

·智能车辆运动规划与控制技术专题·

基于模糊收敛和模仿强化学习的自动驾驶横向控制方法*

郑川 杜煜 刘子健

(北京联合大学,北京市信息服务工程重点实验室,北京 100101)

【摘要】针对自动驾驶车辆各横向控制因素存在强耦合性,依赖理想模型的控制方法难以实现完全解耦且难以从仿真环境迁移到实际车辆,以及强化学习方法在自动驾驶横向控制中收敛速度慢的问题,利用模糊推理器和模仿强化学习在车辆横向控制中的共性,以模糊推理器作为模仿强化学习初始化条件,并进行学习过程指导,从而实现强化学习过程的快速收敛。利用MATLAB/Carla仿真以及实车试验对该控制方法进行了验证,结果表明,该方法在不依赖理想数学模型、不对模糊推理器进行深度优化的基础上,实现了模仿强化学习迭代次数的大幅减少,在500次的全路径迭代过程中完成了更优的车辆横向控制,在仿真和现实环境中均可获得很好的控制效果。

主题词:自动驾驶 横向控制 模仿学习 强化学习 模糊推理

中图分类号:U471.1 **文献标志码:**A **DOI:** 10.19620/j.cnki.1000-3703.20230009

A Lateral Control Method of Autonomous Driving Based on Fuzzy Convergence and Imitative Reinforcement Learning

Zheng Chuan, Du Yu, Liu Zijian

(Beijing Key Laboratory of Information Service Engineering, Beijing Union University, Beijing 100101)

【Abstract】In view of the strong coupling of each control factor in the lateral control of autonomous vehicles, it is difficult for the control method relying on the ideal model to completely decouple and migrate from the simulation environment to the actual vehicle, and the problem that the convergence speed of the reinforcement learning method in the lateral control of autonomous vehicles is not ideal, the fuzzy inference machine and the similarity of the simulation reinforcement learning in the lateral control factors of vehicles are used to combine the two. A fuzzy inference machine is used as the initialization condition for simulated reinforcement learning, and provide guidance for the learning process, thus achieving rapid convergence of the learning process. The MATLAB/Carla simulation and vehicle test are applied to verify the control method. The results show that the method can significantly reduce the number of simulation reinforcement learning iterations, achieve better vehicle lateral control performance in 500 full path iterations, and achieve good control effect in both simulation and real environment on the basis of not relying on the ideal mathematical model and not having to carry out in-depth optimization of the fuzzy inference device.

Key words: Autonomous driving, Lateral control, Imitative learning, Reinforcement learning, Fuzzy inference

【引用格式】郑川,杜煜,刘子健.基于模糊收敛和模仿强化学习的自动驾驶横向控制方法[J].汽车技术,2024(7):29-36.

ZHENG C, DU Y, LIU Z J. A Lateral Control Method of Autonomous Driving Based on Fuzzy Convergence and Imitative Reinforcement Learning[J]. Automobile Technology, 2024(7): 29-36.

1 前言

为确保自动驾驶车辆在路径规划^[1]的基础上实现精准、安全与稳定的横向控制,研究人员根据自动驾驶汽车

的横向运动模型^[2-4]提出了多种横向控制方法。文献[5]基于合适的预瞄距离,采用纯跟踪(Pure Pursuit)算法对斯坦利(Stanley)算法进行改进,提出了一种新的融合控制算法。文献[6]提出了一种基于无模型自适应控制的

*基金项目:北京市朝阳区科技局“智能配送物流机器人协同创新中心”项目。

通信作者:杜煜(1972—),男,博士,教授,主要研究方向为通信网服务质量和智能驾驶相关技术等,duyu@bnu.edu.cn。

无人驾驶汽车横向控制方案。文献[7]、文献[8]设计了基于强化学习的无人驾驶车辆横向控制策略。文献[9]将模型预测控制(Model Predictive Control, MPC)算法和车辆动力学模型相结合并通过仿真验证了MPC算法在车辆横向控制上的准确性和可行性。文献[10]在车辆动力学模型的基础上设计了有前馈补偿的线性二次型调节器(Linear Quadratic Regulator, LQR)算法。文献[11]对基于MPC的智能驾驶汽车横向运动控制进行了研究及仿真验证。然而,这些方法均对模型进行了理想化的假设,如在车辆动力学模型中忽略了车辆横向偏差在距离偏差和方向偏差上的非耦合关系。另外,深度强化学习的方法需要进行基于大量样本的训练,对于以MPC为代表的控制算法,其精度较高的模型往往具有较高的计算代价,难以保证实时性^[7]。由于各控制因素之间具有强耦合性,只考虑部分控制因素并不能实现车辆的精确控制。

建立自动驾驶汽车横向控制模型时,模糊推理^[12]与强化学习^[13-17]的方法有其内在的共性。模糊推理器需寻求输入前件和输出后件语义化并满足特定映射条件。实际上,无模型的强化学习方法在确立其状态空间和行为空间过程中,也在寻求从一个状态到另一个状态的可能性,并利用这种可能性得到最优映射关系。

基于以上分析,本文提出基于模糊推理和模仿强化学习的横向控制方法(A lateral control method based on Fuzzy Inference and Reinforcement Learning, FIRL),在初始的自动驾驶汽车模糊推理器(横向控制收敛)基础上进行强化学习迭代,并在学习过程中以模糊推理器作为强化学习的随机因素,“指导”车辆的学习过程,以大幅减少学习过程中强化学习的迭代次数。

2 模糊推理器设计

图1所示为基于轨迹点的车辆运动模型。其中, $x-y$ 为车辆坐标系, y 轴正方向为车辆前进方向, TH 为轨迹点的航向(轨迹点的切线方向),与 y 轴所成夹角 φ 为方向偏差(Heading), D 为轨迹点与车辆坐标系原点的横向偏差(Deviation)。

从图1中可以看出,车辆沿轨迹行驶时,横向偏差和方向偏差越小,与轨迹的重合度就越高。当车辆由模糊推理器进行横向控制时,在前件的模糊集合中,如果横向偏差和方向偏差处于某个状态,转向盘将对应一个合适的转角。输入这2个变量时,最终通过模糊规则输出转向盘转角实现车辆的横向控制。

以车辆横向偏差和方向偏差作为模糊推理器的输入前件,以转向盘的目标转角(Steering)作为输出后

件,在部分专家经验指导下生成模糊逻辑规则,初步建立一个基于部分专家规则库的模糊推理系统(Fuzzy Inference System, FIS),该系统框架如图2所示。

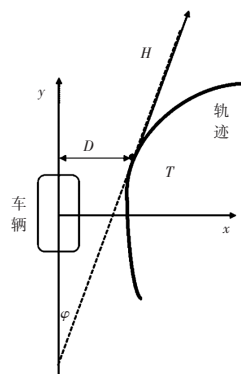


图1 智能驾驶车辆横向控制模型

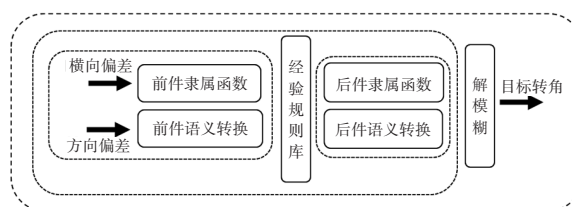


图2 模糊推理系统框架

2.1 前、后件的语义划分

车辆定位点与轨迹点的横向偏差的语义设置为偏左很远(LDFar)、偏左较远(LDNear)、偏左较小(DZZero)、偏右较远(RDNear)、偏右很远(RDFar);车辆行驶方向与轨迹点切线方向的方向偏差的语义设置为偏左很多(LHFar)、偏左较多(LHNear)、偏左较小(HZZero)、偏右较多(RHNear)、偏右很多(RHFar);车辆转向盘的目标转角的语义设置为向左转角很大(LSBig)、向左转角较大(LSSmall)、转角较小(SZZero)、向右转角较大(RSSSmall)、向右转角很大(RSBig)。

2.2 隶属度函数的选取

本文以设计的模糊推理系统作为强化学习的初始解和学习过程的随机因素来源。对整个模糊推理系统的设计要求无需过于严格,只要保证在车辆全路径跟踪时满足收敛条件即可。对前件和后件统一采用三角隶属度函数表述:

$$f(x,a,b,c) = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a < x \leq b \\ \frac{c-x}{c-b}, & b < x \leq c \\ 0, & x > c \end{cases} \quad (1)$$

式中: x 为自变量, a 为控制函数开始增加的点, b 为控制函数达到最大值1的点, c 为控制函数开始减少的点。

2.3 专家规则库的映射

通过部分专家经验建立的25条专家规则如表1所

示,将其映射到规则矩阵(RuleMatrix),其中每一行代表一条规则,前2列分别是车辆定位点与轨迹点的横向偏差 D 和车辆行驶方向与轨迹点切线方向的方向偏差 H (夹角 φ)的数值映射,第3列是车辆转向盘转角 S 的数值映射,第4列是该规则的权重,是前件 D 和 H 的逻辑运算关系“与”逻辑。采用重心法(Centroid)进行解模糊,最终得到目标转角输出,将该解作为后续强化学习的一个初始解和收敛条件。

表1 专家规则库与规则矩阵的建立

序号	横向偏差	方向偏差	目标转角	规则矩阵
1	LDFar	LHFar	RSBig	$\begin{bmatrix} 1 & 1 & 5 & 1 & 1 \\ 1 & 2 & 5 & 1 & 1 \\ 1 & 3 & 4 & 1 & 1 \\ 1 & 4 & 3 & 1 & 1 \\ 1 & 5 & 3 & 1 & 1 \\ 2 & 1 & 5 & 1 & 1 \\ 2 & 2 & 4 & 1 & 1 \\ 2 & 3 & 3 & 1 & 1 \\ 2 & 4 & 3 & 1 & 1 \\ 2 & 5 & 2 & 1 & 1 \\ 3 & 1 & 4 & 1 & 1 \\ 3 & 2 & 4 & 1 & 1 \\ 3 & 3 & 3 & 1 & 1 \\ 3 & 4 & 2 & 1 & 1 \\ 3 & 4 & 2 & 1 & 1 \\ 3 & 5 & 2 & 1 & 1 \\ 4 & 1 & 4 & 1 & 1 \\ 4 & 2 & 3 & 1 & 1 \\ 4 & 3 & 2 & 1 & 1 \\ 4 & 4 & 2 & 1 & 1 \\ 4 & 5 & 1 & 1 & 1 \\ 5 & 1 & 4 & 1 & 1 \\ 5 & 2 & 3 & 1 & 1 \\ 5 & 3 & 2 & 1 & 1 \\ 5 & 3 & 2 & 1 & 1 \\ 5 & 4 & 1 & 1 & 1 \\ 5 & 5 & 1 & 1 & 1 \end{bmatrix}$
2	LDFar	LHNear	RSBig	
3	LDFar	RHZero	RSSmall	
4	LDFar	RHNear	SZero	
5	LDFar	RHFar	SZero	
6	LDNear	LHFar	RSBig	
7	LDNear	LHNear	RSSmall	
8	LDNear	RHZero	SZero	
9	LDNear	RHNear	SZero	
10	LDNear	RHFar	LSSmall	
11	RDZero	LHFar	RSSmall	
12	RDZero	LHNear	RSSmall	
13	RDZero	RHZero	SZero	
14	RDZero	RHNear	LSSmall	
15	RDZero	RHFar	LSSmall	
16	RDNear	LHFar	RSSmall	
17	RDNear	LHNear	SZero	
18	RDNear	RHZero	LSSmall	
19	RDNear	RHNear	LSSmall	
20	RDNear	RHFar	LSBig	
21	RDFar	LHFar	RSSmall	
22	RDFar	LHNear	SZero	
23	RDFar	RHZero	LSSmall	
24	RDFar	RHNear	LSBig	
25	RDFar	RHFar	LSBig	

3 无模型强化学习器设计

3.1 强化学习因素

根据智能体的状态、行为与回报等因素设计强化学习过程,在如图1所示的车辆横向控制模型中,将车辆视为智能体,在一段已知的轨迹中要求其以最小的横向偏差和方向偏差进行寻迹行驶。以已知轨迹点的状态集合(坐标点和点的方向)作为状态空间:

$$S_n = \{ \{p(x_0, y_0), h(x_0, y_0)\}, \dots, \{p(x_n, y_n), h(x_n, y_n)\} \} \quad (2)$$

式中: $S_n(i=0,1,2,\dots,n)$ 为目标轨迹点的状态集合, $p(x_p, y_p)$ 为轨迹点坐标, $h(x_p, y_p)$ 为轨迹点的方向, n 为轨迹点数量。

模糊推理器的前、后件输出结果映射为:

$$FS_n(h, D) \rightarrow FA_n(s) \quad (3)$$

式中: $FS_n(h, D)$ 为输入前件, $FA_n(s)$ 为输出后件, h 为方向偏差, s 为状态集合子集。

将模糊推理作为奖励函数(γ 因子)设为强化学习的随机因素,贪婪策略作为 θ 因子设为强化学习的随机因素,车辆的转向盘转角集合作为行为空间:

$$A_n = \left\{ \{RA_n(S_n)\}_\theta + \{FA_n(s)\}_\gamma \right\} \quad (4)$$

式中: $RA_n(S_n)$ 为贪婪法则因子下的行为空间, $FA_n(s)$ 为模糊推理器的输出行为空间, $\gamma=1-\theta$ 为强化学习的随机因子。

根据车辆横向偏差和方向偏差设定智能体车辆在当前轨迹状态下的回报。车辆在不断试错学习的迭代过程中,最终利用贪婪法实现在误差范围内的寻迹行驶。

3.2 强化学习方法选择

对于自动驾驶车辆,横向控制系统设计时需要考虑如何在尽量少的学习环节中实现最佳横向控制,在同等或者更优的强化学习效果条件下,学习的次数越少,强化学习的收敛越快,在整个操作层面上就越接近实现最优控制。在一些强化学习的方法中:蒙特卡洛(Monre Carlo, MC)方法在完整的轨迹上进行学习,但是学习效率低、收敛速度慢;时序差分(Temporal Difference, TD)法利用样本值估计值函数,属于有方差的学习方法;动态规划(Dynamic Programming, DP)方法无偏差,也无样本估计的方差。如需在建立的模型上针对状态转移概率、回报以及折扣因子进行适当的模型解析,通过与理想轨迹和模糊控制初始解相结合的方式求解,在基于无模型强化学习的方法中,Q学习(Q-learning)较为理想,将其问题归纳到马尔可夫决策过程(Markov Decision Process, MDP),利用贝尔曼方程实现模型的求解^[18-19]。

3.3 自动驾驶车辆横向控制的马尔可夫决策过程

在基于专家经验的模糊推理器上,将车辆控制模型抽象化,根据横向偏差和方向偏差划分车辆状态,形成驾驶环境(状态空间) S_n 。根据车辆的转角输出对其进行划分,形成智能体(行为空间) A_n 。在Q学习中,对于四元组的MDP, $M=\langle S, A, R, \gamma \rangle$,其中 M 代表一个马尔可夫决策, S 为状态空间, A 为行动空间, R 为即时回报, γ 为强化学习的随机因子,即回报因子。在回报函数确定的情况下,利用贝尔曼方程对其建立的MDP进行求解,经策略迭代,可得到所有状态下的最优策略和最优动作。

3.3.1 回报函数设计

根据车辆偏离已知轨迹预设点的横向偏差和方向偏差,设计回报函数为:

$$R = \frac{1}{f_{ac1}|\Delta d| + f_{ac2}|\Delta h|} \quad (5)$$

式中: f_{ac1} 为距离偏差影响因子, f_{ac2} 为方向偏差影响因子, Δd 为即时横向偏差, Δh 为即时方向偏差。

3.3.2 最优策略求解

在状态空间、行为空间和回报函数确定后,用行为回报最大的值更新原有的值。值迭代算法为:

$$Q(s',a') = Q(s',a') + \alpha \{r' + \beta q'_{\max} - Q(s,a)\} \quad (6)$$

式中: $Q(s',a')$ 为下一状态的 Q 值, q'_{\max} 为下一状态下所有行为 Q 值的最大值, $Q(s,a)$ 为当前状态的 Q 值, α 、 β 为迭代因子, r' 为下一状态的回报。

以基于部分专家规则库的模糊推理器对应的模糊解作为强化学习的初始解,进行针对该模糊推理器的强化学习,在此基础上进行数值迭代,找到每个状态对应的最优解。算法设计流程如图3所示。

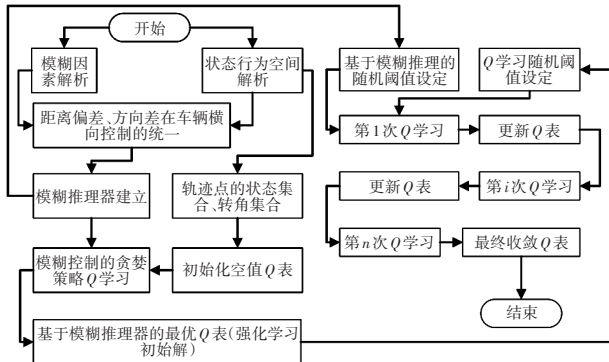


图3 算法设计流程

3.4 模糊推理器作为智能体的模仿学习

在模糊推理器及强化学习状态、行为、回报因素的设计中,模糊推理器实现了横向偏差和方向偏差到转向盘转角的映射。Q-learning以横向偏差和方向偏差作为 Q 值迭代的回报因素,在某个特定的轨迹状态中,模糊推理器输出的行为与强化学习在随机或者贪婪迭代过程中选择的行为空间子集一致,因此可以将模糊推理器作为一个单独的智能体加入到该Q-learning中,通过有限次迭代可以得出基于初步设计的模糊推理器的Q-learning结果。该结果基于模糊推理器的先验经验(部分专家经验库),将模糊推理器的控制结果作为强化学习的初始解,实际上是基于部分专家经验库指导的Q-learning模仿学习过程,后续强化学习基于此结果进行,以大幅减少学习所需的迭代次数。

3.5 模糊推理器指导车辆行为选择

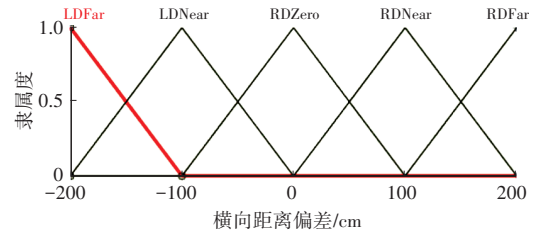
在强化学习过程中无指导地选择随机动作无疑会增加车辆学习的负担,在车辆对已知轨迹寻迹的探索中,很多错误的动作在无指导的情况下都会被选择并进行试错学习。模糊推理器在车辆学习的过程中为车辆提供先验经验,并根据车辆的实时学习状态给予车辆随机选取动作进行试错学习的机会。通过模糊推理器的行为指导,将再次减少学习迭代次数。

4 仿真测试与结果分析

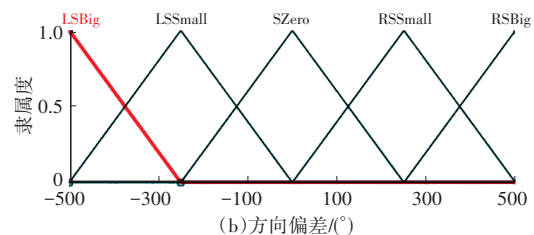
本文在MATLAB^[20]中搭建并仿实现模糊推理器,在Carla^[21]仿真环境中实现模糊推理器的强化学习。

4.1 模糊推理器搭建与仿真

通过MATLAB实现第2节中所设计的模糊推理器,以距离偏差和方向偏差作为输入前件,如图4所示,以转向盘转角作为输出后件,如图5所示。通过部分专家经验以及解模糊函数实现前、后件映射,结果如图6所示。



(a)距离偏差



(b)方向偏差

图4 输入前件

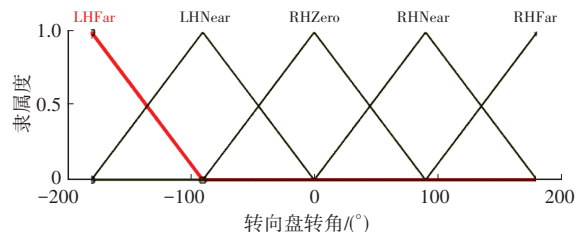


图5 转向盘转角输出后件

图6仿真结果表明,距离偏差为-33.5 cm、方向偏差为-117°时,转向盘输出转角为263°。前件输入范围内的距离偏差和方向偏差均可映射到转向盘转角输出。通过Python编程可实现第2节中的模糊控制器。

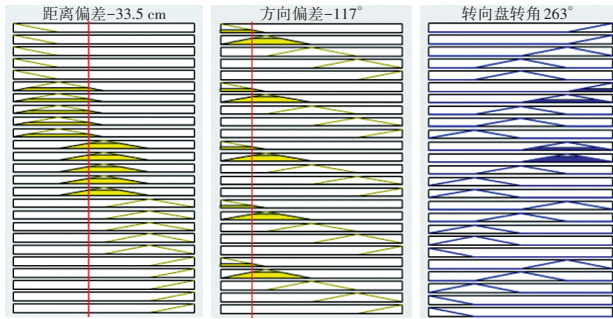


图6 前、后件映射结果

4.2 基于模糊推理器和强化学习的车辆横向控制仿真

Carla基于虚幻引擎(Unreal Engine)进行模拟,并使用OpenDRIVE标准定义道路和城市环境。通过使用Python和C++处理的应用程序编程接口(Application Programming Interface, API)授予对模拟的控制权。仿真环境界面如图7所示。该仿真环境拥有强大的物理引擎,可模拟车辆的机械特性,真实地模拟现实世界。



图7 Carla自动驾驶仿真环境

利用Carla的内置地图,在虚幻引擎中运行并搭建如图8所示的自动驾驶汽车运行环境。在地图环境中选择环岛场景构成的S弯道,道路是由单向双车道组成,环岛场景被分割为4个部分,自动驾驶车辆从第1入口进入环岛并从第3出口驶出环岛,在地图中显示出利用A*算法规划获得的车辆行驶的路径点,并通过Python API构建具备物理特性的自动驾驶车辆。

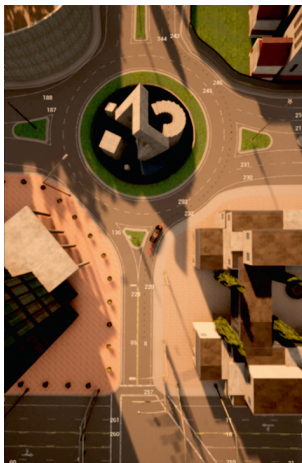
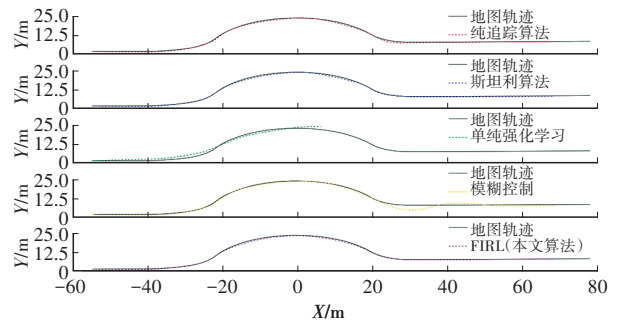


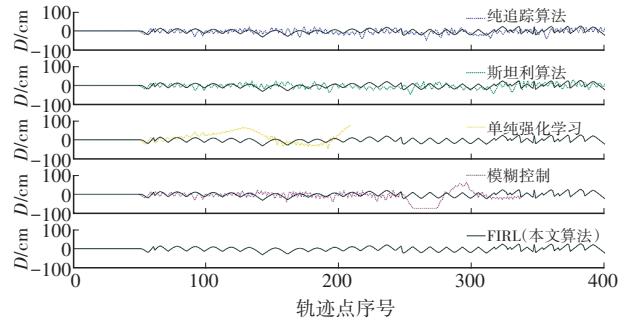
图8 车辆运行环境

该场景下的目标轨迹具备入弯道和出弯道的特点,采用轨迹跟踪时真实轨迹与目标轨迹的重合度来考查模糊推理器和模糊推理的无模型自适应强化学习在自动驾驶车辆横向控制上的性能优劣。

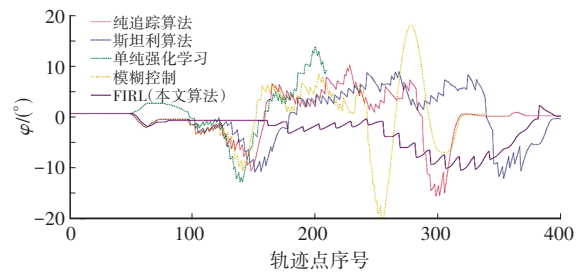
为了与本文提出的FIRL进行对比,利用纯跟踪算法^[22]、斯坦利算法^[5]、第2节中构建的模糊推理器以及第3节中构建的Q-learning(经过500次迭代)进行控制效果对比分析。其中,在Carla仿真环境中采用Python编程可实现单纯的Q-learning对车辆的控制。各方法在全路径的轨迹跟踪中获取轨迹跟踪重合度、横向偏差、方向偏差如图9所示,其中X、Y分别为地图横、纵坐标。



(a) 轨迹跟踪重合度



(b) 横向偏差



(c) 方向偏差

图9 轨迹跟踪误差

由图9a可知:单纯的强化学习方法在500次迭代学习后,车辆并不能完成全轨迹跟踪,仍然需要更多的学习次数进行迭代;采用FIRL时在500次迭代后已经能够很好地进行全路径的轨迹跟踪,说明FIRL能够大幅缩短学习过程。

从图9b中可以看出:斯坦利和纯跟踪方法均能实现较好的横向控制,进而实现车辆的轨迹跟踪,并且斯

坦利算法性能更为优越;FIRL相比较这2种控制方法性能均有所提升,轨迹跟踪的重合度更高。在该仿真中,FIRL仅通过500次的全路径迭代学习即可超越斯坦利算法实现车辆控制。

由图9c可知,模糊控制算法实现了全路径轨迹跟踪的收敛,但是在控制过程中,特别是在进入弯道和出弯道时方向偏差出现了大幅振荡。

将模糊控制、纯跟踪、斯坦利以及FIRL进行500次迭代学习(FIRL-500),轨迹跟踪结果如表2、表3所示。

表2 轨迹跟踪距离误差

算法	最大值/cm	最小值/cm	平均值/cm	方差/cm ²
模糊控制	75.00	0.01	15.20	359.56
纯跟踪	49.79	0.02	10.62	70.62
斯坦利	48.32	0.01	9.38	68.56
FIRL-500	47.67	0.01	9.36	67.88

表3 轨迹跟踪方向误差

算法	最大值/(°)	最小值/(°)	平均值/(°)	方差/(°) ²
模糊控制	22.60	0.08	4.14	21.11
纯跟踪	15.66	0.00	3.30	11.24
斯坦利	15.92	0.01	3.01	12.09
FIRL-500	14.22	0.03	2.93	9.79

由表2、表3可知,相较于模糊控制、纯跟踪、斯坦利算法,FIRL-500使车辆在整个路径的轨迹跟踪过程中控制轨迹偏离的效果和运行的平滑度得到了提升。

5 实车验证

5.1 自动驾驶车辆平台

强化学习结果实车迁移平台使用如图10所示的北京联合大学“京龙”号无人驾驶车,该平台基于长城VV6的基础车辆底盘进行线控改制,使其具备线控制动、线控转向、线控油门、线控挡位的功能,并增加组合导航、32线激光雷达、毫米波雷达、摄像头等传感器,通过CAN总线和用户数据报协议(User Datagram Protocol, UDP)的通信方式实现自动驾驶系统与各线控执行层的通信控制与反馈。

5.2 环岛场景的控制方法迁移

车辆加装组合导航设备对某环岛区域地图进行高精度采集,并绘制可行驶道路的轨迹(道路中心线或车道中心线)。环岛场景对自动驾驶横向控制的要求很高,该场景涉及入弯道、出弯道与弯道行驶等具体场景,如图11所示。车辆从第4出口进入环岛,从第2出口驶出环岛。在仿真场景中构建相同的环岛场景,如图12所示。



图10 自动驾驶车辆平台

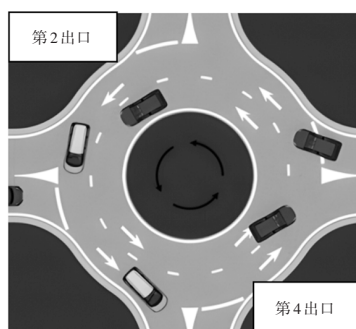


图11 环岛场景



图12 仿真环境中的环岛场景

5.3 仿真车辆与实际车辆关键参数标定

Carla仿真环境拥有精确的车辆体积和真实的核心物理特性,基于虚幻引擎4(Unreal Engine 4)的框架,开启物理特性后,能够很好地模拟车辆在现实环境中的驾驶特性,如车轮摩擦、悬架、质心和车辆起动速度等。本文的控制方法首先在模拟环境中学习得到最优控制结果,在仿真学习过程中需要对仿真场景和真实车辆进行基础设定,标定的关键参数如表4所示。

表4 仿真车辆与实车标定的关键因素

项目	参数描述
车速特性	保持车辆纵向车速在相同的范围内
加速特性	保持车辆加速度在相同的范围内
减速特性	保持车辆减速度在相同的范围内
转向特性	保持车辆转向盘转角的转弯半径一致,最终映射到[-1,1]范围内

车辆一致性标定步骤如下:

a. 最大速度参数设定。将车辆最高速度参数设定为真实场景下的环岛最高限速。实际车辆在环岛场景下,其纵向控制由PID算法对油门的驱动加上速度反馈闭环控制实现,能够实现速度控制精度在 $\pm 1\%$ 范围内。在车辆基础物理特性一致的情况下,仿真车辆与实际车辆具有相似的速度特性。

b. 加、减速特性。仿真车辆和实际车辆均可通过发送加、减速指令对车辆实施加速和减速控制,可忽略整车质量、道路摩擦力的影响,在保证加、减速度范围一致的情况下确保车辆纵向控制的一致性。

c. 转向特性。FIRL对仿真车辆和实际车辆的一致性要求较高,直接决定了横向控制的一致性。在仿真环境中,车辆的转向不是通过控制转向盘转角实现的,而是直接将车轮转角的范围映射到 $[-1,1]$ 范围内。对实际车辆的转向控制通过控制转向盘转角实现。然而,实际车轮转角与转向盘转角并非线性关系,如图13所示。

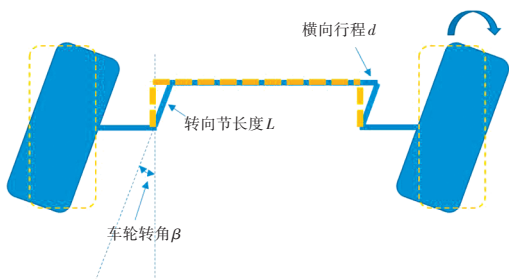


图13 车轮转角与转向盘转角的关系

由图13可知,设转向盘转角为 Φ 时,车轮转角为 β ,前轴横向行程为 d ,转向节长度为 L ,行程 d 与车轮转角的关系为:

$$\beta = \arcsin(d/L) \quad (7)$$

前轴横向行程 l 与转向盘转角 Φ 的关系为:

$$l = \Phi \cdot \Phi_{\max} / L_{\max} \quad (8)$$

式中: Φ_{\max} 为转向盘最大转角, L_{\max} 为车辆在执行最大转向盘转角时的前轴最大横向行程。

可得转向盘转角 Φ 与车轮转角 β 的关系为:

$$\beta = \arcsin\left(\frac{L_{\max} \cdot d}{\Phi_{\max} \cdot L}\right) \quad (9)$$

通过与仿真车辆的转向映射进行对比,最终将转向盘转角映射到 $[-1,1]$ 的范围内,与仿真车辆保持一致。

按照前文在仿真环境中搭建FIRL的方法对车辆进行训练学习,并将最终的学习数据应用于实车横向控制,实现环岛场景下自动驾驶车辆的横向控制,车辆平顺地通过了环岛,表明算法具有良好的控制效果,实车运行可视化界面如图14所示。

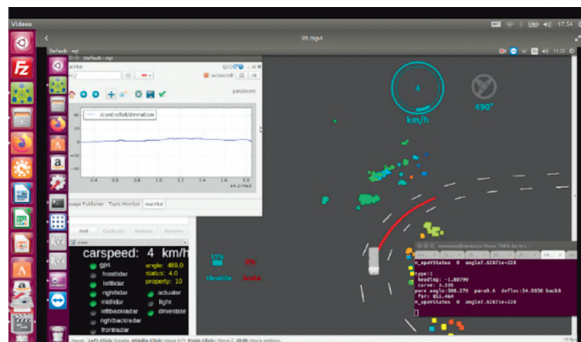


图14 实车运行可视化界面

6 结束语

本文提出了基于模糊推理和模仿强化学习的横向控制方法(FIRL),以模糊推理器作为模仿强化学习的初始解和过程“指导”进行自动驾驶汽车横向控制,并进行了仿真测试验证以及实车试验。通过对模糊控制、纯跟踪、斯坦利、无模仿学习的Q-learning方法以及进行500次迭代学习的FIRL在自动驾驶车辆横向控制方面的控制效果进行对比,得出如下结论:

a. 在无需准确数学模型的前提下,仅通过对系统的输入、输出数据进行系统性分析即可实现车辆横向控制。

b. 通过模仿强化学习的方式对该模型进行优化,基于模糊推理器的自动驾驶汽车横向控制模型可实现更优的横向控制性能。

c. 利用模糊推理器的结果作为强化学习初始解,可大幅缩短学习迭代的过程,同时提高自动驾驶车辆轨迹跟踪的能力。

d. 不依赖理想数学模型、不需要设计复杂的模糊推理器和使用复杂的强化学习方法可实现自动驾驶汽车横向控制。

e. FIRL可以在真实环境中实现自动驾驶汽车的轨迹跟踪,证明了其在真实驾驶场景中的有效性。

参考文献

- [1] 李梓欣,李军. 自动驾驶汽车路径规划算法研究[J]. 汽车工程师, 2021(8): 11-14.
- [2] 狄桓宇,张亚辉,王博,等. 自动驾驶横向控制模型及方法研究综述[J]. 重庆理工大学学报(自然科学), 2021, 35(7): 71-81.
- [3] DI H Y, ZHANG Y H, WANG B, et al. Review on the Method and Model of Intelligent Vehicle Lateral Control[J]. Journal of Chongqing University of Technology (Natural Science), 2021, 35(7): 71-81.

- [3] YANAKIEV D, KANELAKOPOULOS I. Longitudinal Control of Automated CHVs with Significant Actuator Delays [J]. IEEE Transactions on Vehicular Technology, 2001, 50(5): 1289-1297.
- [4] XIA Y Q, PU F, LI S F, et al. Lateral Path Tracking Control of Autonomous Land Vehicle Based on ADRC and Differential Flatness[J]. IEEE Transactions on Industrial Electronics, 2016, 63(5): 3091-3099.
- [5] 王鑫, 凌铭, 饶启鹏, 等. 基于改进Stanley算法的无人车路径跟踪融合算法研究[J]. 汽车技术, 2022(7): 25-31.
WANG X, LING M, RAO Q P, et al. Research on Fusion Algorithm of Unmanned Vehicle Path Tracking Based on Improved Stanley Algorithm[J]. Automobile Technology, 2022(7): 25-31.
- [6] 田涛涛, 侯忠生, 刘世达, 等. 基于无模型自适应控制的无人驾驶汽车横向控制方法[J]. 自动化学报, 2017, 43(11): 1931-1940.
TIAN T T, HOU Z S, LIU S D, et al. Model-Free Adaptive Control Based Lateral Control of Self-Driving Car[J]. Acta Automatica Sinica, 2017, 43(11): 1931-1940.
- [7] 潘峰. 基于驾驶员行为特性的无人驾驶汽车控制方法研究[D]. 北京: 北京化工大学, 2021.
PAN F. Research on the Control Method of Autonomous Vehicle Based on the Characteristics of Human Drivers[D]. Beijing: Beijing University of Chemical Technology, 2021.
- [8] 罗瑞. 基于强化学习的无人驾驶车辆横向控制[D]. 北京: 华北电力大学, 2020.
LUO R. Horizontal Control of Driverless Vehicles Based on Reinforcement Learning[D]. Beijing: North China Electric Power University, 2020.
- [9] 邢斯琦, 冯凯, 薛玲玲. 基于模型预测的自动驾驶横向控制算法研究[J]. 汽车实用技术, 2021, 46(15): 45-47+50.
XING S Q, FENG K, XUE L L. Research on Lateral Control Algorithm of Automatic Driving Based on Model Prediction[J]. Automotive Practical Technology, 2021, 46(15): 45-47+50.
- [10] 武金龙, 谢凡, 尤丽刚, 等. 自动驾驶搅拌车横向控制设计及试验研究[J]. 智能计算机与应用, 2021, 11(9): 184-188.
WU J L, XIE F, YOU L G, et al. Horizontal Control Design and Experimental Research of Automatic Driving Mixer[J]. Intelligent Computer and Application, 2021, 11(9): 184-188.
- [11] 张效宇, 程昆朋, 赵祥磊. 基于MPC的智能驾驶汽车横向运动控制研究[C]// 2018中国汽车工程学会年会论文集. 上海: 机械工业出版社, 2018: 85-90.
ZHANG X Y, CHENG K M, ZHAO X L. Research on Lateral Motion Control of Intelligent Driving Vehicles Based on MPC[C]// Proceedings of the 2018 Annual Meeting of the Society of Automotive Engineers of China. Shanghai: China Machine Press, 2018: 85-90.
- [12] 曹艳玲, 张琦. 自适应神经模糊推理的四轮转向车辆转向控制研究[J]. 机械设计与制造, 2021(3): 224-228+233.
CAO Y L, ZHENG Q. Research on Steering Control of Four-Wheel Steering Vehicle Based on Adaptive Neuro Fuzzy Inference[J]. Mechanical Design and Manufacturing, 2021(3): 224-228+233.
- [13] VALLON C, ERCAN Z, CARVALHO A, et al. A Machine Learning Approach for Personalized Autonomous Lane Change Initiation and Control[C]// 2017 IEEE Intelligent Vehicles Symposium (IV). USA: IEEE, 2017: 1590-1595.
- [14] SUTTON R S. Learning to Predict by the Methods of Temporal Differences[J]. Machine Learning, 1988, 3(1): 9-44.
- [15] 李远哲, 胡纪滨. 强化学习在无人车领域的应用与展望[J]. 信息与控制, 2022, 51(2): 129-141.
LI Y Z, HU J B. Application and Prospect of Reinforcement Learning in the Field of Unmanned Vehicles[J]. Information and Control, 2022, 51(2): 129-141.
- [16] 邹伟, 高玲, 刘昱杓. 强化学习[M]. 北京: 清华大学出版社, 2020.
ZOU W, GE L, LIU Y S. Reinforcement Learning[M]. Beijing: Tsinghua University Press, 2020.
- [17] SUTTON R S, BARTO A G, et al. Reinforcement Learning: An Introduction[M]. Cambridge: MIT Press, 2018.
- [18] GAO Z, SUN T, XIAO H, et al. Decision-Making Method for Vehicle Longitudinal Automatic Driving Based on Reinforcement Q-Learning[J]. International Journal of Advanced Robotic Systems, 2019, 16(3): 141-172.
- [19] WATKINS C J C H, DAYAN P. Q-Learning[J]. Machine Learning, 1992, 8(3/4): 279-292.
- [20] 张森, 张正亮. MATLAB仿真技术与实例应用教程[M]. 北京: 机械工业出版社, 2004.
ZHANG S, ZHANG Z L. MATLAB Simulation Technology and Example Application Tutorial[M]. Beijing: China Machine Press, 2004.
- [21] DOSOVITSKIY A, ROS G, CODEVILLA F, et al. CARLA: An Open Urban Driving Simulator[EB/OL]. (2017-11-10) [2023-01-30]. <https://arxiv.org/abs/1711.03938>.
- [22] 段建民, 夏天, 宋志雪. 基于联合仿真的智能车路径跟踪控制研究[J]. 计算机仿真, 2017, 34(6): 129-134+151.
DUAN J M, XIA T, SONG Z X. Research on Intelligent Vehicle Path Tracking Control Based on Joint Simulation [J]. Computer Simulation, 2017, 34(6): 129-134+151.

(责任编辑 斛 畔)

修改稿收到日期为2023年1月30日。