

分类号: TP242. 6

密 级: 公开

学校代码: 10363

学 号: 2220320 120



安徽工程大学

Anhui Polytechnic University

硕士学位论文

题目 低光条件下基于FFE-Net网络
的视觉SLAM算法研究

论文作者 李鹏飞

校内导师 陈孟元（教授）

校外导师 刘会丙（高级工程师）

专业学位类别 电子信息

研究方向（领域） 人工智能

论文提交日期: 2025 年 05 月 31 日

分类号: TP242.6

单位代码: 10363

密 级: 公 开

学 号: 2220320120

题 目 低光条件下基于 FFE-Net 网络的视觉
SLAM 算法研究

题 目 Research on Visual SLAM Algorithm Based
on FFE-Net Network in Low-Light Conditions

学生姓名: 李鹏飞

校内导师: 陈孟元 (教授)

校外导师: 刘会丙 (高级工程师)

专业学位类别: 电子信息

研究方向 (领域): 人工智能与系统集成

论文答辩日期: 2025 年 05 月 12 日

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结果由本人承担。

学位论文作者签名：李鹏飞

日期：2025 年 5 月 30 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或部分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：李鹏飞

指导教师签名：



日期：2025 年 5 月 30 日

日期：2025 年 5 月 30 日

低光条件下基于 FFE-Net 网络的视觉 SLAM 算法研究

摘 要

移动机器人在运动过程中通常需要沿预定路径进行自主导航，其核心是同步定位与地图构建技术（Simultaneous Localization and Mapping, SLAM），机器人在移动过程中通过搭载传感器获取环境信息，实时构建周围环境模型，同时估算自身的位置和姿态，并生成完整地图。根据传感器类型的不同，可分为激光 SLAM 和视觉 SLAM 两大类，其中视觉 SLAM 凭借其价格优势和较强的信息获取能力，逐渐成为研究的重点。然而在实际场景中，由于光照不均匀或照明条件不足以及曝光时间较短、逆光位置等因素的影响，摄像头捕获的图像容易出现亮度低、物体颜色失真和灰度级窄等问题，导致视觉 SLAM 系统特征提取、跟踪和图像匹配困难等一系列问题，严重影响系统在低光环境下的闭环和定位精度，因此如何提高视觉 SLAM 算法在低光环境下的性能，成为研究的关键问题。

针对低光照条件下视觉 SLAM 算法容易出现跟踪丢失，特征匹配困难的问题，提出一种新的融合 FFE-Net（Fourier Feature Enhanced Network, FFE-Net）网络的视觉 SLAM 算法，通过对频域和时域进行联合处理，优化实现低光图像增强，提高了系统在低光照条件下的特征提取与匹配数量。

针对低光照条件下提取的特征信息表征能力不足的问题，提出了一种局部超像素特征描述符的构建方法。通过多帧联合构建时空关联

组来生成描述符的策略，在减少计算量的同时有效提升了算法对角度变化的鲁棒性，使其更适用于 SLAM 系统的运行需求。

针对 SLAM 算法在低光环境下闭环检测线程中存在闭环精度不足、易产生错误回环的问题，提出一种基于特征描述符的闭环检测方法，通过结合全局特征描述符对闭环候选帧进行预剔除，并利用局部特征描述符进行精细匹配，从而实现对候选回环帧的双重筛选，这一改进显著提升了闭环检测的准确率与 SLAM 系统的整体定位精度。

为验证算法的有效性，实验在室外数据集 KITTI 以及室内 TUM_VI 和 EUROC 上的部分低光序列上进行。结果表明，本文算法在低光条件下提升了闭环检测精度与定位性能，在 EUROC 数据集中，当召回率为 100%时，所提算法的闭环检测准确率达到 95.7%，分别比 VINS-FUSION 和 Maplab 高 19.1%和 5.1%，在 TUM_VI 数据集中，分别提升 5.0%和 12.5%。此外，在 TUM_VI 数据集中，所提算法的平均绝对轨迹误差比 ORB-SLAM3、VINS-FUSION 和 Maplab 分别减少 27.16%、56.05%和 47.75%，在 EUROC 数据集中，比 OV2-SLAM 减少 24.17%。为进一步评估算法的实际应用价值，在真实低光场景中进行了测试，验证所提算法的可行性与鲁棒性，实验结果表明，本文算法在低光条件下显著提升了闭环检测精度与系统整体定位性能，为低光视觉 SLAM 算法的研究提供了重要参考依据。

关键词：同步定位与地图构建，低光增强，联合处理，特征描述符，闭环检测

RESEARCH ON VISUAL SLAM ALGORITHM BASED ON FFE-NET NETWORK IN LOW-LIGHT CONDITIONS

ABSTRACT

Mobile robots usually need to navigate autonomously along a predetermined path in the process of movement, the core of which is Simultaneous Localization and Mapping (SLAM), where the robot obtains environmental information by carrying sensors in the process of movement, constructs a model of the surroundings in real time, and estimates its own position and attitude, and generates a complete map. and generates a complete map. According to the different types of sensors, it can be divided into two categories: laser SLAM and vision SLAM, among which vision SLAM has gradually become the focus of research by virtue of its price advantage and strong information acquisition ability. However, in actual scenes, due to uneven or insufficient lighting conditions as well as short exposure time, backlight position and other factors, the images captured by the camera are prone to problems such as low brightness, object color distortion and narrow gray level, which lead to a series of problems such as difficulty in feature extraction, tracking and image matching of the visual SLAM system, and seriously affect the system's closed-loop and localization accuracy in low-light environments, so how

to improve the performance of visual SLAM algorithms in low-light environments has become a key issue for research.

Aiming at the problem that visual SLAM algorithms under low-light conditions are prone to tracking loss and feature matching difficulties, a new visual SLAM algorithm fusing FFE-Net (Fourier Feature Enhanced Network) network is proposed to optimally achieve low-light image enhancement through joint processing of frequency and time domains, which improves the number of features extracted and matched by the system under low-light conditions.

Aiming at the problem of insufficient characterisation ability of the extracted feature information under low-light conditions, a construction method of local super-pixel feature descriptors is proposed. The strategy of generating descriptors by jointly constructing spatio-temporal correlation groups over multiple frames effectively improves the robustness of the algorithm to angle changes while reducing the computational volume, making it more suitable for the operational requirements of SLAM systems.

Aiming at the SLAM algorithm's problem of insufficient closed-loop accuracy and prone to erroneous loopbacks in the closed-loop detection thread under low-light environments, this paper proposes a closed-loop detection method based on feature descriptors, which realizes double screening of candidate loopback frames by combining the global feature

descriptors with the global feature descriptors for pre-culling the closed-loop candidate frames, and using the local feature descriptors to carry out the fine matching, and this improvement significantly enhances the closed-loop detection accuracy and the overall localization accuracy of the SLAM system.

To verify the effectiveness of the algorithm, the experiments are conducted on the outdoor dataset KITTI and some low-light sequences on the indoor TUM_VI and EUROC. The results show that the algorithm in this paper significantly improves the closed-loop detection accuracy and localization performance under low-light conditions, and the proposed algorithm achieves a closed-loop detection accuracy of 95.7% in the EUROC dataset when the recall rate is 100%, which is 19.1% and 5.1% higher than that of VINS-FUSION and Maplab, and 5.0% and 12.0% higher than that of VINS- in the TUM_VI dataset, respectively. FUSION and Maplab by 5.0% and 12.5%, respectively. In addition, the average absolute trajectory error of the proposed algorithm is reduced by 27.16%, 56.05%, and 47.75% in the TUM_VI dataset compared to ORB-SLAM3, VINS-FUSION, and Maplab, respectively, and by 24.17% compared to OV2-SLAM in the EUROC dataset. In order to further evaluate the practical application value of the algorithm, a hardware experimental platform was built and tested in real low-light scenarios to verify the feasibility and robustness of the proposed algorithm. The experimental

results show that the algorithm of this paper significantly improves the closed-loop detection accuracy and the overall positioning performance of the system under the low-light conditions, which provides an important reference for the research of low-light vision SLAM algorithm.

Li Pengfei (Electronic Information)

Supervised by Professor Chen Mengyuan

Keywords: Simultaneous Localization and Mapping, Low-Light Enhancement, Joint Processing, Feature Descriptors, Loop Closure Detection

目 录

第 1 章 绪论	1
1.1 本文研究背景及意义.....	1
1.2 国内外研究现状.....	3
1.2.1 视觉 SLAM 研究现状	3
1.2.2 低光增强算法研究现状.....	4
1.2.3 闭环检测算法研究现状.....	5
1.3 研究内容与章节安排.....	6
第 2 章 视觉 SLAM 与深度学习基础理论.....	9
2.1 视觉 SLAM 理论介绍	9
2.1.1 传感器数据.....	10
2.1.2 视觉里程计.....	11
2.1.3 后端优化.....	12
2.1.4 闭环检测.....	13
2.1.5 建图.....	14
2.2 深度学习理论基础.....	15
2.2.1 卷积神经网络.....	15
2.2.2 注意力机制.....	17
2.2.3 低光增强算法.....	18
2.4 本章小结.....	19
第 3 章 基于 FFE-Net 网络的低光增强算法	20
3.1 亮度增强模块.....	20
3.2 特征细节恢复模块.....	22
3.3 组合损失函数.....	26
3.4 实验结果与分析.....	27
3.4.1 图像低光增强对比试验.....	27
3.4.2 特征提取对比实验.....	28
3.4.3 特征匹配对比实验.....	30
3.5 本章小结.....	31

第 4 章 基于双特征描述符的闭环检测算法	32
4.1 全局特征描述符.....	33
4.2 局部超像素特征描述符.....	34
4.2.1 光谱变换率.....	35
4.2.2 对比特征值.....	35
4.2.3 色彩饱和度.....	37
4.3 双特征描述符闭环候选帧匹配算法.....	38
4.4 公开数据集实验及分析.....	39
4.4.1 准确率召回率实验.....	39
4.4.2 室外数据集对比实验.....	40
4.4.3 室内数据集对比实验.....	41
4.5 真实场景实验.....	43
4.6 本章小结.....	46
第 5 章 总结与展望	47
5.1 全文总结.....	47
5.2 研究展望.....	48
参考文献	49
攻读学位期间发表的学术论文与取得的科研成果	53
致谢	54

第 1 章 绪论

1.1 本文研究背景及意义

机器人技术随着科技的不断进步在工业、医疗、军事、家庭等领域得到了广泛应用。机器人是一种能够自主执行任务的智能化机械设备，具备高度智能化和自主决策能力，能够模拟甚至替代人类的动作和行为。在工业领域，机器人能够高效完成重复性、危险性或繁琐的任务，显著提升生产效率与产品质量；在医疗领域，机器人可实现高精度手术操作，降低手术风险并缩短恢复时间；在军事领域，机器人能够执行高风险任务，减少人员伤亡；在家庭领域，扫地机器人等智能设备为日常生活提供了极大便利。此外，机器人技术还广泛应用于自动驾驶汽车、无人机和航天探测等领域，自动驾驶汽车通过环境感知与自主导航技术，正在重塑未来的交通模式，无人机则利用机器人技术完成航拍、物流配送及农业监测等任务，而在航天探测中，移动机器人如火星探测器能够在极端环境中执行探测任务，为人类探索宇宙提供重要支持，如图 1-1 所示。



图 1-1 移动机器人技术的应用

移动机器人作为机器人领域的重要组成部分，能够在多样化环境中实现自主移动并高效完成任务，其核心功能包括环境感知、自主导航、数据采集与分析以及任务执行等。它可以通过搭载多种传感器如摄像头、激光雷达、超声波传感器、红外传感器等，实时的感知周围环境，并利用智能算法进行数据分析和决策，从而实现高效的执行任务，在其自主移动的过程中，SLAM 技术是实现自主导航的

核心，它使机器人能够在未知环境中通过传感器同步构建环境地图^[1]并实时定位自身位置与姿态^[2]。自 20 世纪 80 年代提出以来，SLAM 技术已在机器人学、无人驾驶、智能交通等领域得到广泛应用，为移动机器人在复杂环境中的自主导航与定位提供了核心支持。

SLAM 技术主要分为激光 SLAM 和视觉 SLAM^[3]两大类。激光 SLAM 依赖激光雷达传感器，具有较高的精度，但其高昂的成本和较大的体积限制了其大规模应用，相比之下，视觉 SLAM 利用摄像头作为传感器，通过图像信息实现定位和地图构建，具有成本低、设备轻便、信息丰富等优势，因此在移动机器人领域得到了广泛应用。视觉 SLAM 的核心是同时解决定位和地图构建两个互相依赖的问题，定位是指确定机器人在已知地图中的位置和姿态，而地图构建是指从头开始构建环境的地图。SLAM 技术通过不断地观测和更新机器人的位置和地图信息，利用概率滤波、优化算法等方式实现对机器人状态的估计和地图的更新。

在实际场景中移动机器人常常需要在低光环境下工作，例如夜间巡检、地下停车场或光线不足的室内环境，在这些场景中，由于光线不足，摄像头捕获的图像质量较差，存在亮度低、对比度低、颜色失真和灰度级窄等一系列问题，导致 SLAM 系统在低光区域提取不到特征点，在跟踪和图像匹配等方面出现困难，从而严重影响机器人的定位精度和导航性能，如图 1-2 所示。



(a) 低光示例图像 a 特征提取

(b) 低光示例图像 b 特征提取

图 1-2 低光图像特征点提取效果

针对低光环境下视觉 SLAM 系统^{[4][5][6]}性能问题，本文提出了一种基于深度学习增强的低光视觉 SLAM 算法，该算法通过引入深度卷积神经网络对低光图像进行特征增强，提升了图像质量与特征点提取的可靠性，同时在后端线程中采用一种新的特征描述符匹配算法以适应低光条件下的闭环检测，提高机器人在低光环境中的导航与定位能力，实现低光照条件下的视觉 SLAM 轨迹重建，显著提升移动机器人在低光环境下的性能。

1.2 国内外研究现状

1.2.1 视觉 SLAM 研究现状

SLAM 技术是一种通过传感器数据实时获取环境三维地图信息并确定设备位置的关键技术。根据传感器类型、地图表示方法以及定位技术的差异,SLAM 技术可以分为多种实现方式,其中,激光 SLAM 依赖于激光雷达获取三维点云数据,尽管其精度较高,但由于成本昂贵,限制了其广泛应用,相比之下,视觉 SLAM 利用视觉传感器捕捉环境信息,能够提供更为丰富的场景细节,因此在自动驾驶、机器人导航以及室内定位等领域得到了广泛应用。

视觉 SLAM 的实现方法主要分为直接法和特征点法两大类。直接法通过分析图像像素的亮度信息来估计相机运动,无需提取特征点,具有较高的鲁棒性,然而这种方法对相机运动变化和灰度一致性要求较高,根据处理像素数量的不同,直接法可以进一步细分为稀疏直接法、稠密直接法和半稠密直接法,稀疏直接法仅处理少量关键点,计算效率高,适用于实时场景,稠密直接法则处理所有像素,能够生成高精度的稠密地图,但计算成本较大,半稠密直接法则在计算效率和地图精度之间取得了平衡。近年来,直接法视觉 SLAM 取得了显著进展,例如 Newcombe 等人提出的 DTAM 系统实现了稠密三维地图的构建,Jakob Engel 等人开发的 LSD-SLAM 系统则实现了半稠密场景的重建,而 Forster 等人提出的 SVO 算法更是展示了稀疏直接法在实时性和效率方面的优势。

特征点法通过提取和匹配图像特征点实现定位与建图,具有尺度不变性,是目前的主流方法。2007 年 Davision 等人提出的 Mono-SLAM 系统首次实现了单目相机的实时定位与建图,但存在计算复杂度高的问题,同年 Gerorg Klein 等人提出的 PTAM 系统引入了前后端分离框架,显著提高了计算效率。2015 年, Mur-Artal 等人提出的 ORB-SLAM 系统采用多线程并行架构,进一步提升了系统性能,随后 ORB-SLAM2 和 ORB-SLAM3 相继推出,分别增加了对双目和 RGB-D 相机的支持,引入了惯性测量单元,实现了视觉与惯性的融合。

基于视觉的 SLAM 算法主要分为单一点特征和多特征两类。Mur-Artal 等人提出的 ORB-SLAM2 是经典的单特征点算法,但在稀疏纹理场景下易出现跟踪丢失问题,Campos 等人提出的 ORB-SLAM3 通过最大后验估计提高了系统鲁棒

性,但仍存在局限性,为应对这一问题,研究者们提出了多特征提取方法,如 Gomez-Ojeda 等人开发的 PL-SLAM 算法结合了点特征和线段特征,在稀疏纹理场景下表现出色, Li 等人则提出了基于点、线和面的解耦-精化方法,进一步提高了跟踪和映射精度。为减少动态物体的影响, Zhang 等人结合语义分割网络与 K-均值聚类算法,提出了动态特征检测和过滤策略,有效解决了低动态稀疏场景下的动态物体问题。

随着技术的不断发展,视觉 SLAM 在移动机器人、增强现实等领域得到了广泛应用,其研究重点也逐渐从单一传感器向多传感器融合方向发展,以提高系统在复杂环境下的鲁棒性和精度。

1.2.2 低光增强算法研究现状

传统视觉 SLAM 系统在理想光照条件下能够实现较高的定位与建图精度,然而在实际应用场景中,由于光照分布不均、照明条件不足(如地下矿井、深空探测及夜间户外环境)以及曝光时间不足、逆光等因素的影响,往往会导致摄像头捕获的图像质量显著下降,导致采集的图像亮度不足、对比度降低、物体颜色失真以及灰度级范围狭窄等一系列问题,这些缺陷直接导致 SLAM 系统在特征提取、目标跟踪和图像匹配等关键环节面临巨大挑战,进而严重制约了系统在低光环境下的闭环检测精度和定位性能。

现有的低光图像增强方法大致可以分为传统方法和深度学习方法。传统图像增强方法如直方图均衡化是一种简单但有效的图像增强方法,优点是简单易实现,无需额外的参数就可以增加图像的对比度,特别是在局部对比度较低的情况下效果显著,缺点是可能导致图像噪声增加,使图像看起来不自然,对于一些图像,可能会导致细节丢失。伽马校正也是一种传统的图像增强方法,该方法采用数学函数来调整图像亮度,对灰度值进行非线性操作,然而此方法中的参数必须根据场景进行调整,固定值不适用于所有图像,可能会导致颜色失真。另外一种是基于深度学习的图像增强算法,自深度学习发展以来,在视觉跟踪、实例分割和超分辨率重建等多个领域都取得了重大进展, Lore 等人^[7]首次将深度学习应用于低照度图像增强,并提出了 LLNet,通过堆叠稀疏自编码器实现图像增强和去噪。Wu 等人^[8]提出了一种多分支低照度增强网络,该方法旨在利用全卷积网络

提取不同层次的丰富特征，通过多个子网络对其进行增强，最后通过多分支融合生成增强图像。Wei 等人基于 Retinex 理论提出了一种 Retinex 网络模型，该模型由一个分解网络和一个用于光线调整的增强网络组成，在分解的基础上，可以增强图像的亮度，同时抑制噪声。Kuang 等人^[9]提出了一种两阶段的特征细节恢复网络，在细节处理方面取得了较好成果，但在局部区域亮度增强不足。大多数现有的低光图像增强方法都是基于时域信息而很少考虑傅里叶频域信息。最近，Wang 等人^[10]发现低光图像的傅里叶频率信息与图像亮度之间的正相关关系，并证明了通过放大低光图像的振幅分量可以改善低光条件下的图像亮度，但是在特征细节恢复方面仍有一些不足，这样的图像虽然能满足人类视觉感知的需求，但并不一定符合 SLAM 系统的特征提取要求，因此提出一种新的低光增强算法仍然是一个有待研究的课题。

1.2.3 闭环检测算法研究现状

闭环检测^{[11][12]}是 SLAM 系统中的关键部分，对于消除机器人在长时间、大范围导航定位过程中产生的累积误差极为重要。随着 SLAM 技术的不断发展，闭环检测算法也在不断演进，目前闭环检测的研究主要集中在基于深度学习的方法和基于外观的方法两大类。

近年来，深度学习技术在图像处理和特征提取领域取得了显著进展，其中基于深度学习的闭环检测方法通过卷积神经网络自动学习图像特征，避免了传统方法中手工设计特征的局限性，可以在处理光照变化时表现出较强的鲁棒性。2015 年，Gao 等人^[13]提出了基于自编码器的闭环检测方法，通过无监督的方式训练深度卷积神经网络，有效地提取图像特征，解决了传统方法中特征提取耗时的问题，自编码器能够学习到图像的低维表示，从而提高了闭环检测的效率和准确性。2016 年，Xia 等人提出了基于 PCA-Net 的深度网络用于闭环验证，该方法的优点是网络结构简单，易于实现，但在处理大尺寸、多场景数据时，准确率有所下降，PCA-Net 通过主成分分析 PCA 降维处理特征向量，进一步提高了闭环检测的效率。2017 年，Zhang 等人提出了一种基于预训练 Overfeat 网络的特征提取方法，通过 PCA 降维处理特征向量，用于相似性计算。同年，Gao 等人提出了基于层叠降噪自编码器 SDA 的无监督学习方法，通过与 FAB-MAP 2.0 的性能对

比,验证了该算法在闭环检测中的有效性,SDA 通过自动学习数据生成表示,进一步提高了闭环检测的鲁棒性。尽管基于深度学习的方法在闭环检测中取得了显著进展,然而在实际应用中面临多方面的挑战,首先这类模型通常需要大量标注数据进行训练,而在实际应用中,获取大规模标注数据较为困难,其次,由于深度学习模型的计算复杂度较高,在大规模 SLAM 系统中,实时性要求较高的情况下,其推理速度可能难以满足实际需求,此外,如何在低光照等恶劣环境下保持较高的检测精度,仍然是闭环检测领域有待解决的问题。

基于外观的闭环检测方法是目前较为成熟的技术路线之一,其核心思想是通过比较当前帧与历史关键帧的外观相似性来判断是否发生了闭环,这种方法避免了直接对所有关键帧进行逐一比对的高计算复杂度问题,从而提高了系统的实时性。词袋模型^{[14][15]} (Bag-of-Words, BOW) 是视觉闭环检测中最常用的方法之一,它通过将图像中的特征点聚类为“视觉单词”,并利用这些单词的组合来表示图像,从而实现高效的图像检索。FAB-MAP^[16]是词袋模型中的经典算法,它通过树形结构学习离线单词,并使用概率模型来捕捉高频出现的视觉单词组合,从而提高闭环检测的准确性,然而 FAB-MAP 依赖于 SURF 特征提取,导致特征提取阶段耗时较长,为了解决这一问题,DBOW2^{[17][18]}采用了基于 BRIEF 描述符的二进制词袋模型,显著提高了特征提取和匹配的效率,成为当前基于外观的闭环检测的主流框架之一,然而词袋模型在低光照等恶劣环境下的性能表现显著下降,这严重制约了其在复杂场景中的实际应用,因此开发一种能够在低光条件下保持稳定性能的闭环检测方法,构建具有低光适应性的 SLAM 闭环检测系统,对于提升视觉 SLAM 在各类复杂环境中的鲁棒性和实用性具有重要的理论意义和应用价值。

1.3 研究内容与章节安排

低光环境对视觉 SLAM 系统的鲁棒性产生了显著影响,导致系统在低光条件下存在特征提取与匹配困难、闭环精度不足等问题。针对低光场景下移动机器人的定位问题,本文提出了一种结合低光增强网络的视觉 SLAM 方法,在前端结合了傅里叶特征增强网络 (Fourier Feature Enhancement Network, FFE-Net),其设计受傅里叶变换的启发,通过振幅变换图对低光图像的振幅分量进行放大以提

升亮度，同时采用编解码器网络提取多尺度特征，并利用单尺度网络恢复细节信息，通过融合两种网络的优势，将采集到的低光数据进行 FFE-Net 网络处理，可以获得在低光条件下具有良好特征提取效果的输出数据。

处理后的数据与 ORB-SLAM3 系统的前端相融合，在完成正常 SLAM 跟踪线程的同时，根据图像生成特征描述符，通过全局特征描述符对闭环候选帧进行预剔除，并利用局部特征描述符进行精细匹配，从而实现对候选回环帧的双重筛选，这一过程与传统词袋查询相结合，共同完成闭环检测线程，进而构建全局地图。基于上述改进，论文提出了一种更适用于低光条件下的视觉 SLAM 算法，简称为 FFE-SLAM，总体论文技术框架如图 1-3 所示。

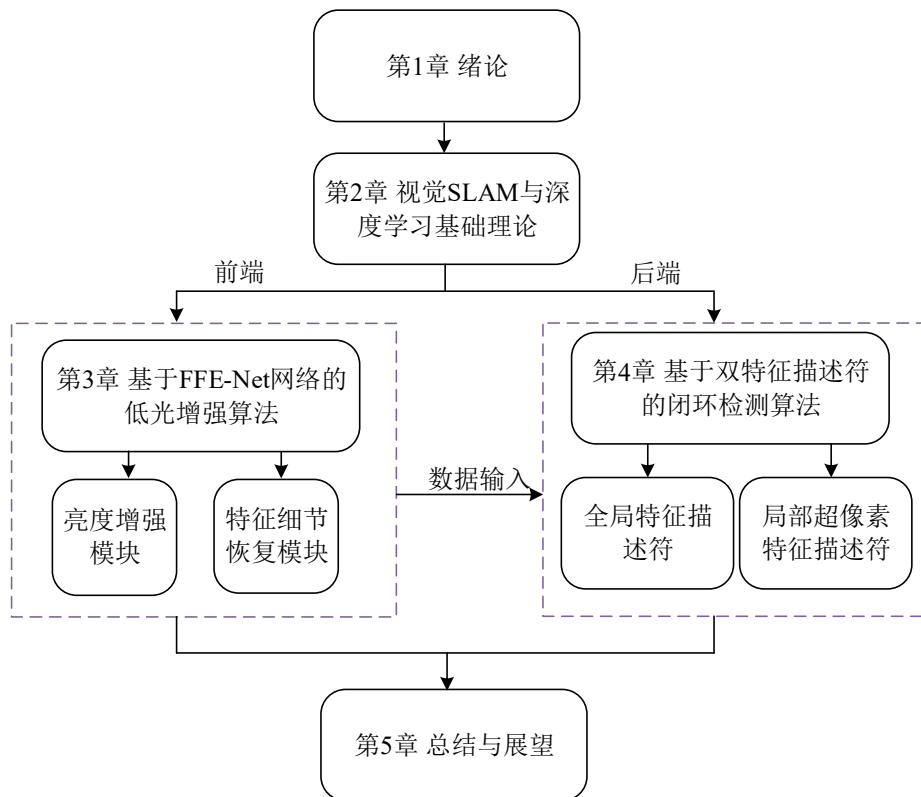


图 1-3 论文总体框架

各章节内容安排如下：

第 1 章，绪论。本章从机器人技术的背景出发，阐述了其在工业、医疗、军事等领域的广泛应用及其重要性，重点分析了移动机器人在环境感知、自主导航和任务执行中的核心功能，并深入探讨了 SLAM 技术的关键作用及其在低光环境下面临的挑战，随后系统综述了低光增强算法和闭环检测算法的国内外研究现状，详细探讨了低光增强方法在提升图像质量与特征提取能力方面的优势，以及闭环检测技术在消除 SLAM 系统累积误差中的重要作用，最后，对本文的研究

内容进行了整体概述，并简要介绍了后续章节的结构安排。

第 2 章，视觉 SLAM 与深度学习基础理论。本章系统阐述了视觉 SLAM 与深度学习的核心概念及其理论基础。首先，从视觉 SLAM 的整体框架出发，详细剖析了其关键组成部分，包括传感器数据采集、视觉里程计、后端优化、闭环检测以及地图构建等五个核心模块，随后深入探讨了深度学习的基础理论，并重点分析了基于深度学习理论构建的低光增强算法，这些内容为深度学习与视觉 SLAM 的深度融合奠定了坚实的理论基础。

第 3 章，基于 FFE-Net 网络的低光增强算法。针对低光照条件下视觉 SLAM 算法存在特征点提取与匹配困难等问题，本文提出了一种基于 FFE-Net 的低光增强算法，通过傅里叶变换对频域与时域进行联合优化，实现低光图像增强。采用公开低光数据集以及 SLAM 领域 KITTI、TUM_VI 和 Euroc 数据集中的部分低光序列分别进行实验验证，结果表明，该算法显著提升了低光环境下的特征提取与匹配性能，降低了低光对 SLAM 系统的影响，增强了机器人在低光照下的感知能力，为低光环境下的自主导航提供了技术支持。

第 4 章，基于双特征描述符的闭环检测算法。针对低光环境下 SLAM 闭环检测精度不足、易产生错误回环的问题，提出了一种基于双特征描述符的改进算法。通过全局特征预筛选和局部特征精细匹配的双重机制，提升了闭环检测准确率和系统定位精度。实验采用室外 KITTI 和室内 TUM_VI、EUROC 等数据集，从准确率、召回率、轨迹精度和误差分布等维度与 ORB-SLAM3^[21]、VINS-FUSION 等主流方法进行对比，并在真实低光场景中验证了算法的有效性。结果表明，本文算法显著提升了低光环境下的闭环检测精度和系统性能，为低光视觉 SLAM 研究提供了重要参考。

第 5 章，总结与展望。总结了本文的工作内容和创新之处，在此基础上讨论了当前研究不足之处并展望了未来工作，指出了下一步可以研究的方向。

第 2 章 视觉 SLAM 与深度学习基础理论

2.1 视觉SLAM理论介绍

视觉同步定位与建图 SLAM^{[19][20][21]}系统凭借其低成本、结构简单以及图像信息丰富等显著优势，已成为机器人领域的研究热点。近年来，随着技术的不断突破和应用场景的扩展，视觉 SLAM 系统在多个领域展现了广泛的应用潜力。例如，在无人驾驶领域，该系统能够实时提供周边环境信息，并精准识别车辆行驶区域，为自动驾驶决策提供关键支持；在无人机领域，视觉 SLAM 系统不仅能够精确定位无人机的位置，还能动态调整其飞行姿态，为路径规划提供可靠依据；在室内机器人领域，该系统通过对室内环境进行高效重构，为机器人的避障与路径规划提供了重要的环境信息支持。

视觉 SLAM 系统的研究历史可追溯至 1986 年，当时研究人员首次尝试利用估计理论方法 ETM 解决机器人构图与定位的难题。一个典型的视觉 SLAM 系统通常由四个核心模块组成：视觉里程计、回环检测、后端优化和建图模块，如图 2-1 所示。这些模块协同工作，共同实现了机器人在未知环境中的精准定位与地图构建功能。

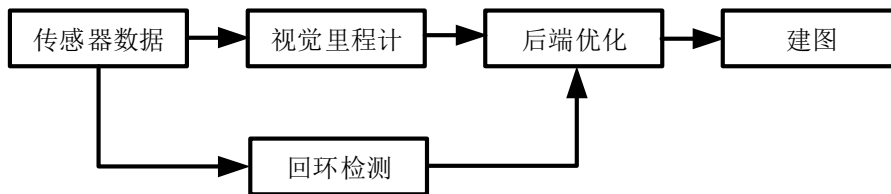


图 2-1 视觉 SLAM 基础框架图

系统启动并初始化完成后，相机实时捕获的场景信息被传输至前端视觉里程计模块，该模块通过估算相邻图像帧间的相对运动关系，分析机器人运动状态并生成初步轨迹和局部地图。由于仅依赖相邻帧信息，视觉里程计会产生累积误差，回环检测模块通过比较当前帧与历史帧的相似性能够识别出相机是否曾到达过某一位置，后端优化模块则利用这些信息进行全局优化，最终获得高精度且全局一致的轨迹和环境地图。

本文将对视觉 SLAM 系统中各个模块的具体理论基础与实现方法进行详细论述。

2.1.1 传感器数据

视觉 SLAM 系统中使用的基于视觉的传感器主要有三种类型：单目相机、双目相机和 RGB-D 相机。单目相机因其结构简单、成本低廉、执行效率高，成为应用最广泛的相机类型，尤其在移动机器人领域得到了广泛应用。

单目相机采用针孔成像模型，针孔模型展示了如何将二维图像上的像素点与三维空间中的点进行映射，其工作原理如图 2-2 所示。

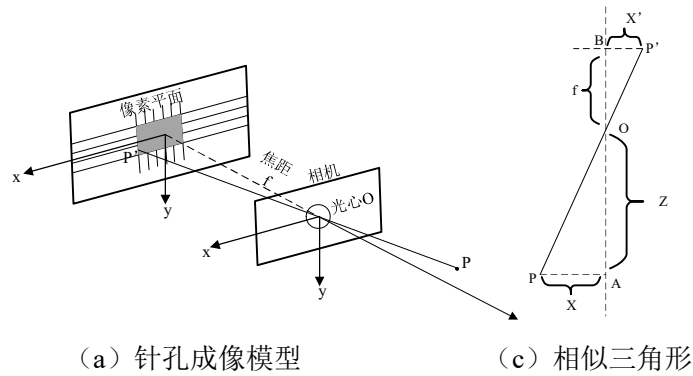


图 2-2 单目相机模型

双目相机，也称为立体相机，由左右两个单目相机组成。该模型模拟人类的视觉系统，通过对比左右两侧不同视角的图像信息来计算深度和距离信息，即可确定物体到相机的距离，如图 2-3 所示。

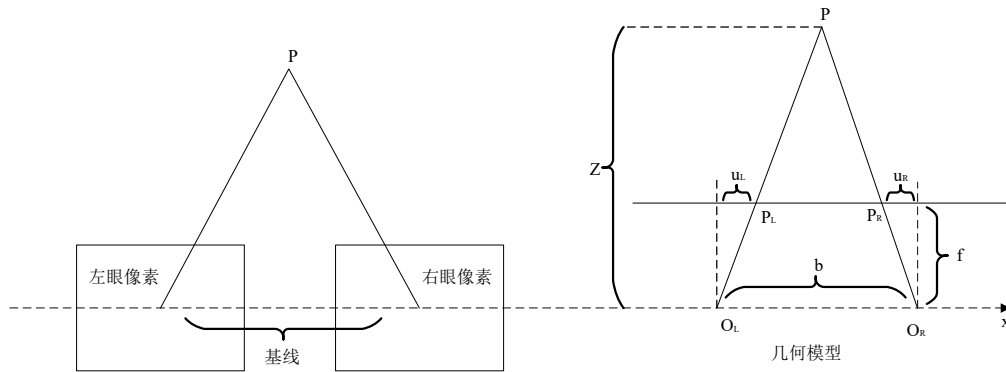


图 2-3 双目相机模型

RGB-D 深度相机是一种新型的传感器，可以捕获每个像素的深度，被广泛应用于语义映射，机器人等领域，通过红外结构光来标定像素间隔，如 Kinetic 1 代、Project Tango 1 代等，原理如图 2-4 所示。

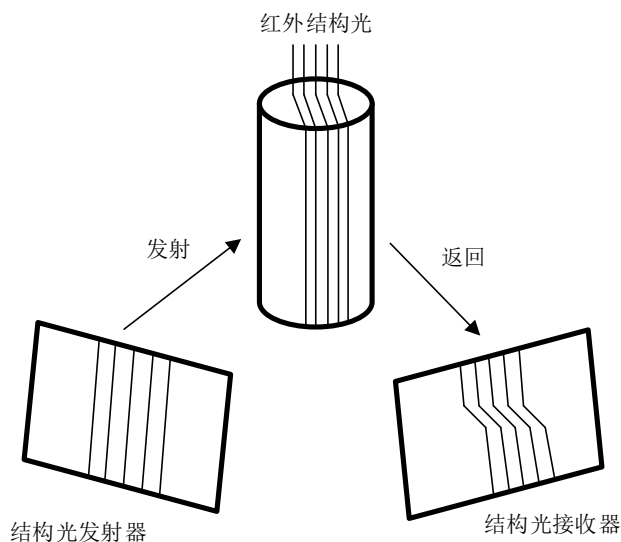


图 2-4 结构光原理图

2.1.2 视觉里程计

视觉里程计是一种通过连续图像序列估计移动机器人运动轨迹的算法，主要包括特征点法^[22]、直接法^[23]和半直接法^[24]。特征点法通过提取和匹配相邻帧图像中的特征点来估计帧间运动，直接法利用图像像素的光强差异计算运动，半直接法则结合了两者的优势，以特征点附近的图像块为匹配单元进行运动估计，目前特征点法仍是 SLAM 系统前端的主流算法，其提取的特征点具备快速性、尺度不变性和旋转不变性，可以满足实时计算需求。

特征点提取包含关键点检测和描述子生成两部分，关键点通常为角点，尽管边缘线段提取方法有所发展，但角点检测仍占主导地位，描述子通过关键点邻域信息生成特征向量，常用算法包括 ORB^[25]、FAST 和 SIFT^[26]等，如图 2-5 所示。其中，SIFT 精度高但计算量大，ORB 由 FAST 角点和 BRIEF 描述子组成，FAST 角点通过比较像素和其邻域的差异可以快速判断角点，显著提升计算效率，然而原始的 FAST 算法缺乏尺度和旋转不变性，ORB 算法通过改进解决了这一问题，ORB 通过计算 FAST 角点所在像素块的旋转方向，实现了旋转不变性，为增强尺度不变性，通过图像金字塔对原图像多次缩放生成多尺度图像，并在每层提取特征点，通常使用 8 层金字塔以平衡计算速度和精度，BRIEF 描述子通过比较关键点及其邻域灰度值的大小关系，生成 128 维二进制向量（由 0 和 1 组成）表示特征，其可视化效果如图 2-6 所示。

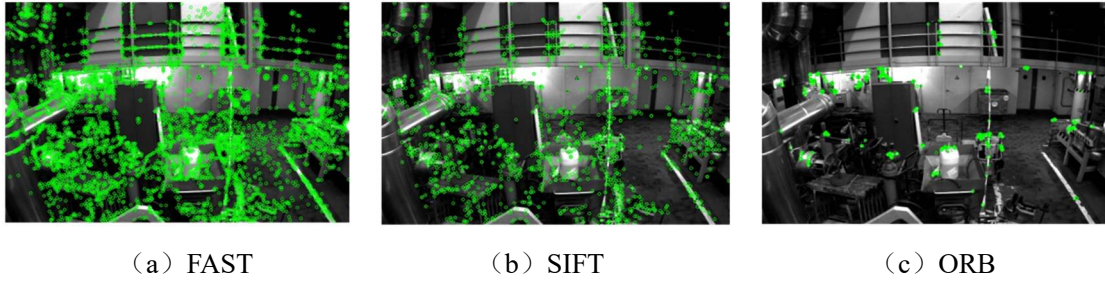


图 2-5 不同种类特征点

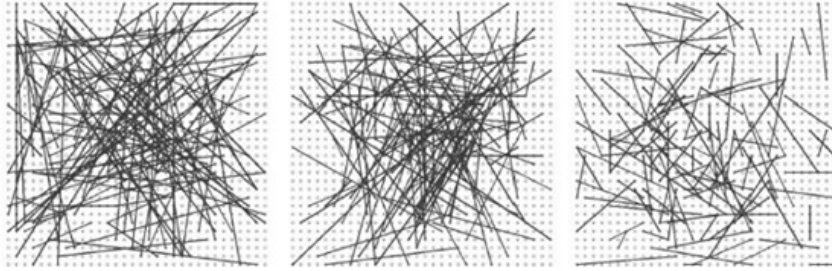


图 2-6 可视化

ORB 通过使用改进的 FAST 算法，特征点提取具有了尺度不变性和旋转不变性，同时具有 BRIEF 描述子的快速计算优势，因此，ORB 特征点提取算法既具有较好的计算速度，又具备较高的精确度，成为视觉 SLAM 中最常用的特征提取方法。

2.1.3 后端优化

在 SLAM 系统中，前端视觉里程计虽能实现瞬时运动轨迹的精确估计，但是随着系统长时间运行，会出现大量的累积误差，后端优化^[27]作为 SLAM 体系的关键环节，不仅提升了地图构建的精度，还强化了机器人的定位准确性，有效解决了数据间的关联难题，增强了系统的鲁棒性。

在后端优化策略中，SLAM 早期主要采用卡尔曼滤波方法，通过融合先验知识与实时传感器数据，并利用递归算法实现了系统状态的精确预测与更新。然而传统卡尔曼滤波方法如扩展卡尔曼滤波 EKF 存在明显的局限性，其基于马尔可夫假设，仅考虑当前时刻的观测信息，而未能充分利用历史数据中的有效信息。这种局限性导致系统误差随时间推移不断累积，最终影响 SLAM 系统的长期稳定性和定位精度，为克服这一问题，非线性优化技术尤其是图优化框架下的光束平差法 BA，为 SLAM 系统提供了较好的解决方案。BA 方法通过对关键帧的精挑细选，构建优化图，通过只纳入满足特定条件的特征点与位姿，有效避免了数

据冗余，减轻了计算负担。优化后的特征点融入地图，形成局部地图，因其包含所有帧的信息，实现了误差的均匀分配，显著提高了地图构建的效率与精度，成为了后端优化的主流方法。

非线性优化方法在后端优化中的应用，不仅克服了传统滤波技术的局限性，更通过高效的计算模型与优化算法，为构建高精度、实时性强的地图提供了技术支撑，推动了 SLAM 技术向更加智能化、自主化的方向发展。

2.1.4 闭环检测

闭环检测是视觉 SLAM 系统的关键模块，主要用于解决视觉里程计因误差累积引起的定位漂移问题。闭环检测通过识别相机是否重返历史场景，提供远距离帧间约束信息，并将这些信息传递至后端优化模块，从而优化全局地图与轨迹，修正位姿偏差，减少漂移误差。在环境地图构建与机器人定位中，闭环检测通过判断机器人是否回到历史位置，检测到闭环后校正地图和位姿，生成全局一致的地图。

传统闭环检测方法主要分为基于几何关系和基于视觉外观两类。基于几何关系的方法依赖于里程计数据，通过判断相机是否接近历史位置来检测回环，但由于里程计误差的限制，其准确性较低。基于视觉外观的方法则通过图像特征匹配实现闭环检测，不依赖里程计数据，使其成为一个独立模块，该方法的核心在于图像特征选择与匹配计算，通常通过特征提取和距离度量，如欧式距离、余弦距离等实现匹配。当前主流方法是词袋模型 BoW，尤其在 ORB-SLAM 系统中表现优异。词袋模型通过手工提取图像特征并进行聚类，利用树状结构提高检索效率，在相似度计算阶段，根据字典生成描述向量，通过向量相似度判断是否闭环。

当相机重返同一场景时，历史数据形成闭环，系统通过重新优化消除大量误差，如图 2-7 所示。图 2-7 (a) 中由于系统未能检测到机器人回到同一地点，无法形成闭环，导致算法估计轨迹与真实轨迹之间存在较大偏差，而在图 2-7 (b) 中，系统成功检测到闭环后，通过闭环校正显著减少了累积误差，提高了地图的一致性与定位精度。正确的闭环检测能够显著提升 SLAM 系统的地图质量与轨迹准确性，保证全局一致性，并在精度要求较高的应用中发挥重要作用。

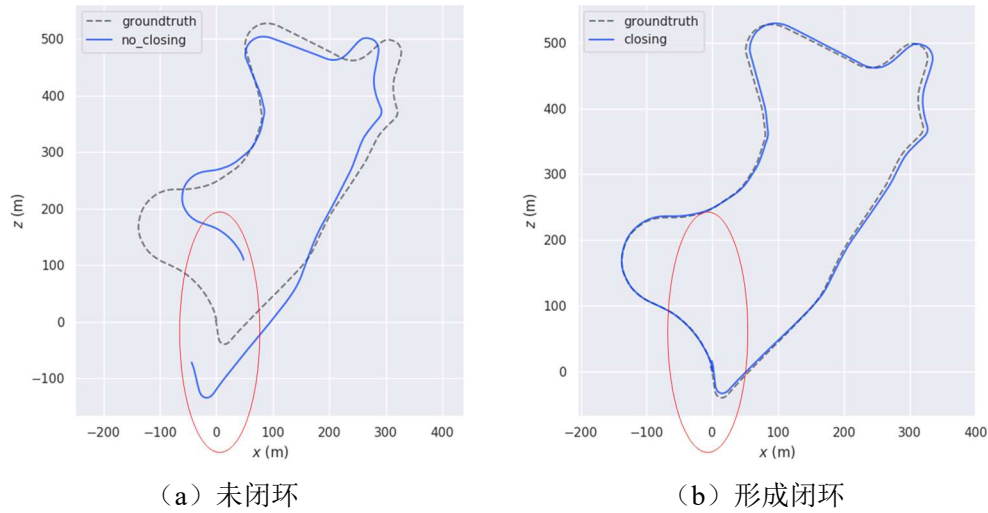


图 2-7 闭环效果

2.1.5 建图

视觉 SLAM 系统通过实时观测与建模环境中的路标点，构建精确的环境地图，该技术框架主要包含三个核心模块：运动估计、后端优化和闭环检测，从而确保系统能够精确定位路标点并整合轨迹数据生成高精度地图。视觉 SLAM 系统构建的环境地图不仅为移动机器人提供了可靠的定位与导航能力，还为避障、三维环境重建以及人机交互等高级功能的实现奠定了重要基础。

在地图构建过程中，定位功能通过实时采集环境信息并与历史地图进行匹配，能够快速确定机器人的当前位置，导航功能则依赖于地图的路径规划能力，机器人通过分析地图信息，能够规划并选择最优路径以实现目标位置的到达。避障功能在动态环境中显得尤为重要，地图结合实时传感器数据，能够帮助机器人有效检测障碍物并动态调整行进路径，确保安全导航。重建功能通过相机采集的环境数据和位姿估计技术，生成高精度的三维地图，不仅支持机器人的导航需求，还可应用于虚拟现实（VR）和增强现实（AR）等领域。交互功能通过在地图中添加高层语义信息，使系统能够更深入地理解环境，从而支持更复杂的任务需求，提升机器人在多样化场景中的适应性和智能化水平。

SLAM 系统通常采用稀疏地图和稠密地图两种形式，稀疏地图由关键点或显著特征组成，适用于计算资源有限或实时性要求高的场景，稠密地图通过网格、像素或点云等形式提供详细环境信息，适用于高精度环境建模的应用。

2.2 深度学习理论基础

2.2.1 卷积神经网络

卷积神经网络^{[28][29]}是深度学习的关键算法，它在视觉和早期自然语言处理中表现卓越，擅长特征提取与学习，广泛应用于特征提取^{[30][31]}、目标识别^{[32][33]}、图像分割^{[34][35]}和分类^{[36][37]}等任务。卷积神经网络的发展可追溯至 Fukushima 等学者提出的神经认知机模型，模拟大脑神经元信息传递，通过逐层提取特征并整合，形成高级数据表示。卷积神经网络的结构通常较为清晰，其核心结构由输入层、隐藏层和输出层三个主要部分构成。在隐藏层中，包含了多个功能单元包括卷积层、激活函数、池化层和全连接层，这些功能单元按照特定的顺序进行组织，其中卷积层与池化层交替排列，如图 2-8 所示。

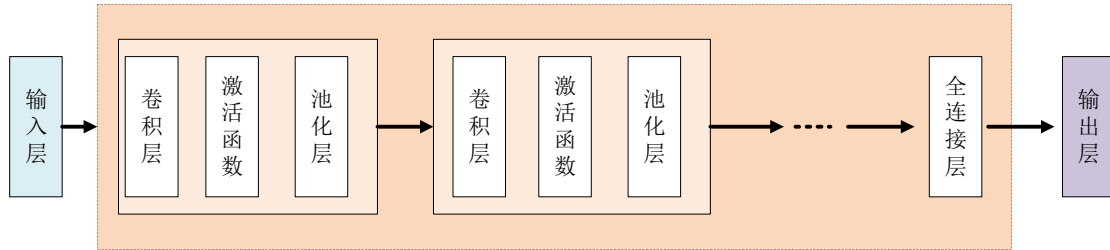


图 2-8 卷积神经网络结构图

卷积层作为卷积神经网络的核心，主要负责图像特征的提取任务。其通过卷积操作生成有序的特征图，其中包含多个卷积核和偏置项，且卷积核的数量与输入通道数保持一致。在计算过程中，卷积核在输入图像上以滑动窗口的方式移动，与图像的局部区域进行逐元素乘积运算，并累加结果后加上偏置项，从而生成特征值，这一过程持续进行，直至卷积核遍历整个输入图像。不同尺寸的卷积核能够提取不同层次的特征，单一尺寸的卷积核识别能力有限，因此采用多尺寸卷积核（如 1×1 、 3×3 、 5×5 、 7×7 等）可以更全面地捕捉图像的局部和全局特征。此外，卷积层的深度直接影响特征的抽象程度，浅层卷积层通常提取基础特征，如纹理、边缘和色彩信息，而深层卷积层则能够识别更加复杂和抽象的特征，如物体形状和语义信息。如图 2-9 所示，以一个 5×5 的输入图像与 3×3 的卷积核为例，在步长为 1 的情况下，卷积操作生成一个 3×3 的特征图，直观展示了卷积过程的计算机制。

在卷积神经网络中，激活函数负责构建输入与输出之间的复杂非线性映射，

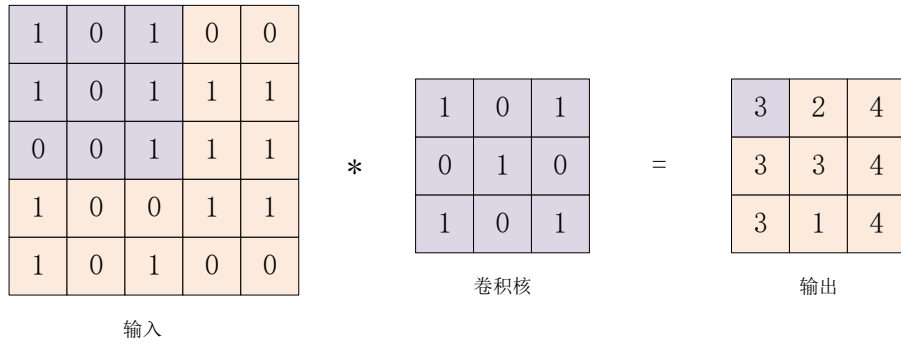


图 2-9 卷积运算过程

若没有激活函数，网络各层之间将仅维持线性关系，导致输出退化为输入的线性变换，即使增加网络深度也无法提升性能，激活函数的引入打破了这种线性限制，赋予网络强大的非线性表达能力，使其能够捕捉和学习更加多样化和复杂的特征，从而显著提升网络的性能。常见的激活函数包括 Sigmoid、ReLU 等，每种函数都有其独特的优点和局限性，适用于不同的应用场景，其公式如（2-1）和（2-2）所示：

$$\text{Sigmoid}(x) = \frac{1}{1+e^{-x}} \quad (2-1)$$

$$\text{ReLU}(x) = \max(0, x) \quad (2-2)$$

池化层又称下采样层，是卷积神经网络中与卷积层交替使用的常见结构，其作用是对特征图下采样，保留关键特征并去除冗余信息，从而降低特征向量的维度和参数，减少计算复杂度，防止过拟合，提升网络泛化能力。池化层还能扩大卷积核的感知范围，便于提取图像整体特征，避免遗漏重要信息。最大池化和平均池化是两种常用操作，图 2-10 展示了其计算过程，对与 4x4 的输入特征图，使用 2x2 核、步长 2 的池化层，分别输出 2x2 的特征图，最大池化输出区域最大值，平均池化输出区域平均值。

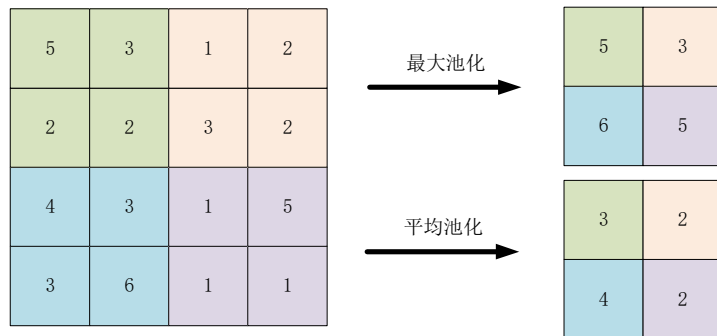


图 2-10 池化运算过程

在卷积神经网络架构中，全连接结构一般作为网络的最终层级，承担着特征整合与结果输出的关键作用。该结构通过融合卷积层和池化层提取的多维特征，将其转换为统一的特征表示，这种转换机制有效降低了特征空间位置对分类决策的干扰，提升了模型的泛化能力。全连接层实现了从特征空间到类别空间的映射，通过计算输入样本属于各个类别的可能性，有效缓解了特征分布差异对识别精度的影响。由于全连接层的每个神经元都需要与前一层所有神经元建立连接，导致其参数量较为庞大。在特征处理过程中，多维特征首先被展平为一维向量，随后与可学习的权重矩阵进行线性变换以强化关键特征，最终通过分类器输出预测结果，如图 2-11 所示。

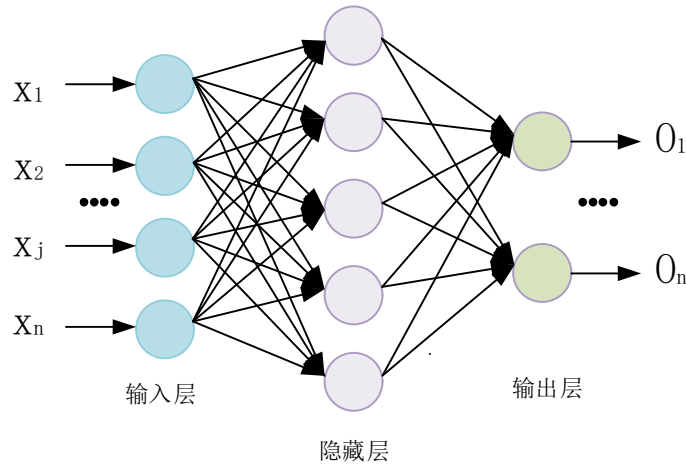


图 2-11 全连接层连接方式

神经网络中的全连接结构通常采用 ReLU 函数来实现特征的非线性转换，然而在网络的输出层，研究者们往往选择使用 softmax 分类器来生成概率形式的预测结果。这种转换机制能够将网络输出映射为概率分布，从而提升模型的分类性能。考虑到全连接结构的参数规模会随着网络深度的增加呈现指数级增长，为了缓解过拟合现象并提高训练效率，学术界提出了使用全局池化策略来替代传统的全连接层。全局平均池化通过计算特征图各通道的均值来实现特征降维，这种方法在显著降低参数数量的同时，还能有效保留特征的空间信息，在视觉识别任务中展现出良好的性能。

2.2.2 注意力机制

注意力机制通过关注主要信息并忽略次要信息，能够显著提升系统性能，它在语义分割、目标检测与图像分析等多个领域均有重要应用。其基本原理是采用

双重加权策略来强化图像中的重要特征,引导深度模型在训练过程中聚焦于关键区域,进而生成了注意力权重分布图。注意力机制依据实现方式可分为两类,强注意力和软注意力。强注意力机制关注动态变化,通过随机选取图像的部分区域作为输入,通常不可微,需依赖增强学习进行训练。相比之下,软注意力机制具有可微特性,可以通过深度神经网络进行梯度计算与训练,其权重经 softmax 归一化后分布在 $[0,1]$ 区间,用于量化各区域的关注程度,此外软注意力机制还可划分为通道注意力、空间注意力以及混合注意力三种类型。

在通道注意力机制方面,该机制主要聚焦于特征图的通道维度,为相关性较强的通道分配更高的权重值,例如 Hu 等人^[38]提出的 SENet 通过捕获特征通道间的相互关系来强化关键特征表达,而 Zhang 等人^[39]设计的 RCAN 在此基础上利用跳跃连接传递低频信息,从而引导网络重点学习高频特征信息。空间注意力机制则更加注重图像的空间特征信息,通过压缩通道数量使模型更多的关注空间信息, Jaderberg 等人提出的空间变换网络利用空间信息进行自适应特征变换,从而提取关键信息, Wang 等人^[40]则通过构建不同空间位置间的相互关系,捕捉长距离依赖特征,以保留更丰富的图像信息。混合注意力结合了多种注意力机制,能够同时考虑空间和通道信息,并根据具体任务需求自适应地选择合适的注意力机制类型,例如通道-空间混合注意力机制结合了通道注意力和空间注意力机制的优势,能够有效应对复杂的任务场景。

2.2.3 低光增强算法

低光增强算法旨在改善低光照条件下图像的可见性和质量,基于深度学习的方法,通过神经网络模型直接从低光图像中恢复细节和色彩信息,这类算法可以分为端到端模型和多阶段增强方法两种,如 FourLLIE 和 EMNet 等,它们具有较高的运行效率,能够满足数据处理的需求,其中 EMNet 采用的是轻量级网络设计,专注于低光图像的高效增强,其网络结构包括编码器、解码器和注意力机制模块,编码器通过卷积层提取低光图像的特征,解码器逐步恢复图像的细节和色彩,注意力机制模块则用于增强关键区域的细节表现,进一步提升图像质量, EMNet 在保证增强效果的同时,显著降低了计算复杂度,适合低光应用场景。图 2-12 展示了低光增强算法的图像增强效果,通过深度学习的方法对低光图像进

行处理，能够显著提升图像质量。

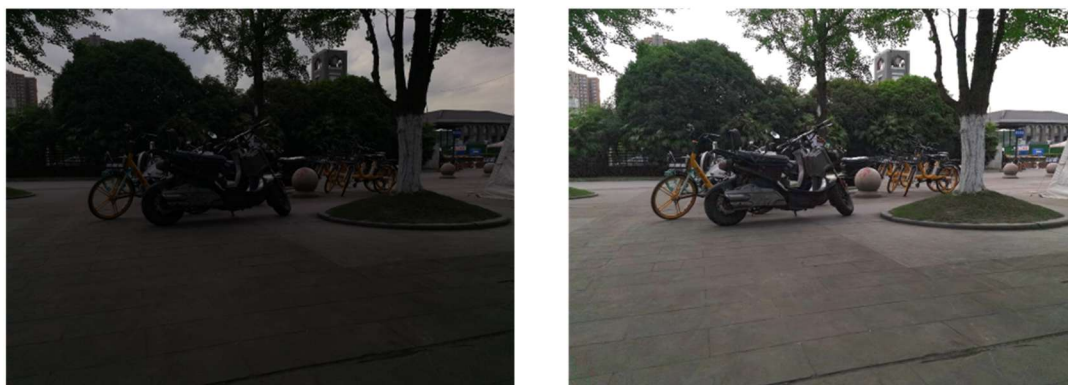


图 2-12 低光增强结果图

2.4 本章小结

本章系统阐述了视觉 SLAM 与深度学习的核心概念及其理论基础。首先，从视觉 SLAM 的整体框架出发，详细介绍了其关键组成部分，包括传感器数据采集、视觉里程计、后端优化、闭环检测以及地图构建等五个核心模块，随后深入探讨了深度学习的理论基础及其构建的低光增强算法，这些内容为深度学习与视觉 SLAM 的深度融合奠定了坚实的理论基础。

第3章 基于 FFE-Net 网络的低光增强算法

移动机器人 SLAM 运行于低光场景时，由于低光条件的影响易出现系统特征点提取困难，匹配精度不高的问题，极大的影响视觉 SLAM 系统位姿估计精度以及定位任务。基于传统视觉 SLAM 算法在低光条件下的局限性，本文提出一种基于深度学习的低光增强算法即 FFE-Net，对原始输入数据进行特征增强。

该网络包含傅里叶空间亮度增强模块和特征细节恢复模块。傅里叶空间中振幅大小与图像亮度之间存在正相关关系，因此本文采用放大低光图像的振幅分量幅度，保留相位分量信息的方法来实现低光图像的亮度提升。由于低光图像的纹理和边缘信息一般存在于图像的高频信息中，而低照度图像往往包含更多的低频信息。为了有效恢复图像的细节信息，分离出图像的高低频分量，并对分离出的分量进行深层特征提取与细节修复，与传统的特征传递不同，本文使用跳转连接编解码器来巩固提取特征。此外，在特征细节恢复模块中还引入了差异权重网络（Differential Weighting Network, DWNet），使生成的图像保留足够的细节，本章技术路线如图 3-1 所示。

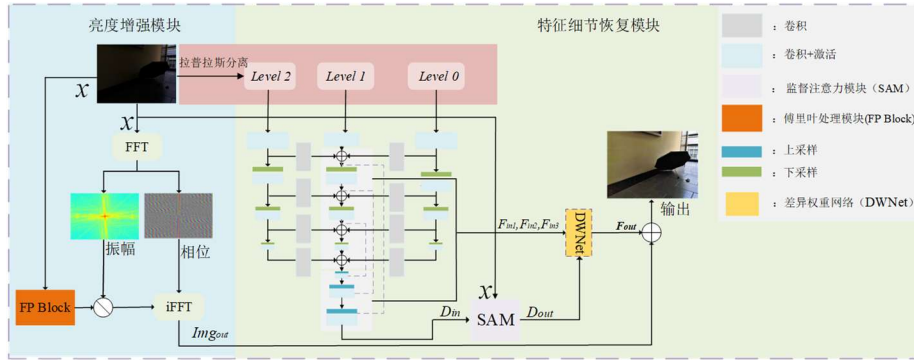


图 3-1 本章技术路线图

3.1 亮度增强模块

对于给定的输入图像 x ，图像参数为 $H \times W$ ，将输入 x 转换到傅里叶空间 X 的变换函数其公式如（3-1）所示：

$$X(u, v) = F(x(h, w)) = \frac{1}{\sqrt{HW}} \sum_{h=0}^{H-1} \sum_{w=0}^{W-1} x(h, w) e^{-j2\pi \left(\frac{h}{H}u + \frac{w}{W}v \right)} \quad (3-1)$$

其中， H 和 W 分别表示图像的长和宽； $F(\cdot)$ 表示傅里叶变换； $x(h, w)$ 为某一像

素点的空间坐标， $X(u, v)$ 是其对应的傅里叶空间坐标。

傅里叶空间中的 $X(u, v)$ 可以用振幅分量 $A(X(u, v))$ 和相位分量 $P(X(u, v))$ 表示，其公式如 (3-2) 和 (3-3) 所示：

$$A(X(u, v)) = \sqrt{R^2(X(u, v)) + I^2(X(u, v))} \quad (3-2)$$

$$P(X(u, v)) = \arctan(I(X(u, v)) / R(X(u, v))) \quad (3-3)$$

其中， $A(\cdot)$ 表示提取振幅， $P(\cdot)$ 表示提取相位； $R(X(u, v))$ 和 $I(X(u, v))$ 分别为 $X(u, v)$ 的实部和虚部。

亮度增强模块将输入图像 x 送入到傅里叶处理模块 (Fourier Processing, FP) 中，以估计其振幅变换图 Q 并用来放大振幅分量，实现低光图像的亮度提升，FP 块的详细结构如图 3-2 所示。

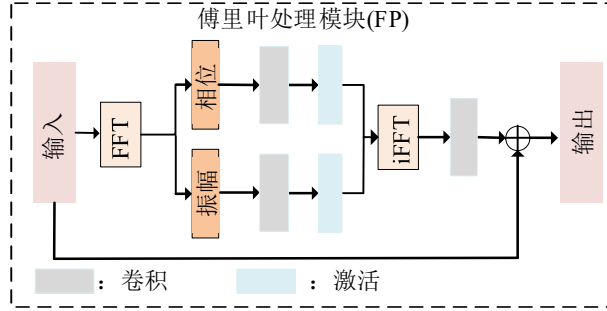


图 3-2 FP 块的详细结构

FP 块将输入转换到傅里叶空间后获取其振幅和相位分量，并对这两个分量分别进行卷积和激活来提取特征。然后这两个分量被转换回时域，并通过一个 3×3 的卷积层进行输入残差处理得到输出的振幅变换图 Q ，将输入图像的振幅分量除以估计的振幅变换图获得放大后的振幅分量，并与相位分量结合进行傅里叶逆变换后得到输出 Img_{out} ，其公式如 (3-4)、(3-5)、(3-6) 和 (3-7) 所示：

$$A_{in} = A(F(x)), P_{in} = P(F(x)) \quad (3-4)$$

$$A_{out} = A_{in} / (Q + \varepsilon) \quad (3-5)$$

$$R_{out} = A_{out} \times \cos(P_{in}), I_{out} = A_{out} \times \sin(P_{in}) \quad (3-6)$$

$$Img_{out} = F^{-1}(R_{out} + jI_{out}) \quad (3-7)$$

其中， $F^{-1}(\cdot)$ 表示傅里叶逆变换； A_{in} 和 P_{in} 分别为输入图像的振幅分量和相位分量； A_{out} 为输出振幅分量； R_{out} 和 I_{out} 分别代表输出 Img_{out} 在傅里叶空间的实部和虚部； ε 为非常小的常数，本文设为 1.0×10^{-8} ，用于避免零做除数。

3.2 特征细节恢复模块

在亮度增强模块中提升了低照度图像的亮度，然而在细节区域损失严重，因此本文引入了特征细节恢复模块。该模块采用拉普拉斯金字塔对图像进行线性分解，得到一组高低频分量，并对这些分量进行全面学习，以恢复丢失的细节信息。

拉普拉斯金字塔^[41]是一种关键的图像处理技术，在图像处理领域具有广泛的应用价值，涵盖图像增强^[42]、融合^[43]、压缩^[44]以及超分辨率重建^[45]等多个方面。在图像增强中，通过处理不同频率的分量，能够在多尺度上优化图像细节，显著提升图像质量。在图像融合任务中，拉普拉斯金字塔支持多分辨率层次的融合操作，能够将来自不同传感器或曝光条件的图像信息有效整合，从而保留更多细节。此外，其多分辨率特性在图像压缩中表现优异，能够大幅降低数据存储需求。在超分辨率重建中，拉普拉斯金字塔的高频分量被用于提升图像分辨率，生成更加清晰的高质量图像。

拉普拉斯金字塔的核心原理是将图像分解为不同频率成分，以便在多尺度上进行图像处理，其中包括高斯金字塔的构建和拉普拉斯金字塔的计算，首先通过高斯平滑和下采样操作生成一系列分辨率逐渐降低的高斯金字塔图像，通过计算高斯金字塔相邻层之间的差异，提取包含图像细节信息的拉普拉斯金字塔层，包括高斯滤波、平滑处理、上采样和卷积等操作，通过对高斯金字塔的相邻层进行上采样和卷积操作，生成与下一层分辨率相同的图像，并计算两者差异，从而构建拉普拉斯金字塔的每一层，其公式如（3-8）所示：

$$L_i = G_i - UP(G_{i+1}) \otimes \zeta_{5 \times 5} \quad (3-8)$$

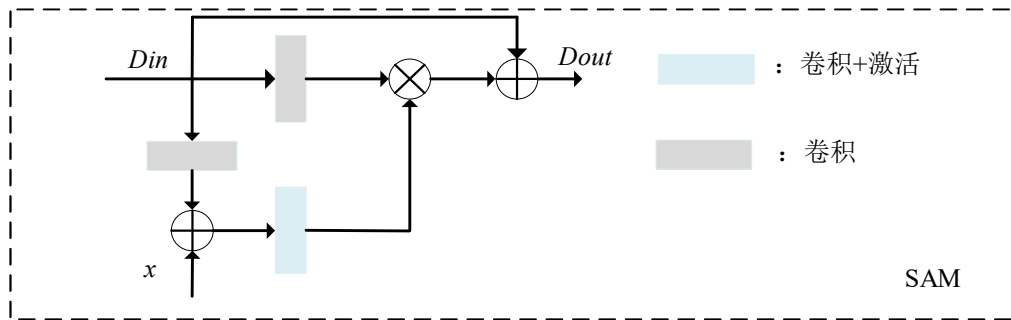
其中 L_i 表示拉普拉斯金字塔的第 i 层， G_i 表示高斯金字塔的第 i 层， $UP(\bullet)$ 表示上采样操作， $\otimes$ 表示卷积操作， $\zeta_{5 \times 5}$ 表示 5×5 的高斯核。

本文利用拉普拉斯金字塔对图像进行线性分解，获得一组包含高低频信息的分量，如图 3-3 所示，图（a）为原始图像，图（b）和图（c）分别表示高频分量，主要用于提取图像的边缘和纹理特征，而图（d）为低频分量，主要用于保留图像的整体色彩和光照信息。本文采用编解码器结构对上下文信息进行采集，充分提取图像的深层特征，从而确保输出高质量图像，使其能够有效恢复图像细节，并在多尺度上优化图像表现，为后续处理任务提供可靠支持。



图 3-3 拉普拉斯金字塔分解低光照图像示例

编解码器从输入图像中提取特征信息时, 也会引起无关特征的响应, 因此添加了监督注意力模块 (Supervision Attention Module, SAM) 来传递特征。SAM 使当前阶段更专注于有用的信息, 抑制无用的信息, 并为增强图像与真实图像进行矫正, 通过计算与参考图像之间的损失, 从而提高网络的整体性能, 具体结构如图 3-4 所示。



将原始图像 x 作为 SAM 的输入之一, 另一个输入是输入特征 D_{in} , 先通过卷积生成残差图像并将其添加到原始弱光图像 x 中, 生成渐进增强图像, 通过卷积加激活运算将渐进增强图像转换为注意力图, 将注意力图与卷积后的 D_{in} 相乘, 然后与输入特征 D_{in} 相加, 得到输出 D_{out} 。SAM 在当前阶段会抑制无关特征和信息量较少特征的响应, 然后将更有价值的特征传播到下一阶段。

目前的低光增强网络存在一个普遍性问题：处理低光图片时对通道和像素平等对待，未能有效处理通道方向光照分布的不均匀性。这种处理方式导致网络无法充分区分图像中通道特征和像素特征的重要性，从而影响了图像增强效果。因此，本文设计了一种差异权重网络 DWNet。该网络能够针对不同的通道特征和像素区域灵活地分配不同的权重，从而更有效地处理各种类型的信息，提高网络性能。

DWNet 网络结构如图 3-5 所示，它是由多个具有跳转连接的残差组 DW 组成的网络结构，每个残差组包含一些具有短跳转连接的差异权重注意力块（Differential Weighting Attention, DWA），差异结构中的残差可以通过多个跳转连接绕过丰富的低频信息，使主网络能够专注于学习高频信息，DWA 主要由通道注意力和像素注意力两大块组成，详细结构如图 3-6 所示。

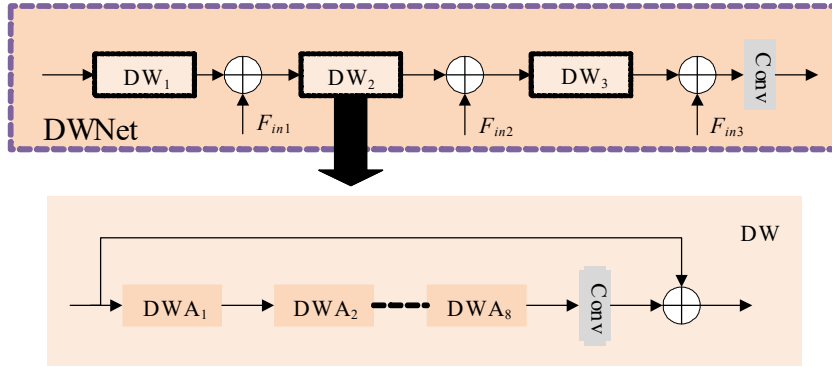


图 3-5 DWNet 网络结构

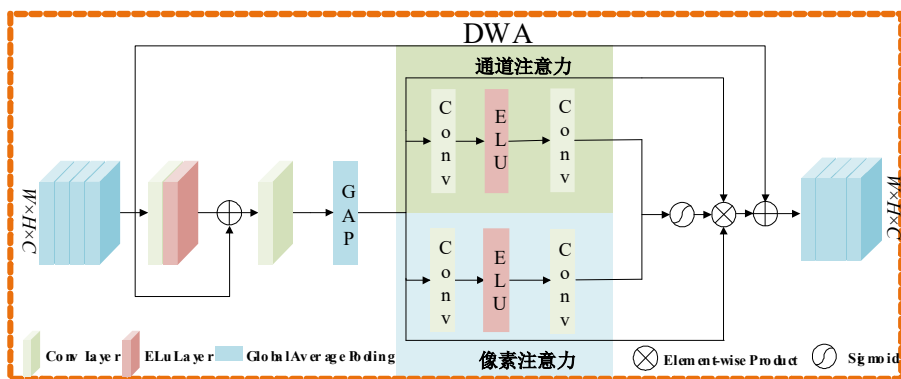


图 3-6 DWA 详细结构

输入通过卷积-激活-卷积运算得到中间特征图，特征图的形状从 $H \times W \times C$ 变为 $1 \times 1 \times C$ （ C 为图像的通道数），使底层网络能够充分使用全局信息，通过执行全局平均池化（Global Average Pooling, GAP）操作，将全局空间信息映射为通道

描述符 I_c 和像素描述符 I_p ，其公式如 (3-9) 和 (3-10) 所示：

$$I_c = H_p(F_{inc}) = \frac{1}{H \times W} \sum_{m=1}^H \sum_{n=1}^W V_c(m, n) \quad (3-9)$$

$$I_p = H_p(F_{inp}) = \frac{1}{H \times W} \sum_{m=1}^H \sum_{n=1}^W V_p(m, n) \quad (3-10)$$

其中， $V_c(m, n)$ 表示图像第 c 个通道 V_c 在 (m, n) 位置的值， $V_p(m, n)$ 表示图像的像素 V_p 在 (m, n) 位置的值； $H_p(\cdot)$ 表示全局池化函数； F_{inc} 表示 DWNet 的图像通道输入， F_{inp} 表示 DWNet 的图像像素输入， F_{inp} 和 F_{inc} 均能从 DWNet 的输入 F_{in} 中获取。

通道注意力模块自适应地重新调整每个通道特征的大小，侧重于不同通道特征，并优先考虑暗通道的情况，为每个通道特征生成不同的加权信息，而像素注意力模块使网络过滤无关的像素信息，更加关注低光像素等重要特征，并为其提供了差异化的权重信息。

为获取差异化的像素权值和通道权值，需经过两个卷积层和 Sigmoid、ELU 激活函数，其中 ELU 是 ReLU 函数的一个改进版本，其公式如 (3-11) 所示：

$$f(\kappa) = \begin{cases} \kappa & \kappa > 0 \\ \alpha(e^\kappa - 1) & \kappa \leq 0 \end{cases} \quad (3-11)$$

其中， $f(\kappa)$ 表示 ELU 激活函数， κ 为输入， α 为常数。与 ReLU 相比，ELU 在负值区域也有一个小的斜率，避免了其在负数区域输出为零可能导致的问题，且平均值接近零，学习更快。

用 CA_c 表示获取到的不同通道的权值，用 CA_p 表示获取到的不同像素的权值，其公式如 (3-12)、(3-13) 和 (3-14) 所示：

$$CA_c = \phi(\text{Conv}(\tau(\text{Conv}(I_c)))) \quad (3-12)$$

$$CA_p = \phi(\text{Conv}(\tau(\text{Conv}(I_p)))) \quad (3-13)$$

$$F_{out} = CA_c \otimes F_{in} \otimes CA_p \quad (3-14)$$

其中， $\tau(\cdot)$ 为 ELU 激活函数； $\phi(\cdot)$ 为 Sigmoid 激活函数；Conv 为卷积；将权重与输入 F_{in} 逐个相乘，得到输出 F_{out} 。

简单地将特征的输出作为下一阶段的输入可能会导致重要信息的丢失。为了在最终生成的图像中保留更多的细节信息（包括亮度、颜色等细节特征），并且

使图像能够更加接近真实图像，本文有效地融合了编解码器子网络和 DWNet 的特征，将编解码器的输出，添加到 DWNet 中的残差组 DW 之后，以减少编解码器重复的上采样和下采样带来的信息损失，并丰富下一阶段的特征，生成了具有丰富空间和清晰纹理的高分辨率特征，从而提高最终生成图像的质量。整体而言，特征增强网络通过有效整合上下文信息和特征融合机制，有力支持了图像生成过程中的细节保留和真实性保持。

3.3 组合损失函数

在训练中，为提升低光照图像的视觉效果，经过多次实验发现，单一损失函数难以显著增强图像亮度。为解决此问题，本文设计了一种组合损失函数，该函数综合运用 Charbonnier 损失、边缘损失以及结构相似性指数 (Structural Similarity Index Measure, SSIM)，全面考虑网络需求、上下文结构信息。其中，Charbonnier 损失作为 L_1 损失的一种变体，可以有效提升模型性能；边缘损失有助于更好地增强网络结构，获取图像边界特征；SSIM 从亮度、对比度和结构三方面对图像进行整体评估。本文使用三种函数的组合来计算网络模型损失，从而进一步优化低光照图像的增强效果。

用 L_{edge} 表示边缘损失， L_{char} 表示 Charbonnier 损失，则 L_1 的损耗为亮度增强模块的损失，其公式如 (3-15)、(3-16) 和 (3-17) 所示：

$$L_{edge} = \sqrt{\|X_M - Y\|^2 + \delta^2} \quad (3-15)$$

$$L_{char} = \sqrt{\|\Delta(X_M) - \Delta(Y)\|^2 + \delta^2} \quad (3-16)$$

$$L_1 = L_{edge} + \lambda_1 L_{char} \quad (3-17)$$

其中， Y 为地面真实图像， X_M 代表恢复的图像； δ 为常数，本文取 10^{-3} ； Δ 表示拉普拉斯算子； λ_1 表示权衡两个损失函数的比重，本文设为 0.05。

对于特征增强模块的损失用 L_2 表示， L_{SSIM} 表示 SSIM 损失，公式如 (3-18) 和 (3-19) 所示：

$$L_{SSIM} = 1 - \frac{1}{N} \sum_{p \in \text{img}} \frac{2\mu_a\mu_b + C_1}{\mu_a^2 + \mu_b^2 + C_1} \cdot \frac{2\sigma_{ab} + C_2}{\sigma_a^2 + \sigma_b^2 + C_2} \quad (3-18)$$

$$L_2 = L_{edge} + \lambda_2 L_{char} + \lambda_3 L_{SSIM} \quad (3-19)$$

其中， a 、 b 分别表示增强图像和参考图像， μ_a 、 μ_b 分别是 a 、 b 的均值， σ_a 、

σ_a 分别是 a 、 b 的方差, σ_{ab} 表示 a 和 b 的协方差; λ_2 、 λ_3 为常数, 本文设为 0.01 和 0.05, 引入常数 C_1 和 C_2 是为了防止出现分母为零的情况。

最后, 总损失 L_{total} 的计算公式如 (3-20) 所示:

$$L_{total} = L_1 + \lambda L_2 \quad (3-20)$$

其中, λ 根据经验设为 0.01。

3.4 实验结果与分析

本章在公开数据集上进行低光增强对比实验以及特征点提取实验和特征匹配实验来对算法进行有效性分析。使用的电脑配置为: CPU 为 12th Gen Intel(R) Core(TM) i5-12400F 处理器, 主频 2.6GHZ, 内存 16.00 GB (2667 MHz), GPU 为 NVIDIA GeForce RTX 3060, 显存 24GB, 系统为 Ubuntu18.04。

3.4.1 图像低光增强对比试验

LSRW (Low-Light Scene Recognition and Weak Illumination Dataset) 低光数据集分为华为(Huawei)和尼康(Nikon)两个不同的子集, 是一个专为测试和评估在低光照条件下图像增强算法性能而设计的专业数据集, 在该数据集上对 IAT (Illumination Adaptive Transformer)、EMNet (Enhanced Multi-scale Network)、FourLLIE (Fourier-based Low-Light Image Enhancement) 三种流行的低光增强算法以及本文算法进行测试, 每种方法增强后的结果如图 3-7 所示, 其中 Img1~4 是选取的低光数据集中的四张图像, GT 表示原始图像, Input 为对应的低光输入图像。由图可以看到: 与原始图像相比, 所有方法都提高了图像的整体亮度, 但 IAT 算法容易出现严重的色彩失真, FourLLIE 方法对于细节恢复不足, EMNet 算法在整体亮度和自然度方面较好, 但局部区域亮度增强不足, 与之相比本文算法在亮度和细节上都取得了更好的恢复效果。



(a) GT/Img1 (b) Input/Img1 (c) IAT/Img1 (d) EMNet/Img1 (e) FourLLIE/Img1 (f) FFE-Net/Img1



(g) GT/Img2 (h) Input/Img2 (i) IAT/Img2 (j) EMNet/Img2 (k) FourLLIE/Img2 (l) FFE-Net/Img2

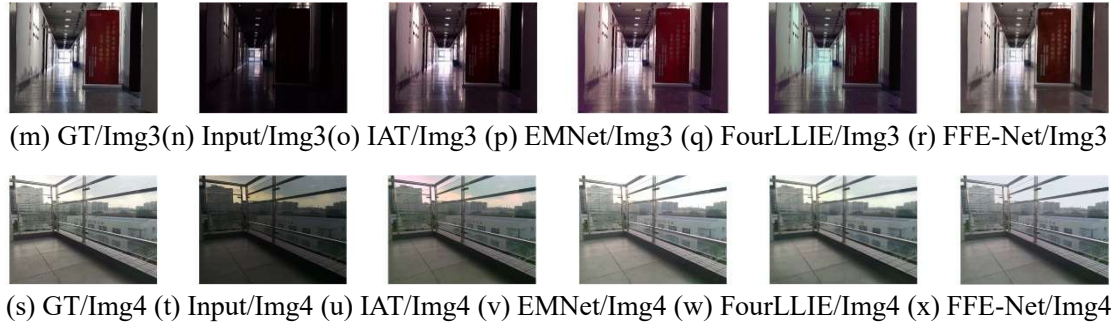


图 3-7 不同方法增强效果对比

除了主观对于图像进行评价，实验分别计算本文算法与其他算法在 8 个低光数据集上的峰值信噪比 PSNR 和 SSIM 评价指标，并将最佳结果以黑体显示，结果如表 3-1 所示。其中，PSNR 通过比较增强图像和原始图像之间对应像素的差异来评估图像质量，PSNR 值越大，表示增强后的图像越接近原始图像，噪声较小，图像的增强效果越好，SSIM 用于评价增强后的图像和原始图像之间的相似性，主要从亮度、对比度和结构信息三个方面对重建后的清晰图像进行评价，SSIM 值越大表示相似程度越高。由表 3-1 可知，本文算法在大部分数据集上均表现最佳，整体上能够确保在大部分数据集中的良好效果。

表 3-1 不同评价指标对公共数据集中的增强图像进行客观评估的结果

数据集	FourLLIE ^[10]		EMNet ^[9]		IAT ^[46]		FFE-Net	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
HUAWEI	27.57	0.64	27.87	0.59	27.49	0.49	27.55	0.60
LOL_Real	27.56	0.82	30.58	0.94	28.41	0.90	34.15	0.94
LOL_Synthetic	27.81	0.56	27.54	0.65	28.88	0.67	28.11	0.69
LSRW_Nikon	28.32	0.44	28.20	0.41	27.86	0.39	28.64	0.42
RELLISUR	27.45	0.88	27.99	0.73	27.49	0.66	29.91	0.76
DPED_iphone	27.38	0.51	27.77	0.25	29.48	0.73	30.07	0.80
DPED_blackberry	27.17	0.44	29.04	0.70	27.77	0.71	27.70	0.53
DPED_sony	27.45	0.25	28.20	0.18	27.45	0.25	27.40	0.37

3.4.2 特征提取对比实验

本节实验将低光图像分别送入 FFE-Net、IAT、FourLLIE 和 EMNet 四种增强算法中进行处理，得到数据集中低光图像及其增强后的结果，分别进行特征提取实验用于验证本文提出的低光照增强网络 FFE-Net 在 SLAM 系统中的有效性。低光照图像增强前后特征点提取对比如图 3-8 所示，其中阴暗部分为主要的差异区域，对比内容包括特征点的数量、提取位置以及图像质量。

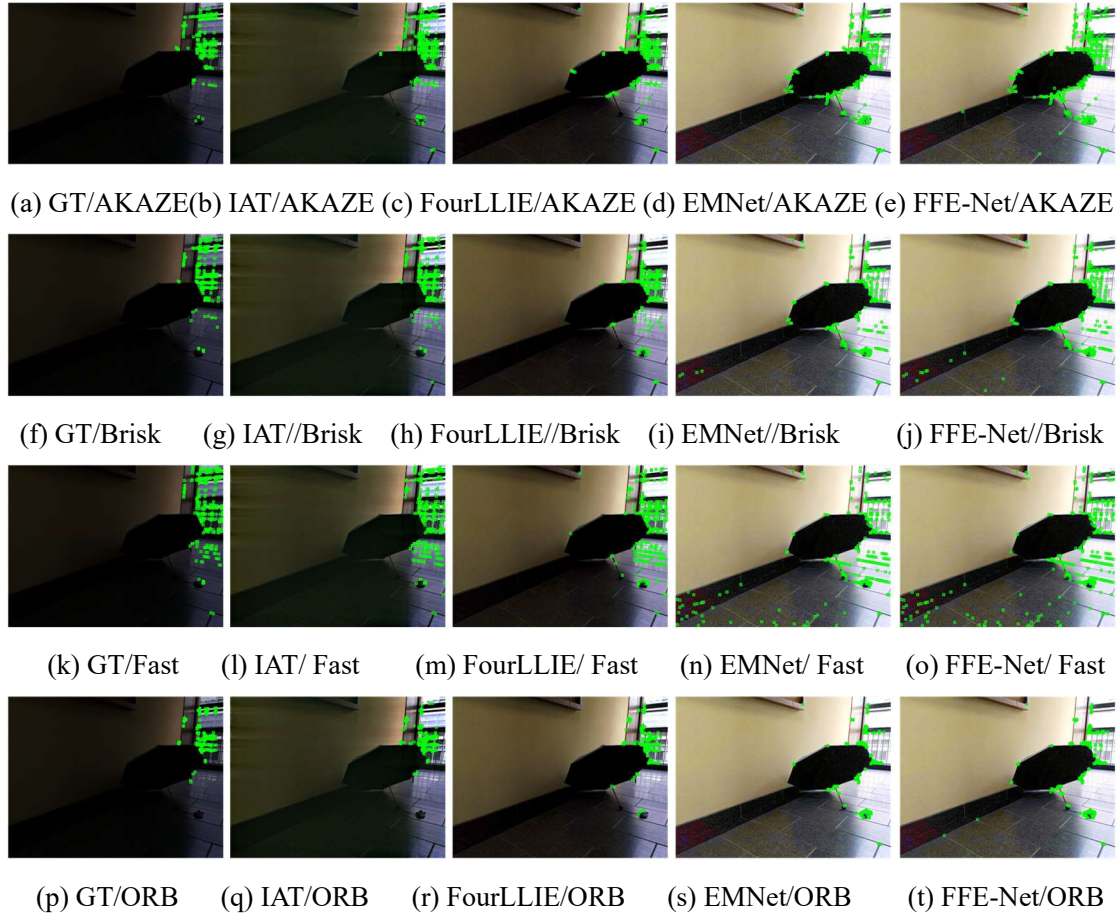


图 3-8 低光增强算法下特征提取效果对比图

由图 3-8 可知，原始低光照图像 GT 由于光照不足，存在细节丢失、纹理模糊以及边缘信息不明显的问题，导致特征点提取数量较少且分布不均匀，描述子也可能因光照不足而失真或不准确。相比之下，FFE-Net 算法在低光照区域表现尤为突出，能够显著增强图像的细节和纹理信息，恢复出更多的边缘和角点特征。实验结果表明，本文算法在 ORB、AKAZE、FAST 和 SIFT 四种特征点提取方法上均表现最优，尤其是在低光区域能够提取到更多的特征点，显著优于 IAT、FourLLIE 和 EMNet 算法。

本文所提算法 FFE-Net 增强后的图像在局部区域更加清晰，特征点数量增加且分布更加合理，更符合角点定义。相比之下，其他增强算法在低光区域的特征点提取能力较弱，导致特征点数量较少且分布不均匀。实验结果表明，FFE-Net 算法不仅能够有效提升低光照图像的视觉质量，还能减少因特征点数量不足或分布不合理导致的误差，增强机器人在低光照下的感知能力，为后续的特征匹配任务提供更可靠的数据支持。

3.4.3 特征匹配对比实验

为验证本文所提出的低光增强算法 FFE-Net 在 SLAM 系统中的有效性，分别选取公开数据集室外 KITTI 和室内 TUM_VI、Euroc 中的部分低光序列进行特征匹配对比实验，并将匹配结果在图 3-9 中进行展示，可以看到在三个数据集上均表现优异，且分布更为均衡，有效提升了 SLAM 系统在低光场景下特征匹配的鲁棒性。

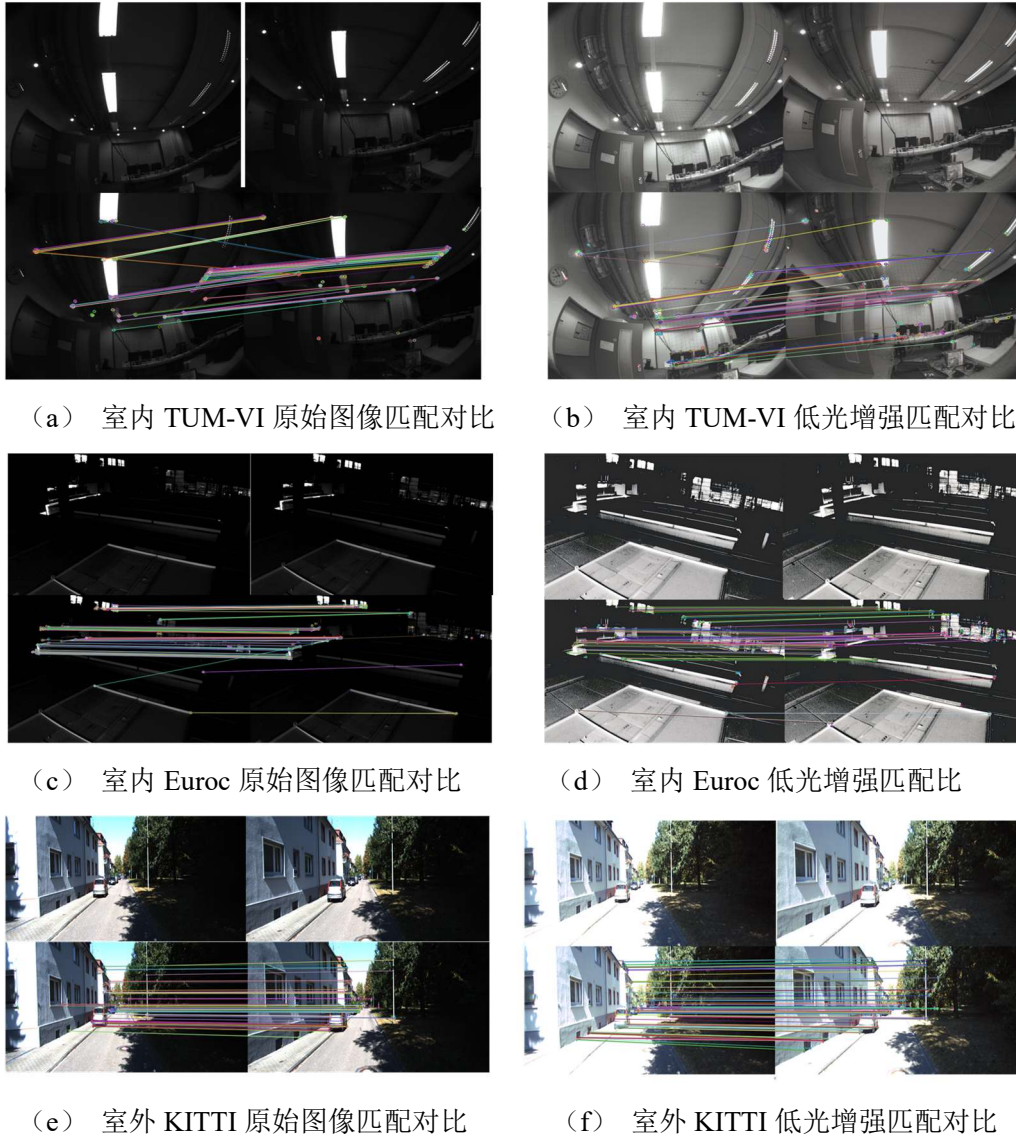


图 3-9 低光增强前后特征匹配对比图

实验分别选取三个数据集中 100 帧数据进行实验验证，平均特征匹配的具体数量如表 3-2 所示，低光增强后的特征匹配数量明显增加，在三个数据集中，平均匹配数量比原来提升 15.67%，显著增强了移动机器人在低光照条件下的环境感知能力，为后续 SLAM 系统位姿计算提供了可靠的数据支持。

表 3-2 不同数据集下匹配数量对比

匹配对数	原始图像	低光增强图像
KITTI	51	58
TUM-VI	69	82
Euroc	83	95

3.5 本章小结

针对低光照条件下视觉 SLAM 算法特征提取困难、跟踪易丢失的问题，本文提出了一种基于 FFE-Net 网络的低光图像增强算法，通过对频域和时域进行联合处理，共同优化进行图像增强。为验证算法有效性，研究选取了 SLAM 领域广泛使用的公开室外 KITTI 数据集以及室内 TUM_VI 和 Euroc 数据集中的部分低光图像序列进行实验验证，实验结果表明，所提算法显著提升了系统在低光照条件下的特征提取与匹配性能，有效降低了低光环境对视觉 SLAM 系统的影响，增强了移动机器人在复杂光照条件下的环境感知能力，为室内外低光环境下的机器人自主导航提供了可靠的技术支持。

第 4 章 基于双特征描述符的闭环检测算法

闭环检测是 SLAM 系统的核心之一，用于识别机器人是否重返历史访问区域，通过检测闭环触发全局优化来消除轨迹估计与地图构建过程中累积的误差，然而传统算法在低光照环境下存在闭环检测精度下降、误检率升高等问题，严重影响系统的实用性与鲁棒性。本文提出了一种基于特征描述符的闭环检测算法，通过优化闭环匹配策略，显著提升了系统在低光照条件下的检测精度与稳定性。

传统的闭环检测算法包含闭环候选关键帧的检索与验证，与现有的方法相比，本文的创新点主要集中在验证阶段的改进。传统方法通常依赖词袋模型，通过查询共享词汇来检索和验证闭环候选关键帧，这种方法在低光照环境下容易因特征信息不足而导致误检率升高。为此，本文提出了一种多层次特征描述符融合策略，通过结合全局特征描述符与局部特征描述符双重验证，显著增强了系统的闭环验证能力，通过引入更严格的筛选机制，确保了候选关键帧的质量，从而有效降低了低光照环境下错误回环的发生概率，算法的详细技术路线如图 4-1 所示。

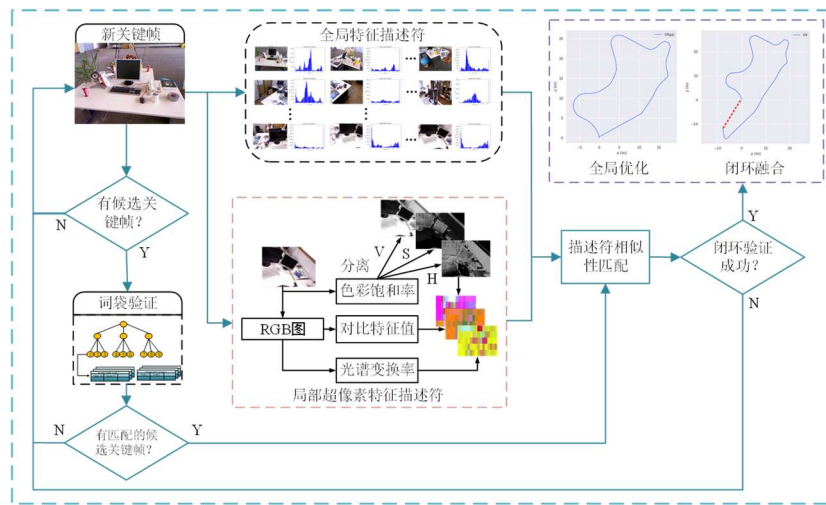


图 4-1 本章技术路线图

对于输入的关键帧，系统会生成其对应的特征描述符，随后进行闭环候选关键帧的检索，当检测到潜在的闭环候选时，系统首先通过词袋模型进行初步相似度验证，对于通过初步验证的候选关键帧，进一步计算其特征描述符进行相似度验证，应用更严格的判定条件来决定是否执行最终的回环融合与全局优化。这种分层验证机制显著提升了系统在低光照条件下的闭环检测可靠性，为 SLAM 系统在复杂环境下的稳定运行提供了有力保障，

4.1 全局特征描述符

基于 ORB-SLAM3 系统的框架，本文考虑计算图像相关图这一全局描述符来对闭环候选帧进行预剔除，通过高效的比较图像间的相似性，从而快速准确的对大量异常值进行消除，提高闭环检测筛选效率，增强 SLAM 系统性能，预剔除示例如图 4-2 所示。



图 4-2 预剔除示例图

图像相关图是一种统计工具，用于描述图像中颜色空间分布的相关性，它通过测量同一颜色像素在一定距离内出现的频率来表示图像的颜色和空间结构，相关图可以很容易地计算并以低成本进行比较且可以很好的提供关于图像颜色分布的全局特征。

图像相关图由一个表格组成，该表格包含在其像素位置之间指定距离内找到两个颜色值的概率，设 I 为一个包含 m 种颜色 $(C_1, \dots, C_m)$ 经过量化的 $r \times r$ 图像。 $I(p)$ 表示在坐标 $(x, y) \in I$ 的像素 p 的颜色 c 。设 $I_c = \{p | I(p) = c\}$ ，则 $p \in I_c$ 表示 $p \in I$ 且 $I(p) = c$ 。像素 $p_1 = (x_1, y_1)$ 和 $p_2 = (x_2, y_2)$ 之间的距离计算为 $|p_1 - p_2| = \max\{|x_1 - x_2|, |y_1 - y_2|\}$ 。表达式 $[r]$ 表示集合 $\{1, 2, 3, \dots, r\}$ 。对于设定的距离 $d \in [r]$ ，图像 I 的相关图定义其公式如 (4-1) 所示：

$$\gamma_{c_i, c_j}^{(t)}(I) = \Pr_{p_1 \in I_{c_i}, p_2 \in I} [p_2 \in I_{c_j} | |p_1 - p_2| = t] \quad (4-1)$$

其中 $i, j \in [m]$ 且 $t \in [d]$ ， $\gamma_{c_i, c_j}^{(t)}(I)$ 表示在距离 t 处从任意颜色为 c_i 的像素找到颜色为 c_j 的像素的概率。

自相关图是相关图的一个特例，它只考虑相同颜色值之间的空间相关性，图像 I 的自相关图定义其公式如 (4-2) 所示：

$$\alpha_c^{(t)}(I) = \gamma_{c, c}^{(t)}(I) \quad (4-2)$$

其中 c_i 和 c_j 颜色值相同，它描述了相同颜色在指定距离内出现的概率。

在所提出的系统中，自相关图被优先用于闭环检测描述符的预筛选，比标准相关图能够更高效地计算和比较图像相似度。系统将从传感器接收到的每一帧彩色图像转换为灰度图，并将其量化为 128 个灰度级，然后根据 1 像素距离计算该帧的自相关图，以保持鲁棒性和低计算成本，闭环候选时会根据当前帧和关键帧之间自相关图的相似性进行预筛选，计算自相关图之间相似性使用曼哈顿距离进行度量，曼哈顿距离的定义为两个图像自相关图对应值之间绝对差值的总和，设 I 和 I' 为将要比较的图像，自相关图之间的曼哈顿距离定义，其公式如（4-3）所示：

$$d_{\alpha M_{anhar tan}}(I, I') = \sum_{i \in [m], l \in [d]} \left| \alpha_{c_i}^{(t)}(I) - \alpha_{c_i}^{(t)}(I') \right| \quad (4-3)$$

自相关图描述了图像分布的空间相关性，可以用于高效的比较图像间相似性，进而用于快速准确的闭环检测预剔除。

4.2 局部超像素特征描述符

对于全局描述符预剔除后保留下来的关键帧，需要使用局部超像素特征描述符进行精细的匹配，由于 SLAM 系统在运行时会产生累计误差，导致同样的场景由于相机采集角度的不同而产生错误回环，为提升对角度的鲁棒性，对于关键帧 P ，选择其前后第 $P-3$ 和 $P+3$ 两张普通帧，共计三帧图像。针对每帧图像，分别划分成 $M \times N$ 的特征区域，并计算其单帧的超像素特征描述符，最终通过均值融合的方式得到关键帧 P 的均值描述符，并将其放入描述符库中，这一过程示例如图 4-3 所示。

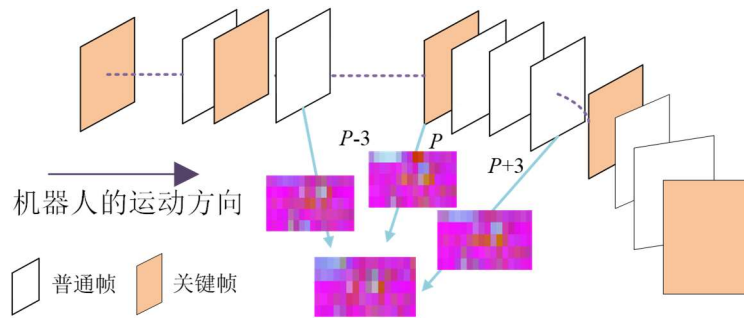


图 4-3 特征描述符生成示例

单帧特征描述符分别依据每个特征区域的光谱变换率，对比特征值以及色彩

饱和率生成，与一般处理所有像素点不同的是，提出一种中心线采样的方式，提取每个特征区域水平中心线和垂直中心线上的像素点，通过关注特征区域中心信息的方式来显著降低计算量，使其能更好的满足 SLAM 系统应用需求。

4.2.1 光谱变换率

光谱变换率是通过参考图像的灰度平均梯度生成，图像的灰度平均梯度 (Grayscale Mean Gradient, GMG) 是图像处理领域中常用的一个指标，用于评估图像的锐利度和清晰度，它可以反映图像像素之间的灰度变化程度，特别是边缘和细节的清晰度，可以通过计算图像上所有的像素灰度值与其邻域像素灰度值一阶差分的加权和得到，其计算公式如 (4-4) 所示：

$$Lt = \frac{\sqrt{2}}{2s} \left(\sum_{\theta=1}^{\beta-1} |f(\theta+1, m_d) - f(\theta, m_d+1)| + \sum_{\psi=1}^{\alpha-1} |f(n_d+1, \psi) - f(n_d, \psi+1)| \right) \quad (4-4)$$

其中， m_d 表示该特征区域水平中心线的纵坐标， n_d 表示该特征区域垂直中心线的横坐标， $f(\theta+1, m_d)$ 表示位于特征区域第 $\theta+1$ 行、 m_d 列的像素点灰度值， $f(n_d, \psi+1)$ 表示位于特征区域第 n_d 行、 $\psi+1$ 列的像素点灰度值， α 和 β 分别为像素点的行数和列数， s 为像素点个数， Lt 为该特征区域的光谱变换率。

4.2.2 对比特征值

图像平均对比度是衡量图像质量的一个关键参数，它直接影响到图像的清晰度和细节表现。对比度高的图像通常在视觉上显得更加鲜明和有层次感，使得不同物体之间的界限更加明显，在计算机视觉和图像处理领域，对比度也是评价算法效果的重要指标之一。图像平均对比度不仅与图像的平均亮度有关，还涉及到图像信息的视觉可见度，即图像细节的呈现和可辨识程度。

对比度可以从不同角度进行定义，其中常见的包括局部对比度和全局对比度。局部对比度侧重于图像小区域内的灰度变化，而全局对比度则关注整个图像的亮度范围。在实际应用中，这两种对比度的综合分析能够更全面地反映图像的质量。

同时对比度描述了在视觉感知中，不同灰度值在空间上的分布差异，它是图

像处理中计算平均对比度的一个重要概念，反映了目标和背景之间的灰度差异，从而影响人眼对图像的感知。其计算公式如（4-5）所示：

$$C_{simul} = |L_t - L_b| \quad (4-5)$$

其中， L_t 和 L_b 分别表示目标区域和背景区域的灰度值。

同时对比度计算方式简单直接，能够有效捕捉图像中显著的灰度变化，图像的平均对比度既与图像的平均亮度有关，也与图像信息的视觉可见度有关，在上式的基础上，计算一副图像的平均对比度 AC，其公式如（4-6）所示：

$$AC = \frac{1}{(M-1)*(N-1)} \sum_{y=0}^{N-2} \sum_{x=0}^{M-2} |Gray(x, y) - Gray(x+1, y)| \quad (4-6)$$

式中 $Gray(x, y)$ 为像素点 (x, y) 的灰度，M，N 为 x，y 方向的像素数。

对于彩色图像，需要分别计算红、绿、蓝三色分量的对比度，然后合成总对比度，其公式如（4-7）所示：

$$TAC = \left(\sqrt{RAC^2 + GAC^2 + BAC^2} \right) / \sqrt{3} \quad (4-7)$$

其中，RAC、GAC 和 BAC 分别为红、绿、蓝三分量的平均对比度，TAC 为总对比度。除以因子 3 是为了不超过数字图像处理中的灰度/色度的值域，以便作对比度图像。

当采用中心线采样的方式提取像素时，即可得到满足 SLAM 系统运行的对比特征值，用来表征每个语义区域的明暗变化情况，显示各部分的亮度差异。由于图像是通过混合 R、G、B 三个通道的变化与亮度来生成人类视觉可感知的颜色信息，因此 RGB 的三个通道分量均与亮度紧密相关，当亮度发生改变时，三个通道的分量会随之调整，进而影响整体的颜色表现，所以同时对 R、G、B 三通道进行中心线采样处理可以使其包含更稳定的特征对比信息，其公式如（4-8）所示：

$$Ce = \frac{1}{\sqrt{3}s^2} \sum_{c=1}^3 \left(\sum_{\psi_c=1}^{\alpha-1} |f(n_d, \psi_c) - f(n_d, \psi_c + 1)| + \sum_{\theta_c=1}^{\beta-1} |f(\theta_c, m_d) - f(\theta_c + 1, m_d)| \right)^2 \quad (4-8)$$

其中， Ce 表示对比特征值， ψ_c 表示在图像 c 通道中第 ψ 列， θ_c 表示在图像 c 通道中第 θ 行， $f(n_d, \psi_c)$ 表示在图像 c 通道中第 n_d 行、 ψ 列的像素点灰度值，

$f(\theta_c, m_d)$ 表示在图像 c 通道中第 θ 行, m_d 列的像素点灰度值。

4.2.3 色彩饱和度

在图像处理与目标检测任务中,光照条件的变化会显著影响检测效果,RGB 图像通过 R、G、B 三个通道的变化与亮度的叠加来表示颜色信息,由于 RGB 三个通道的分量与亮度密切相关,当光照发生变化时,这些分量也会随之调整,从而导致整体颜色信息的改变,这使得 RGB 图像对光照变化较为敏感。

为了降低光照对图像的影响,可以采用 HSV 颜色空间。HSV 模型由色调 (Hue)、饱和度 (Saturation) 和明度 (Value) 三个分量组成。色调 H 将所有颜色映射到 0° 至 360° 的范围内,与 RGB 的三通道配合不同, H 通道只需选择适当的角度即可。饱和度 S 表示颜色接近光谱色的程度,光谱色占比越高,饱和度越高。明度 V 表示颜色的亮度,反映了外界光照的影响,HSV 可以通过 RGB 转换得到,其公式如 (4-9)、(4-10) 和 (4-11) 所示:

$$V \leftarrow \max(R, G, B) \quad (4-9)$$

$$S \leftarrow \begin{cases} \frac{V - \min(R, G, B)}{V}, & V \neq 0 \\ 0, & \text{其他} \end{cases} \quad (4-10)$$

$$H \leftarrow \begin{cases} \frac{60(G-B)}{V - \min(R, G, B)}, & V = R \\ 120 + \frac{60(B-R)}{V - \min(R, G, B)}, & V = G \\ 240 + \frac{60(R-G)}{V - \min(R, G, B)}, & V = B \\ H = H + 360, & H < 0 \end{cases} \quad (4-11)$$

其中, R、G、B 分别代表 RGB 图像的三个通道分量, H、S、V 代表 HSV 图像的三个通道分量。

从上述公式可知, V 通道的值来源于 RGB 图像中的最大分量,因此光照对 HSV 图像的 V 通道影响最为显著,而 H 通道和 S 通道则是通过提取 RGB 图像中的最小分量并结合差值和运算生成的,因此光照对 H 通道和 S 通道的影响较小,色彩饱和度通过从 HSV 颜色空间中分离出抗光照能力较强的 S 通道,并利用中心线采样获取更加稳定的信息,其公式如 (4-12) 所示:

$$Sr = \frac{1}{s} \left(\sum_{\theta=1}^{\beta} \frac{\max(R_C, G_C, B_C) - \min(R_C, G_C, B_C)}{\max(R_C, G_C, B_C)} + \sum_{\psi=1}^{\alpha} \frac{\max(R_C, G_C, B_C) - \min(R_C, G_C, B_C)}{\max(R_C, G_C, B_C)} \right) \quad (4-12)$$

其中, Sr 表示色彩饱和度, RC 、 GC 、 BC 分别代表 RGB 图像的三个通道分量。

4.3 双特征描述符闭环候选帧匹配算法

基于光谱变换率、对比特征值以及色彩饱和度, 每个特征区域的局部超像素特征描述符, 其公式如 (4-13) 所示:

$$T_{ij} = e^{\omega_1 \times Lt} \times \log(\omega_2 \times Ce) + \frac{1}{1 + e^{-\omega_3 \times Sr}} \quad (4-13)$$

其中, T_{ij} 表示图像中第 i 行第 j 列的特征区域的描述符, ω_1 、 ω_2 、 ω_3 分别表示各参数权重。

计算各个特征区域的描述符 T_{ij} 形成 $a \times b$ 的描述符矩阵, 将得到的描述符矩阵行展开成一个向量, 得到描述符向量 T 。对任一关键帧 P , 分别计算其前后第 $P-3$ 和 $P+3$ 两张普通帧, 共计三帧图像的描述符向量, 通过求取均值的方式获得关键帧 P 的均值描述符, 并将其放入描述符库中。采用余弦相似度进行相似度度量, 并将其归一化到 $[0,1]$ 之间, 值越接近 1 代表相似度越高。相似度 $Similarity$ 其公式如 (4-14) 所示:

$$Similarity = \frac{1}{2} + \frac{T_1 \cdot T_2}{2 \|T_1\| \cdot \|T_2\|} \quad (4-14)$$

其中, T_1 和 T_2 分别表示当前关键帧和闭环候选关键帧的均值描述符向量, $\|\cdot\|$ 表示求取向量的模。

对于输入的关键帧, 会生成它的全局特征描述符和局部超像素特征描述符, 同时判断其是否存在闭环候选关键帧, 当存在闭环候选关键帧时, 会通过词袋验证相似度, 对于有匹配的候选关键帧, 使用全局描述符进行快速的匹配计算实现预剔除, 提高筛选效率, 并对于剔除过后保留下来的关键帧进行更精确的局部超像素特征描述符的匹配, 通过更严格的闭环条件来判定是否可以进行最终的回环融合和全局优化, 使其更适用于处理在低光照条件下出现错误回环的问题, 改进的闭环检测流程如图 4-4 所示。

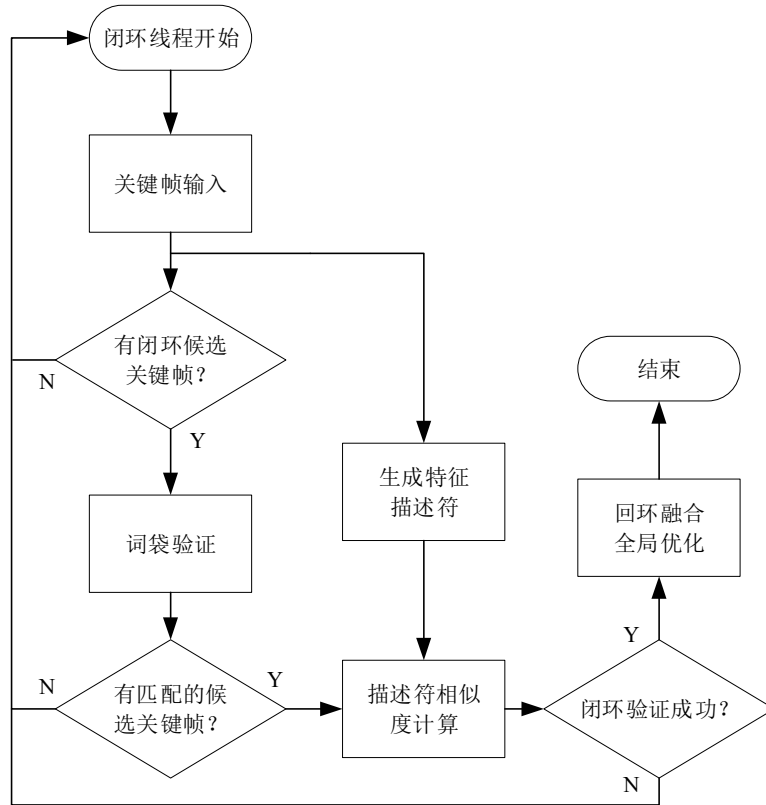


图 4-4 闭环检测流程图

4.4 公开数据集实验及分析

为了验证本章算法的有效性,使用公开室外 KITTI 和室内 TUM_VI、EUROC 数据集中的部分低光序列进行验证。

实验分为四个部分:准确率召回率实验、轨迹对比实验、误差对比实验和真实场景实验,针对不同部分分别与 ORB-SLAM3、VINS-FUSION、Maplab 以及 OV2-SLAM 算法进行了对比。

4.4.1 准确率召回率实验

为了评估闭环检测算法的性能,通常采用准确率和召回率作为核心评价指标。准确率衡量了算法检测到的闭环中真实闭环所占的比例,反映了检测结果的可靠性,召回率则表征了在所有真实闭环中被成功检测出的比例,体现了算法的完整性,两者共同为闭环检测性能提供了全面的量化评估。

在实际应用中,闭环检测系统通常需要在准确率与召回率之间取得平衡,然而这两者往往难以同时达到最优,因为提高召回率往往需要放宽检测条件,从而

引入更多误匹配，导致准确率下降，反之，过于严格的匹配标准虽然能提高准确率，却会漏检许多真实闭环，降低召回率。因此为了全面评估算法性能，通常采用准确率-召回率曲线（Precision-Recall Curve, PR 曲线）作为评价工具。该曲线以召回率为横轴、准确率为纵轴，通过曲线的分布直观反映系统性能，曲线越靠近右上方，表明系统的检测性能越优，PR 曲线为闭环检测算法的优化与比较提供了重要的量化依据。

图 4-5 展示了本文算法与 Maplab、VINS-FUSION 算法在 EUROC 数据集和 TUM-VI 数据集上关于准确率和召回率的对比结果。由图可知，在 100%召回率下时，在 EUROC 数据集中，VINS-FUSION 的准确率为 76.6%，Maplab 的准确率为 90.6%，而本文算法的准确率为 95.7%，闭环精度分别提升 19.1%和 5.1%；在 TUM_VI 数据集中，VINS-FUSION 的准确率为 92.5%，Maplab 的准确率为 85.0%，而本文算法的准确率为 97.5%，闭环精度分别提升 5.0%和 12.5%。通过改进闭环检测算法，本文算法成功剔除了错误的回环，获得了更准确的闭环候选关键帧，从而使闭环位置的确定更为精确。

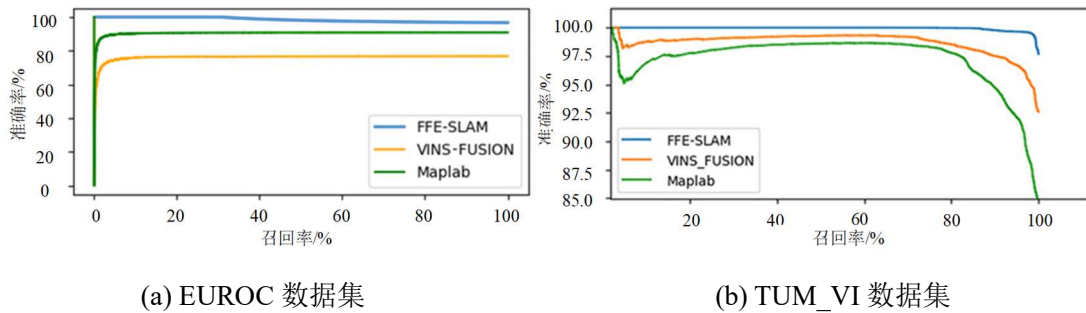


图 4-5 准确率召回率曲线对比

4.4.2 室外数据集对比实验

在本节实验中，使用室外 KITTI 数据集 00-09 总共 10 个序列验证本文算法性能，这些序列涵盖了城市道路、高速公路等多种复杂场景，能够全面评估算法的性能。实验采用 EVO 轨迹精度评估工具进行定量分析，评估结果如图 4-6 所示，由图可见在闭环处本文算法更贴近真实轨迹，表现出更优的位姿修正能力。为了更好的反映实验结果，将 10 个序列的误差 RMSE 值和 MEAN 值分别绘制成折线图，如图 4-7 所示所提算法在 10 个序列中误差普遍低于其他算法，充分证明了在大规模户外环境中的优越性。

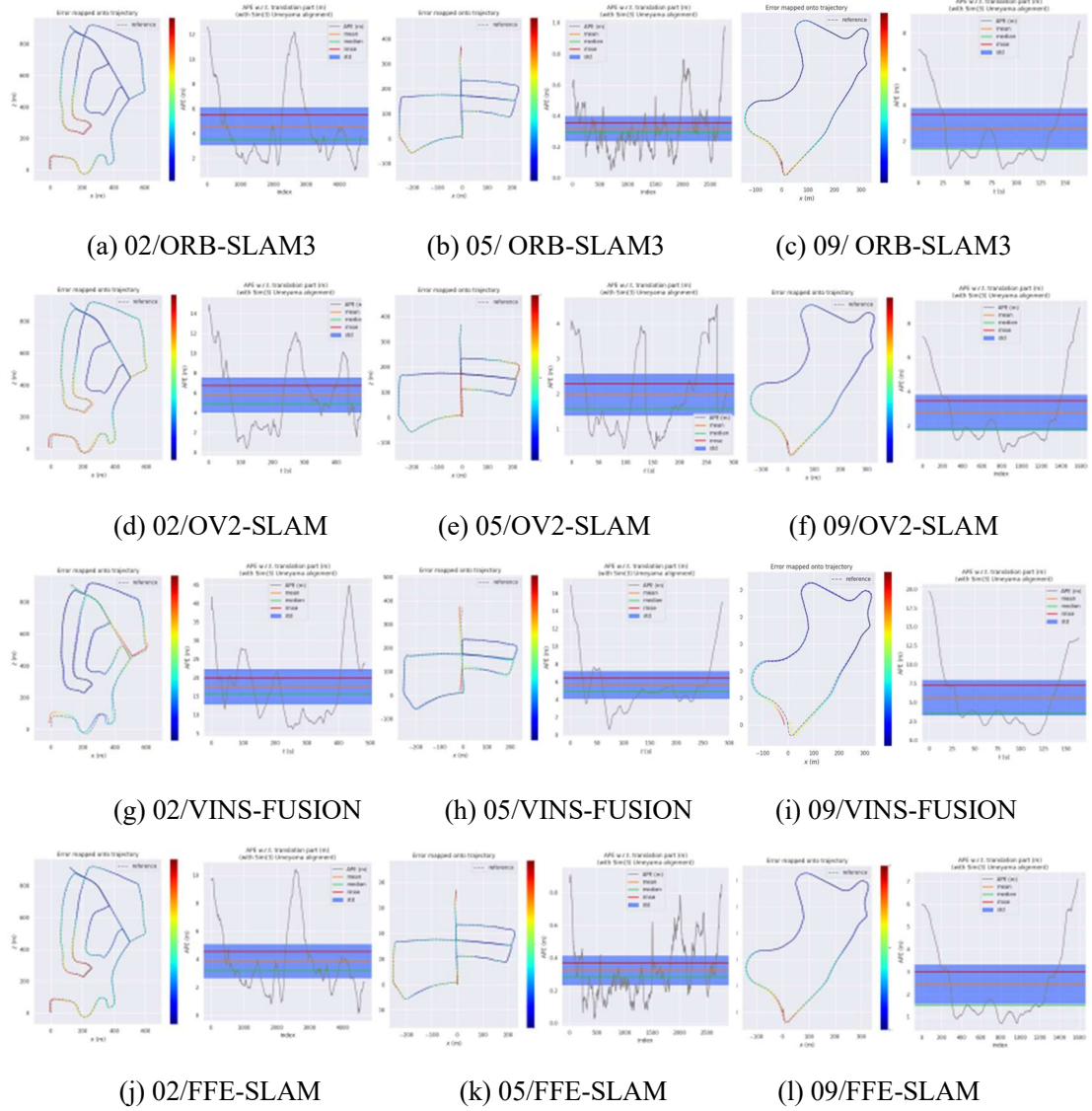


图 4-6 KITTl 数据集轨迹对比

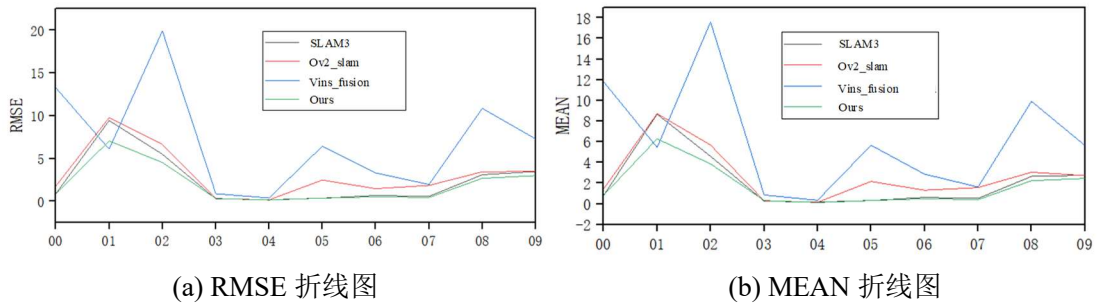


图 4-7 KITTl 数据集误差折线图（单位：米）

4.4.3 室内数据集对比实验

实验采用室内公开数据集 TUM_VI 和 EUROC 中光照较低的序列 room1~room4 和 MH01~MH04 进行测试，将绝对轨迹误差（Absolute Trajectory

Error, ATE) 作为对比指标, 评测工具为数据集自带的 `evaluate_atc.py` 评测脚本。

图 4-8 显示了 TUM_VI 数据集中 room1~room4 的误差图, 其中黑色线条代表真实轨迹, 蓝色线条代表不同算法生成的轨迹, 红色线条代表测量误差。由图可见, VINS-FUSION 算法由于在低光区域中难以提取足够的特征信息, 其轨迹误差较大, Maplab 以及 ORB-SLAM3 算法虽然可以提取基本的特征进行轨迹定位, 但由于闭环修正效果较差, 影响了全局优化后的定位轨迹。

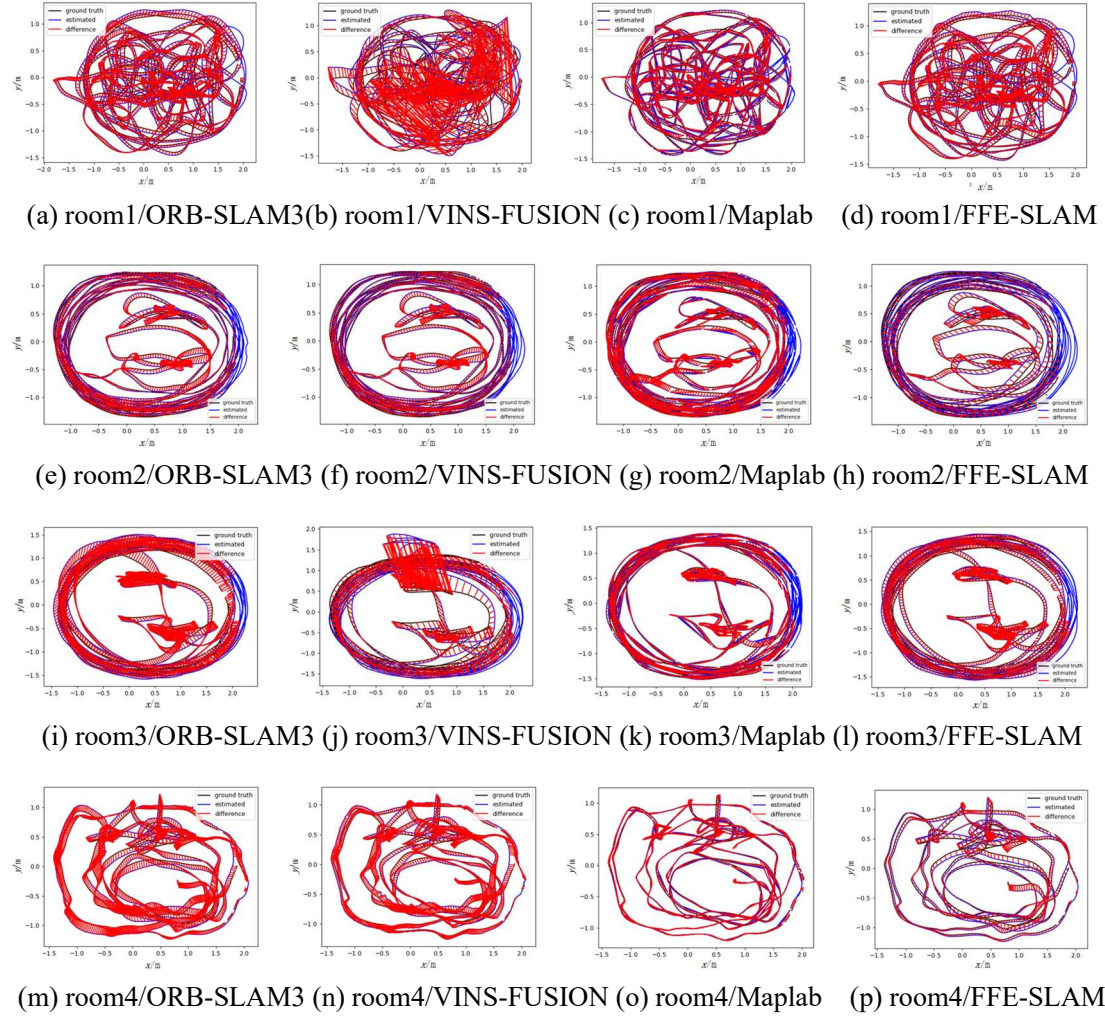


图 4-8 TUM-VI 轨迹对比

实验所测的详细轨迹误差对比数据如表 4-1 所示, 可以看出: 在 TUM_VI 数据集中, 相较于 ORB-SLAM3、VINS-FUSION 和 Maplab 算法, 本文算法的平均绝对轨迹误差分别减少了 27.16%、56.05%和 47.75%, 在 EUROC 数据集中, 相较于 OV2-SLAM 算法, 本文算法的平均绝对轨迹误差减少了 24.17%, 与上述三种算法相比, 本文算法能够提取到更多的特征信息以及特征描述符, 用于匹配验证和位姿计算, 从而使系统整体定位误差达到最小。

表 4-1 不同数据集绝对轨迹误差对比（单位：米）

数据集		ORB-SLAM3 ^[47]	VINS-FUSION ^[48]	Maplab ^[49]	OV2-SLAM ^[50]	FFE-SLAM
TUM_VI	room1	0.079	0.475	0.147	-	0.074
	room2	0.065	0.087	0.157	-	0.052
	room3	0.170	0.335	0.142	-	0.075
	room4	0.087	0.082	0.088	-	0.064
EUROC	MH01	0.040	0.135	0.233	0.053	0.034
	MH02	0.050	0.045	0.147	0.048	0.041
	MH03	0.052	0.197	0.299	0.052	0.043
	MH04	0.103	0.370	0.284	0.076	0.054

4.5 真实场景实验

为检验所提方法在实际环境中的适用性，研究采用移动机器人平台在低光条件下开展性能测试。测试系统基于 Clearpath Robotics 公司研发的 Husky 移动机器人构建，该平台配置了 Intel 第 12 代 i5-12400F 处理器，配备 16GB 运行内存，操作系统采用 Ubuntu 18.04 版本。实验系统集成了 D435i 视觉感知模块，其具体硬件参数详见表 4-2，机器人本体平面图如图 4-9 所示。

表 4-2 主要参数设置

变量名称	参数	数值
最大速度	V	1m/s
重量	kg	50kg
最大坡度	θ	$\pi/4$
相机采样频率	H	30fps

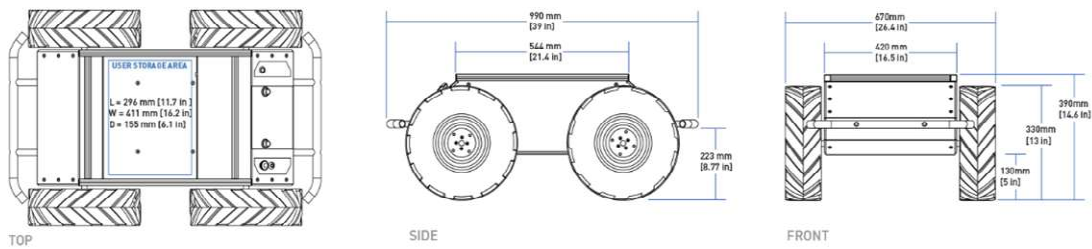


图 4-9 Husky 平面图

真实场景实验分别选取室内与室外场景进行实验，将本文算法 FFE-SLAM 与 VINS-FUSION、ORB-SLAM3 进行对比。室内实验运行环境及运行路径如图 4-10 所示，图 4-11 为其对应不同算法的轨迹对比，如图 4-11（a）所示，由于低光环境的影响，VINS-FUSION 算法的部分低光区域出现特征提取和匹配失败的问题，这会导致跟踪丢失，进而使定位精度产生较大误差，ORB-SLAM3 算法虽

然并未出现跟踪丢失，但在低光场景下进行闭环修正时，未能精准的检测到闭环位置，导致对极几何计算明显失真，闭环修正后全局地图尺度偏大，偏离真实轨迹，图 4-11(b)显示了具体的误差数据，由图可见本文算法的各种误差均为最小。

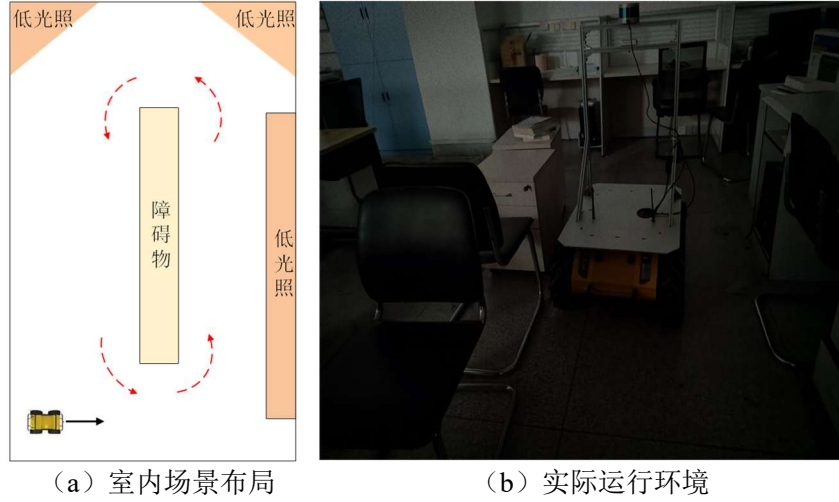


图 4-10 实验机器人室内真实场景实验

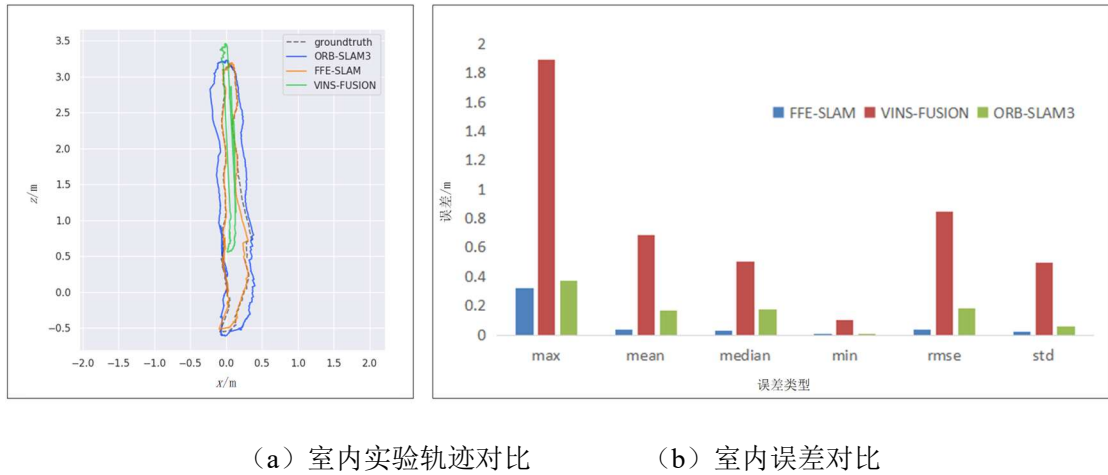


图 4-11 室内不同算法轨迹对比分析

在夜晚的校园环境，通过移动机器人平台开展真实场景实验，实验路径如图 4-12(a)所示，(b)、(c)分别为同一帧图像在本文算法与原始算法上的特征提取对比。结果表明，本文算法提升了低光照条件下的图像质量，解决了传统方法在低照度环境下特征点提取数量不足的问题，使特征点分布更加均匀。这为 SLAM 系统提供了更可靠的环境感知信息，从而增强了移动机器人在低照度环境下的感知能力。

图 4-13 展示了特征匹配性能的对比结果。实验表明，通过本文提出的低光增强算法处理后的图像，其细节信息与边缘特征得到显著增强，从而大幅提升了特征点提取的数量与分布质量，图像序列中相邻帧之间的匹配点数量显著增加，

进而有效提高了系统位姿估计的精度与鲁棒性，为视觉 SLAM 系统在低光环境下的稳定运行提供了重要保障。

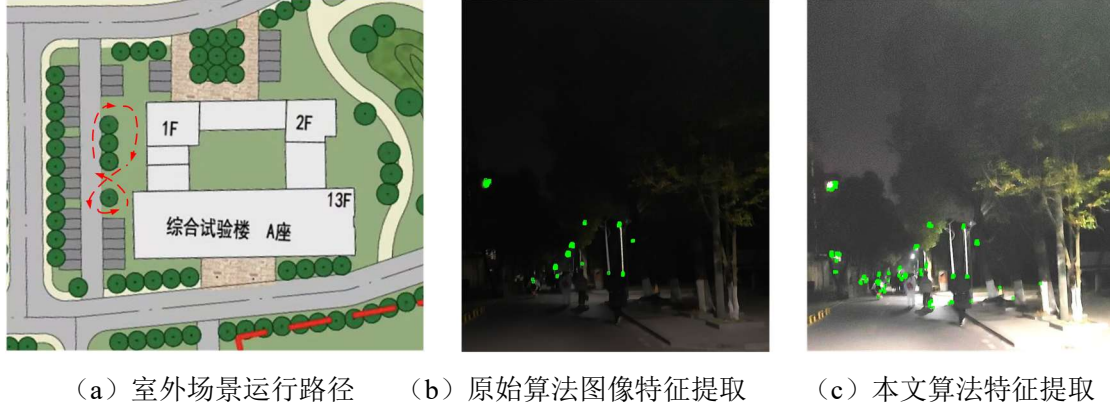


图 4-12 实验机器人室外真实场景实验

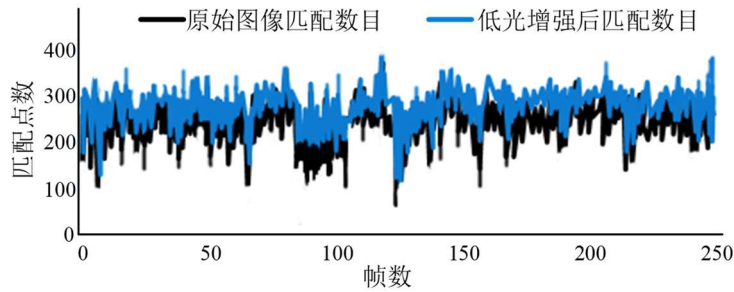
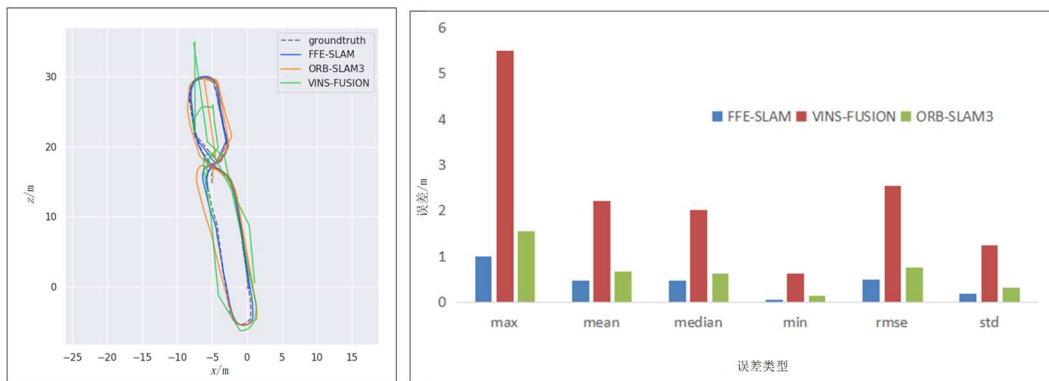


图 4-13 低光增强前后匹配数量对比图

图 4-14 为室外场景下不同算法的轨迹对比结果，与室内实验类似，受低光环境影响，VINS-FUSION 与 ORB-SLAM3 算法在不同区域出现特征点跟踪丢失或未能准确检测到闭环位置，精准修正全局轨迹的问题。图 4-14(b)可见，本文算法在各项误差指标上均表现最优，其轨迹最接近真实值。通过结合低光增强网络和改进的闭环检测算法，本文算法在实际场景中证明了其能在低光环境下精确定位与高精度建图的能力。



(a) 室外实验轨迹对比

(b) 室外误差对比

图 4-14 室外不同算法轨迹与误差对比

4.6 本章小结

针对 SLAM 系统在低光环境下闭环检测线程中存在的闭环精度不足、易产生错误回环的问题，提出了一种基于特征描述符的闭环检测改进算法。通过引入全局特征描述符对闭环候选帧进行预剔除，并结合局部特征描述符进行精细匹配，构建了双重筛选机制，有效提升了闭环检测的准确率与 SLAM 系统整体定位精度。在实验部分，采用 KITTI、TUM_VI 和 EUROC 等公开数据集，分别进行了室外与室内实验，从准确率召回率、轨迹精度和误差分布等多个维度对所提算法进行了系统性评估，并与 ORB-SLAM3、VINS-FUSION、Maplab 和 OV2-SLAM 等主流方法进行了对比分析。此外，搭建了硬件实验平台，在真实低光场景中验证了所提方法的实际应用价值。实验结果表明，本文算法在低光条件下显著提升了闭环检测的精度与系统的整体定位性能，为低光视觉 SLAM 算法的研究提供了重要的参考依据。

第 5 章 总结与展望

5.1 全文总结

SLAM 技术在智能移动机器人和自动驾驶等应用中发挥着重要作用, 尽管近年来视觉 SLAM 技术取得了显著进展, 但其性能仍依赖于光照条件等环境因素的影响, 在低光照场景下往往面临定位精度下降和地图构建效果不佳的问题。为应对移动机器人在低光照环境下的精准定位和地图构建需求, 本文提出一种低光条件下基于 FFE-Net 网络的视觉 SLAM 算法, 通过联合图像增强和双重特征描述符闭环检测策略, 有效解决了低光照场景下特征提取困难、跟踪丢失以及闭环检测精度不足的问题。在公开数据集以及真实场景中验证, 并与多种主流算法对比, 本文算法显著改善了低光照导致的特征点提取较少、匹配困难以及闭环误差增大等问题, 提升了系统的定位精度和鲁棒性。

论文主要研究内容与创新点概括如下:

(1) 针对低光照条件下视觉 SLAM 算法特征提取困难、跟踪易丢失的问题, 本文构建了一种基于 FFE-Net 网络的低光图像增强算法。在亮度增强模块中, 利用傅里叶空间中振幅分量与亮度的正相关特性, 通过放大低光图像的振幅分量实现亮度提升, 同时保留相位分量以确保图像结构完整性。在特征细节恢复模块中, 采用频域分解将图像分离为高低频分量, 通过深层特征提取与跨层特征融合实现细节修复, 并引入通道注意力和像素注意力机制动态优化特征权重提升图像细节恢复效果。实验结果表明, 该方法在 KITTI、TUM_VI 和 Euroc 数据集上使特征点匹配数量平均提升 15.67%, 显著提高了低光环境下视觉 SLAM 系统的鲁棒性, 为机器人自主导航提供了可靠技术支持。

(2) 针对低光照条件下提取的特征信息表征能力不足的问题, 本文提出了一种基于前端 FFE-Net 增强图像的局部超像素特征描述符构建方法。为提升描述符对角度变化的鲁棒性, 采用多帧联合构建时空关联组的生成策略, 针对系统运行过程中的某一关键帧 P , 选择其前后第 $P-3$ 和 $P+3$ 两张普通帧, 共计三帧图片通过均值融合的方式共同生成描述符。同时为降低计算复杂度, 对每帧图像分别划分成 $M \times N$ 的特征区域, 并通过中心线采样的方式分别提取特征区域水平与

垂直中心线上的像素点，计算光谱变换率、对比特征值以及色彩饱和率生成局部超像素描述符。该方法在减少计算量的同时，有效提升了算法对视角变化的稳定性，更适合 SLAM 系统的运行需求。

(3) 针对低光照条件下闭环检测精度不高且易产生错误回环的问题，本文基于 FFE-Net 增强图像与局部超像素特征描述符设计了一种新的闭环检测算法。为提高筛选效率并降低计算复杂度，在传统词袋模型验证之后，通过构建图像自相关图这一全局特征描述符对异常帧快速过滤进行预剔除，之后利用局部超像素描述符进行相似性度量，完成精细匹配过程。实验结果表明，在 EUROC 数据集中，本文算法闭环精度相较于 VINS-FUSION 和 Maplab 分别提升了 19.1% 和 5.1%，在 TUM_VI 数据集中，分别提升了 5.0% 和 12.5%，该算法通过对候选回环帧的双重筛选，降低了错误回环的发生概率，有效提升了闭环检测的精度与鲁棒性，为低光照视觉 SLAM 算法的研究提供了重要的参考依据。

5.2 研究展望

本文围绕“低光条件下基于 FFE-Net 网络的视觉 SLAM 算法”展开研究，通过公开数据集和真实低光照场景实验验证本文所提算法，实验表明，所提算法在低光照环境下具有显著的特征提取能力和较高的闭环检测精度。然而，研究成果仍有进一步完善的空间，主要分为以下几个方向：

(1) 本文重点研究了低光照环境下图像的特征提取与闭环检测部分，但对后端优化和地图构建环节的研究较为有限。后续工作可进一步探索基于语义信息的地图构建方法，结合深度学习技术提取环境中的语义特征，构建更具信息量的语义地图，从而提升移动机器人在复杂环境中的导航能力和场景理解能力。

(2) 本文主要基于纯视觉 SLAM 系统进行研究，但在纹理稀疏的场景中，纯视觉传感器可能无法获取足够的特征信息。未来可考虑引入多传感器融合技术，结合激光雷达、IMU 等传感器，或引入线、面等几何特征，丰富数据来源，避免因特征不足导致的系统失效问题，进一步提升算法的鲁棒性和适用性。

参考文献

- [1] Macario Barros A, Michel M, Moline Y, et al. A Comprehensive Survey of Visual Slam Algorithms[J]. Robotics, 2022, 11(1): 24.
- [2] 刘丰宇,程向红,曹毅.基于深度学习与特征点速度约束的室内动态 SLAM 方法[J].中国惯性技术学报,2023,31(05):438-443.
- [3] 王柯赛,姚锡凡,黄宇等.动态环境下的视觉 SLAM 研究评述[J].机器人,2021,43(06):715-732.
- [4] Zhang S, Zhi Y, Lu S, et al. Monocular Vision SLAM Research for Parking Environment with Low Light[J]. International Journal of Automotive Technology, 2022, 23(3): 693-703.
- [5] Chen X, Yu Y. An Unsupervised low-light Image Enhancement Method for Improving V-SLAM Localization in uneven low-light Construction sites[J]. Automation in Construction, 2024, 162: 105404.
- [6] Quan Y, Fu D, Chang Y, et al. 3D Convolutional Neural Network for low-light Image Sequence Enhancement in Slam[J]. Remote Sensing, 2022, 14(16): 3985.
- [7] Lore K G, Akintayo A, Sarkar S. LLNet: A Deep Autoencoder Approach to Natural low-light Image Enhancement[J]. Pattern Recognition, 2017, 61: 650-662.
- [8] Lv F, Lu F, Wu J, et al. MBLLEN: Low-light image/video Enhancement using cnns[C]. Bmvc. 2018, 220(1): 4.
- [9] Kuang B, Zhang Z. Two-stage Low-light Image Enhancement Network with an attention Mechanism and cross-stage connection[J]. Journal of Electronic Imaging, 2022, 31(5): 053001-053001.
- [10] Wang C, Wu H, Jin Z. Fourllie: Boosting Low-light Image Enhancement by Fourier Frequency Information[C]//Proceedings of the 31st ACM International Conference on Multimedia. 2023: 7459-7469.
- [11] Shi C, Chen X, Xiao J, et al. Fast and Accurate Deep Loop Closing and Relocalization for Reliable Lidar slam[J]. IEEE Transactions on Robotics, 2024, 40: 2620-2640.
- [12] Liso L, Sandström E, Yugay V, et al. Loopy-slam: Dense Neural Slam with loop Closures[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024: 20363-20373.
- [13] Gao X, Zhang T. Loop Closure Detection for Visual Slam Systems using Deep

- neural Networks[C]//2015 34th Chinese control conference (CCC). IEEE, 2015: 5851-5856.
- [14]Zhang Q, Kim J. Shape BoW: Generalized Bag of Words for Appearance-based loop Closure Detection in bathymetric SLAM[J]. IEEE Robotics and Automation Letters, 2024, 9(9): 7405-7412.
- [15]Xi K, He J, Hao S, et al. SLAM loop detection algorithm based on improved bag-of-words model[C]//2022 5th international conference on pattern recognition and artificial intelligence (PRAI). IEEE, 2022: 683-689.
- [16]Chen B, Yuan D, Liu C, et al. Loop Closure Detection Based on Multi-scale Deep Feature Fusion[J]. Applied Sciences, 2019, 9(6): 1120.
- [17]Cui Y, Chen X, Zhang Y, et al. Bow3d: Bag of Words for real-time Loop Closing in 3d lidar slam[J]. IEEE Robotics and Automation Letters, 2022, 8(5): 2828-2835..
- [18]Zhang Q, Xu G, Li N. Improved SLAM closed-loop detection algorithm based on DBoW2[C]//Journal of Physics: Conference Series. IOP Publishing, 2019, 1345(4): 042094.
- [19]陈孟元,钱润邦,郭行荣等.动态场景下基于实例分割与运动一致性约束的 VSLAM 算法[J].中国惯性技术学报,2023,31(10):986-995.
- [20]陈孟元,李鹏飞,符乙,等.低光条件下结合 FFE-Net 网络的视觉 SLAM 算法[J].中国惯性技术学报,2025,33(01):36-45.
- [21]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.
- [22]田野,陈宏巍,王法胜等.室内移动机器人的 SLAM 算法综述[J].计算机科学,2021,48(09):223-234.
- [23]Min Z, Dunn E. Voldor+ slam: For the times when feature-based or Direct methods are not good enough[C]//2021 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2021: 13813-13819.
- [24]Dong X, Cheng L, Peng H, et al. FSD-SLAM: a fast semi-direct SLAM algorithm[J]. Complex & Intelligent Systems, 2022, 8(3): 1823-1834.
- [25]Cai L, Ye Y, Gao X, et al. An improved visual SLAM based on affine transformation for ORB feature extraction[J]. Optik, 2021, 227: 165421.
- [26]Liu X, Zhao X, Xia Z, et al. Secure outsourced SIFT: Accurate and efficient privacy-preserving image SIFT feature extraction[J]. IEEE Transactions on Image

- Processing, 2023, 32: 4635-4648.
- [27]Chen S, Zhou B, Jiang C, et al. A lidar/visual slam backend with loop closure detection and graph optimization[J]. Remote sensing, 2021, 13(14): 2720.
- [28]Alzubaidi L, Zhang J, Humaidi A J, et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions[J]. Journal of big Data, 2021, 8: 1-74.
- [29]Taye M M. Theoretical understanding of convolutional neural network: Concepts, architectures, applications, future directions[J]. Computation, 2023, 11(3): 52.
- [30]Quan Y, Fu D, Chang Y, et al. 3D convolutional neural network for low-light image sequence enhancement in slam[J]. Remote Sensing, 2022, 14(16): 3985.
- [31]Tang H, Tang Y, Su Y, et al. Feature extraction of multi-sensors for early bearing fault diagnosis using deep learning based on minimum unscented kalman filter[J]. Engineering Applications of Artificial Intelligence, 2024, 127: 107138.
- [32]Neupane D, Seok J. A review on deep learning-based approaches for automatic sonar target recognition[J]. Electronics, 2020, 9(11): 1972.
- [33]Qu C, Guo Z, Xia S, et al. Intention recognition of aerial target based on deep learning[J]. Evolutionary Intelligence, 2024, 17(1): 303-311.
- [34]陈孟元,韩朋朋,刘金辉等.动态遮挡场景下基于改进 Transformer 实例分割的 VSLAM 算法[J].电子学报,2023,51(07):1812-1825.
- [35]Gupta M, Mishra A. A systematic review of deep learning based image segmentation to detect polyp[J]. Artificial Intelligence Review, 2024, 57(1): 7.
- [36]Li S, Song W, Fang L, et al. Deep learning for hyperspectral image classification: An overview[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(9): 6690-6709.
- [37]Kumar R, Kumbharkar P, Vanam S, et al. Medical images classification using deep learning: a survey[J]. Multimedia Tools and Applications, 2024, 83(7): 19683-19728.
- [38]Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 7132-7141.
- [39]Zhang Y, Li K, Li K, et al. Image super-resolution using very deep residual channel attention networks[C]//Proceedings of the European Conference on Computer Vision (ECCV). 2018: 286-301.
- [40]Wang X, Girshick R, Gupta A, et al. Non-local neural networks[C]//Proceedings of

- the IEEE conference on computer vision and pattern recognition. 2018: 7794-7803.
- [41]Sumiya Y, Otsuka T, Maeda Y, et al. Gaussian fourier pyramid for local laplacian filter[J]. IEEE Signal Processing Letters, 2021, 29: 11-15.
- [42]Kurinjimalar R, Pradeep J, Harikrishnan M. Underwater Image Enhancement Using Gaussian Pyramid, Laplacian Pyramid and Contrast Limited Adaptive Histogram Equalization[C]//2024 IEEE 3rd World Conference on Applied Intelligence and Computing (AIC). IEEE, 2024: 729-734.
- [43]Luo X, Fu G, Yang J, et al. Multi-modal image fusion via deep Laplacian pyramid hybrid network[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2023, 33(12): 7354-7369.
- [44]Wang J, Duan Y, Tao X, et al. Semantic perceptual image compression with a Laplacian pyramid of Convolutional networks[J]. IEEE Transactions on Image Processing, 2021, 30(7): 4225-4237.
- [45]Liang J, Zeng H, Zhang L. High-resolution photorealistic image translation in real-time: A laplacian pyramid Translation network[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 9392-9400.
- [46]Cui Z, Li K, Gu L, et al. You only need 90k parameters to adapt light: a light weight transformer for image enhancement and exposure correction[J]. arXiv preprint arXiv:2205.14871, 2022.
- [47]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.
- [48]Qin T, Shen S. Online temporal calibration for monocular visual-inertial systems[C]//2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2018: 3662-3669.
- [49]Schneider T, Dymczyk M, Fehr M, et al. maplab: An open framework for research in visual-inertial mapping and localization[J]. IEEE Robotics and Automation Letters, 2018, 3(3): 1418-1425.
- [50]Ferrera M, Eudes A, Moras J, et al. OV2SLAM: A Fully Online and Versatile Visual SLAM for Real-time Applications[J]. IEEE Robotics and Automation Letters, 2021, 6(2): 1399-1406.

攻读学位期间发表的学术论文与取得的科研成果

一、学术论文

- 【1】陈孟元,李鹏飞,符乙,等.低光条件下结合 FFE-Net 网络的视觉 SLAM 算法[J].中国惯性技术学报,2025,33(01):36-45.
- 【2】P. Li, A. Xu, H. Cheng, M. Chen, S. Yang and P. Du, "Optimization algorithm for closed-loop detection based on similarity fusion mechanism," 2023 38th Youth Academic Annual Conference of Chinese Association of Automation (YAC), Hefei, China, 2023, pp. 1340-1344.

二、发明专利

- 【1】安徽工程大学.一种基于深度学习的视觉低光闭环检测算法、存储介质及设备:202410600618.5[P].2025-01-21.

三、科研项目

- 【1】安徽省重点研究与开发计划项目（高新领域），202304a05020073，面向低纹理场景的增强感知视觉移动机器人关键技术与示范应用，2023/08-2026/06，100.0 万元。（皖科高秘〔2023〕250 号）
- 【2】安徽省高校杰出青年科研项目，2022AH020065，基于视觉-惯性的移动机器人复杂环境仿生 SLAM 算法研究，2022/09-2025/08，100.0 万元。（皖教秘科〔2022〕241 号）
- 【3】安徽省高校协同创新项目，GXXT-2021-050，复杂环境下移动机器人位姿自适应优化算法研究及应用，2021/01-2023/12，40.0 万元。（皖教秘科〔2021〕58 号）

四、科技竞赛

- 【1】“游刃有余-低纹理场景下基于增强感知视觉机器人开发”——第十八届“挑战杯”全国大学生课外学术科技作品竞赛国家级三等奖
- 【2】“增强感知视觉机器人-为中国智造插上翅膀（EPV-ROBOT）”——第九届安徽省“互联网+”大学生创新创业大赛总决赛高校主赛道银奖
- 【3】“游刃有余-低纹理场景下基于增强感知视觉机器人开发”——第十届“挑战杯·华安证券”课外学术科技作品竞赛安徽省一等奖

致谢

行文至此，落笔为终。在毕业论文完成之际，我的学生生涯也即将画上句号。回首过往，二十余年的求学之路，尤其是在研究生阶段的三年时光，充满了挑战与成长，也承载了许多人的关怀与帮助。在此，我怀着无比感激的心情，向所有支持我、陪伴我的人致以最诚挚的谢意。

首先，我要深深感谢我的导师陈孟元教授。在研究生阶段，陈老师以其渊博的学识、严谨的治学态度和对科研的无限热情，为我指明了方向。无论是课题的选择、实验的设计，还是论文的撰写，陈老师都给予了我悉心的指导和宝贵的建议。每当我遇到困难时，陈老师总是耐心地为我答疑解惑，帮助我突破瓶颈。他的言传身教不仅让我在学术上受益匪浅，更让我深刻理解了何为责任与坚持。陈老师的教诲与关怀，我将铭记于心。

其次，我要感谢实验室的师兄程浩、郭行荣、钱润邦、龚光强等人。在我初入科研领域时，他们无私地分享了丰富的经验与心得，帮助我快速融入团队并掌握研究方法。他们的鼓励与支持让我在迷茫时找到了方向。同时，我也要感谢我的同门符乙、杨苏朋、徐奥、杜鹏、许瑞珩、丁帅等人，我们一起讨论问题、交流想法，共同度过了许多难忘的时光。无论是深夜的实验，还是激烈的学术讨论，都成为了我研究生生涯中最珍贵的回忆。愿他们在未来的道路上一切顺利，前程似锦。

此外，我还要感谢我的室友丁家乐、李鸿飞和王淼。虽然我们的研究方向不同，但你们的陪伴与支持为我的生活增添了许多温暖与欢乐。无论是学习上的困惑，还是生活中的琐事，你们总是给予我鼓励与帮助。这些点滴的陪伴，让我在繁忙的科研生活中感受到了一份难得的轻松与愉悦。

最后，我要特别感谢我的家人，尤其是我的父母。你们一直以来的默默支持与无私奉献，是我不断前行的动力。无论我遇到怎样的困难，你们总是无条件地相信我、鼓励我，让我能够心无旁骛地追求自己的梦想。你们的爱与关怀，是我一生中最宝贵的财富。在此，我想对你们说一声：谢谢！未来的日子里，我会继续努力，不负你们的期望。

毕业论文的完成，不仅是我学生时代的终点，更是人生新篇章的起点。带着这份感恩与回忆，我将继续前行，迎接未来的挑战与机遇。

尚德敏学 唯实惟新

