

中图法分类号: TP18; TP751 文献标识码: A 文章编号: 1006-8961(XXXX)XX-0001-15

论文引用格式: Du Shuangli, Liu Ziyang, Zhao Minghua, Hu Jing, Shi Cheng, Li Peng. Association-Enhanced Multi-Object Tracking Based on Trajectory State Update for Vehicle Turning Scenes[J/OL]. Journal of Image and Graphics, XXXX: 1-15. DOI: 10.11834/jig.260223. (都双丽, 刘梓扬, 赵明华, 胡静, 石程, 李鹏. 面向车辆转弯场景的轨迹状态更新与关联增强多目标跟踪[J/OL]. 中国图象图形学报, XXXX: 1-15. DOI: 10.11834/jig.260223.) [DOI: 10.11834/jig.260223]

面向车辆转弯场景的轨迹状态更新与关联增强多目标跟踪

都双丽¹, 刘梓扬¹, 赵明华^{1,2}, 胡静¹, 石程¹, 李鹏¹

1. 西安理工大学 计算机科学与工程学院, 陕西 西安 710048; 2. 陕西省网络计算与安全技术重点实验室, 陕西 西安 710048

摘要: 目的 针对复杂交通场景中车辆转弯引起的活动轨迹状态更新不稳定, 以及短期丢失轨迹缺乏针对性更新机制的问题, 提出一种面向车辆转弯场景的轨迹状态更新与关联增强多目标跟踪, 以缓解车辆转弯过程中易出现的漏检、身份切换和轨迹中断现象。方法 首先该方法设计快慢双时间尺度记忆模块, 针对车辆转弯时运动方向和空间位置变化快、外观信息变化相对缓慢的特点, 分别对外观慢变信息与运动快变信息进行递推建模, 并通过状态融合缓解单一状态递推中快慢信息相互干扰的问题。其次该方法设计了短期丢失轨迹状态更新模块, 结合最近更新轨迹嵌入、边界框变化趋势与丢失时长, 对短期丢失轨迹进行状态递推, 缓解了短期丢失轨迹与当前目标的状态偏移现象。结果 该方法在 FastTracker、CityFlow 和 TrafficMOT 数据集上, 在高阶跟踪准确率 (higher order tracking accuracy, HOTA)、关联准确率 (association accuracy, AssA)、身份切换次数 (identity switches, IDs) 以及轨迹碎片化次数 (fragmentations, Frag) 等跟踪相关指标上取得较优表现。结论 所提方法能够有效提升车辆转弯场景下轨迹关联的稳定性与连续性, 增强复杂交通场景中的车辆多目标跟踪性能。

关键词: 车辆多目标跟踪; 车辆转弯; 关联增强; 轨迹状态更新; 快慢双时间尺度记忆

Association-Enhanced Multi-Object Tracking Based on Trajectory State Update for Vehicle Turning Scenes

Du Shuangli¹, Liu Ziyang¹, Zhao Minghua^{1,2}, Hu Jing¹, Shi Cheng¹, Li Peng¹

1. School of Computer Science and Engineering, Xi'an University of Technology, Xi'an, Shaanxi 710048, China; 2. Shaanxi Provincial Key Laboratory of Network Computing and Security Technology, Xi'an, Shaanxi 710048, China

Abstract: Objective With the rapid development of intelligent transportation systems, vehicle multi-object tracking (MOT) has become a fundamental technique for a wide range of traffic-oriented applications, including road video surveillance, traffic event warning, vehicle behavior analysis, and intelligent traffic management. The goal of vehicle MOT is not only to detect multiple vehicles in video frames, but also to maintain their identities over time and generate reliable trajectories under dynamically changing environments. However, in complex traffic scenes, especially those involving vehicle turning, existing tracking methods often suffer from unstable cross-frame association. When a vehicle turns, its motion direction and spatial position may change abruptly within a short period, while its appearance usually evolves more slowly. This inconsistency in temporal variation makes trajectory state propagation difficult and easily leads to missed detections.

收稿日期: 2026-04-17; 修回日期: 2026-07-04

基金项目: 国家自然科学基金 (62571425); 陕西省自然科学基金项目 (2024JC-ZDXM-35); 中国高校产学研创新基金-新一代信息技术创新项目 (2024IT088); 西安理工大学博士创新基金资助项目 (BC202621)

© 中国图象图形学报版权所有

identity switches, trajectory fragmentation, and inaccurate re-association. In addition, turning maneuvers frequently cause short-term occlusion or temporary target disappearance. For these short-term lost trajectories, many existing methods simply preserve historical states or rely on delayed matching, without a dedicated update mechanism. As a result, the trajectory representation gradually deviates from the current target state, reducing the robustness of subsequent association. To address these issues, this study proposes an association-enhanced vehicle multi-object tracking method based on trajectory state update, with the aim of improving association stability and continuity in vehicle turning scenarios. **Method** The proposed method is designed from the perspective of trajectory state updating and focuses on two key challenges in turning scenes: unstable state propagation for active trajectories and insufficient state maintenance for short-term lost trajectories. First, to address the mismatch between fast-changing motion cues and slow-changing appearance cues during turning, a fast-slow dual-timescale memory module is introduced. Instead of updating all trajectory information in a single stream, the proposed module models appearance information and motion information separately according to their different temporal dynamics. Specifically, appearance-related features are treated as slow-varying states, while motion-related cues are regarded as fast-varying states. These two types of information are recursively updated in parallel and then fused to obtain a more stable and discriminative trajectory representation. This design reduces the mutual interference between heterogeneous cues during single-stream propagation and improves the robustness of state updating when vehicles undergo rapid directional changes. Second, to handle the problem of short-term lost trajectories, a short-term lost trajectory state update module is proposed. When a target is not detected for a limited period but has not yet been removed from the tracker, directly preserving its previous embedding may cause the stored state to drift away from the target's actual evolution. To alleviate this issue, the proposed module performs state propagation for short-term lost trajectories by jointly considering three types of information: the most recently updated trajectory embedding, the trend of bounding-box variation, and the duration of disappearance. By integrating these cues, the tracker can estimate a more temporally consistent trajectory state during the missing period, which is more favorable for re-association once the target reappears. In this way, the method enhances the continuity of trajectory representation and reduces the risk of association failure caused by outdated or mismatched states. Overall, the proposed framework improves tracking from the state-update level rather than relying solely on stronger detectors or post-processing association heuristics. **Result** Extensive experiments were conducted on the Fast-Tracker, CityFlow, and TrafficMOT datasets to evaluate the effectiveness of the proposed method under different traffic conditions. The experimental results demonstrate that the proposed approach achieves strong overall performance across all three datasets and shows clear advantages on association-related metrics. In particular, the method obtains the best performance on key tracking indicators such as HOTA and AssA, while also showing superior behavior on identity- and fragmentation-related measures including IDs and Frag. In addition, the method exhibits favorable trends on trajectory quality metrics such as MT and ML, indicating that it can track a larger proportion of targets reliably and reduce the number of severely lost trajectories. These improvements suggest that the proposed trajectory state update strategy is effective in mitigating association instability in turning scenarios. From a qualitative perspective, the method is better able to preserve identity consistency during abrupt maneuvers, maintain track continuity when targets temporarily disappear, and reduce re-association errors caused by state drift. Compared with methods that rely on single-stream trajectory propagation or lack dedicated handling for short-term lost tracks, the proposed approach provides a more stable and adaptive state representation for both active and temporarily missing targets, thereby improving tracking robustness in complex traffic scenes. **Conclusion** This study proposes an association-enhanced vehicle multi-object tracking method based on trajectory state update for complex traffic scenes. By introducing a fast-slow dual-timescale memory module and a short-term lost trajectory state update module, the method improves the stability and continuity of trajectory association during vehicle turning. Experimental results verify its effectiveness, and the proposed method can provide reliable trajectory support for applications such as traffic monitoring, event warning, and intelligent transportation management.

Key words: vehicle multi-object tracking; vehicle turning; association enhancement; trajectory state update; fast-slow dual-temporal memory

Enhanced Multi-Object Tracking Based on Trajectory State Update for Vehicle Turning Scenes — SCID [J/OL]. Journal of Image and Graphics. DOI:10.11834/jig.260223. (都双丽, 刘梓扬, 赵明华, 胡静, 石程, 李鹏. 面向车辆转弯场景的轨迹状态更新与关联增强多目标跟踪—SCID[J/OL]. 中国图象图形学报. DOI:10.11834/jig.260223.)

0 引言

随着汽车产业的迅速发展和机动车保有量持续增长,城市交通规模也在不断扩大。根据中华人民共和国公安部最新统计数据,2025年中国汽车产业年产销首次双超3400万辆(Guo等,2023)。车辆规模快速增长,一方面提升了公众出行效率,另一方面也增加了交通压力,引发了大量的安全事故。世界卫生组织发布道路安全资料表明:近些年全球每年因道路交通死亡的人数仍约为119万人,道路安全问题已成为了亟待解决的重要公共难题。由于道路交通管控难度加大、交通安全风险难以预防,国家层面大力倡导并推进智慧交通系统(intelligent transportation system, ITS)建设(Wen等,2024; Dong等,2025)。ITS使用视频处理与人工智能等技术,实现道路交通状态实时监控、车流动态分析与交通事件快速预警(Li等,2025)。

车辆目标跟踪是ITS的基础核心环节之一。其核心任务是在视频中对车辆进行持续定位,从而形成稳定且具有时序一致性的车辆轨迹。稳定的轨迹能够反映车辆的数量变化、空间分布等信息,可为ITS中车流统计、异常行为分析和风险预警等上层应用提供可靠的数据支撑。

从方法发展角度来看,早期基于传统方法的多目标跟踪主要依赖运动建模与数据关联完成目标轨迹维护。随着ResNet等卷积神经网络的发展,基于检测后关联的深度学习多目标跟踪在目标检测环节取得了明显进展。随着Transformer的引入,基于端到端的多目标跟踪在检测与关联统一建模方面展现出更强潜力。现有方法在通用交通场景中已经取得了较好效果,但在以车辆转弯为主的复杂交通场景中仍存在进一步提升空间(Sun等,2024)。

车辆转弯(Vehicle Turning)对跟踪性能的影响尤为明显。与直行车辆相比,转弯车辆在图像平面

中的运动状态具有更强的非线性特征。车辆进入转弯过程后,其速度方向会持续变化,加速度方向也会随轨迹曲率发生明显改变,导致目标中心点位移方向、边界框位置以及尺度变化不再满足简单的线性递推假设。与此同时,车辆转弯通常发生在路口、匝道或汇流区域,目标之间的相对位置关系变化更加频繁,原本前后分布的车辆可能在图像中转变为交叉、并行或局部重叠关系,从而引起遮挡模式的变化。上述因素会使当前帧观测与历史轨迹状态之间出现较大偏移,进而造成检测框滞后、轨迹状态更新不稳定以及跨帧关联困难。

进一步来看,车辆转弯不仅会引起运动状态的快速变化,还会造成外观信息与运动信息变化速率不一致。车辆颜色、纹理等外观特征通常变化相对缓慢,而运动方向、空间位置和边界框几何状态可能在短时间内快速变化。若现有方法仍采用单一状态流对所有信息进行统一递推,快变运动信息与慢变外观信息容易在更新过程中相互干扰,使轨迹嵌入难以及时适应转弯后的真实状态。此外,车辆转弯过程中产生的局部遮挡和短时漏检还会使部分轨迹进入短期丢失状态。若此类轨迹在失轨期间仅保留旧状态而缺少针对性更新,其轨迹表示会随着时间推移逐渐偏离目标当前的真实位置和运动趋势,最终导致身份切换、轨迹中断和重关联失败等问题。因此,本研究将围绕车辆转弯条件下的轨迹状态更新优化方向展开研究。

基于上述分析,开展复杂交通场景下的车辆多目标跟踪研究具有重要意义。其研究提升复杂交通环境下车辆跟踪能力,并为智能交通系统提供可靠的数据支撑。

1 相关工作

1.1 基于检测后关联的多目标跟踪

基于检测后关联的多目标跟踪方法(tracking-by-detection, TBD)是多目标跟踪研究的经典范式。这类方法通常先使用检测器对视频帧中的目标进行检测,再结合运动信息、外观信息和匹配策略实现跨帧的身份关联。

早期研究重心重点在于对深度外观特征的处理,Zhang等人(2022)提出ByteTrack方法,通过联合高、低置信度检测结果进行分阶段关联,缓解了局部

遮挡造成的目标切换。Cao 等人(2023)提出观察中心排序方法 OC-SORT,重新设计基于观测的运动关联方式,提高了非线性运动场景中跟踪的鲁棒性。

近些年 TBD 多目标跟踪的研究重心转向细粒度轨迹管理。Qin 等人(2024)提出了通用化多目标跟踪方法 GeneralTrack,该方法在跟踪过程中对点到实例级关系进行建模,提升了不同场景下的泛化能力。Du 等人(2024)提出了分层 IoU 的多目标跟踪方法,利用轨迹间隔先验实现统一的分层跟踪。

总体来看 TBD 多目标跟踪在精度和实时性之间取得了较好的平衡。但这类方法高度依赖检测质量,对于车辆转弯带来的非线性运动变化,多数 TBD 方法仍从运动关联或轨迹管理角度进行补救,导致误差在关联阶段逐渐累积增多。

1.2 基于端到端的多目标跟踪

与 TBD 多目标跟踪不同,基于端到端多目标跟踪方法是将目标检测、状态更新与身份关联放在统一框架中进行联合优化。这种方法关注检测与关联协同建模能力。

随着 Transformer 的发展,基于注意力机制的端到端多目标跟踪开始出现。Sun 等人(2021)提出多目标 Transformer 跟踪方法 TransTrack,通过目标查询与跟踪查询同时完成新生目标检测和已有轨迹更新。Zeng 等人(2022)提出端到端多目标 Transformer 跟踪方法 MOTR,引入 track query,进一步提升了跟踪的跨帧时序建模能力。

在端到端框架成熟之后,研究者开始进一步关注长时序信息建模以及检测与关联协同优化问题。Gao 等人(2023)提出长时记忆增强 Transformer 跟踪方法 MeMOTR,在端到端框架中引入长时记忆注入与记忆注意力机制,并通过检测解码器缓解检测查询与跟踪查询之间的语义差异。Yan 等人(2025)提出 CO-MOT,通过改进标签分配与查询训练策略,缓解检测查询与跟踪查询之间的竞争关系。

尽管 MeMOTR 已经利用长时序记忆建模在身份关联与轨迹状态更新中展现出潜力,但对于车辆转弯情况:车辆的颜色、纹理等外观状态信息变化较慢,而速度方向、位移方向、空间位置等运动状态信息变化较快,其仍采用统一状态流进行更新。这会导致两类更新速率需求不一致的信息在更新过程中受到干扰,从而造成轨迹丢失以及身份切换。同时车辆转弯不仅会影响活动轨迹的持续更新,还

容易造成目标轨迹短期丢失。而 MeMOTR 包括现有大多数方法,对于转弯过程中短期丢失轨迹缺少针对性的状态更新机制,从而导致其状态表示会逐渐严重偏离目标的真实位置与运动趋势。这两个问题也制约了车辆转弯是多目标跟踪的轨迹状态更新与身份关联性能,从而影响最终跟踪的稳定性。

为解决车辆转弯产生的问题,本文以 MeMOTR 为基线,提出一种面向车辆转弯场景的轨迹状态更新与关联增强多目标跟踪。针对活动轨迹状态更新不稳定问题,本方法设计快慢双时间尺度记忆模块(fast-slow dual-temporal memory module, FSTM),从当前帧构造外观观测和运动观测,再通过快慢两条独立状态流分别递推,最后统一融合生成下一时刻活动轨迹嵌入。针对短期丢失轨迹缺乏状态更新机制的问题,提出短期丢失轨迹状态更新模块(short-term missing track state update module, SMTU),利用轨迹最近有效更新信息、边界框变化趋势以及丢失时长等线索,对短期丢失轨迹嵌入进行递推优化。通过上述两个模块可以提升轨迹状态更新质量,增强车辆转弯场景下多目标跟踪的稳定性。

2 方法

2.1 概述

本节详细介绍提出的面向车辆转弯场景的轨迹状态更新与关联增强多目标跟踪。本研究基于 MeMOTR,针对车辆转弯引起的活动轨迹状态更新不稳定问题,提出快慢双时间尺度记忆模块 FSTM。针对短期丢失轨迹缺乏针对性更新机制的问题,提出短期丢失轨迹状态更新模块 SMTU。其网络结构图如图 1 所示。

对于整体模型而言:首先利用 ResNet-50 与 Transformer Encoder 对当前帧进行特征提取,得到共享特征。在此基础上, Detection Decoder 以共享特征为输入,生成当前帧检测嵌入。Joint Decoder 接收上一帧的生成的跟踪嵌入与当前检测嵌入,最终输出当前全部保留轨迹嵌入。将全部保留估计嵌入输入预测头得到置信度与边界框。最后轨迹管理模块接收当前全部保留轨迹嵌入、置信度与边界框,进行新生目标初始化、活动轨迹更新以及短期丢失轨迹维护等处理,一方面生成用于可视化展示的目标框与轨迹 ID,另一方面将活动轨迹和短期丢失轨迹分别

送入 FSTM 模块与 SMTU 模块生成活动轨迹嵌入和短期丢失轨迹嵌入。最终这两种嵌入拼接得到下一

时刻的跟踪嵌入, 将其作为下一时刻 Joint Decoder 的输入完成轨迹状态更新递推。

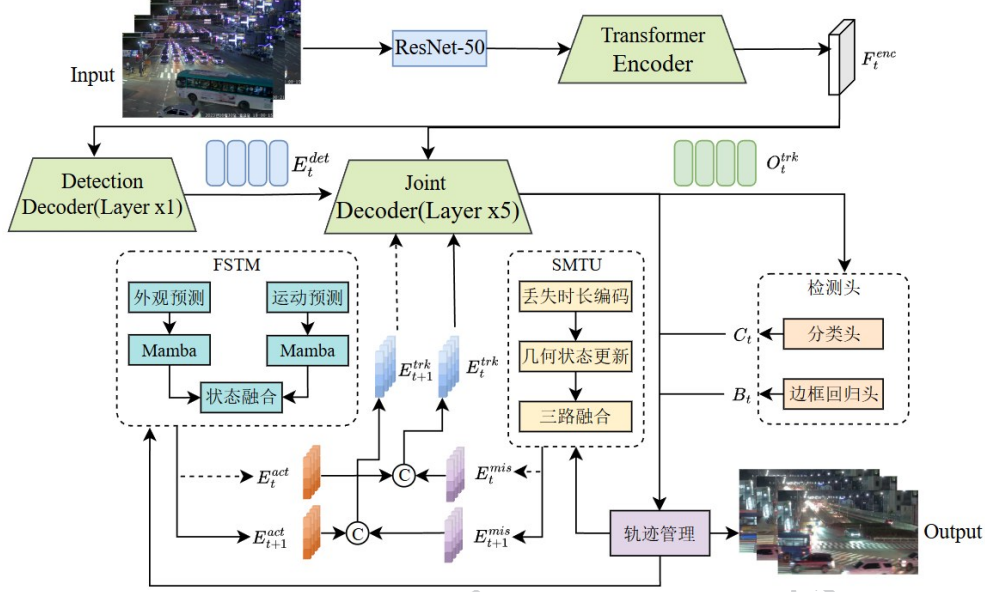


图1 整体网络结构图

Fig. 1 Overall Network Architecture Diagram

2.2 快慢双时间尺度记忆模块

车辆转弯场景下, 目标的运动方向、空间位置往往会在短时间内快速变化, 这种变化会首先影响活

动轨迹的状态递推过程。本研究引入快慢双时间尺度记忆模块 FSTM, 详细结构如图2所示。

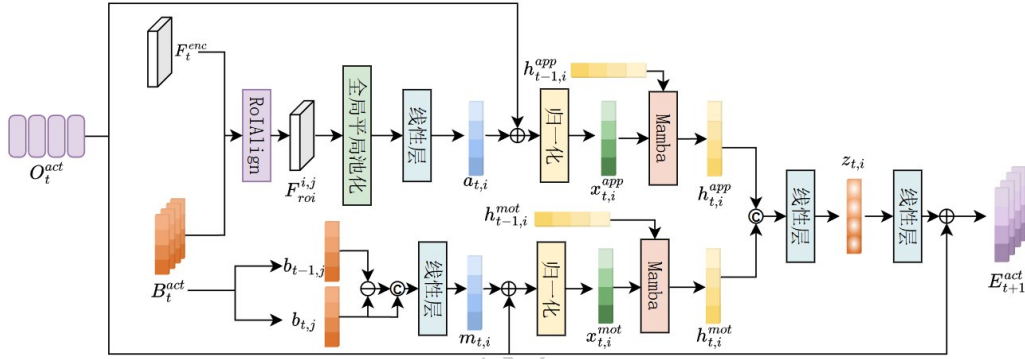


图2 快慢双时间尺度记忆模块详细结构图

Fig. 2 Detailed Structure Diagram of FSTM

Joint Decoder对 E_t^{det} 与 E_t^{trk} 处理后, 得到当前全部保留轨迹嵌入集合:

$$O_t^{trk} = \{o_{t,i}^{trk}\}_{i=1}^{N_t^{all}}, o_{t,i}^{trk} \in \mathbb{R}^D \quad (1)$$

式中 N_t^{all} 表示当前轨迹管理完成后的全部保留轨迹数, 其包含当前活动轨迹以及短期丢失但尚未删除的轨迹。因为在 MeMOTR 中里, 如果跟踪目标短期丢失不会直接删除轨迹嵌入, 而是将其标记为 inactive trajectory, 之后若超过 T_{miss} 帧才彻底移除。

依据推理阶段的轨迹管理规则与预测头得到的分类置信度 C_t , 将完成状态更新并成功匹配到当前观测的轨迹划分进入活动轨迹集合 A_t 。其余划分进短期丢失轨迹集合 Λ_t 。

记活动轨迹嵌入子集为:

$$O_t^{act} = \{o_{t,i}^{trk} | i \in A_t\} \quad (2)$$

同时经过预测头得到活动轨迹对应边界框:

$$B_t^{act} = \{b_{t,i} | i \in A_t\}, b_{t,i} \in \mathbb{R}^4 \quad (3)$$

由主干网络提取出来的共享特征为 $F_t^{enc} \in \mathbb{R}^{H \times W \times D}$ 。对于当前时刻的活动轨迹 $i \in A_t$, 若其在上一时刻存在历史状态, 则取上一时刻对应的快、慢记忆 $h_{t-1,i}^{app} \in \mathbb{R}^D$ 、 $h_{t-1,i}^{mot} \in \mathbb{R}^D$ 。

对任一活动轨迹 $i \in A_t$, 结合当前帧融合特征 $\tilde{F}_t$ 与当前帧边界框 $b_{t,i}$ 执行 RoIAlign 得到目标区域特征 $F_{roi}^{(t,i)} \in \mathbb{R}^{h_r \times w_r \times D}$, 其中, h_r 与 w_r 为固定的 RoI 输出空间尺寸。

然后对 $F_{roi}^{(t,i)}$ 进行全局平均池化 GAP, 并通过线性映射 Linear 得到外观描述向量 $a_{t,i} \in \mathbb{R}^D$ 。

对于运动分支, 建模当前时刻与上一时刻同一轨迹边界框之间的几何变化。设第 i 条活动轨迹轨迹在时刻 $t-1$ 的边界框为 $b_{t-1,i} \in \mathbb{R}^4$, 则当边界框差分为: $\Delta b_{t,i} = b_{t,i} - b_{t-1,i}$ 。对于新生目标, 由于不存在在上一时刻对应边界框, 统一令 $\Delta b_{t,i} = 0$ 。

将当前框 $b_{t,i}$ 与差分向量 $\Delta b_{t,i}$ 拼接, 通过线性映射 Linear 投影到 D 维空间, 得到运动观测投影 $m_{t,i} \in \mathbb{R}^D$ 。再将观测投影 $m_{t,i}$ 与当前活动轨迹嵌入 $o_{t,i}^{rk}$ 相加并归一化 LN, 得到运动观测 $x_{t,i}^{mot} \in \mathbb{R}^D$ 。

设第 $t-1$ 时刻在第 i 条轨迹输出的慢记忆与快记忆分别为:

$$h_{t-1,i}^{app} \in \mathbb{R}^D, h_{t-1,i}^{mot} \in \mathbb{R}^D \quad (4)$$

得到当前外观观测与运动观测后, 由于外观信息与运动信息具有不同的时间变化规律, 进一步对两类观测进行递推更新, 生成慢记忆 $h_{t,i}^{app}$ 与快记忆 $h_{t,i}^{mot}$ 。

由于 Mamba 基于选择性状态空间模型, 能够在递推过程中根据当前输入自适应调节历史状态的保留、遗忘与更新, 既可以保持轨迹身份相关的长时信息, 又能够对当前帧中快速变化的运动观测作出响应。对于车辆转弯场景而言, 目标外观纹理通常变化相对平缓, 而运动方向、空间位置和边界框形态变化更快, 二者在时间尺度上存在明显差异。若采用单一状态流统一递推, 容易造成快变运动信息与慢变外观信息相互干扰。基于此, 本研究采用两个参数相互独立、结构形式一致的 Mamba 递推单元, 分别对外观慢变状态和运动快变状态进行更新, 从而提升活动轨迹在转弯过程中的状态递推稳定性。

为明确 Mamba 递推单元的输入输出维度, 本文将单条活动轨迹的外观观测、运动观测、慢记忆和快记忆均表示为 D 维向量。对于第 i 条轨迹, 外观分支 Mamba 的输入为当前外观观测 $x_{t,i}^{app} \in \mathbb{R}^D$ 与上一时刻

慢记忆 $h_{t-1,i}^{app} \in \mathbb{R}^D$, 输出为更新后的慢记忆 $h_{t,i}^{app} \in \mathbb{R}^D$ 。运动分支 Mamba 的输入为当前运动观测 $x_{t,i}^{mot} \in \mathbb{R}^D$ 与上一时刻快记忆 $h_{t-1,i}^{mot} \in \mathbb{R}^D$, 输出为更新后的快记忆 $h_{t,i}^{mot} \in \mathbb{R}^D$ 。因此, Mamba 递推单元在外部输入输出上保持通道维度不变。若对当前帧全部活动轨迹并行计算, 则外观分支和运动分支的输入可分别组织为 $|A_t| \times D$ 的轨迹状态矩阵, 输出仍为 $|A_t| \times D$ 的更新记忆矩阵。

在时刻 t 的状态更新过程为:

$$\begin{aligned} h_{t,i}^{app} &= \text{Mamba}_{app}(x_{t,i}^{app}, h_{t-1,i}^{app}) \\ h_{t,i}^{mot} &= \text{Mamba}_{mot}(x_{t,i}^{mot}, h_{t-1,i}^{mot}) \end{aligned} \quad (5)$$

式中 Mamba_{app} 与 Mamba_{mot} 采用相同的网络形式, 但参数彼此独立。

对新生目标, 由于不存在可用的上一时刻历史状态, 直接以当前观测初始化对应的快记忆与慢记忆。

得到两路记忆后, 将快记忆与慢记忆拼接, 通过线性层完成状态融合, 得到双时间融合态 $z_{t,i}$ 。将 $z_{t,i}$ 通过线性层 Linear 进行映射, 然后与当前活动轨迹嵌入 $o_{t,i}^{rk}$ 做残差融合, 最后经过 LN, 得到中间表示:

$$e_{t+1,i}^{act} = \text{LN}(W_u z_{t,i} + b_u + o_{t,i}^{rk}) \in \mathbb{R}^D \quad (6)$$

对所有活动轨迹执行上述过程, 得到下一时刻的活动轨迹嵌入集合 E_{t+1}^{act} 。

需要说明的是, 第 1 帧中尚不存在历史轨迹, 因此此时并不存在上一时刻的活动轨迹嵌入, 也不存在短期丢失轨迹嵌入。因此仅依赖检测分支输出当前帧检测结果, 再初始化为首帧活动轨迹, 并以其对应的当前输出表示直接构造活动轨迹集合。

2.3 短期丢失轨迹状态更新模块

现有方法缺少专门的状态更新机制。若仅保留其旧状态而不进行后续递推更新, 则随着时间推移, 该轨迹嵌入会逐渐偏离目标在当前时刻的真实状态。为此本研究引入短期丢失轨迹状态更新模块 SMTU, 其结构如图 3 所示。

读取缓存中的最近两次成功更新的更新帧号、边界框和最近一次成功更新后保留下来的轨迹嵌入。为避免过早历史状态对当前轨迹递推产生干扰, 本文对“最近两次成功更新”设置时间窗口阈值 T_{win} 。具体而言, 只有当两次成功更新帧之间的间隔不超过 T_{win} 时, 才认为其边界框变化能够有效反映短期运动趋势。本文实验中设置 $T_{win} = 10$ 帧, 并要

求该阈值不大于短期丢失轨迹的最大保留帧数 T_{miss} 。若不满足该条件, 则不使用边界框变化趋势进行外推, 而采用零变化量进行状态补偿。

设该轨迹最近两次成功更新所在帧分别为 s_i 和 p_i 。对应地, 最近两次成功更新的边界框分别记为

$b_{s,i}$ 与 $b_{p,i}$ 。其中, $p_i < s_i < t$, 且满足 $s_i - p_i \leq T_{win}$ 。当 $s_i - p_i > T_{win}$ 时, 说明两次成功更新间隔过长, 其几何变化已难以稳定反映当前短期运动趋势, 因此不再使用该变化量进行趋势外推。

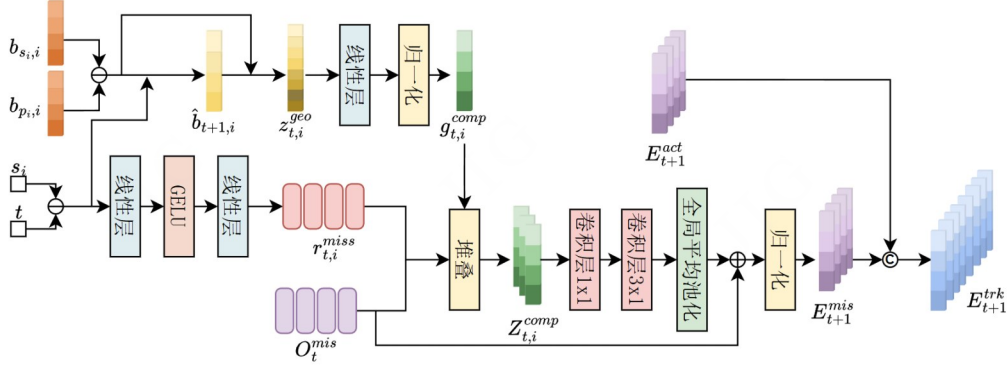


图3 短期丢失轨迹状态更新模块详细结构图

Fig. 3 Detailed Structure Diagram of SMTU

第 i 条短期丢失的轨迹嵌入集合记为:

$$O_t^{mis} = \{o_{s_i+1,i}^{trk} | i \in \Lambda_t\} \quad (7)$$

若轨迹在进入短期丢失状态前仅有一次成功更新记录, 则默认令 $p_i = s_i$, 从而有 $b_{p,i} = b_{s,i}$ 。同样地, 若最近两次成功更新的帧间隔超过 T_{win} , 也令 $b_{p,i} = b_{s,i}$, 使边界框变化量退化为零。这样可以避免利用时间跨度过大的历史边界框对当前短期丢失轨迹进行不可靠外推。

第 i 条轨迹在当前时刻的丢失时长为 $\delta_i = t - s_i$ 。将丢失时长 δ_i 输入前馈网络中, 得到丢失时长编码 $r_{t,i}^{miss}$ 。

短期丢失的目标, 虽然当前帧中缺少稳定匹配结果, 但该轨迹在最近两次成功更新时的边界框仍然保留了转弯前后的几何变化信息。在时间窗口 T_{win} 内, 这些局部几何变化能够较好反映目标短期内的位置偏移和尺度变化趋势; 而超过该窗口后, 历史变化与当前运动状态之间的相关性会下降, 因此不再作为可靠趋势使用。这些信息能够反映目标在转弯过程中的位置偏移和尺度变化趋势, 因此可作为短期丢失轨迹状态更新的重要依据。

根据最近两次成功更新的边界框构造按帧间隔归一化的边界框变化量:

$$\Delta b_i = \begin{cases} \frac{b_{s,i} - b_{p,i}}{s_i - p_i}, & 0 < s_i - p_i \leq T_{win} \\ 0, & s_i - p_i = 0 \cup s_i - p_i > T_{win} \end{cases} \quad (8)$$

采用线性预测, 把边界框变化量近似成恒定的变化, 按丢失时长, 将趋势延续到第 $t+1$ 帧, 生成下一时刻的几何更新状态:

$$\hat{b}_{t+1,i} = b_{s,i} + (\delta_i + 1) \cdot \Delta b_i \quad (9)$$

将下一时刻几何更新状态与边界框变化量拼接, 并通过线性映射对该几何向量进行通道投影, 再经 LayerNorm 归一化, 得到几何更新特征 $g_{t,i}^{comp}$ 。

得到短期丢失轨迹最近一次成功更新后保留的轨迹嵌入 $o_{s_i+1,i}^{trk}$ 、丢失时长编码 $r_{t,i}^{miss}$ 以及几何更新特征 $g_{t,i}^{comp}$ 后, 将三路输入堆叠成长度为 3 的短序列 $Z_{t,i}^{comp}$ 。

随后通过一个 1×1 卷积进行通道对齐, 在通过一个 3×1 间距对三路信息进行融合, 最后经过全局平均池化得到更新残差 $\Delta e_{t+1,i}^{comp}$ 。

最后, 以残差归一化生成第 $t+1$ 时刻的短期丢失轨迹嵌入:

$$e_{t+1,i}^{mis} = \text{LN}(o_{s_i+1,i}^{trk} + \Delta e_{t+1,i}^{comp}) \in \mathbb{R}^D \quad (10)$$

通过递推, 将下一时刻短期丢失轨迹嵌入集合记为 E_{t+1}^{mis} 。

将本节得到的丢失轨迹嵌入与上一节得到的活动轨迹嵌入拼接, 得到 $t+1$ 帧完整的轨迹嵌入。将

该轨迹嵌入送入 Joint Decoder 执行联合解码。

3 实验结果与分析

3.1 数据集

在 FastTracker、CityFlow 与 TrafficMOT 数据集对本研究的方法进行评估。FastTracker 包含约 80 万条标注检测结果,包含多车道路口和汇入道路等场景,本文根据视频片段中的车辆运动方向变化、路口区域位置以及轨迹弯曲程度进行人工抽样统计,约有 45.8% 的样本片段包含较明显的车辆转弯行为,

适合观察复杂交通流下车辆转弯的跟踪表现。City-Flow 包含 3.25 小时同步高清视频,其中路口场景占比较高,车辆转弯、汇流和轨迹交叉现象更加明显,经同样方式统计,其转弯相关样本片段占比约为 58.2%,是三个数据集中转弯与交叉行驶特征最突出的数据集,因此适合验证本研究在转弯场景中的关联稳定性。TrafficMOT 包含 2,102 段短视频,具有较高的目标密度,包含大量车辆转弯场景,其中约 51.5% 的样本片段存在车辆转弯、并线或汇入等非直线运动行为,能够用于评估高密度交通条件下轨迹状态更新与短期丢失轨迹重关联能力。

表 1 各方法在 FastTracker 数据集定量比较结果

Table 1 Quantitative Comparison of Different Methods on the FastTracker Dataset

方法	HOTA	AssA	IDF1	IDs	Frag	MT	ML	DetA
StrongSORT++	40.5	39.5	47.4	512	754	38.5	23.1	43.1
OC-SORT	40.7	39.7	47.0	519	761	38.2	23.5	43.5
MeMOTR	41.6	40.5	48.1	489	721	39.7	22.3	44.9
GeneralTrack	43.7	42.2	49.9	438	655	42.6	19.6	47.1
MOTIP	44.2	42.8	50.8	421	633	43.5	19.4	46.8
本方法	44.3	42.8	50.6	413	621	43.4	19.0	47.0

注:加粗表示最优数据

表 2 各方法在 CityFlow 数据集定量比较结果

Table 2 Quantitative Comparison of Different Methods on the CityFlow Dataset

方法	HOTA	AssA	IDF1	IDs	Frag	MT	ML	DetA
StrongSORT++	42.9	40.9	49.3	571	838	40.2	21.5	45.2
OC-SORT	43.1	41.1	49.0	564	827	40.0	21.4	45.9
MeMOTR	44.0	41.9	50.1	524	781	41.6	20.1	46.9
GeneralTrack	45.4	43.4	51.9	472	707	44.3	18.2	48.2
MOTIP	45.8	44.0	52.8	457	688	45.6	17.7	47.8
本方法	45.9	44.2	52.7	451	677	45.4	17.3	48.0

注:加粗表示最优数据

3.2 评价指标

本研究主要围绕车辆转弯条件下的关联稳定性增强展开,因此对比实验与消融实验统一采用 HOTA、AssA、IDF1、IDs、Frag、MT、ML 和 DetA 作为评价指标。这些指标能够从整体性能、关联质量、身份保持、轨迹完整性以及检测质量等多个角度对方法进行综合评估,适合用于分析车辆转弯场景下轨迹状态更新是否稳定。

其中 AssA 反应目标在跨帧过程中有没有被稳定地关联成同一条轨迹,是关联的综合性指标。IDs 表示身份切换次数。Frag 表示轨迹碎片化次数。MT 表示大部分时间被成功跟踪的目标比例。ML 表示大部分时间处于丢失状态的目标比例。IDF1 用于反映轨迹身份保持一致性。HOTA 用于综合衡量检测与关联的整体性能。DetA 则作为辅助指标。对于本文研究的车辆转弯问题而言,AssA、IDs 和

表3 各方法在TrafficMOT数据集定量比较结果

Table 3 Quantitative Comparison of Different Methods on the TrafficMOT Dataset

方法	HOTA	AssA	IDF1	IDs	Frag	MT	ML	DetA
StrongSORT++	41.5	39.9	48.4	462	697	39.6	22.0	44.2
OC-SORT	41.4	39.7	48.0	470	705	39.2	22.2	44.7
MeMOTR	42.6	40.8	48.7	435	665	40.6	21.2	45.6
GeneralTrack	43.8	42.3	50.2	399	617	42.7	19.4	47.0
MOTIP	44.1	42.9	51.0	384	596	43.6	19.0	46.6
本方法	44.2	42.8	50.8	377	588	43.8	18.5	46.8

注:加粗表示最优数据

Frag更能直接反映轨迹关联稳定性,MT和ML则能够从轨迹完整性角度体现目标是否能够被持续跟踪。

3.3 实验设置

本章方法以ResNet-50与Transformer Encoder作为特征提取骨干网络,模型初始化采用在COCO数据集上预训练的官方DAB-Deformable-DETR权重。训练阶段参考基线MeMOTR的实现细节,采用随机缩放与随机裁剪等基础数据增强方式,并将输入图像短边缩放至800,长边最大限制为1536。优化器采用AdamW,初始学习率设为 2.0×10^{-4} ,并在训练后期进行衰减。视频clip长度采用课程式递增策略,由2帧逐步增加至5帧,以增强模型对时序信息的建模能力。该设置能够在保证与基线方法训练配置基本一致的前提下,使模型逐步适应由短时相邻帧到较长时序片段的轨迹状态递推过程,从而更公平地验证本文所提出模块的有效性。

3.4 对比实验结果分析

1) 定量分析

表1、表2与表3分别表示本方法、基线方法与其他多目标跟踪方法,在FastTracker、CityFlow与TrafficMOT数据集上的性能对比结果。表中加粗表示最优数据。

车辆转弯首先会影响活动轨迹的状态递推过程,进而引发轨迹丢失、ID跳变和检测框滞后等现象,直接导致AssA下降、MT下降和ML上升、IDs增加和Frag升高,并进一步影响IDF1和HOTA等指标。对于短期丢失但尚未删除的轨迹,若缺少针对性的状态更新机制,其轨迹嵌入会逐渐偏离目标当前真实状态,会继续带来AssA下降、IDs增加和Frag升高外,还会使MT下降、ML升高。因此,在分析车

辆转弯场景下的跟踪性能时,不能只关注检测相关指标,还需要重点观察身份保持、轨迹连续性和短期失轨恢复能力。

总体来看,本方法在HOTA、AssA、IDs、Frag、MT和ML等指标上表现更优,说明本方法能够有效增强复杂交通场景下的轨迹关联稳定性。其中,IDs和Frag与车辆转弯场景中的身份保持和轨迹连续性关系更为直接。表1至表3中,本方法在三个数据集上均取得最低IDs和Frag,说明其能够有效减少车辆转弯过程中的身份切换和轨迹碎片化现象。这表明本文方法不仅能够提升整体跟踪性能,而且能够更有针对性地缓解车辆转弯时容易出现的身份跳变和轨迹中断问题。

StrongSORT++和OC-SORT等方法的性能较差,说明利用检测框、外观建模或观测关联策略能够改善轨迹连续性,但依旧难以适应车辆转弯这一类问题。StrongSORT++主要依赖检测结果、运动预测和外观相似度进行关联,其优势在于对常规连续运动目标具有较好的轨迹维护能力。但在车辆转弯场景中,车辆运动方向和边界框尺度变化较快,且同类车辆外观相似度较高,容易削弱外观匹配与运动预测的一致性,从而造成身份混淆或短时失配。OC-SORT强调基于观测结果改进运动关联,对检测框连续可靠的场景较为有效;但当车辆转弯伴随遮挡、检测间断或边界框快速偏移时,观测序列本身会变得不稳定,导致其基于观测的关联优势受到限制。

MeMOTR由于引入了记忆建模,整体表现进一步提升,但其递推过程仍未区分外观缓变与运动快变信息,因此在复杂路口场景中仍存在进一步优化空间。从表1至表3可以看出,MeMOTR相比StrongSORT++和OC-SORT在HOTA、AssA和IDF1等指标

上均有提升,说明长时记忆建模有助于增强轨迹表示的稳定性。然而,在车辆大角度转弯或汇入并线时,目标几何状态和运动方向会发生快速变化,较平滑的历史记忆容易与当前观测之间产生滞后偏差,因此仍可能出现 IDs 和 Frag 偏高的问题。

进一步分析 IDs 和 Frag 的误差来源可以发现,身份切换更容易出现在多车同时转弯、汇入并线以及大角度转弯等场景中。此时车辆运动方向和空间位置变化较快,轨迹在图像平面中容易交叉或接近,导致相邻车辆身份混淆。轨迹碎片化则更多出现在排队转弯、车辆汇流和短时遮挡场景中,目标被遮挡或检测结果短暂中断后,若短期丢失轨迹状态不能及时更新,就容易造成轨迹中断和重新连接。CityFlow 中路口与汇流场景较多, TrafficMOT 中目标密度较高,因此上述问题更加明显。本方法通过快慢双时间尺度记忆模块提升活动轨迹状态递推稳定性,并通过短期丢失轨迹状态更新模块增强失轨目标的重关联能力,因此在 IDs 和 Frag 上取得更稳定的结果。

GeneralTrack 和 MOTIP 是整体性能最强的对比方法,说明近年的方法在表示学习和身份建模方面已经取得了较好效果。但本方法在更能体现轨迹连续性与再关联能力的指标上仍占据一定优势。例如在 CityFlow 这类复杂路口与车辆转弯场景更突出的数据集上,本方法在 AssA、IDs 和 Frag 上的指标较为突出,说明其对车辆转弯这一复杂运动变化和短时失轨问题的针对性更强。在 FastTracker 和 TrafficMOT 上,本方法同样在 IDs、Frag 和 ML 等指标上保持较优结果,说明本方法对车辆转弯场景具有较好适应性。整体来看,本方法的性能提升主要来源于轨迹状态递推能力的改善。

2) 定性分析

图 4、图 5 与图 6 分别展示了不同方法在 FastTracker、CityFlow、TrafficMOT 数据集上的跟踪可视化对比结果。对于三张图而言,第一列、第二列和第三列分别表示数据集对应第 t 帧、第 $t+1$ 帧和第 $t+2$ 帧。第一行为本方法的连续三帧跟踪结果,第二行为 MOTIP 的连续三帧跟踪结果,第三行为基线 MeMOTR 的连续三帧跟踪结果。

从图 4 中可以看出,在 FastTracker 数据集中,本方法(第一行)在连续三帧中始终保持了较稳定的身份关联结果,右侧转弯车辆的轨迹能够连续保持。

MOTIP(第二行)在第 $t+1$ 帧至第 $t+2$ 帧之间,右侧 ID1 车辆在转弯过程中出现了短时失配,表现为轨迹短暂中断,但在第 $t+2$ 帧又恢复了对该目标的正确关联,基线 MeMOTR(第三行)的同一车辆在第 $t+1$ 帧至第 $t+2$ 帧之间同样发生失配,且在第 $t+2$ 帧未能恢复原有身份,而是被重新分配为新的轨迹 ID,表现出更明显的身份切换现象。

从图 5 中可以看出,在 CityFlow 数据集中,本方法(第一行)在连续三帧内始终保持了较稳定的身份关联结果,两辆转弯车辆的轨迹能够随运动方向变化持续更新,边界框也能较好跟随车辆真实位置变化。MOTIP(第二行)在第 t 帧至第 $t+1$ 帧之间,左侧 ID1 车辆在转弯过程中出现了短时失配,表现为原有轨迹短暂中断,但在第 $t+2$ 帧又恢复了对该目标的正确关联,同时 MOTIP 对两辆转弯车辆的边界框都出现了较为明显的滞后偏移现象,即边界框未能及时跟随车辆转向后的真实位置进行更新。基线 MeMOTR(第三行)从第 t 帧到第 $t+2$ 帧,左侧 ID1 车辆在转弯过程中持续处于轨迹丢失状态,表现为原有轨迹中断后未能及时恢复。与此同时,两辆转弯车辆的边界框均出现了比 MOTIP 更严重的滞后偏移。

从图 6 中可以看出,在 TrafficMOT 数据集中,本方法(第一行)在连续三帧内始终保持了较稳定的身份关联结果,转弯车辆的轨迹能够随运动方向变化持续更新,边界框也能较好跟随车辆真实位置变化。MOTIP(第二行)在第 t 帧至第 $t+1$ 帧之间, ID4 车辆在转弯过程中出现了轨迹丢失,同时 MOTIP 对转弯车辆的边界框出现了较为明显的滞后偏移现象,即边界框未能及时跟随车辆转向后的真实位置进行更新。基线 MeMOTR(第三行)在第 t 帧至第 $t+1$ 帧之间, ID5 车辆未能保持稳定的目标表示,其预测框被相邻 ID4 车辆的预测结果覆盖,表现出目标混叠。同时转弯车辆的边界框滞后偏移现象较 MOTIP 更为严重。

综合三组可视化结果可以看出,车辆转弯会使目标的运动方向和空间位置在短时间内快速变化,从而首先影响活动轨迹的状态递推稳定性。当原有轨迹状态不能及时跟随这种变化时,就容易出现短时失配、轨迹中断、边界框滞后偏移以及后续身份切换等现象。MOTIP 将多目标跟踪进一步建模为上下文身份预测问题,其优势在于能够利用较强的身份



图4 FastTracker数据集上的可视化对比结果

Fig. 4 Visual Comparison Results on the FastTracker Dataset

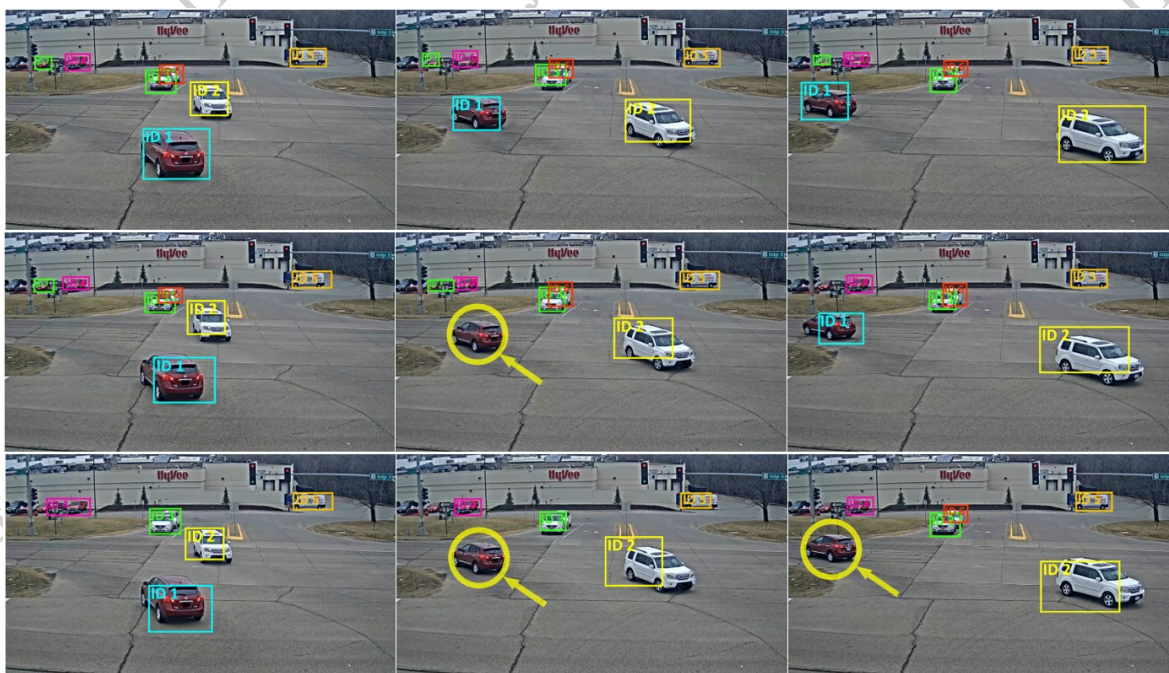


图5 CityFlow数据集上的可视化对比结果

Fig. 5 Visual Comparison Results on the CityFlow Dataset

建模能力完成后续关联判断,因此在部分片段中即使发生短时失配,后续仍能够恢复正确关联。但对于车辆转弯场景而言,问题并不只在于当前目标属于哪个身份,更在于活动轨迹的状态表示能否及时跟随目标真实运动变化。当车辆转弯时目标的运动

信息变化较快,而外观信息变化相对较慢,若缺少针对这种快慢变化差异的专门状态更新机制,轨迹表示就容易滞后于目标真实状态。即使身份预测能力较强,MOTIP仍可能因为当前轨迹状态与目标真实位置、方向不一致,而先出现短时失配和边界框滞后

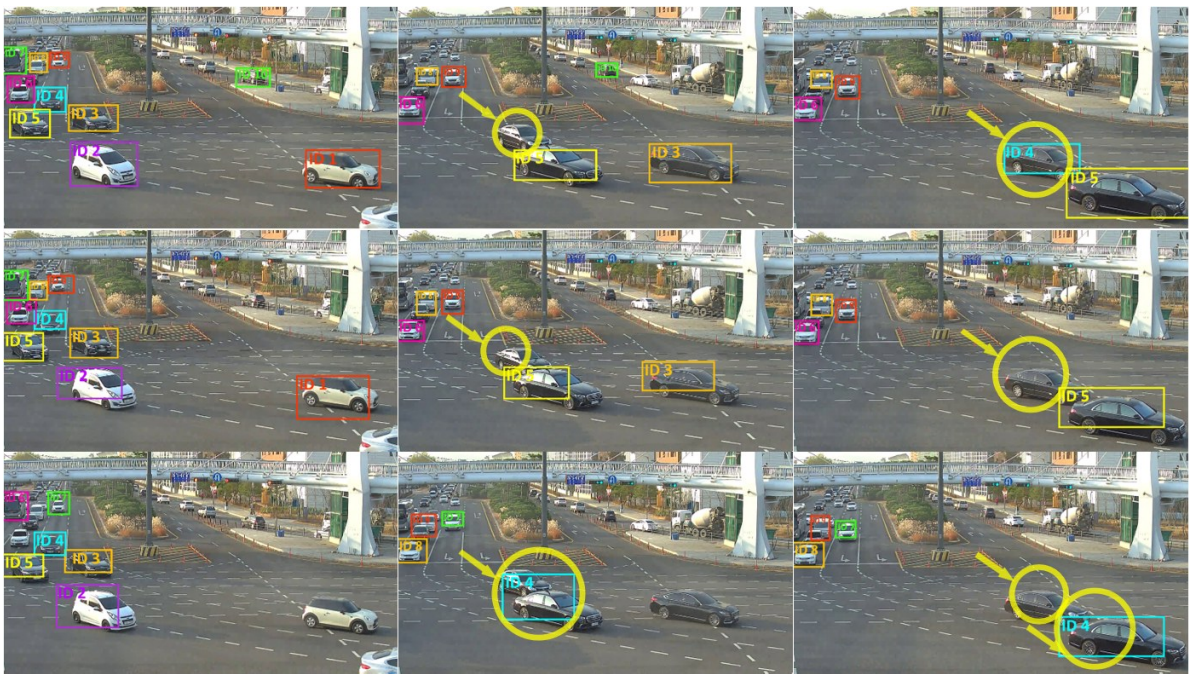


图6 TrafficMOT数据集上的可视化对比结果
Fig. 6 Visual Comparison Results on the TrafficMOT Dataset

偏移,只有在后续观测再次充分时,才有可能重新恢复正确关联。MeMOTR通过 Detection Decoder与长时记忆增强改善了端到端框架中的检测与关联基础,但其时序更新仍主要依赖单一状态流,转弯时快变的运动信息与缓变的外观信息容易相互干扰。MeMOTR对于因转弯而进入短期丢失状态的轨迹,也缺少针对性的状态更新机制,因此更容易出现轨迹持续失配、恢复失败以及身份切换等问题。本方法引入快慢双时间尺度记忆模块,将外观信息与运动信息分支建模并分别递推,缓解了活动轨迹状态更新中的信息干扰,同时通过短期丢失轨迹状态更新模块,对因转弯而短时失配的轨迹进行状态递推

修正,使其在后续重关联时更接近目标当前真实状态。因此,本方法能够在车辆转弯场景下保持更稳定的轨迹连续性、身份一致性和位置跟随能力,这也与表1、表2与表3中 AssA、IDs、Frag、MT和ML等关联相关指标的变化趋势一致。

3.5 消融实验结果分析

为了验证本方法提出的各自模块的效果,分别在 FastTracker、TrafficMOT 数据集上进行消融实验。实验结果如表4与表5所示,表中加粗表示最优数据。表格中 A0表示基线模型 MeMOTR 方法、A1表示快慢双时间尺度记忆模块 FSTM、A2表示短期丢失轨迹状态更新模块 SMTU。

表4 在 FastTracker数据集的消融实验结果
Table 4 Ablation Experiment Results on the FastTracker Dataset

A0	A1	A2	HOTA	AssA	IDF1	IDs	Frag	MT	ML	DetA
√			41.6	40.5	48.1	489	721	39.7	22.3	44.9
√	√		44.0	42.6	50.1	421	634	43.2	19.5	46.8
√		√	43.7	42.2	49.4	438	654	42.6	20.1	46.7
√	√	√	44.3	42.8	50.6	413	621	43.4	19.0	47.0

注:加粗表示最优数据

总体来看,无论单独引入 FSTM 还是 SMTU,模型性能都能够得到提升,而当二者同时引入后,模型

在两个数据集上均取得了最佳结果,说明本方法各组成部分对关联相关性能提升均具有实际贡献,且

表5 在TrafficMOT数据集的消融实验结果

Table 5 Ablation Experiment Results on the TrafficMOT Dataset

A0	A1	A2	HOTA	AssA	IDF1	IDs	Frag	MT	ML	DetA
√			42.6	40.8	48.7	435	665	40.6	21.2	45.6
√	√		44.1	42.6	50.4	384	594	43.4	18.7	46.8
√		√	43.9	42.2	49.8	399	612	42.7	19.2	46.7
√	√	√	44.2	42.8	50.8	377	588	43.8	18.5	46.8

注:加粗表示最优数据

整体设计具有较好的互补性。

与 A0 相比,仅引入 FSTM 后,模型在 HOTA、AssA、IDF1、IDs 和 Frag 等指标上的提升更为明显,说明将外观缓变信息与运动快变信息分开建模,能够有效提升车辆转弯场景下活动轨迹状态递推的稳定性。其原因在于,车辆转弯过程中运动方向和空间位置变化较快,而外观信息变化相对较慢,若采用单一流状态递推,快慢信息之间容易相互干扰,从而影响状态更新效果。在 FastTracker 和 TrafficMOT 上,A0+A1 表现出更高的关联准确率以及更低的身位切换和轨迹碎片化次数,这说明 FSTM 主要解决了可见活动轨迹在连续转弯过程中的状态更新不稳定问题。

仅引入 SMTU 后,模型性能同样出现小幅提升,说明由车辆转弯引起的短期丢失轨迹也需要专门的状态补偿机制。如果短期丢失轨迹在失轨期间缺乏有效更新,其轨迹状态会逐渐偏离目标当前真实状态,从而影响后续重关联效果。SMTU 通过结合最近更新轨迹嵌入、边界框变化趋势以及丢失时长等信息,对短期丢失轨迹进行状态递推,因此能够在一定程度上提升短期失轨轨迹的恢复能力。

当 FSTM 和 SMTU 同时引入后,模型在两个数据集上取得了最佳结果,说明二者并不是作用重复的两个模块。前者主要增强活动轨迹的状态建模能力,后者主要增强短期失轨轨迹的恢复能力。二者协同作用后,能够更完整地抑制车辆转弯带来的关联不稳定问题,从而实现本方法的最优性能。

4 结 论

本方法主要针对车辆转弯导致活动轨迹状态递推不稳定以及转弯引起短期丢失轨迹缺少状态更新两个关键问题,提出了面向车辆转弯场景的轨迹状

态更新与关联增强多目标跟踪。其中快慢双时间尺度记忆模块 FSTM 通过分别建模外观信息与运动信息,缓解了单一状态流递推过程中不同时间尺度信息相互干扰的问题,提升了活动轨迹在车辆转弯场景下的状态更新稳定性。短期丢失轨迹状态更新模块 SMTU 则结合轨迹最近的有效更新信息、边界框变化趋势以及丢失时长,对短期丢失轨迹嵌入进行递推优化,从而增强了模型对短期丢失轨迹后续的重关联能力。

实验结果表明,本方法在 FastTracker、CityFlow 和 TrafficMOT 三个数据集上均取得了较好的综合表现,尤其在 AssA、IDs、Frag、MT 和 ML 等关联相关指标以及可视化结果上优势较为明显,说明所提方法能够有效提升车辆转弯场景下的轨迹关联稳定性与连续性,减少身份切换和轨迹中断现象。

需要指出的是,本文方法仍存在一定局限性。一方面,当车辆在转弯过程中受到长时间严重遮挡时,短期丢失轨迹能够利用的有效历史信息有限,轨迹状态外推仍可能出现偏差;另一方面,在夜间、雨雾或强光反射等复杂成像条件下,车辆外观特征质量下降,可能影响外观慢记忆的稳定更新。此外,本文方法主要围绕车辆转弯场景下的轨迹状态更新与关联质量提升展开研究,重点关注身份保持、轨迹连续性和短期失轨恢复能力,尚未针对模型轻量化和实时推理效率进行专门优化,实时性提升将作为后续工作进一步研究。

未来工作将从以下几个方面进一步展开:首先,引入更加鲁棒的遮挡感知机制,提高模型在严重遮挡和长时间失轨情况下的轨迹恢复能力;其次,结合低光照增强或退化感知特征建模方法,提升模型在夜间和恶劣天气交通场景中的适应性;最后,进一步探索轻量化网络结构和高效推理策略,以增强方法在实际道路监控系统中的部署可行性。

参考文献(References)

- Cao J K, Pang J M, Weng X S, Khirondkar R and Kitani K. 2023. Observation-centric SORT: rethinking SORT for robust multi-object tracking//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, Canada: IEEE: 9686-9696 [DOI: 10.1109/CVPR52729.2023.00934]
- Dong Y H, Tu Q Y, Zhou Z J and Luo J S. 2025. Vehicle trajectory detection technology based on machine vision. *Computer Science and Application*, 15(1): 10-20 (董晔卉, 涂倩怡, 周子俊, 罗建书. 2025. 基于机器视觉的车辆轨迹检测技术. *计算机科学与应用*, 15(1): 10-20) [DOI: 10.12677/csa.2025.151002]
- 王凯, 戴芳, 郭文艳, 王军锋, 王小侠. 2024. 融合目标相似性和作用力的多目标跟踪. *中国图象图形学报*, 29(07): 1984-1997 [DOI: 10.11834/jig.230340]
- 杨荣, 刘偲礪, 王涵, 周壮, 李盛阳. 面向地面匹配场景的空间科学果蝇实验多目标跟踪方法 DN-MeMOTR[J/OL]. *中国图象图形学报*, 2026, 1-11. DOI: 10.11834/jig.250589
- Guo Y Y, Liu P, Yuan Q, Liu P, Xu J and Zhang H. 2023. Review on research of road traffic safety of connected and automated vehicles. *Journal of Traffic and Transportation Engineering*, 23(5): 19-38 (郭延永, 刘佩, 袁泉, 刘攀, 徐进, 张晖. 2023. 网联自动驾驶车辆道路交通安全研究综述. *交通运输工程学报*, 23(5): 19-38) [DOI: 10.19818/j.cnki.1671-1637.2023.05.002]
- Hashempoor H and Hwang Y D. 2025. FastTracker: real-time and accurate visual tracking [EB/OL]. [2026-04-06]. <https://arxiv.org/abs/2508.14370> [DOI: 10.48550/arXiv. 2508.14370]
- Hu S Y, Zhao X and Huang K Q. 2024. Visual intelligence evaluation techniques for single object tracking: a survey. *Journal of Image and Graphics*, 29(8): 2269-2302 (胡世宇, 赵鑫, 黄凯奇. 2024. 单目标跟踪中的视觉智能评估技术综述. *中国图象图形学报*, 29(8): 2269-2302) [DOI: 10.11834/jig.230498]
- Jin S S, Long W, Hu L X, Wang T Y, Pan H and Jiang L H. 2023. Research progress of detection and multi-object tracking algorithm in intelligent traffic monitoring system. *Control and Decision*, 38(4): 890-901 (金沙沙, 龙伟, 胡灵犀, 王天宇, 潘华, 蒋林华. 2023. 多目标检测与跟踪算法在智能交通监控系统中的研究进展. *控制与决策*, 38(4): 890-901) [DOI: 10.13195/j.kzyjc.2021.1763]
- Li Y, Lin F, Chen Y A and Lv S H. 2025. A comprehensive review of deep learning-based visual multi-object tracking. *Science Technology and Engineering*, 25(22): 9211-9223 (李永, 林坊, 陈煜昂, 吕纾涵. 2025. 基于深度学习的视觉多目标跟踪综述. *科学技术与工程*, 25(22): 9211-9223) [DOI: 10.12404/j.issn.1671-1815.2409394]
- Qin Z, Wang L, Zhou S P, Fu P P, Hua G and Tang W. 2024. Towards generalizable multi-object tracking//Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 18995-19004 [DOI: 10.1109/CVPR52733.2024.01797]
- Sun H, Li Y R, Zhang L L, Zhai Y, Chen Z Y and Chen C. 2024. Scenario complexity calculation model of real road test based on operational design condition. *Automotive Engineering*, 46(11): 1983-1992 (孙航, 李雨冉, 张琳琳, 翟洋, 陈振宇, 陈陈. 2024. 基于设计运行条件的实际道路测试场景复杂度计算模型. *汽车工程*, 46(11): 1983-1992) [DOI: 10.19562/j.chinasae.qcgc.2024.11.005]
- Sun P, Cao J, Jiang Y, Zhang R, Xie E, Yuan Z, et al. 2020. TransTrack: multiple object tracking with Transformer [EB/OL]. [2026-04-06]. <https://arxiv.org/abs/2012.15460> [DOI: 10.48550/arXiv. 2012.15460]
- Wen M D. 2024. A study on traffic detection technology and application based on machine vision. *Open Journal of Transportation Technologies*, 13(6): 388-394 (温美东. 2024. 基于机器视觉的交通检测技术及应用研究. *交通技术*, 13(6): 388-394) [DOI: 10.12677/ojtt.2024.136042]
- Yan F, Luo W, Zhong Y, Gan Y and Ma L. 2025. Co-MOT: boosting end-to-end Transformer-based multi-object tracking via cooperation label assignment and shadow sets//The Thirteenth International Conference on Learning Representations. Singapore: OpenReview.net
- Yang R, Liu S L, Wang H, Zhou Z and Li S Y. 2026. Multi-object tracking for drosophila in space science experiments under ground-matching scenarios with DN-MeMOTR. *Journal of Image and Graphics*: 1-11 (杨荣, 刘偲礪, 王涵, 周壮, 李盛阳. 2026. 面向地面匹配场景的空间科学果蝇实验多目标跟踪方法 DN-MeMOTR. *中国图象图形学报*: 1-11) [DOI: 10.11834/jig.250589]
- Zeng F, Dong B, Zhang Y, Wang T, Zhang X and Wei Y. 2022. MOTR: end-to-end multiple-object tracking with Transformer//Computer Vision - ECCV 2022. Tel Aviv, Israel: Springer: 659-675 [DOI: 10.1007/978-3-031-19812-0_38]
- Zhang Y F, Sun P, Jiang Y, Yu D D, Weng F C, Yuan Z H, et al. 2022. ByteTrack: multi-object tracking by associating every detection box//Proceedings of the 17th European Conference on Computer Vision. Tel Aviv, Israel: Springer: 1-21 [DOI: 10.1007/978-3-031-20047-2_1]

作者简介

都双丽, 女, 副教授, 研究方向为计算机视觉. E-mail: dusl@xaut.edu.cn

刘梓扬, 男, 硕士研究生, 研究方向为计算机视觉. E-mail: liulife1109@163.com

赵明华,女,教授,研究方向为计算机视觉。E-mail:zhaoming-hua@xaut.edu.cn

胡静,女,副教授,研究方向为计算机视觉。E-mail:jinghu@xaut.edu.cn

石程,女,副教授,研究方向为计算机视觉。E-mail:chengc_s@163.com

李鹏,男,副教授,研究方向为计算机视觉。E-mail:pengli@xaut.edu.cn