

混合交通环境下自动驾驶车在无信控交叉口运行特征仿真分析

郭雄峰

[上海市城市建设设计研究总院(集团)有限公司, 上海市 200125]

摘要: 针对混合交通环境下自动驾驶车在无信控交叉口的通行效率与安全问题,通过深度强化学习训练了两种车辆行为策略,在仿真环境中搭建了 4 类典型场景并进行测试。结果表明:两种策略下自动驾驶车通过无信控交叉口的平均成功率均达到 99.5% 以上,交叉口车辆的平均刹车时间均维持在较低水平;与基于碰撞时间的传统策略相比,平均通行时间分别缩短了 26.8% 和 18.1%,同时,平均碰撞比例维持在较低水平(0.18%、0.02%)。提出的基于深度强化学习的自动驾驶车行为策略能够高效、安全地协调无信控交叉口的通行,可为类似研究提供参考。

关键词: 混合交通环境;自动驾驶车;无信控交叉口;策略;仿真

中图分类号: U469.79;U412.35;TU997

文献标志码: A

文章编号: 1009-7716(2025)11-0007-04

Simulation Analysis of Operational Characteristics of Autonomous Vehicles at Unsignalized Intersections in Mixed Traffic Environments

GUO Xiongfeng

[Shanghai Urban Construction Design and Research Institute (Group) Co., Ltd., Shanghai 200125, China]

Abstract: Aiming at the traffic efficiency and safety issues of autonomous vehicles at the unsignalized intersections in mixed traffic environments, two vehicle behavior strategies are trained through deep reinforcement learning, and four typical scenarios are built and tested in the simulation environment. The results demonstrate that under both strategies, the average success rate of autonomous vehicles passing through unsignalized intersections reached over 99.5%, and the average braking time of vehicles at intersections remain at a relatively low level. Compared with the traditional strategy based on collision time, the average passage time has been shortened by 26.8% and 18.1% respectively. Meanwhile, the average collision ratio remains at a relatively low level (0.18%, 0.02%). The proposed behavior strategy for autonomous vehicles based on deep reinforcement learning can efficiently and safely coordinate the traffic at unsignalized intersections, which can provide a reference for similar research.

Keywords: mixed traffic environment; autonomous vehicles; unsignalized intersections; policies; simulation

0 引言

交叉口作为城市道路网络的重要组成部分,通常依赖信号灯组织车辆通行,车辆启停频繁,通行效率不高。随着自动驾驶技术的兴起,基于车路协同技术的交叉口信号优化控制研究较为广泛^[1-4],已有研究表明通过优化信号灯配时及自动驾驶车的速度,可以在一定程度上降低车辆行程时间及停车次数,提高交叉口通过能力^[5]。但信号控制引起的车

辆非必要降速仍不可避免,交叉口潜在通行能力并未完全发挥。

自动驾驶技术的发展,使得无信控交叉口成为可能。在无信控交叉口,车辆不再依赖外部信号指示,而是通过自身的智能系统和设备实时感知周围环境并与其他车辆进行信息交换^[6],依据制定的策略决定何时以及如何通过交叉口,从而优化交叉口的通行效率和安全性。宋述乐等^[7]建立了无信控交叉口车速引导模型,并设计仿真实验验证了提高交通效率的有效性;潘福全等^[8]提出了一种基于间隙理论的交叉口车辆控制策略,通过速度诱导车辆有序穿插、安全通过交叉口冲突区域,仿真结果表明所提策略相对信号控制方式效率显著提升;郭瑞军

收稿日期: 2025-05-30

基金项目: 上海市科委“科技创新行动计划”社会发展科技攻关项目(21dz1203804)

作者简介: 郭雄峰(1992—),男,硕士,工程师,从事道路交通设计研究。

等^[9]设计冲突检测及消解模型,引导非通行优先权车辆速度控制,降低了交叉口的冲突发生率;Dresner等^[10]提出了基于预约的交叉口车速控制方法,交叉口管理器(intersection manager, IM)依据冲突协调策略对预约车辆进行引导,保障冲突区域安全;但是基于预约的先到先得策略应对交通环境复杂、车流冲突严重的场景则尚不完善^[11];因此,丁红亮等^[12]提出了一种考虑车辆意图的分层式协同车速引导方法,仿真实验表明所提方法可进一步降低交叉口平均延误。上述研究较多针对完全自动驾驶环境,但从自动驾驶的发展情况来看,实现渗透率100%的纯自动驾驶环境仍需要较长时间,自动驾驶车辆(autonomous vehicles, AV)与人工驾驶车辆(human driven vehicles, HV)共存并交互的混合交通环境是未来相当长一段时间的常态^[13]。混合交通环境中,自动驾驶车辆通过车辆搭载的传感器、雷达感知和预测其他交通参与者的位置与行动意图,人工驾驶车辆则基于驾驶员的观察、感知和经验获取路况信息进而采取行动,在决策区域的冲突点上,两者的通行权模糊,已有研究中,自动驾驶车辆普遍倾向采用防御性驾驶策略,通过减速让行来规避冲突,虽然能够保证低事故率,但也大大降低了交通流运行效率及稳定性。

为解决混合交通流中自动驾驶车在无信控交叉口场景因过度保守策略而导致的通行效率低下问题,本文通过深度强化学习训练自动驾驶车辆在无信控交叉口环境的行为决策逻辑,并进行仿真研究不同行为策略下的通行安全、效率特征。

1 基于强化学习的自动驾驶行为决策模型

混合交通环境中车辆之间难以存在稳定的车间通信,基于碰撞时间(time-to-collision, TTC)的行为决策方法往往效率较低,本文使用深度Q网络(deep Q-network, DQN)来解决自动驾驶车辆在无信控交叉口的通行问题。

1.1 强化学习框架

强化学习是一种基于马尔可夫决策过程(Markov Decision Process, MDP)框架的机器学习范式,其核心机制在于智能体通过与环境的动态交互,逐步优化行为策略,最终实现期望累积奖励的最大化。MDP用元组 (S, A, P, R, γ) 表示,式中: S 为状态集合, $s_t \in S$; A 为动作集合, $a_t \in A$; P 为转移概率,在状态 s_t 下执行动作 a 后转移到状态 s_{t+1} 的概率; R 为

奖励函数,由状态 s_t 下执行动作 a 后转移到状态 s_{t+1} 的即时奖励; γ 为折扣因子,平衡即时奖励与未来奖励, $\gamma \in [0, 1]$ 。

MDP遵循马尔科夫假设,即:

$$P(s_{t+1}|s_t, a_t, \dots, s_0, a_0) = P(s_{t+1}|s_t, a_t) \quad (1)$$

强化学习的目标是让智能体通过与环境的交互,学习一个最优策略 π^* ,以最大化期望累积奖励:

$$\pi^* = \arg \max_{\pi} E \left(\sum_{k=0}^{\infty} \gamma^k r_{t+k} \mid s = s_t \right) \quad (2)$$

式中: r 为奖励, s_t 为 t 时刻的状态。

1.2 DQN 模型

Q-learning算法通过迭代更新值函数,求解MDP的最优策略,其目标是为每个状态-动作对 (s, a) 估计最优策略的期望累积折扣奖励 $Q^*(s, a)$:

$$Q^*(s, a) = E \left[r(s, a) + \gamma \max_{\pi} Q^*(s', a') \right] \quad (3)$$

式中: (s', a') 为执行动作 a 后的下一状态-动作对; $Q^*(s', a')$ 为对应的期望累计奖励。

考虑传统强化学习在处理大规模或复杂环境时的限制,可采用深度学习处理和解析大量的数据及其特征^[14]。在深度Q学习(Deep Q-learning)中,最优价值函数可以近似为一个含有参数 θ 的神经网络:

$$Q(s, a; \theta) \approx Q^*(s, a) \quad (4)$$

动作值函数的学习方法转化为将预期收益与网络预测的状态-行动值之间的误差迭代最小化,即:

$$L(\theta) = E \left[\left(y - Q(s, a; \theta) \right)^2 \right] \quad (5)$$

$$y = r(s, a) + \gamma \max_{a'} Q(s', a'; \theta^-) \quad (6)$$

式(5)及式(6)中: $L(\theta)$ 表示损失函数; y 为目标收益; $Q(s', a'; \theta^-)$ 为期望累计奖励。

1.3 行为策略

为模拟自动驾驶车辆在无信控交叉口交通场景中的不同“激进程度”,选择两种不同的交叉口的行为策略进行强化训练,并以TTC方法作为对照。

(1)时间到出发(time-to-go, TTG)策略。AV事先得得到行动路径,通过一系列“等待”或“出发”的操作来决定出发的时间。

(2)顺序行动(sequential actions, SA)策略。AV事先得得到行动路径,并在行动路径上自主决定加减速或维持恒定速度。

1.4 状态空间及奖励设置

整个交叉口的网络空间状态被离散化为笛卡尔坐标的网格空间。对于TTG策略和SA策略,分别以

18×26 网格及 5×11 网格进行表示,将道路空间离散化为车辆前方 0~20 m 以及两个方向各 90 m 处。空间中的每个自动驾驶车辆都由其航向角、速度和占有项表示,当车辆占据网格中的某一位置时,该占有指标项为 1,否则为 0。

TTG 策略和 SA 策略的每个仿真场景分别经过 25 万次及 100 万次的模拟训练。在网络训练中,安全损失(发生碰撞或减速度过大)设置为-100,成功到达目标点的奖励为 1,步长成本(效率损失)为-0.01。

2 仿真场景设置

2.1 交叉口一般化场景

车辆在无信控交叉口潜在冲突类型包括交叉冲突,换道冲突和追尾冲突。单车道十字路口作为基本平面交叉形式,能够较好反映车辆行驶过程中的冲突情况。本文基于 SUMO (Version 1.16.0) 仿真平台建立 3 个无信控单车道交叉口的混合交通场景,分别是单车道右转、单车道直行、单车道左转,同时,考虑到左转的复杂性,建立了 1 个双车道左转场景,如图 1 所示,其中,红色车辆代表 AV,黄色车辆代表 HV。

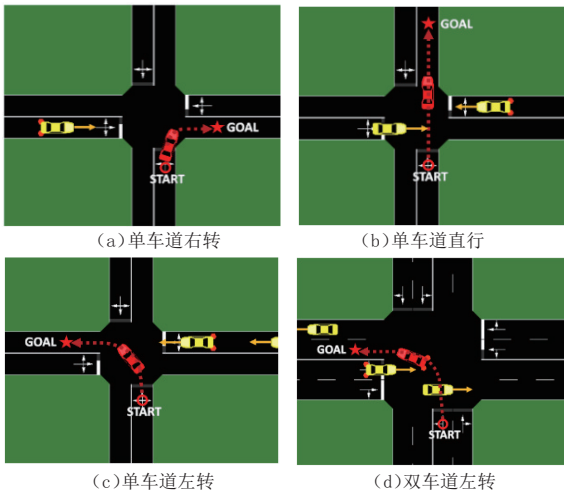


图 1 测试场景示意图

2.1 车辆行为参数

仿真车辆的换道模型采用 SUMO 嵌入模型“SL2015”,跟驰模型采用智能驾驶人模型(Intelligent driver model, IDM)^[15],表示为:

$$\frac{dv}{dt} = a_{\text{exp}} \left[1 - \left(\frac{v}{v_f} \right)^4 - \left(\frac{s_{\text{exp}}}{s} \right)^2 \right] \quad (7)$$

$$s_{\text{exp}} = s_0 + vt_{\text{safe}} + \frac{v\Delta v}{2\sqrt{a_{\text{exp}}b}} \quad (8)$$

式(7)及式(8)中: $\frac{dv}{dt}$ 为车辆输出加速度, m/s^2 ; a_{exp} 为车辆期望加速度, m/s^2 ; v 为当前速度, m/s ; v_f 为自由流速度, m/s ; s 为与前车的实际车间距, m ; s_{exp} 为与前车的期望车距, m ; s_0 为静止安全距离, m ; t_{safe} 为安全时距, s ; Δv 为与前车的相对速度, m/s ; b 为舒适减速度, m/s^2 。

通过遗传算法对 Waymo 自动驾驶车辆运行数据集中的不同跟驰对模型进行标定,确定了仿真中的车辆跟驰行为模型参数设置,如表 1 所示。在 SUMO 仿真中,通过 SpeedFactor 参数模拟不同车辆行驶通过交叉口的异质性。

表 1 车辆跟驰模型参数设置

车辆	t_{safe}	s_{exp}	a_{exp}	b
AV	1.55	5.22	1.87	1.02
HV	1.17	3.28	1.84	1.29

3 仿真测试

3.1 车辆行为假设

所有车道的最高时速定为 40 km/h,设定的 AV 在距离交叉口中心 50 m 处生成并以 10 m/s 的速度逼近交叉口,即 5 s 后到达交叉口中心,车辆的目标点是出口道距离交叉口中心 10 m 处,此时车尾已经完全离开交叉口范围,离开后车辆将以最大速度驶离交叉口,确保在该方向不会对后续的车辆在跟驰方面造成负面影响。

路网中其他方向的车辆均设置为 HV,在距离交叉口 50 m 处生成,遵循路线行驶,不存在临时停车与改变车道的行为,且假设其反应制动时间为 0.15 s。HV 初始速度随机,但是保证了在交叉口内车辆将会产生交互行为。

仿真步长设定为 0.2 s,实验设置了最大仿真步为 100 步,即 20 s。左右两个方向的交通密度由每秒发出的车辆概率设定,在实验中,车辆生成概率为 0.2,整个仿真周期中,进行交互的车辆期望为 4 辆(两个方向一共 8 辆)。

3.1 评价指标

为定量评估不同策略下的交叉口运行特征,定义了如下 4 项测试指标,并以传统 TTC 策略作为基线方法。

(1)成功率:车辆成功穿越交叉口到达目标点的数量比例。车辆发生碰撞、前一辆车未能通过交叉口而对后一辆车的跟驰状态造成影响均视为失败。

(2)碰撞比例:如果未能保证安全距离,或者为保证安全距离而产生的超过设定最大减速度的情况均视为一次“碰撞”。

(3)平均通行时间:车辆成功到达目标点,通行过程使用的时间。

(4)平均刹车时间:环境中其他车辆在运行过程中的刹车时间,体现本车对背景车流的干扰程度。

3.2 仿真结果

考虑到车辆异质性可能带来的未能泛化的一致性结论误差,每个场景至少进行了10 000次实验,4类交叉口仿真场景的运行特征指标如表2至表5所示。

表2 不同场景下交叉口通过成功率对比表 单位:%

场景	TTC策略	SA策略	TTG策略
右转	99.61	99.50	99.96
直行	99.91	99.76	99.78
左转1	99.70	99.99	99.99
左转2	99.42	99.79	99.99
均值	99.66	99.76	99.93

表3 不同场景下交叉口碰撞比例对比表 单位:%

场景	TTC策略	SA策略	TTG策略
右转	0	0.47	0.04
直行	0	0.14	0.01
左转1	0	0	0.01
左转2	0	0.11	0.01
均值	0	0.18	0.02

表4 不同场景下交叉口平均通行时间对比表 单位:s

场景	TTC策略	SA策略	TTG策略
右转	6.46	5.47	4.63
直行	6.19	4.40	4.63
左转1	6.97	5.26	5.24
左转2	7.59	7.13	5.40
均值	6.80	5.57	4.98

表5 不同场景下交叉口平均刹车时间对比表 单位:s

场景	TTC策略	SA策略	TTG策略
右转	0.31	0.88	0.45
直行	0.57	0.61	0.48
左转1	0.52	0.38	0.46
左转2	0.21	0.22	0.20
均值	0.40	0.52	0.40

所有场景中AV的通过成功率均达到99%以上,说明了基于DQN的策略的有效性。与预期一致,执行TTC策略的AV在任何场景下均没有发生碰撞,而基于DQN的两种策略均发生了一定数目的碰撞事故,这也是学习式策略不可避免遇到的问题,如何进一步提高模型泛化能力是要继续研究的问题。基于

DQN的两种策略在通过时间方面都比基于TTC的方法更短,说明了学习式策略在交叉口通行效率的优势。平均刹车时间对比方面,各策略下均维持在较小的区间,差异不大。

总体对比可知,TTG策略的平均成功率最高,且碰撞率远低于SA策略,在总共40 000次仿真测试中仅发生了7次碰撞,说明TTG策略较好地兼顾了安全性。执行TTC策略的AV虽保证了碰撞率为0,但在通行效率方面落后较大,执行TTG策略和SA策略的AV平均通行时间要比TTC策略快分别26.8%和18.1%,因此,提出的DQN方法具有提高交叉口通行效率的潜力。

4 结论

本文应用深度强化学习训练了自动驾驶车辆在无信号交叉口的行为决策模型,并与传统TTC策略进行了对比验证。基于SUMO平台的仿真测试结果表明:

(1)无论是传统TTC通行策略还是基于DQN的策略,均能有效引导自动驾驶车辆成果通过无信控交叉口。

(2)相较于传统TTC策略,基于DQN的通行策略能够有效提升自动驾驶车辆在交叉口的通行效率,并且基本兼顾交叉口运行的安全。

(3)研究针对两条双向两车道道路形成的十字路口,验证了深度强化学习在提升复杂无信控交叉口通行效能方面的潜力。但对于更复杂的交叉口类型仍需进一步研究。

参考文献:

- [1] MA W J, LI J J, YU C H. Shared-phase-dedicated-lane based intersection control with mixed traffic of human-driven vehicles and connected and automated vehicles[J]. Transportation Research Part C: Emerging Technologies, 2022: 135.
- [2] WANG H, MENG Q, CHEN S K, et al. Competitive and cooperative behaviour analysis of connected and autonomous vehicles across unsignalized intersections: A game-theoretic approach[J]. Transportation Research Part B: Methodologies, 2021(149): 322-346.
- [3] WU Y Y, CHEN H P, ZHU F. DCL-AIM: decentralized coordination learning of autonomous intersection management for connected and automated vehicles[J]. Transportation Research Part C: Emerging Technologies, 2019(103): 246-260.
- [4] 宋威龙,熊光明,王诗源,等.基于驾驶员类型分析的智能车辆交叉口行为决策[J].北京理工大学学报,2016,36(9):917-922.
- [5] XU B, JEFF B. Cooperative method of traffic signal optimization and speed control of connected vehicles at isolated intersections[J]. IEEE Transaction on Intelligent Transportation Systems, 2019, 20(4): 1390-

(下转第15页)