

基于视觉的无人机定位与导航方法研究综述

谷美颖^{1,2}, 李 航^{1,2}, 张家伟^{1,2}, 百 晓^{1,2*}, 郑 锦^{1,2}

(1. 北京航空航天大学计算机学院, 北京 100191; 2. 北京航空航天大学江西研究院, 江西南昌 330000)

摘要: 随着无人机(Unmanned Aerial Vehicle, UAV)成本的降低, 无人机引起了越来越多的研究兴趣. 其应用领域广泛, 包括农业、消防、测绘、航拍以及娱乐应用. 这些应用需要无人机在精准的自我定位下进行自主飞行, 通常高度依赖于全球导航卫星系统(Global Navigation Satellite System, GNSS). 然而, GNSS在长距离无线电通信方面存在多种缺陷(如非视距接收、多路径效应、欺骗信号), 这推动了补充或取代卫星导航新方法的发展. 基于视觉的无人机定位与导航方法利用无人机搭载的视觉传感器, 实现自主定位与导航, 成为解决这一问题的重要途径. 本文的贡献在于系统性地梳理了基于视觉的无人机定位与导航技术, 全面总结了该领域的研究现状和发展趋势. 首先, 介绍了无人机视觉定位的方法, 主要分为图像检索和图像匹配两类, 并对其技术特点、适用场景以及相关数据集和评价指标进行了分析. 其次, 详细阐述了无人机视觉导航的方法, 根据导航功能的不同分为障碍物检测与规避方法以及路径规划方法, 揭示了现有技术的优势和局限. 最后, 进一步讨论了基于视觉的无人机定位与导航方法在公共可用数据集、硬件加速、环境复杂性、实时性要求、能源限制以及模拟器到真实世界的泛化等方面可能面临的挑战.

关键词: 无人机; 视觉定位; 视觉导航; 机器视觉; 图像匹配

基金项目: 国家自然科学基金(No.62276016, No.62372029)

中图分类号: TP391.4; V279

文献标识码: A

文章编号: 0372-2112(2025)03-0651-35

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20240699

A Review of Vision-Based UAV Localization and Navigation Methods

GU Mei-ying^{1,2}, LI Hang^{1,2}, ZHANG Jia-wei^{1,2}, BAI Xiao^{1,2*}, ZHENG Jin^{1,2}

(1. School of Computer, Beihang University, Beijing 100191, China;

2. Jiangxi Research Institute, Beihang University, Nanchang, Jiangxi 330000, China)

Abstract: As the cost of unmanned aerial vehicles (UAVs) decreases, they have attracted increasing research interest. UAVs are now widely applied in various fields, including agriculture, firefighting, surveying, aerial photography, and recreational applications. These applications require UAVs to perform autonomous flights with precise self-localization, typically relying heavily on global navigation satellite systems (GNSS). However, GNSS has multiple shortcomings related to long-distance radio communications, such as non-line-of-sight reception, multi-path effects, and spoofing. This has driven the development of new methods to supplement or replace satellite navigation. Vision-based UAV localization and navigation methods, utilizing onboard visual sensors for autonomous localization and navigation, have become crucial in addressing this issue. This review contributes to the field by systematically reviewing vision-based UAV localization and navigation technologies, providing a comprehensive summary of the current research landscape and developmental trends. First, it introduces vision-based UAV localization methods, which are categorized into image retrieval and feature matching approaches. The technical characteristics, applicable scenarios, relevant datasets, and evaluation metrics of these methods are analyzed in detail. Second, this review elaborates on vision-based UAV navigation methods, distinguishing between obstacle detection and avoidance techniques and path planning methods based on their functional objectives, while highlighting the strengths and limitations of existing technologies. Finally, this review further discusses the potential challenges faced by vision-based UAV localization and navigation methods, including the lack of publicly available datasets, the need for hardware acceleration, the complexity of operating environments, real-time processing requirements, energy constraints, and the gap between simulated and real-world environments.

Key words: unmanned aerial vehicle (UAV); vision-based localization; vision-based navigation; machine vision; image matching

Foundation Item(s): National Natural Science Foundation of China (No.62276016, No.62372029)

1 引言

无人机的便携性、机动性、灵活性,为其在可能包含意外障碍物的复杂三维环境中提供了飞行能力.随着GPS、相机和雷达等高精度传感器的飞速发展,无人机具备了定位^[1]、目标检测^[2-4]、自主着陆^[5]和避障^[6,7]的能力,现已成为军事^[8]、农业监测^[9]、交通监管^[10]、测绘^[11,12]、消防^[13]以及搜索与救援^[14]等多个领域的重要工具.在农业领域,无人机能够高效地进行作物监测和病虫害预防;在测绘领域,无人机的使用大大提高了地形测绘的效率和精度;在物流行业,无人机正逐渐成为快递配送的新选择,特别是在偏远或难以到达的地区;在紧急救援中,无人机能够迅速穿越难以通行的地带,进行搜救行动和灾害评估.

尽管无人机的应用前景广阔,但在其执行任务时往往面临着复杂多变的环境和技术挑战.其中,定位和导航是实现无人机高效、安全运行的核心技术.在开阔地区,无人机通常依赖全球定位系统(Global Positioning System, GPS)来确定位置和航向.然而,在城市峡谷、室内、地下或其他全球导航卫星系统(Global Navigation Satellite System, GNSS)信号受限的环境中,GPS信号的丢失或干扰会严重影响无人机的定位精度和操作安全.此外,某些任务需求下,无人机可能还会遭受到人为的信号干扰,这对无人机的定位系统提出了更高的要求.

在这些GNSS拒止环境下,视觉定位与导航技术成为了解决这一问题的关键.如图1所示,通过无人机搭载的视觉传感器(如摄像头)捕捉的实时图像,结合惯性传感测量单元的陀螺仪和加速度计,无人机能够确认自身定位,并通过识别和分析周围环境,进行避障和路径规划,实现在无GPS信号的情况下自主定位与导航.这种技术不仅提高了无人机的操作灵活性和任务适应性,也极大地拓展了其应用场景.然而,视觉定位

与导航技术仍面临许多挑战,包括如何在低光照或直射阳光下有效工作,如何处理动态场景中的飞行障碍,以及如何提高算法的实时性和鲁棒性等等.为了克服这些难题,国内外学者在无人机视觉系统的研究中不断探索和创新,推动了相关技术的发展和完善.

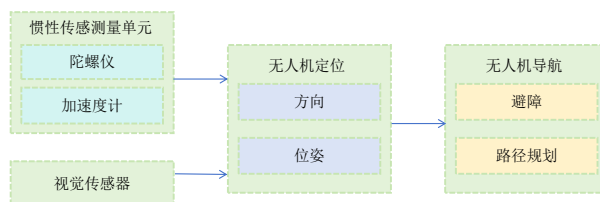


图1 无人机视觉定位与导航流程图

为了进一步提高基于视觉的无人机定位与导航方法的实时性和鲁棒性,使无人机能够实现GNSS拒止环境下的自主定位与导航,本文调研了相关领域研究并对该领域进行了系统的梳理与总结.尽管目前已有无人机定位方法^[1,15]或导航方法^[16-18]的综述(见表1),但是此类工作主要关注传统的方法以及相关理论的介绍,对于近年来关注度高的深度学习方法调研不足.朱得糠等人^[1]系统性地介绍了无人机视觉地理定位过程中的坐标系定义及相互转换关系,并对基于辅助地理信息、距离测量、三角定位和角度跟踪滤波等四类技术方法进行了详尽分析.然而,该综述主要围绕传统方法展开,未涵盖基于深度学习的最新方法,无法反映近年来的技术进展.作为第一篇无人机定位综述,Couturier等人^[15]聚焦于2015至2020年间的视觉定位技术.尽管该文中包含了一些基于深度学习的初步方法,但覆盖范围缺乏近几年的技术,尤其是近年来迅速发展的基于图像检索与图像匹配的定位方法.文献[16]主要综述了基于SLAM(Simultaneous Localization And Mapping)

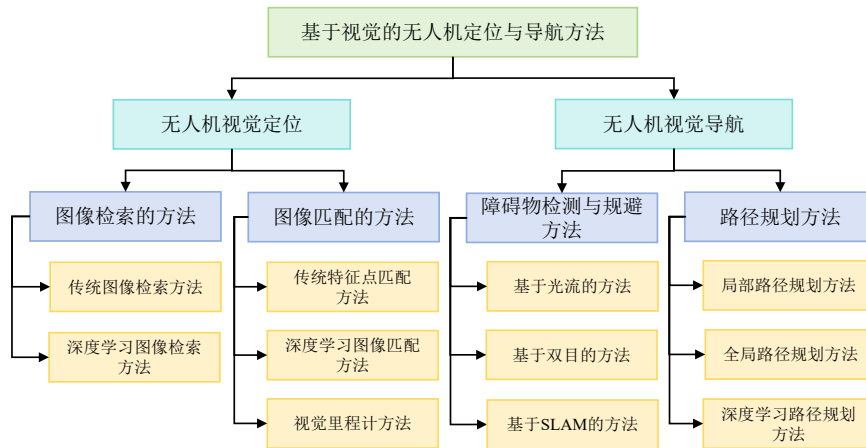
表1 无人机定位与导航现有综述对比

综述	年份	定位	导航	关注侧重点	主要贡献
文献[1]	2023	√	—	视觉地理定位	全面综述了无人机视觉地理定位的技术原理包括各类坐标系定义及坐标转换关系
文献[15]	2020	√	—	绝对视觉定位	第一篇无人机视觉定位综述,比较了基于轨迹长度、环境大小和定位误差的定位方法
文献[16]	2016	—	√	视觉与IMU融合的导航	总结了基于视觉与IMU融合、SLAM的导航方法,得出基于视觉与IMU融合的方法比SLAM可以减少内存和功率的需求
文献[17]	2018	—	√	视觉导航	综述了视觉导航的重要部分,包括SLAM、避障和路径规划
文献[18]	2022	—	√	GPS拒止环境导航	综述了使用多种传感器的完整导航策略,包括感知、定位和运动规划等多个子组成部分
本文	—	√	√	无人机视觉定位与导航	综述了最新无人机视觉定位与导航方法,包括传统方法与深度学习方法,综述了该领域的公开数据集、评价指标以及结果分析,从算法效率、精度、计算复杂度、适用场景等多方面深入对比

和VO(Visual Odometry)的导航方法.文献[17]和文献[18]虽然讨论到了无人机导航中的避障与路径规划等技术,但是未能涵盖基于深度学习的最新研究进展.与上述综述不同,本综述聚焦于GNSS拒止环境下的视觉定位与导航问题,兼顾了定位与导航技术.针对拒止环境下无人机自主飞行面临的复杂需求,本文特别关注了近年来基于深度学习的最新技术进展,包括图像检索、特征匹配、双目避障及强化学习等方法.本文还对这些方法在不同场景中的定量和定性性能进行了系统性梳理和分析,综合比较了不同方法在多个公开数据集上的优缺点,揭示了其适用性、局限性及优化方向.此外,本文深入探讨了无人机视觉定位与导航领域

所面临的核心挑战,例如数据集多样性不足、实时性要求、能源限制以及复杂环境的适应性,并展望了未来研究的关键方向.通过对现有方法进行全面回顾与多维度分析,本综述旨在为研究人员提供深入理解和实践指导,为未来的发展提供参考.

本文介绍了无人机视觉定位方法,并将现在的无人机视觉定位算法分为基于图像检索的方法、基于图像匹配的方法;介绍了无人机视觉导航方法,并根据导航功能的不同将其分为了障碍物检测与规避方法以及路径规划方法.分类法如图2所示.本文对现有无人机视觉定位与导航进行总结,并提供了未来可能的研究方向,为后续研究提供了参考.



2 无人机视觉定位

无人机在空中飞行,首先需要解决自身定位的问题.无人机定位领域的早期工作是使用全局导航卫星系统,如全球定位系统,结合惯性导航系统(Inertial Navigation System,INS)进行位姿估计,这些工作使用卡尔曼滤波器^[19,20]和扩展卡尔曼滤波器(Extended Kalman Filter,EKF)^[21]进行GNSS-INS融合.GNSS-INS系统主要由GPS、加速度计、陀螺仪和磁力计组成,这些传感器成本可控并且十分轻量.GNSS长距离定位精度高,而INS短距离定位精度高,互补的特性使得GNSS-INS系统具备相对较高的定位性能.但是,由于GNSS是通过测量无线电信号^[22]的飞行时间来工作的,无人机和卫星之间的任何障碍都可能影响定位精度.因此,一些学者开始关注室内环境和城市峡谷、森林等多个室外环境下,GNSS的无线电信号接收问题.除此之外,无意干扰和出于某些目的的有意干扰也会诱导信号产生错误,从而影响定位精度,例如当多个反射信号干扰真实信号的接收时.位置估计中的诱导误差通常在几十米以内,但在一些高楼的城市环境中可以超过1 km.

基于视觉的和视觉辅助定位是取代或补充GNSS-INS融合的最主要解决方案.这种方法的主要挑战在于大量的数据分析和解释这些数据的复杂性.在进行定位之前,需要利用先前收集的数据(参考数据)来进行无人机定位.参考数据可能由一组松散的影像或拼接成马赛克的航空影像组成,这些航空影像大多是正射校正的卫星影像.随着像Google Earth这样的免费可用地图系统和像ArcGIS这样的地理信息系统的选择日益增多,这些数据比以往要更容易获得.另一种选择是使用从之前无人机收集的影像,并使用机载GNSS进行地理参考.这种方法要求在数据收集期间全球导航卫星系统必须可靠,但随后的定位不需要使用GNSS.这种类型的数据也因消费级四旋翼无人机的广泛应用而变得易于获取.无人机视觉定位的主要目标是将无人机当前的视图与从上述数据建立的视觉记忆进行匹配或配准,以确定无人机位置.无人机视觉定位任务的主流解决方案如图3所示.这些方法主要的挑战是需要大量参考数据以及进行不同来源的图像匹配或配准.参考图像通常由不同的传感器在不同时间、不同光照条

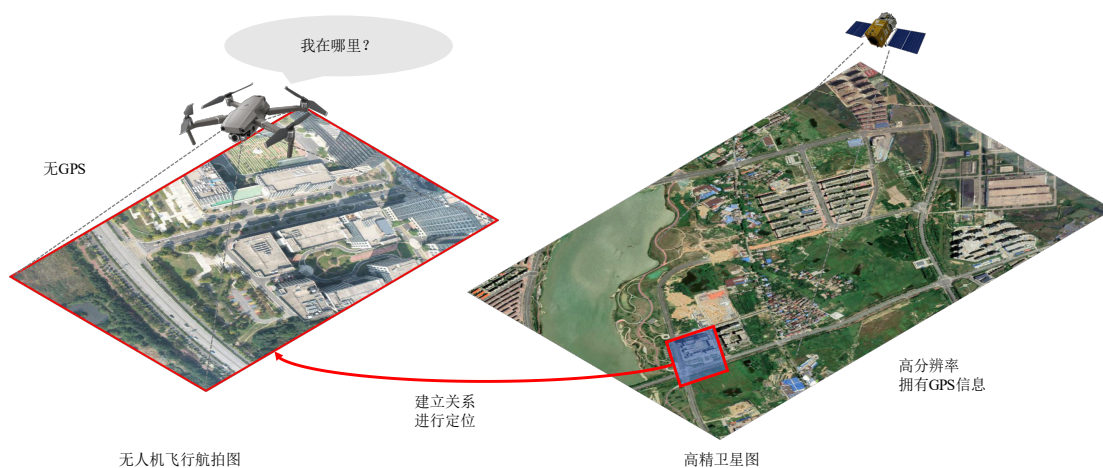


图3 无人机视觉定位任务示意图

件和不同分辨率下拍摄. 单一的参考地图或参考图像集可能包含来自许多来源的图像. 无人机影像与参考图像往往有着不同的视觉效果, 尤其是当这些图像进行正射校正和后处理时. 针对上述问题, 无人机视觉定位领域的工作发展出了多种方式, 其中主流的工作都包括某种形式的图像配准或匹配.

本节将讨论无人机视觉定位中的两大类主要方法: 图像检索的方法和图像匹配的方法. 图像检索方法通过从预先构建的地理数据库中搜索与查询图像相似的图像, 以估计无人机的绝对位置. 此类方法在具有标

注数据的大范围环境中具有显著优势, 尤其适用于目标环境中具有明显地标或参考特征的场景. 相比之下, 图像匹配方法则是通过直接比较无人机当前图像与参考图像之间的特征(如特征点和图像内容)来确定相对位置信息. 此类方法在缺乏全局地理数据库或动态环境中(如飞行器进入GPS信号受限区域)具有优势, 能够通过特征匹配实现位置跟踪, 获得更加精准的位置. 图4展示了这两类方法在无人机定位任务中的不同技术框架. 接下来将分别介绍这两类方法的原理及其最新进展, 包括传统方法与近年来的深度学习方法.

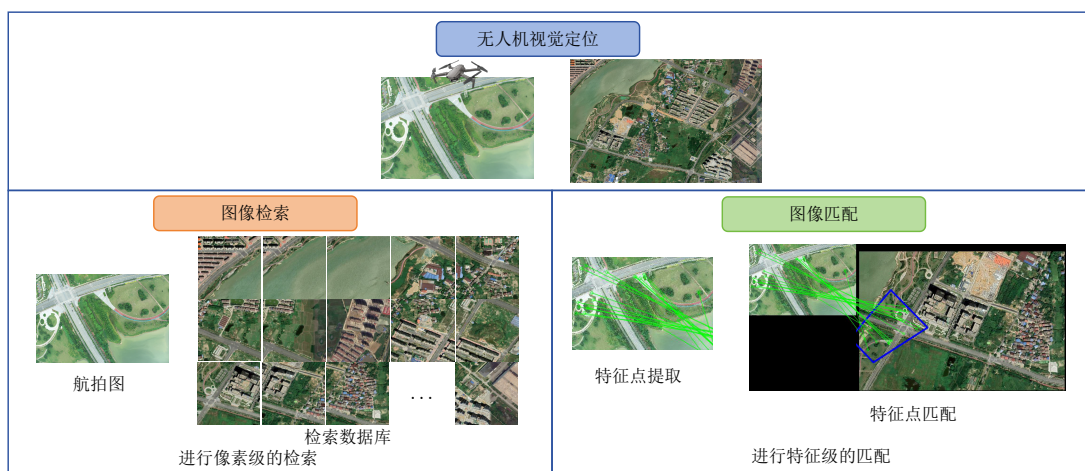


图4 无人机视觉定位分类示意图

2.1 图像检索的方法

图像检索是一种从数据库中搜索和提取与给定图像相似或相关的图像技术. 在视觉无人机定位任务中, 基于图像检索的方法将带有位置信息的卫星地图建立为数据库, 通过分析用户输入的无人机视角查询图像的内容特征, 如颜色、纹理、形状等, 根据内容特征提取与查询图像相似的图像, 表2总结了图像检索的无人机

视觉定位方法的算法类型、核心原理、鲁棒性和适用场景. 从而获得无人机所在位置. 本节将介绍此类方法.

2.1.1 传统图像检索方法

传统的图像检索方法主要基于图像内容的显著特征来实现相似图像的匹配与搜索. 具体来说, 这类方法会提取图像中的颜色、纹理、形状等低层次特征, 通过计算特征之间的相似度, 从而在数据库里进行检索, 实

表2 图像检索的无人机视觉定位方法对比

文献	方法类别	算法类型	核心原理	光照鲁棒性	动态适应性	适用场景
文献[23]	传统图像检索	图像检索+NCC	归一化互相关+粒子滤波框架	—	—	室内/室外
文献[24]	传统图像检索	EKF+SLAM	改进 Cholesky 分解存储协方差矩阵,提高计算效率	—	√	室外
文献[25]	传统图像检索	互信息+MI	差分模板跟踪最大化互信息,结合 sRt 模型处理 3D 运动	√	√	室外,环境变化大
文献[26]	传统图像检索	相位相关+PC	利用相位相关处理光照差异与几何失真,通过分解 PC 矩阵消除光照影响	√	—	室外,不同光照条件
文献[27]	传统图像检索	互信息改进+NID	归一化互信息度量(NID),结合 L-BFGS 优化算法	√	√	室内/室外,多视角任务
文献[28]	深度学习图像检索	CNN+三元组损失	提出数据集 University-1652,并构建任务基线	—	√	室外
文献[29]	深度学习图像检索	局部模式学习方法	局部模式网络(LPN),方形环分区策略	√	√	超视距任务,动态环境
文献[30]	深度学习图像检索	Transformer 建模方法	基于 Transformer 的特征分割与区域对齐(FSRA)	—	—	室外
文献[31]	深度学习图像检索	关键点检测与注意力方法	联合学习表示与关键点检测(RK-Net),USAM 注意模块	√	—	动态场景
文献[32]	深度学习图像检索	交叉熵损失的双分支网络	双分支深度神经网络(DNN)分别提取卫星和无人机视图图像特征	—	—	多高度室外
文献[33]	深度学习图像检索	多分类器集成方法	基于 ConvNeXt 的多分类器方法(MCCG)	—	—	室外
文献[34]	深度学习图像检索	自蒸馏与多尺度方法	自蒸馏框架 (PASS-KD),多尺度补丁分区策略	—	—	室外
文献[35]	深度学习图像检索	分区学习与移位策略方法	基于移位密集分区的表示学习(SDPL)	√	√	室外动态场景
文献[36]	深度学习图像检索	Transformer 特征对齐方法	基于 Transformer 的跨视图特征对齐 (TransFG),特征聚合模块和梯度引导模块	√	—	室外
文献[37]	深度学习图像检索	自适应与风格调整方法	多环境自适应网络 (MuSe-Net),动态调整域迁移	√	√	恶劣天气任务
文献[38]	深度学习图像检索	结合热成像	热生成模块生成对齐的伪热成像数据,卫星-热成像地理定位优化跨域匹配性能	√	√	低光照长距离飞行
文献[39]	深度学习图像检索	图像检索和对齐	共享的深度图像表示进行全局检索和局部对齐	√	√	离线/在线室外

现对无人机的定位. Van dalen 等人^[23]提出采用归一化互相关(Normalized Cross Correlation, NCC)系数作为相似度度量,并结合粒子滤波框架,以生成位置估计和不确定性,将无人机捕获的图像与 Bing Maps 创建的全局参考图进行对齐. 在室内和室外飞行测试中,该方法展示出了良好的定位能力以及较小的偏差. Magree 等人^[24]发现标准 EKF 在同时定位与建图(Simultaneous Localization And Mapping, SLAM)框架中可能会出数值不稳定,为了解决这一问题,他们通过存储协方差矩阵的分解形式来改善稳定性. 此外通过修改的 Cholesky 分解来存储和更新协方差,以达到高效添加和移除特征状态的目的. 并进一步验证了方法的有效性和稳定

性,在计算负荷与标准 EKF 相当的前提下,系统在飞行测试中表现稳定,误差在预期范围内.

过去依赖视觉的定位通常结合 SLAM 存在漂移问题,并且大多数方法对环境变化十分敏感. Yol 等人^[25]提出了一种基于视觉的位置估计算法,来实现无人机捕获的图像与地图中的全局参考图像之间的对齐. 采用互信息(Mutual Information, MI)作为相似度度量,以提高对场景变化的鲁棒性. 使用差分模板跟踪方法,通过最大化互信息来估计相机相对于参考图像的位置变换. 并应用 sRt 模型处理 3D 运动,以适应无人机的俯视图相机配置. 在真实飞行测试中,该方法表现出了对光照和季节变化的鲁棒性,并与 GPS 数据进行了比较,验证

了其准确性。

Wan 等人^[26]提出了一种基于相位相关(Phase Correlate, PC)的照明不变定位算法,该算法十分快速并具有一定的鲁棒性,能处理光照条件差异和相机 3D 运动导致的几何失真。文章首先分析了 PC 的照明不变性,证明了它在太阳角度变化下的稳健性。图像的 PC 矩阵 $Q(u, v)$ 可以分解为两个独立的 PC 矩阵:

$$Q(u, v) = Q_1(u, v) \times Q_2(u, v) \quad (1)$$

其中, Q_1 表示光照带来的影响,它解释了光照引起的差异, $Q_2(u, v) = e^{i(au+bv)}$ 考虑了平移。这种分解意味着可以通过提取 Q_2 来消除 Q_1 对 Q 的影响来实现光照不变下的模板匹配。之后提出一种 PC 扫描方法,生成多个候选位置,无人机的最终位置由扫描中候选位置与其归一化峰值 PC 值的加权确定,增强了定位的鲁棒性和准确性。在不同光照条件下,该方法在有阴影图像和遥感图像上的实验证明了其高精度和快速的定位性能。

在 Yol 等人^[25]工作的基础上, Aluri^[27]提出了一种改进的定位方法,来完成无人机捕获的图像与地图上的参考图像之间的匹配。该方法使用归一化的互信息相似性度量(Normalized Information Distance, NID)作为度量指标,与 MI 相比, NID 不依赖于图像之间的重叠量,因而对视角变化下的定位任务更具鲁棒性。进一步地,关键点 k_i 的图像坐标带有不确定性 R_i , $k_i = \bar{k}_i + \delta k_i$, 其中, δk_i 服从均值为 0, 方差为 R_i 的高斯分布。通过匹配关键帧图像对中的关键点,用三角测量得到关键点对应的 3D 地标位置。此外,为了提高精度,算法使用有限内存拟牛顿法(Limited-memory Broyden Fletcher Goldfarb Shanno, L-BFGS)优化图像配准参数,结合高斯模糊平滑与渐进优化,最终实现了无人机图像到地图图像的全局姿态估计。

尽管传统图像检索方法在特定条件下能够提供有效的定位手段,但仍存在一些显著的优缺点。优势方面,这类方法由于基于视觉特征(如颜色、纹理、形状等)的匹配,在有良好地理标志物的固定场景中能够取得较高的匹配精度;同时,与深度学习方法相比,传统方法对计算资源的要求相对较低,适用于资源受限的无人机平台。然而,这些方法的局限性在复杂环境中尤为明显。首先,传统方法在处理光照变化、图像噪声、视角差异等方面表现出较低的鲁棒性,使其在动态环境或视角变化大的场景下容易失效。此外,传统方法的匹配精度对图像重叠度依赖性较高,难以处理低重叠度或跨尺度的图像对。因此,虽然传统方法为无人机定位任务提供了基础支持,但其适用性在复杂场景下受限,为此需要引入更多鲁棒性更强的方法,如基于深度学习的图像检索方法。

2.1.2 深度学习图像检索方法

随着深度学习在计算机视觉领域的快速发展,基于深度学习的图像检索方法逐渐成为无人机视觉定位中的关键技术。相比传统方法,深度学习方法能够自动学习更具鲁棒性的图像特征,特别是在面对复杂环境、视角变化以及光照不均时表现出更强的适应性。此类方法通常通过卷积神经网络(Convolutional Neural Networks, CNNs)提取高层次的语义特征,从而实现与无人机视角图像与数据库图像之间更准确的匹配。近年来,基于深度学习的图像检索方法已在无人机定位任务中展现出显著优势,为解决复杂场景下的视觉定位提供了新的思路。深度学习离不开数据的支持,Zheng 等人^[28]考虑跨时图的地理定位问题,提出了第一个基于无人机的地理定位数据集 University-1652,包含世界各地的 1 652 个大学建筑的无人机合成航拍图、卫星图和地面视图,将无人机视图作为过渡,减少了跨视图地理定位的困难。他们提出了基于无人机视角的两种任务:无人机视图目标定位和无人机导航。无人机视图目标定位旨在通过无人机视图图像预测目标位置的位置,无人机导航是给定一个卫星视图查询图像将无人机驱动到查询中感兴趣的区域。

Wang 等人^[29]提出了一种局部模式网络 LPN,这是一种端到端的可学习模型,能够从航拍图像和地面图像中提取全局特征。LPN 的关键思想是通过使用方形环分区策略,使网络关注邻近区域的上下文信息。LPN 使用多类交叉熵损失进行训练,在测试期间,分类层之前的特征用于通过余弦距离估计相似性。

Dai 等人^[30]提出了一种基于 Transformer 的特征分割与区域对齐方法(Feature Segmentation and Region Alignment, FSRA),用于无人机视角的地理定位。无人机和卫星图跨视角定位最大的困难是位置偏移以及距离和尺度的不确定性。现有方法主要针对挖掘更全面的细粒度信息,低估了提取鲁棒特征表示的重要性和特征对齐的影响。虽然基于卷积神经网络的方法在跨视图地理定位方面取得了一定的成功,但仍然有一些限制,例如它只能提取邻域内的部分信息,一些尺度缩减操作会使一些细粒度的信息丢失。因此他们引入了一种简单有效的基于 Transformer 的结构 FSRA,以增强模型对上下文信息的理解能力,并理解实例的分布。FSRA 不使用额外的监督信息,根据 Transformer 特征图的热分布划分区域,然后在不同视图对齐多个特定区域,最后将每个区域集成到一组特征表示中。由于 FSRA 是基于特征图的热分布自动划分区域,因此当图像中发生显著变化和尺度变化时,仍然可以划分和对齐特定实例。

受人类视觉系统关注局部模式的启发, Lin 等人^[31]

提出了一个用于联合学习判别表示并使用单个网络检测显着关键点的框架 RK-Net. 该框架引入了一个单元减法注意模块 (Unit Subtraction Attention Module, USAM), 它可以从特征图中自动发现具有代表性的关键点, 并将注意力集中在这些点附近区域. 虽然 USAM 参数量很少, 但性能提升很显著, 并且可以轻松移植到不同的网络. 通过结合 USAM, RK-Net 有助于端到端的联合学习, 而无需额外的注释的先验条件. 表示学习和关键点检测是两个高度相关的任务. 表示学习有助于关键点检测, 关键点检测反过来又丰富了模型对视点变体引起的大外观变化的能力, 在三个挑战性数据集上取得竞争性的地理定位精度.

由于现有的公共数据集不包括无人机在不同高度获得的图像差异, 场景类型比较均匀导致模型无法适应复杂多变的场景. Zhu 等人^[32]提出了一个新的跨视图数据集 SUES-200, 包含无人机在 4 个飞行高度获取的图像和同一目标场景下对应的卫星视图图像. SUES-200 是第一个考虑无人机在不同飞行高度下的空中摄影产生的差异的数据集. 此外, Zhu 等人还构建了一个有效的训练测试和评估跨视图匹配模型, 在 SUES-200 数据集上对不同 CNN 架构的特征提取器的性能进行了全面评估, 并提出了一个稳健的基线模型. SUES-200 可以帮助模型学习在不同高度具有高辨别力的特征. 由于无人机相机姿态和周围环境对航空摄影的影响较小, 匹配系统的评价指标随着无人机飞行高度的提高而提高.

为了提取跨视角的丰富信息, Shen 等人^[33]提出了一种基于 ConvNeXt 的多分类器方法 (ConvNeXt-based Multiple-Classifer Method, MCGG), 用于多视图地理定位. MCGG 通过跨维度交互捕获丰富的判别信息, 获取多个特征表示, 实现综合特征表示. 尽管位置移动或尺度变化, 但具有鲁棒性的 MCGG 提高了利用更多上下文信息的多个特征表示. 在公共基准数据集 University1652^[28] 和 SUES-200^[32] 上进行的大量实验表明, MCGG 方法在无人机视图目标定位和无人机导航应用中相比于其他方法, 性能提升超过 3%.

在图像检索方法中, 检索到的航拍图像只可以与整个查询地面图像的一小部分区域 (即场景上下文变化) 匹配, 而且检索到的航拍图像只能描述粗粒度信息, 但查询地面图像可以捕获细粒度的细节 (即对象尺度变化). Li 等人^[34]提出了一种新的自蒸馏框架 (Patch Similarity Self-Knowledge Distillation, PASS-KD), 该框架将局部和全局知识作为细粒度的位置相关监督, 以自我增强的方式指导跨视图图像特征提取和表示. 具体来说, 该方法在一个辅助图像到图像的检索任务中探索场景上下文的变化, 并设计了一种多尺度补丁分区策略来感知对象跨视图的尺度变化. 在推理阶段中, 将

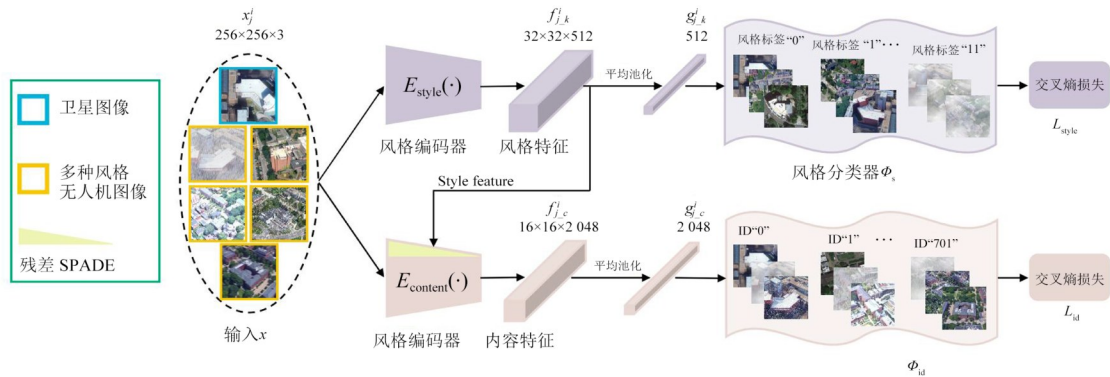
不再使用自蒸馏框架, 以避免额外的计算成本.

前述方法通过特征图分割挖掘更全面的信息, 但是这样不可避免会破坏图像结构, 并且查询图像中的目标移动和尺度变化会严重影响算法的性能. Chen 等人^[35]设计了一种简单而有效的基于部分的表示学习, 称为移位密集分区学习 (Shifting-Dense Partition Learning, SDPL). 该方法通过特征提取、密集分区学习和分类监督三个阶段来实现准确的跨视角地理定位. 其中, 特征提取使用骨干网络提取无人机和卫星视角的高级特征图. 为了提取鲁棒的特征表示以进行跨视角匹配, 作者提出了密集分区策略, 将图像分成多个部分来探索上下文信息, 同时明确维护全局结构. 为了处理非中心目标的场景, 作者进一步提出了移位融合策略, 该策略基于各种分割中心并行生成多组部分, 然后自适应地融合所有特征以选择最佳分区. 实验表明, SDPL 对位置移动和尺度变化具有鲁棒性, 并在两个公开数据集 University-1652^[28] 和 SUES-200^[32] 上具有先进的性能指标. Zhao 等人^[36]提出了一种基于 Transformer 的方法 TransFG, 包含特征聚合模块和梯度引导模块, 在特征表示和对齐方面实现了有效的平衡. 特征聚合模块隐式学习显着特征并使用视觉 Transformer (Vision Transformer, ViT) 动态聚合上下文特征, 梯度引导模块利用局部特征的梯度信息进一步增强跨视图特征表示, 并在不同视图之间对齐特定实例.

无人机在现实不同天气下 (如雨雾) 进行视觉定位时, 性能会显著下降, 因为它们并没有考虑训练数据和多个测试环境之间的域迁移问题. 为了弥补这一点, Wang 等人^[37]提出了一种多环境自适应网络 (Multiple-environment Self-adaptive Network, MuSe-Net), 以动态调整环境变化引起的域转移. 如图 5 所示, MuSe-Net 采用双分支神经网络, 包含一个多环境风格提取网络和一个自适应特征提取网络, 其中多环境风格提取网络是提取与环境相关的风格信息, 而自适应特征提取网络利用自适应调制模块动态最小化与环境相关的风格差距.

Xiao 等人^[38]提出了一种结合热成像和卫星图像的地理定位方法, 针对低光照和长距离飞行场景中无人机的定位问题. 该方法包括热生成模块 (Thermal Generative Module, TGM) 和卫星-热成像地理定位模块 (Satellite-thermal Geo-localization Module, SGM). TGM 通过生成对齐的伪热成像数据, 缓解了数据稀缺问题; SGM 则利用深度学习和域自适应技术优化跨域匹配性能. 实验结果表明该方法在低对比度环境中表现出优异的鲁棒性和较高的定位精度, 为 GPS 受限环境下的无人机地理定位提供了新思路.

He 等人^[39]提出了一种结合图像检索和对齐的深

图5 多环境自适应网络图^[37]

度学习地图辅助视觉地理定位系统,以实现无人机的鲁棒和实时定位.该系统通过共享的深度图像表示完成全局检索和局部对齐两个任务,大大提高了定位效率.该方法包括离线地图预处理和在线定位两个阶段,在不依赖任何位置信息先验的情况下实现了约10 km²的大范围精准定位.实验结果表明,该方法相较于现有技术,在计算效率提升5倍的同时,保持了良好定位精度水平.

相比传统方法,深度学习技术能够提取更为丰富的高层次特征,在光照、视角变化和复杂环境下展现出更强的鲁棒性.通过卷积神经网络(CNN)、图卷积网络(GCN)等深度模型,这些方法在视觉检索中实现了更高的定位精度.同时,诸如自适应学习和域迁移等方法的引入,使得深度学习在不同天气和环境条件下保持了良好的泛化性和可靠性.然而,深度学习在无人机视觉定位中仍面临一些挑战.首先是对计算资源的高需求,这在资源有限的无人机平台上可能受到限制;其次,尽管深度学习在特定环境下的鲁棒性较好,但在极端环境(如严重天气、光照极低等)下的泛化能力仍有待提升.此外,深度学习模型的“黑箱”性质使得解释性和透明性较低,增加了定位误差的诊断难度.

2.2 图像匹配的方法

在无人机视觉定位任务中,基于图像匹配的方法,对获得的卫星图以及无人机视觉图像使用特征点描述信息进行像素级别的匹配,随后根据像素级别的特征点匹配获得无人机相对于卫星图的位姿信息进行无人机定位.相比于图像检索方法,图像匹配方法更注重精确的特征对应,能够提供更高的定位精度.该类方法可以分为基于传统特征点的匹配方法、基于深度学习图像匹配方法以及视觉里程计方法.表3总结了图像匹配的无人机视觉定位方法的算法类型、核心原理、鲁棒性和适用场景.

2.2.1 传统特征点匹配方法

传统特征点匹配方法主要依赖于经典的特征检测和描述技术.传统特征点检测通常涉及角点检测器,如著名的Harris^[57]和FAST^[58].特征点检测的目标是在对同一区域的不同图像进行两次完全独立的检测迭代中找到容易检测到的显著点,这些检测迭代可能在光照、尺度、旋转和视点上有很大差异.在完成特征点检测后,需要进行特征描述符的提取,以便后续的匹配过程.描述符提取是从特征点周围的区域中提取特征向量的步骤.主流技术包括使用梯度直方图的SIFT^[59]和使用二进制测试的Brief^[60].特征提取的目标是构建一个描述符,用于使用欧几里得距离或汉明距离等度量来匹配特征点.

Seema等人^[40]提出了一种将随机抽样一致算法(Random Sample Consensus, RANSAC)用于实时定位无人机图像上目标位置的方法,该方法旨在替代传统的GPS系统.首先从GPS获取经纬度,并构建参考图像.使用SURF算法提取参考图像和无人机捕获的图像特征,应用RANSAC算法筛选匹配点,通过匹配点得到最终的配准结果.RANSAC算法在一定程度噪声下仍能提供准确结果,适合实时应用.但是该算法对于极端噪声和失真扭曲十分敏感,同时对特征点的最少数量有一定要求.Saranya等人^[41]指出全球导航卫星系统容易受到干扰,而惯性导航传感器存在误差累积的问题.他们使用基于视觉的位置估计,通过相机传输的视频图像和地图参考图像进行匹配,来缓解上述问题.比较了归一化互相关系数和RANSAC算法作为匹配算法时模型的性能,其中前者通过边缘检测和归一化互相关矩阵来计算匹配度,后者则通过检测和剔除不匹配特征来找到图像对的对应关系.经过实验得出结论:RANSAC方法具有约0.3 m的误差,适合处理旋转、缩放等变化,但在高噪声下可能失败,而NCC在去除噪声后更准确.

Shan等人^[42]提出了一种基于谷歌地图的无人机导航框架,使用相关性初始化位置,通过光流估计帧间运

表3 图像匹配的无人机视觉定位方法对比

文献	方法类别	算法类型	核心原理	光照鲁棒性	动态适应性	适用场景
文献[40]	传统特征点匹配	RANSAC+SURF	SURF 特征提取+RANSAC 匹配	—	—	室外,静态环境
文献[41]	传统特征点匹配	RANSAC+NCC	归一化互相关(NCC)+RANSAC 剔除匹配	√	√	室外,动态环境
文献[42]	传统特征点匹配	光流估计+HOG	光流估计 + HOG 特征描述 + 粒子滤波器	—	√	室外,复杂环境
文献[43]	传统特征点匹配	滑动窗口因子图	滑动窗口因子图框架,结合 2D-3D 匹配与轨迹特征	√	√	大规模复杂场景
文献[44]	传统特征点匹配	abBRIEF 描述符+MCL	结合色彩空间与降噪的 abBRIEF 描述符 + 蒙特卡洛定位(MCL)	√	—	多种光照条件
文献[45]	传统特征点匹配	地形分类特征	基于地形分类的特征提取(草地、道路、建筑)	√	—	室外自然场景
文献[46]	传统特征点匹配	相对视觉定位(RVL)框架	特征点提取 + 描述符匹配 (SIFT、ORB、AKAZE 等)	—	—	局部视觉场景
文献[47]	深度学习图像匹配	CNN	卷积神经网络(CNN),提取深度城市特征	√	—	室外,城市区域
文献[48,49]	深度学习图像匹配	SIFT + ORB + U-Net	SIFT/ORB 局部特征匹配 + U-Net 语义形状匹配	√	—	室外,密集区域
文献[50]	深度学习图像匹配	多任务多阶段网络	通用特征编码 + 语义分割解码 + 位置回归解码	√	—	城市场景,资源受限设备
文献[51]	深度学习图像匹配	cGAN 分割	条件生成对抗网络(cGAN)分割建筑物、道路和背景+地标提取	√	—	室外,地标丰富场景
文献[52]	深度学习图像匹配	SuperGlue 特征匹配	SIFT 特征提取 + 深度学习特征匹配(SuperGlue)	—	—	室外,长距离飞行
文献[53]	深度学习图像匹配	两阶段方法	全局局部视觉定位,全局检索模块 + 细粒度匹配模块	√	—	室外,稀疏纹理区域
文献[54]	深度学习图像匹配	粗到细的绝对定位	结合了相对 6 自由度(6-DoF)和绝对 2 自由度(2-DoF)定位	√	√	行星环境
文献[55]	视觉里程计	视觉教学与重复	视觉教学与重复(VT&R),结合立体视觉	—	√	室外,复杂路径
文献[56]	视觉里程计	联合优化	视觉里程计与卫星地图对齐联合优化	√	√	城市和农村场景

动,用 HOG 特征描述图像,构建谷歌地图的特征查找表,用粒子滤波器在预测位置附近进行粗到细的搜索,以找到最佳匹配. 具体来说,在通过相关性匹配初始化无人机的位置后,使用光流来预测无人机在后续帧中的位置来减少搜索范围. 之后通过分解单应矩阵获取帧间的平移量,采用 HOG 特征描述图像,与谷歌地图的参考图像进行匹配. 为了加速匹配过程,避免滑动窗口搜索的高计算成本,采用了粒子滤波器. 粒子滤波器通过采样和重采样的过程,逐步更新并集中粒子(代表可能的位置)以逼近真实位置. 最终通过粒子滤波器的迭代进行全局搜索,得到无人机的精确位置. 该方法结合谷歌地图、光流、HOG 特征匹配以及粒子滤波器,解决了传统方法的局限,显著减少了漂移并提高定位准确性,优于仅基于光流的视觉里程计算法. 虽然该方法计

算效率有所提升,但在处理光照变化和场景变化上仍有优化空间.

针对 GPS 信号在复杂环境和电磁干扰下可靠性降低的问题,Chiu 等人^[43]提出了一种新的视觉辅助导航系统. 传统的视觉辅助导航系统依赖于惯性测量单元(Inertial Measurement Unit, IMU)和相机,但在长时间缺乏 GPS 更新的情况下会出现定位漂移. 为解决这一问题,他们开发了一个融合 IMU 测量和两种视觉测量(2D-3D 匹配和轨迹特征)的滑动窗口因子图框架. 利用 2D-3D 匹配技术,将 2D 图像特征与 3D 地形数据库匹配,为导航估计提供绝对信息更新. 通过特征跟踪测量,将 3D 信息从一帧传递到下一帧,提高姿态估计的精度. 设计了一个滑动窗口因子图结构,使用增量式因子图优化算法实现实时的估计计算. 实验结果显示,该

方法在多个大型无 GPS 场景下能够提供准确且一致的空中导航方案。相较于传统滤波方法,滑动窗口因子图框架的优势在于更有效地处理延迟测量和维护估计一致性,尽管它需要较高的计算资源来进行 2D-3D 注册和特征跟踪。该研究的创新之处在于首次仅使用 IMU 和相机实现了精确的 3D 绝对姿态估计,并紧密集成了绝对尺度下的配准信息,显著减少了无 GPS 条件下的位置估计误差。

Mantelli 等^[44]提出了一种基于 abBRIEF 的无人机全球定位与跟踪方法。该方法利用无人机摄像头拍摄的图像,与卫星图像进行比对,实现高度、纬度、经度和航向 4 个自由度的定位。abBRIEF 是一种改进的 BRIEF 描述符,它结合了色彩空间和降噪策略,增强了图像匹配的鲁棒性,尤其在光照变化的环境中。通过应用到蒙特卡洛定位(Monte Carlo Localization, MCL)中,该系统能在不同地图和光照条件下表现出良好的性能,而且不需要实时地图更新。尽管该系统对光照变化和環境变化具有适应性,但可能无法处理大规模、复杂结构环境以及瞬态地面变化,如移动物体或雨天。此外,由于仅支持 4 个自由度定位,可能需要其他传感器来补充完整的位置信息。该系统计算成本低并且具备快速恢复能力,非常适合于 GPS 信号受限的场景。

Masselli 等人^[45]利用在线数据库的航拍图像创建参考地图,并通过匹配无人机拍摄的图像进行定位。他们引入地形分类(如草地、灌木、道路和建筑)作为特征,增强了方法对环境变化的鲁棒性。实验表明,这种方法在模拟和实际飞行中都表现出有效性,但可能会因误分类导致定位出现问题,而且实时处理速度需要优化以适应无人机的应用需求。研究还探讨了视觉地形分类如何帮助无人机在无全球传感器或信标的情况下独立定位,以及如何通过视觉特征提高定位性能,不受天气、时间或季节变化的影响。

Akhroufi 等人^[46]探讨了在 GPS 或其他全球导航卫星系统(GNSS)受限的户外环境中,无人机如何利用本地视觉信息进行相对定位,即相对视觉定位(Relative Vision Localization, RVL)。RVL 框架包括两个主要步骤:数据收集和定位估计。首先从无人机捕获的图像中检测关键点并提取特征,然后丢弃原始图像,只保留特征描述符。在定位时,新图像的关键点与数据库中的描述符进行比较,通过投票过程确定最匹配的场景,从而估计当前位置。选择了三种具有常规描述符的特征(SIFT、SURF 和 KAZE)和三种具有二进制描述符的特征(AKAZE、BRISK 和 ORB)。RVL 的一个创新点是减少了对外部全局信息的依赖,利用本地环境信息进行定位。虽然初步实现 RVL 存在局限,如无法进行全局位置估计,但它展示了实时运行的可能性。

传统特征点匹配方法在无人机视觉定位中提供了高效、可靠的基础手段。通过经典的角点检测和描述符提取,这些方法能够在一定程度上应对光照变化、旋转、尺度变化等挑战,实现跨图像的的稳定特征匹配。然而,这类方法的表现依赖于图像清晰度和环境稳定性,对视角差异较大、光照条件复杂或动态场景的适应性相对较弱。此外,传统方法在处理高维特征描述符时的计算开销较大,难以满足实时性要求。因此,尽管传统特征点匹配方法为无人机视觉定位任务奠定了坚实基础,但其在复杂环境中的局限性也促使了基于深度学习的图像匹配方法的兴起,以更好地应对多变的应用场景和更高的精度需求。

2.2.2 深度学习图像匹配方法

深度学习通过构建复杂的神经网络模型,能够自动从图像中提取更具语义性的高级特征,并进行精确匹配,实现高效、可靠的定位。与传统的图像匹配方法相比,深度学习方法在处理复杂场景和应对环境变化方面表现出色,尤其在光照变化、遮挡和尺度变化等挑战下仍能保持高精度的匹配结果。Amer 等人^[47]提出了深度城市特征的概念,通过卷积神经网络来提取每个城市或区域的视觉特征,使用这些特征完成识别和定位任务。作者利用 Google 地图和 Bing 地图的数据进行了实验,验证了深度城市特征在 GNSS 拒止环境下对无人机定位任务的有效性。

结合传统的计算机视觉技术和深度学习方法,Nassar 等人^[48,49]通过卫星图像和无人机拍摄的图像进行匹配来确定无人机的位置。首先,通过 SIFT 和 ORB 等局部特征进行初步定位。接着用 U-Net 进行语义形状匹配,提取图像中的道路和建筑等信息,以提高定位精度。在两个不同地理区域的测试中,该框架显示出了较高的定位准确性,证明了这种方法对于无人机导航的潜力。然而,这种方法在密集区域和低海拔飞行时可能效果有限,并且对卫星图像质量和标注的准确性有较高要求。

Marcu 等人^[50]设计了一种多任务多阶段神经网络,可以在一次前向传播中同时解决语义分割和定位问题,这对于应用在资源有限的无人机上尤其重要。网络结构包含一个通用特征学习的编码器,一个用于像素级语义标签的分割解码器,以及一个进行精确位置回归的回归解码器。实验结果显示,该模型在城市范围内实现了商业 GPS 精度级别上的定位,并在语义分割任务上超越了当时最好的结果。此外该模型还具有适应嵌入式 GPU 的紧凑设计,能够在资源受限的设备上运行。然而该模型也存在诸多问题,比如使用了道路分割策略,将该模型的应用区域限制在了城市地区;模型中的回归编码器需要用飞行区的图像进行训练,限制了

其泛化性。

Schleiss^[51]利用无人机上的摄像头和 Open Street Maps 的数据,首先将航拍图像进行分割,通过基于条件生成对抗网络(conditional Generative Adversarial Network, cGAN)的图像分割器完成。分割类别一共有三个:建筑物、道路和背景,以此提取出不同类型的地标信息。根据 IMU 提供的比例尺和旋转角度,将分割图像与任务区域的参考地图进行匹配,以此推断无人机位置。该方法允许利用多种类型的地标,如建筑物和道路,提高了匹配的可能性和定位的准确性,比仅依赖建筑物作为地标的更为灵活。实验表明,这种方法在典型任务区域的定位精度接近商业 GPS。与以前的方法相比,它适用于更广泛的场景,并能以更高频率和置信度输出绝对位置。当有适合的地标可用时,该方法可以作为 GNSS 系统的备份。

Gurgu 等人^[52]利用无人机搭载的单目广角相机进行无人机定位,在 GNSS 信号无法可靠使用的情况下能够进行定位。该算法假设无人机的飞行区先利用 SIFT 特征提取方法提取航拍图和卫星图的域是先验已知的,提前构建该特定区域的地图并将其上传到机载计

算机上进行离线使用。定位过程首特征,然后利用深度学习特征匹配方法 SuperGlue^[61]匹配航拍图和卫星图的特征点,将匹配中心点作为无人机的位置。实验结果表明,该定位方法与传统的基于 GNSS 的方法具有相当的准确性,与最先进的视觉里程计(Visual Odometry, VO)方法相比,能够应用于长距离、高空的无人机飞行。

基于匹配的方法虽然可以避免误差累积,但是匹配过程复杂,不能保证实时性。为了加速确定无人机所在区域并进行精确定位, Li 等人^[53]提出了一个全局局部视觉定位(Global Local Visual Localization, GLVL)网络。如图 6 所示,它是一个两阶段视觉定位方法,结合了一个与无人机飞行场景相似的大规模检索模块,以及一个定位精确无人机坐标的细粒度匹配模块,实现了实时和精确的定位。该方法能够同时关注局部结构特征的相似性并关注关键点信息,通过联合训练策略增强检索和匹配精度,以进一步提高模型能力。在包含纹理丰富和纹理稀疏区域的 6 个无人机飞行场景上的实验表明,该方法能够实现无人机实时精确定位要求。特别是在具有稀疏纹理特征的村庄场景中,在 0.48 s 内实现了只有 2.39 m 的定位误差。

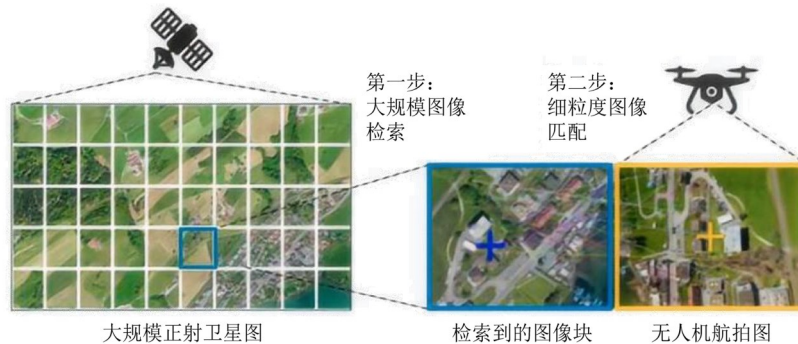


图 6 全局局部视觉定位方法示意图^[53]

Luo 等人^[54]提出了一种名为 JointLoc 的实时视觉定位框架,用于解决行星环境下无人机的定位问题。该方法创新性地结合了相对 6 自由度(6-DoF)和绝对 2 自由度(2-DoF)定位,通过自适应置信机制有效融合两种定位结果,实现了无人机在世界坐标系中的 6-DoF 位姿估计。框架采用粗到细的绝对定位模块,通过划分卫星地图,利用 SuperPoint 提取局部特征,并通过 LightGlue 进行特征匹配,显著减少了定位时间。如图 7 所示的实验结果表明, JointLoc 在模拟火星地形的无人机数据集上表现出卓越的实时性和精度,为行星探索任务中的无人机自主导航提供了可靠的解决方案。

深度学习图像匹配方法在无人机视觉定位中展现了优异的性能,尤其在面对光照变化、遮挡和尺度变化等复杂环境因素时,表现出较高的鲁棒性。通过深层神经网络自动提取高维度、语义丰富的特征,这些方法能

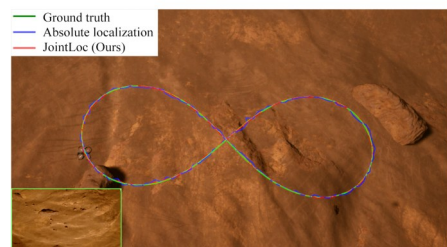


图 7 JointLoc 方法在模拟火星地形的结果图^[54]

够有效捕捉图像的全局和局部信息,弥补了传统特征点匹配方法在应对复杂场景时的局限性。同时,深度学习方法还能够通过大规模数据训练获得跨域适应能力,实现不同地理环境和天气条件下的准确匹配。尽管深度学习方法具有较高的计算需求,但其在 GNSS 拒止环境中的应用潜力巨大,为无人机定位提供了更加可靠的技术支持。未来,随着计算效率的提升和轻量化网

络的开发,深度学习图像匹配方法有望在资源受限的无人机平台上得到更广泛的应用。

2.2.3 视觉里程计方法

视觉里程计是一种基于相机捕获的连续帧比较的方法,用于通过分析相机获取的图像序列来估计相机的运动轨迹。视觉里程计利用无人机视角连续图像帧的视觉信息进行时序上的匹配来推算无人机的相对运动位置和姿态(如位置、方向和速度)。

Warren 等人^[55]探讨了在 GPS 信号失效时,如何使多旋翼无人机能够自主安全返回,其采用了视觉教学与重复(Visual Teach and Repeat, VT&R)技术,通过摄像头捕获的实时视图与之前构建的局部地图进行匹配,实现无人机的视觉定位和控制。首先利用立体视觉和视觉里程来实现视觉定位。这依赖于一个三轴稳定相机,不需要额外的外部传感器或基础设施。在学习阶段,云台控制保持开放状态,由云台内部控制器负责滚转和俯仰的稳定,以及平滑的偏航跟踪。在返回阶段,VT&R 主动控制云台的俯仰和偏航轴,以减少当前视图和最近的特权视图之间的定向误差。特权视图和特权顶点在学习阶段被记录,提供了精确的地标和姿态信息,用于帮助无人机在返回时进行定位和路径跟踪。在该阶段同时使用路径跟随控制器,目的是使无人机尽可能接近原出航路径,同时保持合适的目标速度。通过连接特权顶点生成路径,使用同步轨迹估计和映射轨迹生成算法,拟合出一条常速轨迹。在 5~25 m 的高度和 55 km/h 的速度下,系统性能已经得到验证,能够在复杂路径上实现闭环自主控制。尽管在多种复杂条件下表现出色,但速度限制和对外部条件的敏感性是可能改进的地方。

为了克服卫星和无人机之间图像捕获条件的差异,如季节和透视变化,Goforth 等人^[56]使用现成的卫星数据训练的卷积神经网络表示,只需要在无人机上搭载单目 RGB 相机,利用卫星图像库比较和对齐无人机图像进行无人机定位。为了提升定位精度,Goforth 等人提出一种联合优化,可以联合最小化相邻无人机帧和卫星地图之间的误差。该方法首先利用视觉里程计来确定运动参数的初始估计,然后将最近帧的子集与卫星地图进行比较,以地理定位这些帧,最后使用帧里程计和地图对齐之间的联合优化来细化所有帧的地理定位姿势。实验表明,该方法可以从城市环境推广到具有挑战性的低纹理农村数据集,并且在所有帧上对无人机的姿态进行优化,结合了视觉里程计和卫星地图对齐的约束,即使在没有直接与地图比较的无人机帧也能进行准确的定位。

视觉里程计方法在无人机视觉定位中提供了一种基于连续帧分析的高效路径跟踪方式,通过对运动序

列的连续估计,实现无人机的相对位姿推测。相比基于特征点的传统方法,视觉里程计能够适应更动态的场景,并在无 GPS 信号条件下提供可靠的返回路径规划。此外,结合深度学习模型(如卷积神经网络)和卫星图像数据库的方式,也为跨季节、跨视角的图像对齐提供了新的解决方案。然而,视觉里程计在速度、光照变化和低纹理场景下的鲁棒性仍有提升空间。未来,通过更高效的自适应模型、联合优化策略和更精细的多模态融合,视觉里程计方法有望在更复杂的无人机应用场景中实现更高的稳定性和精度。

2.3 对比与分析

在无人机视觉定位的研究中,图像检索方法和图像匹配方法是两种主要的技术路径。为了深入了解这两种方法的优缺点及其在实际应用中的表现,本节将从定性和定量两个角度对它们进行全面的比较分析。

2.3.1 定性对比

由于无人机航拍图与卫星图之间的极端差异、不同的拍照时间和两个视图之间的不一致分辨率,无人机视觉定位仍是一项艰巨的任务。图像检索方法依赖于全局特征的相似性,这使得它在处理大规模数据库时效率更高,特别适合快速获取相似图像或识别相似场景。然而,由于全局特征的抽象性,这种方法在细节识别和精确定位方面的表现可能不如图像匹配方法。此外,图像检索方法在面对光照变化和部分遮挡等复杂环境时,表现出一定的鲁棒性,但在结构上相似但内容差异较大的图像间,可能会出现误判的情况。

相较之下,图像匹配方法更强调精确的特征点匹配,这使其在细节层面上的识别能力显著增强。其可以通过特征点的精确定位,实现高精度的匹配和定位。虽然图像匹配方法的计算复杂度较高,尤其在处理高分辨率图像或特征点数量众多的情况下,计算资源的消耗也相应增加,但它的高精度特性使其在需要精确分析和决策的场景中不可或缺。然而,图像匹配方法对图像质量和特征点的依赖也意味着,在特征点稀少或图像质量较差的情况下,其匹配效果可能受到影响。表 4 展示了图像检索与图像匹配方法在原理、常用算法、使用场景、匹配精度等维度的对比。

2.3.2 定量对比

在无人机视觉定位领域,大部分方法的实验和验证都是基于自定义数据集进行的。然而,自定义数据集往往在参考地图的大小、参考地图的分辨率、测试环境的大小、飞行距离、飞行高度、无人机的速度和传感器等变量设置上不同,难以对不同的方法进行公平对比。随着深度学习方法的发展,研究者们开发了更为标准化的公共数据集,以便更全面地评估算法的性能。其中,University-1652^[28]、SUES-200^[32]和 DenseUAV^[62]数

表 4 图像检索与图像匹配方法对比

比较维度	图像检索方法	图像匹配方法
原理	基于全局特征向量	基于局部特征点的检测和匹配
常用算法	CNN 特征+最近邻搜索	SIFT,ORB,深度学习特征点检测
使用场景	快速检索	精确定位
匹配精度	在对场景有大致了解或需要快速获取相似图像时表现良好	通过精确的特征点匹配,能够达到高精度的定位和匹配
计算复杂度	通常计算复杂度较低,适合大规模数据库的快速查询	匹配过程涉及复杂的特征点检测和匹配算法,计算复杂度较高
鲁棒性	对光照变化、尺度变化、部分遮挡等情况具有一定的鲁棒性,但可能在结构相似但内容差异较大的图像间出现误差	对图像的细微变化非常敏感,特别适合在光照变化和视角变化较小的情况下使用,但在特征点稀少或质量较差的情况下,匹配效果可能不佳

数据集近年来在无人机视觉定位领域得到了广泛应用。

University-1652^[28]数据集,如图 8 所示,包含由 72 所大学 1 652 栋建筑卫星视图-无人机视图-街道视图组成的图像数据对。University-1652 通常具有一个卫星视点图像、54 个无人机视点图像和该位置的多个街景图像。由于现实世界的无人机飞行成本较高,该数据集中的无人机视点数据通过谷歌地球的模拟飞行获得,其中无人机模拟飞行路线围绕目标场景飞行,高度逐渐下降。

SUES-200^[32]是第一个包含来自不同高度的无人机视图图像的跨视图数据集,可以在 150 m、200 m、250 m、300 m 不同高度收集的数据评估模型指标,如图 9 所示。与仅包含学校场景的 University-1652^[28]数据集不同,SUES-200^[32]包含更广泛的场景类型,包括但不限于校园建筑、公园、学校、湖泊和公共建筑。并且,所有的无人机视图都是在实际场景中获取的。此外,为了解决无人机实际飞行定位问题,无人机必须在同一区域连续飞行,因此一些 SUES-200 的目标场景是在同一区域进行收集的,以保证这些场景是连续的、相似的。

DenseUAV^[62]是一个专为低空城市市场设计的数据集,包括无人机视图图像和卫星视图图像,如图 10 所示。DenseUAV 专注于无人机自定位任务,在真实场景中进行密集采样,以解决在无 GPS 环境中的无人机精确自定位的需要。此外,DenseUAV 引入了不同时期的卫星图像,以促进模型适应不同时期产生的空间变化。

三个数据集 University-1652^[28]、SUES-200^[32]和 DenseUAV^[62]的对比如表 5 所示。

为了更全面地评价不同方法在数据集上的表现,表 5 中列举了用于评估这些方法的常用指标。下面将对每一个评估指标进行详细介绍。

Recall@K: 也称 R@K, 常用于评估检索方法的性能,用于衡量在前 K 个返回结果中正确结果的比例。以 R@1 为例,样本是否为正确匹配可以表示为

$$I(I_q, I_i) = \begin{cases} 1, & I_q = I_i \\ 0, & I_q \neq I_i \end{cases} \quad (2)$$

其中, I_q 对应于查询的类别, I_i 是第 i 个图像对应的类别,按照计算的欧几里得距离的升序排序。如果结果值属于同一类别,则为 1,如果不是,则为 0。对于所有样本, R@1 定义为

$$R@1 = \frac{1}{\|Q\|} \sum_{q \in Q} I(I_q, I_i) \quad (3)$$

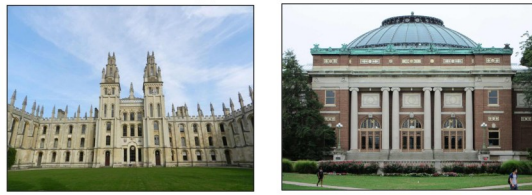
其中, Q 是所有查询图像的集合, $\|Q\|$ 表示 Q 中的图像数。可以看到,当只有查询的类别和图库中最接近的图像的类别相同时, R@1 的值会增加,否则,它们被视为错误匹配的样本。

平均精度 (Average Precision, AP) 用于衡量检索结果的精确度和召回率之间的综合表现。AP 是精确召回曲线下的面积,它考虑了评估中所有真实匹配图像的位置。

$$AP = \sum_{n=1}^N (R_n - R_{n-1}) \times P_n \quad (4)$$

表 5 无人机视觉定位数据集对比

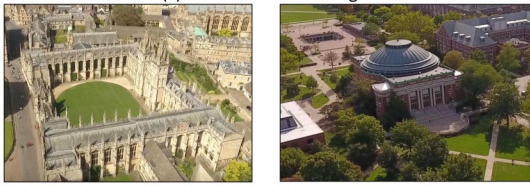
数据集	University-1652 ^[28]	SUES-200 ^[32]	DenseUAV ^[62]
训练数据量	701 类别×71.64 张图像	120 类×51 张图像	2 256 类×9 张图像
视角	无人机、地面、卫星	无人机、卫星	无人机、卫星
每个视角的图片数	54 张无人机视角图片+16.64 张地面视角图片+1 张卫星图片	50 张无人机视角图片+1 张卫星图片	3 张无人机视角图片+6 张卫星图片
无人机视角图片来源	谷歌地球模拟	真实场景	真实场景
不同高度采集	否	是	是
采样策略	稀疏采样	稀疏采样	密集采样
评估指标	Recall@K、AP	Recall@K、AP、RB	Recall@K、SDM@K



(a) 地面视角图像



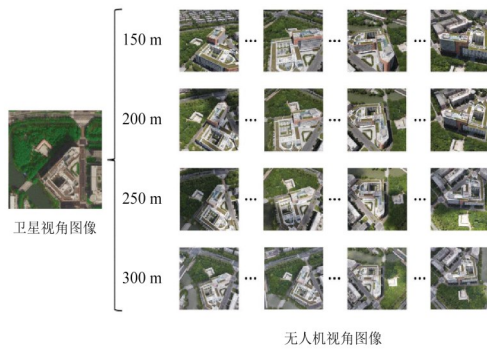
(b) 卫星视角图像



(c) 真实无人机视角图像



(d) 合成无人机视角图像

图8 University-1652数据集示意图^[28]图9 SUES-200数据集示意图^[32]

其中, P_n 是第 n 个检索结果的精度值, R_n 是第 n 个检索结果的召回率值. AP 反映了系统在整个检索过程中的平均表现, 综合考虑了所有返回结果的精确度和召回率. 高 AP 值表示系统在不同阈值下都能保持较高的精度和召回率, 说明其检索性能较为稳定和优秀.

鲁棒性 (Robustness, RB) 用于衡量模型在不同飞行高度的鲁棒性. 首先计算模型在不同高度的平均准确率. SUES-200^[31] 包含 4 种不同高度: 150 m、200 m、250 m、300 m.

$$\bar{y}^i = \frac{(y_{h1}^i + y_{h2}^i + y_{h3}^i + y_{h4}^i)}{4} \quad (5)$$

其中, y_{h1}^i 表示模型 i 在特定高度的 R@1 准确度, \bar{y}^i 表示模型 i 在 4 个高度的平均准确率.

$$c_i = \frac{1}{\sum_{h=1}^4 |y_h^i - \bar{y}^i|} \quad (6)$$

因此, 模型 i 的鲁棒性结果表示为

$$RB = \frac{c_i \times (y_{h1}^i + y_{h2}^i + y_{h3}^i + y_{h4}^i)}{4} \quad (7)$$

空间距离度量 (Spatial Distance Metric, SDM@K) 结合了 Recall@K 的特征, 提供了更准确的定位精度度量. 具体来说, 单个查询样本的 SDM 值定义为

$$SDM_k = \frac{(K-k+1)}{e^{s \times d_k}} \quad (8)$$

其中, $d_k = \sqrt{(x_q - x_i)^2 + (y_q - y_i)^2}$, $(K-k+1)$ 是第 k 个样本的权重, 权重是根据特征距离分配的, 为更接近查询特征的图库图像, 一般设置较大的权重; K 表示图库中与查询特征距离最接近的前 K 个样本; x_q 和 x_i 分别表示查询和图库图像对应的经度, 而 y_q 和 y_i 表示纬度. 简而言之, d_k 表示两幅图像之间的空间欧几里得距离, s 是放大因子. 最后计算所有 K 个样本的 SDM 值, 并通过归一化过程得到最终的 SDM@K:

$$SDM@K = \frac{(K-k+1)}{e^{s \times d_k}} / \sum_{i=1}^K (K-i+1) \quad (9)$$

接下来将展示本节中介绍的无人机视觉定位方法在 University-1652^[28]、SUES-200^[32] 和 DenseUAV^[62] 数据集上的实验结果.

University-1652 数据集^[28] 上的实验结果如表 6 所示. 由于 University-1652 是包含无人机、地面、卫星三个视角的数据集, 可以进行多种视角的检索, 其中涉及无人机视觉定位任务的检索对是无人机-卫星, 因此表 6 中只展示了无人机-卫星对的结果. 表 6 中的“基线方法”指的是 Zheng 等人^[28] 提出的基线模型. SDPL^[35] 由于密集分区学习和分类监督策略, 在 R@1 和 AP 的准确度上都取得了比其他所有方法更好的结果. 值得一提的是, Muse-Net^[37] 虽然在 R@1 和 AP 指标上不够高, 但是它使用生成网络扩充了数据集的环境, 在雨、雾、雪等不同环境下综合表现最好.

SUES-200 数据集^[32] 上的实验结果如表 7 所示, 同样只展示了无人机-卫星对的结果. 由于 SUES-200 数据集包含 150、200、250、300 m 四种不同飞行高度, 因此

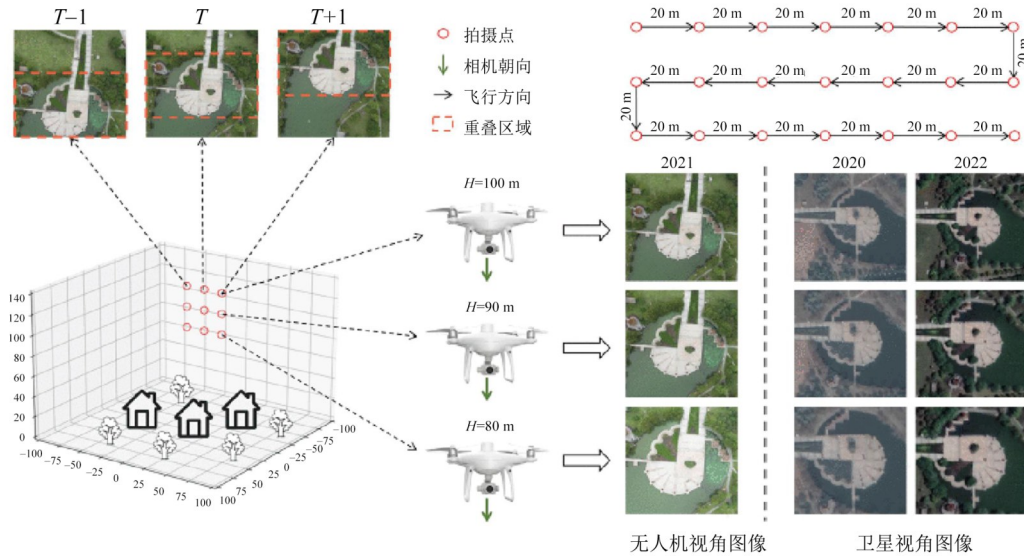


图 10 DenseUAV 数据集示意图^[62]

表 6 University-1652 数据集上的实验结果对比

方法	R@1	AP
基线方法 ^[28]	58.49	63.31
LPN ^[29]	75.93	79.14
FSRA ^[30]	82.25	84.82
RK-Net ^[31]	77.60	80.55
MCCG ^[33]	89.64	91.32
SDPL ^[35]	90.16	91.64
TransFC ^[36]	84.01	86.31
MuSe-Net ^[37]	74.48	77.83

表 7 中列举了不同高度下的结果。其中，“基线方法”指的是 Zhu 等人^[32]提出的基线模型。可以发现，SDPL^[35]的密集分区学习和分类监督策略在不同飞行高度下的 R@1 和 AP 指标都要更好。

表 7 SUES-200 数据集上的实验结果对比

方法	150 m		200 m		250 m		300 m	
	R@1	AP	R@1	AP	R@1	AP	R@1	AP
基线方法 ^[32]	55.65	61.92	66.78	71.55	72.00	26.43	74.05	78.26
LPN ^[29]	61.58	67.23	70.85	75.96	80.38	83.80	81.47	84.53
FSRA ^[30]	68.25	73.45	83.00	85.99	90.68	92.27	91.95	93.46
MCCG ^[33]	82.22	85.47	89.38	91.41	93.82	95.04	95.07	96.20
SDPL ^[35]	82.95	85.82	92.73	94.07	96.05	96.69	97.83	98.05

DenseUAV 数据集^[62]上的实验结果如表 8 所示。由于 DenseUAV 是近期公布的新数据集，目前只有 Zheng 等人^[28]和 Dai 等人^[62]的基线方法在 DenseUAV 上的实验结果。可以发现，Dai 等人的基线方法在各项指标上均大幅高于 Zheng 等人的基线方法，这是由于 Dai 等人的方法是在密集采样的 DenseUAV 数据集上进行训练

的，而 Zheng 等人的方法是在非密集采样的 University-1652 数据集上进行训练的。这也说明了密集采样的重要性，相邻无人机视角图像采集区域的重叠迫使神经网络学习到空间分布信息。

表 8 DenseUAV 数据集上的实验结果对比

方法	R@1	R@5	SDM@1	SDM@5
基线方法 ^[28]	25.77	42.63	33.94	21.07
基线方法 ^[62]	80.18	93.99	84.39	78.02

3 无人机视觉导航

无人机视觉导航是一种利用计算机视觉来实现无人机自主导航和控制的技术。通过摄像头等视觉传感器获取环境信息，利用图像处理和计算机视觉技术进行环境感知、障碍物检测、路径规划等任务，从而实现无人机在复杂环境中的自主飞行。在第 2 节中，已经详细介绍了无人机定位的方法，通过这些技术，无人机能够准确获取当前的位置和姿态信息。然而，仅仅知道位置和姿态并不足以完成飞行任务。在许多实际应用中，无人机需要从起点自主飞行到目标点，这就需要进一步解决导航问题。导航不仅依赖于精准的定位，还需要结合路径规划和障碍物规避等技术。障碍物检测和规避方法用于确保无人机在飞行过程中避免碰撞，保障飞行安全。路径规划是无人机在导航任务中找到从起点到目标点的最优或可行路径的过程，确保任务顺利完成。因此，在本节中，将围绕无人机视觉导航的核心任务展开讨论，包括障碍物检测与规避以及路径规划方法。

3.1 障碍物检测和规避方法

障碍物规避是自主导航的关键组成部分，因为它

可以识别和传递周围障碍物的重要信息,从而减少碰撞和操作失误的风险.因此,优秀的障碍物规避方法有可能显著提升无人机的自主性.障碍物规避是基于检测障碍物并计算无人机与障碍物之间的距离.当障碍物靠近时,无人机会根据障碍物规避模块的指令进行规避或转向.现在常用的一种障碍物规避方法是使用测距仪(如雷达、超声波和红外线)来估算距离.然而,这些测距仪的视野和测量范围较窄,使它们在复杂环境中无法收集足够的信息.与这类方法相比,视觉传感

器可以收集大量的视觉数据,这些数据可以被处理并用于规避障碍物.障碍物规避方法分为三类:基于光流的方法、基于双目的方法和基于SLAM的方法.表9涵盖了基于光流、双目视觉和SLAM三类避障方法的详细对比,包含核心原理、低光照适应性、适用场景等多个维度^[63-76].使用图像处理技术^[77]可以检测规避障碍物.光流的使用可以产生局部信息流并获得图像深度.通过图像处理技术,这些方法能够在不同飞行环境下动态生成障碍物信息,并进行实时路径调整.

表9 基于障碍物检测和规避的无人机导航方法对比

文献	方法类别	核心原理	低光照适应性	复杂场景适应性	适用场景
文献[63]	基于光流的方法	Lucas-Kanade 梯度法,分析像素运动生成深度信息	—	√	室内动态环境
文献[64]	基于光流的方法	昆虫仿生视觉方法,结合局部运动检测生成光流	√	—	微型无人机嵌入式应用
文献[65]	基于光流的方法	通过光强度变化估计距离,模拟昆虫通过缝隙判断障碍物	√	—	微型无人机嵌入式应用
文献[66]	基于光流的方法	运动场估计方法,利用图像块特征分类结合 DCT 分析	—	√	室外复杂环境
文献[67]	基于光流的方法	使用光流场估计结合区域分割与目标识别	—	√	低成本室内外环境
文献[68]	基于光流的方法	二次多项式光流估计方法,基于光流平滑限制与均匀性参数优化,估计像素的运动	√	√	精细化分析的复杂场景
文献[69]	基于双目的方法	基于立体视觉的占据图动态规划,结合概率路线图和 D* Lite	√	√	户外动态障碍环境
文献[70]	基于双目的方法	使用前视立体相机结合 CL-RRT 对轨迹进行碰撞检查	—	√	户外场景
文献[71]	基于双目的方法	恒速避障策略,利用立体视觉估计障碍物	—	—	室内未知环境
文献[72]	基于双目的方法	嵌入式立体视觉算法,利用 U 视差图和 V 视差图检测障碍	√	—	户外场景
文献[73]	基于 SLAM 的方法	自适应特征点生成算法,结合地面控制模块处理复杂室内环境	√	—	室内环境
文献[74]	基于 SLAM 的方法	稀疏地图生成与势场法结合路径规划	√	√	室内复杂环境
文献[75]	基于 SLAM 的方法	基于 NMPC 避障,结合动态和静态障碍物的路径规划	—	√	动态与静态障碍物混合场景
文献[76]	基于 SLAM 的方法	实时视觉多目标跟踪,结合相对位置与速度分析	—	√	货运与移动平台场景

3.1.1 基于光流的方法

基于光流的方法通过分析单个摄像机捕获的图像进行障碍物检测和规避.通常采用基于 Lucas-Kanade 梯度的光流方法,在三维空间环境中提取深度信息,用于检测和避开障碍物.这些方法通过分析连续图像帧中的像素运动来估计障碍物的距离和方向,这类技术一般在 μ Linux 实时操作系统的工作站上开发并实现. Ameli 等人^[63]研究了障碍物大小变化的检测,模拟了人眼在视野中物体因距离变化而放大的机制.它通过比较连续图像来检测障碍物,并确定与障碍物的距离.该研究提出了一种基于检测到的特征点随时间变化的尺寸特征的创新算法.它还使用了无人机运动中从两帧连续图像中找到的特征凸度的尺寸比率.

仿生昆虫视觉也被引入到光流导航方法中.受到蜜蜂视觉的启发, Strübbe 等人^[64]提出了一种简单的非迭代光流方法,用于测量系统的全局光流和自我运动.该方法基于昆虫视觉神经解剖学的 Reichardt 模型^[78]作

为局部运动检测单元.作者还发现,当刺激信噪比 (Signal-to-Noise Ratio, SNR) 增加时,运动处理技术会从 Reichardt 模型切换到梯度检测器.随着信噪比的增加,局部信号调制会减小.作者使用了两种实验来调整输入的信噪比.在一种实验情况下,改变了平均亮度.随着亮度增加,信噪比随着光子发射泊松分布的平方根的增加而增加.在另一种实验情况下,调整了图案对比度.在峰值对峰值比较中,信号及其信噪比随着对比度线性增加.为了将昆虫视觉算法理论验证应用于无人机,文献[65]的团队为微型飞行器 MAV 开发了一种名为 Octave 的视觉驱动自动驾驶仪,作为仿生微型机器人研究的一部分,用于自主飞行器的光学高度控制系统.他们展示了一种基于低复杂度光学速度传感器同时控制高度和速度的系统,该传感器预测向下的光流.这种速度传感器基于在飞行中对基本运动检测器 (Elementary Motion Detectors, EMDs) 的电生理学发现.他们创建了一个简单的 100 g 系留直升机设备,可以在随机图

案表面上跟踪地形. 该系统通过预测光流, 展示了基于轻量级传感器的自主飞行能力.

最近, Darius Merk 开发了一个受昆虫视觉^[79]启发的系统, 该系统利用光速来评估对象之间的距离. 这个系统虽然简单, 但非常有效, 因为在自然界中, 许多昆虫可以根据光的强度来感知附近的障碍物. 在飞行过程中, 昆虫视网膜上的图像运动会产生视觉流信号, 这种光流提供了视觉导航的空间信息. 根据穿过叶片缝隙的光强度, 昆虫可以快速判断障碍物是否可以安全通过. 然而, 基于光流的技术无法确定精确的距离, 这限制了其在某些任务中的应用. 另一方面, 基于 SLAM 的方法可以通过复杂的 SLAM 算法提供精确的度量地图, 使无人机能够在了解更多环境信息的情况下进行导航和避障. Moreno-Armendariz 等人^[80]描述了一种使用 SLAM 系统来绘制未知环境的方法. 为了避免静态和动态障碍物, 使用了一种新颖的人工势场方法. 自主车辆的主要目标是穿越陌生区域. 开发一张地图可以实现这一目标, 车辆可以利用地图在已访问的位置之间建立路径. 获取这样一张地图而无需预先了解周围环境或机器人的初始位置是一项特殊的挑战. 另一方面, 避开静态和动态障碍物需要使用一种革命性的人工势场方法. 这些新颖的设计在现场可编程门阵列 (Field-Programmable Gate Array, FPGA) 上实现. 然后, 在配备计算机视觉系统的差动牵引移动机器人上对这些独特的理念进行了测试, 这些机器人在一个陌生的受控环境中移动. 实验结果表明该系统可以在实时环境中进行应用.

He 等人^[66]提出了另一种基于视觉的无人机导航系统, 该系统可以通过运动场估计来避免障碍物. 假设无人机对其在惯性坐标系 (基于 GPS 数据) 中的线速度有合理的理解, 无人机的陀螺仪可以提供其方向的基本估计. 此外, 暂时假设相机已经被放置在无人机的重心位置, 并且相机的方向与无人机主体的方向一致. 图像帧是使用安装在无人机上的摄像机捕捉的, 通过对这些图像块进行分类, 可以估计建筑物的边缘、角落、道路线和树梢, 因为每个图像块都有自己的一组特征. 作者应用了离散余弦变换 (Discrete Cosine Transform, DCT) 来识别模式或边缘. 此外, 还考虑了图像的像素值以计算无人机与物体之间的距离.

为了应对障碍物检测中传感器重量过大的问题, Lin 等人^[67]提出了一种使用廉价摄像头的基于视觉的目标跟踪方法. 基于地图的运动规划主要关注在任务执行过程中自动转向以避开障碍物, 这在相对低的飞行高度尤其关键^[81]. 研究使用了包含许多障碍物的自然环境, 作者将离线路线规划和在线碰撞规避方案结合用于自主导航管理. 快速扩展随机树 (Rapidly-

exploring Random Tree, RRT) 算法被用于通过离线引导方法构建初始路径. 路径的航路点配有方向指示, 集成摄像头用于在移动过程中捕捉一系列图像, 以创建光流场. 在飞行过程中, 这被用于识别障碍物, 并建立无人机运动与遇到的障碍物之间的关系. 这个过程分为多个步骤, 从图像预处理开始, 通过光流计算、区域分割和目标识别, 最终进行判断, 如图 11 所示. 基于平滑限制和照明均匀性参数, 作者提出了一种光流估计方法^[68]. 该方法使用二次多项式展开来估计附近像素的运动, 从而在精度和速度方面产生准确的结果. 这些研究是在包含不同障碍物的户外区域进行的, 结果显示系统上每帧图像分析的持续时间大约为 6 帧/s 或 150~180 ms. 与稀疏光流算法相比, 该方法每个像素的运动是在两帧连续图像之间更详细地计算的. 然而, 该方法在障碍物规避算法上的应用仍然是有限的.

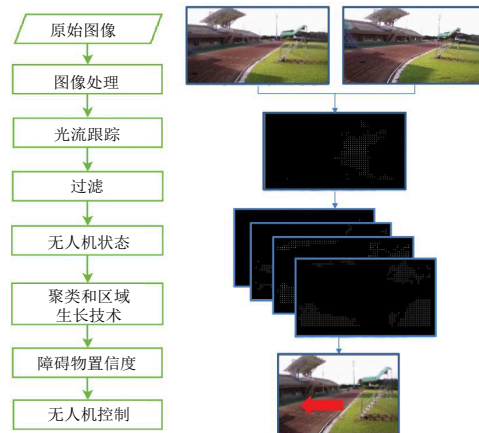


图 11 基于光流的障碍物检测方法^[80]

基于光流的障碍物检测和规避方法通过计算连续图像帧之间的亮度变化来检测运动和避障, 因此在无人机实际应用中会出现以下局限性. 首先, 光照变化会导致图像帧之间的亮度差异增大, 破坏光流算法的基本假设, 从而导致误差增加, 使得光流算法在识别运动和避障时出现误判. 其次, 缓慢的摄像头抖动在图像帧之间引起微小的平移和旋转, 这些抖动会导致光流计算出现误差, 生成虚假光流, 误判环境中的障碍物, 降低避障的准确性和响应速度. 最后, 在户外环境中, 风、雨、灰尘和复杂背景等环境噪声会干扰图像质量, 降低图像的清晰度和对比度, 使得光流算法在特征提取和跟踪时变得困难, 增加误检和漏检的概率, 影响无人机的避障能力和整体性能.

3.1.2 基于双目的方法

基于双目视觉的避障方法利用两台相机从不同角度同时捕捉环境图像, 通过计算图像之间的视差生成深度图, 从而提取环境的三维信息. 双目视觉系统^[82, 83]

首先进行图像对齐和特征点匹配,这一步称为立体匹配^[84-87],其目的是找到两幅图像中相对应的点对。接下来,通过视差计算,即计算这些对应点在两幅图像中的位置差异,获得每个像素点的深度信息。深度信息可以通过三角测量原理来推导,即根据相机之间的已知基线距离和视差,计算出场景中物体的距离。这些深度信息被用于构建环境的三维地图,无人机依靠这张三维地图实时检测和识别障碍物的位置和形状,进行避障决策。

为使无人机在执行电力线检查等任务时能够安全地穿越包含障碍物的未知环境,Hrabar等人^[69]结合了概率路线图和D*Lite进行路径规划,并使用基于立体视觉的占据图进行动态重新规划。这种方法使用高分辨率的占据图和低分辨率的规划图,从而减少规划器的状态空间和规划成本。通过仿真和电缆阵列机器人的实验表明,该方法可以检测并避开电线杆等障碍物以到达目标位置。

户外场景具有高度纹理化的特性,因此使用相机进行场景感知,可以提供宽广的视野,并且尺寸小、重量低和功耗小,同时具备非常宽的双轴视野。Matthies等人^[70]构建一个微型无人机户外飞行避障系统,使用前视立体相机对进行障碍检测,并使用向下看的相机来辅助状态估计。为了提供低内存和计算要求的障碍检测,在图像空间中使用世界的反向范围表示(缩放的立体视差)。通过执行视差图的类似构型空间扩展,将候选轨迹投影到图像空间,并执行类似深度缓冲器的可见性测试,能够非常高效地对闭环快速扩展随机树(Closed-Loop Rapidly-exploring Random Tree, CL-RRT)规划器生成的轨迹进行碰撞检查。这种方法在飞越虚拟森林和穿过虚拟墙壁的狭窄开口的模拟飞行中效果显著。在模拟器中,该方法对飞行器状态的扰动也表现出很强的鲁棒性。在实测中,使用Asctec Pelican四旋翼飞行器在树林中进行测试,由于短期妥协限制了轨迹搜索到可见视野,并限制了偏航偏移以保持朝向目标,导致轨迹存在一些非最优性。

Tijmons等人^[71]关注无人机在未知的室内环境中的导航问题,提出了一种基于立体视觉的避障策略,适用于受限的机载传感和处理设备。该方法通过允许恒速机动,性能优于反应性避障策略,同时计算效率极高,不需要存储先前的图像或地图。该策略处理了大多数固定翼和扑翼平台的非完整运动约束,以及立体相机系统的有限视野,可在没有传感器和电机噪声的情况下无障碍飞行。该方法在模拟器中的实验证明了避障策略的有效性,在无人机实测中也证明了其鲁棒性、可靠性和高效性。

Grinberg等人^[72]开发了一种无人机嵌入式立体视

觉避障算法。该系统基于视差图像进行计算和分析,这些图像编码了捕获场景中物体的距离。计算出的视差图像用于生成所谓的U视差图和V视差图,这些图作为障碍检测和避障算法的输入。虽然避障算法在TULIPP平台的ARM处理器上运行,但立体处理算法,包括图像校正、像素匹配、半全局匹配优化、一致性检查和中值滤波在FPGA上运行。尽管只用C/C++实现并通过高层次综合优化来加速在FPGA上的执行,图像处理显示出适合于无人机实时应用的性能。

总体而言,基于双目视觉的方法具有显著优势,能够生成精确的深度信息并构建三维地图,在高纹理环境中表现尤为出色。然而,其计算复杂度较高,对硬件性能和视野的依赖较大,在动态环境中的适应性仍有待提高。

3.1.3 基于SLAM的方法

基于同时定位与地图构建(Simultaneous Localization And Mapping, SLAM)的方法通过同时构建环境地图和确定自身位置,使无人机能够在未知环境中实现自主导航与避障。SLAM系统集成了传感器数据处理、环境感知、路径规划和运动控制等多个模块,利用激光雷达、视觉传感器或其他多种传感器数据,实时生成环境地图并更新自身位置,从而有效应对复杂环境中的障碍物,如图12所示。

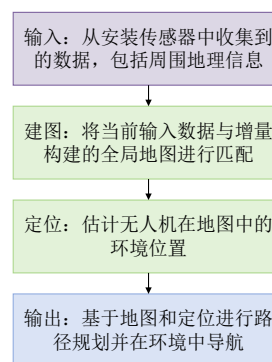


图12 基于SLAM的无人机导航方法示意图

为了应对室内环境光线不足和特征点数量依赖的问题,Bai等人^[73]开发了一种基于PTAM算法的自适应地图特征点生成方法,该方法在GPS信号缺乏的区域内具有实时性能。无人机可以在建筑物内部飞行,用于反恐和救灾。在封闭的室内环境中,GPS信号可能会被屏蔽,无法确定无人机位置。此外,复杂的室内环境对飞行安全构成了挑战。因此,需要实时检测复杂情况并避开障碍物。在这项研究中,PTAM算法被引入到无人机地面控制模块中。他们构建了一种自适应地图特征点生成机制,用于处理室内光线不足、特征点较少等问题,从而降低系统对特征点数量和照明条件的依赖。作

者提出了一种用于地面控制模块的障碍物识别算法和预警机制,并使用小型四旋翼无人机对该方法进行了测试.结果表明,在处理未知室内情况下,自定位方法可以在飞行过程中提供障碍预警,该方法对研究在 GPS 缺失条件下无人机自主定位和飞行安全具有重要意义,并且在实时性方面表现良好.

Esrafilian 等人^[74]提出了一种基于定向快速旋转简要 SLAM (Oriented fast and Rotated Brief SLAM, ORB-SLAM) 的方法.该方法首先通过计算无人机的 3D 位置和生成稀疏点云图来处理视频数据.然后,增加稀疏地图密度以丰富地图.最后,利用势场法快速遍历随机树,产生无碰撞的道路布局.其描述了一种具有单目基于视觉的自主飞行和避障系统的商用四旋翼无人机.前置摄像头的视频馈送和无人机的导航数据通过无线方式传输到地面站笔记本电脑,接收到的数据由 ORB-SLAM 处理,ORB-SLAM 使用视觉计算机器人的 3D 位置和点云形式的环境 3D 稀疏映射.该方法用于丰富重构映射,采用卡尔曼滤波进行传感器融合.线性特征用于计算单目 SLAM 的比例因子,还开发了一个比例-积分-微分控制器 (Proportional-Integral-Derivative controller, PID) 来控制 3D 位置.最后,利用势场法和 RRT 路径规划技术构建了一条无碰撞的路线.

Potena 等人^[75]使用非线性模型预测控制 (Nonlinear Model Predictive Control, NMPC) 控制器通过添加动态和静态避碰功能来改进基于视觉的导航.该方法设计了特定的障碍物,考虑了其速度和不确定性,从而保障在预定轨迹上进行安全操作.作者提出了一种同时集成避碰和感知限制的管理方法,还使用了可配置的障碍表征方法,允许模拟各种障碍物,同时编码障碍物的不确定性和速度.当周围环境充满静态障碍物时,动态障碍物可能会突然出现并引起干扰,会使无人机在检测到的障碍物时能够跨越之前设计的安全路线.在这种情况下,目标识别、视觉伺服和基于视觉的指导是可以利用的方法.这个挑战被表示为最优控制问题 (Optimal Control Problem, OCP),一般采用退缩视界的方式解决. NMPC 在每个驱动周期提供 OCP 的可能解,并且只使用最有效路径的第一个输入来控制机器人. OCP 在几毫秒内通过数值优化解决,确保在识别到意外障碍物后能够重新规划路径,从而实现实时车辆管理.建议的协议还可以将误差和物体速度整合到椭圆体中,使其能够应对动态障碍物.如图 13 所示,无人机在不断瞄准特定目标 (图中红色圆形目标) 的同时达到预期状态,环境中存在应避免的静态障碍物 (图中黑色和黄色条纹物体) 及动态障碍物 (如红色移动框中的其他代理),动态障碍物继续移动会阻挡计划的飞行轨迹 (即蓝色虚线),采用 Potena 等人的方法,无人机会对检测

到的物体做出反应,沿着新的安全轨迹 (即红线) 转向.

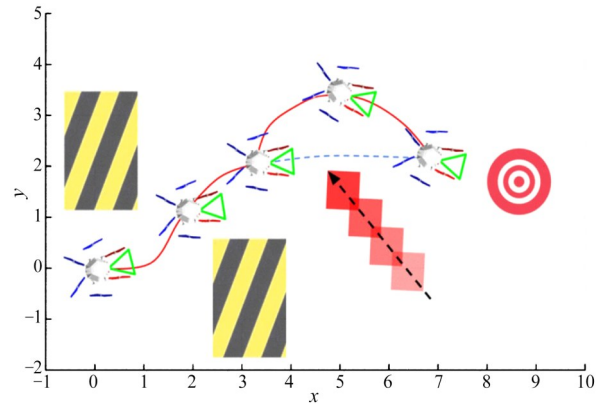


图 13 非线性模型预测控制作用示意图^[75]

Yang 等人^[76]设计了一种可靠的无 GPS 目标识别方法和一种精确的相对位置和速度分析方法,以解决无人机在移动平台上着陆的挑战.该挑战的设置如下:在一个移动平台上放置了 4 个桶,这些桶分别放置在 4 个圆圈内.首先,无人机需要识别移动平台上的圆形目标和圆形附近的黑白 ID 标记,然后无人机选择一个目标桶,追踪并将其运输到另一个移动平台,直到所有 4 个桶从一个移动平台运输到另一个移动平台.为了完成货物运输任务, Yang 等人开发了一种基于实时视觉的无 GPS 多目标跟踪方法.使用 Pixhawk 控制器来实现跟踪,图像用于估计目标与无人机之间的相对距离和速度.这项研究中提出的方法可以用于货运运输和海上救援任务中无人机的目标跟踪和检测.

综上所述,基于 SLAM 的方法在多种复杂场景下表现出色,特别是在 GPS 缺失或环境动态变化显著的情况下.通过集成环境建图与定位能力, SLAM 系统显著增强了无人机的导航与避障性能.

3.2 路径规划方法

路径规划是无人机导航中的一项关键活动,它需要根据一系列性能特征 (如最低工作成本、最短飞行时间和最短路线) 确定从起点到目的地的最有效路径.在这个过程中,无人机还必须避开障碍物,如图 14 所示.根据用于计算理想路径的环境信息类型,这个问题可以分为两类:全局路径规划和局部路径规划.全局路径规划的目标是基于全局地理地图确定最佳路径.然而,全局路径规划方法普遍实时性不足,当遇到需要快速完成其他任务或在飞行过程中出现突发情况时,难以对无人机进行实时控制.而局部路径规划必须不断接收来自周围环境的传感器数据,并实时计算出无碰撞路径.表 10 总结了路径规划方法的特点和适用场景.

3.2.1 全局路径规划方法

全局路径规划方法根据起点和目标点的位置创建

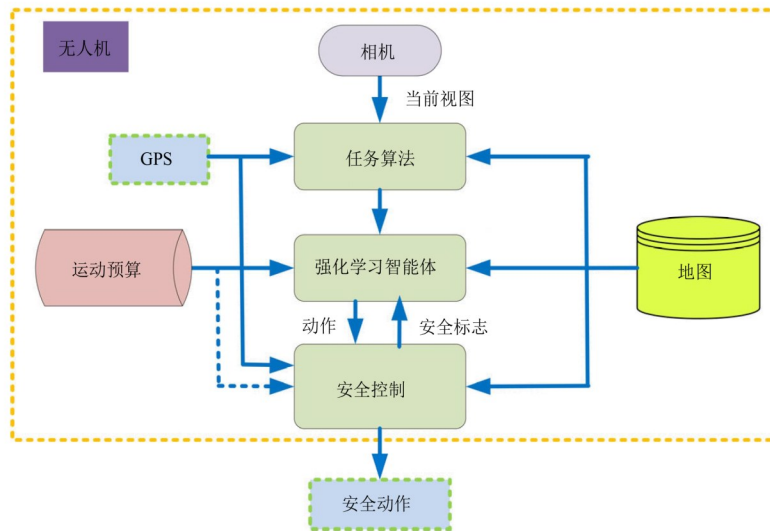


图 14 基于路径规划的无人机导航方法示意图

初始路径,依赖静态地图来进行导航任务.通常,这类方法分为启发式搜索方法和智能优化算法两类.

(1) 启发式搜索方法

启发式搜索方法通过结合启发式信息和搜索策略来高效地寻找最优路径.利用启发式函数评估当前路径的优劣,从而引导搜索过程朝向目标方向,减少计算时间和资源消耗. A*算法是从基本的Dijkstra算法派生而来的,是一种常见的启发式搜索方法.近年来, A*算

法派生了许多增强的启发式搜索方法. Vachtsevanos 等人^[88]设计了一种支持高级控制和任务规划算法的混合硬件/软件平台,从地形数据库构建了一张数字地图,并使用改进的 A*算法选择最佳路径.该平台的核心是使用基于智能模糊逻辑的技术和对对象建模的方法,以解决未建模的动态问题、应对不确定性问题.模糊逻辑程序被应用于关键的车辆模块中,例如路径规划器、模糊导航器、容错工具和飞行控制器.

表 10 基于障碍物检测和规避的无人机导航方法对比

文献	方法类别	算法类型	核心原理	高动态适应性	适用场景
文献[88]	全局路径规划	改进的 A*算法	使用模糊逻辑和对对象建模,结合地形数据库生成最佳路径	—	未建模动态环境
文献[89]	全局路径规划	启发式 A*算法	基于网格点价值函数的最优路径规划,结合模式分类器优化搜索过程	√	空中拦截任务
文献[90]	全局路径规划	稀疏 A*搜索	在路径规划中引入空间搜索限制和实时参数调整,降低计算复杂度	—	空中、陆地和海上场景
文献[91]	全局路径规划	动态 A 算法	在动态或未知环境中通过局部地图更新,实现实时路径重规划	√	动态、部分已知或未知环境
文献[92]	全局路径规划	HHO-GWO 优化算法	结合哈里斯鹰优化和灰狼优化算法,避免局部最小值,生成快速安全路径	—	静态环境
文献[93]	全局路径规划	模拟退火+蚁群优化	使用混合 SA-ACO 算法,结合框架四叉树,提高路径规划效率	√	移动机器人路径规划
文献[94]	全局路径规划	增量启发式路径规划	结合 3D 环境感知和全局增量路径规划,逐步更新占据网格和障碍物信息	√	超视距任务、动态环境
文献[95]	全局路径规划	混沌映射、极化信息素记录、模拟退火	引入了混沌映射、极化信息素记录规则以及模拟退火筛选机制,提高收敛速度和轨迹优化效率	√	复杂海上环境
文献[96]	局部路径规划	分子力场方法	使用虚拟力场控制无人机移动,结合凸多边形初始化优化部署	—	室外静态任务
文献[97]	局部路径规划	人工势场法	采用三方向旋转涡流场,避免局部最优、碰撞及路径振荡	√	动态环境

续表

文献	方法类别	算法类型	核心原理	高动态适应性	适用场景
文献[98]	局部路径规划	遗传算法	多级约束优化3D路径,结合非支配排序算法增强规划能力	—	复杂3D环境
文献[99]	局部路径规划	蚁群优化算法	改进蚁群算法加入惩罚策略,引导搜索历史较差探索未知区域	—	无人车路径规划
文献[100]	局部路径规划	Hopfield网络	结合A*算法初始化神经网络节点,基于能量函数迭代最优路径	—	室外复杂任务
文献[101]	局部路径规划	DEM+体素地图	高程模型和3D立方体网格细化局部模型,结合深度图算法避障	√	复杂3D环境
文献[102]	局部路径规划	合成数据集与体素地图	基于合成数据集构建占据图和3D网格地图,结合最小加速度轨迹生成器规划路径	√	城市搜索与救援任务
文献[103]	局部路径规划	兴趣点+Dijkstra	利用Dijkstra算法从兴趣点中选择允许路径,形成无碰撞轨迹	√	室内/室外未知环境
文献[104]	深度学习路径规划	DQN	双状态输入策略结合DQN,优化奖励函数和Q值函数	√	室内/室外未知环境
文献[105]	深度学习路径规划	模仿学习+强化学习	融合模仿学习与TD3算法,加速训练并解决分布不匹配问题	—	动态任务场景
文献[106, 107]	深度学习路径规划	双深度Q网络	端到端强化学习控制,优化覆盖路径规划和数据收集任务	√	动态环境,随机起点
文献[108]	深度学习路径规划	TD3+深度强化学习	深度强化学习结合指导注意力递归TD3算法,动态调整导航与避障任务的关注权重	√	低空复杂环境
文献[109]	深度学习路径规划	CNN	基于深度神经网络的图像处理,用于无人机在室内环境中的自主导航	—	室内走廊
文献[110]	深度学习路径规划	DNN	优化森林路径环境的高分辨率图像检测,支持无人机路径规划	—	森林路径
文献[111]	深度学习路径规划	RNN+CNN	将AI算法应用于光学传感器图像处理和航天导航	√	航天/无人机任务
文献[112]	深度学习路径规划	实时三维路径规划	基于物体相对位置优化路径,无需存储A*的前沿节点	√	受限环境,动态障碍
文献[113]	深度学习路径规划	自回归网络	混合动作空间优化无人机的路径规划,包括位置和偏航动作,以克服视野限制	√	未知环境

Rouse^[89]使用启发式A*技术完成基于沿估计路径的不同网格点价值函数的最优路径规划。在空中拦截任务规划的框架中创建了一个原型路径规划器。一个真实的地理场景被划分成矩形网格,每个交叉点分配了九个属性。这些特征被组合生成描述每个交叉点属性的模式向量。无人机操作员根据飞越具有特定属性的地点的期望性对这些向量的代表样本进行了评分,使用这些数据构建了一个最小距离模式分类器。然后,使用算法搜索程序创建了一个路径规划器,该路径规划器依赖于结合距离和模式分类器输出的启发式方法,以确定目标的低成本路径。

Szczerba等人^[90]引入了用于路径规划的稀疏A*搜索(Sparse A-star Search, SAS),通过在路径规划期间引入空间搜索限制,有效地降低了计算复杂度。该算法在规划过程中能够引入各种路径约束,并在任务期间随时调整这些参数,使其对几乎所有类型的智能导航系

统都具有重要价值,包括但不限于空中、陆地和海上军事工具。通过引入新的技术来修剪搜索空间,同时约束生成的路径以满足任务参数,该路径规划器的功能性和效率都很高。实时路径规划使其可以适应各种任务限制,例如最小路线段长度、最大转弯角度、路径距离限制和到目标位置的固定进场向量。

Stentz^[91]提出了动态A*方法,通常称为D*算法,适用于部分或未知的动态环境。当在路径上检测到新的障碍物时,它可以更新其对未知环境的地图。该方法是一种经过证明的最优且高效的路径规划算法,适用于配备传感器的机器人。该算法能够处理从完整且准确的地图信息到完全没有地图信息的各种先验地图信息。D*是一个非常通用的算法,可以应用于除了机器人运动规划以外的人工智能问题,可以处理任何路径成本优化问题,其中成本参数在寻找解决方案的过程中发生变化。当这些变化在搜索空间的当前起点附近被

检测到时, D*算法最为高效,能够在未知、部分已知和变化的环境中高效、最优和完整地规划路径。

无人机成功执行任务并避免障碍物的能力取决于无人机的最优路径规划和跟踪性能。无人机能够通过使用飞行参数实现以最小能量和时间消耗规划和跟踪路径。Belge 等人^[92]使用哈里斯鹰优化(Harris Hawk Optimization, HHO)-灰狼优化(Grey Wolf Optimizer, GWO)混合元启发式优化算法进行最优路径规划和跟踪,以使无人机在避免障碍物的情况下完成任务。混合 HHO-GWO 算法可以避免局部最小值和快速收敛的特点,成功地获得可行且有效的路径。此外,确定了无人机质量变化不确定性对最优路径规划和跟踪性能的影响。通过与粒子群优化(Particle Swarm Optimization, PSO)和 GWO 等元启发式群优化算法进行比较,测试了所提出方法的有效性。实验结果表明,所提出的算法能够在不陷入局部最小值的情况下生成快速且安全的最优路径,并且四旋翼无人机能以最小的能量和时间消耗跟踪生成的路径。

(2) 基于优化的方法

基于优化的方法通过建立数学模型和优化算法来寻找最优路径,通常以最小化路径长度、能量消耗或避障代价为目标,利用智能算法和优化技术在复杂环境中进行全局路径规划。近年来,遗传算法(Genetic Algorithm, GA)、模拟退火算法(Simulated Annealing, SA)和蚁群优化算法(Ant Colony Optimization, ACO)等智能优化算法得到了广泛应用,它们通过模拟自然界的进化、物理系统的退火过程和蚂蚁觅食行为来探索解空间,寻找全局最优解。Zhang 等人^[93]结合了框架二叉树表示法和混合模拟退火与蚁群优化算法,提出了一种新的路径规划方法 SAACO,提高了路径规划的效率。利用框架二叉树表示法可以提高环境分解效率并保持地图的表示能力。然而,模拟退火方法依赖于随机起点,如何高效地自动提供更好的初始解集估计仍然是一个研究热点,作者使用蚁群优化为模拟退火提供一个良好的初始解。根据理论分析和模拟实验的结果,SAACO 算法可以成功解决移动机器人路径规划问题,使机器人在无碰撞路径中寻找特定目的地,并提高机器人导航速度。

超视距任务场景或地面控制站访问受限的情况下,需要具备自主安全导航的能力,并需要持续更新现有的和可能过时的障碍物信息。Andert 等人^[94]提出了一种新的 3D 环境感知和全局路径规划技术的综合方法。采用局部约束的传感器融合方法提取稀疏障碍物,以随时进行全局增量路径规划。飞行过程中,立体相机通过分析深度图像检查飞行路径前方的视野,逐步建立 3D 占据网格。为了减少网格类型地图的高数据率和

存储需求,创建了一个近似的多边形世界模型,棱柱和地面平面紧凑了地图表示,使得系统能够不断更新其关于障碍物的知识。增量启发式路径规划器使用先验信息和增量障碍物更新来确保随时有无碰撞路径。飞行测试的映射结果展示了基于真实传感器数据的机载世界建模的功能。在模拟环境中,考虑到世界模型在车辆视野内的变化,展示了路径规划的可行性。

Yan 等人^[95]提出了一种基于改进蚁群优化算法的无人机轨迹生成方案,专注于解决复杂海上环境中不确定障碍物的避障问题。作者首先构建了一个三维海上环境的数学模型,并结合无人机性能指标设计了多维代价函数,用于优化飞行路径。为了克服传统蚁群优化算法易陷入局部最优解的问题,他们引入了混沌映射、极化信息素记录规则以及模拟退火筛选机制,从而显著提高了算法的收敛速度和轨迹优化效率。实验结果表明,该方法不仅在单无人机和多无人机场景下均生成了安全、高效的飞行轨迹,还在轨迹损失率和目标函数值方面优于基线算法,为无人机在复杂环境中的应用提供了理论支持。

总体来看,全局路径规划方法在静态环境和部分动态场景中表现优异。启发式方法在计算效率和路径质量间实现了平衡,适合任务约束明确的场景;智能优化算法通过算法创新,能够应对更复杂的约束和动态环境需求。然而,其局限性在于依赖完整且准确的全局环境信息,对动态障碍物处理能力不足,无法实时应对环境的动态变化。

3.2.2 局部路径规划方法

局部路径规划使用局部环境数据和无人机的状态估计来规划动态避碰的局部路径。在动态环境中进行路径规划受限于不可预测的因素(如物品在动态环境中的移动)而变得极其复杂。在这种情况下,路径规划算法必须通过各种传感器收集有关环境中未知元素的信息(例如大小、形状和位置),以适应环境的动态特性。传统的局部路径规划方法包括空间搜索方法、人工势场技术、模糊逻辑技术和神经网络方法。本节介绍了一些常见的路径规划方法。

受分子间力的启发,Wang 等人^[96]提出了一种基于分子力场的无人机群最优部署算法,使无人机可以最大限度地覆盖指定的工作区域。该方法是一种虚拟力方法,将无人机从其周围环境移动到抽象的人工引力场环境中,不断移动无人机,目标点既有“吸引力”,也有“排斥力”,通过两种力的控制,无人机逐渐向目标位置前进。多旋翼无人机群携带传感器并在指定工作区域内快速构建移动无线传感器网络,根据任何三架无人机的覆盖区域重叠最小的原则,确定所需的最小无人机数量。基于凸多边形的几何特性,提出了两种初始

化方法,使初始部署更加均匀,随后通过所提出的基于分子力场的部署算法优化所有无人机的位置.仿真实验结果表明,与现有的三种算法相比,所提出的算法能够通过所提出的初始化方法获得指定工作区域的最大覆盖率.陷入局部最优的概率和计算复杂度降低,而收敛速度得到了提高.

路径规划是无人机在确定源点和目标点之间最佳路线时面临的最关键问题之一,所选择的路径、安全性和计算负担对于在有障碍物的情况下确保这些策略的成功应用至关重要.考虑到算法的简单性和低计算负担在实际应用中的重要性,Souza等人^[97]提出了一种利用人工势场法计算障碍物路径的方法,采用在三个方向上旋转的涡流场,旨在避免局部最小值和碰撞,以及在狭窄通道和障碍物相关影响阈值中的振荡.此外,作者还开发了一种技术,每架飞机分析其与障碍物和目标的位置关系,单独确定每个障碍物生成的涡流场的最佳旋转方向,以确保工作空间中的安全运动.通过这种方式,代理可以实时适应具有新型障碍物的场景.作者将算法在涉及分层合作以及存在移动目标和障碍物的场景中进行了分析,发现包含具有独立确定方向的涡流场确实有助于减少与障碍物和代理之间碰撞的风险,能够缓解无法到达目标的问题,并减少路径振荡和飞机的行驶距离.

遗传算法是一种解决优化问题的通用方法,特别是那些涉及确定最佳路径的问题.为了实现理想的解决方案,一般采取“生存策略”和“适者生存”的原则.染色体编码、种群规模、适应度函数、遗传操作和控制参数是其中五个主要的组成部分.进化算法已经在现有的多项研究中广泛应用于无人机路径规划和航空路径规划. Shen等人^[98]将路径规划构建为在三维地形场景中的多目标优化问题,并提出了一种基于多级约束处理的进化算法,以规划滑翔无人机的最短无碰撞飞行路径.该算法使用自适应约束处理机制来改进三维环境中的不同路径约束,并使用改进的自适应非支配排序遗传算法来增强算法在复杂环境中的路径规划能力. Yue等人^[99]建立了使用蚁群优化算法解决无人车辆路径规划问题的有效方法.首先,建立无人车辆路径规划的环境模型,处理和描述环境信息,并最终实现问题空间的划分.接下来,描述了蚁群算法的仿生行为.通过添加惩罚策略对蚁群算法进行了改进.该惩罚策略可以提高资源的利用率,并通过使用搜索历史中的较差值来增强信息素的挥发性,从而引导蚂蚁探索其他未知区域.

Liu等人^[100]提出一种改进Hopfield网络进行路径规划,以解决传统Hopfield神经网络的问题.首先,使用传统的A*算法选择搜索区域内符合标准的节点.然

后,将符合标准的节点转换为Hopfield神经网络中的神经元,并利用网络的稳定性迭代选择最优路径.该方法将地形信息转化为Hopfield网络表示的方法,重点创建了一种能够反映三种操作域共通的规划和任务限制的能量函数.使用基于目标点距离、地形梯度和无人机高度数据反馈的能量函数,通过真实地形数据库图像演示了路径规划能力.强调了将路径规划器整合到更大车辆规划方案中的进展.

Yang等人^[101]最初使用环境周围的粗略模型来提供初始知识,然后通过局部在线计算进行精细化.通过处理图像序列并将其与传感器数据结合进行模型更新,并依赖数字高程模型(Digital Elevation Model, DEM)^[144],创建环境的初始3D模型.最后,作者依靠传统优化技术,如Dijkstra算法,来确定计算的最短路径.所使用的代价函数是三个因素的总和,这些因素按目标的距离、地形粗糙度和飞行高度进行缩放.另一方面,精细化方法的主要思想是将局部空间划分为3D立方体单元,然后用网格点表示每个单元.其使用单元的中心作为点,然后使用深度图算法确定每个体素的深度.体素是等大小的立方体体积网格.此外,作者使用体素来更新体积图,将其标记为占用或未占用,然后将地图提供给路径规划算法以避免障碍物.

Mittal等人^[102]提出了无人机城市搜索和救援任务,与上述介绍的其他任务不同,由于可能发生的自然环境结构变化,无法利用预先存在的地图进行导航.他们首次引入了一个包含超过120万张坍塌建筑图像的合成数据集,该数据集包括真实深度、表面法线、语义和相机位置信息.设计的导航系统提供定位、建图和无碰撞路径,可以在没有先验信息的情况下探索非结构化环境.定位通过首先探索区域并使用深度图像构建地图来实现.类似于Yang等人^[101]引入的代价函数,他们提出了一种基于四个评估标准的着陆算法:地形的平整度和倾斜度、深度信息的置信度、区域的陡峭度以及着陆的强度.无人机使用机载摄像头获取的深度图像构建占据地图,构建的地图包括两种,第一种使用开源概率3D映射框架Octomap^[115]实现,是一种用于路径规划的轻量级的基于体素的本地地图.第二种地图基于开源3D网格重建工具Voxblox^[116]进行构建,用于与地面站通信以进行分析和救援规划.最后,为了避免逐帧检测的重复,对着陆点进行了聚类.如果新着陆点在列表中没有邻近的小距离邻居,则更新包括深度信息和无人机姿态的全球着陆点列表.姿态是物体位置和方向的组合.作者使用了带非线性优化的最小加速度轨迹生成器^[117]进行路径规划.通过城市规模的超真实模拟环境和坍塌建筑的实际场景实验,展示了系统的有效性.

Bashir等人^[103]提出了一种针对无人机路径规划的

无碰撞技术,旨在减少计算负载并提供短路径解决方案.如图15所示,该方法假设障碍物为矩形,并在其拐角处定义兴趣点,利用Dijkstra算法从这些兴趣点中选择允许路径,形成无碰撞轨迹.通过模拟测试,该方法验证了其在处理不同数量和形状障碍物时的计算效率和路径规划精确性.此外,该方法通过基于圆盘的不确定性模型,考虑了环境干扰及传感器误差对轨迹跟踪的影响,从而确保了路径规划的可靠性.该技术在保持低计算复杂度的同时,能够在实时规划场景中有效应用.

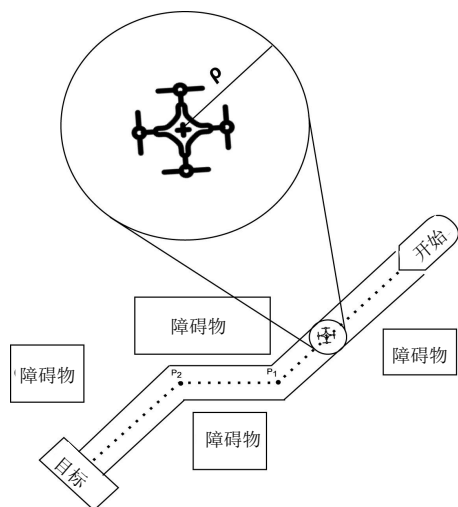


图15 无碰撞无人机路径规划方法示意图^[103]

局部路径规划方法以实时性和动态适应性为核心,适用于快速变化的复杂环境.通过多种技术,如虚拟力、人工势场、优化算法和神经网络,这些方法有效解决了路径选择中的局部最优问题和复杂动态障碍的规避难题.但各类方法也存在局部最优解问题、动态障碍处理能力有限、对环境信息高度依赖以及适用范围受限等局限性.

3.2.3 深度学习路径规划方法

由于有限的能源存储容量和充电电池的低效率,无人机必须使用最有效的机器学习算法来感知、规划和控制,以尽快执行任务.深度学习路径规划利用深度神经网络处理复杂的环境感知和决策问题,通过对海量数据的训练,能够有效提高无人机的自主导航能力.相较于传统的路径规划方法,深度学习方法具有更强的泛化能力和适应性,能够在未知或变化的环境中表现出色.在深度学习路径规划中,主要分为基于强化学习的方法和基于神经网络的方法.

(1) 基于强化学习的方法

强化学习通过与环境的交互不断学习最优策略,实现无人机的自主导航.深度强化学习(Deep Rein-

forcement Learning, DRL)结合了深度学习和强化学习的优势,通过神经网络近似值函数或策略函数,能够在高维状态空间中进行高效的路径规划和决策.基于强化学习的方法最近已成为计算能力有限的无人机的流行方法.在室内和室外环境中,不同的优化算法被应用于局部和全局路径规划策略.这些基于强化学习的方法有助于学习一种控制策略,能够在不同的功率约束下泛化无人机导航系统.如图16所示,Maciel-pearson等人^[104]使用深度强化学习,不断改进无人机代理的学习和理解能力,同时探索部分可见的环境.采用双状态输入策略,将从原始图像中获得的知识与包含位置信息的地图相结合.尽管当前场景的特征图识别出应该避免的拥挤区域,但这些位置信息增强了无人机位置及其与目标点距离的知识.通过适应双状态输入数据的深度Q网络(Deep Q-Network, DQN)变体进行了广泛测试,通过改变奖励和Q值函数,代理能够始终优于改进的深度Q网络、双深度Q网络和深度循环Q网络.结果表明,该方法能够在多个未知环境和恶劣天气条件下导航.

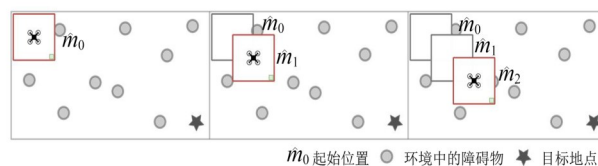


图16 扩展双深度强化学习方法路径规划示意图^[104]

深度强化学习算法受限于数据效率问题,因为通常需要大量数据才能达到合理的性能.为了加速训练过程,He等人^[105]提出了一种新颖的学习框架,将模仿学习和基于RL的方法融合在双延迟深度确定性策略梯度(Twin Delayed Deep Deterministic policy gradient, TD3)算法上^[118],以加速深度强化学习的训练过程.他们在模仿阶段使用专家演示训练策略和Q值网络.在与环境交互时,利用时间差误差和衰减模仿损失来更新预训练网络,以解决从模仿学习到强化学习转变过程中的分布不匹配问题.

为了降低电量消耗,Theile等人^[106]利用端到端的强化学习来学习一种控制策略,使其能够在不同的电力约束下泛化无人机的控制.该方法搭载于配备摄像头的无人机,用于控制执行覆盖路径规划任务,该任务具有随机的起始位置和多个着陆位置选项,并且环境中包含禁飞区.为避免小型无人机功耗的波动带来的飞行范围的变化,他们训练了一个双深度Q网络来为无人机做出控制决策,使用类似地图的输入通道通过卷积网络层代理传输空间信息,平衡受限的电力预算和覆盖目标.该方法可以应用于各种环境,并协调复杂

的目标结构与系统约束。Theile 等人^[107]还进一步拓展了该方法更广泛的场景,包括覆盖路径规划任务和收集数据任务。覆盖路径规划任务的目标是让无人机调查感兴趣的区域,而数据收集任务则是让无人机从分布的物联网传感器设备中收集数据。通过利用环境结构化地图信息,作者在两个完全不同的任务场景中使用相同架构的双深度 Q 网络(DDQN)进行训练,以在平衡各自任务目标与导航约束的情况下做出移动决策。导航控制器仅基于当前传感器数据生成控制信号,不进行优化或配置空间搜索、内存减少或计算要求。该方法利用环境的压缩全局地图结合显示无人机附近区域的裁剪但未压缩的局部地图,可以高效地扩展到大型环境中,控制策略在场景参数变化时无需重新训练。

Yin 等人^[108]提出了一种基于深度强化学习(DRL)的自适应无人机导航方法,以应对低空复杂环境中的导航挑战。研究中引入了指导注意力递归 TD3(Guide Attention Recurrent TD3, GARTD3)算法,该算法通过动态调整导航与避障任务的关注权重,显著提升了无人机的决策能力。此外,其设计了一种速度约束损失函数,用于优化无人机的速度控制,从而在提高避障能力的同时增强导航效率。实验结果表明,在三维模拟环境中,该方法在任务成功率方面提高了 14%,碰撞率降低了 14%,相较于传统 DRL 方法展现了显著的性能提升。

基于强化学习的方法通过不断与环境交互实现无人机的自主学习和导航优化,具有在动态环境中适应性强的优势。然而,这类方法对大规模数据的依赖和高计算成本在一定程度上限制了其应用场景。此外,在快速环境变化中,策略的稳定性和学习效率仍需进一步优化。

(2) 基于神经网络的方法

基于神经网络的方法通过对大量环境数据和路径规划案例的学习,建立从感知到决策的直接映射,基于神经网络的算法也被提出用于室内和室外环境。在一些未构建的环境中,例如森林,自动轨迹导航系统能够跨不同图像分辨率进行泛化。此外,其使无人机能够在具有挑战性的条件下运行广泛的传感器有效载荷能力。

室内无人机应用的主要挑战之一是导航和避障。由于在室内环境中进行,全球定位系统在精确定位和导航方面失效,Chhikara 等人^[109]利用基于深度神经网络的图像处理,促进无人机在建筑物室内走廊中的自主导航。对于深度神经网络,选择良好的超参数组合以获得更好的预测是一个复杂的任务,作者通过使用遗传算法实现了卷积神经网络的超参数调优。在高度非结构化和未知的环境中飞行,对无人机的路径规划能力提出了巨大的挑战。Menfoukh 等人^[110]提出了一种在

森林路径环境中进行自动路径规划的方法。该方法在高图像分辨率下成功泛化,使无人机能够在困难的自然条件下进行管理。通过使用代表无人机平台的高分辨率图像,优化了深度神经网络,以改进对森林路径的检测,从而进行无人机导航。Silvestrini 等人^[111]将人工智能算法应用于航天学制导、导航和控制系统,对比了常用的神经网络架构及其工作原理,并在航天领域的检索时间结构、近似扰动、封装动态行为或执行参数系统识别中进行实验。通常,循环神经网络在近似时间序列方面表现优越,但训练复杂度高。卷积神经网络被用于处理来自光学传感器的图像,将基于神经的方法与图像处理技术结合,来执行光学导航任务。姿态估计和行星着陆方法通常使用卷积神经网络,自主制导和控制的应用通常使用深度强化学习。作者还使用端到端深度学习框架学习航天器制导、导航和控制,将神经技术与传统算法结合以提高性能。为了实现无人机自主导航,Tullu 等人^[112]提出了一种实时三维路径规划器,该规划器能够通过无障碍路径引导无人机到达目的地。该路径规划器具有 A* 算法的启发式特性,但无需像 A* 方法那样在内存中存储前沿节点。该规划器依赖于检测到的物体(障碍物)的相对位置,并确定无碰撞路径。该规划器具有重量轻的优点,适用于实时快速导航。通过在受限环境中的软件、在环路测试(Software-In-The-Loop, SITL)模拟器中严格测试和初步实际飞行测试,证明了其性能效率。

Zhao 等人^[113]提出了一种基于深度学习的自主无人机探索方法(DLAE),专注于在未知环境中实现高效导航。DLAE 通过混合动作空间来优化无人机的路径规划,包括位置和偏航动作,以克服视野限制。该方法引入了无效动作屏蔽和自回归网络,显著提高了学习效率和安全性。实验表明,与传统方法相比,DLAE 在探索效率、决策时间和路径优化方面表现更优,尤其在复杂环境中展示了较强的泛化能力。

基于神经网络的方法通过自动提取高级特征和学习复杂环境中的模式,展现出强大的感知和决策能力。这些方法在面对光照变化、遮挡和非结构化环境时表现出优越性能。然而,这类方法的局限性包括高计算需求和对训练数据质量的依赖。此外,在资源受限的无人机平台上实现神经网络的轻量化部署仍是未来需要解决的问题。

3.3 对比与分析

在无人机视觉导航任务中,障碍物检测和规避方法与路径规划方法承担着不同的功能任务,前者侧重于安全性和实时性,确保无人机能够及时规避障碍物,而后者则更关注路径的全局优化和导航效率。

3.3.1 定性对比

如表 11 所示,在障碍物检测和规避方法中,基于光流的方法由于计算简单,适合实时避障;基于双目的方法在高精度环境感知中表现突出;而基于 SLAM 的方法则在未知和复杂环境中提供了强大的自主导航能力。

如表 12 所示,在路径规划方面,全局路径规划能够提供最优的整体路径,但在动态环境中可能不如局部路径规划灵活。局部路径规划强调实时反应能力,适合动态和复杂的操作环境。而神经网络路径规划方法则通过深度学习实现了环境感知与路径决策的高效融合,适用于非结构化环境下的自主导航。

表 11 障碍物检测和规避方法对比

比较维度	基于光流的方法	基于双目的方法	基于 SLAM 的方法
原理	通过分析图像序列中的像素运动,推测障碍物的位置和运动方向	利用两台相机捕获的图像,通过视差计算获取深度信息	利用传感器数据生成环境地图,识别并规避障碍物
算法精度	精度有限,特别是在纹理较少或环境光照变化明显时,鲁棒性较低	深度信息精度高,能够精确识别障碍物,鲁棒性较好,适合复杂环境	精度较高,但在特征稀少或环境变化剧烈的情况下可能出现误差
计算复杂度	计算复杂度低,适合实时处理	需要较高的计算资源进行视差计算和深度图生成,特别是高分辨率图像	计算复杂度高,需实时处理大量传感器数据和图像
适用场景	适合简单环境中的基础避障任务,如低速飞行或预定路径导航	适合需要高精度感知的复杂环境,如室内导航、森林探测	适合探索未知环境、自动建图与导航任务,如搜救任务、探险

表 12 路径规划方法对比

比较维度	全局路径规划	局部路径规划	神经网络路径规划
原理	基于已知或生成的全局地图,计算从起点到目标点的最优路径	在局部环境中实时规划路径,适应动态变化的场景	通过深度学习直接从传感器数据生成导航决策,实现端到端的路径规划
算法精度	精度高,能找到从起点到终点的最优路径,但在动态环境中的鲁棒性较差	能够快速响应环境变化,鲁棒性强,但可能会陷入局部最优	通过训练数据获得高精度,对环境变化的鲁棒性依赖于训练数据覆盖范围
计算复杂度	计算复杂度取决于地图规模和算法,如 A* 算法计算复杂度较高	相对较低的计算复杂度,主要集中在实时避障和路径优化	依赖训练数据和模型规模,推理阶段可能需要较高的计算资源
适用场景	适合在已知环境中的长距离、精确导航任务,如物流配送、区域覆盖	适合动态环境中短距离导航与实时避障任务,如城市低空飞行、密集场景	适合非结构化环境下的复杂导航任务,如自主飞行、无人驾驶

3.3.2 定量对比

为了深入分析和比较不同无人机导航方法的性能,本节针对现有文献中的实验数据进行整理和总结。尽管导航方法的实验环境在不同研究中存在差异,但通过对这些方法的实验场景(如室内、室外)和任务难度(如静态环境、动态障碍物)进行分类,可以揭示不同方法的性能特点与适用范围。

在无人机导航任务中,评价指标的选择对算法性能的衡量至关重要。这些指标不仅反映了导航系统的精度和效率,还能揭示不同算法在特定任务中的适用性。基于现有研究,下面将对无人机视觉导航常用的评价指标进行详细介绍。

成功率(Success Rate)。成功率衡量算法在给定测试场景中成功避开障碍物并完成任务的比例。成功率越高,说明方法的可靠性越强。

$$\text{Success Rate} = \frac{\text{成功次数}}{\text{总测试次数}} \quad (10)$$

路径长度(Path Length)。路径长度是无人机从起点到终点沿规划路径的总距离,用于评估路径优化的程度。无人机在未发生碰撞的情况下能够飞行的路径长

度越大,其表现越好。

$$\text{PL} = \sum_{i=1}^{n-1} \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2 + (z_{i+1} - z_i)^2} \quad (11)$$

其中, (x_i, y_i, z_i) 为路径上的每个点坐标。

均方根误差(Root Mean Square Error, RMSE)。评估实际路径与规划路径之间的偏离程度,数值越低,路径精度越高。

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (d_i - d_{\text{ref}})^2} \quad (12)$$

其中, d_i 是实际路径点到目标点的距离, d_{ref} 为规划路径点到目标点的参考距离, n 为路径点的数量。

绝对轨迹误差(Absolute Trajectory Error, ATE)。衡量实际路径与目标路径之间的整体误差,用于评估导航精度。

$$\text{ATE} = \sqrt{\frac{1}{n} \sum_{i=1}^n \|p_i^{\text{actual}} - p_i^{\text{target}}\|^2} \quad (13)$$

其中, p_i^{actual} 为实际位置, p_i^{target} 为目标位置。

平均位姿误差(Average Pose Error, APE)。用于评估路径规划过程中无人机的姿态误差,包括位置误差和旋转误差。

位置误差:

$$\text{APE}_{\text{trans}} = \frac{1}{n} \sum_{i=1}^n \|t_i^{\text{actual}} - t_i^{\text{target}}\| \quad (14)$$

旋转误差:

$$\text{APE}_{\text{rot}} = \frac{1}{n} \sum_{i=1}^n \|R_i^{\text{actual}} - R_i^{\text{target}}\| \quad (15)$$

其中, t_i 为位移位置, R_i 为旋转矩阵.

平均超体积 (average Hypervolume, HV). 衡量路径规划中多目标优化的效果, 越高表示算法能够在更多优化目标中表现出色.

平均平方误差 (Mean Square Error, MSE). 用于评估无人机在路径规划中的预测值与真实值之间的误差.

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (16)$$

其中, y_i 为真实值, \hat{y}_i 为预测值.

平均绝对误差 (Mean Absolute Error, MAE). 衡量预测值与真实值偏差的指标, 直观体现预测误差.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i) \quad (17)$$

其中, y_i 为真实值, \hat{y}_i 为预测值.

表 13 展示了不同障碍物检测和规避方法的性能对比, 涵盖了方法类别、测试环境、障碍物动态性、障碍物种类以及实验结果. 基于光流的方法^[80]在室外率达到

表 13 障碍物检测和规避方法性能对比

方法名	方法类别	测试环境	障碍物动态性	障碍物种类	结果
文献[80]	基于光流的方法	室外	静态	围栏、树木、路灯、金属楼梯	平均成功率=70.4%
文献[72]	基于双目的方法	室外	静态	树木	处理速率 51 Mb/s
文献[75]	基于 SLAM 的方法	模拟	静态+动态	—	成功率>75%
文献[76]	基于 SLAM 的方法	室内+室外	静态	靶子	绝对误差<0.06 m, 耗时<0.02 s

全局路径规划方法^[92]的测试在 $15 \text{ m} \times 15 \text{ m} \times 15 \text{ m}$ 的模拟环境中进行, 评估了 3 种地图上的路径规划性能, 其路径长度分别为 50.70 m、52.11 m 和 44.72 m, 对应的 RMSE 分别为 19.57 m、19.35 m 和 16.92 m, 表明其规划路径精度随地图复杂性的变化有所波动. 局部路径规划方法的测试场景更为多样化, 在一个七边凸多边形环境^[96]中实现了 95.57% 的平均覆盖率, 在 $2.2 \text{ m} \times 3.8 \text{ m} \times 2.0 \text{ m}$ 的空间^[97]中完成了路径长度为 5.102 7 m, 目标距离平均为 0.859 0 m 的规划任务, 而在大规模 $300 \text{ km} \times 300 \text{ km} \times 1.5 \text{ km}$ 的模拟环境^[98]中测得平均 HV 值为 0.727 3 m, 文献[99]的平均距离为 63.545 1 m. 此外, 局部路径规划方法还在室外环境中进行了测试, 如文献[101]在 PM2.5 环境中达到 MAE 为 52.04 m 和 RMSE 为 56.38 m 的导航精度, 展示了在复杂条件下的鲁棒性; 而文献[115]在室内 $25 \text{ m} \times 10 \text{ m} \times 10 \text{ m}$ 的环境中实现了绝对轨迹误差仅为 0.015 1 m 的高精度路径规划. 深

度学习路径规划方法主要在模拟和室内外环境中测试. 例如, 文献[104]在 5 km^2 和 80 km^2 的环境中分别用时 368.2 s 和 27 min 完成路径规划, 文献[106]的成功率超过 98%, 展现了深度学习在复杂环境中的高效性和可靠性. 文献[109]在室内飞行高度为 1 m 的情况下, 通过多种误差指标 (如 $\text{MSE}=0.008 2 \text{ m}$, $\text{MAE}=0.024 3 \text{ m}$, $\text{MAPE}=1.958 7 \text{ m}$) 展现了对目标路径的精准控制; 而文献[110]在室外环境中实现了 80% 的成功率, 表明其在动态环境下的适用性.

51 Mb/s, 展示了较高的实时性, 但测试仅涉及单一障碍物种类, 泛化能力有待进一步验证. 基于 SLAM 的方法表现出了对动态场景的更强适应性^[75]. 在模拟环境中对静态和动态障碍物的检测成功率超过 75%, 验证了其对于动态障碍物的适应能力. 然而, 由于其结果基于模拟环境, 可能存在与实际场景的差异. 另一种基于 SLAM 的方法^[76]在室内与室外环境中的测试表明, 该方法对靶子的检测具有较高的精度和实时性, 绝对误差小于 0.06 m, 平均耗时小于 0.02 s, 适合需要高精度导航的任务. 然而, 这一方法的测试对象主要为静态靶子, 其动态适应性尚未被验证.

总体而言, 基于光流和双目的方法在资源受限场景中表现较好, 适用于静态障碍物的检测. 而基于 SLAM 的方法则更适合复杂任务, 尤其是在动态障碍物场景中具有较高的成功率和精度. 未来的研究可以进一步验证这些方法在真实复杂环境中的性能, 尤其是动态障碍物下的表现, 同时优化其计算复杂度, 以满足更多应用场景的需求.

表 14 汇总了不同路径规划方法在性能上的定量对比, 包括全局路径规划、局部路径规划和深度学习路径规划三类, 测试环境涵盖模拟、室内和室外场景. 表 14 中详细描述了每种方法的工作空间、障碍物信息以及具体实验结果, 以展现各类方法在不同环境下的适用性和性能表现.

深度学习路径规划方法主要在模拟和室内外环境中测试. 例如, 文献[104]在 5 km^2 和 80 km^2 的环境中分别用时 368.2 s 和 27 min 完成路径规划, 文献[106]的成功率超过 98%, 展现了深度学习在复杂环境中的高效性和可靠性. 文献[109]在室内飞行高度为 1 m 的情况下, 通过多种误差指标 (如 $\text{MSE}=0.008 2 \text{ m}$, $\text{MAE}=0.024 3 \text{ m}$, $\text{MAPE}=1.958 7 \text{ m}$) 展现了对目标路径的精准控制; 而文献[110]在室外环境中实现了 80% 的成功率, 表明其在动态环境下的适用性.

总体来看, 全局路径规划方法适合更大范围的任务, 局部路径规划方法在复杂小规模环境中表现优异, 而深度学习路径规划方法则在动态环境和高维场景下表现出卓越的适应性与鲁棒性. 这些结果为进一步优化路径规划算法提供了重要参考, 也验证了不同方法在特定任务中的优劣和适用场景.

表 14 路径规划方法性能对比

方法名	方法类别	测试环境	工作空间	障碍物	结果
文献[92]	全局路径规划方法	模拟	15 m × 15 m × 15 m	√	地图 1:路径长度=50.70 m, RMSE=19.57 m 地图 2:路径长度=52.11 m, RMSE=19.35 m 地图 3:路径长度=44.72 m, RMSE=16.92 m
文献[96]	局部路径规划方法	模拟	七边凸多边形	—	平均覆盖率=95.57%
文献[97]	局部路径规划方法	模拟	2.2 m × 3.8 m × 2.0 m	—	路径长度=5.102 7 m, 平均目标距离=0.859 0 m
文献[98]	局部路径规划方法	模拟	300 km × 300 km × 1.5 km	√	平均 HV=0.727 3 m
文献[99]	局部路径规划方法	模拟	—	√	平均距离=63.545 1 m
文献[101]	局部路径规划方法	模拟+室外	室外:80 m × 100 m × 100 m	—	模拟:路径长度=278.33 m, 室外:PM2.5 环境 MAE=52.04 m, RMSE=56.38 m
文献[114]	局部路径规划方法	模拟	70.7 km ²	√	路径长度=8 774 m
文献[115]	局部路径规划方法	室内	25 m × 10 m × 10 m	√	ATE=0.015 1 m
文献[116]	局部路径规划方法	模拟	20 m × 20 m × 8 m × 40 m	√	APE _{trans} =0.3 m, APE _{rot} =0.9° APE _{trans} =0.3 m, APE _{rot} =0.9°
文献[104]	深度学习路径规划方法	模拟	5 km ² , 80 km ²	√	用时=368.2 s, 用时=27 min
文献[106]	深度学习路径规划方法	模拟	—	√	成功率>98%
文献[109]	深度学习路径规划方法	室内	飞行高度 1 m	√	MSE=0.008 2 m, MAE=0.024 3 m, MAPE=1.958 7 m
文献[110]	深度学习路径规划方法	室外	—	—	成功率=80%

4 挑战与未来展望

经过对无人机视觉定位和导航任务进行详细的回顾之后,发现要将这些方法真正应用在 GNSS 拒止环境下还存在一些挑战和机遇。

(1) 公共可用数据集. 无人机视觉定位与导航的有监督学习需要可靠的样本和标签. 然而, 该领域缺乏现实世界和开源的定位导航数据集. 因此, 需要开发数据集来对新作品与文献中现有工作进行基准测试, 这是该领域进步和发展的前提. 近期 Xu 等人^[119]公开了一个无人机定位数据集 UAV-VisLoc, 包含了 11 个场景的卫星图像和真实的无人机飞行数据, 这将为无人机定位的发展带来巨大贡献. 此外, 需要更多种类和规模的数据集, 覆盖不同地形和环境条件, 以便全面评估算法的性能和鲁棒性. 数据集的多样性和质量直接影响模型的泛化能力和实际应用效果. 数据集还需包含多模态信息, 如视觉、激光雷达、惯性测量单元 (IMU) 等多源传感器数据, 以便训练多模态融合模型. 此外, 构建支持多任务学习的数据集 (如同时包含定位、避障和目标识别任务) 将进一步推动该领域的发展。

(2) 硬件加速. 目前的无人机定位与导航方案大多是在离线环境下进行实现的, 很少有作者探索嵌入式和实时定位导航解决方案. 大多数 GNSS 拒止的自主导航, 包含地图构建任务, 例如基于先验的导航, 需要密集的机载计算资源, 导致比无地图导航花费更多的时

间. 为了减少计算和内存需求并提高实时性, 需要使用 FPGA 等软件定义的硬件加速器来加速地图构建过程. 除此之外, 先进的硬件加速技术, 如 GPU、TPU, 也能够显著提升深度学习模型的推理速度, 使无人机能够在飞行中实时处理大量视觉数据, 从而提高导航和定位的精度和效率。

(3) 环境复杂性. 无人机在现实世界中飞行时, 会遇到各种复杂和动态的环境, 如天气变化、障碍物干扰和光照条件变化等. 这些因素都会对无人机视觉定位和导航系统提出极高的要求. 开发能够适应多变环境的鲁棒算法是当前的一大挑战. 例如, 低光环境、强烈的阳光反射、雨雪天气等都可能影响传感器的数据质量, Wang 等人^[37]已经关注到了这个问题, 并使用风格迁移技术生成不同光照和天气的数据, 对原始数据集进行数据增强来提升系统的鲁棒性, 但是生成数据与真实数据的分布之间仍存在差异, 也没有在真实恶劣天气环境中进行测试, 因此环境负责性的问题仍待解决, 需要设计专门的预处理和增强方法, 并在真实环境中进行测试。

(4) 多模态融合的探索. 单一视觉信号可能因天气条件或障碍物遮挡而失效, 未来可探索多模态传感器的融合技术. 结合激光雷达、声波雷达、红外摄像头等多源信息, 不仅可以提升感知精度, 还能显著增强系统的鲁棒性. 开发轻量化多模态感知算法, 使无人机在传感器部分失效的情况下仍能完成任务, 是未来研究的重要方向。

(5)实时性要求.无人机的视觉定位和导航需要在飞行过程中实时处理大量图像和传感数据,以确保其能够及时做出反应.实现高效的实时计算是一个重大的技术挑战.为了保证实时性,除了硬件加速之外,还需要优化算法结构,减少计算复杂度,并利用边缘计算和分布式计算技术,将部分计算任务分配到云端或地面站进行处理,以减轻无人机的计算负担.

(6)能源限制.无人机的电池容量有限,这对其搭载的计算设备和传感器的功耗提出了严格限制.如何在有限的能源条件下实现高效的计算和精确的导航,是一个需要深入研究的问题.为了延长续航时间,需要设计低功耗的算法和硬件,同时可以探索能源管理策略,例如在非关键时刻关闭部分传感器或计算模块,或通过能量回收技术来延长电池寿命.

(7)模拟器到真实世界的泛化.Gazebo和Unreal Engine等环境模拟器结合ROS和AirSim等飞行模拟器能够收集数据,因为它们提供了各种类似真实的环境,也用于评估以降低无人机更换部件的成本.然而,由风干扰、光照变化和运动物体等模拟器提供的虚拟环境与现实世界的环境之间有相当大的差距.因此,通过开发具有各种环境条件和约束的高保真模拟器来减少模拟器到真实世界的泛化限制也是一个重要的挑战.

5 结论

基于视觉的无人机定位与导航方法作为补充或取代卫星导航的有效途径,正处在蓬勃发展阶段.本综述对基于视觉的无人机定位与导航方法进行了全面的回顾和分析,系统梳理了相关技术的研究现状和发展趋势.首先,详细介绍了无人机定位的基于图像检索和图像匹配的两类方法,探讨了它们在不同场景下的适用性及其技术特点,并梳理了常用的公开数据集和评价指标,同时对现有研究结果进行定量与定性分析,为研究者提供了重要的基准和参考.其次,介绍了无人机导航的两类关键技术,即障碍物检测和规避技术以及路径规划技术,揭示了该领域的最新研究进展和主要技术手段,通过对不同方法的定量与定性对比分析,本文明确了现阶段无人机导航技术的优势与局限.同时,本文还讨论了这些方法在公共数据集、硬件加速、环境复杂性、多模态融合、实时性要求、能源限制及模拟器到真实世界的泛化等方面面临的挑战.尽管基于视觉的定位与导航方法在无人机自主飞行中展现出了巨大潜力,但仍有许多技术难题亟待解决.未来的研究应着眼于进一步提高系统的鲁棒性、精度和效率,开发更为先进的算法,并在复杂的实际环境中进行验证和优化.通过不断的技术创新和改

进,基于视觉的无人机定位与导航方法必将推动无人机在各个应用领域中的广泛应用和发展.

参考文献

- [1] 朱得糠,李东泽,郭鸿博,等.无人机视觉地理定位研究综述[J].导航与控制,2023,22(3):21-33,20.
ZHU D K, LI D Z, GUO H B, et al. A review of vision-based geolocation using UAVs[J]. Navigation and Control, 2023, 22(3): 21-33, 20. (in Chinese)
- [2] 冷佳旭,莫梦竟成,周应华,等.无人机视角下的目标检测研究进展[J].中国图象图形学报,2023,28(9):2563-2586.
LENG J X, MO M J C, ZHOU Y H, et al. Recent advances in drone-view object detection[J]. Journal of Image and Graphics, 2023, 28(9): 2563-2586. (in Chinese)
- [3] 张智,易华挥,郑锦.聚焦小目标的航拍图像目标检测算法[J].电子学报,2023,51(4):944-955.
ZHANG Z, YI H H, ZHENG J. Focusing on small objects detector in aerial images[J]. Acta Electronica Sinica, 2023, 51(4): 944-955. (in Chinese)
- [4] 郑锦,蒋博韬,彭微,等.LiDAR点云指导下特征分布趋同与语义关联的3D目标检测[J].电子学报,2024,52(5):1700-1715.
ZHENG J, JIANG B T, PENG W, et al. 3D object detection based on feature distribution convergence guided by LiDAR point cloud and semantic association[J]. Acta Electronica Sinica, 2024, 52(5): 1700-1715. (in Chinese)
- [5] 马宁,曹云峰.面向无人机自主着陆的视觉感知与位姿估计方法综述[J].自动化学报,2024,50(7):1284-1304.
MA N, CAO Y F. A survey on vision-based sensing and pose estimation methods for UAV autonomous landing[J]. Acta Automatica Sinica, 2024, 50(7): 1284-1304. (in Chinese)
- [6] 刘玄冰,周绍磊,肖支才,等.无人机避障方法研究综述[J].兵器装备工程学报,2022,43(5):40-47.
LIU X B, ZHOU S L, XIAO Z C, et al. Review on UAV obstacle avoidance methods[J]. Journal of Ordnance Equipment Engineering, 2022, 43(5): 40-47. (in Chinese)
- [7] 王从宝,张安思,杨磊,等.基于深度视觉的四旋翼无人机自主飞行感知和避障综述[J].无线电工程,2023,53(10):2233-2243.
WANG C B, ZHANG A S, YANG L, et al. A review of deep vision-based autonomous flight perception and obstacle avoidance for quadrotor UAVs[J]. Radio Engineering, 2023, 53(10): 2233-2243. (in Chinese)
- [8] KEANE J F, CARR S S. A brief history of early unmanned aircraft[J]. Johns Hopkins APL Technical Digest,

- 2013, 32(3): 558-571.
- [9] CANDIAGO S, REMONDINO F, DE GIGLIO M, et al. Evaluating multispectral images and vegetation indices for precision farming applications from UAV images[J]. *Remote Sensing*, 2015, 7(4): 4026-4047.
- [10] QU Y Y, JIANG L, GUO X P. Moving vehicle detection with convolutional networks in UAV videos[C]//2016 2nd International Conference on Control, Automation and Robotics (ICCAR). Piscataway: IEEE, 2016: 225-229.
- [11] TURNER I L, HARLEY M D, DRUMMOND C D. UAVs for coastal surveying[J]. *Coastal Engineering*, 2016, 114: 19-24.
- [12] FRASER R H, OLTHOF I, LANTZ T C, et al. UAV photogrammetry for mapping vegetation in the low-Arctic[J]. *Arctic Science*, 2016, 2(3): 79-102.
- [13] AKHLOUFI M A, COUTURIER A, CASTRO N A. Unmanned aerial vehicles for wildland fires: Sensing, perception, cooperation and assistance[J]. *Drones*, 2021, 5(1): 15.
- [14] SCHERER J, YAHYANEJAD S, HAYAT S, et al. An autonomous multi-UAV system for search and rescue[C]// *Proceedings of the First Workshop on Micro Aerial Vehicle Networks, Systems, and Applications for Civilian Use*. New York: ACM, 2015: 33-38.
- [15] COUTURIER A, AKHLOUFI M A. A review on absolute visual localization for UAV[J]. *Robotics and Autonomous Systems*, 2021, 135: 103666.
- [16] BALAMURUGAN G, VALARMATHI J, NAIDU V P S. Survey on UAV navigation in GPS denied environments[C]// *2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPE)*. Piscataway: IEEE, 2016: 198-204.
- [17] LU Y C, XUE Z C, XIA G S, et al. A survey on vision-based UAV navigation[J]. *Geo-spatial Information Science*, 2018, 21(1): 21-32.
- [18] GYAGENDA N, HATILIMA J V, ROTH H, et al. A review of GNSS-independent UAV navigation techniques[J]. *Robotics and Autonomous Systems*, 2022, 152: 104069.
- [19] JUN M, ROUMELIOTIS S I, SUKHATME G S. State estimation of an autonomous helicopter using Kalman filtering[C]// *Proceedings 1999 IEEE/RSJ International Conference on Intelligent Robots and Systems. Human and Environment Friendly Robots with High Intelligence and Emotional Quotients*. Piscataway: IEEE, 1999: 1346-1353.
- [20] OSHMAN Y, SHAVIV I. Optimal tuning of a Kalman filter using genetic algorithms[C]// *AIAA Guidance, Navigation, and Control Conference and Exhibit*. Reston: AIAA, 2000: 4558.
- [21] SASIADEK J, WANG Q, JOHNSON R, et al. UAV navigation based on parallel extended Kalman filter[C]// *AIAA Guidance, Navigation, and Control Conference and Exhibit*. Reston: AIAA, 2000: 4165.
- [22] HOFMANN-WELLENHOF B, LICHTENEGGER H, WASLE E. GNSS--Global Navigation Satellite Systems: GPS, GLONASS, Galileo, and More[M]. Wien: Springer, 2008.
- [23] VAN DALEN G J, MAGREE D P, JOHNSON E N. Absolute localization using image alignment and particle filtering[C]// *AIAA Guidance, Navigation, and Control Conference*. Reston: AIAA, 2016: 0647.
- [24] MAGREE D P, JOHNSON E N. A monocular vision-aided inertial navigation system with improved numerical stability[C]// *AIAA Guidance, Navigation, and Control Conference*. Reston: AIAA, 2015: 0097.
- [25] YOL A, DELABARRE B, DAME A, et al. Vision-based absolute localization for unmanned aerial vehicles[C]// *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*. Piscataway: IEEE, 2014: 3429-3434.
- [26] WAN X, LIU J G, YAN H S, et al. Illumination-invariant image matching for autonomous UAV localisation based on optical sensing[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2016, 119: 198-213.
- [27] ALURI R C. Localization of UAVs Using Computer Vision in a GPS-Denied Environment[D]. Denton: University of North Texas Libraries, 2022.
- [28] ZHENG Z D, WEI Y C, YANG Y. University-1652: A multi-view multi-source benchmark for drone-based geo-localization[C]// *Proceedings of the 28th ACM International Conference on Multimedia*. New York: ACM, 2020: 1395-1403.
- [29] WANG T Y, ZHENG Z D, YAN C G, et al. Each part matters: Local patterns facilitate cross-view geo-localization[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2022, 32(2): 867-879.
- [30] DAI M, HU J H, ZHUANG J D, et al. A transformer-based feature segmentation and region alignment method for UAV-view geo-localization[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2022, 32(7): 4376-4389.
- [31] LIN J L, ZHENG Z D, ZHONG Z, et al. Joint representation learning and keypoint detection for cross-view geo-

- localization[J]. *IEEE Transactions on Image Processing*, 2022, 31: 3780-3792.
- [32] ZHU R Z, YIN L, YANG M Z, et al. SUES-200: A multi-height multi-scene cross-view image benchmark across drone and satellite[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2023, 33(9): 4825-4839.
- [33] SHEN T R, WEI Y M, KANG L, et al. MCCG: A ConvNeXt-based multiple-classifier method for cross-view geo-localization[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2024, 34(3): 1456-1468.
- [34] LI S L, HU M, XIAO X W, et al. Patch similarity self-knowledge distillation for cross-view geo-localization[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2024, 34(6): 5091-5103.
- [35] CHEN Q, WANG T Y, YANG Z H, et al. SDPL: Shifting-dense partition learning for UAV-view geo-localization[EB/OL]. (2024-07-06)[2025-03-31]. <https://arxiv.org/abs/2403.04172v2>.
- [36] ZHAO H, REN K Y, YUE T Y, et al. TransFG: A cross-view geo-localization of satellite and UAVs imagery pipeline using transformer-based feature aggregation and gradient guidance[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 62: 4700912.
- [37] WANG T Y, ZHENG Z D, SUN Y Q, et al. Multiple-environment self-adaptive network for aerial-view geo-localization[J]. *Pattern Recognition*, 2024, 152: 110363.
- [38] XIAO J H, TORTEI D, ROURA E, et al. Long-range UAV thermal geo-localization with satellite imagery[C]//2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Piscataway: IEEE, 2023: 5820-5827.
- [39] HE M F, LIU J C, GU P F, et al. Leveraging map retrieval and alignment for robust UAV visual geo-localization[J]. *IEEE Transactions on Instrumentation and Measurement*, 2024, 73: 2523113.
- [40] SEEMA B, KUMAR H, NAIDU V. Geo-registration of aerial images using RANSAC algorithm[J]. *NCTAESD-2014*, 2014, 25(6): 234-238.
- [41] SARANYA K C, NAIDU V P S, SINGHAL V, et al. Application of vision based techniques for UAV position estimation[C]//2016 International Conference on Research Advances in Integrated Navigation Systems (RAINS). Piscataway: IEEE, 2016: 1-5.
- [42] SHAN M, WANG F, LIN F, et al. Google map aided visual navigation for UAVs in GPS-denied environment[C]//2015 IEEE International Conference on Robotics and Biomimetics (ROBIO). Piscataway: IEEE, 2015: 114-119.
- [43] CHIU H P, DAS A, MILLER P, et al. Precise vision-aided aerial navigation[C]//2014 IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway: IEEE, 2014: 688-695.
- [44] MANTELLI M, PITTOL D, NEULAND R, et al. A novel measurement model based on abBRIEF for global localization of a UAV over satellite images[J]. *Robotics and Autonomous Systems*, 2019, 112: 304-319.
- [45] MASSELLI A, HANTEN R, ZELL A. Localization of unmanned aerial vehicles using terrain classification from aerial images[M]//*Intelligent Autonomous Systems 13*. Cham: Springer International Publishing, 2015: 831-842.
- [46] AKHLOUFI M A, COUTURIER A. Relative visual localization (RVL) for UAV navigation[C]//*Degraded Environments: Sensing, Processing, and Display 2018*. Orlando: SPIE, 2018: 213-226.
- [47] AMER K, SAMY M, ELHAKIM R, et al. Convolutional neural network-based deep urban signatures with application to drone localization[C]//2017 IEEE International Conference on Computer Vision Workshops (ICCVW). Piscataway: IEEE, 2017: 2138-2145.
- [48] NASSAR A, AMER K, ELHAKIM R, et al. A deep CNN-based framework for enhanced aerial imagery registration with applications to UAV geolocalization[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Piscataway: IEEE, 2018: 1-11.
- [49] NASSAR A, ELHELW M. Aerial imagery registration using deep learning for UAV geolocalization[M]//*Deep Learning in Computer Vision*. First edition. Boca Raton, FL: CRC Press/Taylor and Francis, 2020: 183-210.
- [50] MARCU A, COSTEA D, SLUSANSCHI E, et al. A multi-stage multi-task neural network for aerial scene interpretation and geolocalization[EB/OL]. (2018-08-04)[2025-03-31]. <https://arxiv.org/abs/1804.01322v1>.
- [51] SCHLEISS M. Translating aerial images into street-map-like representations for visual self-localization of uavs[J]. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2019, 42: 575-580.
- [52] GURGU M M, QUERALTA J P, WESTERLUND T. Vision-based GNSS-free localization for UAVs in the wild[C]//2022 7th International Conference on Mechanical Engineering and Robotics Research (ICMERR). Pisca-

- taway: IEEE, 2022: 7-12.
- [53] LI H L, WANG J N, WEI Z W, et al. Jointly optimized global-local visual localization of UAVs[EB/OL]. (2023-10-12)[2025-03-31]. <https://arxiv.org/abs/2310.08082v1>.
- [54] LUO X B, WAN X, GAO Y X, et al. JointLoc: A real-time visual localization framework for planetary UAVs based on joint relative and absolute pose estimation[EB/OL]. (2024-05-13)[2025-03-31]. <https://arxiv.org/abs/2405.07429v1>.
- [55] WARREN M, GREEFF M, PATEL B, et al. There's No place like home: Visual teach and repeat for emergency return of multirotor UAVs during GPS failure[J]. *IEEE Robotics and Automation Letters*, 2019, 4(1): 161-168.
- [56] GOFORTH H, LUCEY S. GPS-denied UAV localization using pre-existing satellite imagery[C]//2019 International Conference on Robotics and Automation (ICRA). Piscataway: IEEE, 2019: 2974-2980.
- [57] HARRIS C, STEPHENS M. A combined corner and edge detector[C]//Proceedings of the Alvey Vision Conference 1988. Plessey Research Roke Manor, U K: Alvey Vision Club, 1988: 10-5244.
- [58] ROSTEN E, DRUMMOND T. Machine Learning for High-Speed Corner Detection[M]//Computer Vision-ECCV 2006. Berlin: Springer Berlin Heidelberg, 2006: 430-443.
- [59] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. *International Journal of Computer Vision*, 2004, 60(2): 91-110.
- [60] CALONDER M, LEPETIT V, STRECHA C, et al. BRIEF: Binary robust independent elementary features[M]//Computer Vision-ECCV 2010. Berlin: Springer Berlin Heidelberg, 2010: 778-792.
- [61] SARLIN P E, DETONE D, MALISIEWICZ T, et al. SuperGlue: Learning feature matching with graph neural networks[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2020: 4938-4947.
- [62] DAI M, ZHENG E H, FENG Z H, et al. Vision-based UAV self-positioning in low-altitude urban environments[J]. *IEEE Transactions on Image Processing*, 2023, 33: 493-508.
- [63] AMELI Z, AREMANDA Y, FRIESS W A, et al. Impact of UAV hardware options on bridge inspection mission capabilities[J]. *Drones*, 2022, 6(3): 64.
- [64] STRÜBBES, STÜRZLW, EGELHAAFM. Insect-inspired self-motion estimation with dense flow fields: An adaptive matched filter approach[J]. *PLoS One*, 2015, 10(8): e0128413.
- [65] RUFFIER F, VIOLLET S, AMIC S, et al. Bio-inspired optical flow circuits for the visual guidance of micro air vehicles[C]//Proceedings of the 2003 International Symposium on Circuits and Systems, 2003. ISCAS'03. Piscataway: IEEE, 2003: 3.
- [66] HE Z H, IYER R V, CHANDLER P R. Vision-based UAV flight control and obstacle avoidance[C]//2006 American Control Conference. Piscataway: IEEE, 2006: 5.
- [67] LIN H Y, PENG X Z. Autonomous quadrotor navigation with vision based obstacle avoidance and path planning[J]. *IEEE Access*, 2021, 9: 102450-102459.
- [68] FARNEBÄCK G. Two-Frame Motion Estimation Based on Polynomial Expansion[M]//Image Analysis. Berlin: Springer Berlin Heidelberg, 2003: 363-370.
- [69] HRABAR S. 3D path planning and stereo-based obstacle avoidance for rotorcraft UAVs[C]//2008 IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway: IEEE, 2008: 807-814.
- [70] MATTHIES L, BROCKERS R, KUWATA Y, et al. Stereo vision-based obstacle avoidance for micro air vehicles using disparity space[C]//2014 IEEE International Conference on Robotics and Automation (ICRA). Piscataway: IEEE, 2014: 3242-3249.
- [71] TIJMONS S, DE CROON G C H E, REMES B D W, et al. Obstacle avoidance strategy using onboard stereo vision on a flapping wing MAV[J]. *IEEE Transactions on Robotics*, 2017, 33(4): 858-874.
- [72] GRINBERG M, RUF B. UAV use case: Real-time obstacle avoidance system for unmanned aerial vehicles based on stereo vision[M]//Towards Ubiquitous Low-power Image Processing Platforms. Cham: Springer International Publishing, 2020: 139-149.
- [73] BAI G H, XIANG X J, ZHU H Y, et al. Research on obstacles avoidance technology For UAV based on improved PTAM algorithm[C]//2015 IEEE International Conference on Progress in Informatics and Computing (PIC). Piscataway: IEEE, 2015: 543-550.
- [74] ESRAFILIAN O, TAGHIRAD H D. Autonomous flight and obstacle avoidance of a quadrotor by monocular SLAM[C]//2016 4th International Conference on Robotics and Mechatronics (ICROM). Piscataway: IEEE, 2016: 240-245.
- [75] POTENA C, NARDI D, PRETTO A. Joint vision-based navigation, control and obstacle avoidance for UAVs in dynamic environments[C]//2019 European Conference on Mobile Robots (ECMR). Piscataway: IEEE, 2019: 1-7.

- [76] YANG L Y, XIAO B, ZHOU Y, et al. A robust real-time vision based GPS-denied navigation system of UAV[C]//2016 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER). Piscataway: IEEE, 2016: 321-326.
- [77] GOSIEWSKI Z, CIESLUK J, AMBROZIAK L. Vision-based obstacle avoidance for unmanned aerial vehicles[C]//2011 4th International Congress on Image and Signal Processing. Piscataway: IEEE, 2011: 2020-2025.
- [78] HAAG J, DENK W, BORST A. Fly motion vision is based on Reichardt detectors regardless of the signal-to-noise ratio[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2004, 101(46): 16333-16338.
- [79] BERTRAND O J N, LINDEMANN J P, EGELHAAF M. A bio-inspired collision avoidance model based on spatial information derived from motion detectors leads to common routes[J]. *PLoS Computational Biology*, 2015, 11(11): e1004339.
- [80] MORENO-ARMENDÁRIZ M A, CALVO H. Visual SLAM and obstacle avoidance in real time for mobile robots navigation[C]//2014 International Conference on Mechatronics, Electronics and Automotive Engineering. Piscataway: IEEE, 2014: 44-49.
- [81] PENG X Z, LIN H Y, DAI J M. Path planning and obstacle avoidance for vision guided quadrotor UAV navigation[C]//2016 12th IEEE International Conference on Control and Automation (ICCA). Piscataway: IEEE, 2016: 984-989.
- [82] SANCHEZ-RODRIGUEZ J P, ACEVES-LOPEZ A. A survey on stereo vision-based autonomous navigation for multi-rotor MUAVs[J]. *Robotica*, 2018, 36(8): 1225-1243.
- [83] SHARMA P S, CHITALIYA D N G. Obstacle avoidance using stereo vision: A survey[J]. *International Journal of Innovative Research in Computer and Communication Engineering*, 2015, 3(1): 24-29.
- [84] ZHANG J W, WANG X, BAI X, et al. Revisiting domain generalized stereo matching networks from a feature consistency perspective[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2022: 12991-13001.
- [85] ZHANG J W, LI J H, HUANG L, et al. Robust synthetic-to-real transfer for stereo matching[C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2024: 20247-20257.
- [86] ZHANG J W, HUANG L, BAI X, et al. Exploring the usage of pre-trained features for stereo matching[J]. *International Journal of Computer Vision*, 2024, 132(10): 4305-4326.
- [87] 周晓清, 王翔, 郑锦, 等. 基于自适应空间稀疏化的高效多视图立体匹配[J]. *电子学报*, 2023, 51(11): 3079-3091. ZHOU X Q, WANG X, ZHENG J, et al. Adaptive spatial sparsification for efficient multi-view stereo matching[J]. *Acta Electronica Sinica*, 2023, 51(11): 3079-3091. (in Chinese)
- [88] VACHTSEVANOS G, KIM W, AL-HASAN S, et al. Autonomous vehicles: From flight control to mission planning using fuzzy logic techniques[C]//Proceedings of 13th International Conference on Digital Signal Processing. Piscataway: IEEE, 1997: 977-981.
- [89] ROUSE D M. Route planning using pattern classification and search techniques[C]//Proceedings of the IEEE National Aerospace and Electronics Conference. Piscataway: IEEE, 1989: 2015-2020.
- [90] SZCZERBA R J, GALKOWSKI P, GLICKTEIN I S, et al. Robust algorithm for real-time route planning[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2000, 36(3): 869-878.
- [91] STENTZ A. Optimal and efficient path planning for partially-known environments[C]//Proceedings of the 1994 IEEE International Conference on Robotics and Automation. Piscataway: IEEE, 1994: 3310-3317.
- [92] BELGE E, ALTAN A, HACIOĞLU R. Metaheuristic optimization-based path planning and tracking of quadcopter for payload hold-release mission[J]. *Electronics*, 2022, 11(8): 1208.
- [93] ZHANG Q, MA J C, LIU Q. Path planning based quadtree representation for mobile robot using hybrid-simulated annealing and ant colony optimization algorithm[C]//Proceedings of the 10th World Congress on Intelligent Control and Automation. Piscataway: IEEE, 2012: 2537-2542.
- [94] ANDERT F, ADOLF F. Online world modeling and path planning for an unmanned helicopter[J]. *Autonomous Robots*, 2009, 27(3): 147-164.
- [95] YAN M, AUN CHAN C, GYGAX A F, et al. Efficient generation of optimal UAV trajectories with uncertain obstacle avoidance in MEC networks[J]. *IEEE Internet of Things Journal*, 2024, 11(23): 38380-38392.
- [96] WANG X, TAN G Z, LU F L, et al. A molecular force field-based optimal deployment algorithm for UAV

- swarm coverage maximization in mobile wireless sensor network[J]. *Processes*, 2020, 8(3): 369.
- [97] SOUZA R M J A, LIMA G V, MORAIS A S, et al. Modified artificial potential field for the path planning of aircraft swarms in three-dimensional environments[J]. *Sensors*, 2022, 22(4): 1558.
- [98] SHEN Y, ZHU Y L, KANG H W, et al. UAV path planning based on multi-stage constraint optimization[J]. *Drones*, 2021, 5(4): 144.
- [99] YUE L W, CHEN H N. Unmanned vehicle path planning using a novel ant colony algorithm[J]. *EURASIP Journal on Wireless Communications and Networking*, 2019, 2019(1): 136.
- [100] LIU Y X, XU W. Application of improved hopfield neural network in path planning[J]. *Journal of Physics: Conference Series*, 2020, 1544(1): 012154.
- [101] YANG L K, FAN S R, YU B G, et al. A coverage sampling path planning method suitable for UAV 3D space atmospheric environment detection[J]. *Atmosphere*, 2022, 13(8): 1321.
- [102] MITTAL M, MOHAN R, BURGARD W, et al. Vision-based autonomous UAV navigation and landing for urban search and rescue[M]//*Robotics Research*. Cham: Springer International Publishing, 2022: 575-592.
- [103] BASHIR N, BOUDJIT S, DAUPHIN G, et al. An obstacle avoidance approach for UAV path planning[J]. *Simulation Modelling Practice and Theory*, 2023, 129: 102815.
- [104] MACIEL-PEARSON B G, MARCHEGANI L, AKCAY S, et al. Online deep reinforcement learning for autonomous UAV navigation and exploration of outdoor environments[EB/OL]. (2019-11-11) [2025-03-31]. <https://arxiv.org/abs/1912.05684v1>.
- [105] HE L, AOUF N, WHIDBORNE J F, et al. Deep reinforcement learning based local planner for UAV obstacle avoidance using demonstration data[EB/OL]. (2020-08-06) [2025-03-31]. <https://arxiv.org/abs/2008.02521v1>.
- [106] THEILE M, BAYERLEIN H, NAI R, et al. UAV coverage path planning under varying power constraints using deep reinforcement learning[C]//2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Piscataway: IEEE, 2020: 1444-1449.
- [107] THEILE M, BAYERLEIN H, NAI R, et al. UAV path planning using global and local map information with deep reinforcement learning[C]//2021 20th International Conference on Advanced Robotics (ICAR). Piscataway: IEEE, 2021: 539-546.
- [108] YIN Y F, WANG Z T, ZHENG L L, et al. Autonomous UAV navigation with adaptive control based on deep reinforcement learning[J]. *Electronics*, 2024, 13(13): 2432.
- [109] CHHIKARA P, TEKCHANDANI R, KUMAR N, et al. DCNN-GA: A deep neural net architecture for navigation of UAV in indoor environment[J]. *IEEE Internet of Things Journal*, 2021, 8(6): 4448-4460.
- [110] MENFOUKH K, TOUBA M M, KHENFRI F, et al. Optimized Convolutional Neural Network architecture for UAV navigation within unstructured trail[C]//1st International Conference on Communications, Control Systems and Signal Processing (CCSSP). Piscataway: IEEE, 2020: 211-214.
- [111] SILVESTRINI S, LAVAGNA M. Deep learning and artificial neural networks for spacecraft dynamics, navigation and control[J]. *Drones*, 2022, 6(10): 270.
- [112] TULLU A, ENDALE B, WONDOSSEN A, et al. Machine learning approach to real-time 3D path planning for autonomous navigation of unmanned aerial vehicle[J]. *Applied Sciences*, 2021, 11(10): 4706.
- [113] ZHAO Y M, ZHANG J L, ZHANG C Y. Deep-learning based autonomous-exploration for UAV navigation[J]. *Knowledge-Based Systems*, 2024, 297: 111925.
- [114] LIANG H J, BAI H Y, SUN R, et al. Three-dimensional path planning based on DEM[C]//2017 36th Chinese Control Conference (CCC). Piscataway: IEEE, 2017: 5980-5987.
- [115] LI Z T, ZHAO J N, ZHOU X, et al. RTSDM: A real-time semantic dense mapping system for UAVs[J]. *Machines*, 2022, 10(4): 285.
- [116] CHEN S Y, ZHOU W F, YANG A S, et al. An end-to-end UAV simulation platform for visual SLAM and navigation[J]. *Aerospace*, 2022, 9(2): 48.
- [117] LU S, DING B X, LI Y M. Minimum-jerk trajectory planning pertaining to a translational 3-degree-of-freedom parallel manipulator through piecewise quintic polynomials interpolation[J]. *Advances in Mechanical Engineering*, 2020, 12(3): 168781402091366.
- [118] YU J M, SUN H, SUN J Q. Improved twin delayed deep deterministic policy gradient algorithm based real-time trajectory planning for parafoil under complicated constraints[J]. *Applied Sciences*, 2022, 12(16): 8189.
- [119] XU W J, YAO Y X, CAO J Q, et al. UAV-VisLoc: A large-scale dataset for UAV visual localization[EB/OL]. (2024-05-20) [2025-03-31]. <https://arxiv.org/abs/2405.11936v1>.

作者简介



谷美颖 女,1998年12月出生于山东省德州市.现为北京航空航天大学博士研究生.主要研究方向为视觉定位.

E-mail: gumeiyang@buaa.edu.cn



李航 男,1998年4月出生于河南省洛阳市.现为北京航空航天大学博士研究生.主要研究方向为位姿估计、三维重建.

E-mail: by2306053@buaa.edu.cn



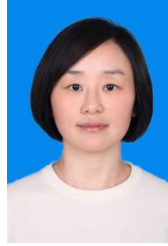
张家伟 男,1998年1月出生于河北省石家庄市.现为北京航空航天大学博士研究生.主要研究方向为立体匹配、三维重建.

E-mail: byzhangjw@buaa.edu.cn



百晓 男,1979年3月出生于北京市.现为北京航空航天大学计算机学院教授、博士生导师.主要研究方向为模式识别、计算机视觉、图像处理等.

E-mail: baixiao@buaa.edu.cn



郑锦 女,1978年10月出生于四川省乐山市.现为北京航空航天大学计算机学院副教授、博士生导师.主要研究方向为计算机视觉、视频图像处理等.

E-mail: JinZheng@buaa.edu.cn