

文章编号:1006-2467(2021)08-1035-14

DOI: 10.16183/j.cnki.jsjtu.2020.387

专家论坛

智能车辆决策方法研究综述

胡益恺, 王春香, 杨明

(上海交通大学自动化系;系统控制与信息处理教育部重点实验室,上海 200240)

摘要: 结合目前国内外智能车辆决策方法的研究现状,分别从决策的输入、输出、周边环境交互方式以及算法类型4个方面对决策方法进行分类归纳、优缺点分析以及适用场景评估;总结归纳现阶段常见决策评估方法以及用于决策研究的数据集;分析现阶段决策方法所面临的技术难点以及未来发展趋势。

关键词: 智能车辆;决策方法;决策评估方法;数据集

中图分类号: TP 242.6

文献标志码: A

Decision-Making Method of Intelligent Vehicles: A Survey

HU Yikai, WANG Chunxiang, YANG Ming

(Department of Automation; Key Laboratory of System Control and Information Processing of the Ministry of Education, Shanghai Jiao Tong University, Shanghai 200240, China)

Abstract: Combined with the current research status of the intelligent vehicle decision-making methods at home and abroad, this paper classifies and summarizes decision-making methods from four aspects: decision input and output, environment interaction, and algorithm types. Besides, it analyzes their advantages and disadvantages, and evaluates applicable scenarios. Moreover, it surveys the common data sets and current evaluation standards which are used for decision-making researches. Furthermore it discusses the technical difficulties faced by current decision-making methods and future development trends.

Key words: intelligent vehicle; decision-making method; evaluation standard; dataset

智能车辆是轮式移动机器人的一种,其利用车载传感器来感知车辆周围环境,获取道路、车辆位置和障碍物等信息,生成车辆的转向与速度控制指令,从而能够安全、高效地到达目的地。经典的智能车辆系统框架由感知模块、定位模块、预测模块、决策模块、规划模块、控制模块以及车身底盘等组成^[1-2]。

智能车辆系统的决策模块输入有感知结果、定位结果、预测结果以及环境先验信息,根据决策方法,计算出合理的车辆行为决策,并发送给后续的执行模块。决策模块将影响智能车辆行驶的安全性与舒适性,是系统的重要中枢。

感知模块通过收集激光雷达、超声波雷达以及

收稿日期:2020-11-23

基金项目:国家自然科学基金(61873165, U1764264),上海汽车工业科技发展基金会(1807)资助项目

作者简介:胡益恺(1996-),男,安徽省合肥市人,硕士生,主要研究方向为机器人。

通信作者:王春香,女,副教授;E-mail: wangcx@sjtu.edu.cn.

摄像机等传感器的信息,获取动态、静态目标的位姿与速度,并对目标的类别进行检测.定位模块根据激光雷达、全球定位系统(GPS)以及车辆底层等传感器的数据,获取智能车辆自身的速度与位姿.预测模块根据感知模块计算获得的周边环境信息,对周边环境进行推演.决策模块作为系统的神经中枢,分别接受高精度地图的先验信息、感知模块以及预测模块的环境物体检测信息、定位模块的自车位姿信息,然后形成车辆行为决策.随后,将生成的决策交由架构下游的规划模块.规划模块根据车辆的决策生成合理的轨迹,随后控制模块根据自身的定位信息,向底层线控输出方向盘转角及油门刹车大小等指令,实现对车辆的横向控制以及纵向控制.

现阶段已有部分学者对智能车辆的决策方法进行了归纳与整理.文献[3]对具体的决策系统方法进行了归纳与介绍,其分别总结了 Boss、Junior、Odin、Talos 等系统的决策方法.然而,该文献从决策系统的角度对决策方法展开归纳,并未对具体决策方法的类型进行分类,并且该文献所包括的决策方法多为传统方法,对现阶段常见的学习类、博弈类等决策方法的论述较少.文献[4]对智能车辆的决策规划方法进行了归纳,其将车辆的决策方法分为基于博弈理论、基于概率、基于部分可观察的 Markov 决策过程以及基于学习 4 个类别,该文献清晰地对现有决策工作进行了总结.然而,该综述并未对基于学习类的方法进行进一步细分介绍,并且基于博弈的决策方法与其他三类决策方法并未形成鲜明的对立关系,存在一定的重合性,从而该文献的分类标准需要进一步优化.文献[5]将智能车决策方法分类为基于规则的决策方法以及基于学习的决策方法,并将学习类算法进一步细分为深度学习方法与其他学习方法.该文献的分类方式较为概括,不利于决策研究者快速了解现阶段的决策方法及研究进展.通过对上述现有决策研究综述的分析发现,其均缺少对不同类型决策方法的适用场景的分析,且目前对于如何评估决策方法的性能以及如何制定合理的评估指标,缺乏文献对其进行归纳与总结.同时,现阶段尚未有文献对智能车辆决策方法研究常用的数据集进行整理与归纳.

考虑到从信息输入到决策生成的信息流链路,本文将从决策模块的输入、输出、车辆周边环境交互方式以及决策模块所采用的算法类型 4 个方面对现有决策科研工作进总结与归纳,从而可以为后续对车辆决策方法进行研究的科研人员提供指引与帮助.

1 决策模块的输入

输入信息根据是否经过计算处理可以分为,其他模块计算结果信息与车载传感器原始数据两类.其中,其他模块计算结果信息主要分为车载感知模块计算信息与路测感知模块信息,其常见形式为本车坐标系下的环境元素的位置、速度、加速度等;而车载传感器原始数据信息则是将车载传感器如视觉传感器的图像数据、激光雷达点云数据直接输入至决策模块.

1.1 车载感知模块计算结果

以车载感知模块计算结果作为模块输入的决策方法,由于遮挡和传感器噪声等因素,对其带来了感知的不确定性和部分可观的特性.感知模块计算结果常见形式为,由视觉传感器、激光传感器等感知信息融合后得到的环境中物体的空间位姿、包围盒尺寸、物体速度加速度以及物体类别.

文献[6]以部分可观信息作为决策模块的输入,运用部分可观察 Markov 决策过程(POMDP)方法实现车辆的决策.文献[7]运用真实的部分感知数据,结合概率估计方法,设计了高速公路场景下针对部分可观环境的决策方法.文献[8]同样运用实车测试中得到的传感器处理数据作为模型输入,训练了基于分层状态机的智能车辆决策模型.

以车载感知模块计算结果作为输入的决策方法,将感知模块与决策模块分离,有较强的可迁移性及实用性.但传感器感知范围有限,且存在遮挡问题,因此感知模块输出的信息具有一定的局限性,从而对决策方法的稳定性提出了挑战.

1.2 路侧感知模块信息

与车载感知模块计算结果所不同,以路侧感知模块信息作为输入的决策方法,常出现于学术研究成果与“车与任何事物的联系”(V2X)场景中.该类型的输入常将环境中参与者的位置、速度、加速度等信息,直接或间接地转换为相对于智能车辆的位姿信息.

路侧感知模块信息常常由高空摄像机或无人机拍摄获得,或者由仿真模拟器直接生成.在这种场景下,智能车辆可以获得所有周边车辆相对于其的位置、速度等信息.文献[9]以高空摄像机生成的下一代仿真(NGSIM)数据集作为模型的输入,运用 pAC (Passive Actor-Critic) 方法完成车辆汇流时的决策.文献[10-11]同样基于完全可观数据,结合博弈论模型,解决了智能车辆变道场景下的决策问题.文献[12]运用路侧感知模块信息作为决策模块的输

入,在强化学习(RL)技术的基础上,提出了一种合作的变道策略.

除了通过高空相机获得路侧感知模块信息,现阶段部分研究成果采用仿真器生成的路侧感知数据作为决策模型的输入.文献[13]以三车道仿真平台的完全可观数据作为系统输入,运用强化学习方法实现速度决策以及变道决策.文献[14]结合基于规则的约束,将基于深度 Q 网络方法用于自动驾驶车道变更决策任务.文献[15]同样以三车道仿真平台的观测数据作为模块输入,提出了一种基于进化学习的智能车辆决策方法.

路侧感知模块信息有着较为直观的特性,属于理想场景下的源数据形式,可以有效改善由于遮挡、噪声等带来的信息不完全可观问题.在新算法的开发与早期验证、以及在不同模型之间进行性能比较的场景中有着较大的优势.然而在真实的车辆行驶环境中,由于基础设施尚未全面部署,难以获得全面的路侧感知信息数据.但随着通讯技术的不断增强,车联网技术逐渐完善,该类决策方法将会得到更为广泛的应用.

1.3 车载传感器原始数据信息

车载传感器通常有激光雷达、毫米波雷达、摄像头,由传感器获得的数据有着数据形式贴合实际车辆行驶环境,数据信息量丰富等特点.

现阶段存在部分决策方法直接以传感器的原始数据作为决策方法的输入信息.文献[16]以视觉传感器获得的图像信号作为系统输入,运用端到端算法设计车辆决策模块,使车辆在即将驶入的路口作出合理的决策.文献[17]基于递归神经网络(RNN)以及生成的对抗网络设计了端到端决策模块.文献[18-20]以车辆摄像头的图像信息作为决策模块的输入,训练出决策指令.文献[21]也同样采用了车载传感器的原始数据作为输入.

以车载传感器原始数据信息作为输入的决策方法是端到端算法的有效应用,其可以获得环境感知的全部信息,并可以提取到环境中影响决策的深层特征.但该种类方法对训练数据要求较高,并且对决策方法的可解释性提出了挑战.

2 决策模块的输出

按照决策方法输出信息的类型,可以将其归纳为离散决策输出以及广义决策输出.离散决策输出的算法所生成的指令为独立的动作,例如左变道、右变道、保持直行等等.广义决策输出的算法,亦可理解为端到端方法.该方法直接生成决策控制序列,

例如方向盘的转角时间序列、油门的时间序列、刹车的时间序列指令.决策模块的输出示意图如图 1 所示.其中: $t_i(i=0,1,\cdots,N)$ 为决策序列中的时间戳.

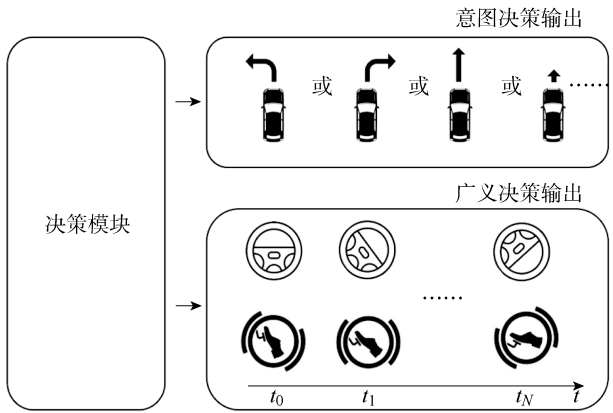


图 1 决策模块输出示意图

Fig. 1 Schematic diagram of output of decision module

2.1 意图决策输出

意图决策输出的指令集,常具有较高的抽象性,例如是否换道、是否避让等,在生成该指令后,需传递给规划模块与控制模块以生成具体的局部路径与控制指令.

文献[22]利用 Bayesian 网络,将决策层的输出划分为采用自适应巡航控制(ACC)跟车策略、向左变道、向右变道以及保持当前速度 4 种离散决策以决定智能车辆是否进行超车行为.文献[13]将强化学习决策模块的输出根据速度进行了更细一步的划分.文献[8]提出的分层状态机决策方法以离散的决策作为输出.文献[23]基于强化学习的决策研究与文献[24]基于 POMDP 的研究同样以离散决策作为模块的输出.类似的,文献[15]将决策模块的输出空间分为保持直行、左变道以及右变道 3 种离散决策.文献[25]基于长短期记忆网络-条件随机场(LSTM-CRF)方法,为车辆变道场景做出决策,决策模块的输出为保持直行、准备变道、左变道和右变道 4 种决策.文献[26]基于博弈论方法,生成 6 个离散的加速度,以实现智能车辆与环境车辆在路口的行为决策.

智能车辆所面临的决策问题可以概括为一个在复杂条件下的优化问题,而意图类型的决策输出让智能车辆在充满非凸问题的复杂环境中,获得将问题转化为凸问题再求解的能力.同时,此时的输出使决策模块更加稳定,因为和单纯的数值优化求解器相比,单纯的数值优化方法很难保证每一帧的解是相对稳定的,但是意图决策输出可以保证决策的连

续性和稳定性. 除此之外, 意图输出也简化了决策模块的输出数据结构, 成为了众多决策方法研究的输出形式. 但意图输出同样存在不足之处, 现实车辆行驶环境较为复杂, 通过离散决策难以枚举全量解空间, 所以可能会丢失最优解.

2.2 广义决策输出

基于广义决策输出的决策方法, 将输出连续的决策指令, 即输出方向盘转角、油门与刹车等控制量的时间序列. 此种决策方法涵盖了智能车辆框架中的决策模块、规划模块以及控制模块的功能, 是端到端算法的应用实现.

文献[27]运用卷积神经网络(CNN)训练了决策模型, 并以方向盘转角作为决策模块的输出. 文献[28-29]将车辆的方向盘转角作为决策模块的输出. 以广义类型作为输出的决策方法, 在解空间上有更广泛的遍历, 该特性赋予了决策方法更多求解优化的空间, 从而可以得出更加贴近实际的决策. 但广义类型的决策输出具有较高的意图不确定性, 如何通过可解释性规则进行约束以提高决策方法的可解释性以及安全性是当下研究的一个热点方向.

综合第 2.1 与 2.2 节内容, 依据输入输出类型, 智能车辆决策方法分类如表 1 所示.

表 1 依据输入输出类型的决策方法分类
Tab. 1 Classification of decision-making methods based on input and output types

类别	优势	劣势	文献编号
路侧感知模块输入	全局可观感知信息	对基础设施有高要求	[6]、[9-15]
感知计算结果输入	数据易处理	丢失丰富的感知信息	[8]
车载原始信息输入	拥有丰富的感知信息	感知范围受限易遮挡	[16-20]
意图决策输出	具有明确的决策意图	难以枚举现实场景	[8]、[13]、[22-24]
广义决策输出	遍历更广的解空间	决策意图不明确	[19-20]、[27]、[30]

3 决策模块与环境的交互方式

除了依据信息输入与输出形式进行分类外, 决策模块与环境的交互方式亦可作为决策方法归纳的重要指标之一. 智能车辆在行驶过程中所进行的决策过程, 可视为车辆与环境进行交互的过程. 现阶段智能车辆决策方法的研究已经度过了简单场景的阶段, 开始聚焦于与环境具有冲突的复杂场景的决策方法研究. 不同的环境交互方式所形成的决策方法在进行决策时会有较大的差异, 本节将从两种环境交互方式来进行归纳分析.

3.1 博弈形式

基于博弈论的决策方法认为智能车辆与环境中的周边车辆存在博弈关系, 常见的建模方式为将存在交互行为的所有决策个体视为博弈中的参与者, 各方的状态改变为博弈策略, 以此构建双方的收益矩阵, 通过求解博弈模型的 Nash 均衡, 作为双方的最优驾驶策略组合. 现阶段研究常用的博弈参与者收益由自车的安全性、效率以及舒适性等指标构成^[26].

文献[10]通过在驾驶员之间运用博弈论方法, 解决了高速公路匝道口汇入的困难场景问题. 文献[26]通过对路口场景的车辆进行建模, 运用加减速决策对建立 Nash 矩阵, 并生成最终决策. 在 Nash 矩阵中, 每一个决策者的收益值会受其余所有

决策者的决策影响, 所以决策 Nash 矩阵的维度将随着决策者数量的增加呈指数型增长, 这将带来计算时间的指数型增加. 为了解决该问题, 可以采用基于树搜索的算法. 文献[31]运用蒙特卡洛树(MCT)加速最优值的搜索. 文献[30]引入了车辆的运动学、动力学限制, 通过递归的方式, 减少了搜索的空间. 文献[28]则通过引入 Stackelberg 博弈来减少智能车辆自身的决策空间. Stackelberg 博弈为两阶段的完全信息动态博弈, 博弈双方都是根据对方可能的策略来选择自己的策略以保证自己在对方策略下的利益最大化, 从而达到 Nash 均衡. 在该文献中让智能车辆作为该博弈场景中的先导决策方, 率先作出即使所有环境车辆做出最恶劣举措下的自身最优决策. 对于其他环境中的参与者, 算法假设其均采取相同的策略, 然后系统通过迭代直到收敛. 该方案将决策方法的计算量减少为随决策参与者线性增长, 实现了运算的加速.

博弈论的思想已融入到众多决策方法中, 多策略决策(MPDM)方法将博弈形式作为决策方法的基础. 该方法为环境中的所有参与者设置了相同的策略空间, 对全体参与者所有可能的选择的决策组, MPDM 设计了一个类似于模型预测控制(MPC)的推演窗口, 计算出不同的子策略集所对应的全局损失函数 cost, 然后在 cost 集里面选择 cost 最小值所对应的决策集作为当前的最优决策^[29]. 文献[32]针

对动态不确定的智能车辆环境,设计了基于 MPDM 的决策方法. 文献[9]将 MPDM 与 pAC 结合,运用 MPDM 计算汇流的候选点,使 pAC 的总体成功率达到了与人类驾驶员相当的 92%.

基于博弈的决策方法,是建立在环境中的决策参与者均为理性决策者的基础上的一种方法. 当智能车辆在真实场景中,遇到非理性驾驶员时,该方法常常会做出错误的决策. 如何对决策者为理性决策者和非理性决策者进行建模分析,如何增强博弈决策算法的稳定性,是基于博弈论的决策方法需要面临的挑战.

3.2 礼让形式

与博弈形式所不同,礼让形式的决策方法更加倾向于以对环境造成影响较小的最优决策. 文献[33]提出一种最大交互防御策略(MIDP),以完成车道汇流场景下的安全决策. 文献[34]将环境车辆的驾驶行为建模为朴素贝叶斯(Naive Bayesian)的形式,智能车辆首先对当前环境车辆进行观测,如果环境车辆加速,那么该车辆礼让智能车辆的可能性较低;如果减速,则该环境车辆有较大概率礼让. 基于此模型,该论文聚焦于高速公路的汇入车流场景,设计出礼让形式的决策方案. 智能车辆对不同的决

策所带来的代价函数进行计算,然后选择最佳策略. 文献[35]使用粒子滤波器来估计环境中智能驾驶员模型的行为参数,例如最大加速度、期望加速度、期望速度、最小间隔距离等. 随后,该决策方法基于此模型,根据当前观测,推演未来空间内各物体的位姿,从而获得最优决策. 文献[36]将环境中的障碍物等信息以语义序列的方式提取出来,然后通过优化目标函数进行决策.

采用礼让形式的决策方法,可以理解为与环境的单向交互,此类型的决策方法通常基于对历史数据中先验信息的提取,根据对未来场景的推演,生成智能车辆当前帧的决策以达到最优目标函数. 该方法较依赖于数据,并且缺乏对环境的主动交互,在长时域对环境车辆的驾驶行为以及运动轨迹的预测不足,短时域内可能会陷于局部最优,从而对决策方法的稳定性带来了挑战. 除此之外,礼让形式下的决策方法在交互场景中更倾向于选择保守的决策,该特性虽然可以有效地提高车辆的安全性,但是在车流量较高的复杂场景中常陷入举步不前的决策结果.

依据与环境的交互方式,智能车辆决策方法分类如表 2 所示.

表 2 依据与环境交互方式的决策方法分类
Tab. 2 Classification of decision-making methods based on interaction with the environment

类别	优势	劣势	文献编号
博弈形式	与环境存在主动交互	假设全为理性决策者	[10]、[26]、[28]、[30]
礼让形式	车辆安全性较高	车辆智能度偏低	[9]、[29-32]

4 决策算法的类型

机器学习是现阶段决策算法研究中常用的方法,是否基于机器学习方法以及所采用算法的类型是对决策方法进行总结归纳的一个重要依据. 通过分析已有的科研工作,本小节将分别从非学习类方法与学习类方法两大类进行介绍.

4.1 基于非学习类的方法

非学习类的决策方法又称为传统决策方法,其主要代表为有限状态机(FSM)方法,该方法由于其稳定性在现阶段仍得到了广泛的应用. 基于状态机的决策方法有着结构简单、可解释性强的特点. FSM 是一种离散输入、输出系统的数学模型,其由有限个状态组成,当前状态接收事件并产生相应的动作,进而引起状态的转移. 状态、事件、转移、动作是有限状态机的四大要素. 根据状态分解以及连接逻辑,将其分为串联式、并联式、混联式 3 种体系架

构^[5].

文献[37]运用决策树方法对智能车辆决策模块进行建模. 决策树方法可以视为状态机方法的抽象方法,决策树方法将状态机中生成的决策以树状形式维护起来,当车辆处于不同的决策场景中时,通过计算车辆状态,在决策树中检索具体的最优行为决策. 文献[8]提出一种基于层次状态机的智能车辆换道决策方法,该决策框架由两部分组成:微型场景信息模型和基于多属性决策的车道变更行为的决策模型.

基于状态机的决策方法拥有清晰的决策意图,从而具备较强的可解释性,但其难以涵盖车辆真实行驶环境中的边缘场景,难以保证决策方法在真实复杂多变的驾驶环境中的稳定性.

4.2 基于学习类的方法

基于学习的决策方法随着车载计算单元计算能力的增强,愈发得到学术界与工业界的关注. 现阶段

学习类的决策方法主要可以分为基于深度监督学习、基于强化学习以及基于逆强化学习(IRL)3种,本节将以此进行分类与介绍。

(1) 基于深度监督学习的方法。基于深度监督学习的决策方法主要形式为从大量的标签驾驶行为数据中学习环境信息到车辆决策的有效映射关系。由于深度监督学习已经在图像等领域得到了广泛的应用与验证,该类型在学习类方法中率先被应用于智能车辆的决策方法研究中。

现阶段已有众多国内外科研机构基于深度监督学习方法对车辆决策进行了研究。文献[38]提出基于注意力机制的车辆决策方法,以实现在拥挤车流中的决策。文献[39]运用多任务学习(MT-LfD)框架实现对环境车辆的决策预测。同样基于监督学习开展科研工作的还有英特尔公司的端到端决策模块^[16],Comma.ai公司所提出的基于RNN的端到端决策模块^[17]。文献[25]以NGSIM数据集作为训练数据,设计了基于长短时记忆(LSTM)神经网络和条件随机场(CRF)模型的类人决策算法,提高了车辆换道决策的成功率。相似的研究还有百度Apollo团队的研究成果^[40],其运用基于深度监督学习的决策方法,实现了对智能车辆的横向控制与纵向控制。Apollo团队运用CNN实现车辆的横向控制,该决策模块以前视摄像机的图片作为决策方法的输入,决策模块直接计算出方向盘的期望偏角,以实现路径的跟踪。同时,Apollo团队采用LSTM模型,从图像时间序列中获取油门与刹车的控制量。

基于深度监督学习的决策方法,依赖于数据集的质量,其可以在特定场景下通过对已有专家数据的学习,做出精准、高效的决策。然而此类型方法仍然存在一些不足,例如当训练数据集出现不均衡的情况或者数据集缺少难例样本时,决策方法难以根据学习到的神经网络做出正确的决策。同时,由于深度监督学习基于深度神经网络以实现车辆决策的计算,其模型的可解释性较低,难以独立应用于车辆的决策系统。

(2) 基于强化学习的方法。强化学习与监督学习不同,监督学习中的决策方法从数据集中直接获得正负样本的标签,而强化学习模型须通过在解空间内的反复试错,才能获得最优模型^[41]。

现阶段已有众多研究成果运用强化学习方法方法生成离散决策或连续决策^[42-45]。文献[46]基于奖励函数和更新函数开发了增强型Q学习算法。文献[47]基于深度强化学习(Deep RL)设计了决策训练

器以进行智能车辆决策训练。文献[48]针对车辆驾驶性能的多目标问题,提出了基于最小二乘策略迭代(LSPI)的多目标强化学习(MORL)算法,并且应用于车辆智能驾驶决策问题。文献[49]将三维点云以及摄像机图片作为决策模块的输入,通过基于多重奖励架构(MRA)强化学习方法,实现车辆在高速公路场景下的控制与决策。随着解空间构型的逐渐复杂,强化学习的学习过程可能会急剧加长甚至不再收敛。现阶段RL研究工作者分别提出了课程式学习^[50-52]与对抗学习^[53-54]。

此类方法的主要优点是该类方法在训练时不需要标注数据集,从而降低了标注的成本。除此之外,其可以处理非结构化数据,例如原始或经过稍微预处理的雷达或基于摄像机的图像信息。然而该方法仍有不足之处,其学习与训练的目标依赖人为设计的奖励函数,该奖励函数设计的好坏将直接决定强化学习决策方法的性能。同时,由于强化学习需要在仿真器内进行迭代训练,仿真器对现实工况的模拟程度将影响决策方法从仿真环境到现实环境的可迁移性。

(3) 基于逆强化学习的方法。针对强化学习需要人为设置奖励函数的弊端,逆强化学习通过从专家演示中学到未知的奖励函数来解决该问题。

文献[55]在高速公路仿真中学习不同驾驶风格,此工作展示了具有未知奖励函数的Markov决策过程在车辆行为决策中的优势。为了避免逆强化学习的过拟合,现阶段研究采用最大熵IRL方法来解决过拟合问题。文献[56]应用了最大熵IRL的原理,因为最大熵分布显示出对数据的最少承诺,这是避免过度拟合的自然选择。文献[57]设计了一个风险敏感的IRL框架,能够明确考虑专家的风险敏感度。该框架能够捕获从中性风险到最坏情况的不同风险偏好范围。基于线性编程的算法可用于推断专家的隐藏风险指标。文献[58]提出了最大熵深度IRL框架,利用深度完全卷积神经网络的表达能力表示潜在的驾驶行为成本模型。

基于逆强化学习的决策方法,弥补了强化学习方法中人为设计奖励函数时的缺陷,是现阶段决策研究的一个新方向,该方法将专家知识与算法的自我学习能力相结合,进一步提高算法的性能。然而和强化学习类方法相类似,逆强化学习类方法对仿真器的环境重现性能有较高的要求。

通过本节的总结与分析,依据其是否采用学习类算法以及所采用学习类算法的类型,智能车辆决策方法的分类归纳如表3所示。

表 3 依据算法类型的决策方法分类

Tab. 3 Classification of decision-making methods based on algorithm types

类别	优势	劣势	文献编号
基于状态机	逻辑清晰	复杂场景表现不佳	[5]、[8]、[37]
基于深度监督学习	灵活性强	对训练数据要求较高	[16-17]、[33]、[38-39]
基于强化学习	运算效率高	局限于人工设定奖惩	[41-45]、[49]
基于逆强化学习	算法生成奖励函数	易造成过拟合	[47]、[55]、[57-58]

5 评估方法与数据集

如前文所述,现阶段对于决策方法的性能评估标准尚未统一,且当前缺乏文献对其进行归纳与总结.同时,现阶段尚未有文献对决策研究中常用的数据集进行整理与归纳.本小结将对决策效果评估方法进行讨论,并将对常用数据集进行总结.

5.1 决策的评估方法

如何评估智能车辆决策方法的优越性,是算法研究中的关键问题.现阶段决策方法的相关成果中,决策方法的评估手段可以分类为自建仿真场景的评估方法与基于数据集的评估方法.

5.1.1 自建仿真场景的评估方法 自建仿真场景并设计智能机器人模仿环境参与者以进行决策方法的评估是现阶段常见的评估方法.该方法中,科研工作者常搭建仿真环境,并定义具体的指标来衡量算法的优越性,例如安全、效率、舒适的量化指标,是否会发生碰撞以及碰撞发生的最小时间等.文献[15]建立了三车道仿真平台,并基于此评估了其提出的基于进化学习的智能车辆决策方法.文献[22]设计了两个评估指标:延长碰撞时间(ETTC)和动态预测距离分布(DPIDP),并在自建的仿真平台上进行了决策方法评估与验证.文献[59]通过设计路口仿真器,验证了其提出的基于可到达集的决策方法.文献[60]设计了 T 型路口仿真器,通过对比碰撞率以及通过时间来评估其决策方法.文献[13]以及[23-24]同样采用自建仿真场景来进行评估.

随着现阶段科研水平的不断提高,众多科研工作者运用智能车辆在现实环境中创建仿真场景对决策方法进行测试.文献[61]在真实车流中评估其决策算法.

5.1.2 基于数据集的评估方法 另一种常见的评估方法是基于已有的数据集对决策方法进行评估.在运用数据集对智能车辆决策方法进行评估时,一个主要特点是数据集中的车辆轨迹是既定的.在众多采用此方法进行评估的研究工作中,常见的方式为将数据集中的目标车辆(例如正在或即将进行换

道、汇流的车辆)替换为采用目标决策方法的智能车辆,随后根据该智能车辆能否安全顺利完成指定目标来进行评估.同时,除了将目标完成度作为度量指标外,在模拟评估过程中车辆的碰撞率、平稳性、最小安全距离等指标也将作为评价函数的子属性对决策方法进行评估.文献[62]在论文中设计了评估指标,并运用 NGSIM 数据集对其提出的基于 pAC 的决策方法的评估.文献[11]运用 NGSIM 数据集中的汇流场景,对其基于博弈论的决策方法进行评估验证,并以汇流是否成功作为评估指标.文献[25]运用 NGSIM 数据集验证其车辆换道决策的类人性与安全性.

该方法的优点在于数据来源于真实环境,较有说服力,缺点在于数据中的车辆缺乏与智能车辆的交互性,不能有效验证涉及交互的决策方法.

5.2 数据集

5.2.1 NGSIM 数据集 美国联邦公路局于 2002 年发起的 NGSIM 计划.通过在路段上方安装高清摄像机的方式采集检测路段上的车辆运行数据,再利用视频处理软件,以 10 帧/s 或 15 帧/s 的频率对车辆轨迹数据进行还原^[63].

该数据集包括 I-80、US-101、Lank、Peach 4 个路段上的车辆轨迹数据.其中,I-80 和 US-101 为高速公路,Lank 和 Peach 为城市道路.此数据集在车辆跟驰、车辆变道、交通流演化和车辆轨迹预测等方面有着广泛的应用^[64-67].针对智能车辆决策研究领域,NGSIM 数据集可用于进行决策方法的功能性验证.同时,由于该数据集采集时间较早且使用广泛,适合研究人员用于不同决策方法之间优越性的对比.但该数据集的不足在于其场景比较单一,并且在车辆的边缘提取上存在一定的误差. NGSIM 数据集的数据结构如表 4 所示.

5.2.2 High-D 数据集 High-D 数据集由德国亚琛工业大学汽车工程研究所提供,为德国高速公路的大型自然车辆轨迹数据集^[68].此数据集由高空悬停无人机拍摄并经由后期图像处理,包括来自 6 个地点的 11.5 h 测量值和 1.1×10^5 辆车,所测量的

车辆总行驶里程为 4.5×10^4 km,还包括 5 600 条完整的变道记录.该数据集定位误差平均小于 10 cm.该数据集的车辆定位精准、车辆边缘识别准确,且车流密度较大,适合于高车速场景下的决策方法研究.但不足之处在于场景较为单一,聚焦于高速公路场景. High-D 数据集的数据结构如表 5 所示.

表 4 NGSIM 数据集数据格式
Tab. 4 Data format of NGSIM dataset

编号	名称	描述
1	vehicle ID	车辆编号
2	frame ID	数据帧号
3	total frames	数据总帧
4	global time	标准时间
5	local x	坐标系 x 值
6	local y	坐标系 y 值
7	global x	标准地理坐标系 x 值
8	global y	标准地理坐标系 y 值
9	vehicle length	车辆长度
10	vehicle width	车辆宽度
11	vehicle class	车辆类型
12	vehicle velocity	车辆速度
13	vehicle acceleration	车辆加速度
14	lane identification	车道编号
15	preceding vehicle	跟驰前车编号
16	following vehicle	跟驰后车编号
17	spacing	车头间距
18	headway	车头时距

表 5 High-D 数据集数据格式
Tab. 5 Data format of High-D dataset

编号	名称	描述
1	id	记录的 ID
2	frameRate	用来录制视频的帧速率
3	locationId	记录位置的 ID
4	speedLimit	行驶车道的速度限制
5	month	录制的月份
6	weekDay	录制的工作日完成
7	startTime	录制的开始时间
8	duration	记录的持续时间
9	totalDrivenDistance	所有履带车辆的总行驶距离
10	totalDrivenTime	所有履带车辆的总行驶时间
11	numVehicles	跟踪的车辆数量,包括汽车和卡车
12	numCars	跟踪的汽车数量
13	numTrucks	跟踪的卡车数量
14	upperLaneMarkings	上车道标记的 y 位置
15	lowerLaneMarkings	下车道标记的 y 位置

5.2.3 INTERACTION 数据集 INTERACTION 是由加州大学伯克利分校机械系统控制实验室(MSC Lab)等建立的一个国际性、对抗性、协作性的数据集^[69].该数据集场景较为丰富,包含了城市道路、公路、匝道合并、带有礼让停车标志的环形路口以及信号灯交叉路口等.同时,该数据集采集地点来自不同的国家,以便自然地包含不同文化中的驾驶偏好和风格,并且该数据集具有完整语义的地图信息,包括物理层、参考线、道路连接以及交通规则.

针对智能车辆决策方法研究领域,该数据集适用于驾驶员决策风格的提取、对抗性场景下决策方法研究以及车流密度较大的、复杂场景下的决策预测与研究.同时,该数据集收集了危险场景、碰撞场景的案例,可以供科研工作者进行危险场景下的研究.但由于该数据集的信息并非由传感器直接收集获得,所以不能应用于端到端的算法开发中. INTERACTION 数据集的数据结构如表 6 所示.

表 6 INTERACTION 数据集数据格式
Tab. 6 Data format of INTERACTION dataset

编号	名称	描述
1	track_id	代表车辆的 ID
2	frame_id	车辆出现在视频中的帧
3	timestamp_ms	时间戳
4	agent_type	跟踪类型:“行人/自行车”
5	x	每一帧的坐标 x
6	y	每一帧的坐标 y
7	v_x	x 方向速度
8	v_y	y 方向速度

5.2.4 Level 5 数据集 Level 5 数据集由美国 Lyft 公司提供,该数据集包含了超过 4 000 个车道段的基础高精度空间语义地图、197 个人行横道、60 个停车标志、54 个停车区、8 个减速带和 11 个减速带^[70].

数据集约有 3 000 个驾驶场景,总计 16.7 h 的视频数据共 6×10^5 帧,以及大约 2.5×10^7 个 3D 边界框和 2.2×10^7 个 2D 边界框. Level 5 由两个子数据集组成,分别是感知数据集与预测数据集.感知数据集的数据格式沿用 nuScenes 格式,由 7 个摄像头和 3 个激光雷达传感器采集生成.预测数据集包含 1.7×10^5 个场景,这些场景捕获了自动驾驶汽车周围的环境,并且每个场景都会在给定的时间点对车辆周围的状态进行编码. Level 5 数据集的数据格式

如表 7 所示.

的数据集,本文对各数据集的特性以及引用情况进行

综上所述,介绍了 4 个在决策方法研究中常用

进行了总结分析,如表 8 所示.

表 7 Level 5 数据集数据格式

Tab. 7 Data format of Level 5 dataset

编号	名称	描述	编号	名称	描述
1	timestamp	时间戳	6	extent	车辆尺寸
2	agent_index_interval	车辆 ID	7	yaw	车辆朝向
3	ego_translation	自车位置转换向量	8	velocity	车辆速度
4	ego_rotation	自车朝向转换向量	9	track_id	轨迹 ID
5	centroid	车辆坐标	10	label_probabilities	车辆类型置信度

表 8 数据集总结表

Tab. 8 Summary of datasets

数据集	场景描述	地点	时长/min	采样率/Hz	数据格式	文献编号
NGSIM	高速、城市道路	美国	90	10	csv	[25]、[64-67]
High-D	高速公路	德国	690	25	csv	[71-78]
INTERACTION	路口场景	世界	990. 98	10	csv	[79-82]
Level 5	城市道路	美国	1002	10	zarr	[83-86]

6 结语

智能车辆决策模块是车辆安全性、舒适性与高效性的决定因素之一,而车辆行驶环境的高度动态性与复杂性对决策方法的研究提出了挑战.近年来,学术界和工业界进行了广泛和深入的研究,并提出了一系列的决策方法.然而,目前的研究依旧存在如下挑战亟待解决.

(1) 数据不均衡问题尚未解决.决策方法的设计与验证离不开数据,然而现阶段智能车辆决策方法研究中所用的真实数据常存在不均衡问题.例如车辆在正常行驶状况下,直道的行为远多于弯道的行为,简单场景的决策会远多于复杂危险场景的决策.不平衡数据集在训练模型时,反馈的梯度分布也会失衡,从而使模型的能力偏向于数据分布较大的数据,而对少分布的数据估计、预测能力较差.

(2) 可解释、稳定性的决策方法尚未实现.现阶段学习类方法在决策方法中得到了较为广泛的应用,为了提高决策方法的可解释性并建立用户与决策模型之间的信任关系,消除模型在实际部署应用中的不稳定性,近年来学术界和工业界已经在可解释性上取得了一定的进展.然而该方向的研究还处于初级阶段,依然存在许多的关键问题尚待解决.

(3) 决策模块中的伦理问题尚未得到深入探讨.自“电车问题”提出以来,众多学者对如何做出符合伦理的决策产生持续的争论.智能车辆的决策模

块由于其功能的特殊性,也常受到伦理问题的挑战.例如在博弈环节中如何设置遵循伦理的目标函数,如何在路权被侵犯时合理地与环境车辆进行交互等.

(4) 噪声、不完全可观场景下决策方法的稳定性需要进一步提高.环境的感知与定位信息对于智能车辆的决策生成而言至关重要.但当车流密度较大时,感知信息中常存在噪声,并且由于环境车辆的遮挡而导致不完全客观环境的问题难以避免.以上问题对决策方法的稳定性提出了挑战.

(5) 实际应用场景中车辆决策实验亟需开展.目前的车辆决策方法研究仍然以实验室环境验证为主,缺乏实际应用场景的验证.

针对现阶段智能车辆决策方法研究中存在的问题,本文从 5 个方面指出未来的研究方向.

(1) 针对边缘场景的决策方法研究与数据集设计工作.随着智能车辆研究工作的展开,现阶段决策方法逐渐聚焦于解决边缘场景、困难场景下的决策生成.因此,设计应对训练数据不均衡,解决现实中不常见场景下的决策问题,将会成为智能车辆决策方法研究的方向之一.同时,由于当前数据集缺乏对危险场景、冲突场景等边缘场景的数据采集,设计具有针对性的数据集亦将成为未来研究工作的方向之一.

(2) 第三代人工智能背景下可解释、鲁棒决策方法研究.现阶段决策方法难以兼顾解释性与稳定

性,原因在于当下的学习类决策方法,较多地依赖于数据驱动,却忽略了专家知识信息.未来车辆决策方法的研究将会把数据驱动与知识驱动结合起来,通过同时利用知识、数据、算法和算力等4个要素实现决策方法的开发.

(3) 考虑伦理道德的车辆决策方法研究.决策是否遵循伦理道德,将会直接影响在车辆遇险等场景下的责任追究与法律判定.同时,决策是否符合伦理规范将会影响用户与公众对决策模型的信任,进而影响智能车辆在社会范围大规模推广的进程.因此,将伦理问题量化,并为决策模块中的伦理问题设计合适的代价函数,亦成为未来决策方法的研究方向之一.

(4) 基于路侧感知与车间通讯的决策方法研究.随着通讯技术的进步以及路侧基础设施的普及,以路侧感知信息和车间通讯信息作为输入的决策方法将有效解决感知噪声干扰和不完全可观测场景下的决策问题,因此该方向将成为未来的决策方法研究热点之一.

(5) 实际复杂场景下的智能车辆决策实验.结合实际的车辆行驶场景,如开放、拥堵场景下的车辆变道行为、汇流场景下的车辆博弈决策行为、路口场景中的车辆决策行为、车辆与行人等多种交通参与者的混合场景下的决策行为等展开实验.

参考文献:

- [1] 胡兵, 杨明, 郭林栋, 等. 基于地面快速鲁棒特征的智能车全局定位方法[J]. 上海交通大学学报, 2019, 53(2): 203-208.
HU Bing, YANG Ming, GUO Lindong, *et al.* Global localization for intelligent vehicles using ground SURF[J]. *Journal of Shanghai Jiao Tong University*, 2019, 53(2): 203-208.
- [2] 杜明博, 梅涛, 陈佳佳, 等. 复杂环境下基于 RRT 的智能车辆运动规划算法[J]. 机器人, 2015, 37(4): 443-450.
DU Mingbo, MEI Tao, CHEN Jiajia, *et al.* RRT-based motion planning algorithm for intelligent vehicle in complex environments[J]. *Robot*, 2015, 37(4): 443-450.
- [3] 陈佳佳. 城市环境下无人驾驶车辆决策系统研究[D]. 合肥: 中国科学技术大学, 2014.
CHEN Jiajia. Research on decision making system of autonomous vehicle in urban environments [D]. Hefei: University of Science and Technology of China, 2014.
- [4] SCHWARTING W, ALONSO-MORA J, RUS D.

- Planning and decision-making for autonomous vehicles[J]. *Annual Review of Control, Robotics, and Autonomous Systems*, 2018, 1(1): 187-210.
- [5] 熊璐, 康宇宸, 张培志, 等. 无人驾驶车辆行为决策系统研究[J]. 汽车技术, 2018(8): 1-9.
XIONG Lu, KANG Yuchen, ZHANG Peizhi, *et al.* Research on behavior decision-making system for unmanned vehicle [J]. *Automobile Technology*, 2018(8): 1-9.
- [6] GONZÁLEZ D S, GARZÓN M, DIBANGOYE J S, *et al.* Human-like decision-making for automated driving in highways [C] // **2019 IEEE Intelligent Transportation Systems Conference (ITSC)**. Auckland, New Zealand: IEEE, 2019: 2087-2094.
- [7] NOH S, AN K. Decision-making framework for automated driving in highway environments[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2018, 19(1): 58-71.
- [8] XIONG G M, KANG Z Y, LI H, *et al.* Decision-making of lane change behavior based on RCS for automated vehicles in the real environment[C] // **2018 IEEE Intelligent Vehicles Symposium (IV)**. Changshu: IEEE, 2018: 1400-1405.
- [9] NISHI T, DOSHI P, PROKHOROV D. Merging in congested freeway traffic using multipolicy decision making and passive actor-critic learning [J]. *IEEE Transactions on Intelligent Vehicles*, 2019, 4(2): 287-297.
- [10] KANG K, RAKHA H A. A repeated game freeway lane changing model [J]. *Sensors*, 2020, 20(6): 1554.
- [11] KANG K, RAKHA H A. Game theoretical approach to model decision making for merging maneuvers at freeway on-ramps [J]. *Transportation Research Record Journal of the Transportation Research Board*, 2017, 2623(1): 19-28.
- [12] GUO J Q, CHENG S L, LIU Y. Merging and diverging impact on mixed traffic of regular and autonomous vehicles[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2020, 22(3): 1639-1649.
- [13] HOEL C J, WOLFF K, LAINE L. Automated speed and lane change decision making using deep reinforcement learning[C] // **2018 21st International Conference on Intelligent Transportation Systems (ITSC)**. Maui, HI, USA: IEEE, 2018: 2148-2155.
- [14] ALIZADEH A, MOGHADAM M, BICER Y, *et al.* Automated lane change decision making using deep reinforcement learning in dynamic and uncertain highway environment[C] // **2019 IEEE Intelligent Trans-**

- portation Systems Conference (ITSC)**. Auckland, New Zealand; IEEE, 2019; 1399-1404.
- [15] LI T T, WU J P, CHAN C Y. Evolutionary learning in decision making for tactical lane changing[C]//**2019 IEEE Intelligent Transportation Systems Conference (ITSC)**. Auckland, New Zealand; IEEE, 2019; 1921-1185.
- [16] CODEVILLA F, MÜLLER M, LÓPEZ A, *et al.* End-to-end driving via conditional imitation learning [C]//**2018 IEEE International Conference on Robotics and Automation (ICRA)**. Brisbane, QLD, Australia; IEEE, 2018; 4693-4700.
- [17] SANTANA E, HOTZ G. Learning a driving simulator [EB/OL]. (2016-08-01) [2020-06-14]. https://www.researchgate.net/publication/305809416_Learning_a_Driving_Simulator.
- [18] YUAN W, YANG M, WANG C X, *et al.* Longitudinal and lateral coupling model based end-to-end learning for lane keeping of self-driving cars[C]//**International Conference on Cognitive Systems and Signal Processing**. Beijing; ICCSIP, 2019; 425-436.
- [19] YUAN W, YANG M, LI H, *et al.* End-to-end learning for high-precision lane keeping via multi-state model[J]. **CAAI Transactions on Intelligence Technology**, 2018, 3(4): 185-190.
- [20] YU H, YANG S, GU W H, *et al.* Baidu driving dataset and end-to-end reactive control model[C]//**2017 IEEE Intelligent Vehicles Symposium (IV)**. Los Angeles, CA, USA; IEEE, 2017; 341-346.
- [21] BOJARSKI M, YERES P, CHOROMANSKA A, *et al.* Explaining how a deep neural network trained with end-to-end learning steers a car [EB/OL]. (2017-04-14) [2020-06-14]. https://www.researchgate.net/publication/316505624_Explaining_How_a_Deep_Neural_Network_Trained_with_End-to-End_Learning_Steers_a_Car.
- [22] IBERRAKEN D, ADOUANE L, DENIS D. Safe autonomous overtaking maneuver based on inter-vehicular distance prediction and multi-level Bayesian decision-making [C] // **IEEE Conference on Intelligent Transportation Systems**. Maui, HI, USA; IEEE, 2018; 1830-8831.
- [23] TRAM T, BATKOVIC I, ALI M, *et al.* Learning when to drive in intersections by combining reinforcement learning and model predictive control[C]//**2019 IEEE Intelligent Transportation Systems Conference (ITSC)**. Auckland, New Zealand; IEEE, 2019; 1919-4092.
- [24] ORZECOWSKI P F, BURGER C, LAUER M. Decision-making for automated vehicles using a hierarchical behavior-based arbitration scheme[C]//**2020 IEEE Intelligent Vehicles Symposium (IV)**. Las Vegas, NV, USA; IEEE, 2020; 2028-6743.
- [25] WANG X, WU J Q, GU Y L, *et al.* Human-like maneuver decision using LSTM-CRF model for on-road self-driving[C]//**2018 21st International Conference on Intelligent Transportation Systems (ITSC)**. Maui, HI, USA; IEEE, 2018; 210-216.
- [26] 成英, 高利, 陈雪梅, 等. 有人与无人驾驶车辆交叉口驾驶博弈模型[J]. **北京理工大学学报**, 2019, 39(9): 938-943.
- CHENG Ying, GAO Li, CHEN Xuemei, *et al.* A driving game model for manned and unmanned vehicles at intersection[J]. **Transactions of Beijing Institute of Technology**, 2019, 39(9): 938-943.
- [27] CHEN Z L, HUANG X M. End-to-end learning for lane keeping of self-driving cars[C]//**2017 IEEE Intelligent Vehicles Symposium (IV)**. Los Angeles, CA, USA; IEEE, 2017; 1856-1860.
- [28] LI N, OYLER D W, ZHANG M X, *et al.* Game theoretic modeling of driver and vehicle interactions for verification and validation of autonomous vehicle control systems[J]. **IEEE Transactions on Control Systems Technology**, 2018, 26(5): 1782-1797.
- [29] CUNNINGHAM A G, GALCERAN E, EUSTICE R M, *et al.* MPDM: Multipolicy decision-making in dynamic, uncertain environments for autonomous driving[C]//**2015 IEEE International Conference on Robotics and Automation (ICRA)**. Seattle, WA, USA; IEEE, 2015; 1670-1677.
- [30] SCHWARTING W, PASCHEKA P. Recursive conflict resolution for cooperative motion planning in dynamic highway traffic[C]//**17th International IEEE Conference on Intelligent Transportation Systems (ITSC)**. Qingdao; IEEE, 2014; 1039-1044.
- [31] LENZ D, KESSLER T, KNOLL A. Tactical cooperative planning for autonomous highway driving using Monte-Carlo Tree Search[C]//**2016 IEEE Intelligent Vehicles Symposium (IV)**. Gothenburg, Sweden; IEEE, 2016; 447-453.
- [32] CUNNINGHAM A G, GALCERAN E, MEHTA D, *et al.* MPDM: Multi-policy decision-making from autonomous driving to social robot navigation[M]// WASCHL H, KOLMANOVSKY I, WILLEMS F. **Control Strategies for Advanced Driver Assistance Systems and Autonomous Driving Functions**. Berlin, Germany; Springer, 2019; 201-223.
- [33] CHEN J Y, TANG C, XIN L, *et al.* Continuous de-

- cision making for on-road autonomous driving under uncertain and interactive environments [C] // **2018 IEEE Intelligent Vehicles Symposium (IV)**. Changshu: IEEE, 2018: 1651-1658.
- [34] WEI J Q, DOLAN J M, LITKOUHI B. Autonomous vehicle social behavior for highway entrance ramp management[C] // **2013 IEEE Intelligent Vehicles Symposium (IV)**. Gold Coast, QLD, Australia: IEEE, 2013: 201-207.
- [35] HOERMANN S, STUMPER D, DIETMAYER K. Probabilistic long-term prediction for autonomous vehicles[C] // **2017 IEEE Intelligent Vehicles Symposium (IV)**. Piscataway, NJ, USA: IEEE, 2017: 237-243.
- [36] DING W C, ZHANG L, CHEN J, *et al.* Safe trajectory generation for complex urban environments using spatio-temporal semantic corridor[J]. **IEEE Robotics and Automation Letters**, 2019, 4(3): 2997-3004.
- [37] 武历颖. 无人驾驶汽车环境信息提取及运动决策方法研究[D]. 西安: 长安大学, 2016.
WU Liying. Research on environmental information extraction and movement decision-making method of unmanned vehicle[D]. Xi'an: Changan University, 2016.
- [38] LEURENT E, MERCAT J. Social attention for autonomous decision-making in dense traffic [EB/OL]. (2019-11-27) [2020-06-14]. <https://arxiv.org/abs/1911.12250>.
- [39] MEHTA A, SUBRAMANIAN A, SUBRAMANIAN A. Learning end-to-end autonomous driving using guided auxiliary supervision [EB/OL]. (2018-08-30) [2020-06-14]. https://www.researchgate.net/publication/327335968_Learning_End-to-end_Autonomous_Driving_using_Guided_Auxiliary_Supervision.
- [40] XU J X, LUO Q, XU K C, *et al.* An automated learning-based procedure for large-scale vehicle dynamics modeling on Baidu Apollo platform[C] // **2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)**. Piscataway, NJ, USA: IEEE, 2019: 5049-5056.
- [41] SUTTON R S, BARTO A G. Reinforcement learning: An introduction[J]. **IEEE Transactions on Neural Networks**, 1998, 9(5): 1054.
- [42] PIETER A, ANDREW Y. Apprenticeship learning via inverse reinforcement learning[EB/OL]. (2004-09-01) [2020-06-14]. https://www.researchgate.net/publication/2947330_Apprenticeship_Learning_via_Inverse_Reinforcement_Learning.
- [43] BOJARSKI M, DEL TESTA D, DWORAKOWSKI D, *et al.* End to end learning for self-driving cars [EB/OL]. (2016-04-01) [2020-06-15]. https://www.researchgate.net/publication/301648615_End_to_End_Learning_for_Self-Driving_Cars.
- [44] XU H Z, GAO Y, YU F, *et al.* End-to-end learning of driving models from large-scale video datasets[C] // **2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**. Piscataway, NJ, USA: IEEE, 2017: 3530-3538.
- [45] HECKER S, DAI D X, GOOL L. End-to-end learning of driving models with surround-view cameras and route planners[C] // **European Conference on Computer Vision**. Glasgow, UK: Springer, 2018: 449-468.
- [46] GAO Z H, SUN T J, XIAO H W. Decision-making method for vehicle longitudinal automatic driving based on reinforcement Q-learning[J]. **International Journal of Advanced Robotic Systems**, 2019, 16(3): 172988141985318.
- [47] YE Y J, ZHANG X H, SUN J. Automated vehicle's behavior decision making using deep reinforcement learning and high-fidelity simulation environment[J]. **Transportation Research Part C: Emerging Technologies**, 2019, 107: 155-170.
- [48] 郑睿. 基于增强学习的无人车辆智能决策方法研究[D]. 长沙: 国防科学技术大学, 2013.
ZHENG Rui. Reinforcement learning-based intelligent decision-making methods for unmanned vehicles [D]. Changsha: National University of Defense Technology, 2013.
- [49] YUAN W, YANG M, HE Y S, *et al.* Multi-reward architecture based reinforcement learning for highway driving policies[C] // **2019 IEEE Intelligent Transportation Systems Conference (ITSC)**. Auckland, New Zealand: IEEE, 2019: 19194107.
- [50] QIAO Z Q, MUELLING K, DOLAN J M, *et al.* Automatically generated curriculum based reinforcement learning for autonomous vehicles in urban environment[C] // **2018 IEEE Intelligent Vehicles Symposium (IV)**. Piscataway, NJ, USA: IEEE, 2018: 1233-1238.
- [51] BOUTON M, NAKHAEI A, FUJIMURA K, *et al.* Cooperation-aware reinforcement learning for merging in dense traffic[C] // **2019 IEEE Intelligent Transportation Systems Conference (ITSC)**. Changshu: IEEE, 2019: 3441-3447.
- [52] KAUSHIK M, PRASAD V, KRISHNA K M, *et al.* Overtaking maneuvers in simulated highway driving using deep reinforcement learning[C] // **2018 IEEE Intelligent Vehicles Symposium (IV)**. Changshu:

- IEEE, 2018; 1885-1890.
- [53] FERDOWSI A, CHALLITA U, SAAD W, *et al.* Robust deep reinforcement learning for security and safety in autonomous vehicle systems[C]//**2018 21st International Conference on Intelligent Transportation Systems (ITSC)**. Piscataway, NJ, USA: IEEE, 2018; 307-312.
- [54] MA X B, CAMPBELL D K, KOCHENDERFER M J. Improved robustness and safety for autonomous vehicle control with adversarial reinforcement learning [C] // **2018 IEEE Intelligent Vehicles Symposium (IV)**. Changshu: IEEE, 2018; 18168097.
- [55] SHARIFZADEH S, CHIOTELLIS I, TRIEBEL R, *et al.* Learning to drive using inverse reinforcement learning and deep Q-networks [EB/OL]. (2016-12-12) [2020-06-14]. <https://arxiv.org/abs/1612.03653>.
- [56] CHEN X L, MONFORT M, ZIEVART B D, *et al.* Adversarial inverse optimal control for general imitation learning losses and embodiment transfer [EB/OL]. (2016-06-01) [2020-06-14]. https://scholar.google.com.hk/scholar?hl=zh-CN&as_sdt=0%2C5&q=Adversarial+Inverse+Optimal+Control+for+General+Imitation+Learning+Losses+and+Embodiment+Transfer&btnG=.
- [57] MAJUMDAR A, SINGH S, MANDLEKAR A, *et al.* Risk-sensitive inverse reinforcement learning via coherent risk models[EB/OL]. [2020-06-14]. <http://www.roboticsproceedings.org/rss13/p69.pdf>.
- [58] WULFMEIER M, ONDRUSKA P, POSNER I. Maximum entropy deep inverse reinforcement learning[EB/OL]. (2015-07-17) [2020-06-14]. <https://arxiv.org/abs/1507.04888>.
- [59] AHN H, BERNTORP K, DI CAIRANO S. Reachability-based decision making for city driving[C]//**2018 Annual American Control Conference (ACC)**. Milwaukee, WI, USA: IEEE, 2018; 3203-3208.
- [60] BOUTON M, NAKHAEI A, FUJIMURA K, *et al.* Safe reinforcement learning with scene decomposition for navigating complex urban environments[C]//**2019 IEEE Intelligent Vehicles Symposium (IV)**. Paris, France: IEEE, 2019; 1469-1476.
- [61] 陈雪梅, 田赓, 苗一松, 等. 城市环境下无人驾驶车辆驾驶规则获取及决策算法[J]. *北京理工大学学报*, 2017, 37(5): 491-496.
- CHEN Xuemei, TIAN Geng, MIAO Yisong, *et al.* Driving rule acquisition and decision algorithm to unmanned vehicle in urban traffic[J]. **Transactions of Beijing Institute of Technology**, 2017, 37(5): 491-496.
- [62] NISHI T, DOSHI P, PROKHOROV D. Merging in congested freeway traffic using multipolicy decision making and passive actor-critic learning [J]. **IEEE Transactions on Intelligent Vehicles**, 2019, 4(2): 287-297.
- [63] ALEXIADIS V, COLYAR J, HALKIAS J, *et al.* The next generation simulation program[J]. **Institute of Transportation Engineers. ITE Journal**, 2004, 74(8): 22.
- [64] LI G, FANG S, MA J X, *et al.* Modeling merging acceleration and deceleration behavior based on gradient-boosting decision tree[J]. **Journal of Transportation Engineering, Part A: Systems**, 2020, 146(7): 1-9.
- [65] LI G, PAN Y Y, YANG Z, *et al.* Modeling vehicle merging position selection behaviors based on a finite mixture of linear regression models[J]. **IEEE Access**, 2019, 7: 158445-158458.
- [66] BIN AL ISLAM S M A, ABDUL AZIZ H M, HAJBABAIE A. Stochastic gradient-based optimal signal control with energy consumption bounds [EB/OL]. (2020-03-18) [2020-06-14]. https://scholar.google.com.hk/scholar?hl=zh-CN&as_sdt=0%2C5&q=Stochastic+Gradient-Based+Optimal+Signal+Control+With+Energy+Consumption+Bounds&btnG=.
- [67] LI G, CHENG J. Exploring the effects of traffic density on merging behavior[J]. **IEEE Access**, 2019, 7: 51608-51619.
- [68] KRAJEWSKI R, BOCK J, KLOEKER L, *et al.* The highD dataset: A drone dataset of naturalistic vehicle trajectories on German highways for validation of highly automated driving systems[C]//**2018 21st International Conference on Intelligent Transportation Systems (ITSC)**. Maui, HI, USA: IEEE, 2018; 18309040.
- [69] ZHAN W, SUN L, WANG D, *et al.* Interaction dataset: An international, adversarial and cooperative motion dataset in interactive driving scenarios with semantic maps [EB/OL]. (2019-09-30) [2020-06-14]. <https://arxiv.org/abs/1910.03088>.
- [70] HOUSTON J, ZUIDHOF G, BERGAMINI L, *et al.* One thousand and one hours: Self-driving motion Prediction dataset [EB/OL]. (2020-06-25) [2020-08-14]. <https://arxiv.org/abs/2006.14480>.
- [71] KIM H, KIM D, KIM G, *et al.* Multi-head attention based probabilistic vehicle trajectory prediction[C]//**2020 IEEE Intelligent Vehicles Symposium (IV)**. Las

- Vegas, NV, USA: IEEE, 2020: 1720-1725.
- [72] RUDENKO A, PALMIERI L, HERMAN M, *et al.* Human motion trajectory prediction: A survey[J]. **The International Journal of Robotics Research**, 2020, 39(8): 895-935.
- [73] ETESAMI J, GEIGER P. Causal transfer for imitation learning and decision making under sensor-shift [C]// **Proceedings of the AAAI Conference on Artificial Intelligence**, 2020, 34(6): 10118-10125.
- [74] MESSAOUD K, YAHIAOUI I, VERROUST-BLONDET A, *et al.* Relational recurrent neural networks for vehicle trajectory prediction [C]// **2019 IEEE Intelligent Transportation Systems Conference (ITSC)**. Auckland, New Zealand: IEEE, 2019: 19193897.
- [75] WANG X, NAIR S, ALTHOFF M. Falsification-based robust adversarial reinforcement learning [EB/OL]. (2020-07-01) [2020-06-14]. <https://arxiv.org/abs/2007.00691>.
- [76] BHATTACHARYYA R P, SENANAYAKE R, BROWN K, *et al.* Online parameter estimation for human driver behavior prediction [C]// **2020 American Control Conference (ACC)**. Denver, CO, USA: IEEE, 2020: 19830660.
- [77] QIAO Z Q, ZHAO J, ZHU J, *et al.* Human driver behavior prediction based on UrbanFlow [C]// **2020 IEEE International Conference on Robotics and Automation (ICRA)**. Paris, France: IEEE, 2020: 10570-10576.
- [78] MESSAOUD K, YAHIAOUI I, VERROUST-BLONDET A, *et al.* Attention based vehicle trajectory prediction[J]. **IEEE Transactions on Intelligent Vehicles**, 2021, 6(1): 175-185.
- [79] SUN L T, ZHAN W, HU Y P, *et al.* Interpretable modelling of driving behaviors in interactive driving scenarios based on cumulative prospect theory [C]// **2019 IEEE Intelligent Transportation Systems Conference (ITSC)**. Auckland, New Zealand: IEEE, 2019: 4329-4335.
- [80] HU Y P, ZHAN W, TOMIZUKA M. Scenario-transferable semantic graph reasoning for interaction-aware probabilistic prediction [EB/OL]. (2020-04-01) [2020-06-14]. https://www.researchgate.net/publication/340500370_Scenario-Transferable_Semantic_Graph_Reasoning_for_Interaction-Aware_Probabilistic_Prediction.
- [81] SHU K Q, YU H L, CHEN X X, *et al.* Autonomous driving at intersections: A critical-turning-point approach for left turns [C]// **2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)**. Rhodes, Greece: IEEE, 2020: 1-6.
- [82] SHU K, YU H L, CHEN X X, *et al.* Autonomous driving at intersections: A critical-turning-point approach for left turns [C]// **2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)**. Rhodes, Greece: IEEE, 2020: 20303092.
- [83] FENG D, HAASE-SCHÜTZ C, ROSENBAUM L, *et al.* Deep multi-modal object detection and semantic segmentation for autonomous driving: Datasets, methods, and challenges [J]. **IEEE Transactions on Intelligent Transportation Systems**, 2021, 22(3): 1341-1360.
- [84] CHANDRA R, GUAN T, PANUGANTI S, *et al.* Forecasting trajectory and behavior of road-agents using spectral clustering in graph-LSTMs [J]. **IEEE Robotics and Automation Letters**, 2020, 5(3): 1974-1183.
- [85] SALZMANN T, IVANOVIC B, CHAKRAVARTY P, *et al.* Trajectron++: Dynamically-feasible trajectory forecasting with heterogeneous data [C]// **European Conference on Computer Vision-ECCV 2020**. Glasgow, UK: Springer, 2020: 683-700.
- [86] GADD M, DE MARTINI D, MARCHEGIANI L, *et al.* Sense-assess-eXplain (SAX): Building trust in autonomous vehicles in challenging real-world driving scenarios [C]// **2020 IEEE Intelligent Vehicles Symposium (IV)**. Las Vegas, NV, USA: IEEE, 2020: 150-155.

(本文编辑:石易文)