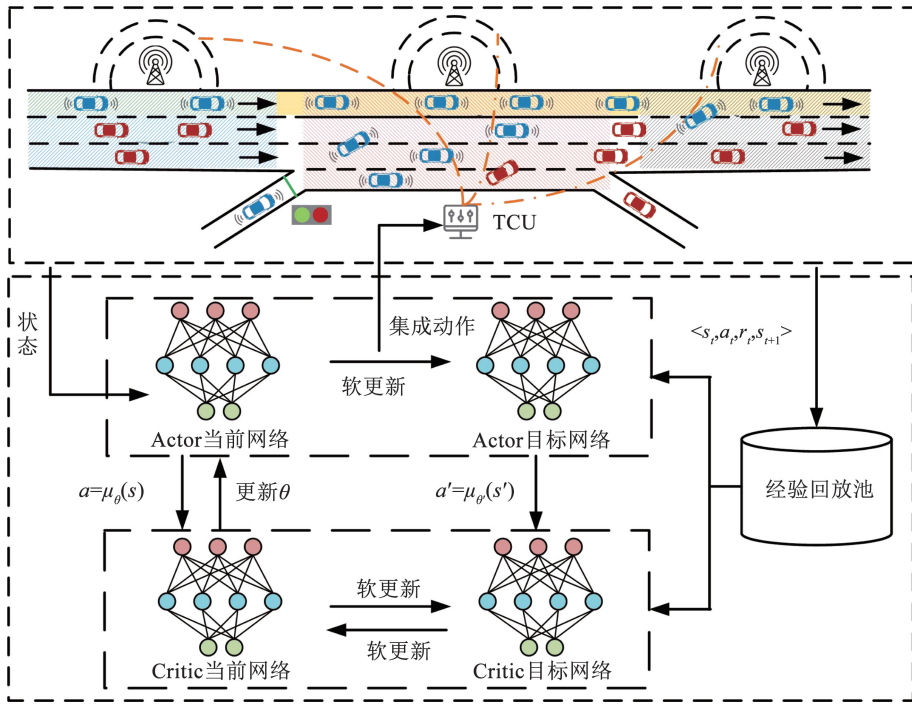


图 4 基于 DDPG 的高速公路交织区集成控制策略流程

Fig. 4 DDPG-based integrated control strategy process for freeway weaving area

算法 2:DDPG 算法

- 1: 初始化经验回放池  $D$
- 2: 随机初始化 Critic 网络  $Q(s, a | \theta^Q)$  和 Actor 网络  $\mu(s, a | \theta^\mu)$
- 3: 初始化目标网络  $Q'(s, a | \theta^{Q'})$  和  $\mu'(s, a | \theta^{\mu'})$
- 4: For episode = 1 to  $M$ , do:
- 5: 初始化动作随机探索噪声  $\xi$
- 6: 初始化 SUMO 仿真环境作为初始状态空间  $s_0$
- 7: For simulation step = 1 to  $T$ , do:
- 8: 根据 Actor 网络  $\mu(s, a | \theta^\mu)$  和探索噪声选择动作  $a_t$
- 9: 使用 SUMO 的 TraCI 接口将  $a_t$  转化为集成控制指令
- 10: 得到奖励  $r_t$ , 更新 SUMO 仿真环境为下一状态空间  $s_{t+1}$
- 11: 储存  $\langle s_t, a_t, r_t, s_{t+1} \rangle$  数据至经验回放池  $D$
- 12: 根据优先经验回放机制从  $D$  中选择  $N_s$  个样本
- 13: 根据式(17)计算最小化损失函数来更新 Critic 网络
- 14: 根据式(18)计算采样策略梯度来更新 Actor 网络
- 15: 更新目标网络  $\theta^{Q'} \leftarrow \tau \theta^{Q'} + (1 - \tau) \theta^Q$ ,  $\theta^{\mu'} \leftarrow \tau \theta^{\mu'} + (1 - \tau) \theta^\mu$
- 16: end for
- 17: end for

### 3 高速公路仿真场景搭建

#### 3.1 混合交通流驾驶行为建模

混合交通流驾驶行为模型可分为跟驰模型和换道模型两部分。本文选取智能驾驶员模型(intelligent driver model, IDM)<sup>[23]</sup>控制 HDV 的纵向驾驶行为;使用协同自适应巡航控制(cooperative adaptive cruise control, CACC)模型<sup>[24]</sup>控制 CAV 的

纵向跟驰行为,需要注意的是,当 CAV 跟驰 HDV 时,由于无法进行车辆间的通信 CACC 将自动退化为自适应巡航模式(adaptive cruise control, ACC)。

HDV 换道模型采用 SUMO 内置的 LC2013 模型<sup>[25]</sup>;当 CAV 的目标车道为普通车道时,由于受到 HDV 的干扰,CAV 无法充分发挥网联优势,此时 CAV 与 HDV 的换道博弈关系通过 LC2013 中的换道合作系数表征,并参考文献[26]中计算得出的换道可接受间隙决定 CAV 是否换道;当 CAV 的目标车道为专用车道时,由于目标 CAV 可捕捉到通信范围内目标车道的全部 CAV 属性,故本文建立了考虑 CAV 多辆后车信息的改进 Mobil 换道决策模型<sup>[27]</sup>,模型表达式为

$$\tilde{a}_{SV}, \tilde{a}_{TFV_i} \geq -b_{\text{safe}}, i \in N_T \quad (21)$$

$$U_V = \tilde{a}_{SV} - a_{SV} +$$

$$p \left[ \sum_{i=1}^{N_T} (\tilde{a}_{TFV_i} - a_{TFV_i}) \right] + q (\tilde{a}_{FV} - a_{FV}) > \Delta a_{\text{th}}, i \in N_T \quad (22)$$

$$N_T = \{TFV_i | 0 \leq |x_{SV} - x_{TFV_i}| \leq L\} \quad (23)$$

式中: $\tilde{a}_{SV}$ 为目标 CAV 换道后的加速度, $\tilde{a}_{TFV_i}$ 为目标车辆换道后的跟随车加速度, $\tilde{a}_{FV}$ 为目标车辆原车道的跟随车加速度, $N_T$ 为目标 CAV 通信范围内的车辆集合, $U_V$ 为目标 CAV 的换道总效益, $a_{SV}$ 、 $a_{TFV}$ 、

$a_{FV}$  分别为换道前目标车辆、原车道跟随车、目标车道跟随车的加速度,  $p$ 、 $q$  为礼让系数,  $a_{th}$  为换道收益阈值,  $x_{SV}$ 、 $x_{TFV_i}$  分别为目标车辆及其目标车道跟随车辆的纵向位置,  $L$  为目标 CAV 的通信范围。

表 1 车辆仿真参数设置

Tab. 1 Vehicle simulation parameter settings

车辆类型	最大加速度/ ( $m \cdot s^{-2}$ )	最大减速度/ ( $m \cdot s^{-2}$ )	车头时距/s	最小车间 距/m	换道合作 系数	战略式换道 系数	不完美驾驶 系数
HDV	2.5	-7.5	1.5	2.0	0.5	1.0	0.5
CAV	2.5	-7.5	1.0	1.5	1.0	1.5	0

### 3.2 仿真场景构建与训练参数设置

本文采用 SUMO 构建高速公路交织区仿真场景,并基于 Python 通过 TraCI 实现二次开发,验证提出的基于 DDPG 的交织区集成控制策略。本文高速公路仿真场景由主线和入口匝道、出口匝道构成,主

为提高 CAV 汇入专用道效率,本文设置由入口匝道驶入的 CAV 在交织区与汇入引导区内只要满足上述换道模型决策条件及安全约束条件,CAV 执行向左换道。SUMO 车辆仿真参数如表 1 所示。

线设置三车道,入口匝道和出口为单车道,各路段设有线圈检测器。仿真场景示意图如图 5 所示,参考 CAV 专用车道入口区域设计的研究<sup>[28]</sup>,并根据本文提出的集成控制策略将主线划分为车道级可变限速区、交织瓶颈区、CAV 汇入引导区。



图 5 仿真场景搭建

Fig. 5 Simulation scenario design

高速公路主线设计速度为 120 km/h,匝道限速 40 km/h,仿真路段全长为 3 km。车道级可变限速区长 500 m,交织区长度为 700 m;设置 CAV 专用道时,CAV 需在满足换道决策安全约束条件下变道两次汇入专用道,CAV 汇入引导区设置为 1 km,以保证 CAV 有足够的车道空间驶入专用道。

为了使智能体即 TCU 适应多种交通需求场景,本文将每一训练回合设置为平峰高峰交替出现,每回合仿真持续时间为 10 800 s,集成动作控制周期设置为 1 min,仿真步长为 0.5 s,交通需求设置如表 2 所示。

表 2 交通需求设置

Tab. 2 Traffic demand setting

时间/s	交通需求	交通需求/(veh · h <sup>-1</sup> )		
		主线→主线	主线→出口匝道	入口匝道→主线
0 ~ 3 600	低需求	1 800	180	500
>3 600 ~ 7 200	高需求	3 000	300	900
>7 200 ~ 10 800	中需求	2 400	240	700

SUMO 仿真过程基于以下假设:1) 不考虑车-车通信、车-路通信的延迟时间;2) 仅考虑小型乘用车,且 CAV 车辆与 HDV 车辆外观尺寸,加减速性能一致;3) 仿真路段不考虑设置 CAV 专用道初始入口,即上游 CAV 在进入交织区之前已驶入专用道;4) 车辆生成遵循泊松分布;5) 假设所有车辆遵循车道级可变限速、匝道信号控制规则,不存在违规驾驶行为;6) CAV 专用道的汇入截止位置为 2 500 m。

经过多次训练调优,本文算法最终选定的超参数取值如下:折扣率为 0.95; Actor 网络学习率  $1 \times 10^{-3}$ ; Critic 网络学习率  $2 \times 10^{-3}$ ; 经验回放池容量为  $2 \times 10^4$ ; 批处理大小为 32; 更新参数  $\tau$  为  $5 \times 10^{-3}$ ; 各隐藏层神经元数量为(256, 128, 64); 动作噪声为 0.5。

## 4 实验结果分析

### 4.1 训练过程分析

本文分别针对基于 DDPG 算法的集成控制策略、CAV 车队动态间隙调整(CAV platoon dynamic gap adjustment, CPGA)的车端控制算法、融合车道级可变限速(lane-level variable speed limit, LVSL)和匝道控制(ramp metering, RM)的路端控制算法 3 种控制策略展开了仿真训练,直至奖励值趋于收敛。不同控制策略下各回合累计奖励值变化曲线如图 6(a)所示,可以看出训练前期由于智能体一直在探索最优控制策略,因此每回合总奖励值波动较大,当训练 120 回合之后,集成控制策略下的奖励值趋于收敛,对比其他策略,集成控制策略可以用更少的

训练回合数达到算法收敛的目标,证明其在收敛效率方面具有一定优势,此外集成控制策略对比单一的路侧控制和车侧控制策略奖励值相对更高,表示智能体能更好地权衡 CAV 专用道汇入需求和主线

通行效率之间的关系,使多目标奖励函数达到相对最优值。由图 6(b)可知,集成控制策略下的各分项奖励函数均得到有效收敛,验证了前文多目标奖励函数设置的合理性。

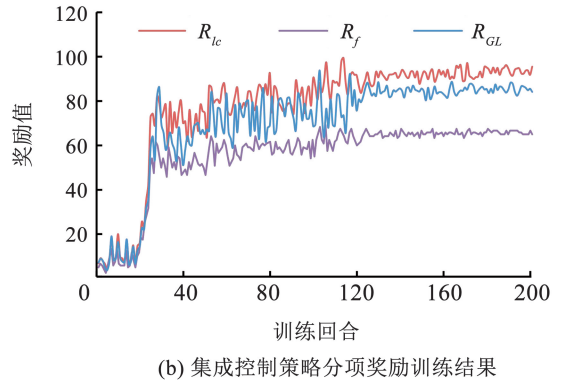
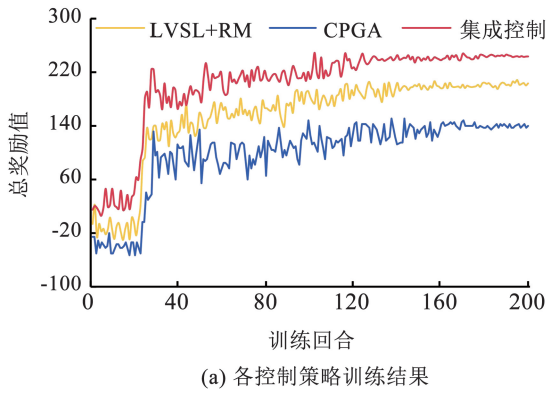


图 6 DDPG 算法训练结果

Fig. 6 DDPG algorithm training results

### 4.2 仿真结果分析

结合文献[29]的研究结果,当 CAV 渗透率为 30%时,设置 CAV 专用道对混合交通通行效率的提升效果最为显著,因此本文选取 CAV 渗透率为 30%场景下的仿真结果进行分析:

#### 4.2.1 CAV 换道至专用道时空分析

CAV 换道时空分布如图 7 所示,CAV 专用道汇入引导区的位置范围为 1 500 ~ 2 500 m,无控制时,CAV 切入专用道位置主要分布于 2 250 m 附近,换道位置十分靠后;相比无控制策略,“LVSL + RM”策

略、“CPGA”策略、“集成控制”策略均对 CAV 换道位置有明显的改善作用,“集成控制”策略可使 CAV 换道位置的高频点位于 1 950 m 左右,较无控制提前 300 m 左右。实验结果说明本文提出的集成策略能够为 CAV 提供充分的换道时空条件,使 CAV 驶入主线后可以尽早驶入专用道行驶,进而实现 CAV 与 HDV 的物理分离,降低混合交通流的复杂性,释放专用道和普通车道通行能力,同时降低交织区事故风险。

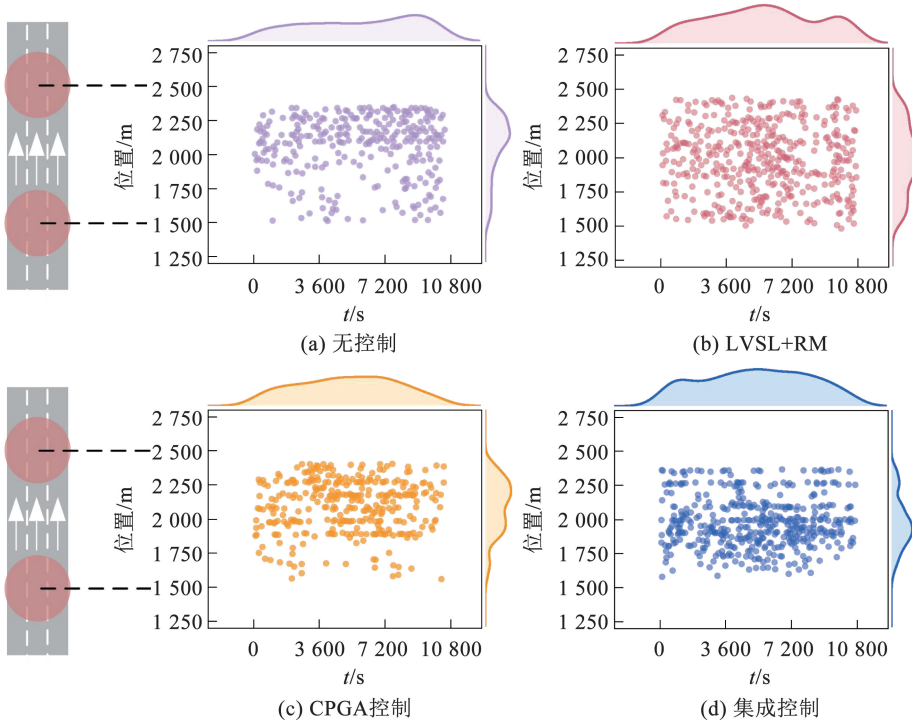


图 7 CAV 换道时空位置分布

Fig. 7 Distribution of spatial and temporal of lane-changing for CAV

#### 4.2.2 CAV 专用道汇入成功率分析

进一步分析各控制策略下 CAV 换道至专用道成功率如表 3 所示,其中,CAV 专用道汇入成功率是指在引导区终点前汇入专用道的 CAV 数量与匝道驶入主线 CAV 数量的比值。无控制策略下,由于交织区车流密集 CAV 难以寻找到合适的换道时空条件,因此汇入成功率较低,“LVSL + RM”策略下由

表 3 CAV 专用道汇入成功率对比

Tab.3 Comparison of the success rate of CAV dedicated lane merging

控制方法	CAV 汇入专用道成功率/%					
	低需求	低需求性能提升	中需求	中需求性能提升	高需求	高需求性能提升
无控制	29.33	—	23.33	—	19.26	—
LVSL + RM	38.67	9.34	35.71	12.38	32.22	12.96
CPGA	31.33	2.00	27.62	4.29	24.44	5.19
集成控制	48.67	19.34	46.19	22.86	44.81	25.55

#### 4.2.3 通行效率及安全

为探究“集成控制”策略在保证 CAV 高效汇入专用道的情况下交织区的实际运行效率,选取仿真 180 个控制周期主线交织区平均车速并绘图,结果如图 8 所示。相比无控制策略,“集成控制”、“CPGA”控制、“LVSL + RM”控制均对主线交织区平均车速均有所提升,“LVSL + RM”由于只设计了路端控制策略,而不考虑 CAV 车队间隙调整,其对于平均车速影响效果相对最佳,“集成控制”策略在高需求时间段(第 60 ~ 120 个控制周期)内平均车速有所下降,但仍优于无控制场景,“CPGA”策略与无控制场景差别不大。

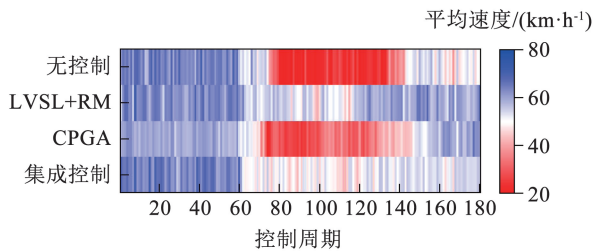


图 8 交织区平均车速对比

Fig.8 Comparison of average vehicle speed in the weaving area

各控制策略对车辆平均行程时间的影响如表 4 所示,由表 4 分析可知,“LVSL + RM”控制策略在低、中、高交通需求场景下对比无控制策略平均行程时间分别降低 9.73%、26.68%、28.31%;“CPGA”控制策略下平均行程时间分别降低 2.69%、2.51%、2.47%。相比“LVSL + RM”控制策略,“集成控制”策略虽为保证 CAV 高效汇入专用道而导致交织区通行效率的优化效果有所下降,但对比无控制,车辆平均行程时间仍分别降低 5.42%、17.41%、20.65%。

于对交织区车辆速度和匝道输入做出了一定限制,因此 CAV 的换道条件得到明显改善,其汇入成功率也大幅提高;“CPGA”控制策略下汇入成功率略有提高;“集成控制”策略对 CAV 的汇入成功率提高效果最为明显,高交通需求时段汇入成功率达 44.81%,较无控制策略性能提升达到 25.55%,充分证明了其引导 CAV 汇入专用道的有效性。

表 4 车辆平均行程时间对比

Tab.4 Comparison of average vehicle travel time

控制方法	平均行程时间/s		
	低需求	中需求	高需求
无控制	159.47	237.86	305.04
LVSL + RM	143.95	174.39	218.69
CPGA	155.18	231.88	297.52
集成控制	150.83	196.46	242.06

本文采用碰撞时间(time to collision, TTC)评价追尾安全风险,TTC 越小说明前后两车发生追尾碰撞的风险越大。TTC 频率分布曲线如图 9 所示,由图 9 可知集成控制策略下 TTC 分布在各统计区间内的概率密度更加均匀,且 0 ~ 3 s 的危险工况数据明显低于无控制场景,说明集成控制策略可有效降低交通流碰撞风险。

综合来看,相较于“LVSL + RM”控制策略,集成控制策略尽管在通行效率优化效果上稍显不足,但从换道时空分布、CAV 汇入成功率来看,集成控制策略优势突出,为 CAV 汇入专用道提供了充分的时空资源及条件,能更好地解决本文研究问题。

#### 4.3 灵敏度分析

根据文献[30]的研究结果,当 CAV 渗透率为 10% ~ 50% 时,设置 1 条 CAV 专用道能在一定程度上改善通行效率,为深入分析不同 CAV 渗透率条件下集成控制策略对 CAV 汇入专用道的影响,设置 CAV 渗透率为 10%、20%、30%、40%、50% 的仿真实验。各控制策略的 CAV 汇入成功率如图 10 所示。

分析图 10 可知,低需求时段下各控制策略随着 CAV 渗透率的增加 CAV 换道成功率也出现增加趋

势;中需求和高需求时段场景下随着 CAV 渗透率的增加,CAV 换道至专用道成功率表现出先增加后降低的趋势,同时高峰时段波动程度更大,出现此类现象原因可能是 CAV 渗透率的增大导致 CAV 专用道

通行能力接近饱和状态,因此 CAV 选择合适的时空间隙汇入专用道的难度有所增加。总体来看,本文所提出的集成控制策略受 CAV 渗透率影响不大,在各仿真实验下均表现出最优性能。

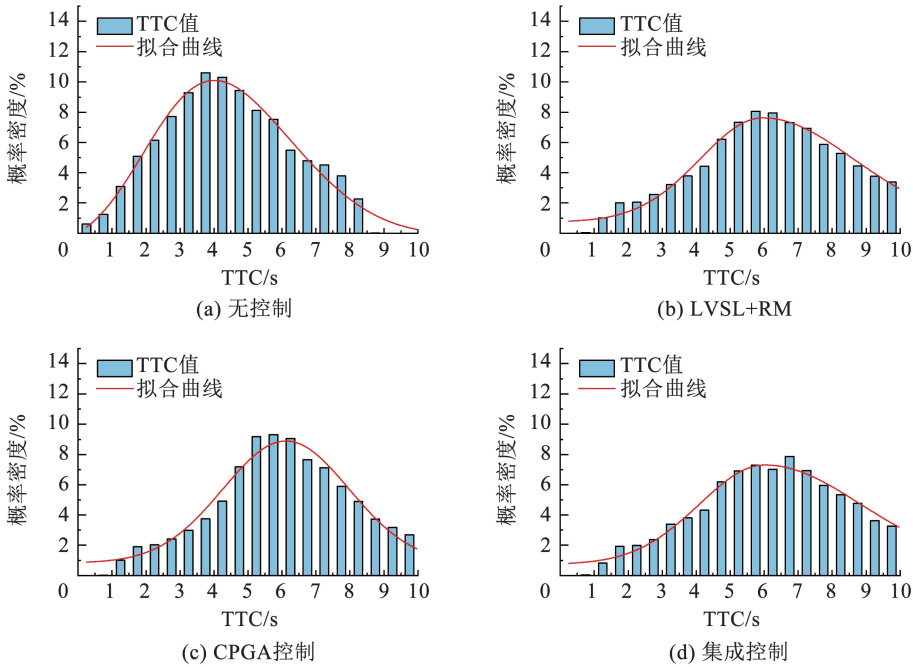


图 9 TTC 数据分布情况

Fig. 9 Distribution of TTC data

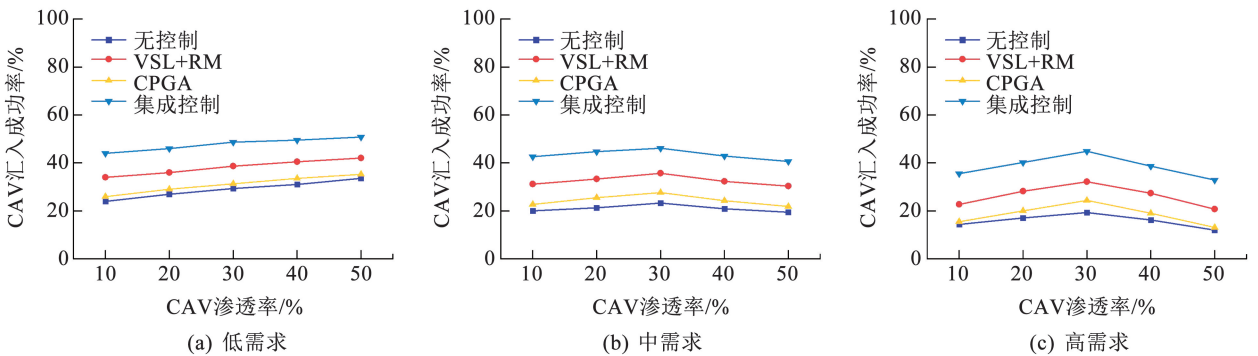


图 10 不同 CAV 渗透率下控制策略性能变化情况

Fig. 10 Variation of control strategy performance with different CAV penetration rates

### 5 结 论

1) 针对 CAV 专用道设置条件下的高速公路混合交通流环境,提出了集成匝道信号控制、主线车道级可变限速、CAV 车队间隙调整的交织区控制策略,将相关车路信息作为共享状态输入,以匝道信号相位时长、主线车道限速值、CAV 专用道车队间隙作为动作空间,提出换道意图奖励函数并结合换道安全约束算法驱动 CAV 高效安全汇入专用道,同时考虑主线通行效率和匝道排队长度设计了多目标奖励函数。

2) 仿真实验结果表明,本文所提集成控制策略

控制效果显著,既在一定程度上有效提高了 CAV 专用道设置条件下交织区时空资源配置的合理性,其 CAV 换道至专用道成功率在低、中、高交通需求控制时段分别为 48.67%、46.19%、44.81%,相比无控制策略分别提高了 19.34%、22.86%、25.55%,又提高了主线交织区平均车速,相比无控制其平均行程时间在低、中、高交通需求的控制时段分别可降低 5.42%、17.41%、20.65%,同时交通流运行安全性也得到一定提升。

3) 本文提出的集成控制策略在不同 CAV 渗透率下性能稳定性较强,可对网联自动驾驶技术各发展阶段有良好的适用性,变化趋势为:在低交通需求

时段下随着 CAV 渗透率增加 CAV 汇入专用道成功率也有所增加,在高交通需求和中交通需求时段下随着 CAV 渗透率的增加,CAV 汇入专用道成功率呈现先增加后减少的变化趋势,在 CAV 渗透率为 30% 时控制效果最佳。

## 参考文献

- [1] 秦严严, 胡兴华, 李淑庆, 等. 智能网联环境下混合交通流稳定性解析[J]. 哈尔滨工业大学学报, 2021, 53(3): 152  
QIN Yanyan, HU Xinghua, LI Shuqing, et al. Stability analysis of mixed traffic flow in connected and autonomous environment [J]. Journal of Harbin Institute of Technology, 2021, 53(3): 152
- [2] YE L, YAMAMOTO T. Modeling connected and autonomous vehicles in heterogeneous traffic flow [J]. Physica A: Statistical Mechanics and its Applications, 2018, 490: 269
- [3] MOHAJERPOOR R, RAMEZANI M. Mixed flow of autonomous and human-driven vehicles: analytical headway modeling and optimal lane management [J]. Transportation Research Part C: Emerging Technologies, 2019, 109: 194
- [4] KIM J, LIM D, SEO Y, et al. Influence of dedicated lanes for connected and automated vehicles on highway traffic flow [J]. IET Intelligent Transport Systems, 2023, 17(4): 678
- [5] YE L, YAMAMOTO T. Impact of dedicated lanes for connected and autonomous vehicle on traffic flow throughput [J]. Physica A: Statistical Mechanics and its Applications, 2018, 512: 588
- [6] 傅泽新, 陈旭梅, 王宇擎, 等. 智能网联环境下管理车道设置策略与影响因素分析[J]. 哈尔滨工业大学学报, 2023, 55(7): 24  
FU Zexin, CHEN Xumei, WANG Yuqing, et al. Managed lane setting strategies and influence factor analysis in intelligent connected environment [J]. Journal of Harbin Institute of Technology, 2023, 55(7): 24
- [7] XIAO L, WANG M, VAN AREM B. Traffic Flow Impacts of converting an HOV lane into a dedicated CACC lane on a freeway corridor [J]. IEEE Intelligent Transportation Systems Magazine, 2020, 12(1): 60
- [8] AMIRGHOLY M, SHAHABI M, GAO H. Traffic automation and lane management for communicant, autonomous, and human-driven vehicles [J]. Transportation Research Part C: Emerging Technologies, 2020, 111: 477
- [9] HAN Y, WANG M, LI L, et al. A physics-informed reinforcement learning-based strategy for local and coordinated ramp metering [J]. Transportation Research Part C: Emerging Technologies, 2022, 137: 103584
- [10] 张健, 李青扬, 李丹, 等. 基于深度强化学习的自动驾驶车辆专用道汇入引导[J]. 吉林大学学报(工学版), 2023, 53(9): 2508  
ZHANG Jian, LI Qingyang, LI Dan, et al. Merging guidance of exclusive lanes for connected and autonomous vehicles based on deep reinforcement learning [J]. Journal of Jilin University (Engineering and Technology Edition), 2023, 53(9): 2508
- [11] LU W, YI Z, GU Y, et al. TD3LVS: a lane-level variable speed limit approach based on twin delayed deep deterministic policy gradient in a connected automated vehicle environment [J]. Transportation Research Part C: Emerging Technologies, 2023, 153: 104221
- [12] 韩磊, 张轮, 郭为安. 混合交通流环境下基于改进强化学习的可变限速控制策略[J]. 交通运输系统工程与信息, 2023, 23(3): 110  
HAN Lei, ZHANG Lun, GUO Weian. Variable speed limit control based on improved dueling double deep Q Network under mixed traffic environment [J]. Journal of Transportation Systems Engineering and Information Technology, 2023, 23(3): 110
- [13] HU X, SUN J. Trajectory optimization of connected and autonomous vehicles at a multilane freeway merging area [J]. Transportation Research Part C: Emerging Technologies, 2019, 101: 111
- [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] 王兴举, 高桂凤, 宫城俊彦. 高速公路单点入口匝道 RLPM 控制方法[J]. 交通运输工程学报, 2012, 12(3): 101  
WANG Xingju, GAO Guifeng, MIYAGI T. RLPM control method of single entrance ramp for highway [J]. Journal of Traffic and Transportation Engineering, 2012, 12(3): 101
- [16] 章立辉, 余宏鑫, 熊满初, 等. 平衡主线和匝道交通运行的强化学习型匝道控制研究 [J]. 重庆交通大学学报(自然科学版), 2023, 42(4): 87  
ZHANG Lihui, YU Hongxin, XIONG Manchu, et al. Reinforcement learning ramp metering to balance mainline and ramp traffic operations [J]. Journal of Chongqing Jiaotong University (Natural science), 2023, 42(4): 87
- [17] VALIDI A, SMIRNOV N, OLAVERRI-MONREAL C. Metamodel-based simulation to assess platooning on battery energy consumption [J]. Transportation Research Part D: Transport and Environment, 2022, 109: 103350
- [18] YADAVALLI S R, DAS L C, WON M. RLPG: reinforcement learning approach for dynamic intra-platoon gap adaptation for highway on-ramp merging [C]//2023 International Conference on Intelligent Robots and Systems (IROS). Detroit: IEEE, 2023: 5514
- [19] ARAMRATTANA M, LARSSON T, ENGLUND C, et al. A simulation study on effects of platooning gaps on drivers of conventional vehicles in highway merging situations [J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(4): 3790
- [20] 王怡琛. 智能网联环境下考虑车辆异质性的高速公路自动车道决策仿真研究[D]. 成都: 西南交通大学, 2022  
WANG Yichen. Simulation research on highway lane change decision considering vehicle heterogeneity for connected and autonomous vehicles [D]. Chengdu: Southwest Jiaotong University, 2022
- [21] 蔺庆海, 何兆成, 谢俊, 等. 基于 AVI 数据与深度强化学习的城市快速路匝道协调控制方法 [J]. 中国公路学报, 2023, 36(10): 224  
LIN Qinghai, HE Zhaocheng, XIE Jun, et al. Urban expressway coordinated ramp metering approach using automatic vehicle identification data and deep reinforcement learning [J]. China Journal of Highway and Transport, 2023, 36(10): 224