

基于强化协作博弈方法的双车道混合交通流特性

郭静秋¹, 方守恩¹, 曲小波², 王亦兵³, 刘洋泽西¹

(1. 同济大学 道路与交通工程教育部重点实验室, 上海 201804; 2. 查尔姆斯理工大学 建筑与土木工程系, 查尔姆斯 41296; 3. 浙江大学 建筑工程学院, 浙江 杭州 310058)

摘要: 对元胞自动机引入 Gipps 跟驰模型, 并结合改进的 Q 强化学习方法分别建立普通车辆及智能网联车的微观行驶策略, 提出了一种新型的混合交通流演化仿真方法. 然后, 利用数值模拟方式对双车道交通环境进行仿真, 探索智能网联车对混合交通流的动态影响. 结果表明, 相比于元胞自动机构建的普通车辆智能体, 改进的 Q 强化学习方法训练的智能网联车智能体具备更强的连续时空环境适应能力, 双车道环境下道路通行能力随着智能网联车渗透率的提升而增大, 最高可提升 45.34%. 此外, 智能网联车渗透率的提高会降低车群低效的换道行为, 拓宽高通行能力水平下的车辆密度范围, 有利于改善交通拥堵.

关键词: 混合交通流; 协作博弈; 元胞自动机; 强化学习

中图分类号: U491

文献标志码: A

for CAV, to strengthen the capability of the simulation system in realistically reproducing CAV lane changing and car following behavior. Finally, a 2-lane freeway stretch was applied to our simulations, and with extensive simulation studies we obtained some promising results. The study results suggest that the impacts of CAV are quite positive. The inclusion of CAV considerably improves traffic flow, mean speed, and traffic capacity. Such understanding is essential for research concerning CAV as well as the CAV implication for future traffic management and control.

Key words: mixed traffic flow; cooperative gaming; cellular automata; reinforcement learning

Characteristics of Mixed Traffic Flow in Two-lane Scenario Based on Cooperative Gaming Method

GUO Jingqiu¹, FANG Shouen¹, QU Xiaobo², WANG Yibing³, LIU Yangzexi¹

(1. Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, Shanghai 201804, China; 2. Department of Architecture and Civil Engineering, Chalmers University of Technology, Chalmers 41296, Sweden; 3. College of Civil Engineering and Architecture, Zhejiang University, Hangzhou 310058, China)

Abstract: This paper aims to explore the impacts of connected and automated vehicles (CAV) on traffic flow efficiency based on in-depth microscopic simulation studies using cooperative gaming method. First, the Gipps car-following models were integrated into an improved cellular automata model to mimic the regular vehicle's driving behavior. Then, cooperative gaming method integrated with enhanced Q-learning was employed as the modeling platform

智能网联车 (connected and automated vehicle, CAV) 是近年来道路交通领域革命性的发展方向, 有望从微观行驶行为层面改善传统交通流特性^[1]. 自适应巡航控制 (adaptive cruise control, ACC) 和协同自适应巡航控制 (cooperative adaptive cruise control, CACC) 是 CAV 技术发展的重要阶段. 然而, 在未来相当长的时间里, CAV 的市场渗透率将逐步增长, CAV 将与普通车辆 (regular vehicle, RV) 长期共享有限的道路资源. CAV 环境下的交通调控和资源整合优化是一项极具挑战的课题. Chen 等人在研究自动驾驶车辆换道决策模型时, 通过层次分析法和逼近最优解的排序思想, 对普通的换道决策进行多属性赋值, 从而实现车辆换道安全和效率的平衡约束^[2]. Talebpour 等人在车联网环境下提出了一种基于博弈论的车辆换道决策模型^[3]. Meng 等人在此基础上, 结合结构平衡理论, 构建了滚动时域控制的博弈换道决策模型^[4]. 他们认为车辆换道决策问题可分解为换道价值和换道安全两个子问题, 并在应用博弈论对车辆间影响、换道安全和驾驶

收稿日期: 2018-07-07

基金项目: 国家重点研发计划资助 (2018YFB1600504)

第一作者: 郭静秋 (1977—), 女, 副研究员, 工学博士, 主要研究方向为智能车交通流及行为建模. E-mail: h1358552@163.com

通信作者: 刘洋泽西 (1994—), 男, 硕士生, 主要研究方向为道路安全. E-mail: 18508230695@163.com

效率综合考虑后给出换道决策。

然而,国内外学者在混合交通流特性研究方面还处于起步阶段。一方面,相比于 RV,CAV 具有更小的反应延迟时间,在行驶过程中与前车保持更小的车头时距,借此可以提升行驶速度;另一方面,CAV 具备与周围同类型车辆相互通信的能力,这一能力可以使得 CAV 在换道操作过程中获得更多信息,有助于生成并执行更加灵活、智能的决策。因此,CAV 有可能对提升道路通行能力发挥积极效能^[2-8]。此外,自动驾驶汽车可能会降低能源消耗和尾气排放,对低碳出行也有一定的推动作用^[9]。

目前,国内外对智能网联环境下的宏观混合交通流特性以仿真研究为主。宏观方面主要依靠不同的车队车辆间距、车辆换道策略分析混合交通流宏观特性^[6, 10]。然而,由于宏观模型通常在该问题上进行了大量的假设,容易使得分析结果与实际条件产生较大的差异。采用均衡交通流模型的文献多数基于流密曲线。微观行为分析是研究此问题的主流途径^[11-12]。通过考虑混合交通流的离散性,分解 CAV 及 RV 不同的跟驰及换道行为来进行仿真演化,并反应混合交通流的整体宏观特性。元胞自动机 (cellular automata model, CA) 是一种经典的 (微) 观交通研究基础模型,它能够通过制定简单的演化规则来有效地模拟并复现微观交通的非线性特征,从而被大量地作为基础模型并应用于各种特殊环境下的微观交通流研究^[13-18]。然而,由于 CAV 与 RV 是两种不同的智能体,传统的 CA 固定规则无法很好地描述 CAV 的智慧跟驰及换道行为,因此难以揭示出逼近真实的混合交通流特性。到目前为止,嵌入 CAV 智能性的混合交通流的仿真研究依然缺乏。

近年来,以强化学习为代表的人工智能领域迅速兴起,并在自然语言处理、图像识别等方面取得重大突破^[19-20]。强化学习是智能体以从环境状态中得到累积奖励值为目标而进行动作选择的映射学习^[21-23]。不同于元胞自动机规则化的行为选择,强化学习通过试错过程来进行最优行为策略映射。Q 学习是一种流行的免模型强化学习方法,通过值迭代的方式逼近马尔科夫决策过程中的最优策略,可以很好地体现 CAV 驾驶行为的不确定性及智能性。尤

其在 CAV 以车群行驶时,映射空间复杂,强化学习方法仍然可以在动作空间上进行无监督模式映射。

鉴于此,考虑一种结合元胞自动机及强化学习的多智能体混合交通流仿真模式。对于 RV,在 CA 强规则行为方式上加入 Gipps 跟驰模型进行更细致的改进^[24-26];对于 CAV,一方面为突出其驾驶行为的不确定性,另一方面为呈现其具备的更高的智能水平,因此通过基于改进 Q 学习来训练不同周围环境下的 CAV,以此训练形成 CAV 的非线性动态驾驶特性。在此基础上对混合交通流的宏观特性进行分析,并对该特性产生的影响进行总结。

1 研究背景

1.1 RV 演化模式

传统的 NaSch 元胞自动机模型遵循线性跟驰思想,认为驾驶员对速度的反应不会反应在跟驰距离上^[27]。之后的学者们对 NaSch 进行改进,揭示了非线性跟驰模型更能合理地反应真实交通状况^[28-30]。Gipps 提出的安全距离模型是一种常见的非线性跟驰模型,该模型认为车辆速度由当前理想速度、最大加速度和安全制动距离决定。考虑将 Gipps 模型引入 CA,即无论前方车辆是否为 CAV, $d_{\text{safe},n}$ 表示第 n 辆普通车与前车在任何时刻都应保持的最小安全跟驰间距。极限情况如图 1 所示。此时,

$$d_{\text{safe},n} = x_{n-1}(t) - x_n(t) - l = \mu \cdot v_n(t) + \frac{1}{2b}(v_n(t)^2 - v_{n-1}(t)^2) \quad (1)$$

式中: $x_{n-1}(t)$ 、 $x_n(t)$ 分别表示 t 时刻前方第 $n-1$ 车辆与本车位置; l 为车辆 n 的长度; μ 为驾驶员反应时间; $v_{n-1}(t)$ 、 $v_n(t)$ 分别表示前方 $n-1$ 车辆与该车在 t 时刻的速度; b 表示车辆 n 的最大减速度。设 l_{cell} 表示单元元胞长度,则在 CA 模型中车辆 n 在 t 时刻的最小安全跟驰间距 $d_{\text{safe},n}(t)$ 应为

$$d_{\text{safe},n}(t) = \left\lceil \frac{d_{\text{safe},n}(t)}{l_{\text{cell}}} \right\rceil \cdot l_{\text{cell}} \quad (2)$$

RV 在跟驰过程 $d_{\text{safe},n}$ 中根据调整下一时间步的车速来避免与前车发生追尾,即存在安全跟驰速度 $v_{\text{safe},n}(t+1)$ 如下:

$$v_{\text{safe},n}(t+1) = \min \left(\left\lfloor v_n(t) + \frac{2 \cdot 5a\mu}{l_{\text{cell}}} \left[1 - \frac{v_n(t)}{v_{\text{max}}} \right] \cdot \sqrt{0.025 + \frac{v_n(t)}{v_{\text{max}}}} \right\rfloor / l_{\text{cell}} \right), \left(\left(\mu b + \sqrt{(\mu b)^2 - b \{ 2[x_{n-1} - x_n - l] - \mu v_n(t) - \frac{2v_{n-1}(t)^2}{b_{n-1}(t) + b_{n-1}(t-1)} \}} \right) / l_{\text{cell}} \right) \quad (3)$$

式中, a 为车辆最大加速度, v_{max} 为车辆最大行驶速

度, $b_{n-1}(t)$ 表示前车在 t 时刻的减速度值。

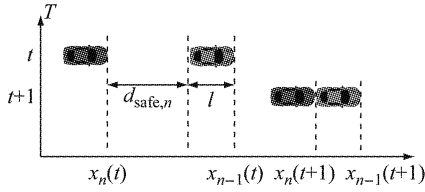


图 1 安全跟驰间距示意

Fig.1 The schematic diagram of safe car-following gap

1.2 基本更新规则

普通车 RV 依照 CA 模型的通用规则框架按序进行 $t \rightarrow t+1$ 更新. 每一规则均对应了特定的车辆操作.

(1)换道规则. 换道行为是车辆在多车道环境下常见的驾驶操作. 基于文献^[31]中的换道规则, 考虑当车辆 n 在式(4)~式(6)环境时会以一定的概率 p_{change} 进行换道操作, 即

$$d_n < \min(v_n(t) + 1, v_{\max}) \quad (4)$$

$$d_{n,other} > d_n \quad (5)$$

$$d_{n+1,other} > v_{n+2}(t) + \delta \quad (6)$$

式中: $d_{n,other}$, $d_{n+1,other}$ 分别表示旁车道前方及后方距离; $v_{n+2}(t)$ 为旁车道后方车 t 时刻车速. δ 衡量车辆 n 的换道操作水平^[32], δ 越大, 表现为越强制性换道, 即在考虑换道时对目标车道后方车辆的间距及速度的要求越低.

(2)加速规则. 车辆在行驶过程中, 当第 n 车辆在每个时间步开始时首先进行按常规加速度进行加速行驶估计. 该步骤速度仅反映驾驶员试图保持高速行驶的意图, 还需在接下来进行安全距离判断, 因此不作为最终速度.

$$v_n \rightarrow \min(v_{\max}, v_n + a) \quad (7)$$

(3)确定性减速规则. 传统 NaSch 模型设置方式不同, 该规则主要保证了车辆间应保持的安全距离. 当第 n 车辆与其前方车辆之间的距离小于该车行驶时所保持的安全距离 $d_{safe,n}$ 、或该车行驶速度在经加速规则后超过安全速度 $v_{safe,n}$ 时, 为确保安全驾驶则需要进行确定性地减速.

$$v_n \rightarrow \min(v_n, v_{safe,n}, d_n, d_{safe,n}) \quad (8)$$

(4)随机慢行. 考虑到驾驶员在行驶过程中可能存在的驾驶行为不稳定性, 在演化规则中引入随机慢化概率 p_{random} ($0 \leq p_{random} \leq 1$). 行驶过程中的车辆按照随机慢化概率进行速度的慢化以更真实反映驾驶员的行驶不确定因素.

$$v_n \rightarrow \max(0, v_n - 1) \quad (9)$$

(5)位置更新. 在速度演化更新规则的基础上, 进行车辆位置的更新.

$$x_n \rightarrow x_n + v_n \quad (10)$$

2 CAV 行为建模

如前所述, CAV 的驾驶行为设计应遵循比 RV 更智慧的跟驰及换道策略. 而目前大多数的 CAV 行为模型是在保证安全的条件下以自我利益最大化为目标、不考虑对周围车辆的影响的建模方式. 随着 CAV 渗透率的提高, CAV 与 RV、CAV 与 CAV 之间的动态交互将对车辆群体产生复杂的影响作用.

2.1 基于 Q 学习的训练方法

在强化学习领域, Q 学习系统是一种典型的离散人工智能学习系统. 在无需任何外界预先知识的情况下可以使学习主体(智能体)从零学起, 直至形成一套足够优化的映射规则, 因此可应用于 CAV 的行驶模式构建. Q 学习系统由 3 个方面组成^[33]: 环境 E 、动作库 A 和奖励值 r . 智能体在状态 S 下选择特定动作 A 的过程称为策略 π , 即 $\pi: S \rightarrow A$. 因此, 在 t 时刻时智能体在状态 s_t 时首先选择动作策略 a , 随后外部环境给予奖励, 智能体接收奖励并评估, 以此决定下一动作并进入下一状态 s_{t+1} . 累积奖励值 V 为未来奖励的折现, 回报折扣因子为 γ ($0 \leq \gamma \leq 1$). 智能体依靠累积奖励值的最大化, 进而由反馈机制引导其在连续时间点中采取智慧高效的动作. 设 $Q^\pi(s, a)$ 表示在状态 s 时根据策略 π 而执行 a 动作的值函数估计, 则

$$V^\pi(S_t) = \sum_{j=0}^{\infty} \gamma^j r_{t+j} \quad (11)$$

$$\pi^* = \operatorname{argmax}_{\pi} V^\pi(s) \quad (12)$$

$$Q^\pi(s, a) = r(s, a) + \gamma \max_{a'} Q(\delta(s, a), a') =$$

$$E^\pi \left[\sum_{t=0}^{\infty} \gamma^t r_t \mid s_0 = s, a_0 = a \right] \quad (13)$$

式中: j 为相对于时刻 t 的未来时间点; $\delta(s, a)$ 为状态转换函数. $Q^\pi(s, a)$ 的更新满足 Bellman 方程如下:

$$Q^\pi(s_t, a_t) = \sum_{s_{t+1}} p(s_t, a_t, s_{t+1}) \cdot r(s_t, a_t, s_{t+1}) + \gamma \sum_{s_{t+1}, a_{t+1}} p(s_t, a_t, s_{t+1}) \cdot Q^\pi(s_{t+1}, a_{t+1}) \quad (14)$$

式中: $p(s_t, a_t, s_{t+1})$ 为状态 s_t 时, 智能体采取动作 a_t 转移到 s_{t+1} 状态的概率; $r(s_t, a_t, s_{t+1})$ 表示动作 a_t 和状态 s_t 转移到 s_{t+1} 的回报值. Q 学习对应的最优动作估计 $Q^{\pi^*}(s, a)$ 和最优策略 $\pi^*(s)$ 为

$$Q^{\pi^*}(s, a) = \max_{\pi} Q^\pi(s, a) \quad (15)$$

$$\pi^*(s) = \operatorname{argmax}_{\pi} [r(s, a) + \gamma V^*(\delta(s, a))] =$$

$$\operatorname{argmax}_a Q(s, a) \quad (16)$$

综上所述, 可以总结基于 Q 学习的 CAV 训练

过程:首先,确定车辆的状态定义和动作选择集合,构建由不同状态和动作选择组合的二维 Q 表;其次,将 CAV 放入仿真环境运行,并混以不同比例的普通车辆,结合式(11)~式(16)迭代更新 Q 表,以形成车辆完整的状态-动作映射;最后,在正式仿真过程中,收集交通微观数据,统计宏观交通特性。

2.2 车辆状态定义

目前在 CAV 的主流仿真研究中,均假设了车辆具备一定的周边交通感知能力及协同能力^[34-35]. 因此,为体现 CAV 应有的智能水平,在跟驰和换道过程中除考虑自身行驶状态,还需要考虑本车所在车道的前方最近车辆 $n-1$ 、相邻车道前后方最近车辆 $n-2$ 、 $n+2$ 的车辆行驶状态,并认为以上 4 车的行驶状态决定了本 CAV 的行驶策略. 图 2 综合考虑以上多变量影响因素在车辆行驶过程中表现出的高度动态性,为了更好地模拟真实状态,车辆 n 通常需要考虑连续若干时间步的状态,并结合自身的最优行驶利益来决定下一时间步的行驶策略。

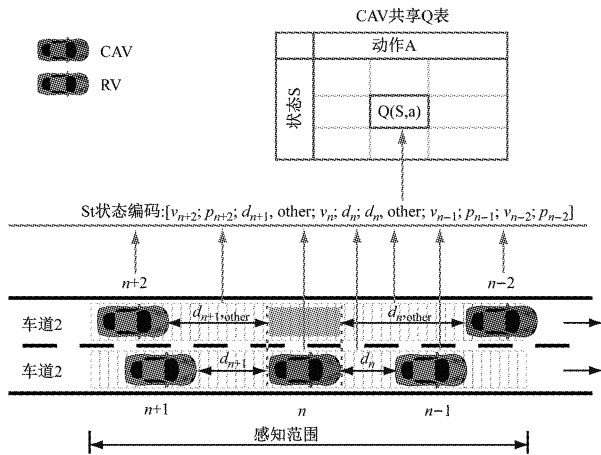


图 2 CAV 状态
Fig.2 CAV's state

根据以上车辆状态定义可以完成 Q 表. Q 表是状态-动作的价值表,纵轴长度等于状态数量,横轴长度等于动作数量. 设 CAV 车辆周围感知区域为车辆 n 纵向方向前后各 $2v_{\max}$ 的数值范围. 因此,第 n 辆 CAV 在时刻 t 时的状态 S 可表示为以下的 10 维向量:

$$S_n(t) = [v_{n+2}; p_{n+2}; d_{n+1, \text{other}}; v_n; d_n; d_{n, \text{other}}; v_{n-1}; p_{n-1}; v_{n-2}; p_{n-2}] \quad (17)$$

其中, p_i 表示 i 号位置对应的车辆类型 ($i \in \{n-1, n-2, n+2\}$, $p_i \in \{\text{CAV}, \text{RV}, \text{None}\}$). 若 i 号位置无车辆,则 $p_i = \text{None}$, $v_i = 0$. 可以看出,在双车道环境下,当第 n 辆 CAV 车辆在跟驰 CAV 或 RV

时,由于 p_{n-1} 取值不同,因此所对应状态表征也不同,据此可以做出不同的动作选择。

2.3 状态动作选择

一般情况下,车辆的动作空间 A_{all} 有 6 个不同动作,分别为:本车道减速“F-”、本车道保持车速“F=”、本车道加速“F+”、换车道减速“C-”、换车道保持车速“C=”、换车道加速“C+”. 为确保车辆间无碰撞无追尾等冲突发生,需要对 CAV 添加一定的先验知识,以避免缺乏合理性的模拟过程,从而显著提高学习效率. 如当 $d_n = 0$ 时车辆 n 不可能采取本车道加速的“F+”动作. 设车辆 n 在状态 S 时可行的非空动作空间为 $A_{\text{feasible}, n}(S)$, 且 $A_{\text{feasible}, n}(S) \in A_{\text{all}}$. 为了充分体现 Q 强化学习方法的在线学习性,采用 ϵ -贪婪策略选取即时动作,即车辆 n 处以 ϵ 的概率执行 Q 表中状态 S 的动作价值最大对应的动作,以 $(1-\epsilon)$ 概率随机执行动作,即

$$\pi(S) = \begin{cases} \operatorname{argmax}_a Q(S, A_{\text{feasible}, n}(S)), & \operatorname{rand}() \leq \epsilon \\ F(A_{\text{feasible}, n}(S)), & \text{其他} \end{cases} \quad (18)$$

其中, $\operatorname{rand}()$ 表示 $[0, 1]$ 中一个随机数, $F(\cdot)$ 表示随机选择函数. 奖励值的设置以行驶目标为准则. 基于所有车辆均以获得最大平均速度为行驶目标的假设,因此 Q 学习中的奖励应引导 CAV 尝试提速操作. 奖励值计算如下:

$$r = v_n(S') - v_n(S) \quad (19)$$

式中: $v_n(S)$ 表示车辆 n 在状态 S 时的车速,且 $S' = S \times \pi(S)$.

2.4 混合训练

CAV 与 RV 在仿真系统中的训练过程如图 3 所示. 由于混合交通流中 CAV 与 RV 共存,两种智能体分别由 Q 学习和 CA 构造,因此考虑对 Q 学习进行改造,取消 Q 学习中的周期,并将 Q 学习中的迭代步与 CA 的时间步训练演化策略相融合. 同时,系统中所有 CAV 共享 Q 表,以显著加速强化学习速度。

3 仿真与数值分析

3.1 仿真设计

仿真平台由 python 语言编写,以道路长度 $L=3$ km 的双车道作为仿真模拟环境. 为更细致地反应车辆在车道上的行驶性质,单元元胞长度 l_{cell} 设置为 1 m,车辆车身长度 l 为 5 m,即单车占用 5 个连续元胞. 车辆最大行驶速度 v_{\max} 为 $25 \text{ 元胞} \cdot \text{s}^{-1}$ ($90 \text{ km} \cdot$

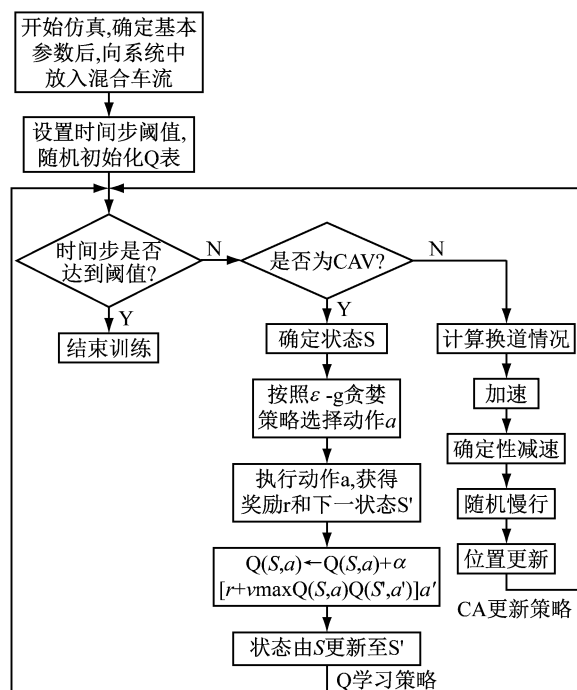


图 3 仿真模拟过程示意图

Fig.3 Schematic diagram of the simulation process

h^{-1}), 最大加速度 a 与最大减速度 b 分别设为 $5 \text{ 元胞} \cdot \text{s}^{-2}$ 、 $10 \text{ 元胞} \cdot \text{s}^{-2}$. RV 的换道操作水平 $\delta = \{-2, -1, 0, 1, 2\}$, 随机慢行概率 $P_{\text{random}} = 0.05$. 为简化分析维度、更大程度地揭示两种车型不同的微观行驶特性、提高仿真效率, 假设换道概率 $P_{\text{change}} = 1$, 即当车辆满足换道条件时便采取换道操作. 设 N 表示车辆总数, β 为 CAV 车辆渗透率, T 为有效仿真时长, 则车流平均速度为单位时期内所有车辆速度总和的平均值, 车流平均密度为每公里每车道平均的车辆数, 流量为单位时间内通过某一道路横截面的车辆数.

$$\bar{v} = \frac{1}{T} \sum_{t=t_0}^{T+t_0-1} \frac{1}{N} \sum_{n=1}^N v_n(t) \quad (20)$$

$$\rho = \frac{N}{2L} \quad (21)$$

$$\rho_i = \frac{N_i}{L} \quad (22)$$

$$Q = \rho \cdot \bar{v} \quad (23)$$

式中: i 为具体车道编号, 即 $i = \{1, 2\}$.

整个仿真过程分为训练过程及正式模拟过程. 在训练过程, 分别在不同密度不同 CAV 渗透率下运行 10^6 时间步用于训练并形成 CAV 的运行模式; 在正式模拟中, 每次演化时间步, 只保留最后 5 000 步作为有效稳定结果. 每种仿真环境均重复运行 20 次, 将每次仿真得到的车道平均密度、车辆平均速度及平均流量再次平均化并以此最终仿真结果, 用以

降低瞬时效应.

3.2 不同 CAV 渗透率下的交通流特征分析

图 4 反应了不同密度及 CAV 车辆渗透率对混合交通流特征的影响程度. 可以清晰看出车辆密度和 CAV 渗透率对混合交通流的通行能力及平均速度的影响效用. 从图 4a 可以看出, 对于一定的 β , 密度与车辆速度呈现反相关关系. 密度越大, 车辆速度越低, 并且当 $30 \text{ veh} \cdot \text{km}^{-1} \leq \rho \leq 40 \text{ veh} \cdot \text{km}^{-1}$ 时影响效果最显著. 另一方面, β 对速度的影响表现出了明显的非线性, 即 Q 学习下 CAV 与 CA 强规则的 RV 具有不同的演化方式. 当 ρ 在 $0 \sim 20 \text{ veh} \cdot \text{km}^{-1}$ 区间(车流稀疏)时, β 对速度的影响程度不大. 当 ρ 在 $20 \sim 60 \text{ veh} \cdot \text{km}^{-1}$ 区间(车流趋于拥堵), 且 β 在 $0 \sim 0.65$ 区间内时 β 对速度的影响程度较弱, 此时车流仍具有较大速度; 当 β 在 $0.65 \sim 1.00$ 时 β 对速度的影响程度加强, 表现为在同一密度下, β 越大, 车流速度越大; 当 ρ 在 $60 \sim 160 \text{ veh} \cdot \text{km}^{-1}$ 区间(车流处于轻微拥堵至较重拥堵状态), β 的提高显著减小了密度对车速的影响程度; 当 ρ 大于 $160 \text{ veh} \cdot \text{km}^{-1}$

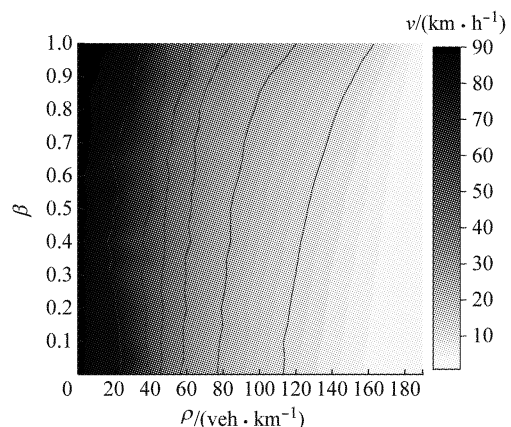
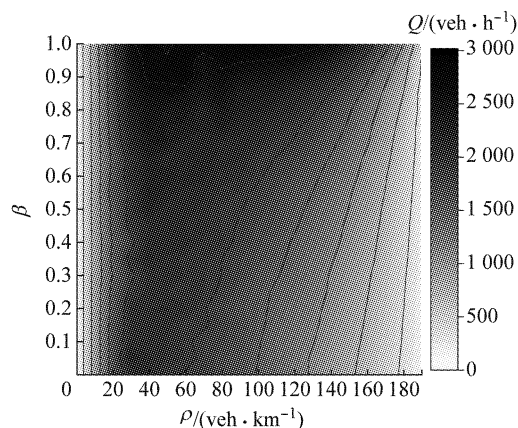
a β - ρ 对速度的影响b β - ρ 对速度的影响

图 4 渗透率及车辆密度对混合交通流的特征影响

Fig.4 Features of the mixed traffic caused by β and ρ

时,即交通处于严重拥堵, β 对车流速度的影响程度降低,但仍然满足正相关关系。

由式(23)可知,图4b与图4a的流量与速度在 β 与 ρ 的变化上具有相似特征,且由图5还可以看出,当 $\beta=0$ 时,道路最大通行能力 $Q_{\max}=2\,073\text{ veh}\cdot\text{h}^{-1}$;当 $\beta=1$ 时, $Q_{\max}=3\,013\text{ veh}\cdot\text{h}^{-1}$,即100% CAV的交通条件下通行能力提升了45.34%。此外,定义 $\Phi_{\eta}(\beta)$ 为在 β 一定时,密度 ρ 对应的车辆流量 Q_{ρ} 大于 $\eta\cdot Q_{\max}$ 的密度区间,即

$$\Phi_{\eta}(\beta) = \left[\arg \min_{\rho} [\rho \mid Q_{\rho} - \eta \cdot Q_{\max} \geq 0], \arg \min_{\rho} [\rho \mid Q_{\rho} - \eta \cdot Q_{\max} \geq 0] \right] \quad (24)$$

$\eta=0.85$ 时不同的 β 所对应的 $\Phi_{\eta}(\beta)$ 如图6所示。可以看出, β 有效地延长了道路高通行能力的适应密度。

由以上混合交通流特征分析可以看出,伴随CAV渗透率的提高,交通流状态有明显改善。分析原因,主要是:

(1)CAV允许更小的车头时距,CAV可以以更紧密的车队集合行驶;

(2)经过充分优化训练的CAV智能体对每个可选动作都事先加以评估,并选择最优驾驶行为,以期在动态交通环境中达到更大速度,从而提升整体交

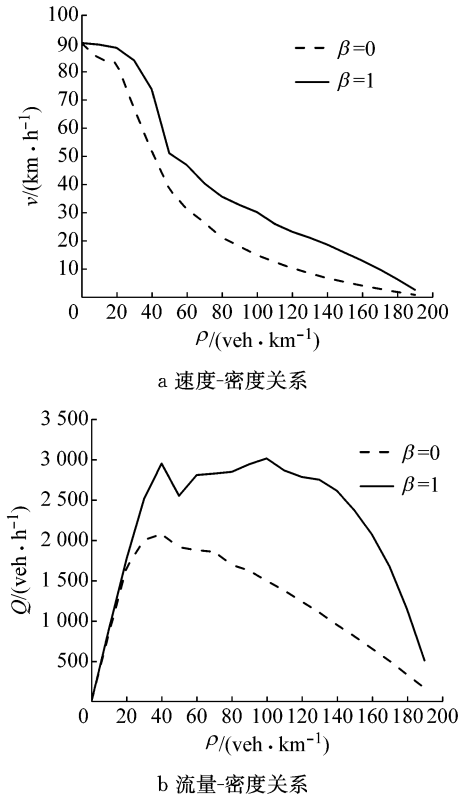


图5 同质性交通流环境

Fig.5 Fundamental diagram of homogeneous traffic

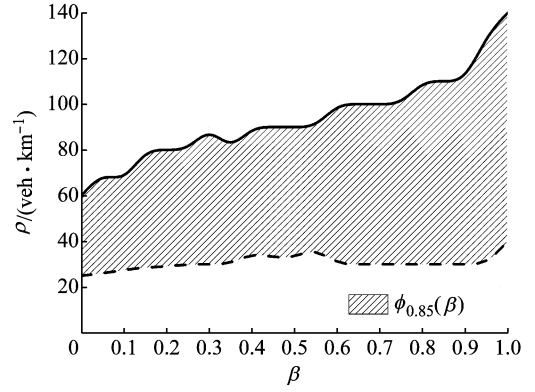


图6 $\Phi_{0.85}(\beta)$ 范围曲线

Fig.6 Scale diagram of $\Phi_{0.85}(\beta)$

通流的通行能力和平均速度。

3.3 换道频率分析

研究表明,频繁的换道是引发交通拥堵及事故的主要成因之一^[36]。换道操作改变了车辆横向稳定性,会对交通流产生重要影响。定义混合流换道频率 f_{LC} 为单位时间单位车辆的换道次数,由普通车辆及CAV车辆的换道频率计算得

$$f_{LC} = \sum_p f_{p,LC} = \sum_p \frac{N_{p,LC}}{T \cdot N_p} \quad (25)$$

式中: $N_{p,LC}$ 为有效仿真过程中 p 类型车辆的换道总次数; N_p 为 p 类型车辆数。仿真结果如图7所示。

一方面可以看出 f_{LC} 、 $f_{CAV,LC}$ 、 $f_{RV,LC}$ 在不同 β 下均呈现类基本图走势。在相同密度条件下, f_{LC} 、 $f_{RV,LC}$ 、 $f_{CAV,LC}$ 随 β 的增大而减小,且对于最大换道频率值,有 $f_{RV,LC}^* > f_{CAV,LC}^*$ 。这是由两方面因素造成的。当 β 增大时,一方面,CAV的驾驶特性使 $f_{CAV,LC}$ 降低的同时能以较高速度行驶,这种高速行驶压缩了RV换道安全条件成立的空间,降低了RV的换道供给,因此 $f_{RV,LC}$ 趋于降低;同时,CAV可以选择相邻时间步中最优的驾驶行为,并在跟驰过程中可以保持更小间距,相比于 $f_{RV,LC}$ 、 $f_{CAV,LC}$,能保持更低的换道需求水平,且能减小低效率的换道需求,即 $f_{CAV,LC}$ 表现为更显著的下降走势。

另一方面,随着 ρ 的增加, f_{LC} 、 $f_{CAV,LC}$ 、 $f_{RV,LC}$ 在不同 β 下均呈现类基本图走势。 ρ 越大,保持的换道频率水平越低。具体而言,当 ρ 低于转折点对应密度时,车辆间仍具有相对充足的空间进行自由换道操作,此时 f_{LC} 、 $f_{CAV,LC}$ 、 $f_{RV,LC}$ 与 ρ 呈现正相关关联性;当高于转折点对应密度后,受道路空间限制的趋势加强, f_{LC} 、 $f_{CAV,LC}$ 、 $f_{RV,LC}$ 表现为与 ρ 呈反相关。此外,相比于CAV,由于RV的换道条件对道路空间要求更高,因此 $f_{RV,LC}$ 表现出对 ρ 变化更加敏感。

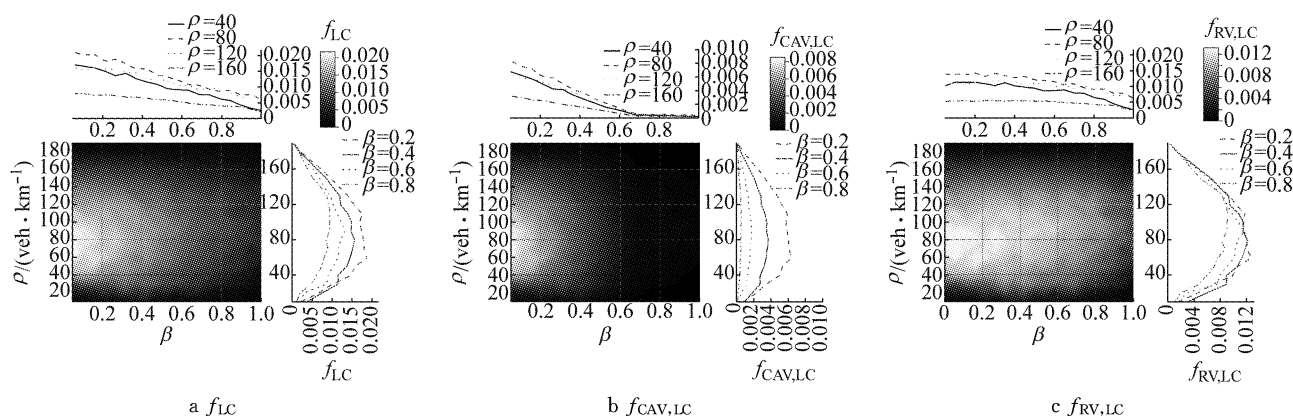


图 7 不同渗透率、不同密度的换道频率

Fig.7 Lane-changing frequency of different β and ρ

4 结论

通过探索一种双车道环境下的强化学习方法与元胞自动机相结合的演化机制,提出了基于改进的Q学习方法,精准模拟普通车和智能网联车辆的微观行驶策略,以此构建了一种针对双车道环境下混合交通流的高效仿真方法.此方法以个体优化为目标,探讨CAV微观驾驶行为所产生的集聚效应是否对交通流有优化作用,得到结论如下:

(1)相比于高度规则化的元胞自动机,强化学习形成的行驶策略具有更高的灵活性及相邻时空环境适应能力,更符合CAV的智慧行为特征;

(2)不同车流密度条件下,道路通行能力及车流平均速度可随着CAV渗透率的提高而增加,且维持高通行能力的密度范围也同步扩大,一定程度上延后了车流拥堵密度;

(3)不同车流密度条件下,随着CAV渗透率的提高,混合车流换道频率降低,车流横向稳定性增强.

由于采用的对称式双车道的道路仿真环境相对简单,对整体交通情况的刻画还不够贴近,因此可能与现实情况还存在一定差距.将来的研究工作需要进一步改进道路模型,也需要对更复杂的道路交通环境下的混合交通流特性进行深入研究.

参考文献:

- [1] SHLADOVER S E. Connected and automated vehicle systems: introduction and overview [J]. *Journal of Intelligent Transportation Systems*, doi: 10.1080/15472450.2017.1336053.
- [2] CHEN Jiajia, ZHAO Pan, LIANG Huawei, *et al.* A multiple

attribute-based decision making model for autonomous vehicle in urban environment [C]// *IEEE Intelligent Vehicles Symposium Proceedings*. [S. l.]: IEEE, 2014: 480-485.

- [3] TALEBPOUR A, MAHMASSANI H S, HAMDAR S H. Modeling lane-changing behavior in a connected environment: a game theory approach [J]. *Transportation Research Part C: Emerging Technologies*, 2015, 59: 216.
- [4] MENG F, SU J, LIU C, *et al.* Dynamic decision making in lane change: game theory with receding horizon [C]// *11th Ukacc International Conference on Control*. Belfast: IEEE, 2016: 1-6.
- [5] KHAN U, BASARAS P, SCHMIDT T L, *et al.* Analyzing cooperative lane change models for connected vehicles [C]// *2014 International Conference on Connected Vehicles and Expo*. Hanoi: [s. n.], 2014: 565-570.
- [6] LEVIN M W, BOYLES S D. A multiclass cell transmission model for shared human and autonomous vehicle roads [J]. *Transportation Research Part C—Emerging Technologies*, 2016, 62: 103.
- [7] MA J Q, LI X P, ZHOU F, *et al.* Parsimonious shooting heuristic for trajectory design of connected automated traffic part II: computational issues and optimization [J]. *Transportation Research Part B—Methodological*, 2017, 95: 421.
- [8] ZHOU F, LI X P, MA J Q. Parsimonious shooting heuristic for trajectory design of connected automated traffic part I: theoretical analysis with generalized time geography [J]. *Transportation Research Part B—Methodological*, 2017, 95: 394.
- [9] SHLADOVER S E, NOWAKOWSKI C, LU X Y, *et al.* Cooperative adaptive cruise control: definitions and operating concepts [J]. *Transportation Research Record Journal of the Transportation Research Board*, 2015, 2489(1): 145.
- [10] CHEN D J, AHN S, CHITTURI M, *et al.* Towards vehicle automation: roadway capacity formulation for traffic mixed with regular and automated vehicles [J]. *Transportation Research Part B—Methodological*, 2017, 100: 196.
- [11] 邱小平, 马丽娜, 周小霞, 等. 基于安全距离的手动—自动驾驶混合交通流研究 [J]. *交通运输系统工程与信息*, 2016, 16(4): 101.
- QIU Xiaoping, MA Lina, ZHOU Xiaoxia, *et al.* The mixed

- traffic flow of manual-automated driving based on safety distance[J]. *Journal of Transportation System Engineering and Information Technology*, 2016,16(4):101.
- [12] 秦严严,王昊. 智能网联车辆交通流优化对交通安全的改善[J]. *中国公路学报*, 2018(4): 202.
QIN Yanyan, WANG Hao. Improving traffic safety via traffic flow optimization of connected and autonomous vehicles[J]. *China Journal of Highway and Transport*, 2018(4):202.
- [13] FOTHERINGHAM A S, ROGERSON P. The SAGE handbook of spatial analysis[M]. [S.l.]:SAGE, 2009.
- [14] KAI N, SCHRECKENBERG M. A cellular automaton model for freeway traffic [J]. *Journal De Physique I*, 1992, 2 (12): 2221.
- [15] MAERIVOET S, MOOR B D. Cellular automata models of road traffic[J]. *Physics Reports*, 2005, 419(1): 1.
- [16] CHOWDHURY D, WOLF D E, SCHRECKENBERG M. Particle hopping models for two-lane traffic with two kinds of vehicles: Effects of lane-changing rules [J]. *Physica A: Statistical Mechanics & Its Applications*, 1997, 235(3/4): 417.
- [17] JIN S, QU Xiaobo, XU Cheng, *et al.* An improved multi-value cellular automata model for heterogeneous bicycle traffic flow [J]. *Physics Letters A*, 2015, 379(39): 2409.
- [18] PEDERSEN M M, RUHOFF P T. Entry ramps in the Nagel-Schreckenberg model [J]. *Physical Review E Statistical Nonlinear & Soft Matter Physics*, 2002, 65(2): 056705.
- [19] 陈启军,肖云伟. 基于行动分值的强化学习与奖赏优化[J]. *同济大学学报(自然科学版)*, 2007,35(4): 531.
CHEN Qijun, XIAO Yunwei. Action values based on reinforcement learning and optimized reward functions[J]. *Journal of Tongji University (Natural Science)*, 2007, 35 (4):531.
- [20] 段艳杰,吕宜生,张杰,等. 深度学习在控制领域的研究现状与展望[J]. *自动化学报*, 2016,42(5):643.
DUAN Yanjie, LÜ Yisheng, ZHANG Jie, *et al.* Deep learning for control: the state of the art and prospects [J]. *Acta Automatica Sinica*, 2016,42(5):643.
- [21] 高阳,陈世福,陆鑫. 强化学习研究综述[J]. *自动化学报*, 2004,30(1): 86.
GAO Yang, CHEN Shifu, LU Xin. Research on reinforcement learning technology: a review[J]. *Acta Automatica Sinica*, 2004,30(1): 86.
- [22] 宋梅萍,顾国昌,张国印. 随机博弈框架下的多agent强化学习方法综述[J]. *控制与决策*, 2005,20(10): 1081.
SONG Meiping, GU Guochang, ZHANG Guoyin. Survey of multi-agent reinforcement learning in Markov games [J]. *Control and Decision*, 2005,20(10): 1081.
- [23] 赵冬斌,邵坤,朱圆恒,等. 深度强化学习综述:兼论计算机围棋的发展[J]. *控制理论与应用*, 2016,33(6): 701.
ZHAO Dongbin, SHAO Kun, ZHU Yuanheng, *et al.* Review of deep reinforcement learning and discussions on the development of computer Go [J]. *Control Theory & Applications*, 2016,33(6): 701.
- [24] CIUFFO B, PUNZO V, MONTANINO M. Thirty years of gipps' car-following model [J]. *Transportation Research Record Journal of the Transportation Research Board*, 2012, 2315(1): 89.
- [25] GIPPS P G. A behavioural car-following model for computer simulation[J]. *Transportation Research Part B*, 1981, 15(2): 105.
- [26] PANWAI S, DIA H. Comparative evaluation of microscopic car-following behavior [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2005, 6(3): 314.
- [27] NAGEL K, SCHRECKENBERG M. A cellular automaton model for freeway traffic [J]. *Journal De Physique I*, 1992, 2 (12): 2221.
- [28] EMMERICH H, RANK E. An improved cellular automaton model for traffic flow simulation [J]. *Physica A Statistical Mechanics & Its Applications*, 1997, 234(3/4): 676.
- [29] LI Keping. Car deceleration considering its own velocity in cellular automata model [J]. *Communications in Theoretical Physics*, 2006, 45(1): 113.
- [30] LI X, WU Q, JIANG R. Cellular automaton model considering the velocity effect of a car on the successive car [J]. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 2001, 64(6): 066128.
- [31] LIU Y Z X, GUO J Q, TAPLIN J, *et al.* Characteristic analysis of mixed traffic flow of regular and autonomous vehicles using cellular automata [J]. *Journal of Advanced Transportation*, 2017,1705:1.
- [32] WATKINS C J C H, DAYAN P. Technical note: Q-learning [J]. *Machine Learning*, 1992, 8(3/4):279.
- [33] ZHU F, UKKUSURI S V. A linear programming formulation for autonomous intersection control within a dynamic traffic assignment and connected vehicle environment [J]. *Transportation Research Part C: Emerging Technologies*, 2015, 55:363.
- [34] LINGUO C, BAIGEN C, WEI S G, *et al.* Basic Simulation Environment for Highly Customized Connected and Autonomous Vehicle Kinematic Scenarios[J]. *Sensors*, 2017, 17(9):1938.
- [35] TOLEDO T, CHOUDHURY C F, BEN A M E. Lane-changing model with explicit target lane choice [J]. *Traffic Flow Theory*, 2005,1934: 157.
- [36] HIDAS P. Modelling vehicle interactions in microscopic simulation of merging and weaving [J]. *Transportation Research Part C—Emerging Technologies*, 2005, 13 (1): 37.