

分类号: TP242.6

密 级: 公开

学校代码: 10363

学 号: 2220320101



安徽工程大学

Anhui Polytechnic University

硕士学位论文

题目 基于 RGB-D SLAM 的无人机
自主导航系统研究

论 文 作 者 王世康

校 内 导 师 王冠凌

校 外 导 师 毛德平

专业学位类别 电子信息 (085400)

研究方向 (领域) 控制工程与自动化技术

论文提交日期: 2025 年 05 月 16 日

分类号：TP242.6

单位代码：10363

密 级：公 开

学 号:2220320101

题 目 基于 RGB-D SLAM 的无人机自主导航系统研究

题 目 Research on Autonomous Navigation system of UAV
Based on RGB-D SLAM

学生姓名： 王世康

校内导师： 王冠凌

校外导师： 毛德平

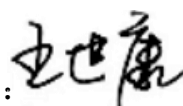
专 业： 电子信息

研究方向： 控制工程与自动化技术

论文答辩日期： 2025.05.12

安徽工程大学学位论文原创性声明

本人郑重声明：我恪守学术道德，崇尚严谨学风。所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已明确注明和引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品及成果的内容。论文为本人亲自撰写，我对所写的内容负责，并完全意识到本声明的法律后果由本人承担。

学位论文作者签名：

日期：2025 年 05 月 16 日

安徽工程大学学位论文版权使用授权书

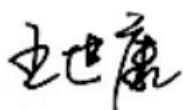
学位论文作者完全了解学校有关保留、使用学位论文的规定，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅或借阅。本人授权安徽工程大学可以将本学位论文的全部内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

保密 ☐，在 年解密后适用本版权书。

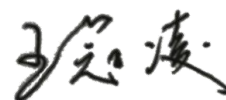
本学位论文属于

不保密 ☒。

学位论文作者签名：



指导教师签名：



日期： 2025 年 05 月 16 日

日期： 2025 年 05 月 16 日

基于 RGB-D SLAM 的无人机自主导航系统研究

摘 要

随着低空经济被纳入国家战略规划，无人机作为其核心载体，在物流配送、智慧城市、应急救援等领域展现出巨大潜力。然而，复杂动态环境下的自主导航技术存在诸多瓶颈，如低纹理场景定位漂移、动态干扰鲁棒性不足、实时路径规划效率低等，严重制约了无人机在低空经济中的规模化应用。本文聚焦于基于 RGB-D 相机的同步定位与建图技术(RGB-D Simultaneous Localization and Mapping, RGB-D SLAM)，设计了一种面向复杂动态环境的无人机自主导航系统，重点解决传统视觉 SLAM 在复杂场景、动态干扰及实时路径规划中的局限性，以提升系统的定位精度、鲁棒性及路径规划效率。

1、针对 ORB(Oriented FAST and Rotated BRIEF)特征提取在高低纹理环境中分布不均的问题，本文提出了一种基于四叉树的改进算法，通过计算局部密度标准差引入自适应分割策略，实现特征点在高纹理区域均匀分布，并尽量保留低纹理区域特征，从而提升位姿估计稳定性。同时，针对无人机资源和功耗限制，本文进一步提出了 YOLOv8_RSS 算法。该算法通过引入 RepGhost 轻量化模块和重构网络框架等技术，在显著提升检测精度的同时有效提高检测速度，具体表现为 mAP50 提高 5.29%、参数量降低 76.25%、计算量减少 9.47%，更适配嵌入式平台的资源约束并满足实时应用需求。在以上基础上，为增强动态环境下的定位精度，本文将 YOLOv8_RSS 与

ORB-SLAM2 的前端视觉里程计结合，通过动态特征剔除，并结合 LK 光流匹配和极线几何约束进一步剔除残留动态特征，从而显著增强视觉 SLAM 的鲁棒性和适应性。

2、针对无人机在复杂环境中的路径规划需求，本文提出了一种改进的生物粒子群优化路径规划 (Bio-inspired Particle Swarm Optimization with Q-Learning, BPSO-QL) 方法。该方法结合粒子群优化与粒子间生物相互作用机制，在传统粒子群优化算法的基础上引入随机角度扰动策略，提升算法在高维空间中的全局搜索能力，并结合 Q 学习(Q-Learning)构建局部路径重规划模块。在不同环境的三维地图中进行路径规划仿真实验，本文改进的生物粒子群优化算法相较于对比算法，达到全局最优解的迭代次数平均减少 21.5%，路径规划长度平均降低 4.23%。实现了无人机在复杂环境中更加灵活、高效的路径规划。

3、为验证本文所设计系统的有效性，在 Ubuntu 18.04 中部署机器人操作系统(Robot Operating System, ROS)，构建了基于嵌入式平台的无人机实验环境。通过 RGB-D 深度相机数据与算法模块，对改进的 ORB 特征提取、YOLOv8_RSS 动态特征剔除、稠密点云地图构建以及改进的生物粒子群优化路径规划方法进行了系统测试与性能评估，以验证系统在复杂动态环境下的稳定性和高效性。

关键词：低空经济，无人机，RGB-D，ORB-SLAM2，YOLOv8，粒子群优化，路径规划

Research on Autonomous Navigation system of UAV

Based on RGB-D SLAM

ABSTRACT

With the integration of the low-altitude economy into national strategic planning, unmanned aerial vehicles (UAVs), as its core carriers, demonstrate immense potential in areas such as logistics and delivery, smart city, and emergency rescue. However, autonomous navigation technology in complex and dynamic environments faces numerous challenges, including localization drift in low-texture scenes, insufficient robustness against dynamic interference, and low efficiency in real-time path planning, which severely restricts the large-scale application of UAVs in the low-altitude economy. This paper focuses on RGB-D SLAM technology and proposes an autonomous UAV navigation system designed for complex dynamic environments. The study aims to address the limitations of traditional visual SLAM in complex scenes, dynamic disturbances, and real-time path planning, ultimately improving the system's localization accuracy, robustness, and path planning efficiency. The main contributions and key research efforts are as follows:

1. To address the issue of uneven feature point distribution in high-texture or low-texture environments during ORB feature extraction, this paper proposes an improved algorithm based on a quadtree structure. By

calculating the standard deviation of local density and introducing an adaptive partitioning strategy, the algorithm achieves a balanced distribution of feature points in high-texture regions while fully preserving feature points in low-texture regions, thereby enhancing the stability of pose estimation. Additionally, to tackle the challenge faced by small object detection algorithms in unmanned aerial vehicles (UAVs)—where resource and power constraints make it difficult to balance detection speed and accuracy—this paper further proposes the YOLOv8_RSS algorithm. By incorporating the lightweight RepGhost module and reconstructing the network framework, this algorithm significantly improves detection accuracy while effectively increasing detection speed. Specifically, it achieves a 5.29% increase in mAP50, a 76.25% reduction in parameter count, and a 9.47% decrease in computational load, making it more suitable for the resource constraints of embedded platforms and meeting the demands of real-time applications. Building on this, to enhance localization accuracy in dynamic environments, this paper integrates the YOLOv8_RSS object detection algorithm with the front-end visual odometry of ORB-SLAM2. Through dynamic feature elimination, combined with LK optical flow matching and epipolar geometric constraints to further remove residual dynamic features, the robustness and adaptability of visual SLAM are significantly improved.

2. To address the path planning requirements of UAVs in complex

environments, this paper proposes an improved bio-inspired particle swarm optimization (BPSO) path planning method. This method integrates particle swarm optimization (PSO) with bio-interaction mechanisms among particles and introduces a stochastic angular perturbation strategy to enhance the global search capability in high-dimensional space. Additionally, a local path re-planning module is constructed using Q-learning. Path planning simulations in three-dimensional maps of various environments demonstrate that, compared to baseline algorithms, the proposed bio-inspired PSO achieves a 21.5% reduction in the average number of iterations required to reach the global optimal solution and a 4.23% decrease in average path length. These improvements enable more flexible and efficient UAV path planning in complex environments.

3. To validate the effectiveness of the proposed system, an experimental UAV environment was built on an embedded platform running the Ubuntu18.04 based ROS (Robot Operating System). The system was evaluated using RGB-D depth camera data and algorithmic modules, including improved ORB feature extraction, YOLOv8_RSS based dynamic feature removal, dense point cloud mapping, and the enhanced bio-inspired PSO path planning method. Experimental results demonstrate the system's stability and efficiency in complex dynamic environments.

KEY WORDS: Low-altitude Economy, UAV, RGB-D, ORB-SLAM2, YOLOv8, particle swarm optimization, path planning

目 录

摘 要	I
ABSTRACT.....	III
第 1 章 绪论	1
1.1 课题研究背景及意义.....	1
1.2 国内外研究现状.....	2
1.2.1 视觉 SLAM 研究现状	2
1.2.2 无人机目标检测研究现状.....	4
1.2.3 无人机路径规划研究现状.....	5
1.3 论文主要研究内容与章节安排.....	6
第 2 章 相关技术介绍	8
2.1 RGB-D 技术概述	8
2.1.1 RGB-D 相机	8
2.1.2 相机成像模型及坐标系变换.....	9
2.2 视觉 SLAM 理论基础	11
2.2.1 前端视觉里程计.....	12
2.2.2 后端优化.....	14
2.2.3 闭环检测.....	15
2.2.4 地图构建.....	16
2.3 YOLOv8 理论基础.....	16
2.4 经典路径规划算法.....	18
2.5 本章小结.....	19
第 3 章 基于语义和几何约束的视觉里程计优化	20
3.1 ORB-SLAM2 算法.....	20
3.1.1 ORB-SLAM2 系统框架.....	20
3.1.2 ORB-SLAM2 的优势与局限	22
3.1.3 改进的 ORB-SLAM2 系统框架	23
3.2 基于 ORB 的特征点提取	23

3.2.1 传统的四叉树 ORB 算法	23
3.2.2 改进的四叉树 ORB 算法	24
3.3 基于 YOLOv8 的目标检测	26
3.3.1 改进的 YOLOv8 框架	26
3.3.2 C2fRepGhost 轻量化模块.....	27
3.3.3 SimAM 注意力机制.....	29
3.3.4 Inner-SIoU Loss.....	30
3.3.5 多尺度特征融合结构.....	33
3.3.6 消融实验分析.....	34
3.3.7 可视化分析.....	36
3.3.8 对比实验分析.....	38
3.4 基于语义和几何约束算法实现动态特征点剔除.....	39
3.4.1 光流法匹配.....	39
3.4.2 对极几何约束.....	41
3.4.3 动态点剔除效果验证.....	43
3.4.4 点云地图构建.....	45
3.5 本章小结.....	46
第 4 章 基于生物粒子群优化的无人机路径规划	48
4.1 粒子群优化基础.....	48
4.1.1 粒子群优化.....	48
4.1.2 Vicsek 模型.....	50
4.1.3 Q-learning	50
4.2 改进的粒子群优化算法.....	51
4.3 算法性能测试实验.....	54
4.3.1 标准测试函数.....	54
4.3.2 BPSO 算法性能对比分析.....	54
4.4 路径规划实验对比分析.....	59
4.5 本章小结.....	63
第 5 章 无人机自主导航系统实验	64

5.1 自主导航系统设计.....	64
5.1.1 系统硬件设计.....	64
5.1.2 系统框架设计.....	67
5.2 相机标定与测试.....	68
5.2.1 D435i 深度相机标定.....	68
5.2.2 D435i 深度相机测试.....	69
5.3 系统部署与验证.....	70
5.3.1 目标检测算法部署.....	70
5.3.2 动态目标特征点剔除.....	71
5.3.3 八叉树地图转换.....	72
5.4 自主导航实验.....	73
5.5 本章小结.....	75
第 6 章 总结与展望	77
6.1 全文总结.....	77
6.2 研究展望.....	78
参考文献	79
攻读硕士学位期间主要研究成果	85
致谢	86

第1章 绪论

1.1 课题研究背景及意义

近年来，我国将低空经济纳入战略性新兴产业体系，并在 2024 年《政府工作报告》中明确提出“要积极打造低空经济等新增长引擎”，标志着低空经济已成为国家战略层面的核心发展方向。无人机作为低空经济中的重要载体，在机器人技术和计算机视觉等领域的不断突破下，已广泛应用于物流配送^[1]、农业植保^[2]、灾害救援^[3]、地理测绘^[4]等场景；又凭借其灵活性、低成本、高效性和高适应性，成为了这些行业中不可或缺的工具。然而，随着应用需求的不断增加，复杂动态环境下的自主导航能力仍是制约其大规模落地的核心瓶颈^[5]。传统依赖 GPS 的导航方式存在信号遮蔽、定位精度低等问题，而激光 SLAM 虽精度高但成本昂贵且难以适配轻量化设备。因此，视觉 SLAM^[6]作为一种低成本、高环境适应性的自主导航技术，成为无人机智能化升级的关键突破口。



图 1-1 四旋翼无人机

视觉 SLAM 技术通过单目/双目相机捕捉环境纹理信息，实现运动估计与地图构建的同步优化。与传统的激光 SLAM 等方法相比，视觉 SLAM 具有较低的硬件成本、更轻便的设备要求、更高的设备集成度等核心优势。这项技术起源于 20 世纪 80 年代，早期受限于处理器性能与算法复杂度，仅能在实验室环境中处理稀疏特征点。2007 年 PTAM 框架首次将跟踪与建图任务拆分为并行线程，奠定了实时视觉 SLAM 的算法基础；2015 年推出的 ORB-SLAM 系统通过引入 ORB 特征点与改进的回环检测机制，显著提升了复杂场景下的鲁棒性^[7]。近年来，随着多传感器融合技术的突破，视觉惯性里程计通过结合 IMU 数据有效解

决了快速运动导致的图像模糊问题，而 RGB-D 相机的普及使得稠密建图成为可能。相较于激光 SLAM，视觉方案不仅将硬件成本降低至千元级别，更能通过图像色彩与纹理识别语义信息，这对无人机执行电力巡检中的设备缺陷识别、智慧城市中的动态目标追踪等任务具有不可替代的价值。

然而，视觉 SLAM 技术仍面临诸多挑战。例如，如何提升低纹理环境中的稳定性、应对动态场景干扰、以及在计算资源受限的情况下实现实时处理等问题。随着计算机视觉与机器学习技术的不断进步，研究人员在这些领域持续提出创新性的方法，尤其是在 RGB-D SLAM 的研究中取得了显著突破。RGB-D 相机的引入彻底改变了视觉 SLAM 的技术路径。不同于双目相机依赖视差计算深度，RGB-D 相机通过主动红外结构光直接获取每个像素的深度值，使得系统在墙面、玻璃等弱纹理区域的稳定性得到了显著提升。以 Intel RealSense D435i 为例，其深度感知误差在 2 米范围内仅为 1%。通过这种方式，视觉 SLAM 不仅提高了无人机的定位和导航精度，还扩展了其在复杂环境下的应用能力。同时，将 RGB-D 相机与无人机路径规划算法相结合，可以构建一个更加高效、智能的自主导航系统。

综上所述，RGB-D 相机的引入为视觉 SLAM 技术提供了新的机遇，其在复杂环境中的稳定性、深度感知的精确性，以及与图像语义融合的能力，使得视觉 SLAM 在无人机导航中的应用前景更加广阔。然而，技术的进步仍需解决当前的一些局限性，如动态场景的干扰及计算资源的优化等问题，这些挑战也为未来的研究与应用提供了丰富的探索空间。

1.2 国内外研究现状

1.2.1 视觉 SLAM 研究现状

SLAM 技术的提出可追溯到 1986 年，Smith 与 Cheeseman 在开创性研究中首次将机器人定位与建图问题统一于概率论框架下，奠定了 SLAM 基础理论体系。2007 年，Davison 等人首次提出一种实时单目视觉 SLAM 系统 MonoSLAM^[8]，基于扩展卡尔曼滤波器实现稀疏特征点跟踪。同年，Klein 等人开发了 PTAM 系统^[9]，首次引入前后端多线程架构和关键帧概念，结合非线性优化提升效率。2014 年，Engel 等人提出 LSD-SLAM^[10]，利用像素梯度跟踪生

成半稠密深度图；同年，Forster 等设计 SVO^[11]，结合特征点法与直接法实现高速稀疏里程计。2015 年，Mur-Artal 等人提出 ORB-SLAM^[12]，在 PTAM 框架上增加闭环检测模块，形成多线程架构，显著提升鲁棒性；2017 年，该团队推出 ORB-SLAM2^[13]，支持双目和 RGB-D 相机。2021 年，Campos 等人提出 ORB-SLAM3 系统^[14]，这是一个支持多摄像头的视觉 SLAM 算法。

近年来，视觉 SLAM 的研究重点逐渐转向复杂动态环境下的鲁棒性与适应性。针对动态物体干扰问题，张等人^[15]提出了通过融合深度语义和不确定度估计的 SUI-SLAM 算法，在动态环境下的视觉 SLAM 任务中表现出较高的定位精度和鲁棒性，但其在实时性、动态物体类别限制、低动态场景提升以及复杂运动状态处理等方面仍存在不足。魏等人^[16]利用立体几何约束和颜色深度信息分割动态区域，显著提升室外场景的实时性与定位精度。崔等人^[17]进一步引入自适应极线约束与长期数据关联方法，通过 ORB-SLAM3 改进框架，在高动态场景下将轨迹误差降低，验证了语义与几何融合的有效性。

深度学习技术的快速发展为视觉 SLAM 开辟了新的研究方向。刘等人^[18]设计基于注意力机制的多任务蒸馏网络，实现在光照剧烈变化场景下提取稳定特征。Mo 等人^[19]提出了一种快速、准确的立体视觉 SLAM 系统 DSV-SLAM，通过扩展单目直接稀疏里程计到立体系统，采用尺度优化最小化光度误差，避免了传统特征检测和匹配，并使用 LiDAR 描述符进行回环检测，显著提高了计算效率和鲁棒性。郭等人^[20]利用 SqueezeNet 优化回环检测，在轻量化计算中实现了召回率的提升，并将建图误差大大降低。蔡等人^[21]针对弱纹理环境提出深度置信模型，结合动态鲁棒核函数优化光束平差法，解决了传统特征点法在低纹理场景下的失效问题。Yang 等人^[22]针对动态目标干扰，提出基于稳定特征匹配约束与新型几何约束的动态特征双重过滤机制，提升动态场景下 SLAM 算法的定位精度。这些方法虽然提高了算法的精度，但对于如何剔除动态物体检测与解决深度学习模型的计算负载限制，仍然存在诸多难题。

总的来说，视觉 SLAM 技术虽取得显著进展，但在动态复杂场景中仍面临实时性不足和鲁棒性弱等挑战，提升其在动态环境中的性能已成为当前研究重点领域。

1.2.2 无人机目标检测研究现状

随着网络技术的快速发展，无人机凭借其灵活性、可靠性和高度可控性，已经成为目标检测领域中一种高效的工具。尤其在实时监控和高效数据处理的需求下，无人机能够搭载高性能传感器和计算平台，快速捕捉并分析图像数据，精确识别地面上的目标。然而，无人机获取的图像往往面临背景复杂、目标尺寸较小以及外形模糊等挑战。传统目标检测方法在这些场景中表现出明显的局限性，特别是在小目标检测精度不足和对复杂背景的适应能力较弱时，常常出现漏检或误检的现象。随着深度学习的发展，基于卷积神经网络的目标检测方法近年来迅速发展起来^[23]，无人机的目标检测能力得到了显著提升，能够更好地应对复杂环境变化和动态目标的识别挑战^[24]。

然而，基于卷积神经网络的无人机的小目标检测方法存在种类繁多、尺寸较小、遮挡严重等问题。针对这些问题，国内外研究者们作出了许多尝试和努力。汪等人^[25]通过将密集连接网络与 YOLOv3 相结合，并引入 DIoU 损失函数作为新的坐标预测计算方法，增强特征传播能力并提高了光学遥感图像目标检测精度。蒋等人^[26]通过研究和分析在模型不同位置融入 scSE 注意力模块所带来的效果，最终将 scSE 注意力机制嵌入到 YOLOv4 的骨干网络和颈部网络中，从而提升了小目标检测的精度。何等人^[27]在 YOLOv5 网络结构中引入 Ghost 模块，并将加权双向特征金字塔网络结构 BiFPN 与注意力机制 CBAM 与算法进行融合，实现了小目标检测速度优异的同时，也保证了检测精度。Wang 等人^[28]提出了 SCP-YOLO 模型，通过引入 SPD 结构、Context-Aggregation 模块和 PConv 方法，显著提升了无人机遥感图像中小麦穗的检测精度，但该模型在检测速度上略有下降，且在处理极高密度目标时仍存在一定的漏检和误检问题。Zhao 等人^[29]在 YOLOv7 算法上进行改进，增加了小目标检测头并添加了无参数的 SimAM 注意力模块，以牺牲计算量为代价增强了对海面小目标的检测性能。Liu 等^[30]针对无人机电力线路检查中的绝缘子等目标，以 YOLOv3-Tiny 为基础，通过引入特征融合结构和 SPP 模块扩展颈部网络，并新增主干网络的输出层，将多尺度特征与 SPP 模块相结合，构建出轻量化的 MTI-YOLO，从而提高复杂场景下绝缘子检测的精度。孙等人^[31]针对无人机在自主飞行且进行目标检测任务时，存在的实时性和准确性需求，对 YOLOv4 作出改进，提出了采用 MobilnetV3 网

络作为主干特征提取网络，并采用 CIoU 作为损失函数计算方法，实验结果表明，改进后的模型损失较低精度但可以有效实现轻量化，且检测速度得到极大提高。Dai^[32]提出了 YOLOD 模型，通过在 YOLOX 的检测头引入高斯先验，改进边界框回归和不确定性估计，提高了绝缘子故障检测的准确性和鲁棒性；但需较高计算资源，对小目标检测存在局限，未针对多种复杂环境进行全面测试。韩等人^[33]通过改进 YOLOv5s 提出了无人机密集小目标检测算法 LSA_YOLO，主要构造多尺度特征提取模块 LM-fem 和提出了新的混合域注意力模块 S-ECA，并设计自适应权重动态融合结构 AFF，取得了小目标平均检测精度的显著提升，不足处在于检测图像输入分辨率较高，可能会对无人机设备造成资源负担。

总体来说，虽然当前的目标检测技术在提高检测精度方面取得了显著进展，但在面对不同应用场景和硬件约束时，如何进一步优化算法，提升算法的适应性和效率仍然是未来研究的关键方向。

1.2.3 无人机路径规划研究现状

无人机凭借其灵活机动、低成本及高环境适应性等优势，已广泛应用于军事侦察、农业植保、灾害救援、物流运输等领域。随着任务场景的复杂化，路径规划技术成为提升无人机自主性与任务效能的核心挑战，其目标是在动态、未知环境中快速生成安全、高效且满足多约束条件的飞行路径。然而，传统算法如 A*^[34]、Dijkstra^[35]和快速探索随机树 RRT^[36]通过特定的策略获得可行的路径，难以应对实时突发障碍、多机协同冲突、复杂地形风场干扰及能量效率优化等问题。尤其在三维动态环境中，路径规划需兼顾避障实时性、多目标权衡、多机协同通信延迟等复杂需求。

针对这些问题，近年来国内外学者从智能算法优化、动态环境适应性、多机协同规划等方向展开深入研究，试图寻找合适的解决方案。王等人^[37]提出 LCGARO 算法，结合了 Levy 飞行、自适应柯西变异和精英群遗传策略，解决人工兔优化算法收敛慢、易陷入局部最优的问题。但实验场景仅验证了静态地形，未涉及动态障碍或突发威胁的适应性。黄等人^[38]通过改进黏菌算法 GSMA，并引入混沌映射和非线性惯性权重，增强全局搜索能力并提升收敛速度。但能量消耗模型未考虑复杂风场等环境干扰，容易受限于实际应用。尹等人^[39]基于

Q-learning 构建多无人机协同路径规划框架, 通过经验矩阵共享和时间协同策略解决碰撞避免问题。但神经网络替代 Q-table 时容易存在训练时间较长等问题。陈等人^[40]提出改进多目标灰狼算法, 结合交叉算子与概率路图算法, 优化信息延迟与飞行距离双目标。然而未考虑动态障碍物的实时避障需求, 仅适用于部分静态场景。Niu 等人^[41]结合改进 IFDS 与人工神经网络, 通过离线训练与实时参数调整实现动态避障。Rosa 等人^[42]基于 Q-learning 平衡路径长度、安全性与能耗, 支持优先级动态调整。但该方法仅适用于静态环境, 未扩展至动态障碍或协同规划场景。Hu 等人^[43]提出 MAHACO 算法, 通过 Said-Ball 曲线生成平滑路径, 结合自适应策略降低多机协同能耗。但未充分考虑通信延迟对协同效率的影响。

尽管当前研究在算法优化、多机协同与动态适应性方面取得显著进展, 但仍存在动态环境泛化能力不足、多目标权衡不充分、硬件与算法协同设计欠缺等问题。另外, 无人机的飞行环境存在复杂性, 在面对路径规划时通常需考虑高维问题。即使算法优化在某些特定场景中取得良好效果, 但是当维数增加时, 其性能可能会下降。

总而言之, 在无人机路径规划领域, 目前的无人机仍然存在诸多问题, 难以计算出全局最优的规划路径, 和难以获取更短收敛路径和更快规划速度。

1.3 论文主要研究内容与章节安排

本文主要研究如何构建一种基于 RGB-D SLAM 的无人机自主导航系统。针对传统视觉 SLAM 系统在复杂动态场景中定位精度和鲁棒性不足的问题, 导致无人机在地图构建和自主导航方面性能受限, 本研究充分利用 RGB-D 相机提供的环境深度信息, 并融合基于深度学习的目标检测模块, 实现高精度的语义地图构建, 同时优化路径规划算法以提升导航效率。最终, 将上述研究成果集成至无人机自主导航系统, 并在实际无人机平台上部署完成功能验证, 证实了系统在复杂动态环境下的优越性能。本文共有六个章节, 各个章节安排及其内容如下:

第一章为绪论。本章阐述了无人机自主导航的研究背景及其意义, 剖析了当前领域内存在的主要问题与挑战, 从而引出本文的研究目标和核心内容。分

别对视觉 SLAM、无人机目标检测以及无人机路径规划的研究现状进行了综述，并对本文的主要研究内容和章节安排进行了介绍。

第二章为相关技术介绍。本章概述了 RGB-D 技术和视觉 SLAM 算法的理论基础，包括前端视觉里程计、后端优化、闭环检测与地图构建，并分析 YOLOv8 算法的整体框架和简要介绍了部分经典路径规划算法。

第三章为基于语义和几何约束的视觉里程计优化。本章概述了 ORB-SLAM2 算法的基本原理和实现过程，并引出当前算法所存在的不足，进而提出了基于四叉树的 ORB 特征提取、YOLOv8_RSS 动态目标检测与语义剔除、LK 光流匹配及极线约束几何滤除的改进方法，通过实验分析算法的可行性，提高视觉 SLAM 在复杂动态环境下的系统稳定性。

第四章为基于生物粒子群优化的无人机路径规划。针对复杂环境下的无人机路径规划问题，提出了一种用于全局路径规划的生物粒子群优化算法，详细阐述了算法的改进思路 and 实现流程，并通过算法性能测试实验证明算法的鲁棒性和可行性，进而稳定实现无人机的路径规划。

第五章为无人机自主导航系统实验。详细阐述所设计无人机自主导航系统的整体架构、功能模块及其实现细节，涵盖无人机硬件平台的选择、传感器配置方案以及控制系统设计的思路，并进一步设计实验以验证本文所提出算法的实用性。

第六章为总结与展望。对本文的研究内容及取得的成果进行概述，分析研究过程中存在的不足，并为后续工作的推进提出整体研究策略与发展方向。

第2章 相关技术介绍

本章围绕无人机自主导航系统的核心技术展开，重点介绍 RGB-D 视觉感知、视觉 SLAM 框架、目标检测算法及路径规划方法的理论基础。这些技术构成后续章节算法改进与系统设计的理论支撑。

2.1 RGB-D 技术概述

RGB-D 技术通过同时获取场景的 RGB 图像和深度图像来增强计算机视觉系统的感知能力。RGB 图像提供了场景的颜色信息，而深度图像记录了每个像素点与相机之间的距离。通过结合这两种信息，RGB-D 技术能够让计算机从颜色与空间两个维度来理解环境。这种技术能够有效解决传统图像处理方法中的一些挑战，如遮挡问题、视角变化和光照干扰。

与仅依赖 RGB 图像的传统方法相比，深度图像提供的空间信息显著增强了场景理解的能力。例如，深度图像帮助系统准确捕捉物体的形状、位置和尺寸，尤其是在复杂环境中，如动态物体识别、机器人导航和手势识别等任务中，RGB-D 技术展现出了巨大的优势。

随着 RGB-D 技术的发展，它在多个领域的应用也不断扩展，特别是在机器人、虚拟现实和自动驾驶等领域。

2.1.1 RGB-D 相机

RGB-D 相机是一种能够同时捕捉彩色图像和深度信息的传感器^[44]，广泛应用于无人机自主导航系统，图 2-1 为几个常见的 RGB-D 相机。与传统的单目摄像机或 RGB 相机相比，RGB-D 相机通过提供环境的三维信息，使得无人机能够在复杂环境中实现更高效的路径规划、障碍物避让和精确定位。



图 2-1 常见的 RGB-D 相机

传统的 RGB 相机仅提供二维的彩色图像数据，这使得无人机在环境感知中主要依赖视觉特征来推断深度和空间关系。然而，RGB 相机对于深度感知的能力有限，尤其在低纹理或结构简单的环境中，可能无法有效识别障碍物或估算距离。与之相比，RGB-D 相机不仅提供彩色图像，还能实时捕捉每个像素的深度信息，使得无人机能够直接获取三维空间的结构数据，从而大幅提升环境感知的精度和可靠性。

在无人机自主导航中，RGB-D 相机的优势尤为显著。通过 RGB-D 数据，无人机能够实时构建环境的三维模型，识别周围的障碍物并进行精确的距离估算。此外，RGB-D 相机能够支持同时定位与建图，通过不断采集深度信息，帮助无人机在未知环境中实时更新地图并进行准确定位。这使得无人机在复杂和动态的环境中能够灵活避障，并根据环境变化动态调整飞行路径。

因此，RGB-D 相机作为环境感知的核心传感器，极大地提高了无人机在复杂环境中的自主导航能力。与传统摄像设备相比，RGB-D 相机通过融合彩色图像与深度数据，提供了更加全面和精确的感知信息，推动了无人机技术向更高的自主性和安全性发展。

2.1.2 相机成像模型及坐标系变换

在 SLAM 系统中，RGB-D 相机结合了 RGB 图像和深度图的优势，可以提供丰富的视觉信息和精确的深度信息，用于构建实时的地图并进行定位。本小节将简要介绍针孔相机模型和坐标系变换原理。

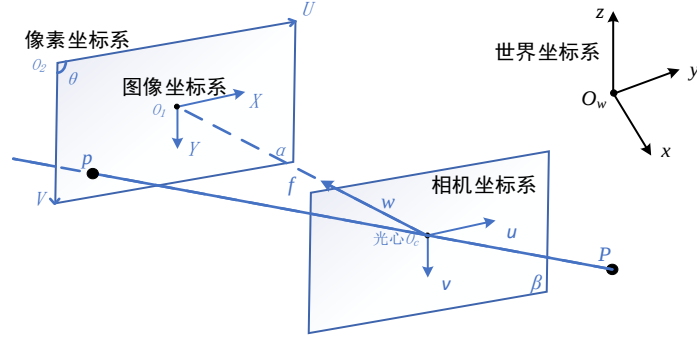


图 2-2 针孔相机成像系统的坐标系

针孔相机的成像系统的坐标系如图 2-2 所示，点 P 存在于三维空间，其穿过焦平面 β 上的相机光心 O_c 并在相机成像平面 α 投影于点 p 。

在相机坐标系和图像坐标系中，对应点的坐标可根据三角形相似原理，满足下列关系：

$$\begin{cases} X = f \frac{u}{w} \\ Y = f \frac{v}{w} \end{cases} \quad (2-1)$$

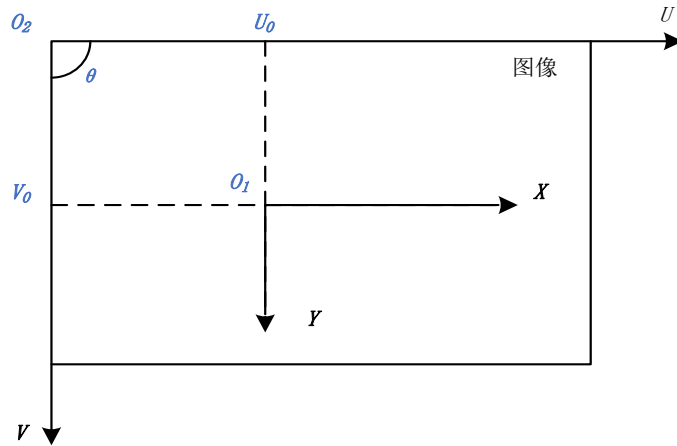


图 2-3 图像坐标系与像素坐标系的坐标转换关系

如图 2-3，两者的坐标转换关系通过坐标可满足下列关系：

$$\begin{cases} U = \frac{X}{aX} + U_0 \\ V = \frac{Y}{aY} + V_0 \end{cases} \quad (2-2)$$

联立公式(2-1)与公式(2-2)，并简化为齐次方程，可得：

$$w \begin{bmatrix} U \\ V \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & U_0 & 0 \\ 0 & f_y & V_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \\ 1 \end{bmatrix} \quad (2-3)$$

上式中， $f_x = \frac{f}{aX}$ 、 $f_y = \frac{f}{aY}$ ，均为有效矩阵，单位为像素。

世界坐标系到相机坐标系的转换是一种刚体变换，两者的三维坐标通过旋转和平移可以实现相互转换，它们的坐标变换关系为：

$$\begin{bmatrix} u \\ v \\ w \\ 1 \end{bmatrix} = \begin{bmatrix} R & T \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2-4)$$

上式中， R 代表旋转矩阵， T 代表平移向量。

联立公式(2-3)与公式(2-4)，可得世界坐标系到像素坐标系的坐标变换关系：

$${}_w \begin{bmatrix} U \\ V \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & U_0 & 0 \\ 0 & f_y & V_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = KM \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2-5)$$

上式中， K 是相机的固有参数，称作内部参数矩阵； M 会随着相机移动而改变用来描述当前相机所处的位姿，称作外部参数矩阵； KM 表示相机的透视投影矩阵。

通过这些坐标系及其之间的变换，可以有效地描述和处理相机成像过程中的几何关系，进而用于物体检测、定位、三维重建等计算机视觉任务。

2.2 视觉 SLAM 理论基础

视觉 SLAM(Simultaneous Localization and Mapping)是计算机视觉和机器人技术中的一个核心技术，即通过摄像头获取的图像信息同时完成环境的建图和定位任务。视觉 SLAM 的目标是利用摄像头提供的图像序列实时估计机器人或相机的轨迹，并同时构建出环境的稀疏或密集地图。其理论基础可以分为几个主要模块：前端视觉里程计、后端优化、闭环检测和地图构建模块^[45]。

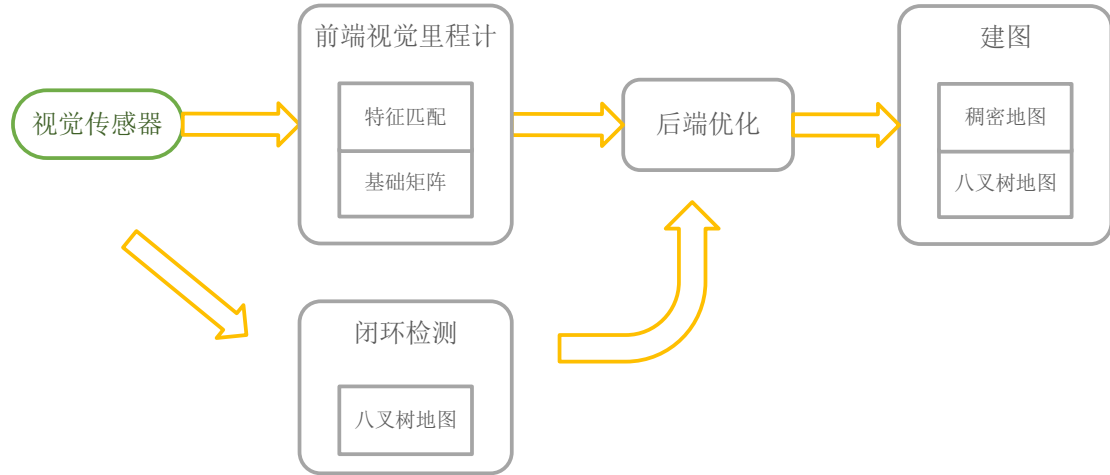


图 2-4 经典 VSLAM 框架

2.2.1 前端视觉里程计

前端视觉里程计(Visual Odometry, VO)负责从连续的图像帧中提取有用的特征并进行匹配,从而估计相机或机器人的位置和姿态。具体过程包括特征提取与匹配、位姿估计、跟踪与优化:

(1) 特征提取与匹配: 首先,前端系统需要从每一帧图像中提取稳定且具有区分度的特征点,常见的描述符包括 Harris^[46]、SURF^[47]、SIFT^[48]、FAST^[49]和 ORB^[50]。然后进行相邻图像帧之间的特征匹配,匹配过程中常用的方法有光流法、直接法、基于描述符的匹配等。其中,ORB 结合了 FAST 角点特征检测和 BRIEF 描述符,并且对其进行了改进,特别是对于旋转不变性和尺度不变性等方面进行了优化。

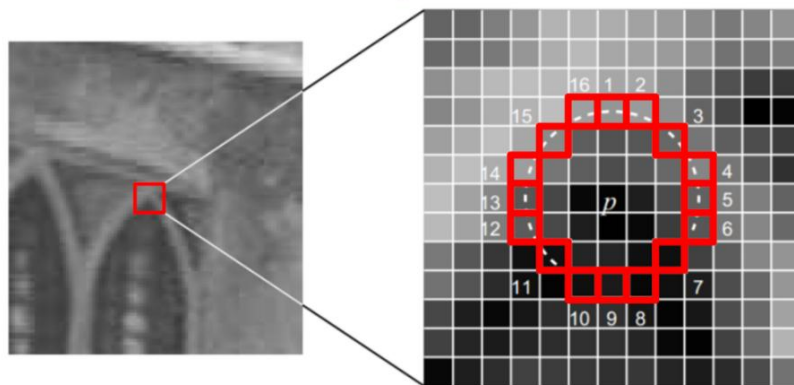


图 2-5 FAST 特征点

FAST 角点提取如图 2-5 FAST 特征点所示。算法检测过程如下: 在图像中的每个像素上进行检测, 首先选择一个像素点 p 作为候选角点。在每个候选点 p

周围，FAST 算法考虑一个包含 16 个像素的通常为半径 3 个像素的圆形邻域，这些 16 个像素按照固定顺序编号，围绕候选点 p 排布。若某个候选像素点 p 的亮度值与其邻域内足够数量的像素的亮度差异较大，则该点被标记为角点。然而，原始的 FAST 算法对图像的旋转不具备鲁棒性，因此 ORB 对 FAST 算法进行了改进，增强了其旋转不变性。具体来说，ORB 为每个检测到的角点分配了一个主方向，从而确保旋转操作不会影响角点的匹配。

特征匹配是指在两帧或多帧图像中识别出相互对应的特征点，这一过程对后续的运动估计具有关键意义。特征匹配通常依赖于特征点的描述符，这些描述符通过编码特征点邻域像素的信息来表征其外观特性。常用的特征匹配方法包括：

- 暴力匹配(Brute-Force Matcher)^[51]是一种简单直接的特征匹配方法，其核心思想是对一帧图像中的每个特征描述子，逐一计算它与另一帧图像中所有特征描述子的距离，如欧几里得距离或余弦距离，然后选择距离最近的描述子作为匹配对，从而建立两幅图像间的特征点对应关系。

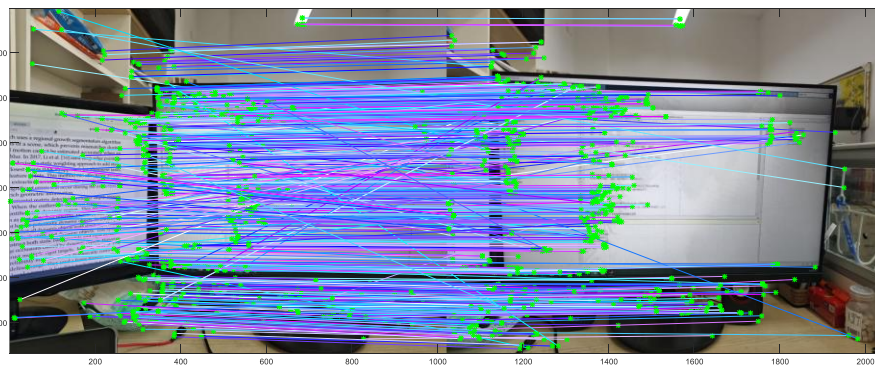


图 2-6 暴力匹配

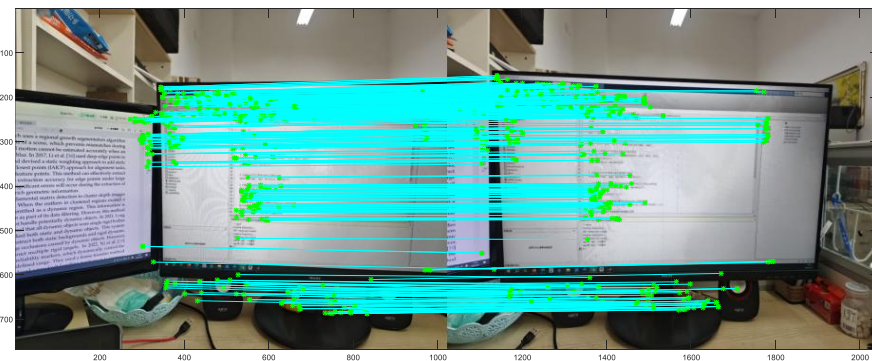


图 2-7 暴力匹配结合比率测试

- HNSW(Hierarchical Navigable Small World)^[52]是一种基于图结构的特征匹

配算法，通过构建多层小世界网络来组织特征描述子，利用图的导航特性高效搜索最近邻，在高维数据中兼顾了匹配速度和精度，适用于现代深度学习特征的快速匹配任务。

•FLANN 匹配(Fast Library for Approximate Nearest Neighbors)^[53]是一种高效的搜索方法，通过构建数据结构，快速找到一帧图像特征描述子在另一帧图像中的近似匹配对，避免逐一比较所有描述子，从而在大规模数据集中显著提升匹配速度。

(2) 位姿估计：通过匹配得到的特征点对，计算相机的相对运动。常见的方法包括基于基础矩阵和单应性矩阵的两视图几何估计，或者使用光束法平差^[54](Bundle Adjustment)来优化位姿。

(3) 跟踪与优化：为了提高鲁棒性，前端通常采用图优化^[55]或滤波方法，如扩展卡尔曼滤波^[56]或者图优化来不断修正相机的位姿。

2.2.2 后端优化

后端优化(Backend Optimization)是视觉 SLAM 系统中的一个关键部分，旨在对前端估计的相机轨迹和地图进行全局优化。前端系统虽然可以提供局部的位姿估计，但由于累积误差如漂移等，因此需要后端系统对轨迹进行优化，并提升定位和建图的精度。后端优化的常见方法有图优化、光束法平差等。

(1) 图优化(Graph Optimization)将 SLAM 问题建模为一个图，其中每个节点表示相机位姿或地图特征点，而边则表示位姿之间的约束，可以来自于前端匹配或者传感器观测，通过建模相机的位姿和地图点之间的约束关系来实现优化，图优化如图 2-8。常见的图优化方法有 g2o(General Graph Optimization)和 iSAM(incremental Sparse Pose Adjustment)。

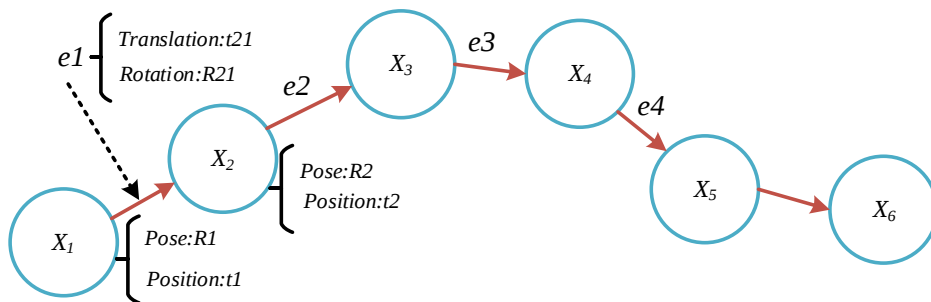


图 2-8 图优化示意图

图优化的目标是最小化整个图中的误差。具体来说，对于每一条边，它有一个与实际测量值的差异，图优化算法的任务就是调整图中的节点，使得所有边的残差最小化，从而优化整个系统。

(2) 光束法平差(Bundle Adjustment, BA)是一种非线性优化方法，用于同时优化相机的位姿和地图中特征点的位置，通过最小化重投影误差来改进地图的精度。通常在 SLAM 的后端阶段使用，可以有效消除由于相机运动产生的误差。光束法平差如图 2-9 所示。

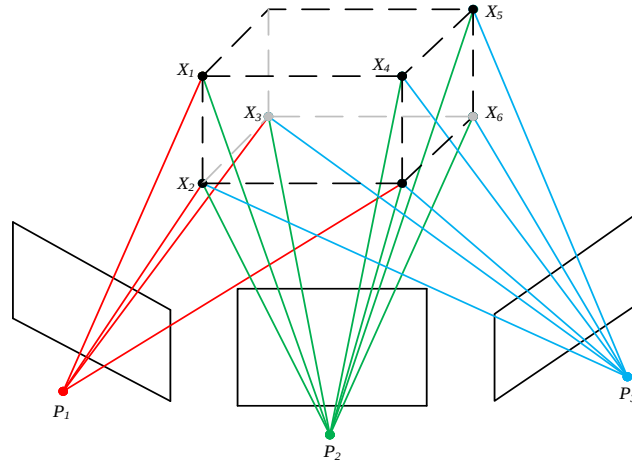


图 2-9 光束法平差示意图

光束法平差的优化目标是最小化从 3D 空间中的地图点投影到 2D 图像中的重投影误差。具体来说，给定相机的位姿和地图点的三维坐标，BA 的目标是调整位姿和地图点的位置，使得通过相机投影模型计算的图像坐标与实际观测到的图像坐标之间的差异最小。优化函数方程如公式(2-6)， u_{ij} 为像点坐标， X_i 是三维点坐标， P_j 是相机投影矩阵：

$$\mathcal{L} = \min \sum_{i=1}^n \sum_{j=1}^m \left(u_{ij} - \pi(P_j, X_i) \right)^2 \quad (2-6)$$

2.2.3 闭环检测

闭环检测(Loop Closure Detection)同样作为 SLAM 系统中的一个关键模块，其目标是在相机走回以前经过的地方时，识别出闭环，即判断当前帧是否与之前的某个位置相同或相似。这可以有效减少因累积误差而导致的漂移问题。闭环检测的主要步骤包括：

(1) 关键帧选择：选取一系列具有代表性的关键帧，这些关键帧通常是地图中具有足够信息的图像帧，用于进行闭环检测。

(2) 特征匹配与相似性度量：通过提取图像特征，使用合适的相似性度量方法，如直方图匹配、描述符匹配等，来检测当前帧和之前关键帧的相似度。

(3) 图优化与后端调整：一旦检测到闭环，后端的图优化算法会重新调整图中的位姿和地图，以消除漂移并改进全局一致性。

2.2.4 地图构建

地图构建(Mapping)是 SLAM 系统的最终目标之一，即根据相机采集到的图像序列构建出环境的地图。根据不同的应用需求，SLAM 的地图可以分为稀疏地图和稠密地图两种形式：

(1) 稀疏地图：稀疏地图是由少量具有代表性的特征点组成的地图。特征点通常是图像中的角点或其他容易识别的兴趣点。这种地图计算量小，但无法提供丰富的细节信息。

(2) 稠密地图：该地图通常使用深度图或者点云来表示，包含了更多的环境细节。对于密集 SLAM，除了位姿估计外，还需要估算每个像素的深度信息，常见方法有立体视觉、结构光、激光雷达等。

在地图构建过程中，通常需要根据相机的运动来更新地图，并确保地图的一致性。通过地图的优化、闭环检测等手段，可以有效去除噪声并提高地图的精度。

2.3 YOLOv8 理论基础

基于深度学习的目标检测技术以其结构灵活、检测精度高、检测速度快等特点成为了研究热点，分为基于候选区域的双阶段网络和基于回归的单阶段网络两类^[57]。双阶段对象检测算法包括 R-CNN^[58]、Fast R-CNN^[59]、Faster R-CNN^[60]、Mask R-CNN^[61]等，该类网络首先生成一系列候选区域框，然后再对目标进行分类和回归。单阶段目标检测算法包括 YOLO 系列算法、SSD^[62]算法等，此类网络不需要生成候选区域框，直接在图像中进行目标特征的提取，同时获得目标的类别和位置信息。与双阶段网络相比，单阶段网络在检测时间上

占有优势，但在检测精度上略显不足。在单阶段目标检测算法中，YOLO 系列算法是目前为止发展最快、最好的算法之一，尤其是原 YOLOv5^[63]团队即 Ultralytics 公司在 2023 年发布的 YOLOv8 算法有着极高的准确率。

YOLOv8 是 Ultralytics 公司在 2023 年初发布的最新 YOLO 系列目标检测模型，相比较上一代 YOLOv5，YOLOv8 在精度和检测速度上都有着更优的提升。在小目标检测方面，YOLOv8 解决了上一代模型算法存在的不足，获得了更优的表现。同时作为 YOLO 系列的最新算法模型，YOLOv8 继承了这一系列传统网络结构，一般分为四个部分：输入层(Input)、骨干网络(Backbone)、颈部结构(Neck)、检测头(Head)。模型结构如图 2-10 所示。

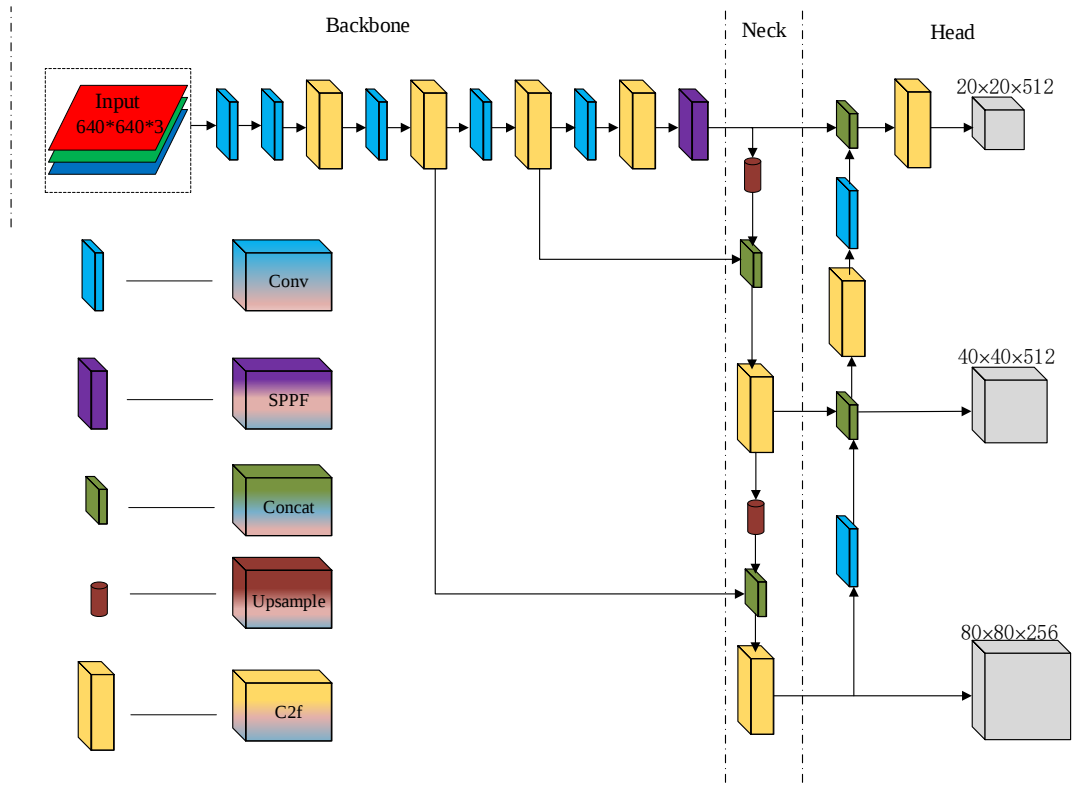


图 2-10 YOLOv8 网络结构

YOLOv8 的输入层采用 Mosaic 进行数据增强，但与 YOLOv5 区别在于引入了 YOLOx 中提出的设计方案，即在最后 10 个 epoch 会关闭 Mosaic；同时抛弃原有的对象预测方法，采用了 anchor-free 机制直接预测对象的中心，进而减少了锚框预测的数量，从而加速了非极大值抑制。YOLOv8 的主干部分中 C2f 模块借鉴了 YOLOv7^[64]中的 ELAN 思想，并将其与 C3 模块融合，形成了 C2f 模块，使 YOLOv8 在维持轻量化的同时，能够获得更丰富的梯度流信息。在骨干网络的末端，仍然采用的是 SPPF 模块即空间金字塔池化层。在颈部结构，

YOLOv8 仍然使用 PAN-FPN 作为特征融合方式, 实验结果表明这种方式更适用于融合和利用不同尺度下特征层信息。在颈部结构的最后一部分, YOLOv8 借鉴了 YOLOx 中解耦头部的方法, 通过置信度和回归框相结合, 以达到更高的精确度。检测头作为输出部分, 同样采用了主流的解耦头结构, 并分为三个检测头, 分别用于检测 20×20 、 40×40 、 80×80 尺寸大小的特征图, 从而利用不同尺度预测目标物体的类别和位置。

2.4 经典路径规划算法

路径规划算法是自主导航、自动驾驶、游戏开发等领域中非常核心的技术, 简单来说, 它的目标是为一个移动对象找到从起点到终点的最佳路径, 同时避开障碍物或者满足特定的约束条件。本文介绍部分常见的路径规划算法。

(1) 迪杰斯特拉算法

迪杰斯特拉算法(Dijkstra)是一种基于图论的单源最短路径搜索方法, 适用于权值非负的图。该算法以起点为初始点, 逐步选取当前距离最小的未访问节点, 并更新其邻接节点的最短路径距离, 直到路径扩展至目标节点或所有节点均被遍历为止。在无人机路径规划中, 算法通过贪心策略搜索带权图的最短路径, 核心是将环境建模为位置和移动方向及成本的网格, 从起点逐步扩展最低成本路径。

(2) 人工势场法^[65]

人工势场法(Artificial Potential Field Method, APF)通过虚拟力场引导路径搜索, 目标点产生吸引力, 障碍物产生排斥力, 合力方向决定移动方向。算法实时计算合力向量, 调整移动步长, 循环迭代直至抵达目标或陷入局部极小。

(3) 蚁群算法^[66]

蚁群算法(Ant Colony Optimization, ACO)模拟蚂蚁群体信息素正反馈机制。蚂蚁在可行路径上释放信息素, 后续蚂蚁根据信息素浓度概率选择路径, 浓度高的路径被强化。算法通过迭代更新信息素和路径选择概率, 最终收敛到最优解。

(4) 粒子群优化算法^[67]

粒子群优化算法(Particle Swarm Optimization, PSO)模拟鸟群协作搜索, 粒子

根据个体历史最优和群体历史最优调整运动速度与方向。算法初始化粒子群后，迭代更新粒子位置，通过适应度函数评估路径质量，直至满足终止条件。

对于不同的应用场景，每种路径规划算法可表现出不同优缺点。本文实验环境具有一定的复杂度，因此需要根据求解要求、适用场景、计算效率等方面选择合适的路径规划算法。根据常见路径规划算法特性如表 2-1 所示。

表 2-1 路径规划算法对比

算法	计算效率	避障能力	全局最优性	实时性	适用场景
Dijkstra	中等	高	保证	低	静态、无负权图
APF	高	中	不保证	高	动态、局部路径规划
ACO	低	高	近似最优	低	复杂、多目标优化
PSO	中等	高	近似最优	中等	动态、复杂环境

从上表可以得出，Dijkstra 适合静态环境，但动态适应性不足；APF 实时性强，但全局规划能力较弱；ACO 优化能力强，但计算效率偏低；PSO 则在多方面表现均衡，能够通过粒子群的协作搜索快速适应复杂环境，同时保证较高的避障能力和路径质量。对比几种常用路径规划算法，PSO 是最适合本文的路径规划方法。为得到更高效的路径规划算法，本文进行相应改进来进一步提升算法性能与求解能力。

2.5 本章小结

本章对 RGB-D 技术和视觉 SLAM 系统进行了系统性的阐述。首先，介绍了 RGB-D 相机的坐标变换原理，明确了深度图与 RGB 图之间的空间关系。接着，深入探讨了视觉 SLAM 系统的核心组成部分，重点分析了前端视觉里程计、后端优化算法、闭环检测机制和环境地图构建方法四个核心模块的主要功能。随后，详细介绍了 YOLOv8 目标检测框架，阐述了其各个模块的协同工作方式，使得系统能够快速而准确地进行目标检测。最后，简要介绍了粒子群优化算法、人工势场法等路径规划算法。以上理论为后续优化视觉 SLAM 中的视觉里程计部分和无人机的路径规划部分提供了坚实的理论基础。

第3章 基于语义和几何约束的视觉里程计优化

ORB-SLAM2 是一种经典的特征点法 SLAM 系统，在静态环境中表现出较高的精度和鲁棒性。然而，在复杂动态环境中，由于依赖 ORB 特征进行跟踪和位姿估计，该系统容易受到动态目标的干扰，导致误匹配、漂移甚至定位失败。其主要局限性包括：动态目标的干扰、特征点匹配的局限性等。针对这些问题，本文提出一种结合语义信息与几何约束的视觉里程计优化方法。通过语义信息和几何约束的有效结合，减少了动态目标对 SLAM 系统的干扰，提高了特征点匹配的可靠性，从而提升了 ORB-SLAM2 在复杂动态环境中的定位精度和鲁棒性。

3.1 ORB-SLAM2 算法

ORB-SLAM2(Oriented FAST and Rotated BRIEF SLAM2)是一种先进的视觉 SLAM 系统，作为 ORB-SLAM 的扩展版本，它在多传感器支持、精度、鲁棒性等方面表现出色，成为当前视觉定位与地图构建领域的领先算法之一。ORB-SLAM2 通过支持单目、双目和 RGB-D 相机，展示了其强大的多传感器处理能力，能够在各种环境中实现高精度的定位和地图构建。

3.1.1 ORB-SLAM2 系统框架

ORB-SLAM2 的系统架构主要由三个独立但密切协作的线程组成：跟踪线程(Tracking Thread)、局部地图构建线程(Local Mapping Thread)和回环检测线程(Loop Closing Thread)。这三个线程分别负责系统的关键任务，确保相机定位、地图构建以及回环检测等任务能够并行高效地进行。

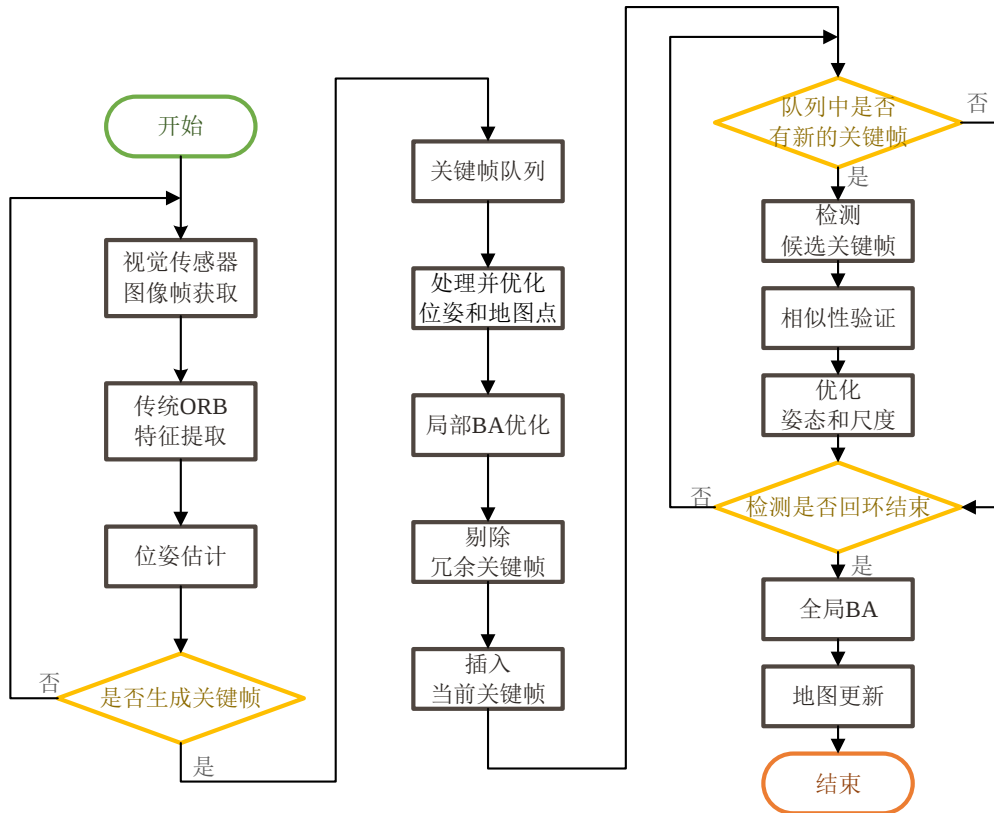


图 3-1 ORB-SLAM2 系统框架

(1) 跟踪线程

跟踪线程的任务是从图像帧中提取 ORB 特征，执行特征匹配，并估算相机的位姿。其核心职责包括：提取特征点并与前一帧或关键帧进行匹配，从而计算出相机的运动轨迹。通过匹配特征点进行局部地图更新，实时估计相机的轨迹。在系统初始化时进行地图的初步构建，并在丢失跟踪时进行重定位。

跟踪线程是整个 SLAM 系统的核心，负责相机定位和局部地图的维护。

(2) 局部地图构建线程

局部地图构建线程负责创建和优化环境中的地图。其主要任务包括：管理关键帧，决定哪些帧成为新关键帧并添加到地图中。生成新的地图点，并通过三角测量优化地图点位置。进行局部优化，使用 BA 方法减少地图误差。

该线程确保 ORB-SLAM2 能够逐步构建和精细化环境地图。

(3) 回环检测线程

回环检测线程通过视觉词袋(Bag of Words, BoW)模型检测回环，即相机返回到先前的地点。其关键任务包括：利用 BoW 模型高效检测当前帧是否与历史帧相似，识别回环。一旦检测到回环，触发回环闭合，使用全局优化修正地图和

轨迹中的误差。

回环检测线程保证了系统长期运行中的地图一致性，消除了累积误差。

3.1.2 ORB-SLAM2 的优势与局限

ORB-SLAM2 的系统架构是其成功的重要因素之一，其采用了多线程设计，包括跟踪线程、局部地图构建线程以及闭环检测线程。这一独特的线程分工使 ORB-SLAM2 能够实时估算相机位姿，动态生成并更新局部地图，同时识别和处理回环，从而有效消除长时间运行可能产生的累积误差。另外，在位姿估计和地图优化方面，ORB-SLAM2 不仅依赖局部优化，还采用全局和局部捆绑调整 BA 来提高地图的精度。通过 BA 优化，ORB-SLAM2 能够在全局范围内调整相机的轨迹与地图点位置，最大程度地减小误差，提升系统的精度和一致性。

尽管 ORB-SLAM2 在多数情况下表现出色，但在以下几个方面存在一定的局限性，特别是在图像纹理信息丰富环境、动态场景和计算资源方面。

（1）图像纹理信息丰富环境下的鲁棒性较差

在纹理信息丰富的环境中，ORB 特征依然依赖图像中的局部信息进行匹配，这可能导致在纹理丰富但几何结构重复的场景中出现匹配困难，造成特征点匹配和位置估计的精度下降。

（2）动态场景的适应性差

在动态场景中，ORB-SLAM2 易将动态目标错误识别为环境的一部分，从而引起地图构建过程中的偏差。例如，当相机经过移动物体时，这些物体的特征可能被错误地融入到地图中，影响系统的定位精度。

（3）计算资源消耗较大

ORB-SLAM2 在多个线程工作中都需要大量计算资源，尤其是在跟踪线程中，实时提取特征点、进行匹配以及位姿估计的计算需求非常高。这种资源消耗在处理高分辨率图像或快速移动的场景时尤为明显。

为了解决上述一系列问题，本文引入了一种改进的四叉树 ORB 算法，通过优化特征点提取策略，确保特征点分布的均匀性，从而更高效地应对复杂且纹理重复的场景，提高系统在此类环境下的鲁棒性。此外，本文融合了基于改进 YOLOv8 的动态目标检测算法，该算法采用轻量化模块设计与多尺度特征融合

结构，能够精准识别并剔除动态环境中物体对定位与建图过程的干扰。在此基础上，系统在实现高精度定位与地图构建的同时，有效降低了计算资源的消耗，显著扩展了系统在资源受限场景中的适用性。

3.1.3 改进的 ORB-SLAM2 系统框架

本文改进的 ORB-SLAM2 系统，主要针对跟踪线程中的视觉里程计进行优化，其系统流程图如图 3-2 所示。

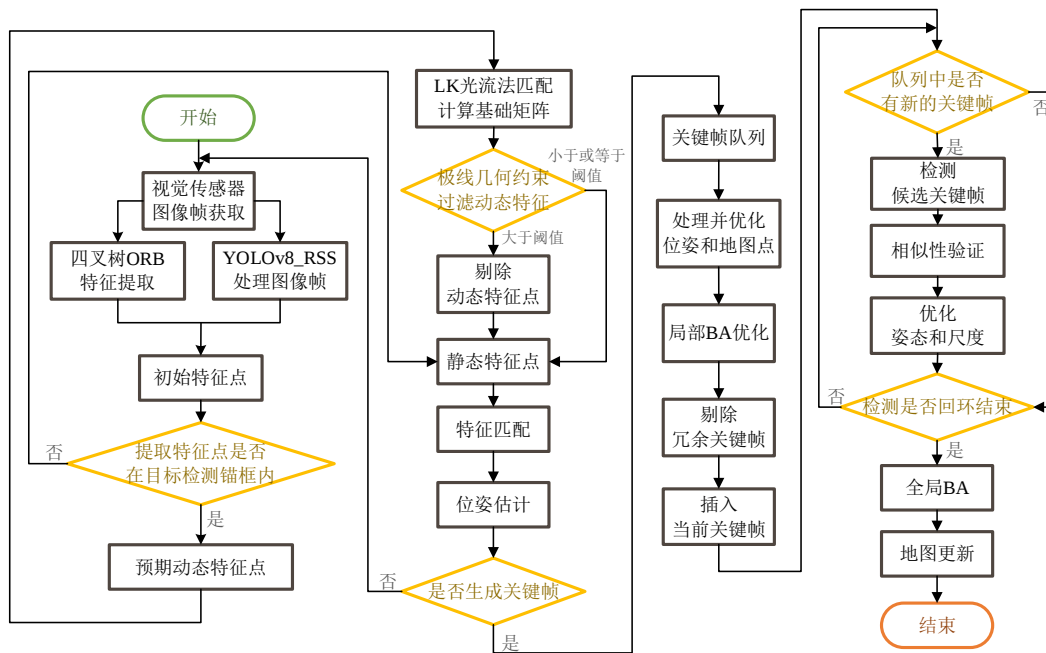


图 3-2 改进的 ORB-SLAM2 系统框架

3.2 基于 ORB 的特征点提取

ORB 算法是图像处理领域中广泛应用的特征点提取方法，它结合了 FAST 特征点检测和 BRIEF 特征描述的优势，为图像特征识别和消除提供了快速稳定的解决方案。ORB 特征点提取的质量与视觉 SLAM 系统对摄像机姿态的估计直接相关。

3.2.1 传统的四叉树 ORB 算法

四叉树算法是一种常用的特征点均匀化方法，适用于离散性较弱的 ORB 特征点的均衡处理。四叉树算法的实现包括以下步骤：

1. 将整个镜像初始化为根节点；
2. 将节点拆分为四个大小相等的子节点；
3. 删除不包含任何特征点的子节点；
4. 重复步骤 2 和步骤 3，直到每个子节点中的特征点的数量小于或等于 1，或者提取的特征点达到所需的数量；
5. 保留每个节点中具有最高响应值的特征点，并删除其余特征点。

四叉树算法有效地平衡了 ORB 特征点的分布，增强了特征点的离散性。然而，四叉树算法在使特征点提取更加均匀的同时，也可能导致两个问题：首先，在特征点累积的情况下，四叉树可能会经历多次分裂，降低特征点提取算法的速度，导致跟踪丢失；其次，该策略只考虑在最后一步保留高响应特征点，导致部分高响应特征点的丢失，特别是在高响应点密集的地区^[68]。如图 3-3 中的红框所示，点的颜色越深，其响应值就越高。在特征点删除的最后一步中，同一节点中的一些高响应点被删除，这意味着图像中的一些重要特征信息的丢失。

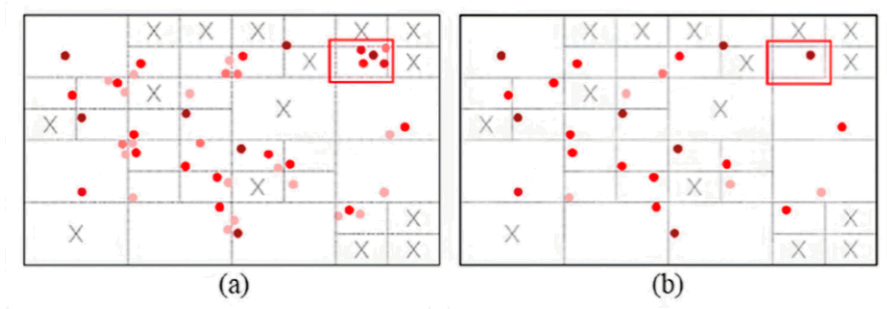


图 3-3 四叉树算法的局限性

总的来说，在图像纹理信息丰富且存在运动物体的区域，ORB 特征点容易堆积聚集，造成局部特征信息冗余。针对 ORB-SLAM2 四叉树算法的这些不足，本文提出了一种改进的四叉树算法，使特征点的分布更加均匀。

3.2.2 改进的四叉树 ORB 算法

对于包含 N 个地图点的四叉树节点，地图点的坐标为 (x_i, y_i) ， i 的取值范围为 1 到 N 。这里的四叉树节点被称为父节点，首先计算父节点的质心坐标 (x_c, y_c) 和这些地图点到质心的平均距离 $distance_aver$ ：

$$x_c = \frac{1}{N} \sum_{i=1}^N x_i, y_c = \frac{1}{N} \sum_{i=1}^N y_i \quad (3-1)$$

$$distance_aver = \frac{1}{N} \sum_{i=1}^N \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2} \quad (3-2)$$

父节点通过四叉树的划分生成子节点。对于每个子节点，计算其质心坐标 (x_{child}, y_{child}) 和每个地图点到子节点质心的平均距离 $distance_{aver_{child}}$ ：

$$x_{child} = \frac{1}{N_{child}} \sum_{i=1}^{N_{child}} x_{i_{child}}, y_{child} = \frac{1}{N_{child}} \sum_{i=1}^{N_{child}} y_{i_{child}} \quad (3-3)$$

$$distance_{aver_{child}} = \frac{1}{N_{child}} \sum_{i=1}^{N_{child}} \sqrt{(x_{i_{child}} - x_{child})^2 + (y_{i_{child}} - y_{child})^2} \quad (3-4)$$

对于每个子节点，计算局部密度标准差 σ_{child} ，即子节点内所有地图点到其质心的距离的标准差：

$$\sigma_{child} = \sqrt{\frac{1}{N_{child}} \sum_{i=1}^{N_{child}} (\sqrt{(x_{i_{child}} - x_{child})^2 + (y_{i_{child}} - y_{child})^2} - distance_{aver_{child}})^2} \quad (3-5)$$

结合平均距离差异和局部密度标准差，计算每个子节点的综合指标 D_{child} ，其中 λ_1 和 λ_2 是调整平均距离差异和局部密度标准差在 D_{child} 中的贡献的权重系数，分别反映了与父节点的分布相比的子节点内的特征点分布的紧密度，和节点内的特征点分布的均匀性或一致性。在本文中 $\lambda_1 = 1$ ， $\lambda_2 = 1$ ，表示平均距离差异和局部密度标准差对 D_{child} 的贡献相同。该综合指标 D_{child} 计算公式为：

$$D_{child} = \lambda_1 \times (distance_{aver_{child}} - distance_{aver}) + \lambda_2 \times \sigma_{child} \quad (3-6)$$

对于当前节点下所有的子节点，计算它们的 D_{child} 值的均值 μ 和标准差 σ ，其中 M 是子节点的总数：

$$\mu = \frac{1}{M} \sum_{i=1}^M D_{child,i} \quad (3-7)$$

$$\sigma = \sqrt{\frac{1}{M} \sum_{i=1}^M (D_{child,i} - \mu)^2} \quad (3-8)$$

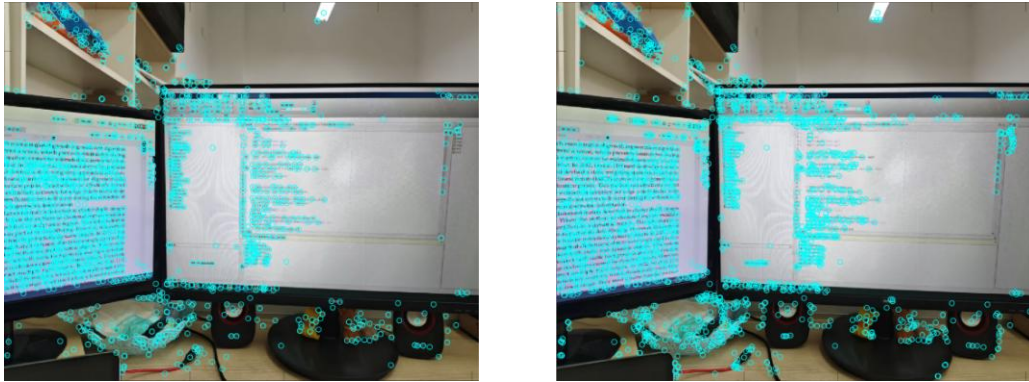
判断是否继续划分子节点，本研究基于统计学原理，利用正态分布的特性使用均值加 2 倍标准差 ($T = \mu + 2\sigma$) 作为分界线进行判断：

当 $D_{child} > T$ 时，说明当前子节点的特征点分布不均匀，需要继续划分该子节点，使得特征点更加集中。

当 $D_{child} \leq T$ 时，说明当前子节点的特征点已经较为集中或均匀，可以停止划分。

通过将改进的四叉树特征提取方法应用到 ORB-SLAM2 的 ORB 特征提取中，更加提取均匀地分布特征点。图 3-4 利用实验室真实环境，其环境复杂，正可以验证本文算法的鲁棒性。从图中，可以看出改进的 ORB 检测可以保证特

征点在图像中的分布既满足密集区域细分的要求，又避免了稀疏区域过度划分。



(a)原始 ORB 特征提取

(b)改进的 ORB 特征提取

图 3-4 四叉树处理后的特征提取与原始 ORB 对比

该方法使得 ORB-SLAM2 在面对复杂的场景时，能够更高效、更准确地提取特征点，为后续的位姿估计和地图构建提供更可靠的支持。

3.3 基于 YOLOv8 的目标检测

3.3.1 改进的 YOLOv8 框架

本研究基于 YOLOv8 提出了一种改进的小目标检测算法 YOLOv8_RSS。首先，本研究引入 RepGhost 模块轻量化网络模型，减少计算量并基本保证检测精度，改善无人机计算资源有限的问题；其次，提出融合 SimAM 注意力机制的 SimAMConv，增强对小目标的关注和学习；再者，提出引入 Inner-IoU 的 SIoU 损失函数加速边界框回归，提高模型泛化能力和小目标检测精度；最后，重构网络框架，增加极小目标检测层，减少大目标检测层，进一步轻量化模型并提升小目标检测精度。改进后的网路结构如图 3-5 所示。

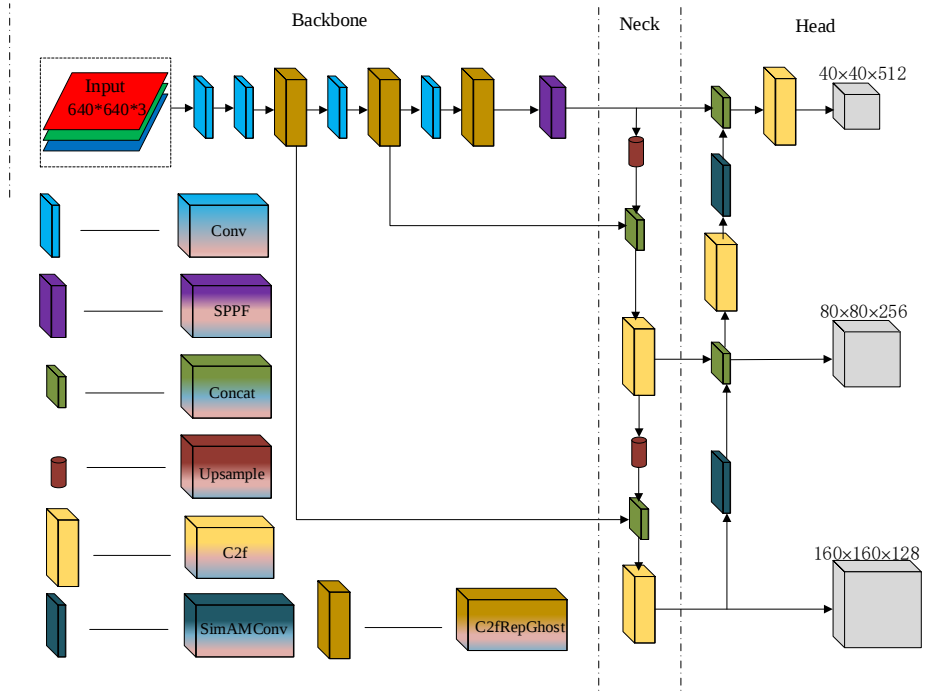


图 3-5 改进的 YOLOv8 框架

3.3.2 C2fRepGhost 轻量化模块

RepGhost 模块作为轻量级网络，是其作者在华为诺亚方舟实验室打造的轻量级网络 GhostNet 的基础之上，提出了一种新的硬件高效的端侧骨干模块。该模块利用结构重新参数化来实现隐式的特征重用，删除了低效的串联运算，并修改了体系结构以满足结构重新参数化的规则，实现功能重用隐式取代低效的连接操作 Ghost 模块。根据其作者的研究，Concat 操作比 Add 操作更加耗时，并且 Add 操作作为一种处理不同特征的简单操作，所需成本较低。简单来说，RepGhost 模块中的张量操作采用了 Add 操作，而不是 Ghost 模块中的 Concat 操作，从而提高了效率；并在 Add 操作后使用 ReLU 激活函数，以满足重新参数化结构的规则。

如图 3-6，RepGhost 模块的结构演化可以简化为两步。如图 3-6-a，首先将 Ghost 模块中 ReLU 激活函数衔接到替换 cat 结构的 Add 之后，并在恒等映射分支中添加 BN 操作得到图 3-6-b，最后融合图 3-6-b 模块得到最终的图 3-6-c 模块，即 RepGhost 模块。可见，RepGhost 模块的推理结构极为简单，仅包含卷积层和 ReLU，从而使得这种结构具有更高的硬件效率。

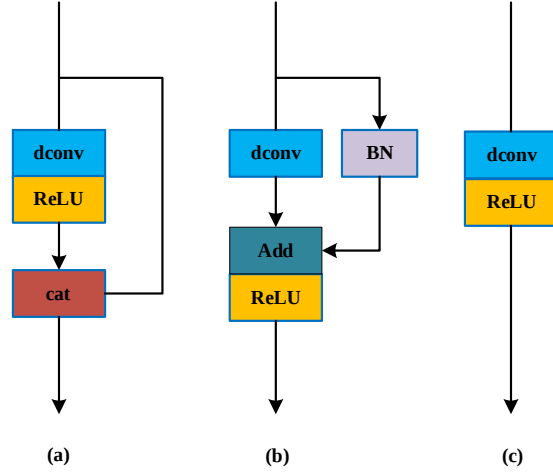


图 3-6 RepGhost 结构演化

如图 3-7，RepGhost Bottleneck 是通过两个 RepGhost 模块堆叠输出后，与输入端进行 Add 操作以完成特征融合，其中 RepGhost 模块分别用于增加和减少通道数量，用来满足与输入端进行特征融合的条件。

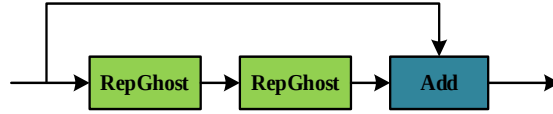


图 3-7 RepGhost Bottleneck 结构图

YOLOv8 网络中的 C2f 模块为特征提取提供了丰富的梯度流信息，为了进一步提高 C2f 模块的特征表征能力，本研究通过将 C2f 模块中 Bottleneck 替换为 RepGhost Bottleneck 设计了新的 C2fRepGhost 模块，如图 3-8 所示。具体来说，在 Bottleneck 中引入 RepGhost 模块，可以利用重参数化方法生成和融合特征图，从而将训练阶段的复杂结构转化为推理阶段的简单结构；同时由于 RepGhost Bottleneck 在特征融合过程中采用了更高效的 Add 操作，减少了网络设计中的复杂性，使整个模块更加轻量化并降低延迟，更符合部分设备资源受限的场景。

在前向传播过程中，C2fRepGhost 模块通过引入结构重新参数化技术，使网络模型在体积以及参数上得到压缩的同时获得硬件高效这一属性，使得模型更易于部署在无人机设备，更适用于基于无人机航拍的小目标检测。

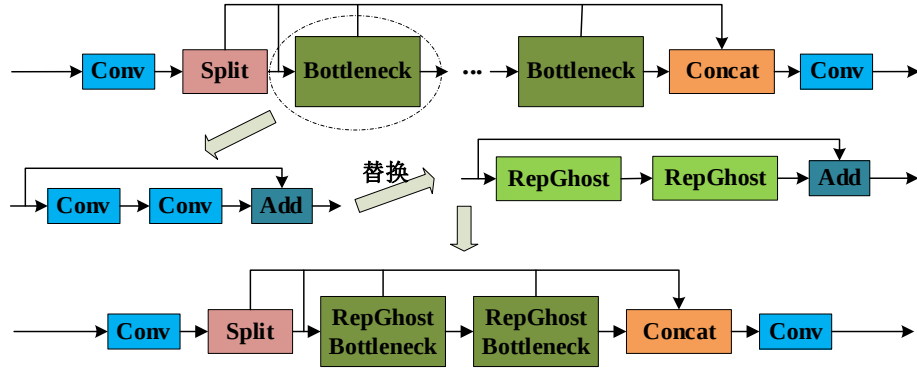


图 3-8 C2fRepGhost 结构演化

3.3.3 SimAM 注意力机制

在检测无人机航拍图像中的小目标时，网络模型往往会因为目标特征表现能力较弱、位置分布不定等缺点而忽略这类小目标，进而导致漏检误检。针对以上问题，为了更好地捕捉图像中的重要特征，提高网络模型的表达能力和性能，本研究引入 SimAM 注意力机制，其结构如图 3-9 所示。

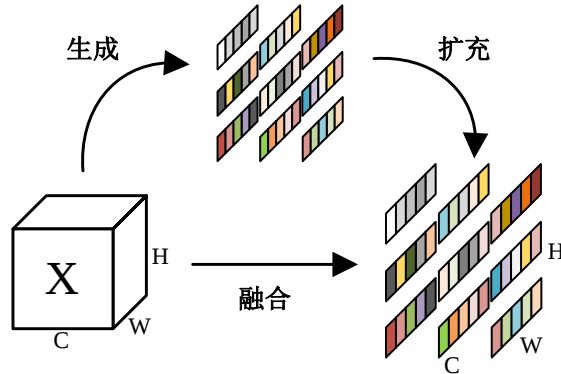


图 3-9 SimAM 注意力结构图

SimAM 作为无参的 3D 注意力机制，与 1D 注意力和 2D 注意力相比存在许多优点。在关注度上，SimAM 注意力在所有位置上以及所有通道上同等对待，弥补了另外两种注意力机制的单一关注的缺点；在计算量上，SimAM 注意力机制无需额外的参数量，能够直接在网络层中推理出三维注意力权重，更加容易在算力资源有限的设备中部署。

其中三维注意力权重可以用公式进行表达，如公式(3-9)至公式(3-12)， X 作为图像特征输入， E 作为每个通道上的最小能量函数， M 可以表示每个通道上值的数量，可以用 $M=H \times W$ 得到， t 代表输入特征的神经元， λ 为权重常数， μ 和

σ^2 代表的是输入特征中每个通道上的均值与方差。

$$\sigma^2 = \frac{1}{M} \sum_{i=1}^M (x_i - \mu)^2 \quad (3-9)$$

$$\mu = \frac{1}{M} \sum_{i=1}^M x_i \quad (3-10)$$

$$E = \frac{4(\sigma^2 + \lambda)}{(t - \mu)^2 + 2\sigma^2 + 2\lambda} \quad (3-11)$$

$$\dot{X} = \text{sigmoid}\left(\frac{1}{E}\right) \odot X \quad (3-12)$$

通过上述公式，利用能量函数直接推导出特征图的三维注意力权重信息，在不引入额外参数的前提下，强化了特征图的空间位置信息表达，从而提升了特征选择的精确性。

本研究将 SimAM 注意力机制与 Conv 模块融合，设计了改进后的注意力卷积 SimAMConv，结构如图 3-10。具体来说，先通过二维卷积操作进行特征提取，再进行归一化处理加速模型收敛速度，随后利用 SiLU 激活函数对归一化后的特征图进行非线性变换，最后进行 SimAM 注意力机制得到经过该注意力机制增强后的特征输出。

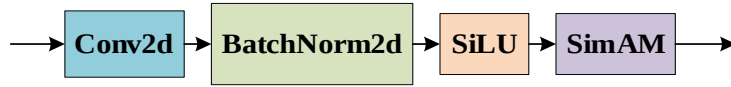


图 3-10 SimAMConv 结构图

改进的注意力卷积模块，通过计算注意力权重抑制无关或噪声特征，从而提高模型对有用信息的关注，增强对航拍图像中小目标与上下文特征中的相关性，从而提升模型性能。

3.3.4 Inner-SIoU Loss

查阅目标检测相关文献不难看出，损失函数的设计在一定程度上可以影响目标检测算法性能。在原 YOLOv8 网络中，采用 CIoU 为损失函数，通过引入真实框和预测框中心之间的距离和两框对角线距离之比来衡量两个框之间的距离，可以提高检测结果的收敛速率。但是当真实框和预测框的长与宽的比值相同但是数值却不同的时候，两个框之间真实的差异就很难准确的表现出来。另外，CIoU 在公式计算中引入反三角函数，增加了设备的计算负荷，不利于一些有限资源环境的设备。CIoU 计算公式如下：

$$Loss_{CIoU} = 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} + \left(\frac{v}{(1-IoU)+v} \right) v \quad (3-13)$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (3-14)$$

其中, b 和 b^{gt} 表示预测框与真实框的中心点, ρ 表示预测框与真实框中心点之间的欧式距离, c 表示包含预测框和真实框的最小矩阵的对角距离, w^{gt} 和 h^{gt} 表示真实框的宽度和高度, w 和 h 表示预测框的宽度和高度。

为提高算法模型的加速收敛能力, 本研究尝试使用 SIoU 作为网络模型的边框损失函数。相较 CIoU 损失函数, SIoU 一样考虑到了真实框和预测框之间的中心点距离、长宽比等要素。但 SIoU 重新定义了惩罚项, 并引入了角度损失, 有效得减少了自由度的数量, 进一步提升目标检测的性能。

角度损失定义公式如下:

$$\Lambda = 1 - 2 * \sin^2 \left(\arcsin(x) - \frac{\pi}{4} \right) \quad (3-15)$$

式中: $x = \frac{c_h}{\sigma} = \sin(\alpha), \sigma = \sqrt{(b_{c_x}^{gt} - b_{c_x})^2 + (b_{c_y}^{gt} - b_{c_y})^2}, c_h = \max(b_{c_y}^{gt}, b_{c_y}) - \min(b_{c_y}^{gt}, b_{c_y})$

距离损失定义公式如下:

$$\Delta = \sum_{t=x,y} (1 - e^{-\gamma \rho^t}) \quad (3-16)$$

式中: $\rho_x = \left(\frac{b_{c_x}^{gt} - b_{c_x}}{c_w} \right)^2, \rho_y = \left(\frac{b_{c_y}^{gt} - b_{c_y}}{c_h} \right)^2, \gamma = 2 - \Lambda$

由上式可知, 距离损失的贡献度随着 α 趋于 0 后显著降低。反之, 当 α 趋于 45° 时, 距离损失的贡献度显著升高。因此, 随着角度的不断增大, γ 被赋予了距离值的时间优先级。值得注意的是, 当 α 趋于 0 时, 距离损失将与常规的距离损失相同。

形状损失定义公式如下:

$$\Omega = \sum_{t=w,h} (1 - e^{-\omega_t})^\theta \quad (3-17)$$

式中, $\omega_w = \frac{|w - w^{gt}|}{\max(w, w^{gt})}, \omega_h = \frac{|h - h^{gt}|}{\max(h, h^{gt})}$ 。 θ 控制损失函数对形状损失的关注程度, 为了避免过度关注形状损失而降低对预测框移动的关注, 将 θ 的取值范围限制在 2 到 6 之间。

综上, SIoU 损失函数的计算公式定义如下:

$$L_{SIoU} = 1 - IoU + \frac{\Delta + \Omega}{2} \quad (3-18)$$

$$\text{式中, } IoU = \frac{|B \cap B^{GT}|}{|B \cup B^{GT}|}$$

SIoU 损失函数考虑到所需回归之间的向量角度, 在处理边界框回归问题上相比 CIoU 有明显的创新和优势。

然而这种添加新的几何约束来加速收敛的办法, 并没有考虑到 IoU loss 本身的合理性。为弥补这一不足, 本研究在 SIoU 损失函数中引入 Inner-IoU Loss, 即使用比例因子比率控制来生成不同尺度的辅助边界框以计算 IoU 损失。如图 3-11 所示, b 、 b^{gt} 表示锚框和真实边界框, (x_c^{gt}, y_c^{gt}) 为真实边界框的中心点, (x_c, y_c) 为锚框的中心点, w^{gt} 、 h^{gt} 、 w 、 h 分别表示真实边界框和锚框的宽与高。

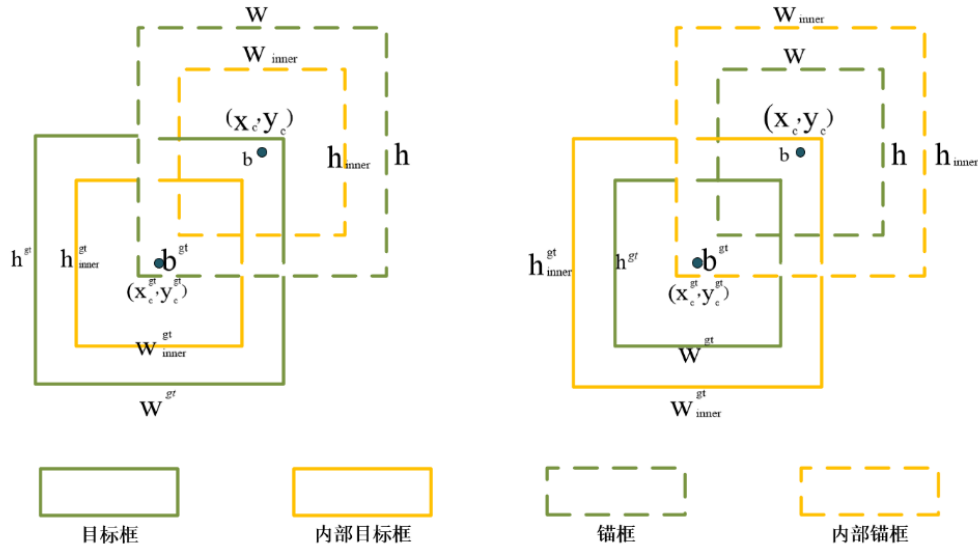


图 3-11 Inner-IoU Loss

将该方法应用于 SIoU 损失函数, 计算方式如下所示:

$$b_l^{gt} = x_c^{gt} - \frac{w^{gt} * ratio}{2}, b_r^{gt} = x_c^{gt} + \frac{w^{gt} * ratio}{2} \quad (3-19)$$

$$b_t^{gt} = y_c^{gt} - \frac{h^{gt} * ratio}{2}, b_b^{gt} = y_c^{gt} + \frac{h^{gt} * ratio}{2} \quad (3-20)$$

$$b_l = x_c - \frac{w * ratio}{2}, b_r = x_c + \frac{w * ratio}{2} \quad (3-21)$$

$$b_t = y_c - \frac{h * ratio}{2}, b_b = y_c + \frac{h * ratio}{2} \quad (3-22)$$

$$inter = (\min(b_r^{gt}, b_r) - \max(b_l^{gt}, b_l)) * (\min(b_b^{gt}, b_b) - \max(b_t^{gt}, b_t)) \quad (3-23)$$

$$union = (w^{gt} * h^{gt}) * (ratio)^2 + (w * h) * (ratio)^2 - inter \quad (3-24)$$

$$IoU^{inner} = \frac{inter}{union} \quad (3-25)$$

$$L_{Inner-SIoU} = L_{SIoU} + IoU - IoU^{inner} \quad (3-26)$$

利用公式(3-19)至公式(3-22)可得, 辅助锚框左边界横坐标 b_l 、右边界横坐标 b_r 、下边界纵坐标 b_t 、上边界纵坐标 b_b , 其中 b_l^{gt} 、 b_r^{gt} 、 b_t^{gt} 、 b_b^{gt} 分别为辅助边界框左边界横坐标、右边界横坐标、下边界纵坐标、上边界纵坐标。

利用公式(3-23)至公式(3-25)可知, inner-IoU 的计算是指辅助边界框之间的IoU 计算。ratio 的取值范围为 0.5 至 1.5。当 ratio 小于 1 时, 辅助边界框小于实际边界框, 导致回归的有效范围小于 IoU 损失, 从而加速高 IoU 样本的收敛。相反, 当 ratio 大于 1 时, 较大尺寸的辅助边界框扩展了回归的有效范围, 提升了对低 IoU 样本的回归效果。

对比传统的 IoU 计算方法, Inner-IoU 通过引入比例因子 ratio 提供了辅助边框的可变化特性, 提升了对图像核心部分的关注度, 提高了算法模型的泛化能力。将 Inner-IoU 与 SIoU 损失函数结合, 新的 Inner-SIoU 有助于提高收敛能力, 可以实现更快更有效的回归结果, 预期在复杂环境更能突出其特点, 更适用于无人机航拍图像的小目标检测。

3.3.5 多尺度特征融合结构

YOLOv8 原始网络结构使用了三个检测头, 尺寸分为 20×20 、 40×40 、 80×80 , 对应检测大型目标、中型目标和小型目标。但是无人机航拍图像中小目标像素极小, 即使采用 80×80 的小型目标检测头, 在识别这类目标时也表现出较低的精度。

针对无人机航拍图像中小目标图像特征, 基于官方发布的 YOLOv8 网络结构, 本研究提出一种适用于小目标检测的新的多尺度特征融合结构, 即在原网络结构的基础上剔除大目标检测头并增加小目标检测头。具体实现: 在骨干网络中剔除特征提取层 P5, 并将金字塔池化层向上衔接于 P4 层之后; 同时将从 P2 层引出新的 160×160 尺度特征层与颈部结构中的上采样特征层继续向上拼接, 从而增强 160×160 尺度融合特征层对小目标语义特征以及位置信息的表达能力, 提升了网络对于无人机航拍图像中小目标的检测能力。

在 Head 结构中, 参考原网络拼接方式, 通过下采样操作将新增的小目标检测头中的特征信息传递到剩下两个尺度特征层, 进一步提升网络的特征融合能力, 从而提高小目标的检测精度。

3.3.6 消融实验分析

为验证所设计的 C2fRepGhost 模块、SimAMConv 模块、Inner-SIoU 损失函数以及网络结构重构的有效性，本研究一共进行了十一项消融实验。本实验采用由中国天津大学机器学习和数据挖掘实验室 AISKYEYE 团队收集的 VisDrone2018 数据集，该数据集包含 10209 张图像，其中 6471 张用于训练，548 张用于验证，3190 张用于测试，包括行人、自行车、公交车等 10 个类别。实验采用随机梯度下降(SGD)方法使用当前梯度信息来调整学习率，详细参数如表 3-1 所示：

表 3-1 数据训练参数

参数名称	值
初始学习率(lr_0)	0.003
最终学习率(lr_end)	0.01
动量(momentum)	0.9
权重衰减(lr_decay)	0.0005
训练预热期轮次	3
比例因子(ratio)	0.8
训练轮次	300
图像分辨率	640×640

实验结果如表 3-2 所示。实验 A 为 YOLOv8 作为基准算法在 VisDrone2018 数据集上进行训练和测试；实验 B 为将骨干网络中 C2f 模块替换为 C2fRepGhost 模块；实验 C 为将检测头部分中 Conv 模块替换为改进的 SimAMConv 模块；结合实验 B 和实验 C 的改进，即在骨干网络中使用 C2fRepGhost，并在检测头中使用 SimAMConv，以同时优化特征提取和表达能力；实验 E 在 YOLOv8 的损失函数中引入 Inner-SIoU；实验 F：在 YOLOv8 的结构中引入小检测头；实验 G 在实验 F 的基础上，进一步将骨干网络中的 C2f 模块替换为 C2fRepGhost；实验 H 在实验 G 的基础上，进一步加入 Inner-SIoU 损失；实验 I 在实验 G 的基础上，进一步在检测头中使用 SimAMConv 模块；实验 J 在实验 I 的基础上，进一步加入 Inner-SIoU 损失；实验 K 在 YOLOv8 的 Neck 和 Head 部分，将 C2f 模块替换为 C2fRepGhost。依此类推共十一项实验。

表 3-2 消融实验

实验	模型方法	实验结果(fused)					
		mAP_50 /%	P /%	R /%	GFLOs	Params /M	Weight size /MB
A	YOLOv8	40.03	51.68	38.33	28.5	11.13	22.5
B	YOLOv8+C2fRepGhost(Backbone)	38.87	51.39	37.08	23.6	9.43	19.2
C	YOLOv8+SimAMConv	40.79	53.23	38.28	28.5	11.13	22.5
D	YOLOv8+C2fRepGhost(Backbone) +SimAMConv	38.93	51.83	37.01	23.6	9.43	19.2
E	YOLOv8+Inner-SIoU	40.41	52.01	38.57	28.5	11.13	22.5
F	YOLOv8+小检测头	45.31	54.63	42.32	29.8	3.30	6.9
G	YOLOv8+C2fRepGhost(Backbone) +小检测头	44.67	54.49	42.54	25.8	2.60	5.6
H	YOLOv8+C2fRepGhost(Backbone) +小检测头+Inner-SIoU	45.16	55.61	42.18	25.8	2.64	5.6
I	YOLOv8+C2fRepGhost(Backbone) +SimAMConv+小检测头	44.89	56.31	42.05	25.8	2.64	5.6
J	YOLOv8+C2fRepGhost(Backbone) +SimAMConv+小检测头+Inner-SIoU	45.32	56.34	42.53	25.8	2.64	5.6
K	YOLOv8+C2fRepGhost(Neck+Head)	38.26	51.03	37.06	22.1	8.72	18.9

由表 3-2 可知，实验 B 使用了改进的 C2fRepGhost 模块替换基准算法 YOLOv8 中骨干部分的 C2f 模块，相较于基准算法，实现了部分轻量化，参数量和计算量分别下降 1.7M 和 4.9GFLOPs，但 mAP50 有所减少。实验 C 使用了融合无参注意力机制的 SimAMConv 替换基准算法网路结构 head 部分中所有 Conv 模块，实验表明，相较于基准算法在计算量、参数量和模型权重大小不变的情况下，提高 0.76% 的 mAP50 和 1.55% 的精确率 P。实验 E 通过将基准算法中 CIoU 替换成 Inner-SIoU，相较于基准算法 mAP50 提高 0.38%、p 提高 0.33%、R 提高 0.24%。实验 F 对基准算法网路结构进行改进，剔除原结构中的大目标检测层，并增加小目标检测层，相较于基准算法 mAP50 提高 5.28%、P 提高 2.95%、R 提高 3.99%，证明了该方法在小目标检测上有明显效果。实验 J 是在多次实验基础上，依次采用上述方法，与基准算法实验结果相比，在提高精度的同时也极大程度上实现了轻量化，参数量降低 8.49M、计算量降低 2.7GFLOPs、模型权重降低 16.9MB，同时 mAP50 提高 5.29%、P 提高 4.66%、R 提高 4.2%。其他实验利用了不同改进方法进行融合，进一步对比出各模块对基准算法的影响，其中实验 B 与实验 K 分别探究了 C2fRepGhost 模块在网络结构的骨干网络，以及在颈部与头部中不同的作用效果，通过两个实验对比可以得出，将

C2fRepGhost 模块应用于 YOLOv8 的骨干网络中，可以体现该模块的有效性。

通过以上实验数据对比可得，本研究算法 YOLOv8_RSS 在有效提高检测精度的同时，大幅降低了参数量、计算量和模型权重，更适用于部署在计算资源较少的无人机设备

3.3.7 可视化分析

为验证本研究改进算法 YOLOv8_RSS 在无人机航拍小目标检测中的性能，改进后的算法模型与基准算法进行实验训练，得出的精确率 P、召回率 R、mAP50 指标对比如图 3-12 所示。

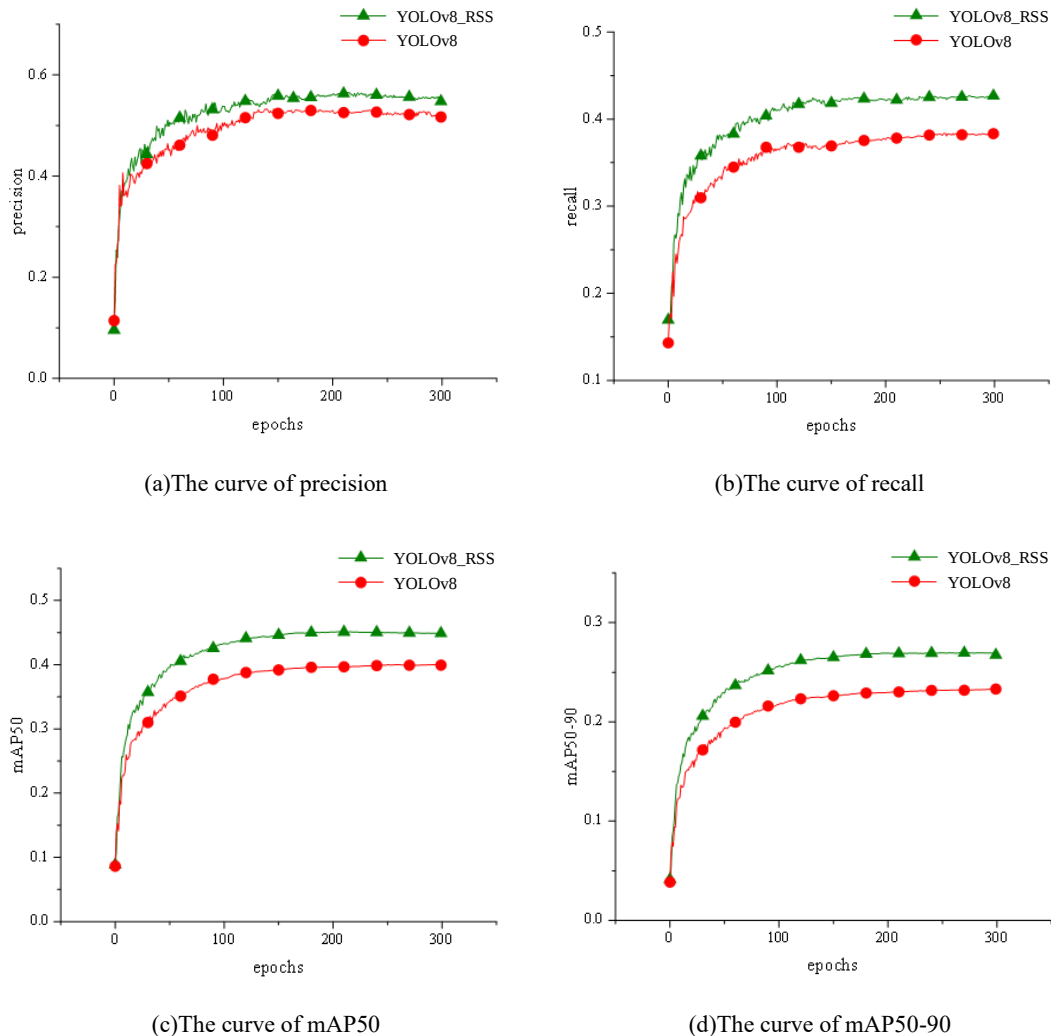


图 3-12 改进算法与基准算法的训练结果

由图可知，两种算法都可以通过上百次迭代达到收敛，但与原模型相比，改进后的网络模型在收敛速度上明显更快，并且有着更好的数值表现。因此证

明了该算法模型在该环境下有着更好的检测性能。

为验证本研究改进算法在实际场景中的检测效果，从几种不同无人机航拍场景中任意选取多张图片，进行基准网络模型与改进网络模型对比实验，检测效果如图 3-13 所示，其中左侧为改进算法，右侧为基准算法。

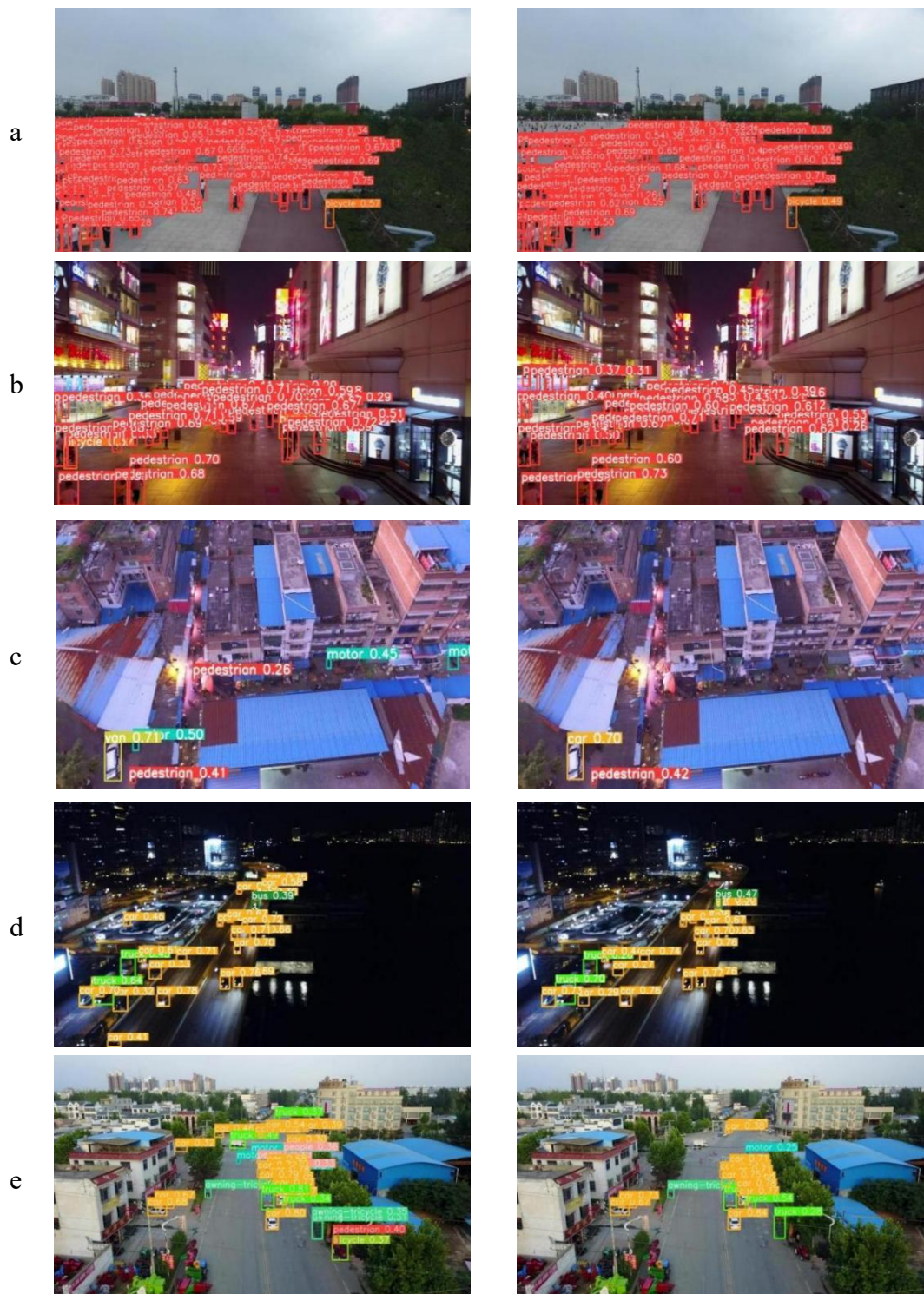


图 3-13 VisDrone2018 数据集上的检测结果可视化。改进算法（左）；基准算法（右）

在 a 类密集环境中可以看出,改进后的算法可以识别更远处更小的目标,反观基准算法的检测效果欠佳;在 b 类复杂背景环境中可以看出,基准算法的误识别率偏高,将图片中左侧的海报检测为行人,虽然改进后的算法也有误识别,但这个概率相比较原算法已经大大降低;在 c 类多尺度环境中可以看出,改进算法与基准算法的表现均不是很好,但改进算法的可识别小目标的个数有小幅增加;在 d 类极暗环境下可以看出,改进算法表现极为优秀,在可识别率和识别准确率上均优于基准算法,大大提升了该种环境下的不同尺度上识别效果;在 e 类目标遮挡环境中可以看出,基准算法不存在优势,尤其对于远处极小目标上表现得更差,但是改进算法无论在远处还是近处都表现良好。

总的来说,在极暗环境下,本研究改进算法对小目标的检测效果表现良好;在多尺度环境下,本研究改进算法可以做到兼顾中目标识别的同时,对小目标和大目标也有着较好的检测效果。因此本研究改进算法 YOLOv8_RSS 对无人机航拍小目标的检测效果优于 YOLOv8 的检测效果,误检率和漏检率得到有效降低。同时针对不同情况下的复杂环境提升了检测能力,也证明了改进算法在多种环境下的鲁棒性和可靠性。

3.3.8 对比实验分析

为验证 YOLOv8_RSS 在无人机小目标检测的有效性,本研究在相同实验环境和数据集 VisDrone2018 上,将该算法与相关领域主流算法进行对比实验,结果如表 3-3 所示。

表 3-3 不同网络模型下对比实验

实验	模型方法	实验结果(fused)			
		mAP_50/%	GFLOPs	Params/M	FPS
A	Faster-RCNN	33.5	206.8	41.19	29.0
B	EfficientNet	35.8	35.2	8.31	73.3
C	YOLOv3	31.4	28.2	10.37	107.7
D	YOLOv5s	34.2	8.4	1.76	135.8
E	YOLOv6	30.3	13.0	4.49	103.2
F	Improved YOLOv8	45.32	25.8	2.64	122.0

由对比实验结果可知,本研究改进算法 YOLOv8_RSS 的综合水平均优于其他模型。与 YOLOv5s 相比,改进算法的 mAP₅₀ 比该算法提高了 11.12%,但计算效率和检测速度方面略低,原因可能包括 YOLOv5s 模型作为 YOLOv5 系列中最小的模型,主要部署于移动设备或边缘设备,该模型虽然检测速度快但准确度相对较低。另外,与其他模型方法相比,改进算法在检测精度和计算效率上均优于其他模型。因此,本研究提出的 YOLOv8_RSS 实现降低参数量和计算量的情况下,大大提高检测速度和精确度。从各项指标可以看出,对于无人机航拍图像的小目标检测,本研究的改进算法有着优异的性能表现。

3.4 基于语义和几何约束算法实现动态特征点剔除

为了使系统适应动态场景、图像纹理丰富以及过多的计算资源消耗的复杂环境,本文在 ORB-SLAM2 的基础上引入了基于语义和几何约束算法的动态 ORB 特征剔除方法。基本方法如下,首先,利用改进的基于四叉树的 ORB 算法提取图像中的 ORB 特征点,并通过改进的 YOLOv8 生成的语义边界框对提取的特征点进行初步筛选,以剔除属于动态对象的特征点。保留的特征点通过光流法与上一帧的特征点进行匹配,形成候选匹配特征对。基于这些匹配特征对,采用鲁棒估计方法计算稳定的基本矩阵,以分析特征点的运动一致性。最终,通过几何约束进一步剔除与静态场景运动模型不一致的动态特征点,保留静态特征点用于后续的定位与建图。此剔除流程充分结合了语义信息和几何信息,显著增强了系统在动态场景下的鲁棒性与效率。

3.4.1 光流法匹配

光流法(Optical Flow)是计算机视觉领域中的一种重要技术,主要用于估计图像中每个像素点在相邻帧之间的运动。其基本假设是,在短时间内,图像中物体的亮度保持恒定,因此通过分析像素亮度的变化可以推测物体的运动。光流法在运动检测中的应用通常分为两种类型:密集光流和稀疏光流。稀疏光流专注于图像中的部分特征点,通常采用 Lucas-Kanade(LK)方法进行计算,它通过在局部区域内假设像素运动的一致性来高效地估计特征点的运动。为了提升计算效率,本文选择仅对经过动态目标过滤的 ORB 特征点计算光流场,并采用

LK 光流法。

本文的动态目标检测方法并未对稀疏光流进行运动一致性分析，而是直接利用稀疏光流法进行特征点匹配。在光流匹配过程中，特征点之间的对应关系可以直接通过光流场获得，无需依赖于传统的描述符计算。通过帧间跟踪，匹配的特征点可用于求解基本相机矩阵。光流图如图 3-14 所示，其中 I_{t1} 、 I_{t2} 、 I_{t3} 分别为图像在 t_1 、 t_2 、 t_3 时刻的灰度值。

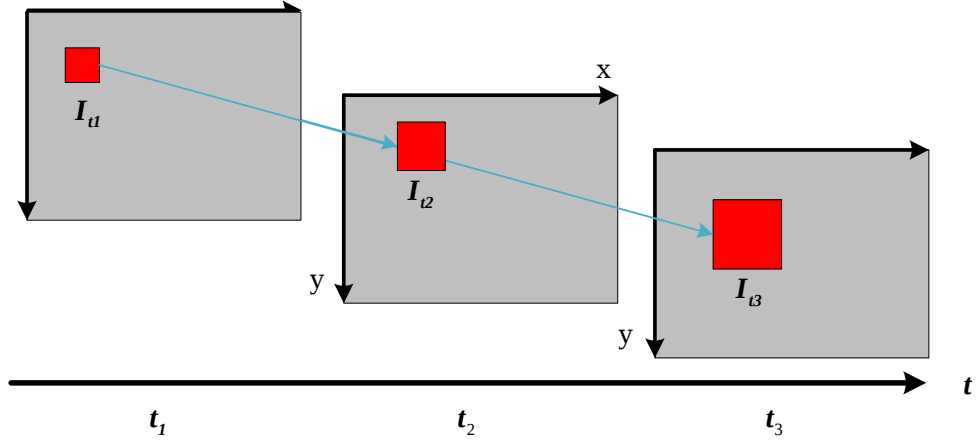


图 3-14 光流图

图像的灰度值可以视为时间的函数。在某一时刻 t ，图像中位于坐标 (x, y) 的 ORB 特征点的灰度值可表示为 $I(x, y, t)$ ，当时间变为 $t + dt$ 时，该特征点的位置移至 $(x + dx, y + dy)$ 。依据灰度值的不变性原理，可以推导出以下结论：

$$I(x, y, t) = I(x + dx, y + dy, t + dt) \quad (3-27)$$

将上述方程进行泰勒展开，并保留至一阶项，可以得出以下结果：

$$I(x + dx, y + dy, t + dt) \approx I(x, y, t) + \frac{\partial I}{\partial x} dx + \frac{\partial I}{\partial y} dy + \frac{\partial I}{\partial t} dt \quad (3-28)$$

在假设灰度值保持恒定的前提下，特征点在时刻 t 以及时刻 $t + dt$ 的灰度值相等，由此可得出：

$$\frac{\partial I}{\partial x} dx + \frac{\partial I}{\partial y} dy + \frac{\partial I}{\partial t} dt = 0 \quad (3-29)$$

两边除以 dt ，可以得出：

$$\frac{\partial I}{\partial x} \frac{dx}{dt} + \frac{\partial I}{\partial y} \frac{dy}{dt} = -\frac{\partial I}{\partial t} \quad (3-30)$$

其中，像素在 X 轴和 Y 轴上的速度分别定义为 dx/dt 和 dy/dt (记为 u 和 v)，图像在 X 轴和 Y 轴上的方向梯度分别为 $\frac{\partial I}{\partial x}$ 和 $\frac{\partial I}{\partial y}$ (记为 I_x 和 I_y)， $\frac{\partial I}{\partial t}$ 则表示灰度值随

时间的变化率(记为 I_t)。基于此, 上式可进一步转化为以下矩阵形式:

$$[I_x I_y] \begin{bmatrix} u \\ v \end{bmatrix} = -I_t \quad (3-31)$$

为了求解式(3-31)中的两个未知数 u 和 v , 我们假定在一个窗口内的像素具有相同的运动特性, 并设定一个大小为 $2w$ 像素的“ $w \cdot w$ ”窗口。由于该窗口内所有像素的运动一致, 由此可以推导出 w^2 方程:

$$[I_x I_y]_k \begin{bmatrix} u \\ v \end{bmatrix} = -I_{tk} (k = 1, 2, \dots, w^2) \quad (3-32)$$

$$A = \begin{bmatrix} [I_x I_y]_1 \\ \vdots \\ [I_x I_y]_k \end{bmatrix}, b = \begin{bmatrix} I_{t1} \\ \vdots \\ I_{tk} \end{bmatrix} \quad (3-33)$$

则式(3-32)可写成:

$$A \begin{bmatrix} u \\ v \end{bmatrix} = b \quad (3-34)$$

利用最小二乘法, 可以得到:

$$\begin{bmatrix} u \\ v \end{bmatrix}^* = -(A^T A)^{-1} A^T b \quad (3-35)$$

在语义过滤后, 保留的特征点通过光流法与上一帧图像中的特征点进行匹配。光流法能够根据特征点的运动信息对其进行匹配, 通过运动一致性判断哪些特征点可能是动态的。经过光流法匹配后, 形成一组候选的匹配特征对, 这些特征对将作为后续几何约束分析的基础。

3.4.2 对极几何约束

考虑到语义约束的局限性, 本文使用几何约束方法来进一步过滤掉动态特征。极线约束, 它是摄影几何学的一部分独立于两个映像之间的外部环境而存在, 并且仅依赖于摄像机的内部参数和相对姿势。极线约束之间的几何关系如图 3-15。

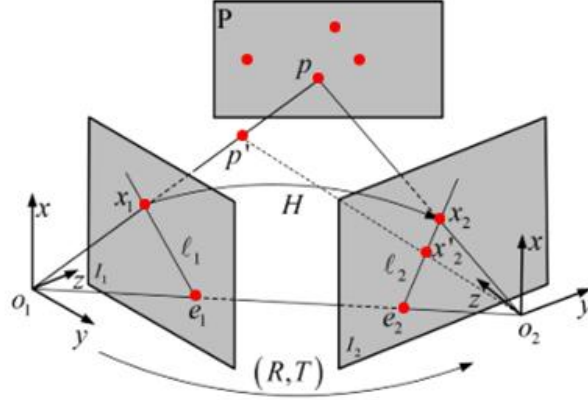


图 3-15 对极几何约束示意图

在图 3-15 中, I_1 和 I_2 是从同一场景捕获的两幅图像, p 为 P 空间中的某一空间点, p 在 I_1 和 I_2 的投影分别为 x_1 和 x_2 ; O_1 和 O_2 是两台相机的光学中心, 它们的连线分别在 I_1 和 I_2 相交于 e_1 和 e_2 。 ℓ_1 和 ℓ_2 分别是 I_1 和 I_2 的极线, 其可以通过将 I_1 、 I_2 和 PO_1O_2 平面相交而获得。

根据针孔相机模型, 在 I_1 和 I_2 两帧图像中, 点像素坐标符合以下约束条件:

$$\begin{cases} s_1 x_1 = Kp \\ s_2 x_2 = K(Rp + T) \end{cases} \quad (3-36)$$

上式中, K 为摄像机的内参矩阵, R 和 T 分别为两个摄像机坐标系之间的旋转和平移矩阵, S 为像素的深度信息。

利用归一化坐标 $\bar{p}_1 = k^{-1}x_1$, $\bar{p}_2 = k^{-1}x_2$, 可以得到:

$$\bar{p}_2 = R\bar{p}_1 + t \quad (3-37)$$

$$(\bar{p}_2)^T [t]_x R \bar{p}_1 = 0 \quad (3-38)$$

上式中, $[t]_x$ 为向量 t 的反对称矩阵, 将上式整理, 可以等到:

$$x_2^T F x_1 = 0 \quad (3-39)$$

$$F = k^{-T} [t]_x R k^{-1} \quad (3-40)$$

上式中, F 为基本矩阵, 于是将计算 x_2 到极线的距离定义 D 为:

$$D = \frac{|x_2^T F x_1|}{\sqrt{\|X\|^2 + \|Y\|^2}} \quad (3-41)$$

θ 为预设阈值, 计算公式为:

$$\theta = \frac{\sum_{i=0}^N e^{-D_i}}{N} \quad (3-42)$$

如果 $D > \theta$, 则该点不满足极坐标约束, 被视为动态特征, 需被剔除。

3.4.3 动态点剔除效果验证

基于语义和几何约束算法实现动态特征点剔除的流程图如下：

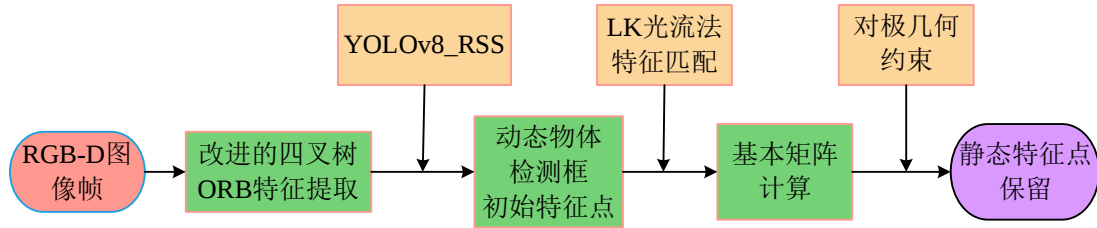


图 3-16 算法实现流程图

首先，利用改进的基于四叉树的 ORB 算法从图像序列中提取稳定且均匀分布的 ORB 特征；再利用 YOLOv8_RSS 生成的语义边界框过滤动态特征；然后将剩余的特征通过 LK 光流法与前一帧进行匹配，匹配的特征对用来计算基本矩阵；最后，使用极线几何约束过滤真正的动态特征。

通过上述流程消除图像中动态物体上的特征点。如图 3-17，原始特征点分布图（左）和去除后的特征点分布图（右）如下：



图 3-17 动态特征点剔除效果图。未剔除（左）；已剔除（右）



续图 3-17 动态特征点剔除效果图。未剔除（左）；已剔除（右）

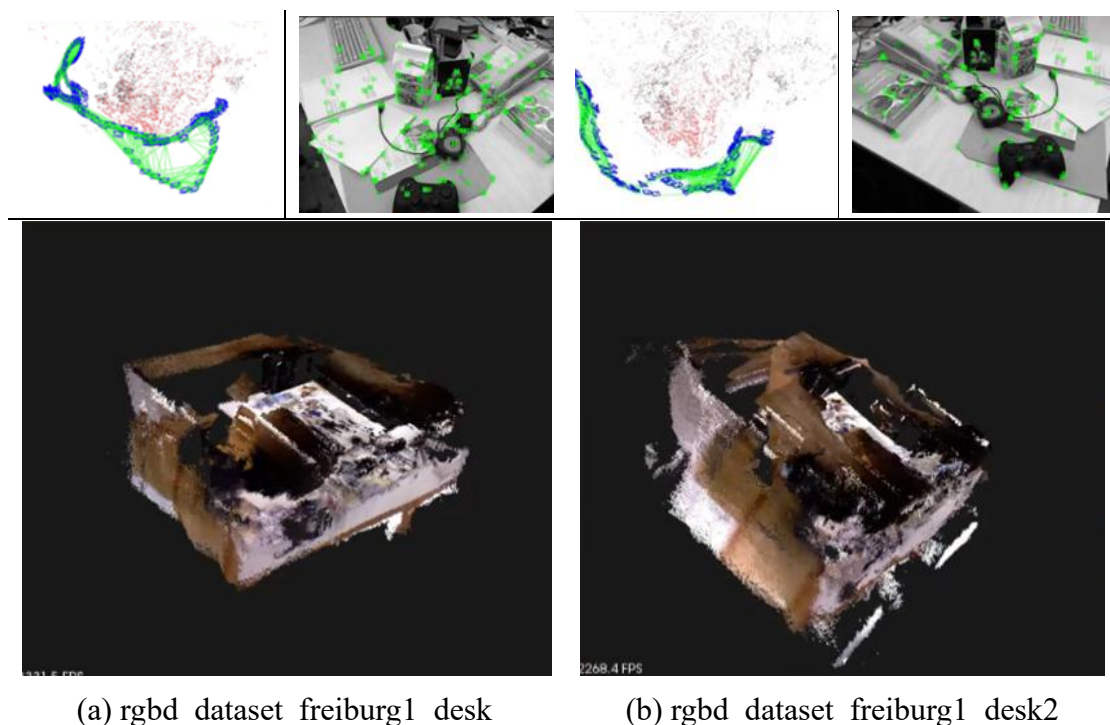
实验结果表明，右图经处理后成功消除了动态物体区域的干扰特征点，同时保持了静态区域特征的完整分布，为后续位姿估计提供了更纯净的输入，有效提升视觉 SLAM 系统在动态环境中的鲁棒性。

3.4.4 点云地图构建

利用改进的 ORB 特征提取与 YOLOv8_RSS 动态目标剔除，结合 RGB-D 数据集生成高质量的三维点云数据。

具体而言，将 RGB 图像输入 YOLOv8_RSS 模型，提取动态目标的类别信息及其对应的二维边界框。对生成的 ORB 特征点进行空间匹配，剔除位于动态目标边界框内的特征点，并保留经过验证的静态特征匹配对及其二维像素坐标。根据深度相机的内参，将二维像素坐标转换为相机坐标系下的归一化图像坐标。结合深度图像提供的深度信息，将归一化坐标转换为相机坐标系下的三维坐标。使用估计的相机位姿将各帧点云变换到全局坐标系，最终合并形成稠密点云地图。

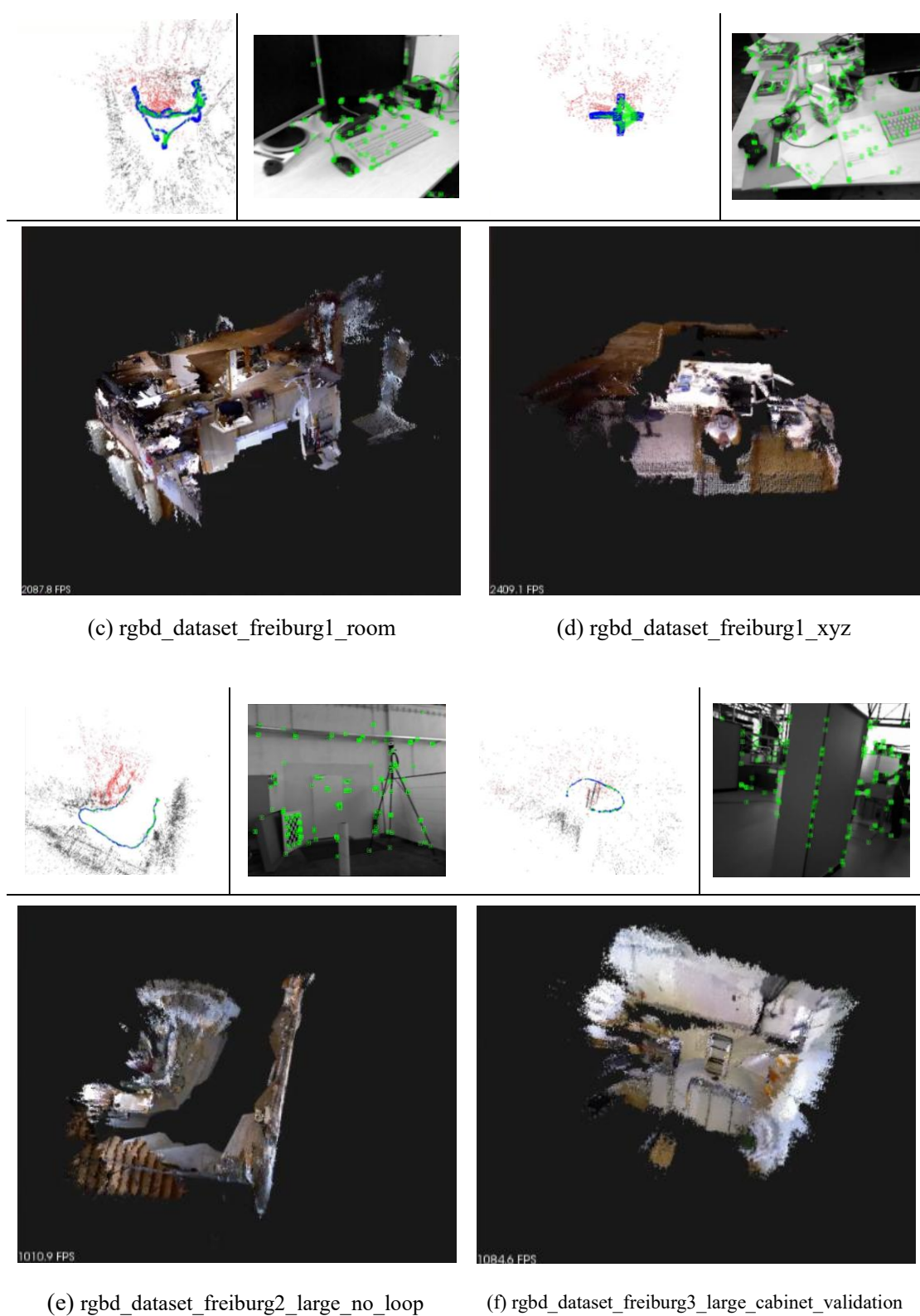
利用 TUM 数据集的场景开展了稀疏点云和稠密点云建图实验。TUM 数据集是一个用于视觉 SLAM 算法测试和开发的 RGB-D 数据集，它由德国慕尼黑工业大学提供，包含彩色图像、深度图和真实轨迹数据。如图 3-18 所示，实验表明语义识别具有高精度，同时实现了语义数据与点云信息的有效整合。



(a) rgbd_dataset_freiburg1_desk

(b) rgbd_dataset_freiburg1_desk2

图 3-18 基于 TUM 数据集的建图实验



续图 3-18 基于 TUM 数据集的建图实验

3.5 本章小结

本文主要对 ORB-SLAM2 系统跟踪线程进行研究，针对动态环境下视觉

SLAM 系统中存在动态目标干扰问题，提出了一种基于四叉树的 ORB 特征点提取方法，旨在提取更稳定和均匀分布的 ORB 特征；并采用基于 YOLOv8_RSS 的动态目标检测，利用语义信息实现动态特征点剔除；随后，采用 LK 光流法将剩余特征点与前一帧进行匹配，以计算稳定的基础矩阵；最后，基于极线约束的几何条件对特征进行筛选，剔除真正的动态特征点。语义信息和几何约束的结合有效地过滤掉了动态特征，提高了视觉 SLAM 系统在动态环境中的定位精度和鲁棒性。在此基础上，基于筛选后的稳定特征点生成稀疏点云地图用于实时定位，再结合深度信息与多帧数据构建稠密点云地图，提供丰富环境细节，为路径规划与导航奠定基础。

第4章 基于生物粒子群优化的无人机路径规划

随着无人机技术的快速发展，特别是在军事、物流、农业、环境监测等领域的广泛应用，无人机路径规划问题成为了无人机系统研究中的一个关键技术。针对无人机在复杂任务环境中的路径规划挑战，本研究提出一种融合生物粒子群优化(Bio-inspired Particle Swarm Optimization, BPSO)算法与强化学习的混合智能规划框架。在全局路径规划中，通过引入随机角度扰动策略重构粒子速度更新机制，有效增强算法在多维解空间中的全局搜索能力；基于动态障碍物场景的实时避障需求，研究进一步设计 Q 学习驱动的局部路径重规划模块，通过强化学习的动态环境感知与自主决策特性实现局部路径的在线优化。由此构建的 BPSO-QL(Bio-inspired Particle Swarm Optimization with Q-Learning)算法，通过群体智能与强化学习的协同架构，兼具全局最优路径生成与动态场景自适应调节的双重优势。相较于传统粒子群优化方法，该算法在路径规划效率、动态避障响应能力及能量消耗优化等方面均展现出显著提升，验证了其在复杂三维空间无人机路径规划中的有效性与鲁棒性。

4.1 粒子群优化基础

4.1.1 粒子群优化

粒子群优化(Particle Swarm Optimization, PSO)是一种基于群体智能的优化方法，其灵感取自鸟群、鱼群等生物群体所展现的社会行为。该算法由 Kennedy 和 Eberhart 于 1995 年提出，广泛应用于函数优化、神经网络训练、工程设计等领域。核心思想在于 PSO 模拟群体中个体的合作与信息共享。每个粒子代表解空间中的一个候选解，通过跟踪个体历史最优解(pbest)和群体全局最优解(gbest)动态调整自身位置和速度，逐步逼近最优解。相较于遗传算法与模拟退火，PSO 无需复杂的交叉、变异操作，且具有天然的并行性。

设在 n 维搜索空间中，粒子总数由 N_{pop} 表示，PSO 算法中每个粒子表示一个候选解，粒子 $i(i \in N_{pop})$ 的属性包括位置向量 $X_i = (x_{i1}, x_{i2}, \dots, x_{iN})$ 、速度向

量 $V_i = (v_{i1}, v_{i2}, \dots, v_{iN})$ 、个体历史最优位置 $pbest_i$ 、群体全局最优位置 $gbset$ 表示。则粒子在第 $t+1$ 次迭代中的速度与位置更新公式为：

$$v_i^{t+1} = w \cdot v_i^t + c_1 r_1 (pbest_i - x_i^t) + c_2 r_2 (gbest - x_i^t) \quad (4-1)$$

$$x_i^{t+1} = x_i^t + v_i^{t+1} \quad (4-2)$$

其中， w 作为惯性权重，用于控制粒子运动惯性； t 作为迭代次数； c_1 和 c_2 作为加速度常数； r_1 和 r_2 作为随机数，取值为 0 到 1，用于引入搜索随机性。

另外，粒子速度和位置还需满足以下范围：

$$v_i^{t+1} = \begin{cases} v_{max}; v \geq v_{max} \\ v_{min}; v \leq v_{min} \end{cases} \quad (4-3)$$

$$x_i^{t+1} = \begin{cases} x_{max}; x \geq x_{max} \\ x_{min}; x \leq x_{min} \end{cases} \quad (4-4)$$

其中， v_{max} 和 v_{min} 表示速度的上限和下限，而 x_{max} 和 x_{min} 表示位置的边界值。

表 4-1 PSO 算法伪代码

算法 1: PSO 算法伪代码

Input: $MaxIt, N_{pop}, d, c_1, c_2, \omega$

Output: G_{best}, F

1: 种群的随机初始化

2: for each $t(t \in MaxIt)$ do,

3: for each $i(i \in N_{pop})$ do,

4: for each $j(j \in d)$ do,

5: 通过公式(4-1)更新速度

6: 通过公式(4-2)应用速度边界

7: 通过公式(4-3)更新位置

8: 通过公式(4-4)应用位置边界

9: end

10: Fitness evaluation ($F(x_i) = J(x_i)$)

11: If $F(x_i(t+1)) < F(pbest_i(t))$ then,

12: $pbest_i(t) = x_i(t+1)$

13: end

续表 4-1 PSO 算法伪代码

 算法 1: PSO 算法流程

```

14:      If  $F(Pbest_i(t)) < F(G_{best})$  then,
15:           $G_{best} = Pbest_i(t)$ 
16:      end
17:  end
18: end
  
```

4.1.2 Vicsek 模型

Vicsek 模型由 Vicsek 等人提出的粒子系统中自序运动的动力学模型，用于解释鸟群、鱼群等生物群体的自组织行为。Vicsek 模型是多智能体系统的基本模型，包括一些关键特征，例如改变邻域、局部交互和动态行为。其核心思想是：个体仅通过局部感知邻居的状态（如速度方向），无需全局信息或中央控制，即可实现群体运动的有序协同。这一特性使其成为研究群体智能和复杂系统的重要工具。粒子在表面上以生物动力相互作用连续移动，并以恒定的绝对速度移动。粒子以与它们在具有生物动机相互作用的正方形区域中的邻居相同的方向旋转。粒子的位置和角度根据可由下列方程得到更新。

$$x_i^{t+1} = x_i^t + v_i^t \Delta t \quad (4-5)$$

$$\theta_i^{t+1} = \langle \theta^t \rangle_r + \Delta \theta \quad (4-6)$$

其中， x_i^{t+1} 为粒子 i 在第 $t+1$ 次迭代中的位置， v_i^t 为该粒子速度， θ_i^{t+1} 表示粒子的角度， $\langle \theta^t \rangle_r$ 表示半径为 r 的圆内粒子的平均方向， $\Delta \theta$ 为来自 $\left[-\frac{\eta}{2}, \frac{\eta}{2}\right]$ 区间的噪声。

4.1.3 Q-learning

Q-learning 是一种无模型强化学习算法，通过 Q 值更新策略优化决策，该方法可以适应移动机器人在工作空间中的行为，以进行路径优化。Q-learning 的目标是找到最优动作价值函数 $Q(s, a)$ ，表示在状态 s 选择动作 a 后能获得的最大累积奖励。可以根据 Bellman 方程更新 Q 值。

$$Q(s_t, a_t) = Q(s_t, a_t) + \alpha \left[r(s_t, a_t) + \gamma \cdot \max_{a \in A} Q(s_{t+1}, a) - Q(s_t, a_t) \right] \quad (4-7)$$

其中, α 表示学习率, 用来控制新信息覆盖旧知识的速率。 $r(s_t, a_t)$ 是在采取行动后的奖励。 γ 是 $(0, 1)$ 之间的折扣因子, 用来评估未来奖励的重要性, 当 $\gamma = 0$ 仅考虑即时奖励, 并且 $\max_{a \in A} Q(s_{t+1}, a)$ 表示下一状态 s_{t+1} 的最大 Q 值。

4.2 改进的粒子群优化算法

在动态环境中, 无人机通常以获得最优路径和实现安全运行的避障为目标。本文基于粒子群优化算法和粒子间生物相互作用的启发, 提出了一种全局路径规划的 BPSO 算法, 在公式(4-1)的基础上引入随机生成的角度, 用来提升搜索多样性提出。更新速度调整后的计算方程如下。

$$v_i^{t+1} = w \cdot v_i^t + c_1 r_1 (pbest_i - x_i^t) + c_2 r_2 (gbest - x_i^t) + (\theta, \varphi) \Delta t \quad (4-8)$$

其中, (θ, φ) 为三维球面坐标系中的随机方向角。

改进后的 BPSO 算法流程如表 4-2。

表 4-2 改进的 PSO 算法伪代码

算法 2: BPSO 算法伪代码

Input:

Particlemax: 粒子数量

Iterationmax: 迭代次数

f : 目标函数

w : 惯性重量

Wdamp: 惯性重量阻尼比

Output:

Global Best: 全局最优路径

$f(x^*)$

1: 种群的随机初始化

2: for particle = 1: *particle_{max}* do

3: particle. θ = $2 \cdot \pi \cdot (\text{rand} - 0.5)$

4: particle. φ = $\pi \cdot (\text{rand} - 0.5)$

续表 4-2 改进的 PSO 算法伪代码

算法 2: BPSO 算法伪代码

```

5:   Initialization of particle.position
6:   particle.velocity = 0
7:   Evaluation of the particle's fitness value  $f(x_i^t)$ 
8:   if  $f(x_i^t) < pbest_i^t$  then
9:       Update Personal Best  $pbest_i^t$ 
10:    if  $pbest_i^t < gbest^t$  then
11:        Update Global Best  $gbest^t$ 
12:    end
13: end
14: end
15: for  $iteration = 1 : iteration_{max}$  do
16:     for  $particle = 1 : particle_{max}$  do
17:         通过公式(4-9)更新粒子速度
18:         通过公式(4-3)应用速度边界
19:         通过公式(4-2)更新粒子位置
20:         通过公式(4-4)应用位置边界
21:         Evaluation of the particle's fitness value  $f(x_i^t)$ 
22:         if  $f(x_i^t) < pbest_i^t$  then
23:             Update Personal Best  $pbest_i^t$ 
24:             if  $pbest_i^t < gbest^t$  then
25:                 Update Global Best  $gbest^t$ 
26:             end
27:         end
28:     end
29:     Update Global Best  $gbest^t$ 
30:      $w \leftarrow w * wdamp$ 
31: end

```

在 BPSO 算法中, 种群的随机初始化包括随机分布的位置和方向角 (θ, φ) , 速度设置为 0。如果速度或位置在搜索空间之外, 则粒子的速度或位置将被限制到最近的边界值。然后, BPSO 算法更新粒子的速度和位置, 然后评估的基础上的目标函数的解决方案。如果新解的适应值小于当前粒子历史最优值 $pbest$, 则更新 $pbest$ 。如果 $pbest$ 小于当前群体全局最优值 $gbest$, 则更新 $gbest$ 。如果全局最优值在 10 次迭代中没有更新, 则将生成的路径作为最优全局路径, 否则, 在迭代结束后生成全局路径。

另外, Q-learning 方法在动态决策方面表现优异, 但其在大规模状态-动作空间中的收敛速度较慢, 而群体智能算法则不存在此问题。本文融合了群体智能算法与强化学习技术的优点, 提出的 BPSO-QL 方法, 该方法首先支持全局路径规划的 BPSO 算法, 然后实现局部路径规划的 Q-learning 方法。在全局路径规划的基础上, 当发现移动障碍物时, BPSO-QL 算法将使用 Q-learning 方法进行局部路径规划, 以避免静态障碍物并达到目标位置。BPSO-QL 算法流程如图 4-1 所示。

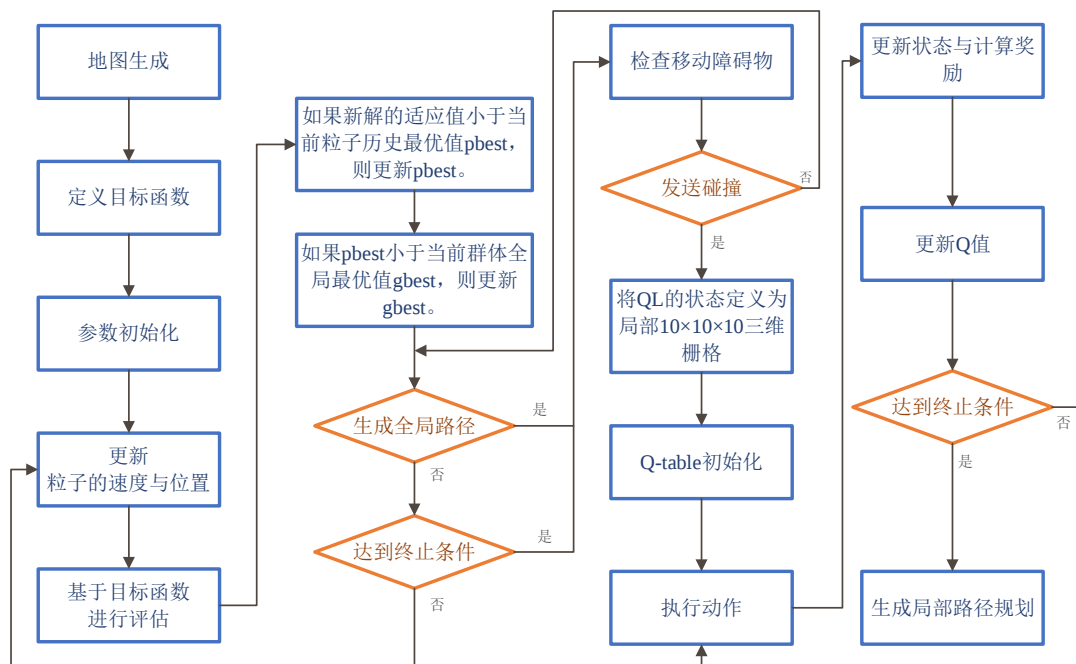


图 4-1 BPSO-QL 算法流程

4.3 算法性能测试实验

4.3.1 标准测试函数

在启发式算法的性能评估体系中，基准测试函数作为标准化分析工具发挥着关键作用，其通过构建具有明确数学特征的问题空间，为算法在搜索能力、收敛特性及鲁棒性等维度的量化比较提供理论依据。国际计算智能领域广泛采用的 CEC(Congress on Evolutionary Computation)基准测试函数集，依据函数形态特征划分为三大类，单模态函数(F1-F7)、多模态函数(F8-F13)和复合函数(F14-F23)。

单模态函数以单峰结构为特征，其变量间存在强相关性，导致函数梯度在远离全局最优区域时呈现误导性。此类特性易使算法陷入局部极值陷阱，尤其在采用梯度依赖策略时，难以突破局部单调性约束。多模态及复合型函数则通过构造多峰拓扑结构模拟复杂优化场景，多模态函数包含指数级增长的局部极值，且相邻极值适应度差异显著，对算法的模态跃迁能力构成挑战；复合型函数通过混合非线性耦合维度与参数自适应机制，进一步强化解空间的崎岖性。这两类函数要求算法具备动态平衡勘探与开发的能力，尤其需克服维度间非线性相关导致的搜索效率衰减问题，并在高维解空间中维持全局收敛性。

4.3.2 BPSO 算法性能对比分析

基于 CEC 测试集的数学特征，本研究将采用平均误差(Mean)与标准差(Std)双指标量化评估 BPSO 算法的性能差异。同时选取四类典型群体智能算法构成对比体系：PSO、NGO、SSA、WOA。PSO 是基于种群的优化算法，通过个体与全局最优的双重引导实现快速收敛，在单模态函数(F1-F7)中易受变量强相关性导致的 Hessian 矩阵非对称性干扰，引发梯度误导性停滞，但其线性收敛特性在低维单峰场景仍具时效优势；NGO 模拟北方苍鹰捕猎行为，采用"侦察-俯冲-攻击"三阶段策略，其动态惯性权重机制可有效应对多模态函数(F8-F13)的非凸地形特征，其阶段搜索强度自适应机制能抑制局部极值的引力效应；SSA 通过麻雀种群的角色分工（发现者/跟随者/警戒者）与 Levy 飞行扰动，平衡全局探索与局部开发，但其维度敏感特性在复合型函数(F14-F23)的高维非线性耦合空

间中易引发搜索轨迹震荡；WOA 以座头鲸气泡网捕食为原型，融合螺旋包围与收缩攻击机制，参数简洁但难以适应复合函数参数空间的动态缩放需求，固定收敛因子在复合问题中易早熟。

在本研究的对比实验中，实验参数设置保持统一性以控制变量：特征维度统一设定为 $d=30$ ，迭代终止条件设置为最大迭代次数 1000 次。为确保实验结果的统计显著性和算法性能评估的客观性，采用重复独立实验设计，针对每个被测算法在基准测试函数集上均实施 30 次独立实验，以充分消除随机因素干扰并获取具有统计学意义的性能数据分布特征。测试结果如表 4-3 所示。

表 4-3 基准函数测试结果

Function	Mean/Std	BPSO	PSO	NGO	SSA	WOA
F1	Mean	4.26E-194	2.77E-06	5.52E-179	1.34E-08	2.18E-153
	Std	3.22E-189	1.15E-05	0.00E+00	2.56E-09	8.60E-153
F2	Mean	6.74E-134	2.000230	1.38E-92	0.808220	5.07E-105
	Std	3.52E-131	4.103811	1.66E-92	0.758183	1.95E-104
F3	Mean	4.21E-163	1385.283236	8.04E-49	223.864539	23277.32678
	Std	5.79E-161	2014.276655	3.59E-48	84.909945	9652.733175
F4	Mean	8.02E-84	3.777294	4.92E-76	7.317847	38.391568
	Std	7.34E-83	0.759146	1.04E-75	4.708529	30.797344
F5	Mean	2.011E-02	4568.166039	23.492749	346.982811	27.260510
	Std	0.027518647	20128.55178	0.616583	733.526626	0.591408
F6	Mean	0.003004818	1.15E-06	8.45E-08	1.20E-08	0.084338
	Std	0.003055587	2.73E-06	7.98E-08	2.98E-09	0.099860
F7	Mean	0.000104063	0.160126522	0.000297205	0.082202	0.001606
	Std	7.24E-05	0.603772522	0.000138076	0.035113	0.001778
F8	Mean	-2.03E+04	-8.73E+03	-7.96E+03	-4.97E+04	-1.17E+04
	Std	1497.603002	512.452175	348.6716801	9515.054033	1248.07598
F9	Mean	0.00E+00	46.91582336	0.00E+00	58.60296631	2.84E-15
	Std	0.00E+00	11.83763476	0.00E+00	15.07506504	1.27E-14
F10	Mean	8.88E-16	0.067184634	6.22E-15	2.173098548	4.80E-15
	Std	0.00E+00	0.299689068	1.82E-15	0.605283783	1.59E-15

续表 4-3 基准函数测试结果

Function	Mean/Std	BPSO	PSO	NGO	SSA	WOA
F11	Mean	0.00E+00	0.012800206	0.00E+00	0.008986259	0.002624252
	Std	0.00E+00	0.016846615	0.00E+00	0.011466047	0.01173601
F12	Mean	8.24E-05	0.06738261	4.90E-09	5.646251799	0.007047647
	Std	0.000159825	0.10781107	4.39E-09	2.578598655	0.006000719
F13	Mean	0.001264	0.013114	0.144373	2.288209	0.214285
	Std	0.003367	0.023788	0.290625	5.373143	0.132915
F14	Mean	0.894542	0.998004	0.942842	0.9346733	1.982288
	Std	1.19E-14	0.00E+00	0.00E+00	1.69E-16	2.243888332
F15	Mean	0.000307487	0.002635126	0.000307486	0.001218247	0.000637698
	Std	1.35E-09	0.006081349	3.61E-15	0.000945029	0.000403053
F16	Mean	-1.031628	-1.031628	-1.031628	-1.031628	-1.031628
	Std	7.26E-13	1.34E-16	2.22E-16	8.85E-15	1.50E-10
F17	Mean	0.397887	0.397887	0.397887	0.397887	0.397888
	Std	2.13E-10	0.00E+00	0.00E+00	5.66E-15	3.44E-07
F18	Mean	3.000000	3.000000	3.000000	3.000000	3.000009831
	Std	5.88E-12	6.60E-16	2.50E-16	9.27E-14	1.80E-05
F19	Mean	-3.862782	-3.862782	-3.862782	-3.862782	-3.859746
	Std	3.46E-09	2.28E-15	2.28E-15	5.62E-14	0.002956027
F20	Mean	-3.302937	-3.255401	-3.321995	-3.246568	-3.219363
	Std	0.045377272	0.078782682	4.56E-16	0.063445963	0.134330586
F21	Mean	-10.153200	-7.025480	-9.153199	-6.371186	-9.107126
	Std	8.23E-08	3.615752159	2.13E-06	2.993992482	2.102079588
F22	Mean	-10.402941	-8.990922	-13.402941	-9.638538	-9.751657
	Std	8.19E-08	2.934037508	5.21E-08	2.35278238	2.027619326
F23	Mean	-10.536410	-6.164682	-8.536410	-8.576921	-8.836338
	Std	7.57E-08	3.769984816	1.63E-15	3.138818792	3.136627127

从测试结果上来看，BPSO 算法在 CEC 测试集上展现出显著的综合性能优势。从收敛精度来看，其在单峰函数，如 F1、F3 中平均误差极低，标准差趋近

于零，证明其全局寻优能力和稳定性高于 PSO、NGO 等对比算法；在多峰及复合函数，如 F8、F22 中，BPSO 的误差均值与理论最优值高度接近，且在复杂问题，如 F21-F23 中表现出强适应力，进一步验证了本文改进策略的有效性。

尽管 NGO 和 SSA 在部分多峰函数，如 F12、F20 中表现较优，但 BPSO 在绝大多数场景下保持较高性能，且其误差与最优结果的差距普遍可控。因此，BPSO 通过粒子间生物相互作用而启发的动态平衡机制，有效解决了传统群体智能算法易陷入局部最优、收敛精度不足等问题。

为深入探究 BPSO 算法的动态收敛特性，本研究绘制了所有对比算法的迭代收敛曲线。如图 4-2 所示，收敛曲线分析涵盖 CEC 测试集全部 23 个基准函数，其中横轴代表迭代次数，纵轴对应目标函数值。通过收敛曲线对比各算法在不同函数空间中的收敛速率与精度稳定性，可直观揭示 BPSO 的改进效果。

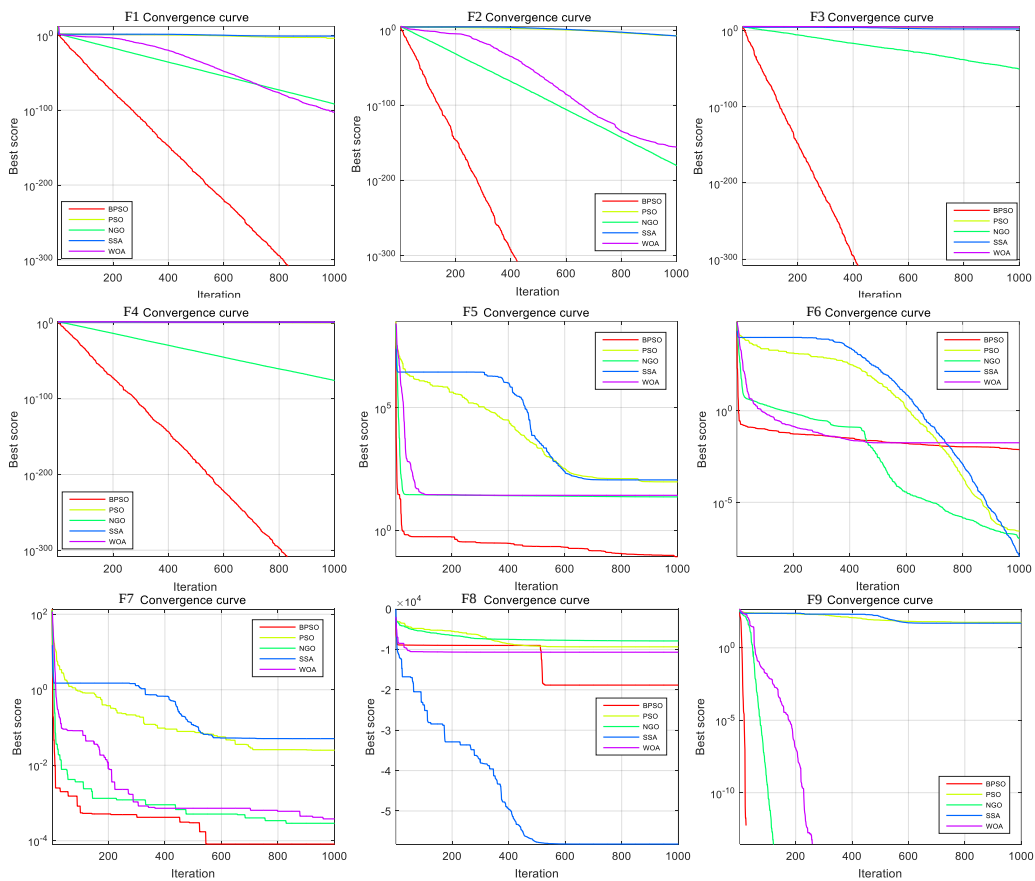
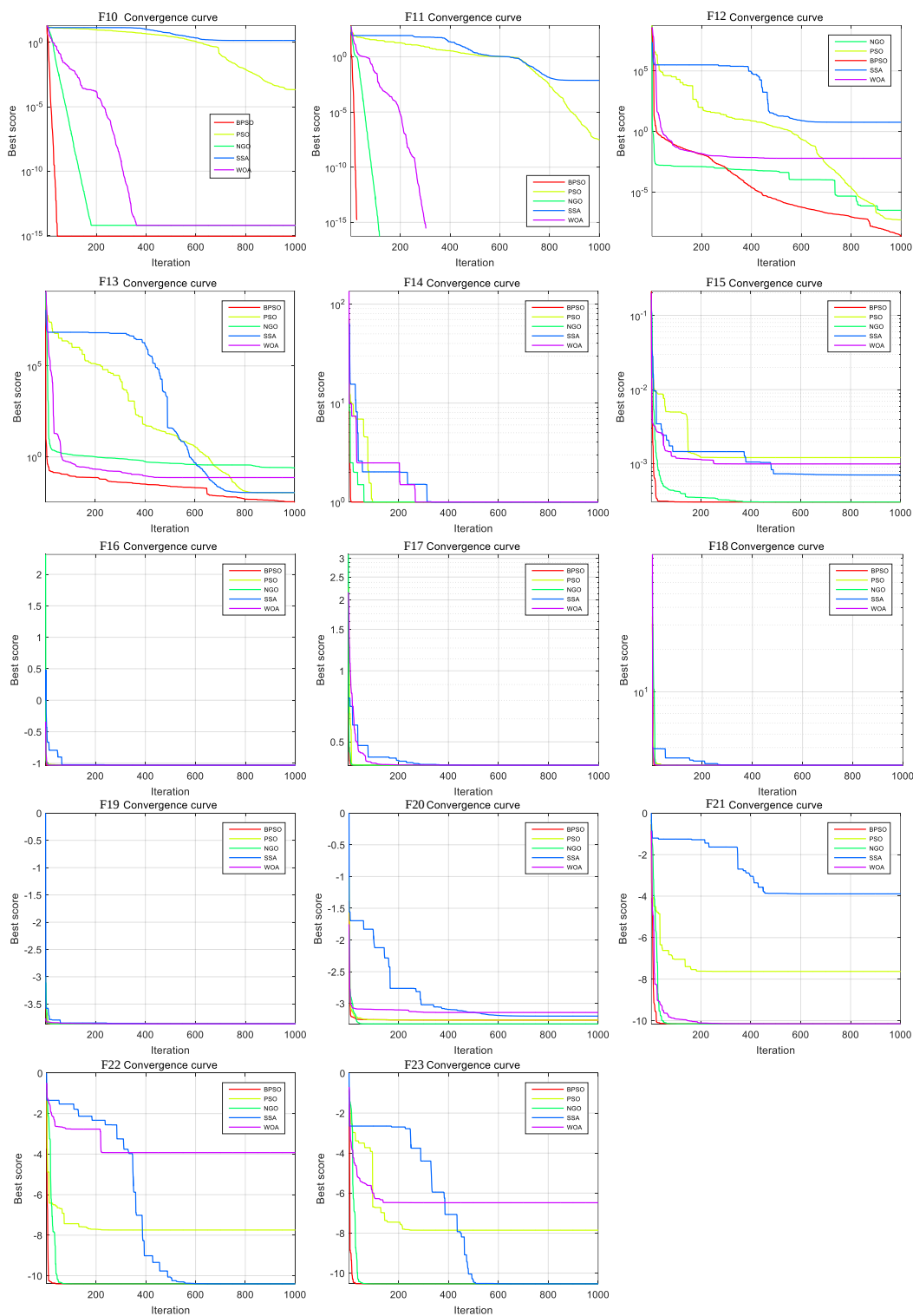


图 4-2 基准函数收敛曲线



续图 4-2 基准函数收敛曲线

通过 23 幅收敛曲线图分析, 实验结果表明 BPSO 在 F1、F2、F3、F4、F8、F10、F13、F17、F18、F21、F23 这些基准函数中展现出显著加速的收敛特性。同时, 在 F1、F2、F3、F4 中, BPSO 在迭代初期即突破 PSO、WOA 对比算法的局部极值停滞现象, 且全程保持近似线性的收敛斜率, 验证了其生物行为启

发的全局搜索机制有效抑制早熟收敛。在 F12 多峰函数中,传统算法 SSA、NGO 在迭代中期陷入局部最优,而 BPSO 持续收敛最后接近最优解。在 F16、F17、F18 中, BPSO 的性能与其他算法相似,但其优势在于能够避免陷入局部最优,从而获得更好的全局搜索能力。

综合来看, BPSO 在收敛曲线上表现出快速逼近全局最优、极低波动性的特点,尤其在复杂多峰中显著优于对比算法。

4.4 路径规划实验对比分析

为系统验证本文提出的 BPSO-QL 算法在无人机三维航迹规划中的有效性和适用性,在保证客观性和公平性的前提下,选取 GWO、PSO、NGO、HHO 四种经典群体智能算法作为基准参照体系,全面验证了本文改进算法的航迹优化性能。分别在复杂密集环境、简单环境和复杂稀疏环境中进行了实验仿真,仿真环境如图 4-3 所示。

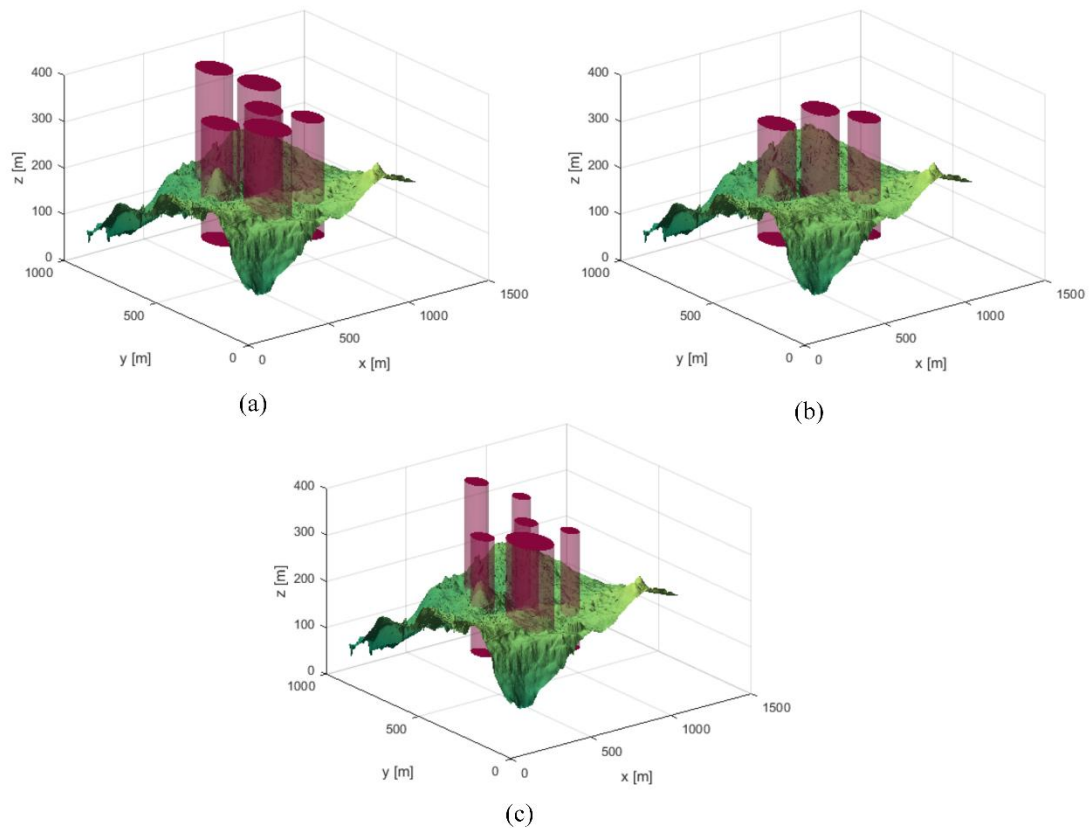


图 4-3 仿真环境

每种算法在每种情况下都进行了 30 次仿真。在实验控制环节,本次实验采

用统一的初始化策略：所有参与对比的算法均采用相同的初始路径生成算法和参数配置，从而保证算法性能对比的公平性。通过设置标准化的评价指标体系，从路径可视化、收敛速度、路径长度等多个维度开展结果分析。

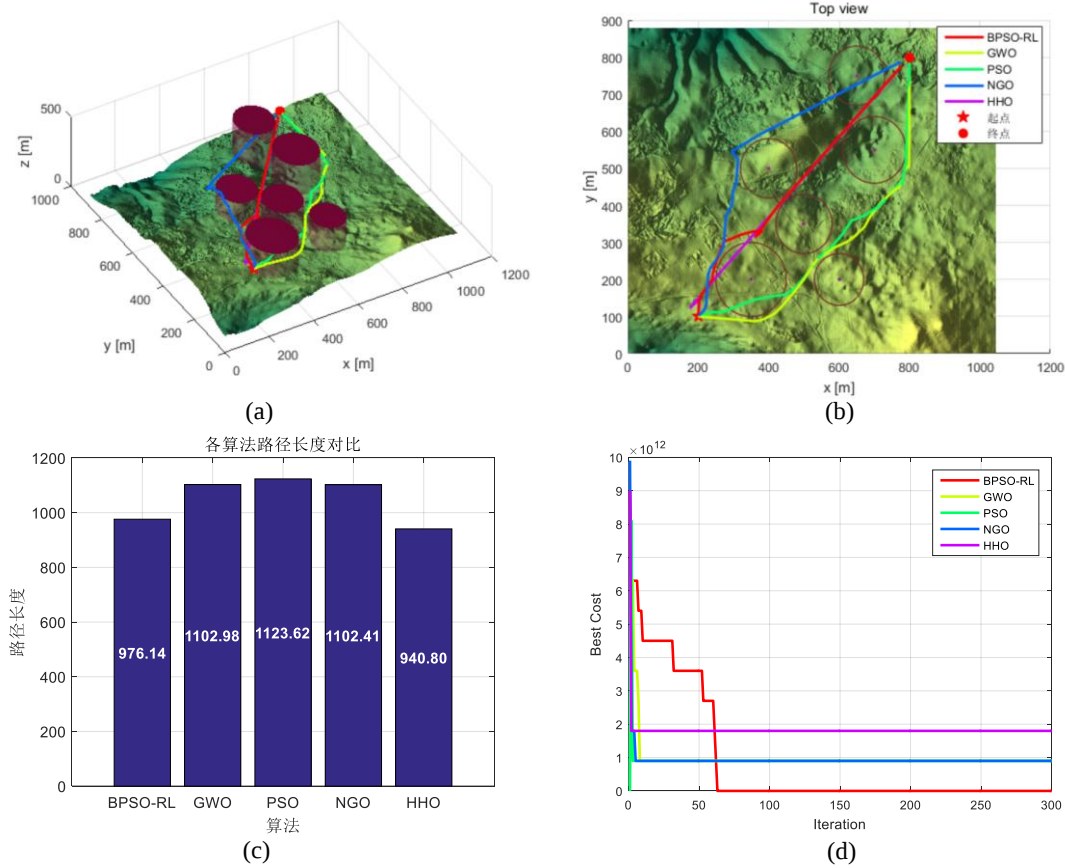


图 4-4 复杂密集环境下的仿真结果图

如图 4-4，五种算法在三维复杂且密集环境中的路径规划结果。在该环境中，障碍物数量设置较多，而且障碍物之间相对紧密。从轨迹结果上可以看出，BPSO-QL 算法的路径规划效果最佳；虽然 HHO 算法的路径规划长度最短，但其规划路径明显穿越了第一个障碍物，其余对比算法则表现出明显的路径冗余与无效绕行特征，因此在轨迹上本文算法占最佳优势。从收敛曲线上可以看出，BPSO-QL 算法在迭代过程中展现出独特的局部最优逃脱机制。尽管在 20-45 次迭代期间多次陷入局部极值，但其通过强化学习引导的粒子群优化策略成功实现全局探索，并于第 60 次迭代收敛至全局最优解，相比之下，其他算法在迭代初期即陷入局部最优且无法跳出。因此，在三维复杂且密集环境中，BPSO-QL 算法性能较高，具有良好的路径规划性能，且拥有较快的收敛性能和较高的鲁棒性。

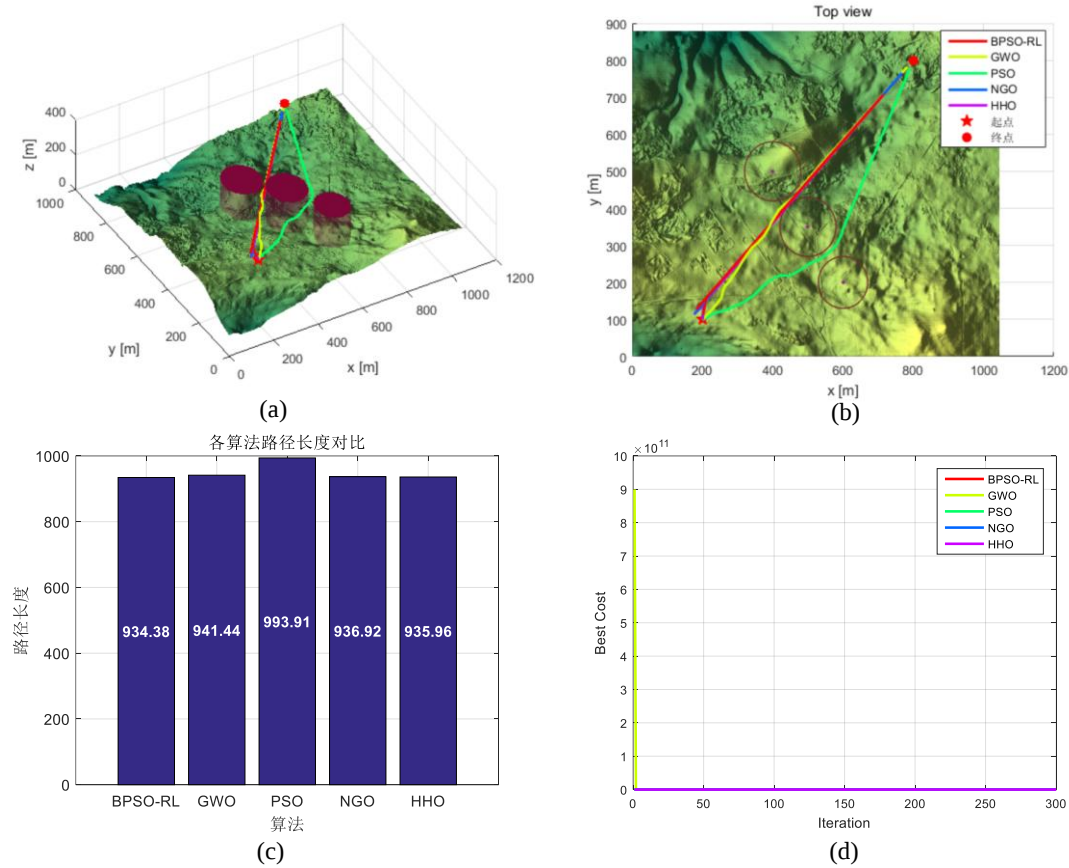


图 4-5 简单环境下的仿真结果图

如图 4-5，五种算法在三维简单环境中的路径规划结果。在该环境中，障碍物数量设置最少。从轨迹结果上可以看出，BPSO-QL 算法表现出了较高的水平。其生成的路径在细节上处理得当，未出现明显的路径冗余现象。尽管在规划初期，该算法有一小段绕行的情况，但迅速进行了自我纠正，确保了路径的有效性和高效性。从收敛曲线上可以看出，这五种算法均展现出了良好的收敛特性。在迭代初期，它们就能够迅速接近并找到全局最优解，这体现了算法的高效性和稳定性。因此，在简单环境下，本文提出的 BPSO-QL 算法以及对比的四种算法均取得了良好的效果，具有较高的收敛速度和较好的规划效果。

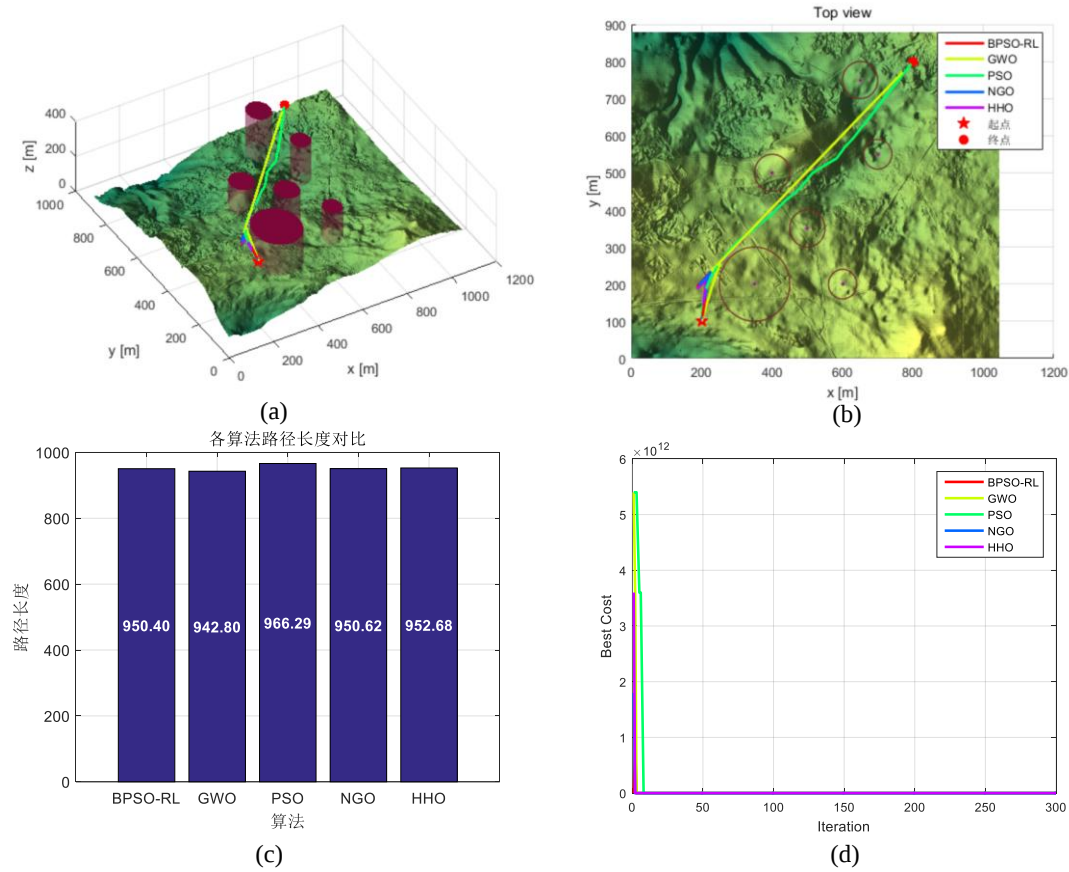


图 4-6 复杂稀疏环境下的仿真结果图

如图 4-6，五种算法在三维复杂稀疏环境中的路径规划结果。在该环境中，障碍物数量设置较多但障碍物之间较为稀疏。从航迹轨迹上来看，PSO 算法效果最差，且中途出现明显拐点，加长了路径长度；BPSO-QL 与其他算法的轨迹基本上相似，可以满足路径规划需求。从收敛曲线上看，PSO 算法的收敛速度明显较慢，在多次迭代后，才接近全局最优解。而 BPSO-QL 算法，包括其余四种算法，则表现出了较快的收敛速度。它们在五次迭代内就能够找到全局最优解。因此，在三维复杂稀疏环境中，BPSO-QL 算法以及其他几种对比算法在路径规划和收敛速度方面均表现出了较好的性能。而 PSO 算法则存在明显的不足，需要进一步改进和优化。

综上所述，基于实验结果的深入分析，可以明确得出以下结论：在局部最优问题、收敛速度以及飞行距离的优化方面，BPSO-QL 算法相较于其他四种算法展现出了显著的优势。具体而言，该算法不仅能够有效避免陷入局部最优解，确保寻优过程的全局性；同时，其收敛速度之快也充分证明了算法的高效性，能够在较短时间内达到理想的优化效果。此外，在飞行距离的优化上，

BPSO-QL 算法同样表现出色，能够为无人机规划出更为短捷的飞行路径。可见，BPSO-QL 算法凭借其出色的性能，可以为无人机设定了安全可行、平滑且高效的飞行路径。

4.5 本章小结

针对无人机在复杂环境中的路径规划问题，本研究提出了一种融合生物粒子群优化与强化学习的混合智能规划框架，旨在提升无人机在三维空间中的导航与避障性能。在全局路径规划方面，通过引入随机角度扰动策略改进粒子速度更新机制，增强了算法在多维解空间中的全局搜索能力，有效克服了传统粒子群优化方法易陷入早熟收敛的局限，从而提升全局路径的优化效果。在局部路径规划中，针对动态障碍物场景的实时避障需求，设计了基于 Q 学习的路径重规划模块。该模块利用强化学习的环境感知与自主决策特性，通过在线学习动态调整路径，实现了局部路径的实时优化。由此构建的 BPSO-QL 算法通过群体智能与强化学习的协同作用，兼具全局最优路径生成和动态场景自适应的双重优势。为验证算法性能，本研究采用基准测试函数集及三种不同规模障碍物场景开展仿真实验。结果显示，与传统粒子群优化方法相比，BPSO-QL 算法在路径规划效率、动态避障响应能力及能量消耗优化等方面均表现出显著优势，充分证明了其在复杂三维环境中无人机路径规划任务中的有效性与鲁棒性。

第5章 无人机自主导航系统实验

无人机自主导航是复杂环境下任务执行的关键技术。基于 RGB-D SLAM 的系统通过融合深度与视觉数据，实现精准的环境感知与路径规划。本章设计并验证一套自主导航系统，涵盖硬件平台搭建、模块化软件框架、相机标定测试及系统实现。选用 Jetson Xavier NX、Intel RealSense D435i、PX4 控制器和 T-motor 2216 电机，结合改进的 ORB-SLAM2 与生物粒子群路径规划算法，优化地图构建与导航性能。实验证明系统在动态环境中具备高鲁棒性与可靠性，为无人机导航技术提供实践参考。

5.1 自主导航系统设计

5.1.1 系统硬件设计

为了验证本文提出的基于 RGB-D SLAM 的无人机自主导航系统的可行性，本文在实际四旋翼无人机上进行部署，并搭建了一个无人机硬件平台，如下图所示。下面将介绍构成这种系统的关键硬件组件。



图 5-1 无人机硬件平台

Jetson Xavier NX 是 NVIDIA 推出的一款超小型却功能强大的 AI 计算平台，专为机器学习和深度学习应用设计，如图 5-2。它搭载 6 核 Carmel ARM 64 位 CPU、384 个 NVIDIA CUDA®核心以及 48 个 Tensor 核心，提供卓越的计算性

能，能够高效处理复杂的视觉 SLAM 任务，使无人机得以实时进行图像处理和环境语义分析，因此成为自主导航系统的核心组件。

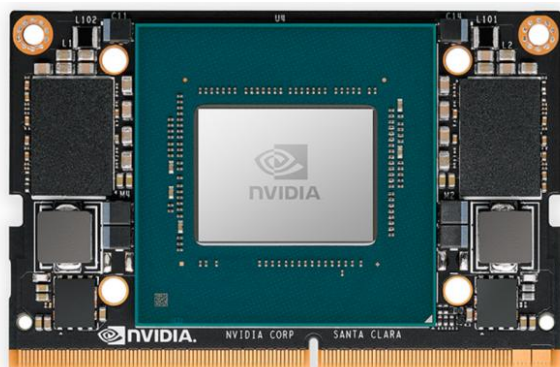


图 5-2 Jetson Xavier NX

Intel RealSense D435i 深度相机是基于 RGB-D SLAM 的自主导航无人机系统的另一个关键组件，如图 5-3。它集成了左右成像器、红外投影仪、颜色传感器，能够在多种照明条件下捕捉高质量的三维环境数据，为无人机提供精准的立体视觉信息。通过测量物体的大小、形状和空间位置，D435i 使无人机能够有效感知和理解周围环境，支持复杂场景下的自主导航和避障功能，与 Jetson Xavier NX 的强大计算能力相辅相成，共同提升系统的整体性能。

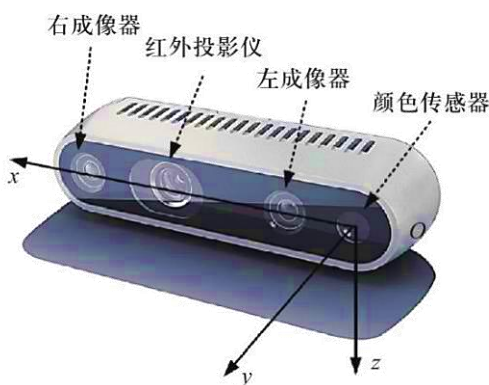


图 5-3 英特尔®实感™深度摄像头 D435i

PPX4 是一款开源无人机飞控系统，可以为无人机提供稳定、灵活且可靠的飞行控制解决方案，如图 5-4。通过精确的飞行操作和高阶导航策略，使无人机能够在动态环境中保持稳定的飞行状态。PX4 飞行控制器与 Jetson Xavier NX 通过高速通信链接紧密协作，能够快速响应环境变化和导航指令，确保无人机在执行任务时具备出色的适应性和实时性，为复杂场景下的自主飞行提供了强有力的支持。



图 5-4 PX4 飞行控制器

T-motor 2216 电机是一款专为多旋翼无人机设计的高性能无刷电机，以其高效、可靠和稳定的动力输出而著称，如图 5-5。其采用先进的无刷技术，能够在各种负载条件下提供强劲的推力，同时保持低能耗和高效率，确保无人机在执行任务时的持久续航能力。该电机的设计经过优化，可以为无人机在运行 VSLAM 算法时提供稳定的飞行平台，从而保证传感器数据采集的精度。搭配适当的电调和螺旋桨，T-motor 2216 能够实现精准的操控响应，使无人机在复杂环境中执行细微的飞行指令，成为自主导航系统中不可或缺的动力支持。



图 5-5 T-motor 2216

以上为本文通过在实际四旋翼无人机上部署并搭建硬件平台，用来验证基于 RGB-D SLAM 的无人机自主导航系统可行性。该系统整合了多个关键硬件组件，确保无人机在复杂环境中实现精准导航和自主飞行。Jetson Xavier NX 作为核心计算平台，凭借其 6 核 CPU 和强大的 CUDA 及 Tensor 核心，为实时图像处理和 VSLAM 任务提供了高效的计算能力。Intel RealSense D435i 深度相机通过深度感测和立体视觉功能，使无人机能够准确感知周围环境的三维信息。PX4 飞行控制器提供了稳定可靠的飞行控制，并通过与 Jetson Xavier NX 的高速通信，快速响应导航需求。T-motor 2216 电机则为无人机提供了高效动力，确保飞行稳定性和操控精度。

5.1.2 系统框架设计

基于 RGB-D SLAM 的无人机自主导航系统的软件框架采用模块化设计，主要包括以下核心模块：

(1) 传感器数据采集模块：负责从 Intel RealSense D435i 获取 RGB 图像和深度数据。

(2) SLAM 处理模块：基于 RGB-D 数据在 Jetson Xavier NX 上运行改进的 ORB-SLAM2 算法，构建环境地图并实现无人机的路径规划，并利用规划结果生成控制指令。

(3) 飞行控制模块：通过 PX4 飞行控制器将导航指令转化为具体的电机控制信号。

(4) 动力模块：可监控 T-motor 2216 电机的状态，确保动力输出的稳定性和效率。

(5) 通信与协调模块：采用 MAVLINK 无人机通信协议和，ROS 机器人操作系统和 MAVROS 通信接口实现各模块间的数据传输和与系统的交互。

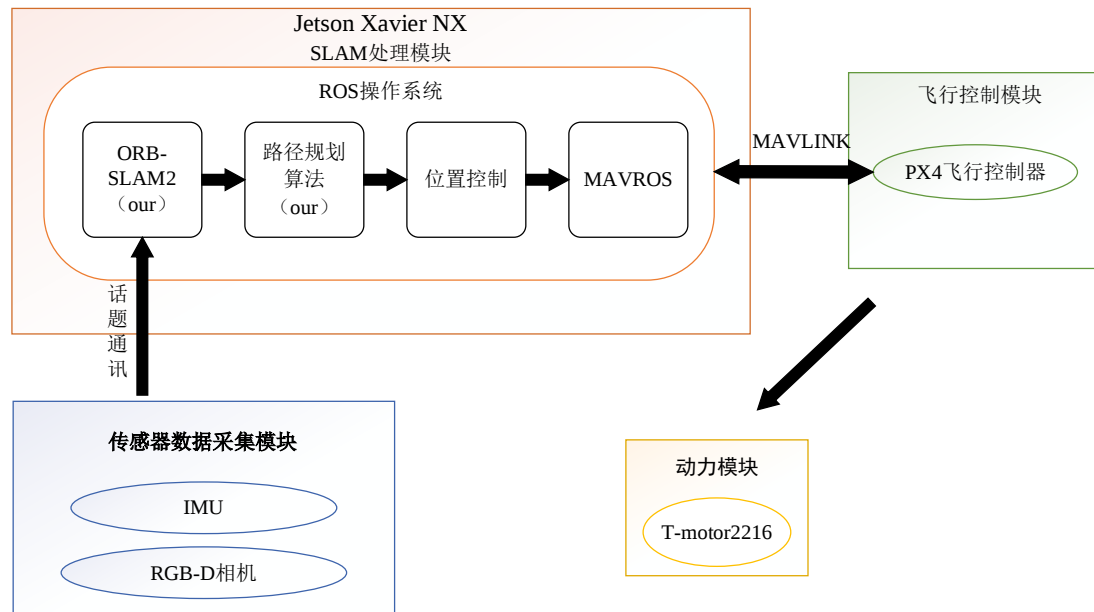


图 5-6 自主导航系统框架结构

如图 5-6 所示，该系统框架充分利用 Jetson Xavier NX 的计算能力、D435i 的感知能力、PX4 的控制精度和 T-motor 2216 的动力稳定性，实现了基于 RGB-D SLAM 的无人机自主导航。通过模块化设计和 ROS 集成，系统具备高效性、可扩展性和实时性，适用于无人机室内外复杂环境下的自主飞行任务。

5.2 相机标定与测试

5.2.1 D435i 深度相机标定

Intel RealSense D435i 深度相机作为多模态传感器，其标定需同时考虑 RGB 相机与红外相机的内参，例如焦距、主点、畸变参数及多相机间的外参。本实验采用 ETH Zurich 开发的 Kalibr 工具箱，基于张正友标定法^[69]实现高精度多相机标定，并通过 ROS 框架完成数据采集与处理。

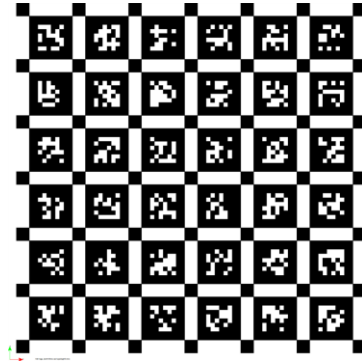
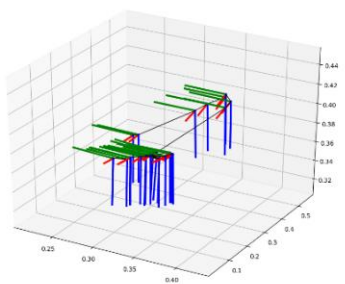


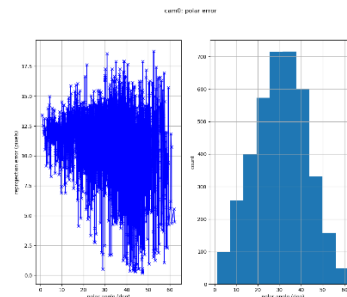
图 5-7 AprilGrid 标定板

如图 5-7 所示，实验采用 AprilGrid 标定板， 6×6 布局，其由均匀分布的 AprilTag 标签组成，每个标签尺寸为 0.08 米，标签间的空白区域间距为 0.024 米。

将标定板垂直固定于均匀光照环境，相机正对标定板并沿三轴执行平移及旋转运动进行录包。经过 Kalibr 工具箱处理后，实验结果如图 5-8 所示。



(a)估计位姿



(b)极坐标误差

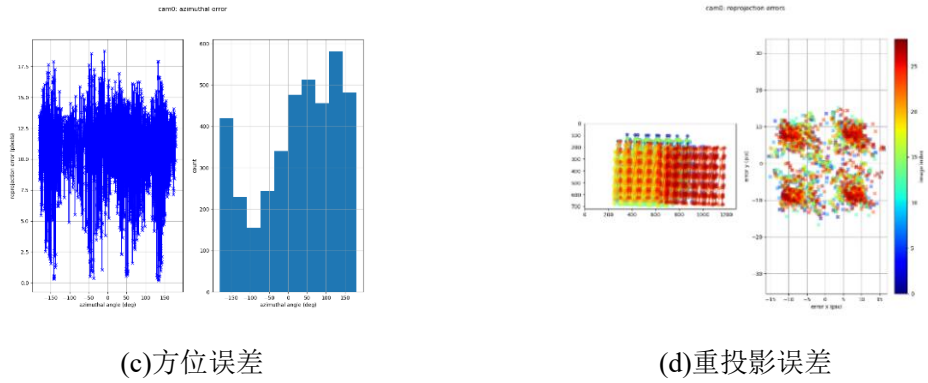


图 5-8 相机标定结果

标定后的相机参数如表 5-1 所示：

表 5-1 相机参数

参数	数值	标准差
相机模型	畸变针孔模型	-
焦距(f_x, f_y)	[538.73, 535.60]	[11.60, 11.49]
主点(c_x, c_y)	[630.47, 339.81]	[1.62, 1.57]
径向畸变(k_1, k_2)	[-0.0742, 0.0041]	[0.0032, 0.0006]
切向畸变(p_1, p_2)	[0.0054, -0.0018]	[0.0002, 0.0002]
重投影误差	[-0.000014, 0.000005]	[8.214691, 8.079933]

从结果上可分析，标定结果精度较高，相机内参与畸变参数估计可靠，可以满足视觉 SLAM 的应用需求。

5.2.2 D435i 深度相机测试

深度相机作为自主导航系统中感知环境的核心组件，其性能直接影响导航算法的精度和可靠性。本节旨在对深度相机的功能进行测试，验证其在不同环境条件下的数据采集能力，并为后续的导航实验奠定基础。测试主要包括深度数据的准确性、视场范围、帧率稳定性以及在不同光照条件下的表现。

实验采用 VMware Workstation 17 Pro 虚拟化环境，主机硬件配置为 Intel Core i7-13650HX 处理器、NVIDIA GeForce RTX 4060 Laptop GPU、16GB 内存，虚拟机分配 4 核 CPU、4GB 内存及 50GB 磁盘空间。虚拟机操作系统为 Ubuntu 18.04 LTS，运行 ROS Melodic 框架，并通过 USB3.2 直通模式接入 Intel RealSense D435i 深度相机。分别在不同环境中进行实验。

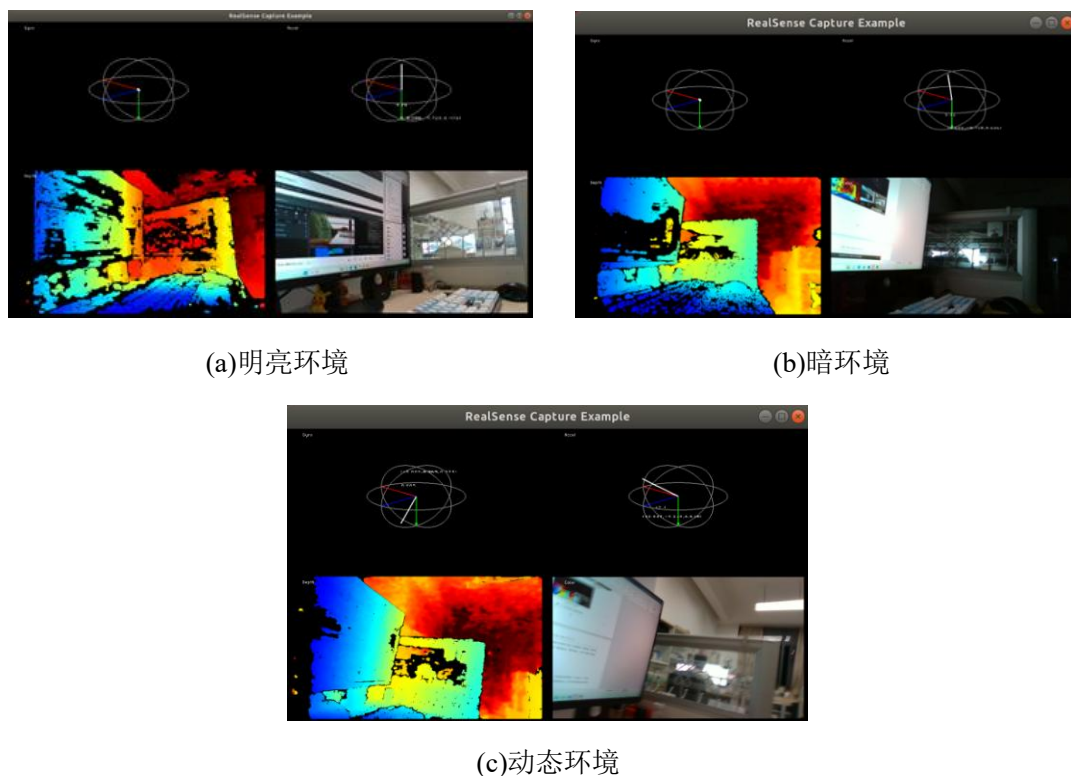


图 5-9 不同环境下的相机测试

实验结果表明，Intel RealSense D435i 深度相机在不同环境下的性能表现存在显著差异。在明亮环境下，该相机展现出最佳性能，深度测量精度高，噪声水平极低，深度图清晰稳定，完全满足无人机室内导航场景对高精度环境感知的需求。在暗环境下，测量精度略有下降，噪声有所增加，但仍能提供可靠的深度数据，表明其在弱光条件下具备一定的适应性。在动态环境下，图像延迟保持较低水平，未出现丢帧现象，显示出良好的实时性和稳定性，可适用于动态导航任务。

5.3 系统部署与验证

5.3.1 目标检测算法部署

本文在第三章提出了基于 YOLOv8 的目标检测算法，改进后算法在检测精度与检测能力上有明显提升。在无人机上完成部署，利用 Intel RealSense D435i 深度相机进行目标检测，检测结果如图 5-10 所示。



图 5-10 实际场景中目标检测

从检测结果上看，本文改进算法在实际应用中表现依旧出色，在检测大小目标时均有较高的识别精度，在面对背景杂乱、强光环境下同时保持着较低的漏检率和误检率。

5.3.2 动态目标特征点剔除

本文在第三章提出了改进的四叉树 ORB 特征提取与 YOLOv8_RSS 算法，在无人机系统中实现部署，动态点剔除效果如图 5-11 所示。



图 5-11 动态特征点剔除

从动态特征点剔除效果来看，改进的四叉树 ORB 特征提取与 YOLOv8_RSS 算法在无人机系统中表现良好。改进的 ORB 特征提取能够快速且准确地识别目标区域中的关键特征点。同时，YOLOv8_RSS 算法的引入进一步优化了动态点剔除过程，尤其是在高速移动或背景变化剧烈的场景中，系统仍能保持较高的检测精度和实时性。

5.3.3 八叉树地图转换

为了优化地图在无人机导航中的应用效率并降低存储需求，本研究采用了基于八叉树结构的概率映射框架 Octomap。传统地图存储方法在处理大规模点云数据时面临存储容量庞大的问题，但 Octomap 通过其独特的压缩机制，显著减少了存储空间占用。具体而言，Octomap 将三维空间按照八叉树结构进行递归分割，直至形成最小的正方形单元。每个节点在这些单元中存储其是否被占用的信息，从而实现对空间的高效表示。另外，当某一区块的所有子节点均处于被占用或未被占用状态时，系统无需进一步扩展该节点，这一机制进一步提升了存储效率。此外，每个小正方形通过 0 到 1 之间的浮点数来量化其被占用的概率，这一设计不仅增强了模型的灵活性，还为环境观测中的不确定性提供了量化依据。然而，在实际应用中，由于传感器噪声或环境中动态物体的干扰，节点的概率信息需要不断更新，以确保地图的准确性和实时性。

计算公式为：

$$P(n|z_{1:T}) = \left(1 + \frac{1-P(n|z_T)}{P(n|z_T)} \frac{1-P(n|z_{1:T-1})}{P(n|z_{1:T-1})} \frac{P(n)}{1+P(n)} \right)^{-1} \quad (5-1)$$

其中 n 代表子节点， z_T 代表在时间 T 的测量值， $P(n|z_{1:T})$ 代表子节点的占用概率。

利用概率的对数形式对上述公式进行转换，以满足概率值 $P \in (0,1)$ ，将公式(5-2)带入公式(5-1)：

$$y = \log \left(\frac{x}{1-x} \right) \quad (5-2)$$

可以得到：

$$L(n|z_{1:t+1}) = L(n|z_{1:t-1}) + L(n|z_t) \quad (5-3)$$

5.4 自主导航实验

为进一步探究改进的 ORB-SLAM2 算法与优化的生物粒子群路径规划算法在无人机自主导航系统中的实际应用价值与性能表现，本文设计和实施无人机的自主导航实验。该实验的主要目标在于对无人机自主导航系统在复杂环境条件下的整体性能与可靠性进行验证，重点关注其在三维环境感知能力、环境建模精度以及路径规划有效性等方面的表现。

本实验选取了具有复杂布局与结构特征的实验室环境作为测试场景，以评估系统在实际应用中的适应性。无人机飞行定高 $2m$ ，实验空间内配置了多元化的室内要素，包括但不限于工作台、显示器、座椅、门、锥桶和电脑等实体元素，其空间分布如图 5-12 所示，大小为 $7m \times 7m \times 4m$ 。该环境构建充分模拟了现实场景的复杂性。



图 5-12 实验场景

实验开始时，操控无人机绕实验环境周围飞行一圈，同时利用无人机上的传感器构建高精度三维语义点云图，点云构建过程如图 5-13 所示。

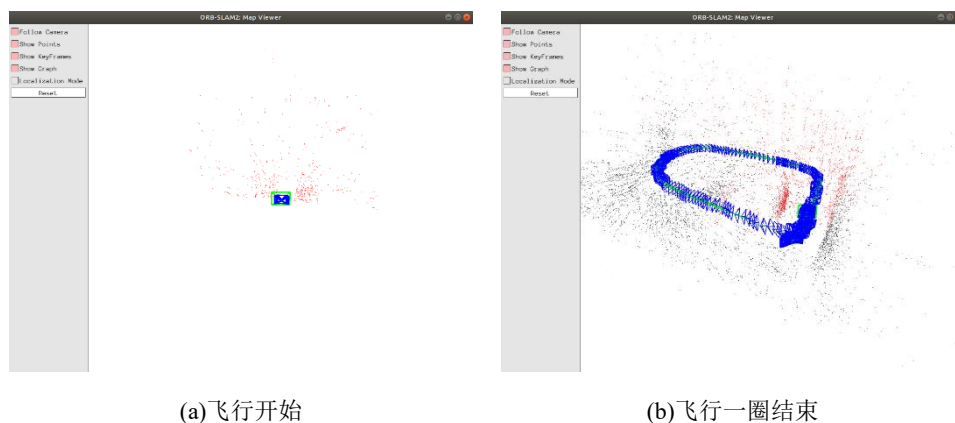


图 5-13 稀疏点云地图构建

随后，通过导入是视觉 SLAM 算法对这些数据进行处理，生成详尽的三维环境模型。这些地图不仅清晰地展现了环境的三维结构，还包含了环境中物体的语义信息，为后续的构建八叉树地图和路径规划提供了精确的环境模型和必要的信息。构建的稠密点云地图如图 5-14 所示。



图 5-14 稠密点云地图构建

在成功生成高精度的三维语义点云和八叉树地图后，无人机系统的自主导航模块被激活，并融合了经过优化的粒子群优化路径规划算法。在为无人机指定明确目标后，路径规划模块随即开始运作，驱动无人机按照算法计算得出的最优航线进行自主飞行。飞行期间，系统持续实时更新环境数据和飞行轨迹，以确保无人机能够快速识别并避开新出现的障碍物，同时以高效且精确的方式到达预定目标点。实验结果如图 5-15 所示。

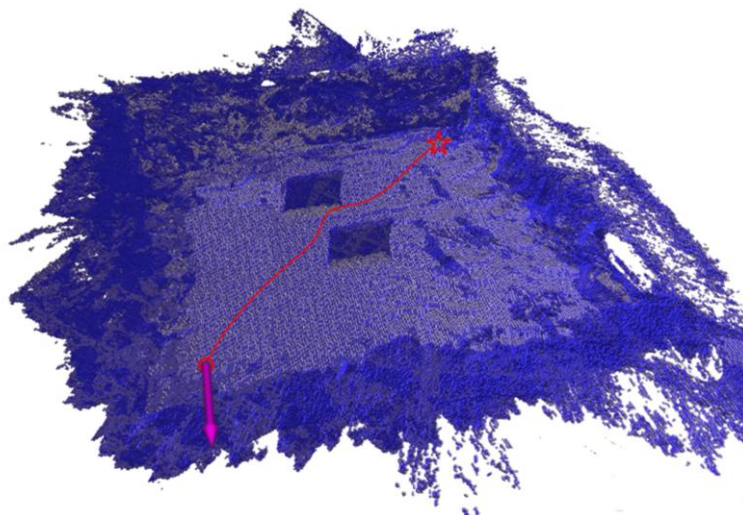


图 5-15 基于八叉树的无人机导航规划

实验结果表明，利用 RGB-D SLAM 技术构建的三维语义点云图和八叉树地图能够精确反映实验环境的三维结构和语义信息，为无人机的路径规划提供了强有力的支持。



图 5-16 无人机自主飞行过程

无人机飞行过程如图 5-16 所示。经过多轮实验测试，无人机能够成功执行自主导航任务，表现出优异的稳定性与出色的适应能力。

5.5 本章小结

通过一系列实验验证了基于 RGB-D SLAM 的无人机自主导航系统的可行性与性能。实验涵盖硬件平台搭建、系统框架设计、相机标定与测试、目标检测、地图构建及自主导航等内容。实验表明，系统能够在复杂动态环境中准确感知

环境三维结构、拥有较高的建图精度，并有效避障并快速到达目标点，展现出极高的鲁棒性与可靠性。

第6章 总结与展望

6.1 全文总结

本研究聚焦于 RGB-D SLAM 在无人机自主导航中的应用，旨在提升无人机在复杂动态环境下的定位精度、鲁棒性及路径规划效率。针对传统视觉 SLAM 在低纹理环境、动态场景干扰以及计算资源受限等问题上的不足，本文提出了一系列创新性改进，主要包括视觉里程计优化、基于生物粒子群优化的路径规划改进，有效提高了无人机自主导航系统在真实环境中的稳定性、精确度和适应性，使其能够在复杂环境中高效执行自主任务。

(1) 改进 ORB 特征提取方法：本文提出了一种基于四叉树的 ORB 特征提取改进算法。在高纹理或低纹理环境下，原始四叉树策略存在特征点过度分裂或分布不均的问题。本文通过计算局部密度标准差，引入自适应分割策略，使特征点在高纹理区域不会过度集中，而在低纹理区域仍然能够保留足够的特征点，保证位姿估计的稳定性。

(2) 融合语义目标检测进行动态特征剔除：为了提高无人机在动态环境中的定位精度，本文引入基于 YOLOv8_RSS 的目标检测算法，在视觉里程计前端对图像进行动态目标检测与特征剔除。此外，为了提高检测效率，本文在 YOLOv8 模型的基础上进行了轻量化优化，引入 RepGhost 模块减少计算开销，并优化了目标检测的损失函数，使其能够更快速地检测小型动态目标。改进的 YOLOv8 实现了 mAP50 增长了 5.29 个百分点，参数数量下降到 8.49M，减少了 76.25%；计算量下降到 2.7GFLOPs，缩减了 9.47%。在此基础上，结合 LK 光流匹配和极线几何约束，进一步剔除了残留的动态特征点，从而显著增强了 SLAM 系统的鲁棒性和稳定性。

(3) 改进的生物粒子群优化的路径规划方法：BPSO 方法结合了粒子群优化算法与粒子间生物相互作用机制，在传统 PSO 的基础上引入随机角度扰动策略，重构粒子速度更新方式，从而提升算法在高维空间中的全局搜索能力，使其能够更快地收敛至最优路径。在三维环境下的路径规划仿真实验表明，相较于现有方法，该算法在达到全局最优解时的平均迭代次数减少 21.5%，路径规

划长度平均降低 4.23%，显著提升了收敛速度与路径优化效果。此外，为满足无人机在动态环境中的实时避障需求，本文进一步结合 Q 学习构建局部路径重规划模块，利用强化学习的自适应特性，使无人机能够根据环境变化动态调整航线，从而实现更加灵活、高效的自主导航。实验结果验证了该方法在复杂动态环境下的优越性，进一步提升了无人机路径规划的鲁棒性与实际应用价值。

6.2 研究展望

本文围绕 RGB-D SLAM 在无人机自主导航中的应用展开，针对传统视觉 SLAM 在动态场景干扰及计算资源受限等问题，提出了一系列创新性改进。然而，尽管本文的方法在提升无人机自主导航能力方面取得了一定进展，仍然存在诸多值得进一步研究和优化的方向。

(1) 提升视觉 SLAM 的鲁棒性：尽管本文采用了基于语义和几何约束的特征点优化策略，提高了动态环境中的稳定性，但在极端低纹理、光照变化剧烈或存在大面积透明和反光表面的环境下，仍然存在特征点提取不足、回环检测失败等问题。未来可以将视觉与 IMU、激光雷达、超声波等传感器进行融合，在特定场景下进一步优化定位精度和提升系统稳健性。

(2) 优化计算资源与提升实时性：本文采用了 YOLOv8_RSS 轻量化目标检测模型，在提高系统在动态环境下的鲁棒性的同时，仍然增加了一定的计算负担。因此，未来研究可借助边缘计算或云计算架构，在保证实时性的前提下提高无人机在复杂环境中的自主决策能力。

(3) 增强路径规划的智能程度和自适应性：本文改进的生物粒子群优化算法虽然提高了全局路径搜索能力，但在更复杂的三维环境中仍存在一定的局限性。未来可以引入强化学习与传统路径规划的结合优化，利用深度强化学习进行端到端的导航决策，提高无人机在多约束场景下的路径规划能力。

无人机自主导航作为低空经济高质量发展的核心技术之一，在国家政策的大力推动下，正迎来多维度、深层次的变革与突破。未来，随着技术、基础设施与政策体系的协同推进，无人机自主导航系统将更安全、更智能，并深度融入城市管理、物流网络、应急救援等经济社会领域，实现“安全高效”的空域生态和“全场景覆盖”的商业价值。

参考文献

- [1] 伍国华, 毛妮, 徐彬杰, 等. 基于自适应大规模邻域搜索算法的多车辆与多无人机协同配送方法[J]. 控制与决策, 2023, 38(01): 201-210.
- [2] Guo S, Li J, Yao W, et al. Optimization of the factors affecting droplet deposition in rice fields by rotary unmanned aerial vehicles (UAVs)[J]. Precision Agriculture, 2021, 22(6): 1918-1935.
- [3] Wang Y, Liu W, Liu J, et al. Cooperative USV-UAV marine search and rescue with visual navigation and reinforcement learning-based control[J]. ISA transactions, 2023, 137: 222-235.
- [4] Stone T C, Davis K J. Using unmanned aerial vehicles to estimate body volume at scale for ecological monitoring[J]. Methods in Ecology and Evolution, 2025, 16(2): 317-331.
- [5] Li J, Xiong Y, She J. UAV path planning for target coverage task in dynamic environment[J]. IEEE Internet of Things Journal, 2023, 10(20): 17734-17745.
- [6] Merzlyakov A, Macenski S. A comparison of modern general-purpose visual SLAM approaches[C]//2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2021: 9190-9197.
- [7] Saputra M R U, Markham A, Trigoni N. Visual SLAM and structure from motion in dynamic environments: A survey[J]. ACM Computing Surveys (CSUR), 2018, 51(2): 1-36.
- [8] Davison A J, Reid I D, Molton N D, et al. MonoSLAM: Real-time single camera SLAM[J]. IEEE Transactions on pattern analysis and machine intelligence, 2007, 29(6): 1052-1067.
- [9] Klein G, Murray D. Parallel tracking and mapping for small AR workspaces[C]//2007 6th IEEE and ACM international symposium on mixed and augmented reality. IEEE, 2007: 225-234.
- [10] Engel J, Schöps T, Cremers D. LSD-SLAM: Large-scale direct monocular SLAM[C]//European conference on computer vision. Cham: Springer

- International Publishing, 2014: 834-849.
- [11]FORSTER C, PIZZOLI M, SCARAMUZZA D. SVO: Fast semidirect monocular visual odometry[C]//Proceedings of the IEEE International Conference on Robotics and Automation. Piscataway: IEEE Press, 2014: 15-22.
- [12]Mur-Artal R, Montiel J M M, Tardos J D. ORB-SLAM: a versatile and accurate monocular SLAM system[J]. IEEE Transactions on robotics, 2015, 31(5): 1147-1163.
- [13]Mur-Artal R, Tardós J D. Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras[J]. IEEE Transactions on robotics, 2017, 33(5): 1255-1262.
- [14]Campos C, Elvira R, Rodríguez J J G, et al. Orb-slam3: an accurate open-source library for visual, visual-inertial, and multimap slam[J]. IEEE Transactions on Robotics, 2021, 37(6): 1874-1890.
- [15]张玮奇, 王嘉, 张琳, 等. SUI-SLAM: 一种面向室内动态环境的融合语义和不确定度的视觉 SLAM 方法[J]. 机器人, 2024, 46(06): 732-742.
- [16]魏彤, 李绪. 动态环境下基于动态区域剔除的双目视觉 SLAM 算法[J]. 机器人, 2020, 42(03): 336-345.
- [17]崔岸, 张新颖, 马耀辉. 复杂环境下基于自适应极线约束的 AGV 视觉 SLAM 算法[J]. 中国惯性技术学报, 2024, 32(03): 234-241.
- [18]刘冬, 于涛, 丛明, 等. 基于深度学习图像特征的动态环境视觉 SLAM 方法[J]. 华中科技大学学报(自然科学版), 2024, 52(06): 156-163.
- [19]Mo J, Islam M J, Sattar J. Fast direct stereo visual SLAM[J]. IEEE Robotics and Automation Letters, 2021, 7(2): 778-785.
- [20]郭烈, 葛平淑, 王肖, 等. 基于卷积神经网络优化回环检测的视觉 SLAM 算法[J]. 西南交通大学学报, 2021, 56(04): 706-712+768.
- [21]蔡显奇, 王晓松, 李玮. 一种室内弱纹理环境下的视觉 SLAM 算法[J]. 机器人, 2024, 46(03): 284-293+304.
- [22]Yang S, Fan G, Bai L, et al. MGC-VSLAM: A meshing-based and geometric constraint VSLAM for dynamic indoor environments[J]. IEEE Access, 2020, 8: 81007-81021.

-
- [23]罗会兰, 陈鸿坤. 基于深度学习的目标检测研究综述[J]. 电子学报, 2020, 48(06): 1230-1239.
 - [24]江波, 屈若锟, 李彦冬, 等. 基于深度学习的无人机航拍目标检测研究综述[J]. 航空学报, 2021, 42(04): 137-151.
 - [25]汪鹏, 辛雪静, 王利琴等. 基于 YOLOv3 的光学遥感图像目标检测算法[J]. 激光与光电子学进展, 2021, 58(20): 509-517.
 - [26]蒋镭圻, 彭月平, 谢文宣等. 嵌入 scSE 模块的改进 YOLOv4 小目标检测算法[J]. 图学学报, 2021, 42(04): 546-555.
 - [27]何宇豪, 易明发, 周先存, 等. 基于改进的 Yolov5 的无人机图像小目标检测[J]. 智能系统学报, 2024, 19(03): 635-645.
 - [28]Wang L, Miao Z, Liu E. UAV remote sensing detection and target recognition based on SCP-YOLO[J]. Neural Computing and Applications, 2024, 36(28): 17495-17510.
 - [29]Zhao H, Zhang H, Zhao Y. Yolov7-sea: Object detection of maritime uav images based on improved yolov7[C]//Proceedings of the IEEE/CVF winter conference on applications of computer vision. 2023. 233-238.
 - [30]Liu C, Wu Y, Liu J, et al. MTI-YOLO: a light-weight and real-time deep neural network for insulator detection in complex aerial images[J]. Energies, 2021, 14(5): 1-19.
 - [31]孙伟, 潘森, 黄恒. 基于改进 YOLOv4 模型的无人机目标检测算法[J]. 传感技术学报, 2023, 36(03): 456-461.
 - [32]Dai Z. Uncertainty-aware accurate insulator fault detection based on an improved YOLOX model[J]. Energy Reports, 2022, 8: 12809-12821.
 - [33]韩俊, 袁小平, 王淮, 等. 基于 YOLOv5s 的无人机密集小目标检测算法[J]. 浙江大学学报(工学版), 2023, 57(06): 1224-1233.
 - [34]Liu H, Wei X, Li J, et al. A star identification algorithm based on simplest general subgraph[J]. Acta Astronautica, 2021, 183: 11-22.
 - [35]Dijkstra E W. A note on two problems in connexion with graphs[M]//Edsger Wybe Dijkstra: His Life, Work, and Legacy. 2022: 287-290.
 - [36]Wang J, Li B, Meng M Q H. Kinematic Constrained Bi-directional RRT with

- Efficient Branch Pruning for robot path planning[J]. *Expert Systems with Applications*, 2021, 170: 114541: 1-7.
- [37]王文涛, 叶晨, 田军. 基于多策略改进人工兔优化算法的三维无人机路径规划方法[J]. *电子学报*, 2024, 52(11): 3780-3797.
- [38]黄鹤, 高永博, 茹锋, 等. 基于自适应黏菌算法优化的无人机三维路径规划[J]. *上海交通大学学报*, 2023, 57(10): 1282-1291.
- [39]尹依依, 王晓芳, 周健. 基于 Q 学习的多无人机协同航迹规划方法[J]. *兵工学报*, 2023, 44(02): 484-495.
- [40]陈洋, 钟树成, 陈志环. 具有最小信息延迟的多无人机路径规划方法[J]. *中国惯性技术学报*, 2024, 32(05): 521-530.
- [41]Niu Y, Yan X, Wang Y, et al. 3D real-time dynamic path planning for UAV based on improved interfered fluid dynamical system and artificial neural network[J]. *Advanced Engineering Informatics*, 2024, 59: 102306: 1-12.
- [42]de Carvalho K B, de OB Batista H, Fagundes-Junior L A, et al. Q-learning global path planning for UAV navigation with pondered priorities[J]. *Intelligent Systems with Applications*, 2025: 200485: 1-8.
- [43]Hu G, Huang F, Shu B, et al. MAHACO: Multi-algorithm hybrid ant colony optimizer for 3D path planning of a group of UAVs[J]. *Information Sciences*, 2025, 694: 121714: 1-39.
- [44]Joo K, Kim P, Hebert M, et al. Linear RGB-D SLAM for structured environments[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 44(11): 8403-8419.
- [45]Taheri H, Xia Z C. SLAM; definition and evolution[J]. *Engineering Applications of Artificial Intelligence*, 2021, 97: 104032: 1-25.
- [46]Harris C, Stephens M. A combined corner and edge detector[C]//*Alvey vision conference*. 1988, 15(50): 10-5244.
- [47]Bay H, Tuytelaars T, Van Gool L. Surf: Speeded up robust features[C]//*Computer Vision—ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006. Proceedings, Part I 9*. Springer Berlin Heidelberg, 2006: 404-417.
- [48]David L. Distinctive image features from scale-invariant keypoints[J].

- International journal of computer vision, 2004, 60: 91-110.
- [49]Rosten E, Drummond T. Machine learning for high-speed corner detection[C]//Computer Vision–ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006. Proceedings, Part I 9. Springer Berlin Heidelberg, 2006: 430-443.
- [50]Rublee E, Rabaud V, Konolige K, et al. ORB: An efficient alternative to SIFT or SURF[C]//2011 International conference on computer vision. Ieee, 2011: 2564-2571.
- [51]Jakubović A, Velagić J. Image feature matching and object detection using brute-force matchers[C]//2018 International Symposium ELMAR. IEEE, 2018: 83-86.
- [52]Malkov Y A, Yashunin D A. Efficient and robust approximate nearest neighbor search using hierarchical navigable small world graphs[J]. IEEE Transactions on pattern analysis and machine intelligence, 2018, 42(4): 824-836.
- [53]Muja M, Lowe D G. Fast approximate nearest neighbors with automatic algorithm configuration[J]. VISAPP (1), 2009, 2(2): 331-340.
- [54]Zach C. Robust bundle adjustment revisited[C]//Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13. Springer International Publishing, 2014: 772-787.
- [55]Kümmerle R, Grisetti G, Strasdat H, et al. g 2 o: A general framework for graph optimization[C]//2011 IEEE International Conference on Robotics and Automation. IEEE, 2011: 3607-3613.
- [56]Li M, Mourikis A I. High-precision, consistent EKF-based visual-inertial odometry[J]. The International Journal of Robotics Research, 2013, 32(6): 690-711.
- [57]Wei W, Cheng Y, He J, et al. A review of small object detection based on deep learning[J]. Neural Computing and Applications, 2024, 36(12): 6283-6303.
- [58]Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2014: 580-587.
- [59]Girshick R. Fast r-cnn[C]//Proceedings of the IEEE international conference on computer vision. 2015: 1440-1448.

-
- [60]Ren S, He K, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on pattern analysis and machine intelligence, 2016, 39(6): 1137-1149.
 - [61]He K, Gkioxari G, Dollár P, et al. Mask r-cnn[C]//Proceedings of the IEEE international conference on computer vision. 2017. 2961-2969.
 - [62]Liu W, Anguelov D, Erhan D, et al. Ssd: Single shot multibox detector[C]//Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14. Springer International Publishing, 2016. 21-37.
 - [63]Zhu X, Lyu S, Wang X, et al. TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios[C]//Proceedings of the IEEE/CVF international conference on computer vision. Montreal:IEEE, 2021. 2778-2788
 - [64]Wang C Y, Bochkovskiy A, Liao H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2023: 7464-7475.
 - [65]Khatib O. Real-time obstacle avoidance for manipulators and mobile robots[J]. The international journal of robotics research, 1986, 5(1): 90-98.
 - [66]Dorigo M, Maniezzo V, Colorni A. Ant system: optimization by a colony of cooperating agents[J]. IEEE Transactions on systems, man, and cybernetics, part b (cybernetics), 1996, 26(1): 29-41.
 - [67]Kennedy J, Eberhart R. Particle swarm optimization[C]//Proceedings of ICNN'95-international conference on neural networks. iee, 1995, 4: 1942-1948.
 - [68]Yao J, Zhang P, Wang Y, et al. An adaptive uniform distribution ORB based on improved quadtree[J]. IEEE Access, 2019, 7: 143471-143478.
 - [69]ZHANG Z. A flexible new technique for camera calibration[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(11): 1330-1334.

攻读硕士学位期间主要研究成果

学术成果

- [1] 王世康, 王坤相, 江自豪等. 基于改进 YOLOv8n 的无人机航拍图像小目标检测[J]. 传感技术学报. (已录)
- [2] 王坤相, 王正刚, 王世康. 基于改进 YOLOv8 的交通标志检测算法[J]. 哈尔滨商业大学学报. (已录)

竞赛成果

以竞赛队长身份获得第十八届全国大学生智能汽车竞赛, 室外 5G 远程驾驶, 全国总决赛一等奖。

以竞赛队长身份获得第十八届全国大学生智能汽车竞赛, 室外 ROS 竞速, 全国总决赛二等奖。

以竞赛队长身份获得第十九届全国大学生智能汽车竞赛, 地平线智慧医疗, 全国初赛二等奖。

致谢

记得研一刚来到 405 实验室时，师兄们就说这三年会很快，如今作为研三即将毕业的我，也会对刚来的师弟师妹们说上这句话。

时光荏苒，又是毕业。过去的近二十载校园岁月里，我要感谢父母的无私付出，包容我，鼓励我，支持我，让我可以在安稳的状态下去做我想做的一切。无忧无虑，甚是有趣。

过去的三年硕士研究生涯里，我首先要感谢我的导师王冠凌老师和校外导师毛德平老师，感谢他们带我接触到智能车竞赛，更要感谢他们在我学术研究、学科竞赛和企业实习上的谆谆教导。其次，我要感谢实验室的同仁们，正因为有了他们，才让这三年充满了欢声笑语和激情澎湃，欢乐是一起在竞赛中夺取总决赛胜利的喜悦，激情是一起在竞赛中过五关斩六将的热血。此外，我还要感谢 418 寝室的兄弟们，感谢他们在我生活中带来的种种乐趣，也感谢他们在我低落时期给予的温暖问候。另外，我还要感谢在考研路上和研究生生涯里与我一起努力的人。最后，由衷感谢在我生活中给我帮助的每个人。这些都是我不断前行的宝贵动力。

学习之途，道远且长，然校园之途，便到此为止吧。从此，笔锋转向未知的章节，墨水将洇染悠长的诗页。愿你历尽千帆，归来仍是少年。

尚德敏学 唯实惟新

