

·同济大学智能汽车决策技术专题·

# 无人驾驶车辆行为决策系统研究\*

熊璐 康宇宸 张培志 朱辰宇 余卓平

(同济大学,上海 201804)

**【摘要】**调研了国内外无人驾驶车辆行为决策系统的研究现状,对行为决策系统进行分类,基于国内外行为决策系统研究实例,对基于规则和基于学习算法的不同行为决策系统的实现方式、适用条件及优缺点进行比较,分析了现阶段无人车行为决策系统的研究水平、技术难点和发展趋势,为无人驾驶车辆行为决策系统的设计提供参考。

**主题词:**无人驾驶车辆 行为决策算法 有限状态机 学习算法

**中图分类号:**U461 **文献标识码:**A **DOI:** 10.19620/j.cnki.1000-3703.20180752

## Research on Behavior Decision-Making System for Unmanned Vehicle

Xiong Lu, Kang Yuchen, Zhang Peizhi, Zhu Chenyu, Yu Zhuoping

(Tongji University, Shanghai 201804)

**【Abstract】**The research status of behavior decision-making system of unmanned vehicle in China and other countries was investigated, and the behavior decision-making system was classified. Based on the research samples of behavior decision-making systems in China and other countries, the implementation and application conditions of different behavior decision-making systems based on rules and learning algorithms were discussed, and the advantages and disadvantages were compared. The research level, technical difficulties and development trends of the current unmanned vehicle behavior decision-making system were analyzed, which provided a reference for the design of the unmanned vehicle behavior decision-making system.

**Key words:** Unmanned vehicle, Behavior decision-making system, Finite state machine, Learning algorithm

### 1 无人车系统架构

无人驾驶车辆是可以自主行驶的车辆,其系统结构核心在于“智能”,不仅具备加减速、转向等常规的汽车功能,还集成了环境感知、行为决策、路径规划、车辆控制等系统功能,能够综合环境及自车信息,实现类似人类驾驶的行为。

典型的无人驾驶车辆系统架构如图1所示,总体分为环境感知、决策规划和运动控制三大部分。

环境感知系统的主要目的是获取并处理环境信息,利用多传感器目标检测与融合等技术,获取周围环境态势,为系统其他部分提供周围环境的关键信息。感知层将处理后的信息发送给决策层。

决策规划系统综合环境及自车信息,使无人车产生

安全、合理的驾驶行为,指导运动控制系统对车辆进行控制<sup>[1]</sup>。行为决策系统是狭义的决策系统,其根据感知层输出的信息合理决策出当前车辆的行为,并根据不同的行为确定轨迹规划的约束条件,指导轨迹规划模块规划出合适的路径、车速等信息,发送给控制层。

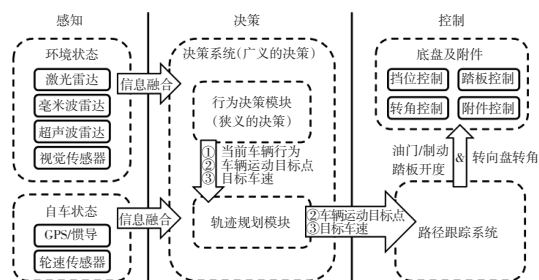


图1 典型无人驾驶车辆系统架构

运动控制系统接收决策规划层的指令并控制车辆

\*基金项目:上海市科委项目(17DZ1100202;16DZ1100700)。

通讯作者:余卓平(1960—),教授,博士生导师,博士,研究方向为车辆系统动力学与控制,yuzhuoping@tongji.edu.cn。

响应,保证控制精度,对目标车速、路径等进行跟踪<sup>[2]</sup>。

## 2 无人车行为决策系统概述

### 2.1 设计准则

无人车行为决策系统的目标是使无人车像熟练的驾驶员一样产生安全、合理的驾驶行为。其设计准则可总结为:良好的系统实时性;安全性最高优先级(车辆具备防碰撞、紧急避障、故障检测等功能);合理的行车效率优先级;结合用户需求的决策能力(用户对全局路径变更、安全和效率优先级变更等);乘员舒适性(车辆转向稳定性、平顺性等)。

对于适用于城市道路和高速公路工况的行为决策系统,设计准则还包括<sup>[2]</sup>:右侧车道通行优先;保持车道优先;速度限制;交通标志及交通信号灯限制等。

### 2.2 输入与输出

无人车行为决策系统的上层是感知层,其上层输入包括:

a. 局部环境信息,其基于相机、雷达等车载传感器,通过多传感器目标检测<sup>[3-4]</sup>与融合等技术<sup>[5]</sup>,融合处理各传感器数据,获取车辆周边环境态势,输出关键环境信息,交由决策系统处理。

b. 自车定位信息,其基于GPS/惯性导航系统、高精度地图定位<sup>[6]</sup>、即时定位与地图构建(Simultaneous Localization And Mapping, SLAM)<sup>[7-8]</sup>等方法,主要功能是对车辆进行高精度定位。

c. 地理地图信息和任务信息,包括先验全局路径、道路高精地图、交通标志信息等。

广义的决策系统的输出是运动控制信息。本文讨论狭义的行为决策系统,其下层是运动规划模块,行为决策系统的输出包括当前车辆行为、车辆运动局部目标点与目标车速等。

### 2.3 系统分类

无人车行为决策系统主要有基于规则和基于学习算法两大类。

基于规则的行为决策,即将无人驾驶车辆的行为进行划分,根据行驶规则、知识、经验、交通法规等建立行为规则库,根据不同的环境信息划分车辆状态,按照规则逻辑确定车辆行为的方法。其代表方法为有限状态机法<sup>[9-10]</sup>,代表应用有智能先锋II<sup>[11]</sup>、红旗CA7460、Boss<sup>[12]</sup>、Junior<sup>[13]</sup>、Odin<sup>[14]</sup>、Talos<sup>[15]</sup>、Bertha<sup>[16]</sup>等。

基于学习算法的行为决策,即通过对环境样本进行自主学习,由数据驱动建立行为规则库,利用不同的学习方法与网络结构,根据不同的环境信息直接进行行为

匹配,输出决策行为的方法,以深度学习的相关方法及决策树等各类机器学习方法<sup>[17-18]</sup>为代表。代表应用有英伟达(NVIDIA)<sup>[19]</sup>、Intel<sup>[20]</sup>、Comma.ai<sup>[21]</sup>、Mobileye<sup>[22]</sup>、百度、Waymo、特斯拉等。

## 3 基于规则的无人车行为决策系统

### 3.1 有限状态机法

基于规则的行为决策方法中最具代表性的是有限状态机法,其因逻辑清晰、实用性强等特点得到广泛应用。

有限状态机是一种离散输入、输出系统的数学模型。它由有限个状态组成,当前状态接收事件,并产生相应的动作,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<sup>[23-24]</sup>。

有限状态机的核心在于状态分解。根据状态分解的连接逻辑,将其分为串联式、并联式、混联式3种体系架构。

串联式结构的有限状态机系统,其子状态按照串联结构连接,状态转移大多为单向,不构成环路。并联式结构中各子状态输入、输出呈现多节点连接结构,根据不同输入信息,可直接进入不同子状态进行处理并提供输出<sup>[25]</sup>。如果一个有限状态机系统下的子状态中既存在串联递阶,又存在并联连接,则称这个系统具有混联结构。

### 3.2 基于有限状态机的行为决策系统

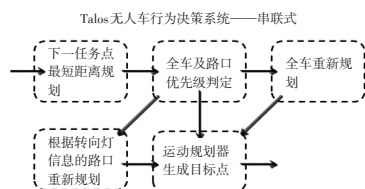
有限状态机法是经典的决策方法,因其实用性与稳定性在无人车决策系统中广泛应用,目前已比较成熟,美国国防高级研究计划局(Defense Advanced Research Projects Agency, DARPA)城市挑战赛中各队使用的决策系统为其典型代表。

#### 3.2.1 串联结构

麻省理工大学的Talos<sup>[15]</sup>无人车如图2所示,其行为决策系统总体采用串联结构。该无人车以越野工况挑战赛为任务目标,根据逻辑层级构建决策系统。其系统分为定位与导航、障碍物检测、车道线检测、路标识别、可行驶区域地图构建、运动规划、运动控制等模块,其中导航模块负责制定决策任务。



(a) Talos 无人车



(b)行为决策系统

图2 Talos无人车及其行为决策系统

串联式结构的优点是逻辑明确、规划推理能力强、问题求解精度高。其缺点在于对复杂问题的适应性差,某子状态故障时,会导致整个决策链的瘫痪。串联结构适用于某一工况的具体处理,擅长任务的层级推理与细分解决。

### 3.2.2 并联结构

斯坦福大学与大众公司研发的Junior无人车<sup>[13]</sup>行为决策系统结构如图3所示,其具备典型的并联结构。该系统分为初始化、前向行驶、停止标志前等待、路口通过、U形弯等13个子状态,各个子状态相互独立。

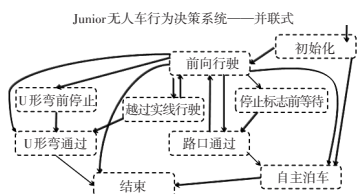


图3 Junior无人车的行为决策系统

Junior决策系统是并联划分子系统最多的系统之一,但在实际场景测试中,依然存在其有限状态机没有覆盖的工况,且对真实场景的辨识准确率较差。这说明单纯地应用并联式场景行为细分并不能提高场景遍历的深度,相反容易降低场景辨识准确率。

梅赛德斯奔驰公司研发的Bertha无人车<sup>[16]</sup>行为决策系统如图4所示。该系统分为路径规划、目标分析、交通信号灯管理、放弃管理4个独立并行的子状态模块。其中,放弃管理模块通过换挡操纵杆信号进行无人和有人驾驶的切换。

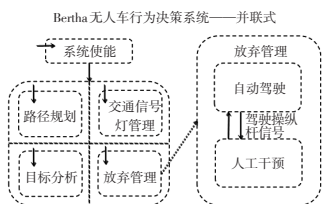


图4 Bertha无人车的行为决策系统

国防科学技术大学研发的红旗CA7460<sup>[2]</sup>行为决策系统如图5所示,其具备典型的并联结构。该系统适用于高速公路工况,其决策系统划分为自由追踪行车道、自由追踪超车道、由行车道换入超车道、由超车道换入

行车道等模式。

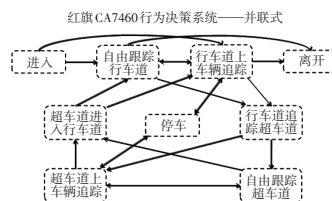


图5 红旗CA7460的行为决策系统

红旗CA7460对车辆行驶的安全性指标和效率指标进行了衡量,根据交通状况和安全性指标选出满足条件的候选行为,再根据效率指标决策出最优行为。

卡尔斯鲁厄工业大学的AnnieWAY无人车团队<sup>[26]</sup>建立了并行层次状态机,构建决策系统应对环境中出现的各类情况。布伦瑞克工业大学提出的移动导航分布式体系结构,包含一系列驾驶行为(跟随道路点、车道保持、避障、行驶在停车区等),通过投票机制决策驾驶行为并应用于Caroline号无人车<sup>[27-28]</sup>上。

并联式结构将每一种工况单独划分成模块进行处理,整个系统可快速、灵活地对输入进行响应。但在复杂工况下,由于遍历状态较多导致的算法机构庞大,以及状态间的划分与状态冲突的解决是难点。并联结构适用于场景较复杂的工况。相较于串联结构,并联结构的优点是具备场景遍历广度优势,易于实现复杂的功能组合,具有较好的模块性与拓展性,缺点是系统不具备时序性,缺乏场景遍历的深度,决策易忽略细微环境变化,状态划分灰色地带难以处理,从而导致决策错误。

### 3.2.3 混联结构

串、并联结构具备各自的局限性,混联式结构可较好地结合两者优点,层级式混联结构是比较典型的方法。

卡耐基梅隆大学与福特公司研发的Boss无人车<sup>[12]</sup>行为决策系统如图6所示,其具备典型的层级式混联结构。系统顶层基于场景行为划分,底层基于自车行为划分。3个顶层行为及其底层行为分别为:车道保持(车道选择、场景实时报告、距离保持器、行为发生器等)、路口处理(优先级估计、转移管理等)和指定位姿。

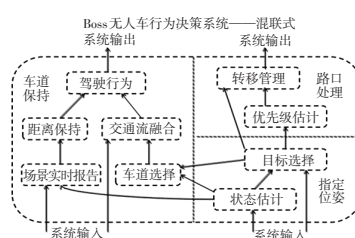


图6 Boss无人车的行为决策系统

弗吉尼亚理工大学研发的Odin无人车<sup>[14]</sup>行为决策



系统如图7所示,该系统引入决策仲裁机制,其决策系统划分为车道保持、超车、汇入交通流、U型弯、拥堵再规划等模块。每个子决策模块输出的结果均交由决策融合器进行决策仲裁。各模块具备不同优先级,优先级低的模块必须让步于优先级高的模块。

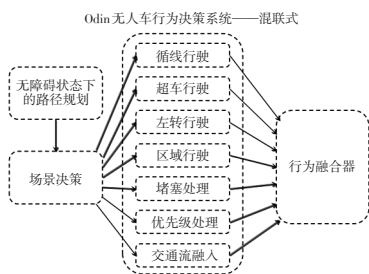


图7 Odin无人车行为决策系统

中国科学技术大学研发的智能驾驶Ⅱ号<sup>[11]</sup>行为决策系统如图8所示,其具备典型的混联结构。该系统进行了专家算法和机器学习算法的融合,顶层决策系统采用并联式有限状态机,分为跟车巡航、十字路口、U型弯、自主泊车等模块。底层采用学习算法(ID3决策树法),用以得出车辆的具体目标状态及目标动作。

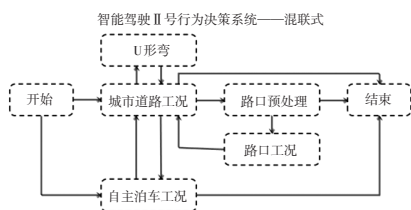


图8 智能驾驶Ⅱ号行为决策系统

这种层级处理的思想还衍生出各种具体的应用方法。康奈尔大学的Skynet号无人车<sup>[29]</sup>通过结合交通规则和周围环境,建立了3层规划系统,划分驾驶行为、规划策略与底层操作控制,实现对车辆行为的决策处理。清华大学的THMR-V无人车<sup>[30]</sup>采用分层式的体系结构,将系统分为协调级和执行级,通过协调级对环境和车辆状态进行局部规划,最高时速可达百公里。佛罗里达大学的Knight Rider号无人车<sup>[31]</sup>包含人工智能模块,建立了层次结构驾驶员模型,将驾驶任务分为战略、战术和操作层级,通过战略层设定次级目标来生成序列最优战术,并经过操作层输出控制信号。来自宾州大学的Little Ben号无人车研发团队<sup>[32]</sup>建立了简单的规则对比赛给定的任务文件进行驾驶行为的划分,结合路网信息确定车辆行为,以此作为车辆的上层决策系统。

#### 4 基于学习算法的无人车行为决策系统

近年来,人工智能技术迅猛发展,学习算法越来越多地运用于无人驾驶车辆环境感知与决策系统<sup>[33]</sup>。基

于学习算法的无人车行为决策系统研究目前已取得显著成果,根据原理不同主要可分为深度学习相关的决策方法<sup>[34-35]</sup>与基于决策树等机器学习理论的决策方法。NVIDIA端到端卷积神经网络决策系统<sup>[19]</sup>与中国科学技术大学应用的ID3决策树法<sup>[11]</sup>是其典型案例。

##### 4.1 深度学习方法

深度学习方法因其在建模现实问题上极强的灵活性,近年来被许多专家、学者应用于无人车决策系统。NVIDIA研发的无人驾驶车辆系统架构是一种典型架构,其采用端到端卷积神经网络进行决策处理,使决策系统大幅简化。系统直接输入由相机获得的各帧图像,经由神经网络决策后直接输出车辆目标转向盘转角。

该系统使用NVIDIA DevBox作处理器,用Torch 7作为系统框架进行训练,工作时每秒处理30帧数据,其训练系统框架如图9所示。图像输入到卷积神经网络(Convolutional Neural Networks, CNN)计算转向控制命令,将预测的转向控制命令与理想的控制命令相比较,然后调整CNN模型的权值使得预测值尽可能接近理想值。权值调整由机器学习库Torch 7的反向传播算法完成。训练完成后,模型可以利用中心的单个摄像机数据生成转向控制命令。

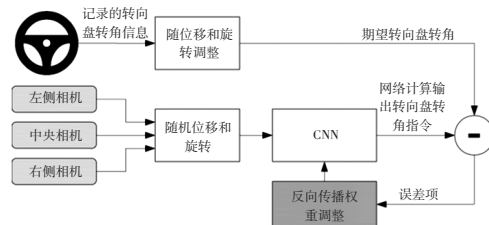


图9 NVIDIA无人车决策系统训练模型

其深度学习系统网络结构如图10a所示,共9层,包括1个归一化层、5个卷积层和3个全连接层。输入图像被映射到YUV颜色空间,然后传入网络。

仿真结果表明,其神经网络能完整地学习保持车道驾驶的任务,而不需要人工将任务分解为车道检测、语义识别、路径规划和车辆控制等。CNN模型可以从稀疏的训练信号(只有转向控制命令)中学到有意义的道路特征,100 h以内的少量训练数据就足以完成在各种条件下操控车辆的训练。

百度端到端系统实现了对车辆的纵向和横向控制:纵向控制采用堆叠卷积长短期记忆(Long Short Term Memory, LSTM)深度学习模型,提取帧序列图像中的时空特征信息,实现特征到纵向控制指令的映射;横向控制采用CNN深度学习模型,从单幅前视相机的图像直接计算出横向控制的曲率。模型主要关注视觉特征的

提取、时序规律的发现、行为的映射等方面。

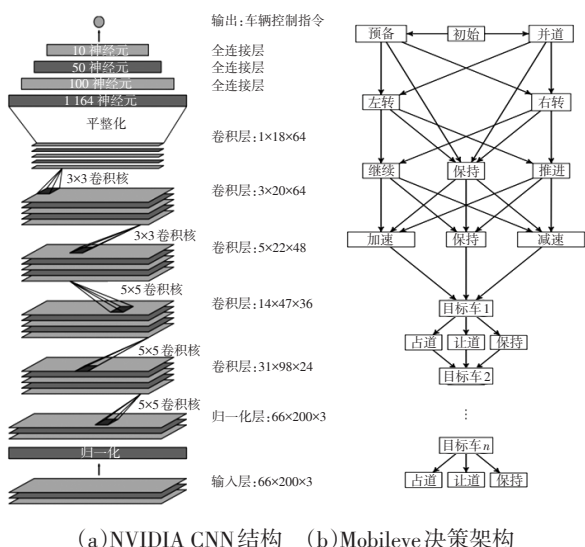


图 10 NVIDIA 与 Mobileye 决策架构

其中,纵向控制被看作时空序列预测问题,输入单元为最近 5 帧图像(图像采集频率是每秒 8 帧),每帧图像均缩放为 80 像素×80 像素的 RGB 格式。LSTM 模型的第 1 层有 64 个通道,其内核大小为 5 像素×5 像素,后续层拥有更多的通道和更小的内核,最后一个卷积层为 2 个完全连通的层。输出单元是线性单元,损失函数是 MSE,优化器是 rmsprop。横向控制由 1 个预处理层、5 个卷积层和 2 个全连接层组成。输入为 320 像素×320 像素的 RGB 格式图像。

Intel<sup>[20]</sup>利用已有控制数据训练网络,完成端到端的自动驾驶,如图 11 所示。其在网络中考虑了方向性的控制指令(直行、左转、右转),使得网络可以在车道保持的同时完成转弯等操作。图 11 为 2 种不同的结合控制指令的结构:一种作为网络的输入,另一种将指令分为 3 个不同的输出层,根据控制指令选择不同的输出。

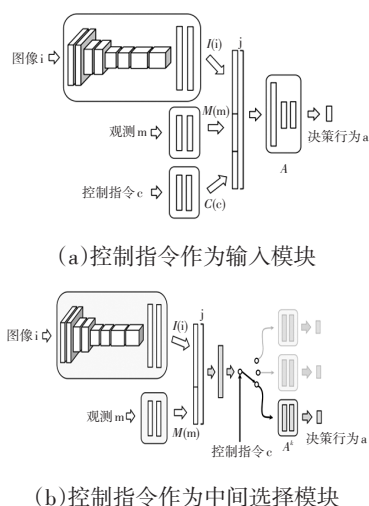


图 11 Intel 决策网络架构

Comma.ai<sup>[21]</sup>利用 CNN,并且几乎只用 CNN 来构造决策系统。该系统将图像导入网络,通过网络输出命令调整转向盘和车速,从而使车辆保持在车道内。根据 Comma.ai 公布的数据,该公司目前已经累计行驶约  $1.35 \times 10^6$  km,累计行驶时间为 22 000 h,累计用户 1 909 人。

美国伍斯特理工学院<sup>[36]</sup>提出了一种自动驾驶汽车的端到端学习方法,能够直接从前视摄像机拍摄的图像帧中产生适当的转向盘转角。主要使用卷积神经网络将特征级的图像数据作为输入,驾驶员的转向盘转角作为输出进行训练和学习,使用 Comma.ai 公布的数据集进行训练和评估。试验结果表明,该模型能够实现相对精确的转向控制,很好地完成了车道保持动作。

Mobileye<sup>[22]</sup>把增强学习应用在高级驾驶策略的学习上,感知及控制等模块则被独立出来处理,其系统结构如图 10b 所示。相对于端到端学习大大提高了决策过程的可解释性和可操作性,很好地适配了传统机器人学中感知-决策-控制的系统架构。

Drive.ai 获准在美国加州公共道路上测试无人驾驶汽车,其在感知和决策上都使用深度学习,但避免整体系统的端到端,而是将系统按模块分解,再分别应用深度学习,同时结合规则、知识确保系统的安全性。

Waymo 通过模拟驾驶及道路测试获取了大量的数据对其行为决策系统进行训练。该系统不仅能对物体进行探测,还能对障碍物进行语义理解。对不同道路参与者的行为方式建立准确的模型,判断它可能的行为方式以及对汽车自身的道路行为产生的影响,输入到决策系统,保证决策行为的安全性。

卡耐基梅隆大学<sup>[37]</sup>提出了一种基于预测和代价函数算法(Prediction and Cost function-Based algorithm, PCB)的离线学习机制,用于模拟人类驾驶员的行为决策。其决策系统针对交通场景预测与评估问题建立模型,使用学习算法,利用有限的训练数据进行优化。训练数据来源于人类驾驶员在实际道路的跟车场景,主要包含自车车速、前车车速以及两车之间的距离等。在 120 km 的低速跟车测试中,PCB 和人类跟车的车速差异仅为 5%,能够很好地完成跟车操作。

国防科技大学的刘春明教授等人<sup>[38]</sup>构建了 14 自由度的车辆模型,采用模型控制预测理论,利用基于增强学习理论的方法,基于仿真数据得到了无人车的决策模型。该方法利用多自由度车辆模型对车辆的实际动力学特性进行考量,有利于满足车辆行驶稳定与乘员舒适的要求。

麻省理工大学<sup>[39]</sup>在仿真器中模拟单向 7 车道工况,

利用定义好的深度强化学习(Deep Q-Learning, DQN)模型调整网络结构,可在浏览器上进行训练工作,完善决策系统。

## 4.2 其他机器学习方法

除基于学习的决策方法外,还有很多机器学习方法在决策系统中加以应用。决策树法为机器学习理论中一种具有代表性的方法,中国科技大学的智能驾驶Ⅱ号将其用于决策系统。其应用的ID3决策树法适用于多种具体工况,如路口、U形弯工况等,其先由顶层有限状态机决策出具体场景,再进入决策树进行相应的计算。

以十字路口工况为例,首先确定当前工况的条件属性(即系统输入,如自车车速、干扰车车速等)和决策属性(即系统输出,如加速直行、停车让行等)。选取若干样本数据进行基于灰关联熵的条件属性影响分析,获得如图12所示的基于ID3算法的行为决策树。

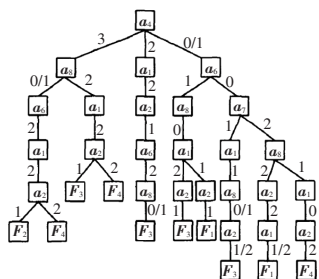


图12 基于ID3算法的行为决策树

该行为决策树即机器通过学习后自主获得的行为规则库的一种表现形式。无人车运行时,将驾驶环境信息转化成条件属性,交由决策树进行计算,最终得出决策指令,指导无人车的行为操作。

ID3决策树法具有知识自动获取、准确表达、结构清晰简明的优点,其缺点同样明显,即对于大量数据获取的难度较大,数据可靠性不足,数据离散化处理后精度不足。

德国宝马和慕尼黑工业大学<sup>[40]</sup>提出了一种基于部分可观测马尔科夫决策过程(Partially Observable Markov Decision Processes, POMDP)的决策模型,主要解决动态和不确定驾驶环境下的决策问题,其不确定性主要来源于传感器噪声和交通参与者行驶意图的不确定性。

POMDP将其他车辆的驾驶意图作为隐藏变量,建立贝叶斯概率模型,可求解出自车在规划路径上的最优加速度。在复杂交叉路口的仿真测试下,能够较好地根据其他车辆的驾驶行为调整自车的最优加速度,保证安全性与行车效率。

丰田公司的Urtasun等人<sup>[41]</sup>提出了一种基于概率的生成式模型,提取室外环境的语义信息作为输入,并输

出行为决策。

## 5 无人车行为决策系统发展趋势

### 5.1 现阶段研究水平

无人车行为决策系统研究水平的衡量标准主要体现在实车应用性、实现功能的复杂程度、应用场景的复杂程度、决策结果正确性与系统复杂性。

实车应用是对决策系统的基本要求。现阶段,有限状态机法作为基于规则的代表方法,已经较为成熟并广泛应用于国内外诸多无人车上。基于学习算法的决策算法,如深度学习相关方法、决策树法、粗糙集理论等学习算法近年来发展迅速,亦具有很多相关应用实例并取得很好效果。截至2017年11月,谷歌自动驾驶汽车部门Waymo已在美国6个州开展了自动驾驶测试,宣称其测试用无人车已行驶 $5.6 \times 10^6$  km并完成了 $4 \times 10^9$  km的模拟驾驶。特斯拉的量产车中安装了Autopilot硬件系统,可通过空中下载(Over-The-Air, OTA)技术进行从L2到L4的升级。2017年上半年以来,陆续开启了车道偏离警告、自动转向(Autosteer)、召唤(Summon)以及自动车道变换等功能。

在功能复杂程度上,现阶段的有限状态机决策技术除实现简单的循线行驶、车道保持、紧急避障等功能外,也可实现较为复杂的组合功能<sup>[42-43]</sup>,如路口处理、自主泊车,乃至一定工况范围内的长距离自动驾驶。状态机方法可并行遍历多个场景,且擅长在特定场景内对任务进行拆分决策。而现阶段学习算法的功能应用依赖于训练集的丰富程度,一般在指定工况与任务下效果较好,多场景与多任务协调能力仍相对较差。如北京理工大学的BYD-Ray<sup>[44]</sup>应用粗糙集理论主要解决换道决策的功能,智能驾驶Ⅱ号的1个决策树只能解决1个功能下的决策任务。以谷歌为代表的端到端相关方法取得了较好效果,但其对于训练数据较为依赖,决策网络可解释性与可调整性较差。

在应用场景的复杂程度上,现阶段的有限状态机决策技术可以适应复杂的综合城市环境,其优势在于场景广度的遍历。Boss、Junior、Odin等在DARPA无人车挑战赛中于6 h内完成96 km的路测,其场景完全还原了真实城市道路环境,需要车辆与其他车辆交互、避让,同时遵守交通规则,具有很高的环境复杂度。而学习算法更加着重于场景深度的遍历,对于同一场景下的各种细微变化更加应变自如<sup>[45-46]</sup>。NVIDIA无人车利用训练后的CNN模型,可以使车辆在同一类场景下在面对不同道路材质、车道线线型、各种天气状况等细微环境变化



时做出正确反应。Drive.ai 共有 4 辆车在旧金山湾区进行路测,在夜路、雨天、有雾等复杂情况下也基本可以实现完全自动驾驶。

在决策结果的正确性上,学习算法与规则算法各有优势,在实现较高的功能复杂度和适应较高场景复杂度的基础上,两种算法均能够保证决策系统的正确性,完成相应的行驶任务。但对于复杂工况,规则算法面临状态划分界限确定问题。学习算法需要遍历程度高的数据作为支持。

在系统复杂性上,基于学习的算法舍弃了规则算法的层级架构,更加简洁直接,大幅简化了决策系统结构<sup>[47]</sup>。规则算法面临复杂工况,不可避免因遍历工况、完善逻辑而导致算法规模庞大冗杂。并且对于复杂工况下算法性能的提高问题,规则算法因其逻辑结构问题,难以大幅提升,存在瓶颈。

根据国际自动机工程师学会(SAE)的自动驾驶分级,现阶段基于专家和学习算法的无人车基本可达到 L3 级有条件自动驾驶水平,部分可以达到 L4 级高度自动驾驶水平。通用汽车于 2018 年 1 月发布了新一代无人驾驶汽车——Cruise AV,这是世界首辆无需驾驶员、转向盘和踏板即可实现安全驾驶的量产汽车,可以达到 L4 级别。百度于 2015 年 12 月完成了北京中关村软件园的百度大厦至奥林匹克森林公园并原路返回的道路测试,全程均无人驾驶,最高车速达 100 km/h。2018 年 2 月的春节联欢晚会中,百度 Apollo 无人车队在港珠澳大桥上以无人驾驶模式完成了 8 字交叉跑等高难度驾驶动作,标志着我国无人驾驶技术亦取得了长足进步。

## 5.2 技术难点

基于规则和基于学习算法的行为决策技术各自具备较为鲜明的优缺点。

基于专家规则的行为决策系统的优点是:算法逻辑清晰,可解释性强,稳定性强,便于建模;系统运行对处理器性能要求不高;模型可调整性强;可拓展性强,通过状态机的分层可以实现较为复杂的组合功能;在功能场景的广度遍历上存在优势。其缺点是:由于状态切割划分条件导致车辆行为不连贯;行为规则库触发条件易重叠而造成系统失效;有限状态机难以完全覆盖车辆可能遇到的所有工况,通常会忽略可能导致决策错误的细节;场景深度遍历不足导致系统决策正确率难以提升,对复杂工况处理及算法性能的提升存在瓶颈。

基于学习算法的行为决策系统的优点是具备场景遍历深度的优势,针对某一细分场景,通过大数据系统更容易覆盖全部工况;利用网络结构可简化决策算法规模;部分机器具备自学习性能,机器能够自行提炼环境

特征和决策属性,便于系统优化迭代<sup>[48-49]</sup>;不必遍历各种工况,通过数据的训练完善模型,模型正确可率随数据的完备得以提升。其缺点是:算法决策结果可解释性差,模型修正难度大;学习算法不具备场景遍历广度优势,不同场景所需采用的学习模型可能完全不同;机器学习需要大量试验数据作为学习样本;决策效果依赖数据质量,样本不足、数据质量差、网络结构不合理等会导致过学习、欠学习等问题。

根据上述决策系统的优缺点,提出现阶段决策系统需要解决的具体技术难点:

a. 基于有限状态机决策模型的状态划分问题。有限状态机的状态划分需要依据明确的边界条件。但实际驾驶过程中,驾驶行为间存在某些“灰色地带”,即同一场景下可能有 1 个以上合理的行为选择,使驾驶状态存在冲突。对于决策系统而言,一方面要避免冲突状态强行划分而造成的无人车行为不连贯,另一方面也要能够判断处于“灰色地带”的无人车不同行为的最优性。通过在决策系统中引入其他决策理论,如 DS 证据理论(Dempster-Shafer Evidence Theory)等、决策仲裁机制、博弈论法、状态机与学习算法结合等方法可帮助解决该问题。

b. 基于有限状态机决策模型的复杂场景遍历问题。有限状态机需要人工设定规则库以泛化无人车行驶状态,这种模式使得状态机具有广度遍历优势。但随着环境场景的增多与复杂,有限状态机的规模也不断庞大,使得算法臃肿;而同一场景的深度遍历,即由于环境细节变化导致的决策结果变化,也很难通过状态遍历。对于有限状态机决策系统而言,利用学习算法的优点,采用状态机与学习算法结合的方法,结合学习算法深度遍历的优势,有利于有效解决该问题。

c. 基于学习算法决策模型的正确性与稳定性问题。学习算法的训练结果与样本数量、样本质量和网络结构有关。如果学习模型过于复杂,会造成无法区分数据本质和噪声的情况,泛化能力减弱;反之,如果学习模型过于简单,或者样本数量不足、对场景遍历不足,则会导致规则提炼不精准,出现欠学习问题。此外,学习算法模型逻辑解释性较差,实际应用中调整与修正很不方便。对于基于学习算法的决策系统而言,一方面要收集大量可靠、高质量的试验数据,另一方面要选择合理的学习算法,配置合理的试验参数,调整网络结构,以提高训练结果的正确性与稳定性。

## 5.3 技术发展趋势

随着计算机科学的迅猛发展,近年来掀起了机器学习技术的研究浪潮,学习算法越来越多地运用于无人车

行为决策系统,计算机硬件平台的发展也为深度学习网络训练提供了强有力的支持。

基于规则的决策方法相对较为成熟,其在场景遍历广度上具备优势,逻辑可解释性强,易于根据场景分模块设计,国内外均有很多应用有限状态机的决策系统实例。然而其系统结构决定了其在场景遍历深度、决策正确率上存在一定的瓶颈,难以处理复杂工况<sup>[50]</sup>。

基于学习算法的决策系统因具有场景遍历深度的优势,将被越来越多地用作决策系统的底层,即针对某一细分场景,采用学习算法增强算法的场景遍历深度,使其能够在环境细微变化中仍然保证较高的决策精度。然而其算法可解释性差、可调整性差、场景广度遍历不足等劣势导致了仅采用学习算法的决策系统仍存在应用局限,较难处理复杂的功能组合<sup>[51]</sup>。

根据上述基于两种算法的优、劣势,现阶段无人车决策系统的发展趋势可归纳为:

a. 采用基于规则算法的行为决策算法仍会在决策系统中广泛应用,将作为决策系统的顶层架构与某些具体问题的细分解决方案,并将更多地采用混联结构,发挥规则算法基于场景划分模块处理及针对具体问题细分处理时逻辑清晰、调整性强的优势,可同时兼顾场景遍历的广度与深度。采用该方法的研究重点将在于解决状态划分“灰色地带”的合理决策问题,以及行为规则库触发条件重叠等问题。

b. 无人车决策系统将更多地采用规则算法与学习算法结合的方式。顶层采用有限状态机,根据场景进行层级遍历;底层采用学习算法,基于具体场景分模块应用,可发挥学习算法优势,简化算法结构、增强场景遍历的深度,并可减小数据依赖量,保证决策结果的鲁棒性与正确性。采用该方法的研究重点在于如何合理对接有限状态机与学习算法模型,以及学习算法的过学习、欠学习等问题。

c. 端到端方法将更多作为决策子模块的解决方案,而非将决策系统作为一个整体进行端到端处理。通过这种方式可发挥学习算法的优势,将决策模块拆解也可提高系统的可解释性与可调节性。

d. 目前行为决策系统的设计准则主要考量安全与效率,对车辆特性与乘员舒适性考虑较少。在保证安全与效率的基础上,可通过加入对车辆动力学特性的考量,筛选更合理的驾驶数据等方式,对行为决策系统进行优化。

#### 参 考 文 献

- [1] 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.
- [2] 孙振平. 自动驾驶汽车智能控制系统[D]. 长沙: 国防科学

技术大学, 2004.

- [3] 张荷芳, 豆菲菲. 多传感器信息融合的无人车行驶策略[J]. 计算机与数字工程, 2015(3): 392-395.
- [4] Choi J, Ulbrich s, Lichte B, et al. Multi-Target Tracking Using a 3D- Lidar Sensor for Autonomous Vehicles[C]// International IEEE Conference on Intelligent Transportation Systems. IEEE, 2013: 881-886.
- [5] Cho H, Seo Y W, Kumar B V K V, et al. A Multi-Sensor Fusion System for Moving Object Detection and Tracking in Urban Driving Environments[C]// IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2014: 1836-1843.
- [6] Davison A J. FutureMapping: The Computational Structure of Spatial AI Systems[J]. 2018.
- [7] Thrun S. Probabilistic Robotics[J]. Communications of the Acm, 2005, 45(3): 52-57.
- [8] Zhang J, Singh S. LOAM: Lidar Odometry and Mapping in Real-Time[C]// Robotics: Science and Systems. 2014, 2: 9.
- [9] 谭同超. 有限状态机及其应用[D]. 广州: 华南理工大学, 2013.
- [10] 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, 2015, 59: 216-232.
- [11] 杜明博. 基于人类驾驶行为的无人驾驶车辆行为决策与运动规划方法研究[D]. 合肥: 中国科学技术大学, 2016.
- [12] Urmson C, Anhalt J, Bagnell D, et al. Autonomous Driving in Urban Environments: Boss and the Urban Challenge[J]. Journal of Field Robotics, 2008, 25(8): 425 - 466.
- [13] Montemerlo M, Becker J, Bhat S, et al. Junior: The Stanford entry in the Urban Challenge[J]. Journal of Field Robotics, 2008, 25(9): 569-597.
- [14] Bacha A, Bauman C, Faruque R, et al. Odin: Team VictorTango's entry in the DARPA Urban Challenge[J]. Journal of Field Robotics, 2008, 25(8): 467-492.
- [15] Leonard J, How J, Teller S, et al. A Perception- Driven Autonomous Urban Vehicle[J]. Journal of Field Robotics, 2008, 25(10): 727-774.
- [16] Ziegler J, Bender P, Schreiber M, et al. Making Bertha Drive- An Autonomous Journey on a Historic Route[J]. IEEE Intelligent Transportation Systems Magazine, 2015, 6 (2): 8-20.
- [17] Ross S, Melik-Barkhudarov N, Shankar K S, et al. Learning Monocular Reactive UAV Control in Cluttered Natural Environments[C]// IEEE International Conference on Robotics and Automation. IEEE, 2013: 1765-1772.
- [18] Isele D, Rahimi R, Cosgun A, et al. Navigating Occluded Intersections with Autonomous Vehicles Using Deep Reinforcement Learning[J]. 2017.



- [19] Bojarski M, Del Testa D, Dworakowski D, et al. End to End Learning for Self-Driving Cars[J]. 2016.
- [20] Codevilla F, Müller M, López A, et al. End-to-end Driving via Conditional Imitation Learning[J]. 2017.
- [21] Santana E, Hotz G. Learning a Driving Simulator[J]. 2016.
- [22] Shalevshwartz S, Shammah S, Shashua A. Safe, Multi-Agent, Reinforcement Learning for Autonomous Driving[J]. 2016.
- [23] 熊光明, 李勇, 王诗源. 基于有限状态机的智能车辆交叉口行为预测与控制[J]. 北京理工大学学报, 2015, 35(1): 34-38.
- [24] 陈雪梅, 田庚, 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): 492-496.
- [25] Furda A, Vlacic L. Enabling Safe Autonomous Driving in Real-World City Traffic Using Multiple Criteria Decision Making[J]. IEEE Intelligent Transportation Systems Magazine, 2011, 3(1): 4-17.
- [26] Kammel S, ren, Ziegler J, et al. Team AnnieWAY's Autonomous System for the 2007 DARPA Urban Challenge [J]. Journal of Field Robotics, 2008, 25(9): 615-639.
- [27] Rauskolb F W, Kai B, Lipski C, et al. Caroline: An Autonomously Driving Vehicle for Urban Environments[J]. Journal of Field Robotics, 2008, 25(9): 674-724.
- [28] Miller I, Campbell M, Huttenlocher D, et al. Team Cornell's Skynet: Robust perception and Planning in an Urban Environment[J]. Journal of Field Robotics, 2008, 25(8): 493-527.
- [29] Miller I, Campbell M, Dan H, et al. Team Cornell's Skynet: Robust Perception and Planning in an Urban Environment [J]. Journal of Field Robotics, 2008, 25(8): 493-527.
- [30] 张朋飞, 何克忠, 欧阳正柱, 等. 多功能室外智能移动机器人实验平台——THMR-V[J]. 机器人, 2002, 24(2): 97-101.
- [31] Patz B J, Papelis Y, Pillat R, et al. A Practical Approach to Robotic Design for the DARPA Urban Challenge[J]. Journal of Field Robotics, 2008, 25(8): 528-566.
- [32] Bohren J, Foote T, Keller J, et al. Little Ben: The Ben Franklin Racing Team's Entry in the 2007 DARPA Urban Challenge[J]. Journal of Field Robotics, 2008, 25(9): 598-614.
- [33] Pan X, You Y, Wang Z, et al. Virtual to Real Reinforcement Learning for Autonomous Driving[J]. 2017.
- [34] Sebastian B, Tobias G, Rüdiger D. Probabilistic Decision-Making under Uncertainty for Autonomous Driving using Continuous POMDPs[J]. IEEE 17th International Conference on Intelligent Transportation Systems (ITSC), 2014: 392-399.
- [35] Barto A G, Mahadevan S. Recent Advances in Hierarchical Reinforcement Learning[J]. Discrete Event Dynamic Systems, 2003, 13(4): 341-379.
- [36] Chen C, Seff A, Kornhauser A, et al. DeepDriving: Learning Affordance for Direct Perception in Autonomous Driving [C]// IEEE International Conference on Computer Vision. IEEE, 2015: 2722-2730.
- [37] Dolan J M. A Learning-Based Autonomous Driver: Emulate Human Driver's Intelligence in Low-Speed Car Following [J]. Proceedings of SPIE - The International Society for Optical Engineering, 2010, 7693.
- [38] Zheng R, Liu C, Guo Q. A Decision-Making Method for Autonomous Vehicles Based on Simulation and Reinforcement Learning[C]// International Conference on Machine Learning and Cybernetics. IEEE, 2014: 362-369.
- [39] Fridman L, Jenik B, Terwilliger J. DeepTraffic: Driving Fast through Dense Traffic with Deep Reinforcement Learning[J]. 2018.
- [40] Hubmann C, Becker M, Althoff D, et al. Decision Making for Autonomous Driving Considering Interaction and Uncertain Prediction of Surrounding Vehicles[C]// Intelligent Vehicles Symposium. IEEE, 2017: 1671-1678.
- [41] Zhang H, Geiger A, Urtasun R. Understanding High-Level Semantics by Modeling Traffic Patterns. IEEE International Conference on Computer Vision. 2013: 3056-3063.
- [42] 陈平. 辅助驾驶中控制与决策关键技术研究[D]. 上海: 上海交通大学, 2011.
- [43] 黄子超. 基于有限状态机的车队纵向协作策略建模与仿真[D]. 武汉: 武汉理工大学, 2013.
- [44] 田庚. 复杂动态城市环境下无人驾驶车辆仿生换道决策模型研究[D]. 北京: 北京理工大学, 2016.
- [45] Xu H, Gao Y, Yu F, et al. End-to-end Learning of Driving Models from Large-scale Video Datasets[J]. 2016: 3530-3538.
- [46] 张文明, 韩泓冰, 杨珏, 等. 基于驾驶员行为的神经网络无人驾驶控制[J]. 华南理工大学学报, 2016, 44(12): 74-80.
- [47] Argall B D, Chernova S, Veloso M, et al. A Survey of Robot Learning from Demonstration[J]. Robotics & Autonomous Systems, 2009, 57(5): 469-483.
- [48] Broad A, Arkin J, Ratliff N, et al. Real-Time Natural Language Corrections for Assistive Robotic Manipulators[J]. International Journal of Robotics Research, 2017, 36(5-7).
- [49] Dosovitskiy A, Koltun V. Learning to Act by Predicting the Future[J]. 2016.
- [50] 陈慧, 徐建波. 智能技术发展趋势[J]. 中国集成电路, 2014, 186(11): 64-70.
- [51] 伦一. 自动驾驶产业发展现状及趋势[J]. 电信网技术, 2017(6): 33-36.

(责任编辑 斛 畔)

修改稿收到日期为2018年7月20日。