

DOI: 10.19650/j.cnki.cjsi.J2514374

融合多视角投影与时空拓扑的多无人机 多目标跟踪方法

党昭洋, 孙晓永, 郭润泽, 周沛达, 孙 备
(国防科技大学智能科学学院 长沙 410072)

摘要:为了增强无人机对运动目标的持续跟踪能力,弥补单架无人机在目标跟踪时的性能不足,从无人机自身特点出发,利用多无人机协同感知优势,提出了一种融合多视角投影与目标时空拓扑特征的多机多目标跟踪方法。首先,利用无人机与机载光电吊舱的位置姿态信息而不依赖于图像特征,通过融合无人机位姿与目标高度的一致性约束,实现无人机多视角间的快速投影变换,进行多机动态互补视角下的目标初步关联;其次,利用不同无人机视角下目标之间的时空拓扑特征进行双向关联匹配,并使用空间线索与时间线索对初步关联结果进行细致优化,进一步提高跨视角目标关联精度,提升无人机在遮挡场景下的目标持续跟踪能力;同时,重点针对无人机在拉升、下降、盘旋以及快速运动等多种运动状态下的目标遮挡场景,构建了包含位姿数据的多机多目标跟踪数据集(DP-MDMT),并结合真实任务场景开展相关实验验证。实验结果表明,该方法在DP-MDMT数据集上的关联性能指标召回率(Recall)、精确率(Precision)、多设备目标关联得分(MDA)分别为60.2%、85.6%和47.1%,相较于主流的多机多目标跟踪算法MIA-Net结果分别提升了6.4%、13.1%和7.4%,跟踪性能指标多目标跟踪精度(MOTA)和身份F1分数(IDF1)分别为80.1%和85.1%,且算法平均运行效率为29.7 fps,满足多无人机对地目标跟踪的实时性需求。

关键词:多目标跟踪;多机视觉协同;多视角投影;时空拓扑特征;目标遮挡

中图分类号: TP391.4 TH701 文献标识码: A 国家标准学科分类代码: 510.4050

Multi-object tracking for multi-drone systems integrating multi-view projection and spatiotemporal topology

Dang Zhaoyang, Sun Xiaoyong, Guo Runze, Zhou Peida, Sun Bei

(College of Intelligence Science and Technology, National University of Defense Technology, Changsha 410072, China)

Abstract: To enhance the persistent tracking capability of drones for moving objects and overcome the limitations of single-drone systems, this paper proposes a multi-drone multi-object tracking method that leverages collaborative perception. The approach integrates multi-view projection and the spatiotemporal topology of objects. By utilizing the positional and attitude data of the drones and their onboard photoelectric pods—without relying on image features—rapid projection between views is achieved through a consistency constraint between drone pose and object height. This enables preliminary object association under dynamic, complementary perspectives from multiple drones. Furthermore, bidirectional association matching is performed using the spatiotemporal topological features of objects from different viewpoints. Spatial and temporal cues refine the initial associations, improving cross-view object matching accuracy and enhancing tracking robustness in occluded scenarios. Focusing on occlusions during various drone maneuvers such as climbing, descending, circling, and rapid motion, a dedicated multi-drone multi-object tracking dataset (DP-MDMT) incorporating pose data was constructed. Experiments in real-task scenarios show that the proposed method achieves recall, precision, and multi-device association (MDA) score of 60.2%, 85.6%, and 47.1%, respectively, on the DP-MDMT dataset, representing improvements of 6.4%, 13.1%, and 7.4% over the MIA-Net algorithm. The tracking metrics, including the multiple object tracking accuracy (MOTA) and ID F1-score (IDF1) reach 80.1% and 85.1%, respectively, with an average processing efficiency of 29.7 fps, meeting the real-time requirements for multi-drone ground object tracking.

Keywords: multi-object tracking; multi-drone visual collaboration; multi-view projection; spatiotemporal topological features; object occlusion

0 引言

2024年中国无人机行业市场前景预测研究报告中指出,无人机凭借其体积小、造价低、动作灵活、易于操控、战场生存能力较强等优点,已成为世界各国装备发展的重要方向,被广泛用于农业监测、灾害救援、战场侦察等多个民用以及军用领域^[1]。目标跟踪作为无人机执行任务中的核心环节,能够对高价值空中或地面目标进行识别和跟踪,形成独特的非对称优势,从而大幅提升无人机应用和作战效能。随着人工智能技术的不断发展和无人机任务环境的扩展延伸,基于无人机的多目标跟踪技术具有广阔的空中视野,可捕获更为丰富的信息,进而实现强大的数据收集能力^[2-3],可为进一步的目标意图识别提供有效信息,为目标威胁度评估、袭扰意图等态势识别等提供数据支撑,逐渐成为目前无人机领域的热点研究方向之一。

相较于其他目标跟踪任务,无人机目标跟踪任务环境与任务内容更为复杂多样,视场内的目标或场景易发生快速变化,面临更为严重的目标遮挡、视角变换、运动模糊等目标跟踪难题^[4-6],例如目标或平台的运动、交互、雨、雾、日夜交替等天气变化等,大多数情况下凭借单架无人机难以实现对运动目标的持续精确跟踪。多无人机系统理论上可通过多视角互补解决此问题^[7],但是实际中涉及无人机跨视角实时目标关联的难点问题,具体为:

1) 无人机飞行速度快,不同无人机拍摄视角差异大,基于图像背景特征的跨视角关联匹配困难。由于无人机在执行任务的过程中往往具有高速移动的特征以及显著的视角变化,显著的视角差异使得互补视角下的重叠图像往往存在严重遮挡和剧烈的光照变化^[8],难以提取完备的图像特征。同时,在弱光照或在夜晚条件下无人机拍摄的图像往往分辨率较低具有较多噪点,缺少足够的纹理特征。这都使得多无人机视角间的图像关联匹配更具挑战性。

2) 不同视角下目标外观特征不一致,运动特征差异大,基于目标自身特征的跨视角关联匹配困难。在无人机拍摄的图像中,目标的外观特征由于视角的频繁变化而表现出不一致性,这使得传统的外观和运动特征在无人机互补视角下的目标关联问题中效果有限。当无人机飞行在一定高度时,图像影幅变大,分辨率和清晰度变低,地面上的待跟踪目标尺度变得很小^[9],目标特征和纹理变得稀少,同时在复杂场景中还易受背景噪声干扰,难以提取有效特征。

3) 多无人机协同探测数据基准少,现有方法计算复杂度高,算法实际应用部署困难。目前,针对无人机

动态互补视角下的多目标场景视频协同分析相关研究仍处于起步阶段,仅有的少数研究也未建立针对无人机其特点的相关基准数据集和通用算法框架,现有的多摄像头多目标跟踪算法采用离线跟踪的方式,算法复杂度高,计算资源开销需求大,无法满足无人机具体应用部署的实时性。

本研究针对多无人机互补视角下的多目标时空关联感知问题开展研究,主要工作包括:首先利用无人机位姿信息而不依赖于图像特征,通过融合位置与高度的一致性约束,实现无人机多视角间的快速投影变换。在此基础上,进一步分析初步关联结果,挖掘目标间的时空拓扑特征,并使用空间线索与时间线索进行细致优化。同时,重点针对目标遮挡场景,构建了包含位姿数据的多机多目标跟踪数据集 DP-MDMT。实验结果表明,该方法具有较好的多无人机多目标跟踪性能,实现了可靠的多无人机互补视角下的多目标时空关联感知。

1 多无人机多目标跟踪相关工作

多无人机多目标跟踪 (multi-drone multi-object tracking, MDMOT) 任务的核心在于如何从多个无人机视角中生成一致的多目标跟踪轨迹,无人机目标跟踪典型应用如图 1 所示。相较于 MDMOT,多摄像头多目标跟踪 (multi-camera multi-object tracking, MCMOT) 任务得到了广泛研究^[10-11],MDMOT 与重叠视角^[12]下的 MCMOT 任务类型与任务流程相似,都通过融合多视角下的目标互补信息,解决目标跟踪丢失以及目标特征变化等问题,其唯一区别就在于图像信息的获取方式。该多视角任务的主要难点在于构建不同无人机视角下图像中目标的相关性^[13-14],并设计模型融合多视角互补信息,以提高目标跟踪性能^[15]。



图1 无人机目标跟踪典型应用

Fig. 1 Typical applications of drone object tracking

在多摄像头多目标跟踪任务中,单应性约束方法被广泛应用于提升单摄像头多目标跟踪任务的跟踪精度与鲁棒性。该方法通过将不同视角下的检测跟踪结果投影至统一的地面平面,并寻找多个视角投影的交点,以此作为目标的估计位置。然而,在实际的多摄像头多目标跟踪应用场景中,目标与目标之间可能会发生严重遮挡,导致出现大量的假阳性目标。此前的多摄像头多目标跟踪任务大多针对于监控场景下的行人或车辆目标^[16],为了解决假阳性目标以及行人身高差异所带来的影响,Peng 等^[17]提出了一种多摄像头视角下的贝叶斯网络模型,其核心在于使用预设的行人高度信息,结合现有的检测算法,以生成初步的检测数据。进一步地,模型采用贝叶斯网络对每个摄像头视野内候选目标间的遮挡关系进行建模,并通过单应性投影关系将地面平面与摄像头视角相连,构建出完整的 MvBN 推理结果,并在 PETS09 S2L1^[18], PETS09 City Center^[18], APIDIS^[19]、EPFLTerrace^[20] 这 4 个多摄像头多目标跟踪数据集上进行了实验验证。

近年来,基于 Transformer 的端到端方法在跨摄像头多目标跟踪任务中展现了巨大潜力,尤其在处理目标关联问题方面得到了广泛应用。Hou 等^[21]提出了一种简洁而高效的无锚框特征透视变换网络(MVDet),专用于跨摄像头多目标跟踪任务,并支持端到端的训练流程。该网络接收多个视角的三通道图像作为输入,利用 ResNet 进行特征提取,并通过投影变换将 3D 视角的特征图映射到 2D 平面。进一步地,这些 2D 特征与坐标信息相结合,生成鸟瞰视图(bird's eye view, BEV),从而全面地表示场景信息。通过对 BEV 特征的聚合和卷积操作,MVDet 能够预测行人的位置并给出检测结果。这种方法显著降低了遮挡对跟踪性能的不利影响。然而,MVDet 在摄像头数量较少时的性能可能会下降,且由于其基于离线处理,无法满足实时处理的需求。为应对跨摄像头多目标跟踪任务中的高计算复杂性,You 等^[22]提出了一种名为 DMCT(deep multi-camera tracking)的实时 3D 多目标跟踪方法。该方法设计了一个深度学习网络,旨在估计每个目标在虚拟地平面上的投影位置,并将透视投影效应整合入地平面热图,构建了一个轻量级的深度掠影网络 DGN,用以捕捉人体行为。DMCT 能够同时处理视频流中的多帧图像,其操作方式与人体关键点检测类似,能够在 8 个摄像头之间达到 15 fps 的处理速度。此外,通过 Transformer 的空间级联操作,DMCT 允许不同模态之间的信息超越局部限制进行有效交互。

然而,如图 2 所示,相较于无人机任务场景,多摄像头多目标跟踪算法都是专门针对于静态摄像机设计,在多摄像头多目标跟踪任务场景中,目标尺寸较大且具有显著的外观特征。而在多无人机多目标跟踪任务中,由

于无人机飞行高度较高,待跟踪目标尺寸较小,特征提取困难,且极易出现目标遮挡、移动速度变化等跟踪难点情况,难以提取可靠的外观特征和运动特征,传统的基于固定几何关联、轨迹段匹配或者基于外观特征重识别的 MCMOT 方法,在面对 MDMOT 任务场景下传感器持续运动、视角频繁切换、目标尺度剧烈变化等挑战时难以直接应用,MCMOT 与 MDMOT 任务特性对比如表 1 所示。

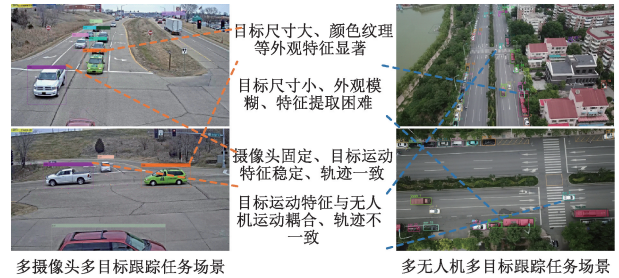


图 2 不同任务场景对比

Fig. 2 Comparison of different task scenarios

表 1 MCMOT 与 MDMOT 任务特性对比
Table 1 Comparison of MCMOT and MDMOT characteristics

对比维度	MCMOT 任务	MDMOT 任务
传感器动态特性	摄像头位置视角固定,背景变化缓慢;可预先进行相机标定,建立相对稳定的几何关系	每架无人机位置、姿态、高度实时快速变化,视角变化频繁且不好预测
视角变化频率	视角变化主要源于目标运动,摄像头视角本身固定,变化缓慢	无人机快速运动,视角变化频繁,同一目标的表现特征在不同无人机视角下差异大
目标尺度变化	目标到摄像头的距离相对稳定,目标尺度变化主要源于其径向运动,变化范围相对可控	高空俯拍导致目标像素分辨率低,无人机与目标的复合运动导致目标尺度在连续帧内变化剧烈

此外,专门针对于多无人机多目标跟踪任务的相关研究和方法目前还在起步阶段,Liu 等^[23]收集并构建了第 1 个多无人机多目标跟踪数据集,并提出了一种名为 MIA-Net 的网络结构,通过全局匹配和局部匹配方法来关联和跟踪多个无人机视角中的目标,该工作为多无人机多目标检测和跟踪领域提供了一个新的基准数据集和基线算法,极大的促进了跨无人机视角下目标关联方法的发展。

2 多无人机多目标跟踪数据集构建

在无人机探测感知任务中,高质量且大规模的数据

集对于开发和推动先进高效的目标感知算法的发展至关重要。近年来,已有大量的基准数据集专门用于单无人机目标检测与跟踪任务的评估与优化,如 VisDrone、UAVDT、DroneCrowd 以及 UAV123 等,而针对多无人机条件下的多目标检测与跟踪任务的数据集还相对较少。因此,本文从无人机目标感知任务的特点和优势出发,构建了一个带有无人机实时位姿数据的多无人机多目标跟踪数据集 DP-MDMT。

该数据集利用两架无人机进行采集,并始终确保其具有重叠视角。为了提高数据集的有效性和多样性,本文在多种光照条件(如晴朗、多云以及夜晚)和不同场景条件(如街道、广场、公路等)下进行数据采集。在飞行过程中,无人机伴有拉升、下降、盘旋以及快速运动等多种运动情况,数据采集时无人机不同运动状态示意如图 3 所示。同时,大部分视频序列均在飞行高度 100 m 以上并使用 1 倍焦距进行拍摄,目标尺寸较小且外观特征相近,使得所采集数据更加贴近于真实任务场景,数据集更为全面且更具有挑战性。

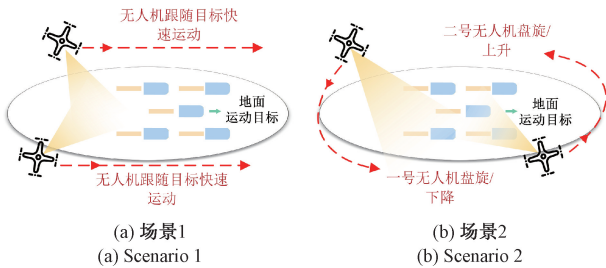


图 3 数据采集时无人机不同运动状态

Fig. 3 Different motion states of the UAV during data collection

部分数据集采集场景如图 4 所示。在对所拍摄的视频数据进行细致筛选和整理后,本文共采集 15 组视频,其中训练集 9 组,验证集 4 组,测试集 2 组,每组视频均由两架无人机从不同的视角和飞行高度采集,每组视频包含两个视频序列,总计 30 段视频序列,23 938 帧高清图,并标记了 95 200 个独立的目标,主要包含汽车、摩托、行人这三类。



图 4 多无人机多目标跟踪数据集采集场景

Fig. 4 Multi-drone multi-object tracking dataset collection scenario

所拍摄视频采用 VATIC 工具进行数据标注,目标标注格式为矩形框(XTL, YTL, XBR, YBR),其中(XTL, YTL)和(XBR, YBR)分别是标注框的两个对角点。同时,对标注数据进行逐帧验证,以确保每个目标被设置为不同的 ID 号,且没有重复的 ID 标签。最后,将标注结果保存在 XML 文件中。

DP-MDMT 着重针对跟踪任务中的目标遮挡等问题进行数据采集,涵盖了多样化的遮挡情境,这些场景中许多目标的部分或完全遮挡极大地增加了目标检测与跟踪的难度,而在另一无人机视角下这些目标则完全不受遮挡,能够被完整拍摄,DP-MDMT 数据集中的遮挡目标补充如图 5 所示。同时,数据集中不仅包含了众多遭受遮挡的目标,还涵盖了行人、汽车与自行车等目标类别,这些目标在体积和移动速率上呈现出显著的差异性,进一步增加了无人机视角下目标关联的复杂度。这种多样性和复杂性为无人机多目标跟踪算法的开发和评估提供了更为真实和具有挑战性的环境。

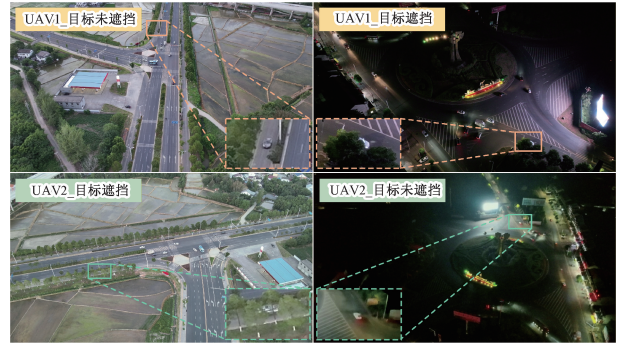


图 5 DP-MDMT 数据集中的遮挡目标补充示意
Fig. 5 Supplementary schematic of occlusion objects in the DP-MDMT dataset

3 无人机多视角快速投影变换

在无人机视角下,往往因远端拍摄、背景遮挡或光照影响等因素使得所拍摄的图像特征不明显或质量较差,如要使用图像特征进行关联匹配则对无人机工作环境以及拍摄条件都具有较高要求。因此,本文从无人机自身特点出发,直接使用无人机位姿信息来实现多视角间的快速投影变换,而不依赖于图像特征。相较于传统固定相机,无人机及其机载吊舱能够直接从机载 GPS 以及惯性导航单元 IMU 中获取无人机航拍时的实时位置坐标与姿态角,能够更加快速有效的实现互补视角下的多目标初步关联。

如图 6 所示,假设两架无人机分别在位置 O_1 和 O_2 对地面场景执行探测任务,且两机之间存在重叠视场,分别获取到图像 I_1 和 I_2 。场景点 $P(X_w, Y_w, Z_w)^T$ 在图像中的投影分别为 $P_1(X_{c_1}, Y_{c_1}, Z_{c_1})^T$ 和 $P_2(X_{c_2}, Y_{c_2}, Z_{c_2})^T$ 。

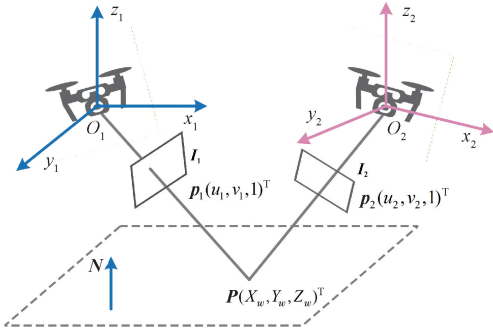


图 6 不同视角下的航拍图像

Fig. 6 Aerial images from different perspectives

单应性矩阵 \mathbf{H} 满足:

$$\mathbf{p}_2 = \lambda \mathbf{H} \mathbf{p}_1 \quad (1)$$

其中, λ 为一个非 0 比例常数。设世界坐标系原点为 O_1 , 相机成像模型为:

$$\mathbf{Z}_{c_1} \begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} = \mathbf{K} \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} \quad (2)$$

其中, \mathbf{K} 是相机的内参矩阵:

$$\mathbf{K} = \begin{pmatrix} f/dx & 0 & u_0 \\ 0 & f/dy & v_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (3)$$

由此可推导世界坐标, 即:

$$\begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} = \mathbf{Z}_{c_1} \mathbf{K}^{-1} \begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} \quad (4)$$

若两无人机 O_1 和 O_2 之间的运动关系为 $\mathbf{R}[I, -\mathbf{t}]$, 其中 \mathbf{R} 为 3×3 的旋转矩阵, \mathbf{t} 为三维平移向量。则有:

$$\mathbf{Z}_{c_2} \begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = \mathbf{K} \mathbf{R} [I, -\mathbf{t}] \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} = \mathbf{K} \mathbf{R} \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} - \mathbf{K} \mathbf{R} \mathbf{t} \quad (5)$$

结合地面平面方程 $\mathbf{n}^T \mathbf{X} = d$, \mathbf{n} 为地面的单位法向量; d 为坐标原点到地面平面的距离, 两个无人机视角间的投影变换矩阵则可表示为:

$$\mathbf{H} = \mathbf{K} \mathbf{R} \left(\mathbf{I} - \frac{\mathbf{t} \mathbf{n}^T}{d} \right) \mathbf{K}^{-1} = \mathbf{K} \mathbf{R} \mathbf{K}^{-1} - \frac{1}{d} \mathbf{K} \mathbf{R} \mathbf{t} \mathbf{n}^T \mathbf{K}^{-1} \quad (6)$$

对于各无人机所在的探测位置 \mathbf{O} 可以用 6 个参数表示, 记为 $\mathbf{O}_i(\varphi_i, \theta_i, \psi_i, x_i, y_i, z_i)$ 。无人机上装有惯性导航单元和 GPS, 可以实时读取并记录无人机拍摄时的姿态角和空间坐标。根据无人机在两个视点间 3 个姿态角的变化量 φ, θ, ψ 和两个视点坐标 (x_1, y_1, z_1) 、 (x_2, y_2, z_2) 就可以确定 \mathbf{R} 和 \mathbf{t} , 即:

$$\mathbf{R} = \mathbf{R}_\varphi \mathbf{R}_\theta \mathbf{R}_\psi \quad (7)$$

$$\mathbf{R}_\varphi = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & \sin \varphi \\ 0 & -\sin \varphi & \cos \varphi \end{pmatrix}$$

$$\mathbf{R}_\theta = \begin{pmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{pmatrix} \quad (8)$$

$$\mathbf{R}_\psi = \begin{pmatrix} \cos \psi & \sin \psi & 0 \\ -\sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{t} = [x_2 - x_1, y_2 - y_1, z_2 - z_1]^T \quad (9)$$

同时, 多视角快速投影变换还需满足目标间的高度一致性约束, 即无人机在视点 (x_1, y_1, z_1) 与 (x_2, y_2, z_2) 下所拍摄图像中的所有场景都需在同一成像深度中。

4 基于时空拓扑特征的双向关联匹配

在通过无人机的位姿数据执行初始的目标关联分析之后, 所获得的初步关联结果, 记为 H_1 。然而, 由于无人机以及机载吊舱位姿测量的固有误差, H_1 中不可避免地包含了一些错误的关联。此外, 地面目标的高密度分布会进一步加剧测量误差对目标关联的不利影响。因此, 本章通过计算跨视角下目标时空拓扑特征之间的映射关系, 进一步提高初步结果的准确性, 目标时空拓扑特征构建如图 7 所示。

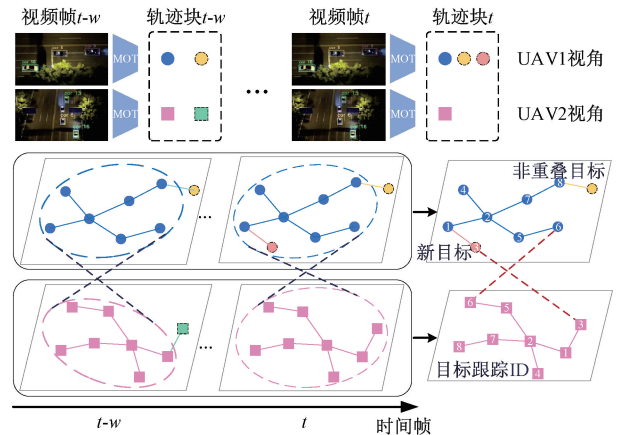


图 7 目标时空拓扑特征构建

Fig. 7 Construction of object spatiotemporal topology features

4.1 目标时空拓扑特征提取

时空拓扑特征在多目标群体中描述了个体目标本身以及与邻近目标在连续视频帧中的空间几何关系与时间动态信息, 能够更好地处理遮挡问题与目标运动轨迹的不确定性。

首先, 无人机拍摄数据 V_i 作为输入传入到目标跟踪

器 OMCTrack^[24] 中,在每个视频中检测并提取潜在目标,生成包含目标 ID 与对应检测边界框的单无人机多目标跟踪结果。选取目标边界框的中心点坐标及其目标 ID 构造单个节点的时空拓扑向量 \mathbf{F}_{it}^k , 其公式为:

$$\mathbf{F}_{it}^k = \{x_{it}^k, y_{it}^k, ID_{it}^k\}, \quad x_{it}^k \in \mathbf{R}; y_{it}^k \in \mathbf{R}; ID_{it}^k \in \mathbf{R} \quad (10)$$

其中, x_{it}^k 与 y_{it}^k 为第 i 个无人机视角下视频帧 t 中节点目标 k 的中心点坐标, ID_{it}^k 为其身份 ID。互补视角 i 与 m 间的变换关系为:

$$\begin{bmatrix} x_{it}^k \\ y_{it}^k \\ 1 \end{bmatrix} = \mathbf{T}_t^{i \rightarrow m} \begin{bmatrix} x_{mt}^j \\ y_{mt}^j \\ 1 \end{bmatrix}, \quad ID_{it}^k = ID_{mt}^j \quad (11)$$

当目标间时空拓扑特征不少于 4 个匹配节点时,可通过式(12)~(13)来求解变换矩阵 $\mathbf{T}_t^{i \rightarrow m}$, 公式如式(12)所示。

$$\mathbf{T}_t^{i \rightarrow m} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \quad (12)$$

$$\begin{cases} x_{mt}^j = \frac{a_{11}x_{it}^k + a_{12}y_{it}^k + a_{13}}{a_{31}x_{it}^k + a_{32}y_{it}^k + 1} \\ y_{mt}^j = \frac{a_{21}x_{it}^k + a_{22}y_{it}^k + a_{23}}{a_{31}x_{it}^k + a_{32}y_{it}^k + 1} \end{cases} \quad (13)$$

此外,虽然目标间的拓扑特征关系在短时间具有一定的稳定性,但在动态开放环境下,目标可能产生不规则运动以及目标新增或消失可能导致某个节点拓扑特征向量发生突变。由此,本文从多目标跟踪器 OMCTrack 中提取目标的时空拓扑特征而非从目标检测器中提取,这是由于目标跟踪器能够对拓扑特征向量进行一定的滤波与预测,计算得到拓扑特征的最优估计,以此得到稳定且适用的时空拓扑特征表示,提升多无人机多目标跟踪任务中目标关联匹配过程的鲁棒性。

4.2 互补视角多目标时空联合优化

1) 空间线索优化

为深入利用目标时空拓扑特征提升匹配精度,本文提出一种双向匹配策略,建立正向与反向双重投影匹配流程,对单向匹配结果进行特征向量补偿以增强匹配完整性。设定 UAV1 视角下提取的目标时空拓扑特征集合为 \mathbf{P} , UAV2 视角下的特征集合为 \mathbf{Q} 。执行前向匹配得到特征向量集合 \mathbf{PQ} , 执行后向匹配得到特征向量集合 \mathbf{QP} 。具体地,从后向关联结果提取特征向量 $\mathbf{q}_s \mathbf{p}_t$, 其中 $1 \leq s \leq n$ 且 $1 \leq t \leq m$, 检查前向特征匹配集合 \mathbf{PQ} 中是否存在对应点对。若缺失,则判定该点对在前向匹配中被遗漏,将其作为有效匹配加入 \mathbf{PQ} , 并利用初步关联矩阵进行约束以剔除错误匹配。该策略能够识别并补偿前

向缺失的目标特征,提升匹配覆盖率和准确性,尤其能够有效应对目标分布不均或特征漏检,增强跨视角关联鲁棒性。空间线索优化方法伪代码见算法 1。

算法 1: 空间线索优化方法伪代码

输入: 目标时空拓扑特征集合 $\mathbf{P} = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n\}$, $\mathbf{Q} = \{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_m\}$, 初步关联矩阵 \mathbf{H}_1
 Step1: 提取前向匹配特征 $\mathbf{PQ} = \{\mathbf{p}_i \mathbf{q}_j, 1 \leq i \leq n \text{ and } 1 \leq j \leq m\}$
 Step2: 提取后向匹配特征 $\mathbf{QP} = \{\mathbf{q}_s \mathbf{p}_t, 1 \leq s \leq m \text{ and } 1 \leq t \leq n\}$
 for $\mathbf{q}_s \mathbf{p}_t$ in \mathbf{QP}
 Step3: 对前向匹配结果集合 \mathbf{PQ} 进行检查
 if $\mathbf{q}_s \mathbf{p}_t$ not in \mathbf{PQ}
 Step4: Add $\mathbf{q}_s \mathbf{p}_t$ to \mathbf{PQ}
 Step5: 剔除错误匹配 ($\mathbf{H}_1, \mathbf{PQ}$)
 end for
 输出: 最终的匹配结果

2) 时间线索优化

在目标关联过程中,由于目标检测能力的限制和以及单应矩阵估计过程中的偏差,可能会得到错误的拓扑映射关系。通过目标间拓扑特征的有效性分析可得,拓扑特征分量值变化微小,目标的拓扑特征具有一定的稳定性。基于此,本文利用时间线索对其进行优化,通过计算相邻时刻单应矩阵的余弦相似度来剔除错误的映射关系,余弦相似度计算方法如式(14)所示。

$$\begin{cases} \cos(\vartheta) = \frac{\sum_{i=1}^n A_i \times B_i}{\|\mathbf{A}\| \times \|\mathbf{B}\|} \\ \|\mathbf{A}\| = \sqrt{\sum_{i=1}^n A_i^2} \\ \|\mathbf{B}\| = \sqrt{\sum_{i=1}^n B_i^2} \end{cases} \quad (14)$$

其中, \mathbf{A} 与 \mathbf{B} 为相邻时刻根据目标时空拓扑特征所得到单应矩阵按行或按列展开的向量, A_i 和 B_i 分别为向量 \mathbf{A} 和 \mathbf{B} 的第 i 个元素。 $\cos(\vartheta)$ 值越接近于 1, 代表两个矩阵越相似, 接近 0, 说明两矩阵正交, 接近负 1 则代表两矩阵相反。

如图 8 所示,为多无人机多目标关联过程中相邻时刻关联矩阵的余弦相似度变化曲线,相邻时刻矩阵之间的余弦相似度大都在 0.999 以上。本文选取 0.99 作为时间线索优化阈值,当 $\cos(\vartheta) < 0.99$ 时,则认为此时刻的关联匹配矩阵得到的为错误的目标拓扑映射关系,对其进行剔除,使用此时刻的初步关联矩阵 \mathbf{H}_1 进行重新关联。通过时间线索优化,能够有效地避免因单应矩阵估

计不准确而导致的关联错误,从而确保了最终结果的准确性和稳定性。

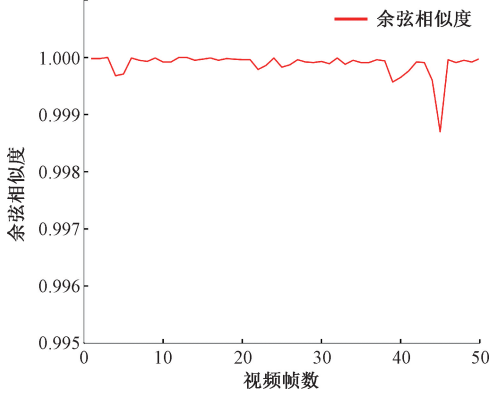


图 8 相邻时刻关联矩阵的余弦相似度变化曲线
Fig. 8 Cosine similarity change curve of adjacent time correlation matrix

3) 目标 ID 动态分配

ID 分配的核心任务是通过变换矩阵对尚未在多无人机系统中进行有效关联的目标的 ID 进行重新分配,根据其特性,这些未匹配目标可主要分为 3 类:新出现的目标、非重叠的目标以及匹配不一致的目标。对于新出现的目标,一是在视频序列中首次出现,二是跟踪丢失后重获 ID 的目标,其关键在于确认其是否在其他视角存在对应项,以便进行匹配或维持其未匹配状态。如式(15)所示,利用变换矩阵 $T_t^{i \rightarrow m}$,可将无人机视角 i 中的特征节点 (x_{it}^k, y_{it}^k) 准确映射无人机视角 m 对应的节点坐标 $(x_{it}^{k'}, y_{it}^{k'})$ 。

$$\begin{bmatrix} x_{it}^{k'} \\ y_{it}^{k'} \\ 1 \end{bmatrix} = T_t^{i \rightarrow m} \begin{bmatrix} x_{it}^k \\ y_{it}^k \\ 1 \end{bmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{bmatrix} x_{it}^k \\ y_{it}^k \\ 1 \end{bmatrix} \quad (15)$$

随后,计算映射坐标 $(x_{it}^{k'}, y_{it}^{k'})$ 与视角 m 中的跟踪目标集合的距离,在满足特定的匹配准则的前提下,将最近的节点关联并分配相同 ID,即:

$$ID_{it}^k = \begin{cases} ID_{m}^j, & (dist(k', j) < SA_{th}) \\ ID_{it}^k, & (dist(k', j) \geq SA_{th}) \end{cases} \quad (16)$$

其中, $j = (x_{m}^j, y_{m}^j)$ 是与 $k' = (x_{it}^{k'}, y_{it}^{k'})$ 匹配的无人机视角 m 下跟踪目标集合中距离最近的节点。在多无人机跟踪探测场景下,目标在不同帧或不同无人机视角下可能呈现出显著的视觉差异, IoU 在多视角变化场景下适应性差,本文采用欧式距离作为匹配度量,以衡量节点目标之间的相似性,从而提高匹配的准确性和鲁棒性。目标 ID 分配流程如图 9 所示。

在执行上述匹配流程后,部分新出现的目标得已成功匹配并为其分配 ID,其余则维持为未匹配状态。在非

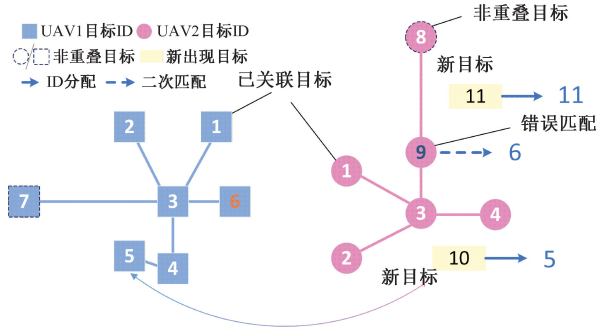


图 9 目标 ID 分配
Fig. 9 Object ID assignment

重叠区域首次出现的目标则被标记为非重叠目标,随着新视频帧的捕获,非重叠目标的 ID 将尝试与现有跟踪目标进行匹配,以形成匹配对,或继续保持未匹配状态。为避免误匹配,本文在式(16)中设定了一个严格的对准阈值 SA_{th} 以确保匹配精度。

此外,被遮挡目标由于其部分或完全被背景或其他目标遮挡,通常难以被检测算识别,且检测置信度往往较低,在跟踪过程中大多被视为误检,并被剔除。然而,当这些目标在两个摄像头视角中同时出现时,其大概率为真实目标。为此,本文将未匹配的目标和低置信度的边界框进行视角映射,与另一个视角中检测到的所有边界框进行匹配。同时,在 ID 分配过程中还应遵循两个约束:一为互斥性,来自同一视角下的目标不能关联在一起;二是唯一性,每个视角下的目标只能分配一个全局目标 ID。当出现多个无人机视角中的目标 (k_1, k_2, \dots) 经过变换矩阵映射到无人机视角 m 后,与视角 m 中的同一目标 j 距离最近且都小于阈值 SA_{th} 的冲突情况时,将遵循“最小距离优先”原则,即仅将目标 ID 分配给那个映射后与目标 j 欧式距离最小的目标。在此基础上,利用目标的时空一致性,实现了互补视角下的多无人机多目标跟踪任务。

5 实验与分析

5.1 实验设置

1) 数据集

为验证所提出的多无人机多目标跟踪算法的有效性,本文在自制数据集 DP-MDMT 以及 MDMT 公开数据集上进行相关验证实验。

2) 实验环境设置

对于多机多目标跟踪网络,本文采用 OMCTrack 网络的多目标跟踪结果作为输入。对于目标关联相关参数,本文将空间对准阈值 SA_{th} 设置为 50。在低置信度目标补充方面,低置信度阈值设置为 $IoU = 0.1$,对于整体

跟踪结果,使用 $IoU = 0.3$ 进行非最大值抑制,以避免产生重复跟踪框。

3) 评价指标

对于最终的多无人机多目标跟踪性能本文采用 IDF1、MOTA、HOTA 等评价指标进行评估。对于多设备下的多目标跟踪算法,需要根据多设备融合结果来评价算法的 ID 关联匹配的准确性,本文采用多设备目标关联得分^[23] (multi-device object association score, MDA) 对多无人机多目标跟踪场景下的目标 ID 关联程度进行评估。

5.2 实验结果与分析

1) 与其他方法的性能对比

为了验证所提出的多无人机多目标跟踪方法的有效性,本节将本文方法与其它多视角多目标跟踪方法在自制数据集 DP-MDMT 以及公开数据集 MDMT 上进行了评测。同时,无人机快速投影变化(以下简称为 RPT)与基于目标时空拓扑特征的双向关联方法(以下简称为 BST)为本文算法的两大核心模块,需要说明是,由于 MDMT 数据集不包含无人机位姿信息,因此在 MDMT 数据集上只对 BST 模块进行性能测试,初始变换矩阵通过读取数据集标签中的目标初始 ID 获得,对于时间线索优化模块,利用 LightGlue^[25] 特征点提取方法对错误估计进行纠正。

(1) 目标关联性能测试

为验证互补视角下多无人机多目标跟踪方法在跨视角目标关联任务中的有效性,本节将本文方法与其他跨视角目标关联方法在 DP-MDMT 以及 MDMT 数据集上进行了对比,被比较的方法包括目标重识别方法^[26] (Re-ID)、基于三角序列的无人机多目标关联方法^[27] (TTS)以及多匹配身份认证网络 (MIA-Net),目标关联性能对比结果如表 2 所示。

表 2 不同方法的目标关联性能对比

Table 2 Comparison of object correlation performance of different methods

数据集	方法	Recall	Precision	MDA	FPS
DP-MDMT	Re-ID	24.3	39.6	13.7	11.3
	TTS	26.5	53.1	15.2	33.5
	MIA-Net	53.8	72.5	40.5	23.6
	本文	60.2	85.6	47.1	29.7
MDMT	Re-ID	27.6	41.5	18.5	11.5
	TTS	26.8	53.7	16.1	33.6
	MIA-Net	55.3	74.1	42.3	23.5
	本文	57.4	80.7	43.8	23.9

在 DP-MDMT 数据集上,本文方法取得了 60.2% 的召回率,85.6% 的准确率以及 47.1% 的多设备关联得分 MDA,并且处理速度仅次于 TTS,达到了 29.7 fps,实现了良好的实时性。在 MDMT 数据集上,本文使用 LightGlue 算法替代了无人机多视角快速投影模块,由于两无人机拍摄视角差异较大,利用空间特征点进行不同视角间的关联匹配性能相对受限,因此目标关联性能与处理速度略有下降。

表 3 显示了不同方法在 DP-MDMT 数据集不同场景序列中的 MDA 得分。在 DP-MDMT 数据集中,序列 SEQ_01 无人机飞行高度较高,目标分布密集且运动速度快,几何分布近似为直线。序列 SEQ_05 无人机拍摄视角持续变化,无人机伴随拉升、下降以及盘旋等运动。序列 SEQ_10 则为夜间场景,图像分辨率较低含有较多噪点,且两架无人机飞行高度不同,视角差异较大。由表 3 可得,在各种复杂动态环境下,本文方法实现了更优的跨视角目标关联性能和更好的鲁棒性。

表 3 不同方法在 DP-MDMT 数据集不同序列中的 MDA 得分

Table 3 The MDA scores of different methods in different sequences of DP-MDMT datasets

方法	序列 SEQ_01	序列 SEQ_07	序列 SEQ_10
Re-ID	2.53	3.45	10.4
TTS	8.67	9.42	12.5
MIA-Net	31.30	30.50	40.9
本文	38.20	46.30	48.2

(2) 多无人机多目标跟踪性能测试

为了验证多无人机多目标跟踪算法的优势,与几种单无人机多目标跟踪算法进行了比较,并使用 MOTA 和 IDF1 分数进行评测,各类方法在 DP-MDMT 数据集上的多目标跟踪性能对比结果如表 4 所示。可以看出,本文所提出的多无人机多目标跟踪算法的 MOTA 与 IDF1 分数都优于其他单无人机目标跟踪算法以及基于 Re-ID 的多无人机多目标跟踪算法,实现了更好的目标跟踪性能。而基于 ReID 的多无人机多目标跟踪算法由于其在无人机拍摄条件下的跨视角目标关联性能不佳,反而会造成关联混淆、ID 频繁切换等问题,使得目标跟踪性能下降,所获得的 MOTA 以及 IDF1 分数低于其基线方法。

2) 可视化结果

为了更好的证明本文方法对于处理遮挡目标的有效性,补充的遮挡目标的可视化结果如图 10 所示,其中的虚线框为局部放大图,在 1 号无人机的视角中,图中的白色卡车被树木遮挡,难以被有效检测。而在 2 号无人机

表 4 不同方法在 DP-MDMT 数据集上的多目标跟踪性能对比

Table 4 Comparison of multi-object tracking performance of different methods on DP-MDMT datasets

数据集	1 号无人机		2 号无人机		整体		MDA
	MOTA	IDF1	MOTA	IDF1	MOTA	IDF1	
DeepSORT ^[28]	73.8	80.2	71.3	77.6	72.6	78.9	
UAVMOT ^[29]	76.2	81.7	73.5	79.3	74.9	80.5	
ByteTrack ^[30]	78.3	84.1	76.6	81.8	77.5	83.0	
BoTSORT ^[31]	75.9	80.9	72.8	78.5	74.4	79.7	
OMCTrack	79.6	85.2	77.3	82.7	78.5	84.0	
Re-ID+OMCTrack	68.3	73.7	68.1	72.5	68.2	73.1	11.3
本文+OMCTrack	80.9	85.8	79.2	84.3	80.1	85.1	47.1

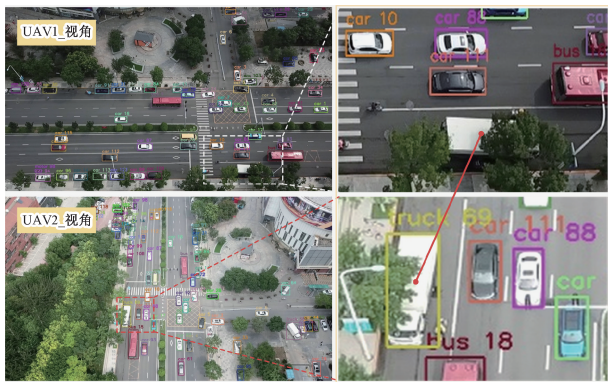


图 10 遮挡目标补充的可视化结果

Fig. 10 Visualization of occlusion Object replenishment

视角中,该目标则可以被有效观测。该可视化结果表明,多视角间的信息融合能够有效地补充被遮挡目标和漏检目标,可以更好地解决单无人机在目标被遮挡时由于单无人机视角地局限性而造成地跟踪丢失的问题。

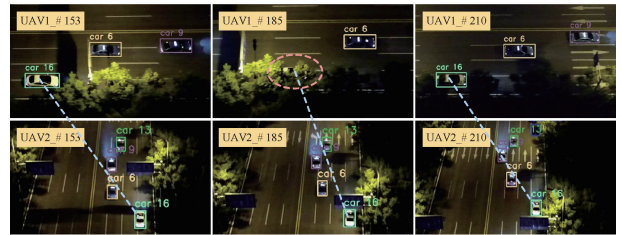
单无人机多目标跟踪算法与多无人机多目标跟踪算法的可视化结果对比如图 11(a)和(b)所示,在本文所提出的单无人机多目标跟踪算法中,白车的初始 ID 为 2,从第 153~210 帧经历了路边树木的长时遮挡,当目标在第 210 帧重新进入视野中时,目标的跟踪 ID 从 2 变为了 32,发生了 ID 切换。在多目标跟踪角度,第 153 与 210 帧的白色车辆目标被认定为了不同的两个目标。

相比之下,在多无人机多目标跟踪策略中,白车目标的初始 ID 为 16,即使经历了长达接近 60 帧的长时遮挡,当目标重新出现时仍能保持其 ID 不变,具有鲁棒、稳定的多目标跟踪性能。由此可以看出,本文所提出的多无人机多目标跟踪方法能够有效利用多个无人机的不同视角对目标进行同时跟踪,能够有效的解决当目标经历长时遮挡情况下跟踪失败的问题,提升复杂动态环境下的无人机多目标跟踪能力。

如图 12 所示为该目标在各视角下的目标状态及其



(a) 单无人机多目标跟踪结果
(a) Single UAV multi-object tracking results



(b) 多无人机多目标跟踪结果
(b) Multi-UAV multi-object tracking results

图 11 单机与多机目标跟踪算法的结果对比

Fig. 11 Comparison of results of single UAV and multi-UAV object tracking algorithms

跟踪结果对比,其中虚线代表目标经历遮挡,不同线宽代表不同的目标跟踪 ID。从图 12 中可以直观的看出,在单无人机视角下,目标经历遮挡后,目标的跟踪 ID 发生了显著变化,使得跟踪失败。而对于多无人机互补视角,能够对遮挡目标能够进行有效补充,极大的解决了无人机跟踪过程中的目标遮挡问题。

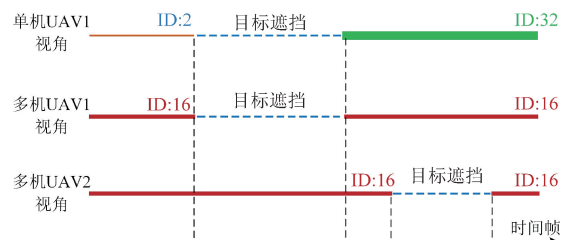


图 12 各视角下目标状态以及跟踪结果对比

Fig. 12 Comparison of object status and tracking results from each perspective

为了更直观地展示多无人机多目标跟踪方法的有效性,在MDMT数据集和DP-MDMT数据集上绘制了跟踪结果,如图13和14所示。可视化结果表明本文所提出

多无人机多目标跟踪方法能够很好的适应各种复杂动态环境,可以有效改善由于目标遮挡造成的跟踪损失,提高多目标跟踪性能。



图 13 MDMT 数据集结果可视化

Fig. 13 Visualization of MDMT dataset results



图 14 DP-MDMT 数据集结果可视化

Fig. 14 Visualization of DP-MDMT dataset results

3) 消融实验

(1) 各模块有效性分析

为了更好的分析本文方法的性能,在 DP-MDMT 数据集上进行了一系列消融实验,以验证其中每个模块的有效性。在消融实验中使用 OMCTrack 算法作为基线模型。表 5 为添加不同模块时算法性能对比,Baseline 的 UAV1 以及 UAV2 视角下的 MOTA 和 IDF1 指标为 79.6%、85.2%、77.3% 和 82.7%,添加无人机多视角快速投影方法(RPT)后,UAV1 以及 UAV2 视角下的 MOTA 和 IDF1 指标均分别有所提高,相较于 Baseline 分别提高了 0.5%、0.1%、0.6% 和 0.5%。融合无人机多视角快速投影变换和基于时空拓扑特征的双向关联匹配方法(BST)后的性能达到最优,UAV1 以及 UAV2 视角下的 MOTA 和 IDF1 指标分别达到了 80.9%、85.8%、79.2% 和 84.3%。由此可得,本文提出的无人机快速投影变换方法和基于目标时空拓扑特征的双向关联方法都对目标跟踪性能带来了提升,两者融合后能进一步提高跟踪性能。

表 5 添加不同模块时算法性能分析

Table 5 Algorithm performance analysis when adding different modules

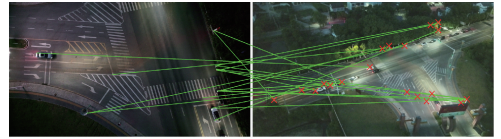
基线方法	RPT 模块	BST 模块	1 号无人机		2 号无人机		MDA
			MOTA	IDF1	MOTA	IDF1	
√			79.6	85.2	77.3	82.7	-
√	√		80.1	85.3	77.9	83.2	40.3
√	√	√	80.9	85.8	79.2	84.3	47.1

由于无人机以及光电吊舱的位姿信息的测量值存在一定误差,单纯基于无人机快速投影方法进行跨视角间的目标关联可能会产生错误的目标 ID 分配结果,尤其是当目标密集分布时,因此,有时对于无人机多目标跟踪性能的提升有限。融合目标间的时空拓扑特征进行重新目标关联,能够有效的消除由于位姿信息测量误差所带来的错误估计,实现了更好的跨视角目标关联效果和更优的多无人机多目标跟踪性能。

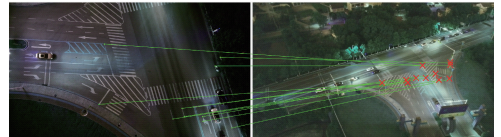
(2) 无人机多视角快速投影方法性能对比

本小节将无人机多视角快速投影方法与经典的基于空间特征点匹配的方法进行了性能对比,具体可视化结果如图 15 所示。由于两无人机视角差异巨大,且存在大量背景相似目标,难以提取到可靠的图像特征点。如图 15(a) 和 (b) 所示,基于 ORB 特征^[32] 以及基于 LightGlue 方法的关联结果绝大部分为错误匹配。对于无人机多视角快速投影变换方法,为了可视化目标关联结果,在处理过程中,本文给出了 1 号无人机视角下的目标源点,并利用所得投影变换矩阵对关联结果进行了绘制。

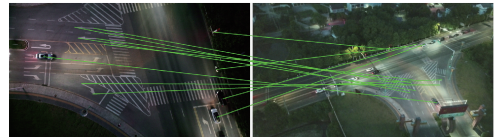
图 15(c) 中的结果表明,本文方法基本实现了精确的跨视角多目标关联匹配,避免了由于图像特征不可靠造成的误匹配问题。



(a) 基于ORB特征的无人机互补视角关联匹配结果
(a) Complementary perspective correlation matching results of UAV based on ORB features



(b) 基于LightGlue的无人机互补视角关联匹配结果
(b) Correlation and matching results of UAV complementary perspectives based on LightGlue



(c) 基于无人机多视角快速投影变换关联匹配结果
(c) Association and matching results based on fast projection transformation of multi-UAV multi-view

图 15 无人机多视角快速投影方法性能对比

Fig. 15 Comparison of performance of multi-view fast projection methods for UAV

(3) 空间对准阈值参数设置

在多无人机多视角关联过程中,通过设定最大空间投影误差阈值,可以判定两个局部目标在空间维度上是否满足关联条件,此阈值为时空拓扑特征关联环节的关键空间参数,其参数值直接影响系统的整体性能。为此,进行了一系列实验,对不同空间对准阈值参数对目标关联以及目标跟踪任务的性能影响进行了深入分析,系统的调整了其从 30 ~ 80 pixels,图 16(a) 和 (b) 体现了在不同阈值下 MOTA、IDS 以及 MDA 性能指标的变化曲线。实验结果表明,随着阈值不断增大,目标关联得分指标 MDA 不断提升。然而,针对于目标跟踪任务的相关性能指标,在空间对准阈值介于 30 ~ 55 pixels,算法的跟踪性能整体维持在较高水平,MOTA 指标整体呈上升趋势。同时,IDS 不断下降,代表算法在该阈值区间内能够较好的实现跨视角间的目标关联匹配,能够有效减少因目标遮挡等因素所产生的目标 ID 切换等现象,但随着阈值的不断增大,使得更多本不应匹配的局部目标对被错误地认定为匹配,不可避免地导致了目标跟踪 ID 的不稳定性,从而对算法的目标跟踪性能产生了负面影响。

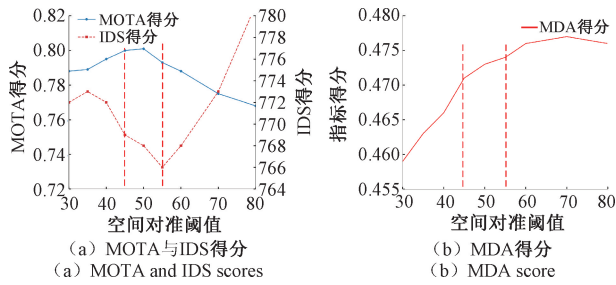


图 16 不同空间投影误差阈值对算法性能的影响

Fig. 16 The influence of different spatial projection error thresholds on the performance of the algorithm

基于上述实验分析,本文确定了空间对准阈值的适宜取值范围为45~55 pixels,在此区间内,算法能够在保持较高匹配准确性的同时,有效控制IDS的数量,确保了跨摄像头目标关联的稳定性和准确性,提升了多无人机多目标跟踪性能。

6 结 论

针对复杂动态环境下无人机目标跟踪易受目标遮挡、光照变化、运动模糊等问题影响,单靠优化单平台目标感知算法难以解决极端条件下的现实问题,本文从无人机自身特点出发,提出了适用于多无人机动态互补视角下的多目标关联感知算法。首先,提出了一种不依赖于图像特征的无人机多视角快速投影变换方法,通过实时读取记录无人机自身位姿参数,融合无人机位姿与目标高度一致性约束,计算无人机互补视角间的投影变换矩阵,进行目标间初步关联匹配。在此基础上,利用不同视角下目标之间的时空拓扑特征,使用空间线索与时间线索对初步关联结果进行细致优化,实现了更好的多无人机多目标跟踪性能。结果表明,本文算法能够在复杂动态环境下实现稳定鲁棒的互补视角下多目标精确关联匹配,能够更好的解决目标在遮挡、视野进出等场景下的丢失恢复问题,提高了多无人机多目标跟踪性能。

未来的工作将重点围绕以多运动目标场景为中心的多模态、多协同等方面展开,一方面,进一步探索跨视角跨模态条件下的多机多目标跟踪方法,提升无人机全天候、多场景下的侦察探测能力;另一方面,进一步研究基于混合视角下的多无人机多目标跟踪方法,使其不仅能够互补视角能够有效跟踪,在非重叠视角的场景下也能够有良好的多目标跟踪性能,进一步提升算法在不同任务场景中的适用性。

参考文献

[1] 杨洋,宋品德,钟春来,等. 无人机视角下基于深度学习的多目标跟踪研究进展[J]. 计算机工程与应用

用, 2023, 59(23): 48-62.

YANG Y, SONG P D, ZHONG CH L, et al. Research progress of multi-target tracking based on deep learning from perspective of UAV[J]. Computer Engineering and Applications, 2023, 59(23): 48-62.

[2] 苑玉彬,吴一全,赵朗月,等. 基于深度学习的无人机航拍视频多目标检测与跟踪研究进展[J]. 航空学报, 2023, 44(18): 6-36.

YUAN Y B, WU Y Q, ZHAO L Y, et al. Research progress of UAV aerial video multi-object detection and tracking based on deep learning[J]. Acta Aeronautica et Astronautica Sinica, 2023, 44(18): 6-36.

[3] WU P N, LI Y X, XUE D. UAV target tracking: A survey[J]. Artificial Intelligence Review, 2025, 58(11): 1-62.

[4] 王指辉,周嘉麒,廖万斌,等. 基于分散注意力机制的无人机目标跟踪算法[J]. 计算机与数字工程, 2024, 52(5): 1287-1292.

WANG ZH H, ZHOU J Q, LIAO W B, et al. Target tracking algorithm for UAV based on split attention[J]. Computer & Digital Engineering, 2024, 52(5): 1287-1292.

[5] 范云生,张凯,牛龙辉,等. 海上无人机对运动船舶的长期检测跟踪算法[J]. 仪器仪表学报, 2024, 45(3): 326-335.

FAN Y SH, ZHANG K, NIU L H, et al. Long-term detection and tracking algorithm for moving vessels by maritime UAVs[J]. Chinese Journal of Scientific Instrument, 2024, 45(3): 326-335.

[6] 王文胜,何君尧,黄民,等. 融合梯度改进YOLO和KCF模型的无人机目标识别跟踪算法[J]. 仪器仪表学报, 2025, 46(2): 221-233.

WANG W SH, HE J Y, HUANG M, et al. Fusion gradient improved YOLO and KCF models for UAV target recognition and tracking algorithm[J]. Chinese Journal of Scientific Instrument, 2025, 46(2): 221-233.

[7] 崔阳洁,屠展,杨彬淇,等. 无人机遮挡目标检测与协同跟踪方法[J]. 机器人, 2025, 47(3): 427-437.

CUI Y J, TU ZH, YANG B Q, et al. Occluded target detection and multi-UAV cooperative tracking method[J]. Robot, 2025, 47(3): 427-437.

[8] YAN B B, WEI Y X, LIU SH X, et al. A review of current studies on the unmanned aerial vehicle-based

- moving target tracking methods[J]. *Defence Technology*, 2025, 51: 201-219.
- [9] 陶海波, 冯瑞娜, 阙启正, 等. 复杂背景大尺度变化的无人机目标跟踪[J]. *南昌航空大学学报(自然科学版)*, 2025, 39(1): 39-49.
- TAO H B, FENG R N, QUE Q ZH, et al. Object tracking for UAV with scale changes and complex background[J]. *Journal of Nanchang Hangkong University: Natural Sciences*, 2025, 39(1): 39-49.
- [10] 伍瀚, 孙浩, 计科峰, 等. 时序信息引导跨视角特征融合的多无人机多目标跟踪方法[J]. *电子学报*, 2025, 53(3): 728-743.
- WU H, SUN H, JI K F, et al. Temporal-guided cross-view feature fusion network for multi-drone multi-object tracking[J]. *Acta Electronica Sinica*, 2025, 53(3): 728-743.
- [11] 乔羽, 范慧杰, 付生鹏, 等. 高空无人机跨场景多目标跟踪方法[J]. *小型微型计算机系统*, 2025, 46(8): 1993-1999.
- QIAO Y, FAN H J, FU SH P, et al. High-altitude unmanned aerial vehicle multi-target tracking across scenes[J]. *Journal of Chinese Computer Systems*, 2025, 46(8): 1993-1999.
- [12] AMOSA T I, SEBASTIAN P, IZHAR L I, et al. Multi-camera multi-object tracking: A review of current trends and future advances[J]. *Neurocomputing*, 2023, 552: 126558.
- [13] 李洪瑶, 李小强, 韩心中, 等. 基于决策融合的多无人机协同目标检测识别算法[J]. *系统工程与电子技术*, 2022, 44(3): 746-754.
- LI H Y, LI X Q, HAN X ZH, et al. Cooperative object detection and recognition algorithm for multiple UAVs based on decision fusion[J]. *Systems Engineering and Electronics*, 2022, 44(3): 746-754.
- [14] HAN R Z, FENG W, ZHANG Y J, et al. Multiple human association and tracking from egocentric and complementary top views[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 44(9): 5225-5242.
- [15] WU H, SUN H, JI K F, et al. Temporal-spatial feature interaction network for multi-drone multi-object tracking[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2025, 35(2): 1165-1179.
- [16] 周翰祺, 方东旭, 张宁波, 等. 基于深度学习的多无人机多目标跟踪[J]. *计算机工程*, 2025, 51(4): 57-65.
- ZHOU H Q, FANG D X, ZHANG N B, et al. Multi-UAV multi-object tracking based on deep learning[J]. *Computer Engineering*, 2025, 51(4): 57-65.
- [17] PENG P X, TIAN Y H, WANG Y W, et al. Robust multiple cameras pedestrian detection with multi-view Bayesian network[J]. *Pattern Recognition*, 2015, 48(5): 1760-1772.
- [18] FERRYMAN J, SHAHROKNI A. *Pets2009: Dataset and challenge* [C]. 2009 Twelfth IEEE International Workshop on Performance Evaluation of Tracking and Surveillance. *IEEE*, 2009: 1-6.
- [19] DE VLEESCHOUWER C, CHEN F, DELANNAY D, et al. Distributed video acquisition and annotation for sport-event summarization [J]. *NEM Summit*, 2008, 8(10.1016).
- [20] FLEURET F, BERCLAZ J, LENGAGNE R, et al. Multicamera people tracking with a probabilistic occupancy map[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, 30(2): 267-282.
- [21] HOU Y ZH, ZHENG L, GOULD S. Multiview detection with feature perspective transformation [C]. *Computer Vision-ECCV 2020: Part VII*, 2020: 1-18.
- [22] YOU Q Z, JIANG H. Real-time 3D deep multi-camera tracking[J]. *ArXiv preprint arXiv:2003.11753*, 2020.
- [23] LIU ZH H, SHANG Y Y, LI T M, et al. Robust multi-drone multi-target tracking to resolve target occlusion: A benchmark [J]. *IEEE Transactions on Multimedia*, 2023, 25: 1462-1476.
- [24] DANG ZH Y, SUN X Y, SUN B, et al. OMCTrack: Integrating occlusion perception and motion compensation for UAV multi-object tracking [J]. *Drones*, 2024, 8(9): 480.
- [25] LINDENBERGER P, SARLIN P E, POLLEFEYS M. Lightglue: Local feature matching at light speed [C]. 2023 IEEE/CVF International Conference on Computer Vision, 2023: 17627-17638.
- [26] ZHOU K Y, YANG Y X, CAVALLARO A, et al. Omniscale feature learning for person re-identification [C]. 2019 IEEE/CVF International Conference on Computer

Vision, 2019: 3702-3712.

- [27] LI X D, WU L ZH, NIU Y F, et al. Multi-target association for UAVs based on triangular topological sequence[J]. Drones, 2022, 6(5): 119.
- [28] WOJKE N, BEWLEY A, PAULUS D. Simple online and realtime tracking with a deep association metric [C]. 2017 IEEE International Conference on Image Processing, 2017: 3645-3649.
- [29] LIU SH, LI X, LU H CH, et al. Multi-object tracking meets moving UAV[C]. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 8876-8885.
- [30] ZHANG Y F, SUN P Z, JIANG Y, et al. Bytetrack: Multi-object tracking by associating every detection box [C]. Computer Vision-ECCV 2022: 17th European Conference, 2022: 1-21.
- [31] AHARON N, ORFAIG R, BOBROVSKY B Z. BoT-SORT: Robust associations multi-pedestrian tracking[J]. ArXiv preprint arXiv: 2206.14651, 2022.
- [32] RUBLEE E, RABAUD V, KONOLIGE K, et al. ORB: An efficient alternative to SIFT or SURF [C]. 2011

International Conference on Computer Vision, 2011: 2564-2571.

作者简介



党昭洋, 2024 年于国防科技大学获得硕士学位, 现于国防科技大学攻读博士学位, 主要研究方向为无人平台智能感知。

E-mail: dzy0329@nudt.edu.cn

Dang Zhaoyang received his M. Sc. degree from the National University of Defense Technology in 2024. He is currently pursuing a doctorate at the National University of Defense Technology. His main research interest includes intelligent perception of unmanned platforms.



孙晓永(通信作者), 2020 年于国防科技大学获得博士学位, 现为国防科技大学助理研究员, 主要研究方向为无人平台智能侦测。

E-mail: sunxiaoyong14@nudt.edu.cn

Sun Xiaoyong (Corresponding author) received his Ph. D. degree from National University of Defense Technology in 2020. He is currently an assistant researcher at the National University of Defense Technology. His main research interest includes the unmanned platform intelligent detection.