

基于自然驾驶数据的交互轨迹基元表征与提取*

李子睿^{1,2}, 王浩闻¹, 龚建伟¹, 吕超¹, 赵晓聪³, 王猛²

(1. 北京理工大学机械与车辆学院, 北京 100081; 2. 德累斯顿工业大学, 德国 01067;

3. 同济大学, 道路与交通工程教育部重点实验室, 上海 201804)

[摘要] 在共享道路空间中不同流向道路使用者间存在通行路径冲突, 为规避碰撞风险, 道路使用者须通过驾驶交互进行路权协商, 从而消解潜在冲突。对交互行为的表述和建模, 对于准确理解和预测动态环境具有重要意义。为此, 本文提出一种以交互基元为分析单元的多车驾驶交互行为语义级表征和提取方法。首先, 利用非参数贝叶斯方法对交互驾驶行为进行分割, 得到具有显著行为模式的驾驶交互片段。然后, 利用黏性层次狄利克雷-隐马尔可夫模型, 从驾驶交互片段中提取得到交互基元。最后, 对规范化处理后的交互基元进行无监督聚类, 以获得驾驶交互场景的语义级行为特征。基于NGSIM高速公路数据集中20 797组多车交互数据的实证研究, 本文提出的方法可提取并分析多个体参与的复杂交互场景, 突破了已有研究中只针对两车交互场景构建交互基元的局限性, 可支撑对多交通参与者交互的本质进行分析。实验结果表明, 本文所提出的方法可以将连续的驾驶行为划分为离散的交互基元。且聚类划分结果可以与实际交互场景相对应, 用于不同交互轨迹基元中车辆之间的交互行为特性分析。同时, 该方法对于复杂场景下游驾驶任务具有提升作用。在车辆多步轨迹预测任务中, 相比于基线方法, 本文所提出的交互基元提取方法在与基线预测方法融合后可以将平均预测误差和终点预测误差分别降低19.3%和14.6%。

关键词: 交互行为; 轨迹基元; 非参数化贝叶斯方法; 无监督聚类

Interactive Trajectory Primitives Representation and Extraction Based on Naturalistic Driving Data

Li Zirui^{1,2}, Wang Haowen¹, Gong Jianwei¹, Lü Chao¹, Zhao Xiaocong³ & Wang Meng²

1. School of Mechanical Engineering, Beijing Institute of Technology, Beijing 100081;

2. TU Dresden, Germany 01067;

3. Tongji University, Key Laboratory of Road and Traffic Engineering, Ministry of Education, Shanghai 201804

[Abstract] In shared road space, there are path conflicts between different road users moving in various directions. Road users must negotiate right-of-way through driving interactions to avoid collision risks, thus resolving potential conflicts. The description and modeling of interactive behaviors is crucial for accurately understanding and predicting the dynamic environment. Therefore, a semantic-level representation and extraction method for multi-vehicle interactive behaviors is proposed in this paper, taking interactive trajectory primitives as analysis units. Firstly, a nonparametric Bayesian method is utilized to segment interactive behaviors, obtaining interaction segments with significant behavior patterns. Then, the sticky hierarchical Dirichlet-Hidden Markov Model is employed to extract interaction primitives from these interaction segments. Finally, unsupervised clustering is applied to the normalized interaction primitives to obtain semantic-level behavioral features of interaction scenarios. An empirical study based on 20 797 pairs of multi-vehicle interaction data from the NGSIM highway dataset shows that the method proposed in this paper can extract and analyze complex interactive scenarios involving multiple participants, breaking through the limitation of existing research that only constructs interaction primitives for two vehicle interac-

* 国家自然科学基金(61703041, U19A2083)资助。

原稿收到日期为2024年01月12日, 修改稿收到日期为2024年04月01日。

通信作者: 龚建伟, 教授, 博士生导师, 工学博士, E-mail: gongjianwei@bit.edu.cn。

tion scenarios, and supporting the analysis of interaction among multiple traffic participants. The experimental results show that the proposed method can segment continuous driving behaviors into discrete interaction primitives. The clustering results correspond to actual interaction scenarios and can be used to characterize the interaction behaviors among vehicles in different interactive trajectory primitives. Furthermore, the method can enhance performance of downstream driving tasks in complex scenarios. In multi-step vehicle trajectory prediction, by integrating with baseline prediction methods, the proposed method can reduce the average prediction error and final position error by 19.3% and 14.6%, respectively, compared to baseline methods.

Keywords: interactive behaviors; trajectory primitive; non-parametric Bayesian method; unsupervised clustering

前言

驾驶行为建模对于发展高级驾驶辅助系统(advanced driving assistance system, ADAS)和自动驾驶都具有重要作用。现阶段研究驾驶行为建模的主要方法有微观交通仿真模型^[1]、统计学习模型^[2]、深度学习模型^[3]、强化学习模型^[4]等,利用上述方法所建立的模型已广泛应用于车辆跟驰、换道、超车等场景中。随着车路协同和计算机技术的发展与进步,在系统中考虑多个交通参与者的交互行为可以模拟真实道路环境中参与者的相互作用与影响,进而提升强交互环境中车辆、行人以及非机动车的安全性。同时,对于交互行为的精确建模也有助于更加合理地规划和使用潜在冲突区域,以提高交叉路口等场景中的交通效率。交互行为模型可以刻画交通参与者之间的影响程度,并可有效应用于轨迹预测、跟踪、意图辨识等实际问题中。

交通参与者的交互及决策行为可用连续运动及离散状态共同描述,连续运动体现交通参与者在时间上的连续控制行为,离散状态表征连续动作下交通参与者的潜在决策行为。在反映连续运动及离散状态关系的众多模型中,最常用的方法是将决策行为建模为离散状态系统,并将各离散状态系统中的运动表征为连续时间过程^[5-6]。例如在交叉路口场景中,车辆的决策可划为直行、左转、右转、停止这4个离散状态^[7];在换道场景中,车辆可选择左换道、右换道和保持车道这3种决策行为^[8];在跟车场景中,车辆可选择加速、减速或匀速行驶。这类方法的不足是需要预先指定离散状态,且仅可对单一个体的决策行为进行归类。

实际道路环境中的驾驶交互往往交通参与者较多,因此对多个交通参与者交互工况下决策行为的

研究尤为重要。在多个个体交互工况中,个体间的决策行为存在相互影响,因此单个个体的决策状态无法直接独立预设。针对这类个体间动作强耦合的复杂交互场景,移动机器人领域常采用基于强化学习和动态优化类的方法进行决策问题建模和求解^[9-10]。这类算法通常需要在较大搜索空间内探索最优解,因此当应用于自动驾驶车辆决策系统时,面临解空间过大或所建模的优化问题复杂性过高的问题,难以满足对于决策实时性的要求。

针对上述两种方法的问题,一些研究中提出基于基元的研究方法。基元是指能表征机器人基本决策和运动规划的基本运动单元。该类方法用最基本的运动单元对运动形式进行表述,并在较小的解空间内进行搜索以提高时间效率。Kober等^[11]建立了完整的动态运动基元理论,并成功应用于多自由度机械臂,在此基础上研究者们提出了一系列将动态运动基元与概率预测、强化学习、模仿学习、迁移学习^[12-14]进行结合的方法。此类方法利用基元表征提取建立知识库,并进行泛化迁移,在关节小幅度运动的机械臂中取得了优异的性能表现。但是,相较于机械臂应用场景中所采集的高重复性的示范轨迹,智能交通系统应用中所采集的自然驾驶数据具有低重复、高动态的特点。针对此问题,Wang等^[15]使用基于动态运动基元方法对驾驶员的轨迹基元和操控基元进行联合建模和表征,并以基元类别为依据建立驾驶员跟踪行为模型,提升了转向操控量的建模精度。

上述方法的特点是将提取到的基元与智能体的控制相结合,但并未对多智能体交互场景中的运动基元进行分析。Wang等^[16]提出利用非参数贝叶斯方法对两车交互这一典型场景中的轨迹基元进行聚类提取和统计学分析。基于此,相关研究提出了一系列关于轨迹基元生成的方法^[7, 17]。这其中包括利

用层次狄利克雷-隐马尔科夫模型对有信号灯交叉路口的两车交互轨迹进行分割和基元提取。但上述研究仅针对两车交互场景进行建模和分析,没有对多车参与的复杂交互场景进行统一建模和分析。针对多车交互行为的建模问题,在多车轨迹预测领域已被广泛研究。Li等^[18]提出将交通参与者的语义级别意图信息融入到交互行为预测模型中。Zhang等^[17]将基于速度信息提取到的驾驶基元融入到人工势能场并生成轨迹预测结果,所建立的人工势能场在模型构建中可以与最小作用力等方法结合并用于风险评估与安全控制^[19]。这体现了驾驶基元对于下游轨迹预测任务效果的提升。基于自然驾驶数据的交互行为表征、提取与分析可以解决常见的单一指标(例如车头时距、TTC等)无法精确表征多车之间交互关系的问题。但是,上述方法未将交互驾驶基元以高级语义信息的形式融入到下游轨迹预测中。

因此,现有研究需要一种可以体现车辆交互信息的、具有通用性的、对下游任务有提升效果的基元提取方法。针对上述研究的不足,本文提出了多车交互基元的表征、建模和提取方法,对多车交互环境下的交互驾驶行为进行表述。本文研究框架如图1所示。利用黏性层次狄利克雷-隐马尔可夫模型对多车轨迹进行分割并提取交互基元。采取图像像素信息对多车(车辆数目不固定)交互过程进行统一表征,最终使用无监督聚类方法处理基元以获得交互行为的本质特征。对NGSIM^[20]高速公路自然驾驶数据集进行预处理,实现了对交互驾驶行为的系统分析。在本研究中,交互基元是在中心车辆与一个或多个周围车辆组合情况下的多车交互演化过程的描述,也是对从中心车视角下的多车交互行为的通用化表征。通过本文所提出方法得到的交互基元可以将复杂、动态演变的车辆间交互简化为最基本的单元(基元)。交互基元有助于中心车辆对存在交互的动态环境进行场景理解^[21]。在本文的实验中,展示了交互基元对于下游轨迹预测精度的提升作用。

1 交互基元的表征与提取

1.1 变量及符号说明

在实际情况中,由于道路环境的多样性与复杂性,车辆处于多车的交互环境中,即中心车辆(human vehicle, HV)附近某一范围内会存在一辆或多辆的周围车辆(surrounding vehicles, SV)。该范

围内车辆的行为会相互影响,称为一个交互环境。当某一交互环境中的车辆行为发生改变,该交互环境内影响的发出者和接收者都会因此而改变,原交互环境即会被破坏——交互进入下一阶段。这样的交互阶段称为一个交互组。某一HV的完整驾驶过程可表达为

$$S = \{S_1, S_2, \dots, S_N\} \quad (1)$$

式中 N 为交互组的数量,多个交互组 $S_i (1 \leq i \leq N)$ 连续组成某一HV的完整驾驶过程 S 。假设 t 时刻的交互环境内SV的数量为 $n(t)$,各SV的ID所组成的集合为 $D(t)$,该交互场景处于交互组 S_j 中。则 $t+1$ 时刻交互场景隶属于新交互组 S_{j+1} 的(充分必要)条件是:

$$[n(t+1) \neq n(t)] \vee [D(t+1) \neq D(t)] \quad (2)$$

在某一交互组 S_i 持续的时间范围内,交互环境中车辆实体不会发生变化, S_i 可以被描述为环境中HV和SV特征的组合:

$$S_i = \{X_1, \dots, X_r, \dots, X_T\} \quad (3)$$

$$X_t = [l_t^{(HV)}, l_t^{(SV_1)}, \dots, l_t^{(SV_n)}, v_t^{(HV)}, v_t^{(SV_1)}, \dots, v_t^{(SV_n)}, a_t^{(HV)}, a_t^{(SV_1)}, \dots, a_t^{(SV_n)}] \in \mathbb{R}^{(4n+4)} \quad (4)$$

式中: n 为交互环境中SV的总数; $l_t^{(HV)} \in \mathbb{R}^2, l_t^{(SV_k)} \in \mathbb{R}^2 (1 \leq k \leq n)$ 分别表示HV和第 k 辆SV在 t 时刻纵向和横向的位置; $v_t^{(HV)} \in \mathbb{R}$ 和 $v_t^{(SV_k)} \in \mathbb{R} (1 \leq k \leq n)$ 分别表示HV和第 k 辆SV在 t 时刻的速度; $a_t^{(HV)} \in \mathbb{R}$ 和 $a_t^{(SV_k)} \in \mathbb{R} (1 \leq k \leq n)$ 分别表示HV和第 k 辆SV在 t 时刻的加速度; T 是 S_i 中各车信息取样点的数量。

交互组 S_i 还可被描述为交互基元的组合,这种表达方式侧重于表达车辆的驾驶特征。交互基元是含有相同特征的驾驶行为的集合,交互基元中包含了这些驾驶行为的全部信息。此处的驾驶行为是指广义的、包括多个车辆交互的多车驾驶行为,并非指单个车辆的驾驶行为。现阶段有关多车交互行为分析、建模和预测的研究大多基于公开发表的车辆轨迹数据集,如利用无人机和道路摄像头的俯视视角数据集(NGSIM^[22], HighD^[23], INTERACTION^[24]等)和利用拥有多传感器采集车所构建的数据集(Waymo^[25], Nuscenes^[26]等)。多车交互过程的变化最终会体现为车辆轨迹的变化。但是从最终所观测到的轨迹无法精确对应驾驶员意图的变化。因此,参考文献[7]、文献[16]、文献[17]和文献[27]等有关驾驶基元的研究,本研究假设当驾驶意图发生变化时,驾驶员的操控行为和车辆轨迹会随之发生改变。在多车交互环境中,当研究范围内某一车辆的

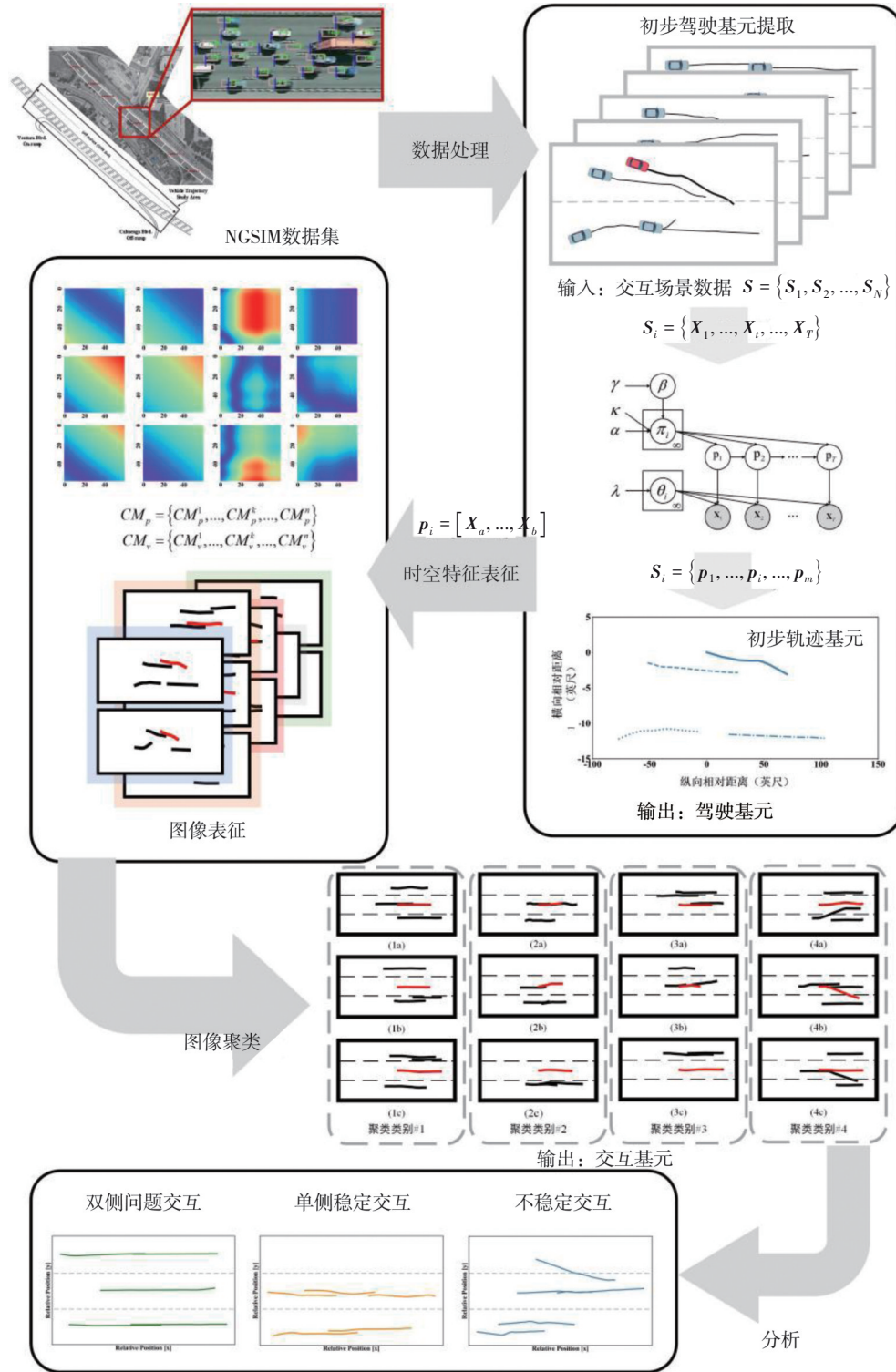


图1 交互基元的表征与提取流程图

驾驶特征(意图)发生变化时,改变时刻前后的驾驶过程可被划分为两种交互基元;在同一交互组内,周围车辆交互基元的切割划分是跟随中心车辆进行。根据驾驶特征的不同,交互组可被拆分为一个或多个交互基元的连续组合:

$$S_i = \{p_1, \dots, p_i, \dots, p_m\} \quad (5)$$

$$p_i = [X_a, \dots, X_b] \quad (6)$$

式中: $1 \leq a \leq b \leq T$; $p_i (1 \leq i \leq m)$ 为该交互组中的第 i 个基元; $m (m \geq 1)$ 为某个交互组所包含的基元数目。时间区间 $[a, b]$ 内所有时刻都属于 p_i 基元,这

些时刻的驾驶特征相同;不同基元 p_i 与 $p_j(i \neq j)$ 间驾驶特征则存在差异。

1.2 交互基元划分

基于上述对交互组及交互基元的讨论,针对任意一个交互过程 S_i ,本文将之拆分为交互基元的连续组合,以研究其中的驾驶特征。交互基元的提取本质上是针对驾驶特征的聚类,但针对的对象是连续时间序列信息而非离散的特征点。传统方法应用于聚类问题时都须预设类别数目,因此在缺少类别先验知识的问题中是无法实现的。本文中引入黏性HDP-HMM方法来处理复杂的车辆行驶信息并划分驾驶特征。传统的HMM方法须提前设定隐层的种类,由于交互组的差异性 & 先验经验的缺失,无法预先给定基元划分的种类及数目;而HDP方法则可以预先定义HMM中隐层的先验分布,一些转换参数将会影响到隐状态的选择和传递。将上述两种方法结合起来的HDP-HMM方法可以在无先验条件的前提下自动提取基元,但是基元的划分结果往往过于密集,这不符合基元持续性的要求。引入自转换偏置系数的黏性HDP-HMM方法则可解决上述问题,由此得到较为理想的基元划分结果。

1.2.1 隐马尔可夫模型(HMM)

在本问题中,交互过程中车辆各时刻的行驶信息已知(称为观测变量),希望推断出各时刻所隶属的基元(称为隐性变量)。各隐性变量间存在时间上的联系,可认为构成了马尔可夫链,因此引入HMM方法推断各时刻隐性变量的组成。

假设所有可能的基元种类共同构成一个有顺序的基元集 $P \in R^m$, m 为基元种类的数目。在关注的交互组中, t 时刻的驾驶特征 $p_t = P_i \in P(1 \leq t \leq T)$,其中 P_i 表示这是 P 中的第 i 个基元。定义 $q_i(1 \leq i \leq m)$ 为 p_t 取 P_i 的概率,由于驾驶行为的随机性,一般认为 $q_i = 1/m$ 。 P 中由第 i 个特征转变到第 j 个特征的转移概率为 $\pi_{ij}(i, j \leq m)$,转移概率构成转移矩阵 $\Pi_{m \times m}$,转移矩阵中任意一行 $[\pi_{i*}]$ 的转移概率之和为1,即 $\sum_{k=1}^m \pi_{ik} = 1$ 。 t 时刻的行驶信息 X_t 是由该时刻的驾驶特征 p_t 生成的,因此HMM模型可概括为

$$p_t | p_{t-1} \sim \pi_{p_{t-1} *}, \quad (7)$$

$$X_t \sim \Gamma(p_t, \theta_{p_t}) \quad (8)$$

式中 $\Gamma(\cdot)$ 称为发射函数,表示由 p_t 生成 X_t 的概率; θ_{p_t} 称为发射参数。

若基元集 P 已知,即可根据交互组 S_i 中的行驶

信息推断出各时刻隶属于的驾驶特征,进而将相同的驾驶特征归于同一类基元。在实际应用中,HMM模型是通过Viterbi算法^[28]实现的,其核心思想是选择生成某一观测变量序列概率最高的隐性变量组合作为实际的隐性变量划分,实质是一种递归算法。定义 $\delta_t(i)$ 为第一个驾驶特征 p_1 取 P_i 后又生成 X_1 的概率,即

$$\delta_1(i) = q_i \Gamma(p_1, \theta_{p_1}) \quad (9)$$

则根据Viterbi算法的核心思想及上述初始条件有:

$$\delta_t(j) = \max_i \{ \delta_{t-1}(i) \pi_{ij} \} \Gamma(p_t, \theta_{p_t}), \quad 2 \leq t \leq T \quad (10)$$

由于最终希望得到的是基元的组成而非概率的数值,因此须回溯基元的最优路径组合。本文中利用 $bp_t(j)$ 来保存每个时刻的参数信息:

$$bp_t(j) = \arg \max_i \delta_{t-1}(i) \pi_{ij} \quad (11)$$

上式的含义为当 $t-1$ 时刻已经为最优情况时,考察隐层变量 p_t 转为 p_j 时的参数设置。

当利用Viterbi算法迭代到 T 时刻后,即回溯之前每个时刻的基元组成:

$$p_T = \arg \max_i \delta_T(i) \quad (12)$$

$$p_t = bp_{t+1}(p_{t+1}), \quad t = T-1, T-2, \dots, 1 \quad (13)$$

在基元的研究中, P 为未知,故须引入下面介绍的HDP方法提供 P 中元素的先验分布。

1.2.2 层次狄利克雷过程(HDP)

狄利克雷过程^[29]是描述概率测度分布的随机过程,假设 p_t 是从某一分布中产生的,该分布可由参数 θ 确定,同时 θ 又遵循另一分布 $H(\theta)$,引入离散参数 γ 后,这类随机过程可被描述为 $DP(\gamma, H)$ 。若 $G_0 \sim DP(\gamma, H)$,则利用截棍构造(stick-breaking construction)可将在 DP 中随机采样得到的概率测度 G_0 表述为如下形式:

$$v_k | \gamma \sim \text{Beta}(1, \gamma), \quad k = 1, 2, \dots \quad (14)$$

$$\beta_k = v_k \prod_{\ell=1}^{k-1} (1 - v_\ell), \quad k = 1, 2, \dots \quad (15)$$

$$G_0 = \sum_{k=1}^{\infty} \beta_k \delta_{\theta_k}, \quad \theta_k | H \sim H, k = 1, 2, \dots \quad (16)$$

式中 β 是由截棍构造取样得到的、受 γ 影响的权重参数,记为 $\beta \sim \text{GEM}(\gamma)$, β_i 的和恒等于1。

X_2 由 β_i 的性质及上面对于HMM方法的讨论,HMM的转移概率测度可以利用HDP方法表示为

$$G_i = \sum_{j=1}^m \pi_{ij} \delta_{\theta_j}, \quad \theta_j | H \sim H, j = 1, 2, \dots, m \quad (17)$$

符号含义同上,令 $G_i \sim \text{DP}(\alpha, G_0)$, 其中 $G_0 \sim \text{DP}(\gamma, H)$ 。

1.2.3 黏性 HDP-HMM

基于上述讨论,本文又引入 $\kappa > 0$ 来增加自转移的概率,完整的黏性 HDP-HMM 模型^[16, 21, 30]如图2所示。黏性 HDP-HMM 方法可概括为

$$\beta | \gamma \sim \text{GEM}(\gamma) \quad (18)$$

$$\pi_i | \alpha, \beta, \kappa \sim \text{DP}\left(\alpha + \kappa, \frac{\alpha\beta + \kappa\delta_i}{\alpha + \kappa}\right), \quad i = 1, 2, \dots, m \quad (19)$$

$$p_t | p_{t-1} \sim \pi_{p_{t-1}}, \quad t = 1, 2, \dots, T \quad (20)$$

$$X_t \sim \Gamma(p_t, \theta_{p_t}), \quad t = 1, 2, \dots, T \quad (21)$$

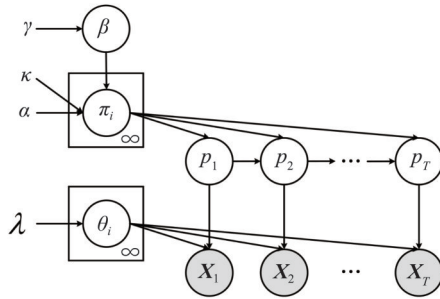


图2 黏性 HDP-HMM 方法原理图

2 交互基元的聚类

黏性 HDP-HMM 方法对连续时间数据进行分割,各数据点在时间维度相互关联;基元聚类则将提取到的交互基元看作整体,各交互基元间不存在时间上的关联。在多车的交互环境,为同时考虑所有 SV 对 HV 的影响,本文以图像形式反映交互基元中所蕴含的驾驶信息。传统的聚类方法,如 k-means 聚类、谱聚类、层次聚类等都对与图像相关的聚类问题展示出较好的结果,由于 k-means 聚类原理简单、实现容易且收敛速度快,本文中利用该方法对图像进行聚类。

本文描述了基本交互基元的划分过程。为进行基元的聚类与合并,首先通过 HDP-HMM 方法得到的基本交互基元转化为基于像素信息的图像。生成的图像信息主要由两部分组成:静态道路信息和动态车辆信息。所有图像中车辆行驶方向均为自左向右,并规定左上为图像坐标原点。经过神经网络模

型处理后将4类相同大小的卷积信息平均得到最终输出特征用于 k-means 聚类算法。从第一阶段提取到的驾驶基元到标准图像生成的过程中,只涉及尺度变换,并不涉及对于图片剪切处理。因此,在像素精度足够的前提下,不会造成原本在上游处理中存在的交互车辆在转化为图片后被遗漏。由于图片的精确程度与像素相关,而像素的大小又将直接影响聚类的效率,为尽可能多地保留图片中的信息并提升聚类速度,本文采用预训练模型对图片进行先行处理。

在视觉任务或涉及图像处理的任务中,由牛津大学视觉几何组 (visual geometry group, VGG) 提出的系列模型具有结构简单、适应性好、调用方便、稳定支持下游任务的优点^[31]。在本文模型开发阶段,作者对比了包括 VGG16 在内的多种 VGG 模型 (VGG11、VGG13、VGG16、VGG19)。测试结果表明,对本研究中的过程输出基本没有影响。本文希望可以恰当地选择 k 个聚类的中心以使得下面损失方程的结果最小:

$$\text{loss} = \sum_{\varphi \in \Phi} \min_{c \in C} \|\varphi - c\|^2 \quad (22)$$

3 实验与分析

利用黏性 HDP-HMM 方法对 NGSIM 数据集中的车辆交互过程进行交互基元的划分,并将提取到的交互基元进行聚类以获得车辆交互行为的基本组成部分。

3.1 数据集及数据处理

使用公开的 NGSIM-US101 数据集来验证模型效果。该数据集中的车辆行驶信息取自2005年6月15日7:50-8:35的美国洛杉矶101号公路路段,采样频率为10 Hz,数据均处于绝对坐标系中。该数据集中的车辆大多行驶于笔直路段,处在同向、多车道的交互环境中,且存在一定比例的超车和换道等行为。数据处理流程如图3所示。

由于交互仅发生在有限的空间范围内,因此本文仅关注中心车辆附近的驾驶行为。将 NGSIM-US101 数据集中的数据表示在二维坐标系中后,将每一辆车依次作为中心车辆,研究的横向范围划定为中心所处车道及其相邻车道。在文献[22]中,以中心车为原点,纵向前后选取90英尺作为中心车辆的感兴趣区域。对于基于 NGSIM 数据集中交互行为分析、建模和预测的后续研究多以此为参考。因

此,本研究中纵向感兴趣区域的选取范围确定为90英尺,如图3所示。同时,数据预处理过程中的统计分析表明,中心车所在车道线前方90英尺内平均车辆数目为1辆(中心车后方同理)。这避免了中心车前方有多辆车造成交互影响的传递。因此,90英尺纵向感兴趣区域的选择在数据处理阶段降低了问题

的复杂度。基于横纵向空间距离的周围车辆选取本质上是选取中心车辆周围环境中存在潜在交互的SV。此空间选取原则是参照了大多数基于NGSIM数据集的交互行为建模研究。此类空间选取原则也有利于和其他下游任务基线方法进行公平对比。

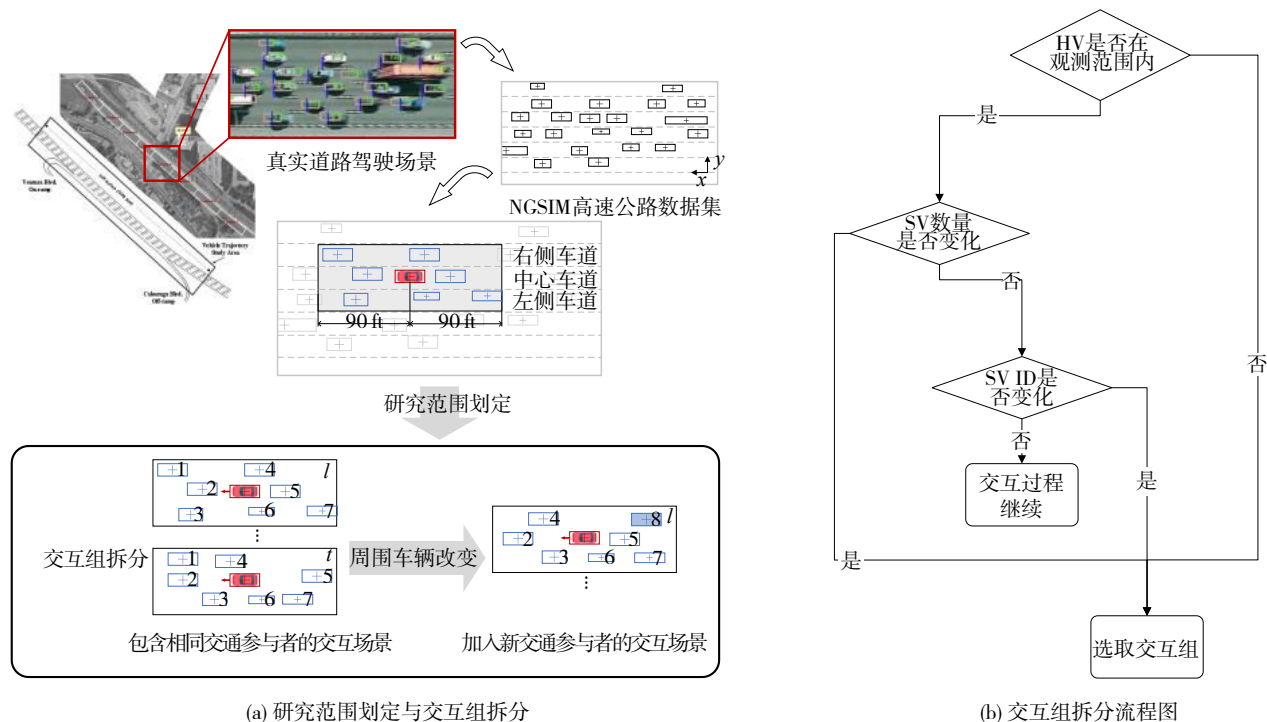


图3 数据处理流程

当所选取中心车辆位于最左侧或最右侧车道时,将只考虑一个相邻车道的交互情况。这并不会影响后续基于图像的信息表征。生成的图像信息主要由两部分组成:静态道路信息和动态车辆信息。其中,静态道路信息包括可跨过车道线(如车辆换道过程中穿过的车道线)和不可跨过车道线组成(如最左、右侧车道的边界线)。动态车辆信息包括中心车辆(红色)和周围车辆(黑色)。同时,考虑到并行行驶车辆的横向位移对交互影响更为明显。因此,将每类信息都整定为包含 120×90 个像素点的图片,输出到神经网络每一张图片特征大小 $[120, 90, 3]$ 。其中120为图片长度,90为图片宽度,3为相关信息的RGB特征。组成高维图像信息的过程整合了交互过程的动态、静态信息。此图像构成方法并不会造成交互信息的缺失。

3.2 交互基元提取实验及结果分析

3.2.1 初步轨迹基元提取

由图4可知,持续时间较长的基元占比较小,持

续时间为0.3 s的基元在不同周围车辆环境下占比均最高。由于持续时间极短的交互基元所包含的驾驶信息极少,同时为保证基元提取的数据总量,仅考虑持续时间大于0.3 s的交互基元。当移除掉小于0.3 s的交互片段后,剩余有效交互片段约为75%,舍弃片段约为25%。

由于SV数量不同时交互组的基数不同,为方便比对,本文选取基数较多的3组SV数量情况(SV=3, SV=4, SV=5),各随机抽样1 000例交互组,统计SV数量不同时各交互组的交互基元组成数目,结果如图5所示。由图可见,当SV数目不同时,交互基元组成数量的分布却大致相同,其中基元数目为1的情况出现概率最高,约为75.4%,而交互基元数目大于5的情况占比仅为1.3%,这说明平直道路上大部分情况中驾驶特征没有改变(稳定的驾驶过程),极少数情况下会出现驾驶特征的连续和多次改变。

3.2.2 聚类实验结果及分析

上文中首先将数据集拆分为多个交互组,然后

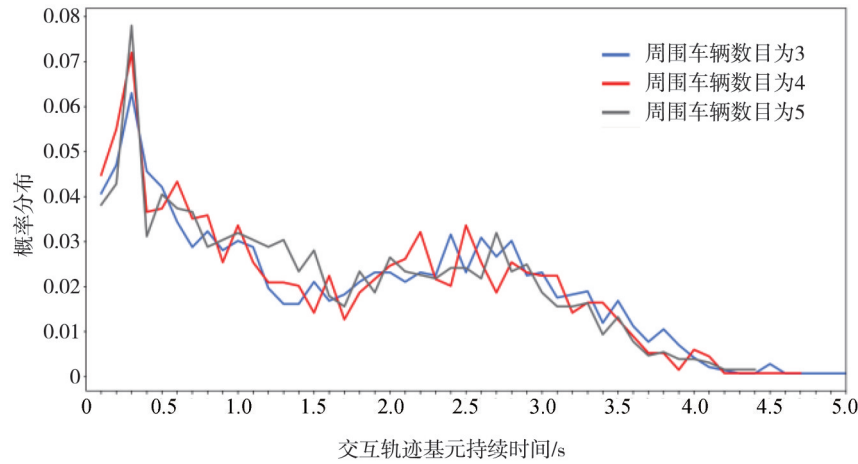


图4 不同周围车辆数目环境中初步轨迹基元持续时间分布情况(样本抽样数量为1 000组)

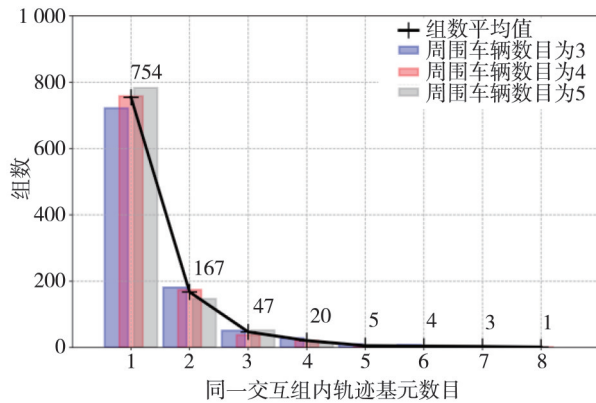


图5 交互组中交互基元组成数量的分布情况(样本抽样数量为1 000组)

利用黏性HDP-HMM方法在交互组中自主地将完整驾驶过程划分为交互基元的组合。为将具有相同特质的交互基元归为一类,本文采用聚类的手段探究驾驶特征的本质组成。

(1) 实验数据预处理

由于轨迹包含交互过程的较多信息,因此本文对交互基元的轨迹图像进行k-means聚类。为使参与聚类的交互基元之间具有相似性,将提取到的交互基元轨迹统一到同一坐标尺度下,即以HV的起始位置为坐标原点(坐标原点位于图像正中央)。由于车辆行驶于平直道路上,其在纵向车道内的行驶距离远大于横向车道之间的换道距离,因此本文将图片的长宽比设定为2:1,使轨迹具有良好的可视化效果。根据经验可知,换道行为虽然绝对行驶距离小,但是其对于周围车辆的潜在影响却明显高于直线行驶时某一车辆对周围车辆的影响,因此本文

将车辆横向的相对距离放大1.5倍以表征这种影响。由于图像聚类对RGB颜色敏感,为明确研究对象为HV,本文将HV的轨迹用红色表示,而SV的轨迹用黑色。上述预处理步骤如图6所示,首先将轨迹按交互基元拆分,针对每一张交互基元图像,进行图像归一化操作,并进行轨迹间距放缩与色彩更改。

(2) 聚类结果及分析

k-means方法为经典的无监督聚类方法,所以聚类的类别数目须预先设定。为确定类别数目并比较不同方法对聚类效果的影响,本文对已提取的交互基元分别应用不同的预训练模型,利用模型的预设参数对交互基元轨迹图像的特征进行提取,再将卷积后的结果进行k-means聚类。图7以SV为3的情况为例,展示当聚类数目从2变化至20时,各种预训练模型的轮廓系数(silhouette score)。图7中灰色折线为不使用预训练模型、提取图像全部像素信息时的k-means聚类结果。可以发现,不使用预训练模型的轮廓系数低于使用预训练模型的轮廓系数,这表示预训练模型有助于提升聚类的准确度。对比图7中的各种预训练方法,当聚类数目较少时,使用残差网络(residual network, ResNet)中ResNet50模型^[32]的聚类效果明显更好,在聚类数为4时使用ResNet50预训练模型得到的聚类结果最好,因此本文将聚类的类别数目取为4。

进一步地,根据图8中4组聚类结果所代表的驾驶特征,将单向平直道路中车辆的交互情况分为以下3类。

(1) 周围车辆位于双侧且稳定的交互行为。如

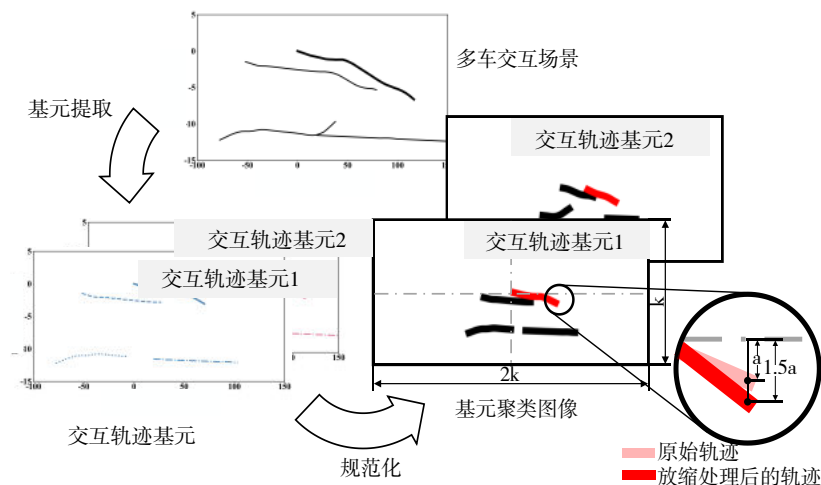


图6 聚类实验图像预处理

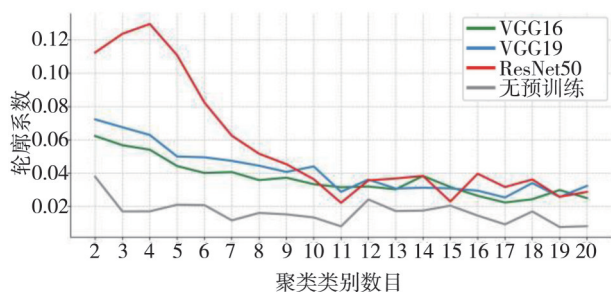


图7 不同聚类数目下各种预训练模型的轮廓系数

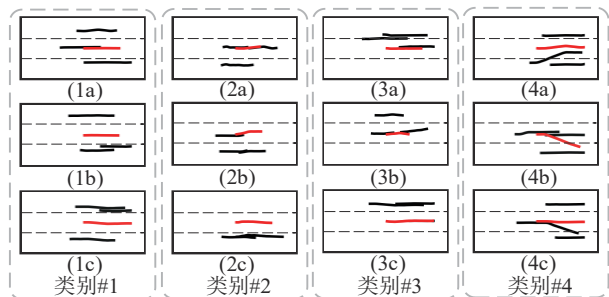


图8 聚类类别数目为4且SV=3时的聚类结果

类别#1所示,在这一类交互基元中,HV的左右车道内都有车辆,各车之间的相对位置变化不大,速度差异也不明显。此时车辆都处于稳定的交互环境中,高速公路上行驶的车辆大部分时间都表现为这种驾驶特征。

(2)周围车辆位于单侧且稳定的交互行为。如类别#2和类别#3所示,HV左侧或右侧的某一个车道存在车辆,此时车辆仍处于稳定的交互环境中。这种驾驶特征与上述第1类的区别主要在于:本交互基元结束后的下一个交互基元若存在换道等不稳

定的驾驶行为,汇入车辆仅可能来自于单一方向。SV换道时的来向的不同会导致驾驶员应对决策的不同,而第1类驾驶特征中驾驶员须同时关注两侧的车道。

(3)不稳定的交互行为。如类别#4所示,此时各个车道内都可能存在车辆,当交互环境内的某一车辆有换道意愿或处于换道过程中时,车辆间的相对位置和速度差异将会明显改变。换道等行为将会对交互环境内的所有车辆带来巨大影响,尤其是高速公路等车辆行驶速度很高的路段,换道等操作带来的风险明显高于稳定状态时的,这类驾驶特征将会影响周围驾驶员的预判和决策,从而导致各车驾驶特征的变化,此时的交互行为由于时刻存在动态变化,所以是不稳定的。

从提取到的类别#1和类别#4两类交互基元中,统计相关驾驶行为特征,如图9所示。考虑的驾驶行为特征包括基元截取时间内中心车速度、中心车加速度、中心车与前后相邻车辆间距离和此类别中交互基元(片段持续时间)。其中,交互基元1代表类别#1,交互基元2代表类别#4。根据统计结果可知,代表双侧稳定交互的基元1相对于包含换道行为的不稳定交互基元(类别#4),车辆之间有更长的车间距、更长交互过程持续时间、更高的平均车速和更小的绝对加速度。结果表明,不稳定的交互过程会导致感兴趣区域内的车辆平均速度下降并有可能出现加减速行为。同时,根据基元提取结果,不稳定交互过程还会导致较短的交互持续时间。在代表“不稳定交互”的基元中,持续时间较短(小于0.5 s, 0.5-1.0 s)的比例会高于代表“稳定”交互的基元。

此种持续时间较短的交互多存在于两个持续时间较

长的交互基元切换过程中。

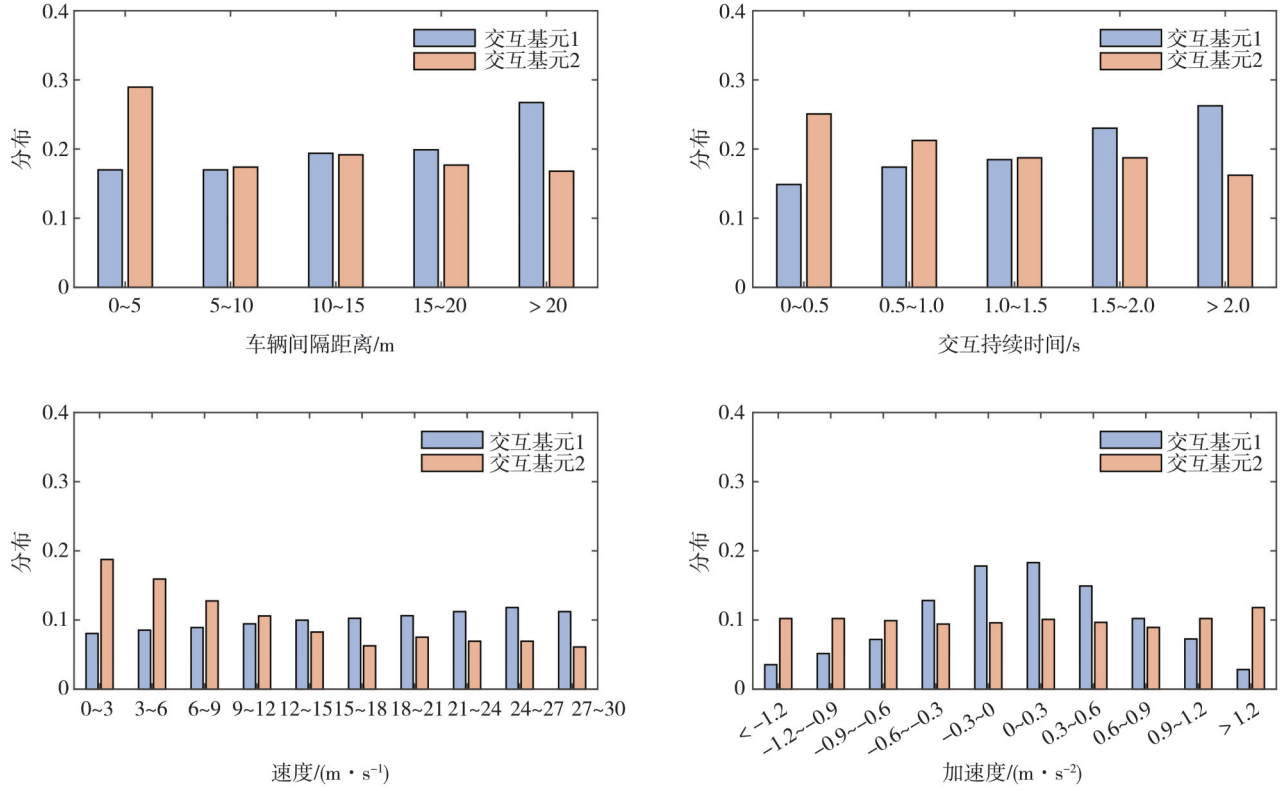


图9 交互基元类别#1(交互基元1)和类别#4(交互基元2)统计分析结果

需要指出的是,在其他SV值中, k 值的选取会产生变化,本文仅以SV=3为例展示所提出方法的基本原理和分析流程。在此案例中,可以发现,聚类方法生成的4类结果对应于真实交通场景中的:双侧车辆稳定交互、单侧车辆稳定交互(左)、单侧车辆稳定交互(右)、不稳定交互。这表明类内样本在交互行为层面具有较高的相似性。

3.2.3 讨论

现阶段针对驾驶基元的自动化提取过程与方法,存在难以精确评估各个方法差异的问题。在本研究中,为体现本文所提出方法的优势,在原本自动化基元提取框架的基础上,在实验中衔接了对换道场景中的车辆轨迹预测。采用多步轨迹预测长短时神经网络(LSTM)作为轨迹预测基础模型^[33]。通常的LSTM轨迹预测模型输入特征包括历史轨迹信息(位置、速度、加速度等)。为对比交互驾驶基元的提取效果,在LSTM轨迹预测模型输出信息特征中又加入了所属于基元类别(如类别#1,类别#2等)。用于评估轨迹预测的指标为轨迹平均预测误差(ADE)、轨迹终点预测误差(FDE),具体公式为

$$ADE = \frac{1}{T_{\text{pred}}} \sum_{t=1}^{T_{\text{pred}}} \|p_t - \hat{p}_t\| \quad (23)$$

$$FDE = \|p_{T_{\text{pred}}} - \hat{p}_{T_{\text{pred}}}\| \quad (24)$$

式中: $\|\cdot\|$ 是真实位置 p 和预测位置 $\hat{p}$ 的欧氏距离; T_{pred} 是预测步长。 ADE 与 FDE 是车辆轨迹预测尤其是交互行为预测中最为通用的评价指标。其中, ADE 反映了所预测轨迹的平均效果,是对于预测轨迹的整体评价; FDE 反映了对于车辆经过一段时间运动后终点到达位置的预测效果,这是对车辆终点到达性预测能力的专门评价。

在多步轨迹预测实验中,输入轨迹长度为2 s,预测(输出)轨迹长度为3 s。共计对比3种方法,其中LSTM-1是采用基本LSTM方法^[33],输出特征包括所有车辆的位置、速度和加速度。LSTM-2是在LSTM-1方法的基础上加入了朱冰等^[27]提出的方法所提取到的驾驶基元类别。LSTM-3是在LSTM-1方法的基础上加入了本文所提出方法得到的交互驾驶基元。实验结果如表1所示。结果表明,在轨迹预测模型的输入信息中考虑驾驶基元可以减小预测误差。基元类别的引入可以作为模态信息引导轨迹

预测过程。同时,对于朱冰等^[27]提出的基元提取方法,本文方法对于轨迹预测效果的提升更为显著,并将 *ADE* 和 *FDE* 两类预测误差分别降低 19.3% 和 14.6%。

表1 换道场景车辆轨迹预测误差结果

方法	<i>ADE</i> /m	<i>FDE</i> /m
LSTM-1	3.31	3.80
LSTM-2	2.84	3.21
LSTM-3	2.29	2.74

4 结论

提出了一种交互基元表征和提取框架,用于对连续交互行为的离散状态进行提取和分析。非参数贝叶斯方法 HDP-HMM 被用于从 NGSIM 高速公路数据集中提取交互基元,该方法可在无预设条件的前提下对连续时间信息进行分割;无监督的聚类方法被用于对提取出的交互基元进行分类,以获得驾驶行为的本质特征。实验结果表明,本文所提出的方法可以将连续的驾驶行为划分为离散的交互基元,且聚类划分结果可以与实际交互场景相对应,可用于不同交互轨迹基元中车辆之间的交互行为特性分析。同时,本文所提出方法对于复杂场景下游驾驶任务具有提升作用。在车辆多步轨迹预测任务中,相比于基线预测方法,本文所提出的交互基元提取方法在与基线预测方法融合后可以将平均预测误差和终点预测误差分别降低 19.3% 和 14.6%。

本文提取得到的交互基元还可用于场景生成和测试工作,本文主要关注场景中基元的表述和提取,未来工作将聚焦于智能网联环境下场景生成问题,以建立网联车辆交互基元“提取-生成”闭环。同时,本文所构建的模型没有显示引入安全约束,下一步工作将同样关注交互基元安全性评价。

参考文献

- [1] WAGNER C. Asymptotic solutions for a multi-anticipative car-following model [J]. *Physica A: Statistical Mechanics and Applications*, 2012, 260(1-2): 218-224.
- [2] LI Z, GONG C, LU C, et al. Transferable driver behavior learning via distribution adaption in the lane change scenario [C]. 2019 IEEE Intelligent Vehicles Symposium (IV). IEEE, 2019: 193-200.
- [3] XING Y, LV C, WANG H, et al. An ensemble deep learning approach for driver lane change intention inference [J]. *Transportation Research Part C: Emerging Technologies*, 2020, 115:

102615.

- [4] LU C, WANG H, LV C, et al. Learning driver-specific behavior for overtaking: a combined learning framework [J]. *IEEE Transactions on Vehicular Technology*, 2018, 67(8): 6788-6802.
- [5] GADEPALLY V, KRISHNAMURTHY A, OZGUNER U. A framework for estimating driver decisions near intersections [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2013, 15(2): 637-646.
- [6] HAVLAK F, CAMPBELL M. Discrete and continuous, probabilistic anticipation for autonomous robots in urban environments [J]. *IEEE Transactions on Robotics*, 2013, 30(2): 461-474.
- [7] ZHANG W, WANG W. Learning V2V interactive driving patterns at signalized intersections [J]. *Transportation Research Part C: Emerging Technologies*, 2019, 108: 151-166.
- [8] XING Y, LV C, WANG H, et al. Driver lane change intention inference for intelligent vehicles: framework, survey, and challenges [J]. *IEEE Transactions on Vehicular Technology*, 2019, 68(5): 4377-4390.
- [9] 朱冰, 蒋渊德, 赵健, 等. 基于深度强化学习的车辆跟驰控制 [J]. *中国公路学报*, 2019, 32(6): 53-60.
ZHU B, JIANG Y, ZHAO J, et al. A car-following control algorithm based on deep reinforcement learning [J]. *China Journal of Highway and Transport*, 2019, 32(6): 53-60.
- [10] 赵健, 宋东鉴, 朱冰, 等. 基于自学习和监督学习混合驱动的智能汽车跟驰控制策略 [J]. *中国公路学报*, 2022, 35(3): 55-65.
ZHAO J, SONG D, ZHU B, et al. Intelligent vehicle-following control strategy based on self-learning and supervised-learning hybrid-driven framework [J]. *China Journal of Highway and Transport*, 2022, 35(3): 55-65.
- [11] KOBER J, PETERS J. Policy search for motor primitives in robotics [C]. *Advances in Neural Information Processing Systems*, 2008: 849-856.
- [12] GOMEZ-GONZALEZ S, NEUMANN G, SCHÖLKOPF B, et al. Adaptation and robust learning of probabilistic movement primitives [J]. *IEEE Transactions on Robotics*, 2020, 36(2): 366-379.
- [13] KOERT D, PAJARINEN J, SCHOTSCHNEIDER A, et al. Learning intention aware online adaptation of movement primitives [J]. *IEEE Robotics*, 2019, 4(4): 3719-3726.
- [14] LIOUTIKOV R, NEUMANN G, MAEDA G, et al. Learning movement primitive libraries through probabilistic segmentation [J]. *The International Journal of Robotics Research*, 2017, 36(8): 879-894.
- [15] WANG B, GONG J, CHEN H. Motion primitives representation, extraction and connection for automated vehicle motion planning applications [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2019, 21(9): 3931-3945.
- [16] WANG W, ZHANG W, ZHU J, et al. Understanding V2V driving scenarios through traffic primitives [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2020, 23(1): 610-619.

- [17] ZHANG C, ZHU J, WANG W, et al. Spatiotemporal learning of multivehicle interaction patterns in lane-change scenarios [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2021, 23(7): 6446–6459.
- [18] LI Z, LU C, YI Y, et al. A hierarchical framework for interactive behaviour prediction of heterogeneous traffic participants based on graph neural network [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2021, 23(7): 9102–9114.
- [19] 田洪清, 丁峰, 郑讯佳, 等. 基于势能场虚拟力的智能网联车辆运动规划 [J]. *汽车工程*, 2021, 43(4): 518–526.
TIAN H, DING F, ZHENG J, et al. Motion planning based on virtual force of potential field for intelligent connected vehicles [J]. *Automotive Engineering*, 2021, 43(4): 518–526.
- [20] PUNZO V, BORZACCHIELLO M T, CIUFFO B. On the assessment of vehicle trajectory data accuracy and application to the Next Generation SIMulation (NGSIM) program data [J]. *Transportation Research Part C: Emerging Technologies*, 2011, 19(6): 1243–1262.
- [21] FOX E, SUDDERTH E, JORDAN M, et al. A sticky HDP-HMM with application to speaker diarization [J]. *The Annals of Applied Statistics*, 2011: 1020–1056.
- [22] DEO N, TRIVEDI M M. Convolutional social pooling for vehicle trajectory prediction [C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018: 1468–1476.
- [23] 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), 2018: 2118–2125.
- [24] ZHAN W, SUN L, WANG D, et al. INTERACTION Dataset: an INTERNATIONAL, adversarial and cooperative moTION dataset in interactive driving scenarios with semantic maps [J]. *arXiv preprint arXiv:191003088*, 2019.
- [25] ETTINGER S, CHENG S, CAINE B, et al. Large scale interactive motion forecasting for autonomous driving: the waymo open motion dataset [C]. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021: 9710–9719.
- [26] CAESAR H, BANKITI V, LANG A H, et al. nuScenes: a multi-modal dataset for autonomous driving [C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 11618–11628.
- [27] 朱冰, 孙宇航, 赵健, 等. 自动驾驶汽车测试场景基元自动提取方法 [J]. *汽车工程*, 2022, 44(11): 1647–1655.
ZHU B, SUN Y, ZHAO J, et al. Automatic extraction method for autonomous vehicle test scene primitives [J]. *Automotive Engineering*, 2022, 44(11): 1647–1655.
- [28] FORNEY G D. The viterbi algorithm [J]. *Proceedings of the IEEE*, 1973, 61(3): 268–278.
- [29] TEH Y, JORDAN M, BEAL M, et al. Sharing clusters among related groups: hierarchical dirichlet processes [J]. *Advances in Neural Information Processing Systems*, 2004, 17: 1385–1392.
- [30] WANG W, XI J, ZHAO D. Driving style analysis using primitive driving patterns with Bayesian nonparametric approaches [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2018, 20(8): 2986–2998.
- [31] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J]. *arXiv preprint arXiv:14091556*, 2014.
- [32] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 770–778.
- [33] ALTCHÉ F, DE LA FORTELLE A. An LSTM network for highway trajectory prediction [C]. 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC). *IEEE*, 2017: 353–359.