基于决策-规划迭代框架的智驾车换道行为建模

肖 雪[1],李克平[1],彭 博[2],昌满玮[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. 吉林大学 交通学院，长春 130012）

摘 要：行为决策和轨迹规划是智驾车完成驾驶任务的关键环节，着眼于二者之间的参数传递关系和新型混合交通流的动态特征，建立了智驾车换道模型。采用主从博弈思想描述智驾车换道决策过程，采用负指数函数量化博弈代价函数。将博弈决策结果作为轨迹规划模型输入之一，采用多项式分别描述车辆的横向及纵向运动，利用模拟退火算法寻找最优轨迹。设置了不同驾驶特性的环境车辆，并对决策规划过程进行了仿真验证，结果表明，决策规划模型能够快速生成安全、舒适的可行轨迹。

关键词：交通运输系统工程；博弈论；轨迹规划；模拟退火

中图分类号：U491.2 **文献标志码**：A **文章编号**：1671-5497(2023)03-0746-12

DOI：10.13229/j.cnki.jdxbgxb20220560

Integrated lane-changing model of decision making and motion planning for autonomous vehicles

XIAO Xue[1], LI Ke-ping[1], PENG Bo[2], CHANG Man-wei[1]

(1. *Key Laboratory of Road and Traffic Engineering*, *Ministry of Education*, *Tongji University*, *Shanghai* 201804, *China*;
2. *College of Transportation*, *Jilin University*, *Changchun* 130012, *China*)

Abstract: Decision-making and motion planning are the crucial modules of autonomous vehicles. Focusing on the driving parameter relationship between decision making module and motion planning module, and the dynamic characteristics of new hybrid traffic flow, lane-changing model of autonomous vehicles was established. The decision-making process of lane-changing of autonomous vehicles was described by Stackelberg game theory, and the game cost function is quantified by negative exponential function. The decision results of the autonomous vehicle were taken as one of the inputs of the motion planning module. Polynomial was used to describe the lateral and longitudinal trajectories of the vehicle, and simulated annealing algorithm was used to find the optimal trajectory. The simulation results show that the proposed decision programming model can quickly generate safe, comfortable and feasible trajectory.

Key words: engineering of communications and transportation system; game theory; motion planning; simulated annealing

---

收稿日期：2022-05-12.

基金项目：国家重点研发计划项目(2019YFB1600200).

作者简介：肖雪(1995-)，女，博士研究生. 研究方向：微观交通行为仿真. E-mail：2010770@tongji.edu.cn

0　引　言

自动驾驶是当前工业界和学术界研究的热点问题,被认为能够大幅提升交通系统的运行效率,其规划流程依次为环境感知、行为决策、轨迹规划和跟踪控制。其中,行为决策和轨迹规划是决定智驾车运动状态至关重要的两个部分,二者间呈现强耦合关系[1,2]。决策模块基于感知系统的输出数据对当前驾驶环境进行分析,决定智驾车的行驶策略,如换道、跟驰等决策;规划模块则在决策模块的指导下,负责向控制模块输出具体的行驶轨迹参数。轨迹规划本质是在车辆运动学约束和决策指令约束下的求极值问题,其关键决策参数的规划空间为整个道路可行空间和运动学参数空间,如车辆横向位移、纵向行驶速度、加速度等[3]。在决策命令下进行轨迹规划,相当于进一步界定了轨迹规划空间,新的规划空间变为运动学可行参数集与行为决策结果可行参数集合的交集。决策范围过小,能够极大减小轨迹规划的规划空间,但可能产生交集过小或是空集的情况,若范围过大,则难以体现行为决策和轨迹规划的高效衔接过程和优势。故如何合理设计运动参数变量的约束和传递,实现决策模块与轨迹模块之间的协同控制是自动驾驶系统开发中的一个关键问题。

近年来,学者们在行为决策和轨迹规划方面展开了大量研究,大多研究专注于对单一模块进行深入探索,鲜少研究二者联合规划模型。在行为决策方面,主要包括概率论、机器学习和博弈论。基于概率论的决策方法主要通过贝叶斯推理识别交通个体的驾驶意图,通过基于高斯回归和部分观测马尔可夫决策过程估计不同环境下的车辆运动[4,5]。Gindele 等[6]和 Schulz 等[7]将动态贝叶斯网络用于完整的驾驶情境评估,不同个体对于场景的状态观察、轨迹规划、交互意图、不同交通参与者的相互依赖性和动作的因果关系体现在人工构造的贝叶斯网络中。Althoff 等[8]采用高斯回归过程估计具有随机行为的车辆状态,并通过蒙特卡洛仿真实现车辆运动决策,但该方法在实现由各种可能驾驶意图引起的多模态车辆运动特征方面具有局限性。Kamrani 等[9]采用马尔可夫过程描述车辆的决策过程,采用强化学习估计不同驾驶特性的马尔可夫转移概率。

基于机器学习决策方法的主要思想为将本车和周围车辆的行驶状态、历史轨迹、道路几何等交通环境数据作为输入,输出未来时刻的驾驶决策。Vallon 等[10]采用支持向量机模拟车辆的换道决策过程,模型输入为本车纵向行驶速度及与周围环境车辆的纵向位置差,输出为是否换道的二项决策。Lenz 等[11]采用基于多层全连接神经网络学习的二维高斯混合模型预测本车纵向加速度和速度。Lee 等[3]基于非参数化的预测结构,采用循环神经网络和卷积神经网络生成具有多模态特性的备选轨迹,并采用多层循环神经网络选取最终行驶轨迹。基于机器学习[12]和概率论[13]的决策方法虽然能够初步体现车辆的多模态特性,但没有明确分析交通个体之间的相互作用关系,且基于数据驱动的模型训练过程依赖大量实际数据。

基于博弈论的决策方法则明确讨论了在交互过程中交通个体间的相互影响,其决策过程也可显式地体现。在路段场景下,Stackelberg 博弈模型被认为能够有效模拟具有先后位置关系的车辆博弈行为。Yu 等[14]从驾驶动机出发,采用 Stackelberg 非合作博弈描述路段场景下智驾车的换道过程,并基于跟随车辆预测博弈收益小于其真实收益的原则预测跟随车辆驾驶激进性,赋予智驾车类人驾驶逻辑,理论上保证了车辆决策结果的合理性。胡益恺等[15]采用施塔克尔贝格(Stackelberg)合作博弈讨论了智驾车在汇流场景中的决策行为,利用贝塞尔曲线生成多条备选轨迹,选择将最优博弈收益的轨迹作为车辆最终行驶轨迹。Kang 等[16]将存在交互行为的所有决策个体视为非合作博弈中的参与者,构建了重复博弈机制寻找最优驾驶策略组合。Hang 等[17]构建了路段场景下的 Stackelberg 非合作博弈模型,采用重力势场函数参数化道路环境并规划车辆行驶轨迹。

自动驾驶轨迹规划方法主要包括图解法、状态空间法、人工势场法、采样法[18]等。其中,采样法[19,20]由于其简单有效、适应性强的特点被广泛采用。其核心思想为通过预设采样点的车辆运动参数约束车辆轨迹,通过路径-速度或横向轨迹-纵向轨迹的时空解耦方式生成多条备选轨迹,通过穷举的方式寻找最优轨迹。然而,在真实场景中为满足规划轨迹的最优性,采样空间通常维度较高,使得算法复杂性极大增加,难以满足实时性计算需求。如何在采样空间中快速找到最优解是该方向亟待解决的关键问题。

一种解决思路是将采样空间进行降维和离散

化处理。对于离散化采样空间而言,离散空间仅能够得到次优解,根本原因在于固定的离散点限制了车辆横向偏移、曲率等参数。Xu等[21]通过解除已找到最优轨迹的运动学参数限制,对最优轨迹进行微调优化,但优化后的最优轨迹安全性无法得到保证。Ziegler等[22]对交通感知数据进行了特征化处理,设定道路多边形安全空间约束以减少采样空间。另一种解决思路是基于已经生成的轨迹集合,采用快速搜索算法找到一条最优轨迹。Jiang等[23]将碰撞代价从总代价函数中剥离出来,按总代价从高到低对轨迹进行排序,从最优轨迹开始顺序检查是否满足碰撞要求。该方法减少了绝大多数备选轨迹碰撞代价的计算,在采样空间较为密集时能够显著提高效率。

综上所述,现有研究多专注于对单一模块的优化建模。已有的博弈模型研究中都是假设基于预设的车辆轨迹,而已有轨迹规划模型大多是基于指定的驾驶决策,较少考虑二者之间的有机联系和相互依赖性关系。智驾车行为决策和轨迹规划是一种高度相关的强耦合过程,其输出结果应具有高度协同性和一致性,单独对其中一个模块进行讨论难以适应真实交通环境的运行规律。故本文提出了基于决策-规划协同迭代框架的智驾车路段驾驶行为模型。先将驾驶行为解构为高度关联的决策和规划行为,决策行为确定了驾驶意图,规划行为确定了未来一段时间内连续的运动行为,进而构建了决策-规划迭代框架。通过车辆关键运行参数的流动和反馈,表现了决策和规划行为的有机联系,即决策主导规划,规划辅助决策。具体而言,在决策模型中,采用Stackelberg博弈模型构建基于双层优化的最优策略选择机制;以安全性、舒适性、高效性为原则设计博弈收益函数,对所有运动学参数进行指数化处理,保证各代价分量量纲的一致性、关键参数的敏感性以及由不同驾驶动机引起的多模态特性。将博弈模型决策结果作为轨迹规划模块生成候选轨迹的约束条件,采用模拟退火算法寻找最优轨迹,将实际轨迹行驶信息作为博弈模型输入,由此形成连续规划、滚动优化的仿真过程。

1 车辆博弈模型

1.1 主从博弈模型

在路段场景下,车辆的驾驶行为可分为跟驰或换道,速度行为体现为是否加速。考虑到轨迹规划所需的关键决策参数,选择车辆纵向加速度和是否换道作为博弈决策关键参数,博弈双方的策略如表1所示。博弈决策过程为:跟随车辆假定处于前方的领导车辆先确定驾驶策略,再根据领导车辆策略进行最优化调整,如式(1)所示;领导车辆的最终决策也会考虑跟随车辆的最优决策,如式(2)所示,约束条件如式(3)所示。

$$\Phi(a_{\mathrm{av}}, L_{\mathrm{av}}) = \left\{\zeta \middle| U_{\mathrm{hv}}(a_{\mathrm{av}}, L_{\mathrm{av}}, \zeta) \geqslant U_{\mathrm{hv}}(a_{\mathrm{av}}, L_{\mathrm{av}}, a_{\mathrm{hv}})\right\} \tag{1}$$

$$(a_{\mathrm{av}}^*, L_{\mathrm{av}}^*) = \underset{(a_{\mathrm{av}}, L_{\mathrm{av}})}{\mathrm{argmix}}\left[\max_{a_{\mathrm{av}} \in \Phi(a_{\mathrm{av}}, L_{\mathrm{av}})} U_{\mathrm{av}}(a_{\mathrm{av}}, L_{\mathrm{av}}, a_{\mathrm{hv}})\right] \tag{2}$$

$$\mathrm{s.t.}\begin{cases} a \in [a_{\min}, a_{\max}] \\ \zeta \in [a_{\min}, a_{\max}] \\ L_{\mathrm{av}} \in \{l-1, l, l+1\} \end{cases} \tag{3}$$

式中:下角标av表示智驾车,hv表示交互车辆;a为加速度;U为代价函数值;L为换道符号,$l-1$、l、$l+1$分别表示向右换道、保持车道、向左换道;$a_{\min}$、$a_{\max}$分别为车辆最大、最小加速度限制;ζ为使交互车辆代价函数值最大时的加速度;Φ为所有ζ组成的集合;符号"*"表示最优参数。

表1 博弈策略

Table 1 Game strategies

代价函数		跟随车辆	
		加速a	减速d
领导车辆	换道c	$(U_{\mathrm{av},c}, U_{\mathrm{hv},a})$	$(U_{\mathrm{av},c}, U_{\mathrm{hv},d})$
	跟驰f	$(U_{\mathrm{av},f}, U_{\mathrm{hv},a})$	$(U_{\mathrm{av},c}, U_{\mathrm{hv},d})$

从建模过程可知,该过程是一种双层博弈优化过程,下层为跟随车辆根据假定的领导车辆策略选取其最优策略,上层为领导车辆将跟随车辆的可能决策考虑进策略选择过程,以跟随车辆最优策略集为约束,保证了决策结果均是博弈双方在当前驾驶环境中的最优驾驶策略。该过程考虑了跟随车辆驾驶行为不确定性,更加符合真实交通场景车辆间的博弈交互逻辑。

1.2 博弈代价

博弈代价函数将显式地量化博弈参与者的驾驶目标和动机,并最终决定博弈参与者的策略选择。博弈代价函数有两种评估思想:一是以未来预测时刻代价值优于当前时刻代价值为博弈策略评估原则,即以博弈对象当前行驶环境为评估标

准，希望未来驾驶环境优于当前驾驶环境[14]；二是直接计算未来预测时刻代价函数值，取最优代价值作为最终执行的博弈策略[24]。本文采用第2种评估标准，希望博弈双方都以本车未来的最优行驶策略为目标进行决策。

考虑到环境车辆驾驶行为在未来一段时间内存在不确定性，直接根据驾驶参数预测行为存在一定误差，故将模型预测控制思想应用到代价函数中预测未来一段时间内的代价值，并增加阻尼因子，以提高代价函数的容错性和鲁棒性。以安全、舒适、高效为原则设计博弈代价函数，如下所示：

$$U=\sum_{t=1}^{T}\gamma^{t}U(k),\quad \gamma\in[0,1] \tag{4}$$

$$U(k)=\omega_1U_{\mathrm{sa}}(k)+\omega_2U_{\mathrm{co}}(k)+\omega_3U_{\mathrm{ef}}(k) \tag{5}$$

$$\omega_1+\omega_2+\omega_3=1 \tag{6}$$

式中：T为预测时间窗；γ为阻尼因子，其含义为越靠近当前时刻的代价函数分量对于总代价函数的影响越大；t为预测时间间隔；$U_{\mathrm{sa}}(k)$、$U_{\mathrm{co}}(k)$、$U_{\mathrm{ef}}(k)$分别为k时刻的安全代价、舒适代价、效率代价；ω_1、ω_2、ω_3为权重项，体现不同博弈者的驾驶特性。

式(4)表示根据每一个时刻的车辆运行参数，计算其代价函数，并乘以该时刻的阻尼因子，最后将所有时刻代价函数值相加求得使其取最小值时下一时刻的运动参数。

安全性是车辆行驶过程中考虑的最重要因素。本文从时间安全和空间两个角度，将未来预测时刻博弈车辆间的纵向距离、车头时距以及速度差作为衡量驾驶安全性的关键因素，如下所示：

$$U_{\mathrm{sa}}(k)=\left(\mathrm{e}^{-\frac{s^2}{\alpha_1}}+\mathrm{e}^{-\frac{t_{\mathrm{head}}^2}{\alpha_2}}+\mathrm{e}^{-\frac{1}{\Delta v^2\cdot\alpha_3+\varepsilon}}\right)\Big/3 \tag{7}$$

$$P_{\mathrm{p}}(k)=P_{\mathrm{p}}(k-1)+v_{\mathrm{p}}(k-1)t+\frac{a_{\mathrm{p}}(k)\cdot t^2}{2} \tag{8}$$

$$P_{\mathrm{f}}(k)=P_{\mathrm{f}}(k-1)+v_{\mathrm{f}}(k-1)t+\frac{a_{\mathrm{f}}(k)\cdot t^2}{2} \tag{9}$$

$$s=P_{\mathrm{p}}(k)-P_{\mathrm{f}}(k) \tag{10}$$

$$v_{\mathrm{p}}(k)=v_{\mathrm{p}}(k-1)+a_{\mathrm{p}}\cdot t \tag{11}$$

$$v_{\mathrm{f}}(k)=v_{\mathrm{f}}(k-1)+a_{\mathrm{f}}\cdot t \tag{12}$$

$$t_{\mathrm{head}}=\begin{cases}\dfrac{P_{\mathrm{p}}-P_{\mathrm{f}}}{v_{\mathrm{p}}(k)},\ \text{if}\ P_{\mathrm{p}}\geqslant P_{\mathrm{f}}\\[2ex] \dfrac{P_{\mathrm{f}}-P_{\mathrm{p}}}{v_{\mathrm{f}}(k)},\ \text{其他}\end{cases} \tag{13}$$

$$\Delta v=v_{\mathrm{p}}(k)-v_{\mathrm{f}}(k) \tag{14}$$

式中：下角标p、f分别表示领导车辆、跟随车辆；P为车辆横向位置；v为速度；s、Δv、t_{head}分别为k时刻博弈车辆间的位移差、速度差、车头时距；α_1、α_2、α_3为参数项；ε为极小值。

将不同指标通过负指数函数进行归一化处理，使安全项更加关注较小的车间距离、车头时距，在跟随车辆低速行驶时降低对速度差的关注度，而在高速行驶时更加关注速度差，同时设置参数项权衡不同指标之间的重要程度。

决策结果需符合舒适性要求，本文采用急动度$U_{\mathrm{co}}(k)$衡量驾驶舒适性，如下所示：

$$U_{\mathrm{co}}(k)=\frac{\exp\left(-\dfrac{1}{j_x(k)\cdot\beta_1+\varepsilon}\right)+\exp\left(-\dfrac{1}{j_y(k)\cdot\beta_2+\varepsilon}\right)}{2} \tag{15}$$

$$j_x(k)=\frac{a_x(k+t)-a_x(k)}{t} \tag{16}$$

$$j_y(k)=\frac{a_y(k+t)-a_y(k)}{t} \tag{17}$$

式中：j_x、j_y分别为博弈车辆的横向、纵向急动度；a_x、a_y分别为车辆横向、纵向加速度；β_1、β_2为参数项。

急动度是加速度在单位时间内的变化率。研究表明[25]，当急动度介于[−1,1]时，认为舒适性较好。

若仅采用安全性和舒适性作为评价指标，博弈结果将倾向于选择较小的行驶速度和加速度，导致驾驶行为过于保守。需在保证安全性和舒适性的前提下，提高驾驶高效性。本文将车辆纵向行驶速度作为衡量驾驶效率的关键因素，效率代价函数$U_{\mathrm{ef}}(k)$为：

$$U_{\mathrm{ef}}(k)=\exp\left(-\frac{1}{|v_{\max}-v(k)|\cdot\eta+\varepsilon}\right) \tag{18}$$

式中：$v_{\max}$为最大速度；η为参数项。

本节从安全性、舒适性与高效性3个方面定义了博弈代价函数，将各分量分别通过负指数函数进行处理，保证了各代价分量量纲的一致性和关键参数的敏感性。通过调整总代价函数中的权重项，体现了不同驾驶动机以及由不同驾驶动机引起的多模态特性。同时，代价函数各分量均具有具体显式的物理含义，决策结果可解释性较强。

若环境中存在多个博弈交互对象，则分别计

算智驾车与其他博弈交互对象的博弈代价函数值,将具有最小代价函数值的车辆作为当前交互对象。以非合作博弈方式,在备选策略集中进行离散化采样求解,将博弈决策结果(即自动驾驶纵向加速度和换道位置)作为轨迹规划模块的约束条件。

2 轨迹规划模型

本节将自动驾驶行驶轨迹进行横向和纵向分解以解决时空耦合问题。分别采用5次、4次多项式拟合横向、纵向轨迹;基于安全性、舒适性、高效性原则设计代价函数,采用模拟退火算法求解最优轨迹参数。

2.1 车辆备选轨迹生成

2.1.1 车辆横向轨迹

智驾车的理想行驶路径是沿着当前道路行驶,当执行变道和避障等操作时,车辆需要在当前车道中心线上进行一定的横向偏移。由于轨迹初始状态为当前状态,因此横向轨迹由终点状态和规划周期确定,将车辆横向状态关键参数定义为$\{x_0,\dot{x}_0,\ddot{x}_0,x_t,\dot{x}_t,\ddot{x}_t,t\}$。其中,$x_0$、$\dot{x}_0$、$\ddot{x}_0$分别为初始时刻车辆的横向位置、速度、加速度;$x_t$、$\dot{x}_t$、$\ddot{x}_t$分别为规划周期末时刻车辆的横向位置、速度、加速度。为保证横向轨迹的平顺性和连续性,在实际规划时$\dot{x}_t$、$\ddot{x}_t$均设为0值;t为规划时间。基于上述分析可见,在横向规划过程中需采样的末状态参数为x_t和t。其中,参数x_t的采样区间由上层博弈决策结果确定。采用5次多项式描述车辆横向位置随时间变化关系,如下所示:

$$L_A(t)=a_0+a_1t+a_2t^2+a_3t^3+a_4t^4+a_5t^5 \tag{19}$$

$$\dot{L}_A(t)=a_1+2a_2t+3a_3t^2+4a_4t^3+5a_5t^4 \tag{20}$$

$$\ddot{L}_A(t)=2a_2+6a_3t+12a_4t^2+20a_5t^3 \tag{21}$$

$$\text{s.t.}\begin{cases}L_A(0)=x_0,\quad L_A(t)=x_t\\ \dot{L}_A(0)=\dot{x}_0,\quad \dot{L}_A(t)=\dot{x}_t\\ \ddot{L}_A(0)=\ddot{x}_0,\quad \ddot{L}_A(t)=\ddot{x}_t\end{cases} \tag{22}$$

式中:$\dot{L}_A(t)$、$\ddot{L}_A(t)$分别为式(19)对时间的一阶导数、二阶导数,其物理含义分别为横向速度、加速度;a_0、a_1、a_2、a_3、a_4、a_5为系数,如下所示:

$$a_0=x_0 \tag{23}$$

$$a_1=\dot{x}_0 \tag{24}$$

$$a_2=\ddot{x}_0/2 \tag{25}$$

$$a_3=\frac{-(20x_0-20x_t+12t\dot{x}_t+8t\dot{x}_t+3t^2\ddot{x}_0-t^2\ddot{x}_t)}{2t^3} \tag{26}$$

$$a_4=\frac{30x_0-30x_t+16t\dot{x}_0+14t\dot{x}_t+3t^2\ddot{x}_0-t^2\ddot{x}_t}{2t^4} \tag{27}$$

$$a_5=\frac{-12x_0+12x_t-6t\dot{x}_0-6t\dot{x}_t-t^2\ddot{x}_0+t^2\ddot{x}_t}{2t^5} \tag{28}$$

2.1.2 车辆纵向轨迹

与横向行驶过程相比,车辆在纵向上行驶更希望在保证安全平稳的前提下保持合理速度,故将速度作为纵向规划时的关键因素,将车辆纵向状态关键参数定义为$\{y_0,\dot{y}_0,\ddot{y}_0,\dot{y}_t,\ddot{y}_t,t\}$。其中,$y_0$、$\dot{y}_0$、$\ddot{y}_0$分别为车辆初始时刻纵向位置、速度、加速度,$y_0$、$\dot{y}_0$由车辆当前状态确定,$\ddot{y}_0$由上层博弈决策结果确定;$\dot{y}_t$、$\ddot{y}_t$分别为规划周期末时刻纵向速度、加速度,为保证轨迹具有最小的急动度,$\ddot{y}_t$设为0。在纵向规划过程中需采样的末状态参数为$\dot{y}_t$、t。采用4次多项式表示车辆纵向位置随时间的变化关系,如下所示:

$$L_O(t)=b_0+b_1t+b_2t^2+b_3t^3+b_4t^4 \tag{29}$$

$$\dot{L}_O(t)=b_1+2b_2t+3b_3t^2+4b_4t^3 \tag{30}$$

$$\ddot{L}_O(t)=2b_2+6b_3t+12b_4t^2 \tag{31}$$

$$\text{s.t.}\begin{cases}L_O(0)=y_0\\ \dot{L}_O(0)=\dot{y}_0,\quad \dot{L}_O(t)=\dot{y}_t\\ \ddot{L}_O(0)=\ddot{y}_0,\quad \ddot{L}_O(t)=\ddot{y}_t\end{cases} \tag{32}$$

式中:$\dot{L}_O(t)$、$\ddot{L}_O(t)$分别为式(29)对时间的一阶导数、二阶导数,其物理含义分别为纵向速度、加速度;b_0、b_1、b_2、b_3、b_4为系数,如下所示:

$$b_0=y_0 \tag{33}$$

$$b_1=\dot{y}_0 \tag{34}$$

$$b_2=\ddot{y}_0/2 \tag{35}$$

$$b_3=\frac{-(3y_0-3y_t+2t\ddot{y}_0+t\ddot{y}_t)}{3t^2} \tag{36}$$

$$b_4=\frac{2y_0-2y_1+t\dot{y}_0+t\dot{y}_1}{4t^3} \tag{37}$$

本文研究对象为直行路段,故上述横、纵向轨迹均在笛卡尔坐标系下讨论。对于道路中心线曲率不为零的弯道路段,可先将当前交通环境转化为Frenet坐标系[26,27],再按上述方法生成备选轨迹。

2.2　轨迹代价函数设置

采样规划周期末状态关键运行参数,并进行多项式轨迹拟合,即可生成多条备选轨迹,每条备选轨迹都针对一种特定的运动方案。以博弈决策代价函数的设置原则为基准选取安全性、舒适性、高效性作为轨迹代价函数,代价函数 C 如下所示:

$$C = w_c C_c + w_{J_x} C_{J_x} + w_{J_y} C_{J_y} + w_o C_o + w_v C_v + w_t C_t \tag{38}$$

式中:C_c 为规划车辆与其他车辆之间的安全代价;C_{J_x}、C_{J_y} 分别为车辆横向和纵向急动度,为舒适性代价;效率指标由 C_o、C_v、C_t 组成,其中,C_v 为规划车辆速度与最大行驶速度之间的差值,C_o 为规划车辆横向位置与目标横向位置之间的差值,C_t 为行驶时间;w_c、w_{J_x}、w_{J_y}、w_o、w_v、w_t 均为权重项。

代价分量如表 2 所示。其中,D 为规划周期内规划车辆与其他车辆纵向距离差值的平方和。

表 2　轨迹规划代价函数

Table 2　Cost function of motion planning

含义	权重	代价分量
安全性	w_c	$C_c = 1 \Big/ \sum_0^t D$
舒适性	w_{J_x}	$C_{J_x} = \int_0^t \dddot{L}_A^2(t)\,\mathrm{d}t$
	w_{J_y}	$C_{J_y} = \int_0^t \dddot{L}_O^2(t)\,\mathrm{d}t$
高效性	w_o	$C_o = \int_0^t L_O^2(t)\,\mathrm{d}t$
	w_v	$C_v = \left(\dot{L}_A(t) - v_{\max}\right)^2$
	w_t	$C_t = t$

2.3　求解过程

模拟退火和粒子群等启发式寻优算法在求非线性函数的极值问题方面具有显著优势[28]。模拟退火算法的求解思路是在当前唯一解的基础上,以概率的形式进行迭代搜索最优解,直到找到满足要求的极值[29]。而粒子群算法则是预先设置一定规模的初始解,迭代求解最优值[30]。与粒子群算法相比,模拟退火算法因其基于一个可行解进行启发而具有更高的计算效率,以概率的形式接受新解使得全局收敛能力更强,而当其接近稳定状态时,更不容易接受差解,故局部搜索能力更为精准。鉴于模拟退火算法在计算效率和求解准确性方面的优势,本文采用模拟退火算法选取车辆运行最优轨迹。

3　仿真实验及结果分析

3.1　仿真场景及参数设置

本节分析智驾车在路段环境中与不同类型驾驶员交互过程中的博弈决策和轨迹规划行为,并采用模拟退火算法进行优化求解,算法流程图如图 1 所示。首先,根据车辆驾驶环境信息确定博弈交互对象与博弈策略,计算博弈代价求得博弈均衡解,即车辆行驶最优决策参数;在此基础上,进行轨迹规划生成多条备选轨迹,采用模拟退火算法选取最优轨迹并执行,更新车辆状态继续进行博弈规划,直到车辆完成换道过程。

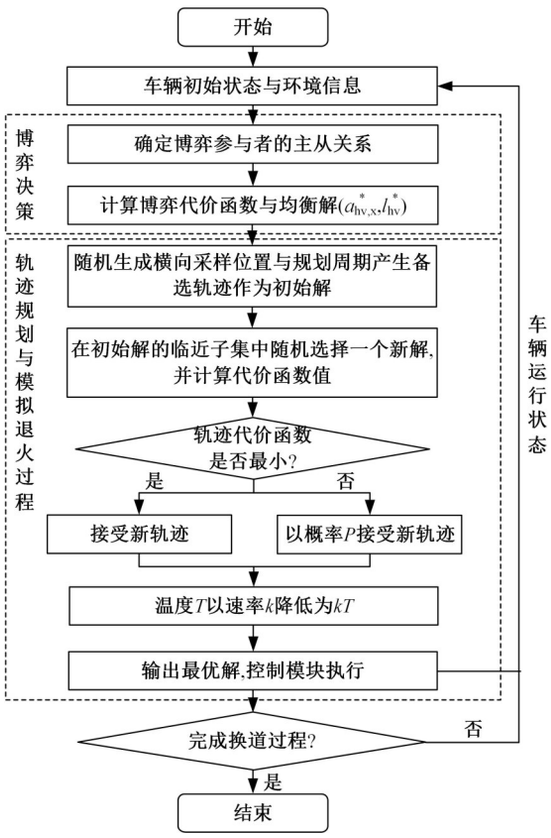

图 1　决策规划流程

Fig. 1　Decision making and motion planning workflow

仿真场景为单向 3 车道路段,车辆相对位置如图 2 所示。道路长度总为 200 m,均为直行路段,车道宽度为 3.6 m。仿真过程均在处理器为 AMD Ryzen 5 4600H (3.0 GHz)的计算机上完成。仿真开始时,智驾车附近存在 3 辆不同驾驶特性、不同速度行驶的人工驾驶车辆,初始行驶状态参数如表 3 所示。其中,驾驶特性参数集合中第 1 项为博弈效率权重,第 2 项为博弈舒适性权重,第 3 项为安全性权重,当博弈效率指标权重大于等于 0.7 时,驾驶类型为激进型驾驶员,当安全

权重大于等于0.5时,为保守型驾驶员,其他情况则为普通型驾驶员。仿真测试共设置两个场景,主要目的为当选择车1作为智驾车的主要交互对象且呈现不同驾驶特性时,测试智驾车的驾驶表现。其中,智驾车采用本文提出的基于博弈-规划循环迭代框架的换道行为模型,其他车辆采用预设的驾驶策略。

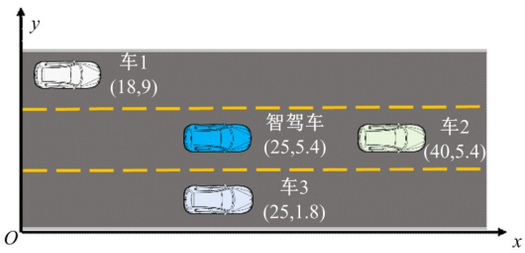

图2 仿真场景及车辆初始位置

Fig. 2 Simulation scenario and initial positions of vehicles

表3 车辆运行参数设置

Table 3 Driving parameter settings of vehicles

参 数	数 值
车辆长度/m	4.2
车辆宽度/m	2
最小碰撞半径/m	3
加速度/(m·s^{-2})	[−3,3]
车辆可行速度/(m·s^{-1})	[0,17]
智驾车速度/(m·s^{-1})	12.5
智驾车加速度/(m·s^{-2})	0
智驾车驾驶特性参数	0.5,0.3,0.2
车1速度/(m·s^{-1})	12.5
车1加速度/(m·s^{-2})	0
车1驾驶特性参数	0.1,0.1,0.8
车2速度/(m·s^{-1})	11
车2加速度/(m·s^{-2})	0
车2驾驶特性参数	0.3,0.5,0.2
车3速度/(m·s^{-1})	12.5
车3加速度/(m·s^{-2})	0
车3驾驶特性参数	0.5, 0.3, 0.2

3.2 仿真结果分析

3.2.1 博弈决策-轨迹规划结果分析

当车1为保守型驾驶特性、智驾车与其他车辆均为普通型时,换道过程中车辆的运行参数如图3所示。仿真开始时,处于智驾车前方的车2行驶速度为11 m/s,小于智驾车行驶速度,智驾车为追求更高的行驶速度而与各车辆进行博弈决策以期换道到速度较快的车道上。分别计算智驾车与车1、车2、车3博弈的博弈代价值分别为

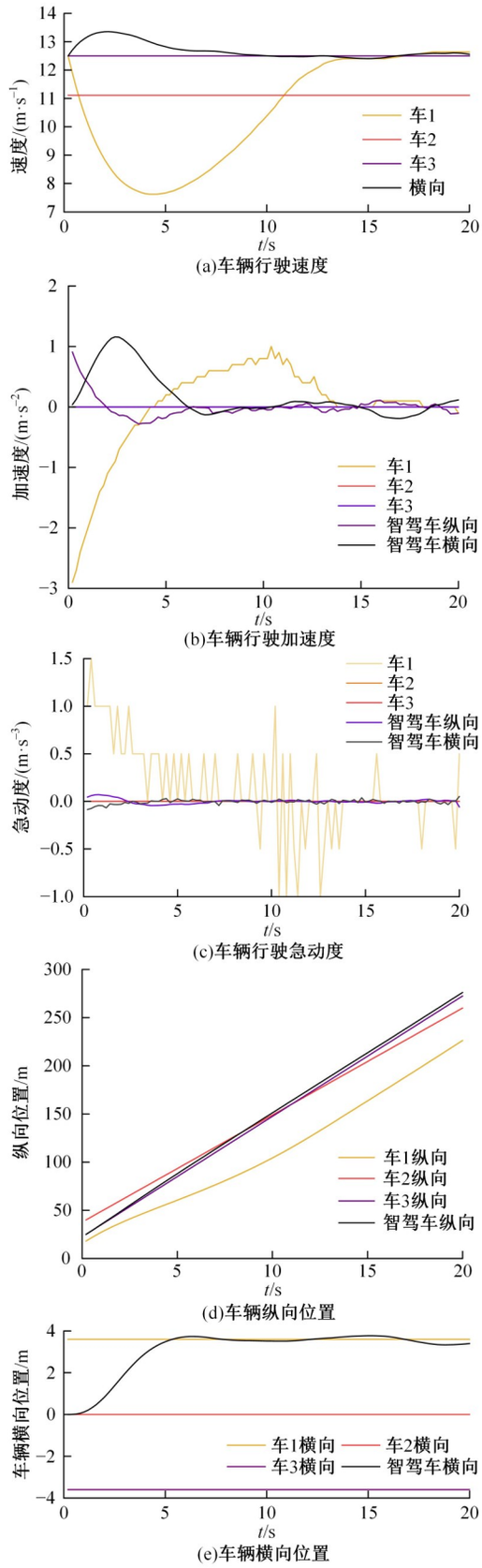

图3 车1为保守型、智驾车普通型时的行驶参数

Fig. 3 Driving parameters when vehicle1 is conservative and AV is normal

0.55、0.63、0.87,故选择代价最小的车 1 作为博弈对象,并期望换道至其所在车道上。因车 1 为保守型,故当车 1 接收到智驾车有换道意图时,决定减速让行,二者间初始车头时距为 0.56 s,故车1采取较大的加速度以保持安全车头时距,而智驾车则开始缓慢加速,以期在保证安全的前提下快速完成换道过程。当智驾车顺利换道后,两车均以其期望行驶速度行驶。在整个换道过程中,智驾车急动度始终介于[−0.2,0.2]之间,小于舒适度边界值 1。交互过程中车 1 与智驾车的博弈代价值如图 4 所示,其中,图 4(a)表示当智驾车的加速度介于[0,1] m/s²之间时,博弈代价最小,对应的驾驶策略最优,而图 4(b)表示当车 1 的加速度介于[0,−1.5] m/s²之间时,车 1 的博弈代价最小。

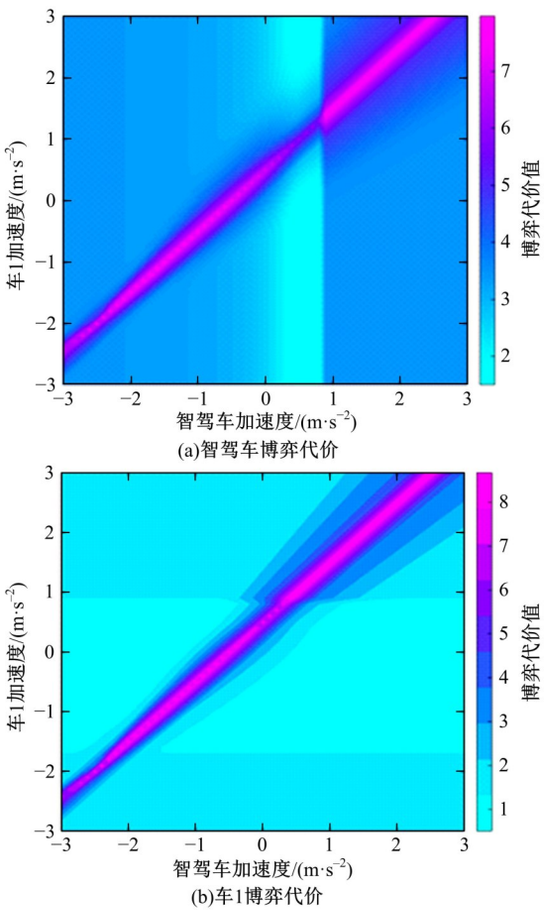

图 4　博弈代价热力图

Fig. 4　Heat map of game cost

当车 1 呈激进型驾驶特性、智驾车与其他车辆均为普通型时,智驾车的换道过程如图 5 所示。仿真开始时,智驾车分别与车 1、车 2、车 3 进行主从博弈,其博弈代价分别为 0.85、0.63、0.87,故

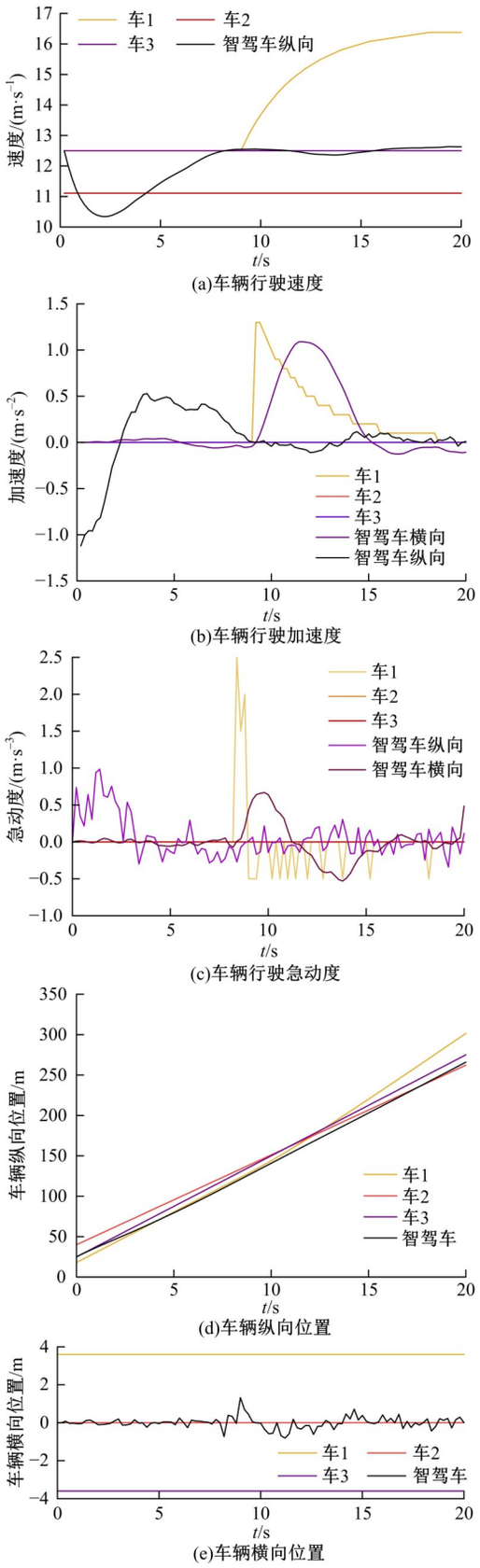

图 5　车 1 为激进型、智驾车普通型时的行驶参数

Fig. 5　Driving parameters when vehicle1 is aggressive and AV is normal

选择代价最小的车2作为博弈对象,并在原车道上跟驰行驶。二者间初始车头时距为1.20,小于3 s安全车头时距准则[31],故智驾车采取较大的减速度以保持安全车头时距,急动度始终小于边界值1。因车1较为激进,故其在与车2并排行驶时采取加速通过。车2为智驾车前方车辆,其最优博弈代价即为保持当前行驶状态。

3.2.2 仿真结果与实际轨迹数据对比分析

为验证本文决策规划算法在实际应用中的有效性,选取真实换道过程中的加速度、急动度和换道车辆车头时距与本文仿真结果进行对比,如图6所示。选取NGSIM数据库中US101路段上的108个车辆单次换道过程,识别关键换道时间点,并选取关键时间点前10 s和后10 s共20 s的换道过程运动参数[32]。图6(a)表明在整个换道过程中,实际轨迹加速度分布在[−4,4] m/s²之间,智驾车的加速度分布集中在[−1.5,1.5] m/s²内,加速度变化范围更小;图6(b)表明智驾车的急动度介于舒适范围之内,远远小于实际轨迹急动度,表明本文算法与实际人工驾驶车辆相比舒适度更佳;图6(c)表明,实际换道过程车头时距分布集中在2.2 s附近,且分布更加分散,智驾车车头时距分布集中在1.3 s附近,相比于实际车头时距分布,总体呈现更小的车头时距,在保证安全车头时距的前提下,运行效率更高。

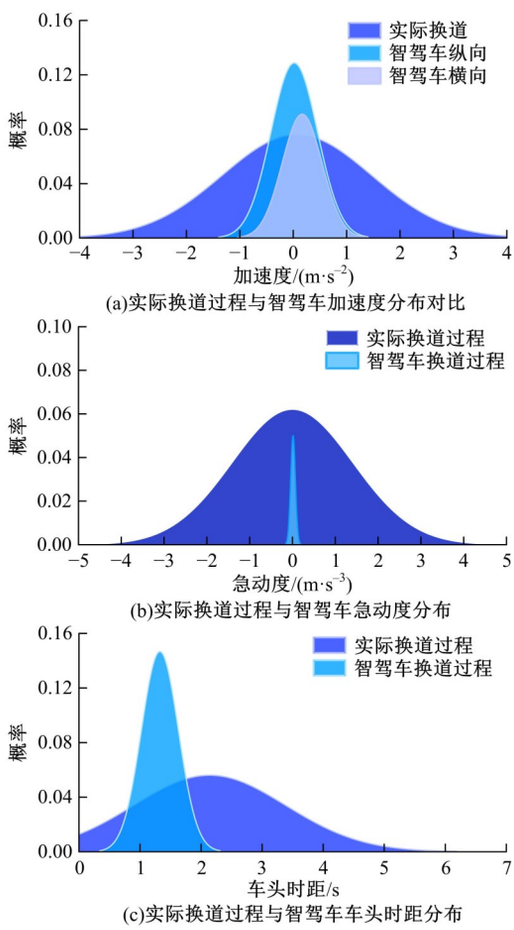

图6 实际轨迹与智驾车运行参数对比

Fig. 6 Driving parameters comparison between real lane changing process and AV

3.2.3 寻优算法对比分析

为验证不同寻优算法对整个决策规划过程中的效率提升效果,对比分析了遍历算法、模拟退火(SA)算法和粒子群(PSO)算法在轨迹规划采样过程以及求解最优轨迹过程中的表现,结果如图7所示。3种求解过程采用相同的可行采样空间,如表4所示,模拟退火参数设置如表5所示,粒子群算法参数设置如表6所示。

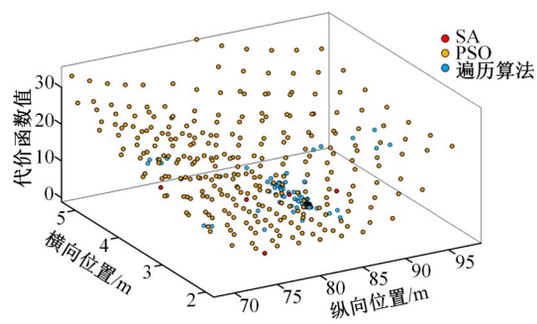

图7 遍历、模拟退火、粒子群算法可行解空间对比

Fig. 7 Comparison of feasible solution space between ergodic method, PSO, and SA

表4 轨迹规划采样空间设置

Table 4 Setting of sampling space for trajectory planning

参数		取值
横向位置采样范围/m	左转	[1.8,5.4]
	直行	[−1.8,1.8]
	右转	[−5.4,−1.8]
横向位置采样间隔/m		0.1
规划时间采样范围/s		[4,5]
规划时间采样间隔/s		0.1
横向加速度采样范围/(m·s⁻²)		[−3,3]
横向加速度采样间隔/(m·s⁻²)		0.3
实际规划时间长度/s		0.1

图7和图8分别为采用遍历算法、模拟退火算法、粒子群算法和模拟退火-粒子群(SAPSO)结合算法下的可行解空间和最优轨迹参数实际搜索时间。图7中数据来源于在仿真时刻开始时第1个规划周期内智驾车搜索最优行驶轨迹的过程。

表 5 模拟退火算法参数设置

Table 5　Parameters setting of SA

参数	数值
初始温度	100
马尔科夫链长度	5
温度下降率	0.9
结束温度	3

表 6 粒子群算法参数设置

Table 6　Parameters setting of particle swarm optimization

含义	参数值
种群数量	5
最大迭代次数	20
惯性因子	0.95
加速常数	3

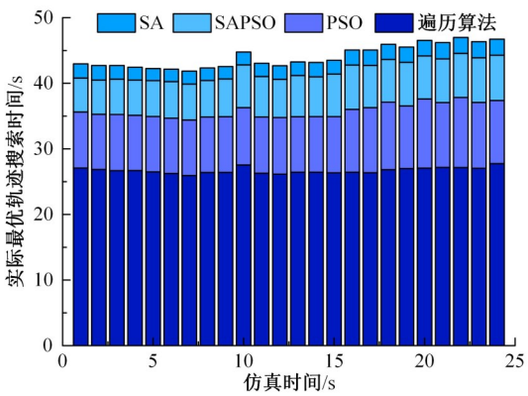

图 8　仿真运行时间对比

Fig. 8　Comparison of simulation running time

从图 7 中可以看出:遍历算法涵盖了整个离散采样空间内每个采样点,每个迭代周期需计算 336 次,相比于模拟退火算法和粒子群算法,其在理论上更加接近实际最优解,但同时也搜索到一些不必要的高代价轨迹。粒子群算法和模拟退火-粒子群结合算法在每个最优轨迹求解周期需计算 100 次。模拟退火算法平均每周期需计算 12 次,其通过随机产生可行解并对其产生随机扰动,其搜索过程仅基于当前最优轨迹,相当于对一个粒子进行迭代优化。

根据图 8 可知,每个规划周期内遍历算法耗时为 9.91～30.50 s,模拟退火算法耗时为 1.39～5.15 s,粒子群算法、模拟退火-粒子群结合算法耗时分别为 4.23～7.15 s、8.12～10.04 s。可见,模拟退火算法在寻找最优轨迹时用时最短,其原因是,模拟退火算法在迭代寻优的过程中计算的轨迹数最少,其仅需在一条轨迹的基础上进行启发搜索,而粒子群和模拟退火-粒子群结合算法由于需计算 100 次轨迹代价导致搜索时间增加。同时,由于在模拟退火-粒子群结合算法中增加了全局最优值判断机制,耗时比纯粒子群算法稍长。

图 9 中遍历算法求得的最小值为最优代价值,模拟退火算法、粒子群算法和模拟退火-粒子群结合算法最终找到的轨迹最小代价与最优值之间的差值分别为 0.097、0.143 和 0.025,基本接近遍历算法求得的最优值。虽然模拟退火算法在迭代过程中搜索到的最优轨迹代价值稍有浮动,但其最终能以较少的迭代次数求得近似最优解。

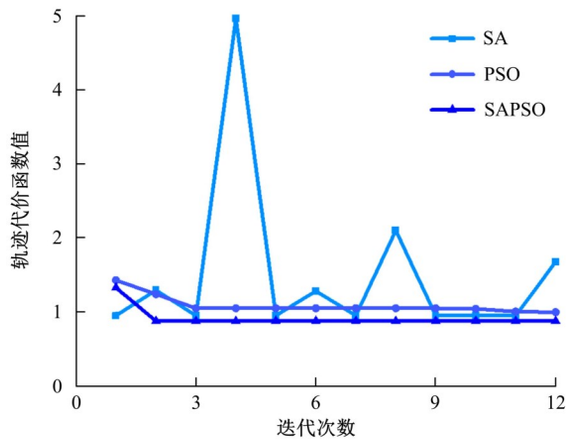

图 9　三种寻优算法收敛过程对比

Fig. 9　Convergence process comparison of SAPSO, PSO, and SA model

3.2.4　模拟退火算法不同参数组合分析

模拟退火算法参数对求解质量和计算时间有显著影响,故本文设置不同起始温度(本文中涉及的温度仅为理论上的温度概念,并不具有实际物理温度含义,故无单位)和降温比例,分析其对求解质量和计算时间的影响,并选取最优参数,测试参数及仿真结果如表 7 所示。结果表明,降温比例为 0.9,起始温度为 90 时求得的解质量最优,且耗时最少。

表 7 不同模拟退火算法参数组合仿真结果

Table 7　Simulation results of different SA parameters

组别	起始温度	终止温度	降温比例 0.9		降温比例 0.95	
			最优解	耗时/s	最优解	耗时/s
1	70	1	1.02	1.86	1.31	2.02
2	80	1	1.29	1.71	1.22	2.06
3	90	1	1.02	1.80	1.15	1.80
4	100	1	1.05	1.82	1.17	1.77
5	110	1	1.23	1.88	1.31	1.78
6	120	1	1.06	1.72	1.22	1.82
7	130	1	1.18	1.78	1.39	1.97

4 结束语

本文基于智驾车决策主导规划、规划辅助决策的研究思路,提出了一种基于博弈决策-轨迹规划联合规划框架的智驾车轨迹规划方法。该方法基于主从博弈思想构建了双层决策博弈模型,用以解决路段动态场景下智驾车的换道决策问题。设计了安全性、舒适性和高效性的线性组合博弈代价函数,将其分量分别通过负指数函数进行归一化处理,增加了模型的可解释性和可迁移性。进一步地,采用多项式描述车辆实际运行轨迹,并将博弈决策结果作为轨迹规划采样空间的约束条件。为兼顾轨迹求解效率的同时保证轨迹最优性,采用模拟退火算法求解最优轨迹。最后,验证了不同驾驶场景下本文方法的有效性。结果表明,博弈决策和轨迹规划联合规划算法能够帮助智驾车顺利完成驾驶任务,同时能够有效提高决策合理性和轨迹规划高效性。

参考文献:

[1] Hang P, Lv C, Huang C, et al. An integrated framework of decision making and motion planning for autonomous vehicles considering social behaviors[J]. IEEE Transactions on Vehicular Technology, 2020, 69(12): 14458-14469.

[2] Lu Y, Xu X, Zhang X, et al. Hierarchical reinforcement learning for autonomous decision making and motion planning of intelligent vehicles[J]. IEEE Access, 2020, 8: 209776-209789.

[3] Lee N, Choi W, Vernaza P, et al. Desire: distant future prediction in dynamic scenes with interacting agents[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, America, 2017: 2156-2174.

[4] Xu W, Pan J, Wei J, et al. Motion planning under uncertainty for on-road autonomous driving[C]//2014 IEEE International Conference on Robotics and Automation (ICRA), Hong Kong, China, 2014: 2507-2512.

[5] Schmerling E, Leung K, Vollprecht W, et al. Multimodal probabilistic model-based planning for human-robot interaction[C]//2018 IEEE International Conference on Robotics and Automation (ICRA), Brisbane, Australia, 2018: 1-9.

[6] Gindele T, Brechtel S, Dillmann R. Learning driver behavior models from traffic observations for decision making and planning[J]. IEEE Intelligent Transportation Systems Magazine, 2015, 7(1): 69-79.

[7] Schulz J, Hubmann C, Löchner J, et al. Interaction-aware probabilistic behavior prediction in urban environments[C]//2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Madrid, Spain, 2018: 3999-4006.

[8] Althoff M, Mergel A. Comparison of markov chain abstraction and monte carlo simulation for the safety assessment of autonomous cars[J]. IEEE Transaction on Intelligent Transportation System, 2011, 12(4): 1237-1247.

[9] Kamrani M, Srinivasan A R, Chakraborty S, et al. Applying markov decision process to understand driving decisions using basic safety messages data[J]. Transportation Research Part C: Emerging Technologies, 2020, 115: 1-16.

[10] 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), Los Angeles, America, 2017: 1590-1595.

[11] Lenz D, Diehl F, Le M T, et al. Deep neural networks for Markovian interactive scene prediction in highway scenarios[C]//Intelligent Vehicles Symposium, Los Angeles, America, 2017: 685-692.

[12] Xie D F, Fang Z, Jia B, et al. A data-driven lane-changing model based on deep learning[J]. Transportation Research Part C: Emerging Technologies, 2019, 106(2019): 41-60.

[13] Morton J, Wheeler T A, Kochenderfer M J. Analysis of recurrent neural networks for probabilistic modeling of driver behavior[J]. IEEE Transaction on Intelligent Transportation System, 2017, 18(5): 1289-1298.

[14] Yu H T, Tseng H E, Langari R, et al. A human-like game theory-based controller for automatic lane changing[J]. Transportation Research Part C: Emerging Technologies, 2018, 88(5): 140-158.

[15] 胡益恺, 庄瀚洋, 王春香, 等. 基于主从博弈的智能车汇流场景决策方法[J]. 上海交通大学学报, 2021, 55(8): 1027-1034.
Hu Yi-kai, Zhuang Han-yang, Wang Chun-xiang, et al. Stackelberg-game-based intelligent vehicle decision method for merging scenarios[J]. Journal of Shanghai Jiaotong University, 2021, 55(8): 1027-1034.

[16] Kang K, Rakha H A. A repeated game freeway lane

changing model[J]. Sensors, 2020, 20(6):1554-1588.

[17] Hang P, Lv C, Huang C, et al. An integrated framework of decision making and motion planning for autonomous vehicles considering social behaviors[J]. IEEE Transactions on Vehicular Technology, 2020, 69(12): 14458-14469.

[18] Katrakazas C, Quddus M, Chen W H, et al. Real-time motion planning methods for autonomous on-road driving: state-of-the-art and future research directions[J]. Transportation Research Part C: Emerging Technologies, 2015, 60: 416-442.

[19] Gonzalez D, Pérez J, Milanes V, et al. A review of motion planning techniques for automated vehicles[J]. IEEE Transactions on Intelligent Transportation Systems, 2016, 17(4): 1135-1145.

[20] Lim W, Lee S, Sunwoo M, et al. Hybrid trajectory planning for autonomous driving in on-road dynamic scenarios[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 22(1): 341-355.

[21] Xu W D, Wei J Q, Dolan J M, et al. A real-time motion planner with trajectory optimization for autonomous vehicles[C]//2012 IEEE International Conference on Robotics and Automation, St Paul, America, 2012: 2061-2067.

[22] Ziegler J, Bender P, Dang T, et al. Trajectory planning for Bertha-A local continuous method[C]//Intelligent Vehicles Symposium, Michigan, America, 2014: 450-457.

[23] Jiang Y, Jin X, Xiong Y, et al. A dynamic motion planning framework for autonomous driving in urban environments[C]//In Proceedings of the 2020 39th Chinese Control Conference, Shenyang, China, 2020: 5429-5435.

[24] Liao X, Zhao X, Wu G, et al. A game theory based ramp merging strategy for connected and automated vehicles in the mixed traffic: a Unity-SUMO integrated platform[DB/OL]. [2022-01-27]. https://arxiv.org/abs/2101.11237

[25] Zhou D, Ma Z, Sun J. Autonomous vehicles turning motion planning in multi-interaction area of mixed-flow intersections[J]. IEEE Transactions on Intelligent Vehicles, 2020, 5(2): 204-216.

[26] Chu K, Lee M, Sunwoo M. Local path planning for off-road autonomous driving with avoidance of static obstacles[J]. IEEE Transactions on Intelligent Transportation Systems, 2012, 13(4): 1599-1616.

[27] Werling M, Kammel S, Ziegler J, et al. Optimal trajectories for time-critical street scenarios using discretized terminal manifolds[J]. International Journal of Robotics Research, 2012, 31(3): 346-359.

[28] 陈华根, 吴健生, 王家林,等. 模拟退火算法机理研究[J]. 同济大学学报:自然科学版, 2004, 32(6): 802-805.
Chen Hua-gen, Wu Jian-sheng, Wang Jia-lin, et al. Mechanism study of simulated annealing algorithm [J]. Journal of Tongji University (Natural Science), 2004, 32(6): 802-805.

[29] Kirkpatrick S, Gelatt C D, Vecchi M P. Optimization by simulated annealing[J]. Science, 1983, 220 (4598): 671-680.

[30] 唐俊. PSO算法原理及应用[J]. 计算机技术与发展, 2010, 20(2): 213-216.
Tang Jun. Principle and application of PSO algorithm [J]. Computer Technology and Development, 2010, 20(2): 213-216.

[31] Qiao F, Rahman R, Li Q, et al. Safe and environment-friendly forward collision warning messages in the advance warning area of a construction zone[J]. International Journal of Intelligent Transportation Systems Research, 2017, 15: 166-179.

[32] Vishal M, Christos K, Constantinos A. Prediction of lane-changing maneuvers with automatic labeling and deep learning[J]. Transportation Research Record, 2020, 2674(7): 336-347.