

高速路侧雷达和相机融合的目标检测与跟踪

曾永强¹,张永华^{2*},赵辉¹,刘鑫¹,许铭生¹,沈建伟¹,胡志伟¹,陈珍泉¹

(1.福建省高速公路集团有限公司莆田管理分公司,福建莆田 351100;

2.清华大学 计算机科学与技术系,北京市 100084)

摘要:相机作为智能交通系统(Intelligent Transportation Systems, ITS)中应用最广泛的传感器,在目标遮挡和复杂环境干扰下常出现检测和定位精度下降的问题,制约了ITS系统性能的进一步发展。为解决此问题,该文设计了一种基于高速路侧毫米波雷达(Millimeter-Wave Radar, MWR)和相机的多模态融合目标检测与跟踪方法。与相机相比,高分辨率MWR具有更优的测量精度和天气鲁棒性,能够有效弥补相机在感知能力上的局限。该方法采用基于中心点的雷达-相机融合算法进行目标检测,并借用贪婪算法实现目标关联。试验表明:在公开数据集nuScenes和自建的高速路侧多模态数据集上,所提方法在nuScenes数据集中实现了69.1%的 A_{MOTA} (多目标跟踪精度)性能,优于所有基于视觉的3D跟踪基准方法;在自建数据集验证了其良好的适用性与准确性,单图处理时间为35 ms。

关键词:智能交通系统;多模态融合;目标检测与跟踪;高速路侧雷达;毫米波雷达

中图分类号:U495

文献标志码:A

Object Detection and Tracking Based on Fusion of Radar and Camera on Highway Side

ZENG Yongqiang¹, ZHANG Yonghua^{2*}, ZHAO Hui¹, LIU Xin¹, XU Mingsheng¹,

SHEN Jianwei¹, HU Zhiwei¹, CHEN Zhenquan¹

(1. Putian Management Branch, Fujian Expressway Group Co., Ltd., Putian, Fujian 351100, China;

2. Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China)

Abstract: Camera is the most widely used sensor in intelligent transportation systems (ITS), but the decline of detection and positioning accuracy caused by target occlusion and external environment interference has always been an important factor restricting the development of ITS. In order to solve this problem, a multi-mode fusion target detection and tracking method based on millimeter-wave radar (MWR) and camera on highway side was designed. Compared with cameras, high-resolution MWR had better measurement accuracy and weather robustness, serving as a better complement to camera perception. The proposed method used the center point-based radar and camera fusion algorithm for target detection and adopted the greedy algorithm for target association. The test results show that on the public dataset nuScenes and the self-built multi-modal dataset of highway side, the proposed method achieves an A_{MOTA} (multi-object tracking accuracy) performance of 69.1% on the nuScenes dataset, outperforming all visual-based 3D tracking benchmark methods; the self-built dataset verifies its good applicability and accuracy, and the processing time for a single image is 35 ms.

Keywords: intelligent transportation system; multi-modal fusion; object detection and tracking; radar on highway side; millimeter-wave radar

0 引言

相对于普通公路,高速公路是现有公路体系的最高技术等级。在高速公路养护施工中,常采用占

道施工的方式,临时封闭某一车道以满足作业需求,这在一定程度上给过往车辆的正常行驶造成了影响,从而产生安全隐患^[1-2]。高速公路车速极快,一旦发生事故,极易导致车毁人亡的悲剧,经济损失也动

收稿日期:2024-08-16 修回日期:2025-05-10

基金项目:福建省交通运输科技项目(编号:202244)

作者简介:曾永强,男,工程师.E-mail:289783517@qq.com

*通信作者:张永华,男,博士,助理研究员.E-mail:zhangyonghua@tsinghua.edu.cn

辄上百万元^[3]。近十年来,随着信息、通信、传感器技术的快速发展,公路智能化经历了大量的技术研发、示范和推广,实现了工业互联网、5G、传感器技术的有机融合,提高了智能交通系统(Intelligent Transportation Systems, ITS)精细化水平。现代公路系统架构的重点是以综合交通流量监测为主要特征,充分利用雷达、激光雷达、摄像头等先进传感器对车辆轨迹进行跟踪,为解决交通系统的完整性、效率和环境问题提供了有效途径^[4]。利用技术手段建立检测与跟踪系统,对车辆运动的准确、鲁棒和可靠的定位跟踪,是实现ITS及时消除事故隐患等功能和应用的基础。对高速公路道路信息实现交通流状况、车辆路径和行车安全等实时空间感知,保障高速公路的安全通行已经成为高速公路运营公司的迫切需求^[5]。

传统的目标检测与跟踪系统算法首先利用目标检测器获取目标的信息、外观或运动特征。然后,通过关联算法计算这些特征与现有轨迹状态的相似度,从而实现目标的动态跟踪。基于相机的目标检测与跟踪系统算法方法^[6-8]主要应用于2D数据,通过丰富的视觉纹理信息来实现稳定的跟踪。SORT算法通过线性速度模型卡尔曼滤波器预测物体运动,然后度量检测框和跟踪框的交并比(Intersection-over-Union, R_{IOU})距离,最后利用匈牙利算法将检测结果关联在一起,实现多目标检测与跟踪^[9]。Redmon等^[10]、谢海波等^[11]构建了基于Yolo网络的多车辆跟踪框架,可用于实时交通应用。然而,在交通场景中,车辆形态变化、运动模糊、背景干扰及恶劣天气等因素,会显著降低特征的可靠性,使基于单目视觉的跟踪技术仍具有挑战性。因此,Wu等^[12]提出了相机激光雷达融合方法,能以较低成本弥补单一传感器的局限,是当前的研究热点。根据融合位置,这些方法可以分为前端融合和后端融合。由于图像和点云之间存在差异,在前端融合方法中,两种模态数据的特征对齐相当困难^[13]。后端融合方法^[14-15]利用图像和激光雷达的2D和3D检测结果进行匹配,但由于检测器的错误检测以及交通场景的复杂性和类别多样性,经常发生不同检测目标之间的身份ID交错不匹配的跟踪错误问题^[16]。基于激光雷达的MOT方法受天气影响较大,激光雷达在雨、雪、雾等恶劣天气条件下性能会显著下降,影响其可靠性。与激光雷达相比,毫米波雷达能够在各种恶劣天气条件下稳定工作并且具有较远的探测距离,能够覆盖更大的检测范围,适用于全天候、长距离的多目标

检测与跟踪任务。此外,毫米波雷达的成本较低,更适合在ITS等大规模应用场景中部署。

为了解决上述问题,本文基于雷达传感器和图像传感器多模态融合技术,扩展了自动驾驶中的3D目标检测算法CenterFusion^[17],利用雷达和相机多模态融合数据,提出了一个端到端的目标检测和跟踪方法,并命名为RCFTrack(Radar and Camera Multimodal Fusion Tracking, RCFTrack)。另外,提出了一种基于加权成本函数的贪婪算法,该算法根据目标的估计深度、速度及其2D位移来计算,获取的深度和速度信息可区分不同目标的重要信息,显著提高了网络正确关联被遮挡和重叠目标的判别能力。RCFTrack将前一帧图像和检测到的目标以及当前图像帧和雷达检测结果作为输入,将3D目标检测结果和检测到的目标的跟踪ID作为输出。与仅使用相机传感器信息相比,RCFTrack使用融合的雷达和图像特征能够更准确预测连续帧中目标的位移。基于具有挑战性的nuScenes数据集^[18]的试验结果表明:RCFTrack在nuScenes基准测试上实现了69.1%的 A_{MOTA} ,优于所有其他基于图像的跟踪方法以及nuScenes基于激光雷达的基线方法AB3DMOT^[19]。在自建数据集验证了适用性以及准确性,其性能相比CenterTrack^[20]高出6倍,且每张图像的运行时间为35 ms,非常适合ITS应用。

本文的主要贡献有三方面:①提出了一种用于端到端训练的深度神经网络RCFTrack,用于3D目标检测与跟踪的雷达和相机传感器融合框架;②提出了一种用于目标跟踪关联的贪婪算法,该算法将目标的深度、速度和2D位移相结合,实现了更准确的目标跟踪;③遵循nuScenes数据集格式,采集并标注了高速公路场景路侧雷达和相机目标检测与跟踪多模态数据集。

1 模型概述

首先RCFTrack使用改进的毫米波雷达与相机融合检测网络,实现目标的精准检测与定位。其次,利用融合后的雷达和图像特征估计物体在连续帧间的位移,以此进行时序关联。最后,结合前一帧的相机和雷达数据,有效增强对当前帧中被遮挡目标的跟踪能力。处理跟踪问题时,即使目标离开帧,仍会在连续帧中保留目标的身份ID;当目标再出现时无需建立新的跟踪轨迹。在关联步骤中,提出了一种贪婪算法,该算法除利用物体的2D位移

外,还利用物体的速度和深度信息来实现随时间的精确关联。

RCFTrack网络总体架构如图1所示,在端到端训练的单个深度网络中完成目标的检测和关联跟踪。RCFTrack的输入显示在左侧,是当前和之前的图像帧 $I^{(t-1)}, I^{(t)} \in R^{W \times H \times 3}$,当前和之前的雷达探测 $P^{(t-1)}, P^{(t)} \in R^{N \times 5}$ 。其中, W, H 为图像帧的尺寸; N 为雷达探测次数,以及上一帧中的跟踪目标 $T^{(t-1)} = \{b_0^{(t-1)}, b_1^{(t-1)}, \dots, b_n^{(t-1)}\}$ 。跟踪的目标由 $b =$

(p, d, v, w, N_{ID}) 表示。其中, $p \in R^2$ 为目标的中心位置; $d \in R$ 为目标的深度; $v \in R^2$ 为目标的速度; $w \in [0, 1]$ 为检测置信度; N_{ID} 为跟踪目标唯一标识的整数。在模型中添加了一个额外的回归头(“Dis”),它使用融合的雷达和图像特征来预测连续帧中的物体位移。关联步骤中的贪婪算法使每个目标的位移、深度和速度与以前的检测相关联。输出目标是检测和跟踪目标 $T^{(t)} = \{b_0^{(t)}, b_1^{(t)}, \dots, b_n^{(t)}\}$,并为连续帧中的目标分配一致的 N_{ID} 。

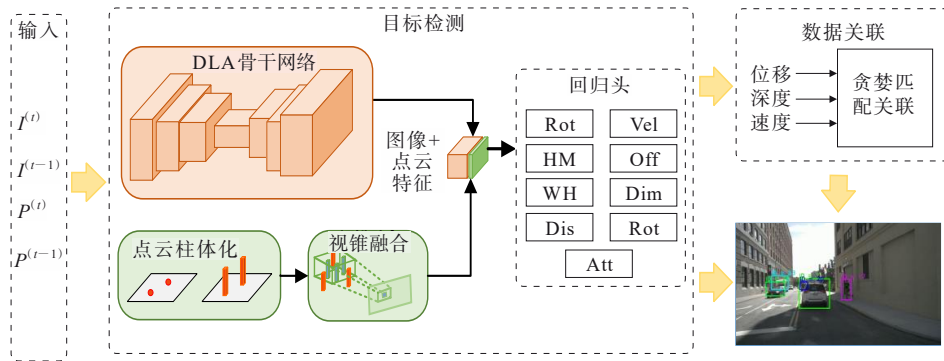


图1 RCFTrack网络总体架构图

Figure 1 Overall architecture of RCFTrack network

1.1 改进检测网络

CenterFusion是雷达和相机融合进行3D目标检测的最先进算法之一,其提供的目标速度估计,对目标跟踪任务具有很大帮助。CenterFusion采用图像 $I \in R^{W \times H \times 3}$ 和一组雷达检测, (x^i, y^i, z^i) 为雷达点云中点 I 的坐标, (v_x^i, v_y^i) 分别为 x 和 y 方向的径向速度。这些坐标是根据路测雷达坐标系定义,其中, x 向前, y 向左, z 向下(从路测雷达的角度来看)。

改进的CenterFusion采用以深层聚合网络为骨干网络的中心点检测方法,将当前图像帧 $I^{(t)}$ 与雷达探测 $P^{(t)}$ 、前一时刻的图像帧 $I^{(t-1)}$ 与雷达探测 $P^{(t-1)}$ 以及探测到的目标作为输入,所有检测到的目标的3D边界框和每个目标的绝对速度作为输出。再将先前的检测结果表示为2D高斯核的单通道热图。然后加入先前的图像、雷达检测和检测到的目标有助于网络更好地估计目标在当前帧中的位置。同时,利用前一帧的雷达信息进一步提高网络检测目标的能力,有效改善目标的被遮挡问题。此外修改后的网络通过级联的雷达和图像特征来估计检测到的目标在当前帧和先前帧之间的2D位移。除了来自当前帧和先前帧的图像特征,雷达深度和速度信息还有

助于网络生成更准确的物体位移预测。

总体而言,通过估计热图 $\hat{Y} \in [0, 1]^{\frac{W}{R} \times \frac{H}{R} \times C}$ 来检测目标的中心点。其中, R 为下采样因子; C 为数据集中目标类别的数量。所估计的热图 \hat{Y} 中的局部最大值对应于图像中检测到的目标的中心。地面实况热图 Y 通过每个目标的中心点处渲染高斯形状的峰值生成,该峰值为其相应边界框的中心点。

改进的目标检测网络通过回归层为所有目标生成初步的3D边界框。然后,采用基于视锥体的关联方法,将雷达检测与这些初步检测进行匹配。具体实施步骤如下:先将雷达探测扩展为预定义体柱,在每个目标周围形成视锥体,并将内部的雷达体柱关联至该目标。如果1个视锥体内有多个雷达体柱时,系统择近保留1个,余者弃用。

基于视锥关联的结果,将雷达检测的深度和速度映射到图像上对应的目标。这些值表示为单独的热图通道,并连接到基于图像的特征。然后,这些融合特征被用于两方面:①通过重新计算深度、大小及方向等属性来优化初步检测结果;②为每个检测到的目标估计速度向量。

本部分使用基于焦点损失的目标函数,定义为:

$$L_i = \frac{-1}{N} \sum_{x_{yc}} \begin{cases} (1 - Y_{x_{yc}}^\Delta)^a \log(Y_{x_{yc}}^\Delta), \text{ if } Y_{x_{yc}} = 1 \\ (1 - Y_{x_{yc}})^\beta (Y_{x_{yc}}^\Delta) \log(1 - Y_{x_{yc}}), \text{ 其他} \end{cases} \quad (1)$$

式中: N 为目标的数量; $Y \in [0, 1]^{\frac{w}{R} \times \frac{h}{R} \times c}$ 为带注释目标的真实热图; α 和 β 为焦点损失的超参数。基于检测到的目标中心点,利用不同的回归头来预测其尺寸、方向、深度及速度。

1.2 基于贪婪算法的跟踪

RCFTrack采用贪婪算法实现目标的跨帧关联。不同于CenterTrack^[3]中的简单贪婪算法,CenterTrack仅在连续图像中使用目标的2D位移进行关联。本文提出了一种基于加权成本函数的贪婪算法,该算法根据目标的2D位移来计算其估计深度和速度,提高了网络正确关联被遮挡和重叠目标的能力。RCFTrack将前一帧图像和检测到的目标以及当前图像帧和雷达检测结果作为输入,3D目标检测结果和检测到的目标的跟踪ID作为输出。检测到的目标由 $a = (p, d, v, c)$ 表示。其中, $p \in Z^2$ 为物体的像素中心; $d \in R$ 为物体的深度; $v \in R^2$ 为物体的速度; $c \in C$ 为物体类别。与文献[4]类似,通过一个回归层计算位移,其形式为2个输出通道: $\hat{Y}^{(t)} \in R^{\frac{w}{R} \times \frac{h}{R} \times c}$,表示图像目标中心的位移。与其他回归头类似, L_1 损失被用作训练该层的目标函数。

为了在时间上关联目标,根据目标在图像上的深度、速度和位移定义了一个成本函数:

$$\text{Cost}_{t,t-1} = \begin{cases} \alpha \cdot L_{\text{pixel}} + \beta \cdot L_{\text{depth}} + \delta \cdot L_{\text{velocity}}, & c_t = c_{t-1} \\ \infty, & c_t \neq c_{t-1} \end{cases} \quad (2)$$

$$L_{\text{depth}} = (x_t - x_{t-1})^2 + (y_t - y_{t-1})^2 \quad (3)$$

$$L_{\text{depth}} = (d_t - d)^2 \quad (4)$$

$$L_{\text{velocity}} = (v_{x_t} - v_{x_{t-1}})^2 + (v_{y_t} - v_{y_{t-1}})^2 \quad (5)$$

式中: x, y 为物体的中心; d 为物体的深度; v_x, v_y 分别为物体在 x 和 y 方向上的速度。 $\alpha, \beta, \delta \in R^+$ 为可调参数。

对于位置 p 处的每个检测到的物体,RCFTrack在 $p - D_p$ 的半径 r 内寻找先前的检测。如果在该位置存在不匹配的先前检测,RCFTrack计算上述成本函数来确定这些检测之间的距离,并以最低的成本将目标与先前的检测相匹配。对于每一个不匹配的

检测,都会创建一个新的轨迹。

2 试验

2.1 实施详细信息

根据CenterFusion,RCFTrack使用 800×448 像素的输入分辨率,并采用水平翻转和随机移位进行数据正则化。使用深层聚合网络用作提取图像特征的主干,并通过Adam进行了优化。RCFTrack网络训练周期(epoch)设置为60,并使用预先训练了170个epoch的CenterFusion网络权重,批量大小为32,学习率为 1.2×10^{-4} 。试验平台配备了Intel Xeon E5-1650 CPU和2个Quadro A6000 GPU计算加速服务器,系统为Ubuntu 20.04,软件环境是Python 3.9和Pytorch 1.8。RCFTrack在CenterFusion的基础上重新实现了视锥融合算法,并将其并行运行,可缩短检测网络的整体运行时间。

2.2 数据集和评估指标

本研究首先在极具挑战性的nuscnescs数据集上进行评估。nuscnescs是一个用于自动驾驶的大规模公开数据集,其提供了相机、雷达和激光雷达数据,并包含用于3D物体检测和跟踪的精细注释。它提供了1000个不同序列,其中700个序列用于训练,150个序列用于验证,150个序列用于测试。每个序列由40个带注释的帧组成,每个帧包含相机、雷达和激光雷达样本。帧中的每个注释都包含一个3D边界框和一个唯一的ID。轨迹目标框的交并比 R_{IoU} 大于0.5被视为真阳性, R_{IoU} 小于0.5的轨迹目标框被视为误报,没有匹配轨迹的目标框视为假阴性。nuscnescs基准测试中采用的主要评估指标有 N_{FRAG} 、 N_{IDSW} 、 A_{MOTA} 及 \bar{A}_{MOTA} 。其中, N_{FRAG} 为轨迹被分割(即在跟踪过程中中断)的总次数, N_{IDSW} 为身份切换的数量, A_{MOTA} 为多目标跟踪精度:

$$A_{\text{MOTA}} = 1 - \frac{N_{\text{FNr}} + N_{\text{FPt}} + N_{\text{IDSw}}}{GT} \in (-\infty, 1) \quad (6)$$

A_{MOTA} 可以是负的,因为算法可能会出现检测或跟踪错误数量大于地面真值框的数量情况。其中, N_{IDSw} 为正确检测到对象但重新分配ID错误的次数; N_{FPt} 为整个视频中的误报总数; N_{FNr} 为整个视频中的漏报总数。

\bar{A}_{MOTA} 是不同召回阈值下再调用归一化多目标跟踪精度(A_{MOTA})指标的加权平均值:

$$A_{\text{MOTA}} = \max \left[0, 1 - \frac{N_{\text{FNr}} + N_{\text{FPt}} + N_{\text{IDSW}} - (1-r)P}{rP} \right] \quad (7)$$

$$\bar{A}_{\text{MOTA}} = \frac{1}{n-1} \sum_{r \in \left\{ \frac{1}{n-1}, \frac{2}{n-1}, \dots, 1 \right\}} A_{\text{MOTA}} \quad (8)$$

式中: r 为召回阈值; P 为所有帧中注释目标的总数。

此外,为验证所提方法在基于路侧相机和雷达进行车辆检测与跟踪时的准确性和可靠性,在真实高速公路场景下进行了一系列现场测试。测量数据由福建省交通监控指挥中心采集。在莆田至泉州路段某路边的电线杆上,装有1个毫米波雷达传感器和高分辨相机。其中,毫米波雷达具体型号为77 GHz Conti-ASR-408-21,采样率约为15 Hz;相机分辨率为1080P(1 920×1 080),采样率为30 Hz。权衡标注成本和模型验证需要,并遵循nuScenes数据集标注格式,共采集并标注400个序列,其中300个序列用于训练,100个序列用于测试。每个序列包含35个带注释的帧,每个帧包含相机和雷达样本。

2.3 试验结果

提出的RCFTrack在nuScenes数据集目标跟踪基准中与其他已发布方法的比较结果如表1所示(“↑”表示各指标值越大模型性能越好,“↓”与“↑”相反)。所对比的方法有使用2D相机数据的CenterTrack、DEFT,使用3D点云数据的AB3DMOT以及2D相机与3D点云数据的CBMOT、EagerMOT。为进行消融研究,设置了3个RCFTrack版本:RCFTrack^{cf.ga}表示CenterFusion与贪心算法的简单组合,RCFTrack^{icf.ga}表示改进的CenterFusion与贪心算法的组合,RCFTrack在RCFTrack^{icf.ga}之上改进了贪心算法,即最终提出的跟踪模型。RCFTrack 3个不同版本的试验结果表明:简单组合CenterFusion与贪心算法跟踪效果优于基于2D或3D单模态输入的方法,该方法性能虽与主流融合相机-雷达融合方法接近,但仍未超越。使用改进的CenterFusion作为检测器,能够进一步提升整体的跟踪性能,但是跟踪轨迹碎片化以及ID切换的问题没有得到有效解决。RCFTrack的跟踪性能对比RCFTrack^{icf.ga}, N_{IDSW} 和 N_{FRAG} 指标显著降低,验证了本文针对CenterFusion与贪心算法所提改进的有效性。

由表1可知:RCFTrack获得了最高的 \bar{A}_{MOTA} (69.1%)和 A_{MOTA} (57.6%);同时可以发现使用2D相

机和3D点云数据的多模态方法明显优于只使用2D或者3D的单模态方法;与最先进的CBMOT对比,RCFTrack的 \bar{A}_{MOTA} 和 A_{MOTA} 提升分别达到2.1%和6.9%。同时,该方法能够达到良好的跟踪效果。在nuScenes数据集的结果验证了所提改进的有效性。此外,对比同样基于贪婪匹配思想的CenterTrack方法,RCFTrack性能提升达到了6倍以上。鉴于RCFTrack和CenterTrack跟踪算法的相似性,这些结果证明了在检测和跟踪阶段利用雷达数据的效果。这两种方法都使用贪心算法来关联目标,但RCFTrack还利用检测到的目标的深度和速度来随着时间的推移更好地关联它们。雷达提供的速度数据使网络能够更准确地预测图像中物体的位移,进一步提高物体关联。

表1 RCFTrack与其他代表性方法在nuScenes测试集的评估结果

Table 1 Evaluation results of RCFTrack and other representative methods in nuScenes test set

模型	输入类型	\bar{A}_{MOTA} ↑	A_{MOTA} ↑	N_{IDSW} ↓	N_{FRAG} ↓
CenterTrack ^[20]	2D	0.108	0.850	7 608	3 107
DEFT ^[21]	2D	0.180	0.160	6 901	3 420
AB3DMOT ^[19]	3D	0.151	0.154	9 027	2 557
CBMOT ^[22]	2D+3D	0.677	0.539	709	1 015
EagerMOT ^[23]	2D+3D	0.676	0.568	1 560	601
RCFTrack ^{cf.ga}	2D+3D	0.649	0.545	2 557	1 113
RCFTrack ^{icf.ga}	2D+3D	0.681	0.553	1 897	1 202
RCFTrack	2D+3D	0.691	0.576	718	798

为了进一步验证RCFTrack在跟踪场景中的普适性,本研究将RCFTrack的3个版本与复现的CBMOT和EagerMOT在自建路侧数据集上进行了训练与测试对比,评估结果如表2所示。

由表2可知:在图像数据分辨率更高的场景中,将改进的CenterFusion作为检测器, \bar{A}_{MOTA} 指标RCFTrack^{icf.ga}(0.582)比RCFTrack^{cf.ga}(0.531)提升了9.6%。RCFTrack算法在 N_{IDSW} 、 A_{MOTA} 和 \bar{A}_{MOTA} 上均取得最优结果,相对CBMOT和EagerMOT的优势扩大,同时其 N_{IDSW} 指标与EagerMOT的差距显著缩小。自建数据集的试验结果表明:RCFTrack对于高速路测场景更具有适应性。此外,RCFTrack每张图像的运行时间为30 ms(33 Fps),这使得它适用于ITS。

表 2 RCFTrack 与其他代表性方法在自建数据集的评估结果

Table 2 Evaluation results of RCFTrack and other representative methods in self-built datasets

模型	输入类型	\bar{A}_{MOTA} ↑	A_{MOTA} ↑	N_{IDSW} ↓	N_{FRAG} ↓
CBMOT ^[19]	2D+3D	0.561	0.543	980	879
EagerMOT ^[20]	2D+3D	0.548	0.536	1 956	788
RCFTrack ^{cf.ga}	2D+3D	0.531	0.542	1 778	1 097
RCFTrack ^{icf.ga}	2D+3D	0.582	0.558	1 406	972
RCFTrack	2D+3D	0.601	0.565	886	842

为更直观地体现 RCFTrack 算法的性能,选取 nuScenes 测试集中十字路口场景的序列 17 和降水天气场景的序列 101 进行跟踪结果的可视化分析。

图 2(a)展示了序列 17 的跟踪结果。在第 5 帧中,如 1、2 号虚线框所示,准确识别对向车辆并分配 ID 号;第 11 帧中,4 号虚线框对 1 号虚线框的目标中保持了稳定的 ID 跟踪;第 13 帧中,6 号虚线框位置本应为 3 号虚线框中 ID=17 的行人,但该行人被 ID=7 的车辆完全遮挡,而在第 15 帧中,如 8 号虚线框所示,RCFTrack 成功恢复了该被遮挡行人的原始 ID 号。同时,7 号虚线框与 5 号虚线框的跟踪目标及 ID 号一致。序列 17 的跟踪结果表明:RCFTrack 在实际道路行驶中具有良好的识别与遮挡应对能力。图 2(b)为序列 101 在雨天环境下的跟踪结果。如 9、12、14、15 号虚线框所示,尽管因目标远离导致 ID=6 的车辆暂时丢失,但当该目标再次被检测到时,系统仍能恢复其原有 ID,实现稳定跟踪。11、13 号虚线框表明,RCFTrack 在车辆密集的拥挤场景中也能实现多目标的准确识别与 ID 保持。综合序列 17 和序列 101 的可视化分析可知,本研究针对 CenterFusion 的改进以及基于加权成本函数的贪婪算法,能够有效处理复杂目标运动与严重遮挡问题。RCFTrack 在实现高精度目标识别的同时,显著减少了 ID 切换次数,达到了预期的应用效果,具有较高的实际应用价值。

3 结语

该文提出了一种基于雷达和相机传感器融合的端到端目标检测与跟踪方法——RCFTrack。为实现跨时间的目标关联,设计了一种基于加权成本函数的贪婪算法,该函数综合考虑了目标深度、速度和位移信息,从而显著提升了对重叠和遮挡目标的跟

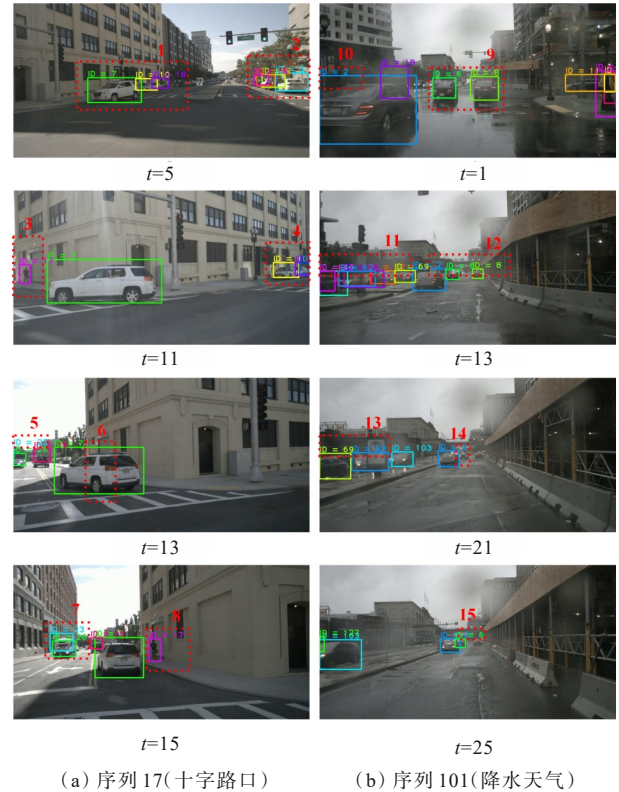


图 2 RCFTrack 在十字路口和降水天气场景下跟踪结果可视化

Figure 2 Visualization of RCFTrack's tracking results in intersections and precipitation weather scenarios

踪鲁棒性。在多目标跟踪基准 NuScenes 数据集和自建数据集上对 RCFTrack 进行了验证,结果证明了该方法具有较高的跟踪准确性,为智能交通系统中的路侧感知与检测提供了一种可行方案。但 RCFTrack 在长周期跟踪任务中存在跟踪中断的问题。后续研究将聚焦于多模态信息的深度挖掘与融合,以提升对远距离目标识别的准确性,并增强长时间跟踪轨迹的稳定性。

参考文献:

References:

- [1] 管丽萍,尹湘源.交通事件管理系统研究现状综述[J].中外公路,2009,29(3):255-261.
GUAN Liping, YIN Xiangyuan. Summary of research status of traffic incident management system[J]. Journal of China & Foreign Highway, 2009, 29(3): 255-261.
- [2] 李薇,蒋响轩,熊睿,等.基于自然驾驶的山区高速公路综合交通风险评价方法[J].公路,2025,70(6):284-292.
LI Wei, JIANG Yunxuan, XIONG Rui, et al. Comprehensive traffic risk evaluation method for

- mountainous expressways based on natural driving[J]. Highway, 2025,70(6):284-292.
- [3] 吴志敏,黄觉,向崎. 基于K-Means聚类算法的高速公路事故多发路段鉴别[J]. 中外公路, 2022, 42(6): 260-264.
WU Zhimin, HUANG Jue, XIANG Qi. Identification of freeway accident-prone sections based on K-Means clustering algorithm[J]. Journal of China & Foreign Highway, 2022, 42(6): 260-264.
- [4] 李海亮. 面向BIM设计的复杂工点三维场景建模方法研究[J]. 中外公路, 2021, 41(6): 306-309.
LI Hailiang. Research on 3D scene modeling method for complex worksites based on BIM design[J]. Journal of China & Foreign Highway, 2021, 41(6): 306-309.
- [5] 虞将苗,唐嘉明,张肖宁,等. 基于三维探地雷达的沥青路面厚度动态调整技术研究[J]. 中外公路, 2020, 40(3): 70-75.
YU Jiangmiao, TANG Jiaming, ZHANG Xiaoning, et al. Research on dynamic adjustment technology of asphalt pavement thickness based on 3D ground penetrating radar [J]. Journal of China & Foreign Highway, 2020, 40(3): 70-75.
- [6] FANG K, XIANG Y, LI X C, et al. Recurrent autoregressive networks for online multi-object tracking [C]// Lake Tahoe: 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), 2018: 466-475.
- [7] 李旭辉,黄奕辉. 基于互联网的公路施工工序质量跟踪系统构建[J]. 公路,2019,64(10):215-218.
LI Xuhui, HUANG Yihui. Construction of highway construction process quality tracking system based on Internet[J]. Highway, 2019, 64(10): 215-218.
- [8] 王孜健,张涵,么新鹏,等. 高速公路二维图形与三维形貌病害数据的构建[J]. 中外公路,2025,45(2):238-246.
WANG Zijian, ZHANG Han, YAO Xinpeng, et al. Construction of disease dataset comprising 2D images and 3D morphologies of highways[J]. Journal of China & Foreign Highway, 2025, 45(2): 238-246.
- [9] BEWLEY A, GE Z Y, OTT L, et al. Simple online and realtime tracking[C]//Phoenix: 2016 IEEE International Conference on Image Processing (ICIP), 2016: 3464-3468.
- [10] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]//Las Vegas: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 779-788.
- [11] 谢海波,朱玮峻,张璧,等. 面向高强度螺栓检测的YOLOv5-Ganomaly联合算法研究[J]. 中外公路, 2024, 44(4): 171-179.
XIE Haibo, ZHU Weijun, ZHANG Bi, et al. Research on YOLOv5-Ganomaly joint algorithm for high-strength bolt detection[J]. Journal of China & Foreign Highway, 2024, 44(4): 171-179.
- [12] WU H, HAN W K, WEN C L, et al. 3D multi-object tracking in point clouds based on prediction confidence-guided data association[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(6): 5668-5677.
- [13] HU H N, CAI Q Z, WANG D Q, et al. Joint monocular 3D vehicle detection and tracking[C]//Seoul:2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019: 5389-5398.
- [14] ZHANG W W, ZHOU H, SUN S Y, et al. Robust multi-modality multi-object tracking[C]//Seoul:2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019: 2365-2374.
- [15] MENG Z L, XIA X, XU R S, et al. HYDRO-3D: Hybrid object detection and tracking for cooperative perception using 3D LiDAR[J]. IEEE Transactions on Intelligent Vehicles, 2023, 8(8): 4069-4080.
- [16] 杨浩冉,李辉,艾晓雪,等. 图像与点云三维体信息交互的3D多目标跟踪网络[J]. 控制与决策, 2024, 39(12): 4127-4135.
YANG Haoran, LI Hui, AI Xiaoxue, et al. 3D multi-object tracking network based on 3D volume information interaction between image and point cloud[J]. Control and Decision, 2024, 39(12): 4127-4135.
- [17] NABATI R, QI H R. CenterFusion:Center-based radar and camera fusion for 3D object detection[C]//Waikoloa:2021 IEEE Winter Conference on Applications of Computer Vision (WACV), 2021: 1526-1535.
- [18] CAESAR H, BANKITI V, LANG A H, et al. NuScenes: A multimodal dataset for autonomous driving[C]//Seattle: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 11621-11631.
- [19] WENG X, WANG J, HELD D, et al. AB3DMOT: A baseline for 3D multi-object tracking and new evaluation metrics [J]. arXiv preprint arXiv, 2020:2008.08063.
- [20] ZHOU X Y, KOLTUN V, KRÄHENBÜHL P. Tracking objects as points[C]//Cham: Computer Vision - ECCV 2020. Springer International Publishing, 2020: 474-490.
- [21] CHAABANE M, ZHANG P, BEVERIDGE J R, et al. Deft: Detection embeddings for tracking[J]. arXiv preprint arXiv, 2021:2102.02267.
- [22] BENBARKA N, SCHRODER J, ZELL A. Score refinement for confidence-based 3D multi-object tracking[C]//Prague: 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2021: 11315-11321.
- [23] KIM A, OŠEP A, LEAL-TAIXÉ L. EagerMOT: 3D multi-object tracking via sensor fusion[C]//Xi'an: 2021 IEEE International conference on Robotics and Automation (ICRA), 2021: 11315-11321.