

电子科技大学

专业学位研究生学位论文开题报告表

攻读学位级别： ☐ 博士 ☒ 硕士

培养方式： ☒ 全日制 ☐ 非全日制

专业学位类别及领域： 电子信息

学 院： 信息与软件工程学院

学 号： 202222090537

姓 名： 马逸逍

论 文 题 目： 室内动态场景下的

V-SLAM 算法研究

校内指导教师： 殷光强

校外指导教师： 梁熙

填 表 日 期： 2023 年 12 月 12 日

电子科技大学研究生院

一、学位论文研究内容

课题类型	☐ 应用基础研究 ☒ 应用研究
课题来源	☐ 纵向 ☐ 横向 ☒ 自拟
学位论文研究内容	学位论文的研究目标、研究内容及拟解决的关键性问题（可续页） 1. 研究目标 SLAM（Simultaneous Localization and Mapping）指利用传感器数据实现系统自身的定位和环境地图的构建。V-SLAM 是基于相机的 SLAM 系统，通过摄像头获取环境信息，并进行特征提取、匹配和跟踪，以实现相机运动估计和地图构建。然而，传统的 V-SLAM 算法假设场景是静态的，因此来自动态区域的特征通常被视为异常值。如果这些动态区域的特征被跟踪，相机姿态估计的精度就会受到影响，甚至可能导致轨迹偏差和系统崩溃。因此，处理动态特征点成为 V-SLAM 中动态场景的一个重要问题。 在动态室内场景中同时包含了静态特征和动态特征，单纯的根据实例分割算法按照类别剔除先验动态物体上的特征点将有可能出现误删静态特征点的情况。拟研究在室内动态场景下如何完全提取真正处于运动状态的动态物体，只使用静态特征点进行相机位姿估计，从而进行更精准的定位、跟踪和建图。 2. 研究内容 本文以室内动态场景为研究对象，利用 YOLOv8 实例分割结合多视图几何区分正处于运动状态的先验动态物体，以实现在复杂场景中准确检测和区分处于运动状态的先验动态物体。再通过关联遗漏的动态物体检测算法补全未检测到的运动物体，从而实现全面的动态物体检测，拟解决动态物体漏检问题。之后去除动态物体上的特征点使用静态特征点进行准确的相机位姿估计，拟解决动态特征点对相机位姿估计的干扰问题。并通过静态信息融合的背景补全算法合成一个没有移动物体的 RGB 图像从而提高建图的质量，拟解决移动物体对建图结果的影响问题。主要研究内容包括： （1）基于 YOLOv8 和多视图几何的先验动态物体运动状态检测算法 通过 YOLOv8 实例分割算法提取出先验动态物体，再结合多视图几何检测处于先验动态物体中的特征点是否真正处于运动状态。拟研究一种能准确判断先验动态物体运动状态的算法，实现对先验动态物体的检测和准确的运动状态估计。 （2）关联遗漏的动态物体检测算法 在实例分割中，关联遗漏是指未能正确关联或分割与已检测目标相关的其他物体或部分。为了解决动态物体关联遗漏造成动态特征点剔除不完全的问题，拟研究一种能够解决关联遗漏的动态物体检测算法以准确地检测出与已检测目标相连的动态物体。 （3）研究基于静态信息融合的背景补全算法 拟补全动态物体去除后大片空白的背景，以合成一个没有移动物体的逼真图像。补全后的 RGB 图像可以提供更准确的环境结构信息并提供更准确的场景几何信息，从而提高系统重定位的准确性。 3. 拟解决的关键问题 （1）解决先验动态物体的运动状态不明确的问题 先验动态物体在实际情况中不一定处于运动状态，例如室内静止不动的人。在纹理信息

丰富的情况下可以提取较多的特征点,如果仅依靠类别将先验动态物体上的特征点全部去除,虽然可能会误删静止物体上的特征点但仍然有大量特征点可用于匹配。然而在纹理信息较少的背景下,例如静止的人靠在室内白墙上,简单地按照类别删除特征点会导致丢失大量静态特征点从而导致特征点数量不足等问题。针对上述问题,本文拟研究基于 YOLOv8 和多视图几何的先验动态物体运动状态检测算法,通过 YOLOv8 实例分割识别出先验动态物体再判断先验动态物体中的特征点的真实运动状态,解决先验动态物体真实运动状态不明确的问题。

(2) 解决关联的动态物体遗漏问题

在室内动态场景下,除了能够主动运动的物体之外还有一些与先验运动物体存在某些关联的物体可能是处于运动状态的,例如:被人拿着的书本。如果错误地将这些遗漏的动态物体上的特征点标记为静态点并进行特征点匹配,这可能导致相机位姿估计的准确性降低。针对上述问题,本文拟研究关联遗漏的动态物体检测算法,通过将动态特征点进行更完整的剔除,解决动态特征点剔除不完全的问题。

(3) 解决动态对象剔除后影响重定位准确性的问题

重新定位是指在已经建立地图的情况下,通过匹配当前帧与地图中的关键帧确定相机位姿。在存在动态物体的情况下,使用剔除动态物体之后的图像信息进行重定位可能会缺少动态物体所占据区域的几何信息,这可能导致重新定位时无法准确地估计相机的位置和姿态。针对这个问题,本文拟研究基于静态信息融合的背景补全算法,通过填补 RGB 图像中由于动态物体去除后留下的空白区域,解决动态物体剔除造成几何信息缺失的问题。

二、学位论文研究依据

学位论文的选题依据和研究意义，国内外研究现状和发展态势；选题在理论研究或实际应用方面的意义和价值；主要参考文献，以及已有的工作积累和研究成果。（2000 字）

1. 选题依据和研究意义

SLAM (Simultaneous Localization and Mapping, 即同时定位与地图构建) 是一种计算机视觉和机器人领域的技术, 旨在通过使用传感器数据在没有先验知识的情况下同时实现系统自身的定位和环境地图的构建。SLAM 通常依赖于多种类型的传感器, 如相机、激光雷达、惯性测量单元 (IMU) 等, 以获取环境中的感知信息^[5]。如果 SLAM 的传感器主要使用的是相机那么又可以称为 V-SLAM (Visual SLAM)^[6]。V-SLAM 利用摄像头获取环境信息, 并通过前端进行特征提取、特征匹配和帧间跟踪, 后端则负责进行回环检测和全局位姿优化, 以完成精确的相机运动估计和地图构建。

传统的 V-SLAM 算法在运行过程中假设场景是静态的, 这导致来自动态区域的特征通常被视为异常值。在基于特征点的 V-SLAM 系统中, 若这些来自动态区域的特征被跟踪, 则相机姿态估计的精度会受到严重影响从而导致轨迹偏差甚至系统崩溃^[7]。因此对于 V-SLAM 来说, 静态特征点的提取对相机跟踪的稳定性和地图构建的质量至关重要。

如何处理动态特征点就成为了 V-SLAM 中室内动态场景的一个重点关注问题。V-SLAM 前端通常采用随机样本一致性算法 (Random Sample Consensus, RANSAC)^[8] 来进行离群值的判断, 当存在少量运动特征点的时候 RANSAC 可以将其检测为 outlier, 但当运动特征点较多时 RANSAC 算法往往会失效^[9]。最近的一些研究通常基于光流估计或深度学习来应对场景中的动态物体^[9-13]。光流检测通过计算连续帧之间动态特征点由运动产生的光流向量判断特征点的运动方向和速度, V-SLAM 系统通过光流检测得到特征点的运动进行相机的位姿估计^[10]。但是相较于室外物体而言, 室内场景下物体的运动速度相对更慢。当物体的运动较慢时, 像素之间的位移较小, 这会导致光流估计的不准确^[11]。由于光流估计的精度取决于像素的位移大小, 低速物体的光流估计可能会产生较大的误差。深度学习引入目标检测或者实例分割来识别物体类别, 剔除视野中识别到运动物体类别范围上的特征点^{[9][12]}, 只使用静态特征点进行相机的位姿估计。现阶段通过深度学习识别物体类别剔除动态点可能会误删位于静态物体上的特征点, 若要减少静态特征点的误删除需要进一步的运动检测, 这样就会造成更多时间消耗。虽然现阶段轻量化的目标检测速度很快, 但是矩形框必然会引入背景静态点造成静态特征点的错误剔除。

在 V-SLAM 中, 实例分割对场景中的图像进行像素级别的语义分类, 将不同的像素分配给不同的语义类别。它通过将图像分割为语义对象的集合从而提供更丰富的外界环境信息, 进而筛选出对相机位姿估计有利的特征点。但是仅仅依靠实例分割对先验运动物体这一类别上的特征点进行剔除显然不合理, 这可能会导致处于静止情况下的先验动态物体上的特征点也被错误剔除, 使得在室内纹理信息较低的情况下特征点数目不足从而导致难以进行跟踪。因此可以结合几何信息准确区分运动物体, V-SLAM 可以更好地剔除运动物体上的特征点并进行相机位姿估计, 从而提供更精确的定位和地图建模。

2. 国内外研究现状和发展态势

V-SLAM 是一种基于视觉信息的 SLAM 技术, 已经在机器人导航、无人机等领域得到广泛的应用。以下是近年来与 V-SLAM 相关的一些研究工作:

(1) 语义 V-SLAM

语义 V-SLAM 是将 V-SLAM 与对环境的语义理解相结合的一项技术。它通过将语义信息 (如物

体类别、场景语义或语义分割)纳入定位和建图过程来优化传统的 V-SLAM 系统。语义 V-SLAM 的一个研究方向是将语义信息整合到建图过程中。这些方法利用语义分割技术提取环境中具有意义的区域,并将其与相应的地图模块关联起来。通过将语义标签整合到地图中,使得构建的地图可以区分不同类别的物体。2018 年 Furrer F 等人^[13]利用语义信息通过边缘和凸度信息对传感器提供的深度图像进行分割,并通过合并对象模型的方法完成场景的重建。2023 年 Cheng Shuhong 等人^[14]将实时 RGB-D 语义可视化信息加入 V-SLAM 系统,在语义线程中生成三位点云以及三维语义对象进而发布到机器人操作系统(ROS)进行可视化,最终生成用不同颜色表示不同物体的语义八叉树地图。

语义 V-SLAM 的另一个方向是利用对象检测模块进行物体识别。通过目标检测或者实例分割在像素级别识别对象,增加 V-SLAM 对环境的理解。现阶段在语义 V-SLAM 中使用的实例分割方法大多都是深度学习的方法。[15]中提出的 Dynamic-SLAM 框架基于卷积神经网络,结合物体的先验知识利用 SSD 对象检测器在语义层面检测线程中的物体提高 SLAM 的性能。2020 年 Xi Zhihong^[7]将 ORB-SLAM2^[16]与 PSPNet 实例分割网络相结合,利用实例分割结果以及光流检测消除某一类别上的特征点,从而提高相机位姿估计的精度。现阶段常常将语义信息加入处理动态场景下的 V-SLAM 问题,V-SLAM 中加入对语义信息的理解也是现在处理动态场景的一项极具挑战的任务。

(2) 基于光流法的动态 V-SLAM

传统的 V-SLAM 算法假定场景是静态的,在面对移动物体时可能会失效或产生不准确的结果,近年来动态 V-SLAM 也成为人们的重点研究问题之一^[17-19]。

动态 V-SLAM 的一种方法是利用光流法对动态物体的运动进行跟踪。2019 年 Yang 等人^[20]用立方体建模物体并利用 2D KLT 稀疏光流算法直接在像素层面进行特征点的跟踪,并通过将各项约束加入 BA 优化来改进相机位姿估计。Zhang 等人^[21]的 VDO-SLAM 在预处理模块中首先进行场景中可能移动物体的实例分割,再利用密集光流跟踪分割结果上动态物体的角点特征从而进行相机位姿估计。2022 年, Qiu Yuheng 等人^[22]设计的动态 V-SLAM 系统 AirDOS 通过 Mask-RCNN 执行的实例分割将动态物体行人提取出来,再进行 Alpha-Pose 提取人体关键点后通过光流法跟踪动态目标,最终通过人体关键点的三角测量以及光流法的动态目标跟踪进行相机位姿估计。但是基于光流法的动态 V-SLAM 系统易受到光照等因素的影响,在情况复杂的动态环境下很难达到预期效果。

(3) 基于深度学习的动态 V-SLAM

动态 V-SLAM 的另一个方向是利用深度学习来处理动态场景^[23-25]。这些方法利用目标检测或实例分割算法来识别和跟踪环境中感兴趣的物体。通过将物体级信息与 V-SLAM 系统集成,这些方法可以有效处理动态物体,并保持场景的一致性。

2018 年 Zhong 等人^[26]将 V-SLAM 系统与目标检测 SSD 算法相结合从而去除运动物体上的特征点。由于目标检测算法检测动态物体会引入很多非动态物体范围内的背景点,因此再通过 Grab-cut 算法进行背景点的剔除以达到只准确剔除动态特征点从而提高动态环境下的 V-SLAM 系统定位性能的目的。Yu 等人^[12]设计了名为 DS-SLAM 的动态 V-SLAM 系统,将实例分割与运动一致性检测相结合减少动态目标的影响。尽管 DS-SLAM 的定位精度大大提高但是仍存在实例分割的速度限制以及计算特征密集的问题。2022 年 Cui LinYan 等人^[27]提出的语义光流 SOf-SLAM 使用 SegNet 生成逐像素语义分割结果作为输入掩码,再通过计算出基本矩阵过滤掉动态特征。然而在动态环境中使用传统的特征点匹配的方法以及仅仅依赖前后连续的两帧中的数据使得 SOf-SLAM 任有较大提升空间。

总之,将语义信息结合到动态场景是近年来 V-SLAM 处理动态问题的主流方式,有助于提高 V-SLAM 系统区分动态物体和静态物体进而提高 V-SLAM 在动态环境中的场景理解和鲁棒性方面的能

力。

3. 选题在理论研究或实际应用方面的意义和价值

在理论研究方面，室内动态 V-SLAM 的研究可以推动 SLAM 技术的进一步发展。传统的 V-SLAM 方法主要处理静态环境，而动态 V-SLAM 则致力于在存在动态物体的情况下实现定位和地图构建。这需要解决动态物体的检测、跟踪和建模等关键问题。研究动态 V-SLAM 可以推动相关算法和技术的创新和提高，为实际应用提供更准确、鲁棒的解决方案。

在实际应用方面，室内动态 V-SLAM 具有广泛的应用前景。动态 V-SLAM 可以应用于室内机器人导航和家政机器人等领域，提供对动态环境的感知和理解能力，为人们提供更智能和便捷的服务。

综上，室内动态 V-SLAM 的理论研究和实际应用价值都非常重要。通过深入研究室内动态 V-SLAM 的理论基础和算法，可以推动 V-SLAM 技术的发展。同时，将室内动态 V-SLAM 应用于实际场景中，可以为机器人和自主系统提供更准确、鲁棒的环境感知和地图构建能力，推动相关领域的发展和创新。

4. 参考文献

- [1] 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.
- [2] Antonio D, Medeiros H R, Macedo D, et al. SegNetRes-CRF: A Deep Convolutional Encoder-Decoder Architecture for Semantic Image Segmentation[C]// 2018:1-6.
- [3] Library W P. International Conference on Computer Vision[J]. 2007.
- [4] Redmon J, Divvala S, Girshick R, et al. You Only Look Once: Unified, Real-Time Object Detection[C]// Computer Vision & Pattern Recognition. IEEE, 2016.
- [5] Cui L, Ma C. SOF-SLAM: A semantic visual SLAM for Dynamic Environments[J]. IEEE Access, 2019, PP (99):1-1.
- [6] Tourani A, Bavle H, Sanchez-Lopez J L, et al. Visual SLAM: what are the current trends and what to expect?[J]. Sensors, 2022, 22(23): 9297.
- [7] S. Han and Z. Xi, "Dynamic Scene Semantics SLAM Based on Semantic Segmentation," in IEEE Access, vol. 8, pp. 43563-43570, 2020, doi: 10.1109/ACCESS.2020.2977684.
- [8] M.A. Fischler, R.C. Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography.Commun. ACM 24 (6) (1981) 381-395.
- [9] Li A, Wang J, Xu M, et al. DP-SLAM: A visual SLAM with moving probability towards dynamic environments[J]. Information Sciences, 2021, 556: 128-142.
- [10] Qiu Y, Wang C, Wang W, et al. AirDOS: Dynamic SLAM benefits from articulated objects[C]//2022 International Conference on Robotics and Automation (ICRA). IEEE, 2022: 8047-8053.
- [11] Liao X, Cai Z, Chen J, et al. Physics-based optical flow estimation under varying illumination conditions[J]. Signal Processing: Image Communication, 2023, 117: 117007.
- [12] Yu C, Liu Z, Liu X J, et al. DS-SLAM: A semantic visual SLAM towards dynamic environments[C]//2018 IEEE/RSJ international conference on intelligent robots and systems (IROS). IEEE, 2018: 1168-1174.
- [13] Furrer F, Novkovic T, Fehr M, et al. Incremental object database: Building 3D models from multiple partial observations[C]//2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE,

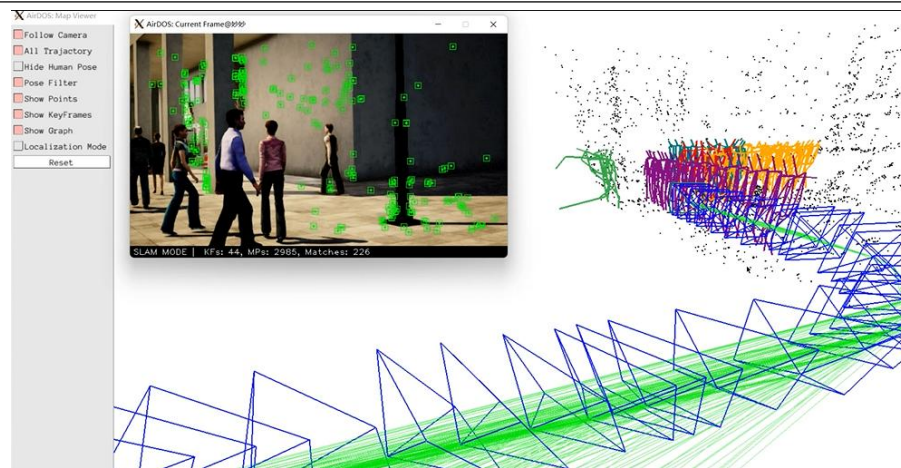
2018: 6835-6842.

- [14] Cheng S, Sun C, Zhang S, et al. SG-SLAM: a real-time RGB-D visual SLAM toward dynamic scenes with semantic and geometric information[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 72: 1-12.
- [15] Xiao L, Wang J, Qiu X, et al. Dynamic-SLAM: Semantic monocular visual localization and mapping based on deep learning in dynamic environment[J]. Robotics and Autonomous Systems, 2019, 117: 1-16.
- [16] 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.
- [17] Chen W, Shang G, Hu K, et al. A Monocular-Visual SLAM System with Semantic and Optical-Flow Fusion for Indoor Dynamic Environments[J]. Micromachines, 2022, 13(11): 2006.
- [18] Hu B, Luo J. A Robust Semi-Direct 3D SLAM for Mobile Robot Based on Dense Optical Flow in Dynamic Scenes[J]. Biomimetics, 2023, 8(4): 371.
- [19] Theodorou C, Velisavljevic V, Dyo V. Visual SLAM for Dynamic Environments Based on Object Detection and Optical Flow for Dynamic Object Removal[J]. Sensors, 2022, 22(19): 7553.
- [20] Yang S, Scherer S. Cubeslam: Monocular 3-d object slam[J]. IEEE Transactions on Robotics, 2019, 35(4): 925-938.
- [21] Zhang J, Henein M, Mahony R, et al. VDO-SLAM: a visual dynamic object-aware SLAM system[J]. arXiv preprint arXiv:2005.11052, 2020.
- [22] Qiu Y, Wang C, Wang W, et al. AirDOS: Dynamic SLAM benefits from articulated objects[C]//2022 International Conference on Robotics and Automation (ICRA). IEEE, 2022: 8047-8053.
- [23] Chen L, Ling Z, Gao Y, et al. A real-time semantic visual SLAM for dynamic environment based on deep learning and dynamic probabilistic propagation[J]. Complex & Intelligent Systems, 2023: 1-25.
- [24] Lee J, Back M, Hwang S S, et al. Improved real-time monocular SLAM using semantic segmentation on selective frames[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 24(3): 2800-2813.
- [25] Li X, Shen Y, Lu J, et al. DyStSLAM: an efficient stereo vision SLAM system in dynamic environment[J]. Measurement Science and Technology, 2022, 34(2): 025105.
- [26] Zhong F, Wang S, Zhang Z, et al. Detect-SLAM: Making object detection and SLAM mutually beneficial[C]//2018 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, 2018: 1001-1010.
- [27] Cui L, Ma C. SOF-SLAM: A semantic visual SLAM for dynamic environments[J]. IEEE access, 2019, 7: 166528-166539.

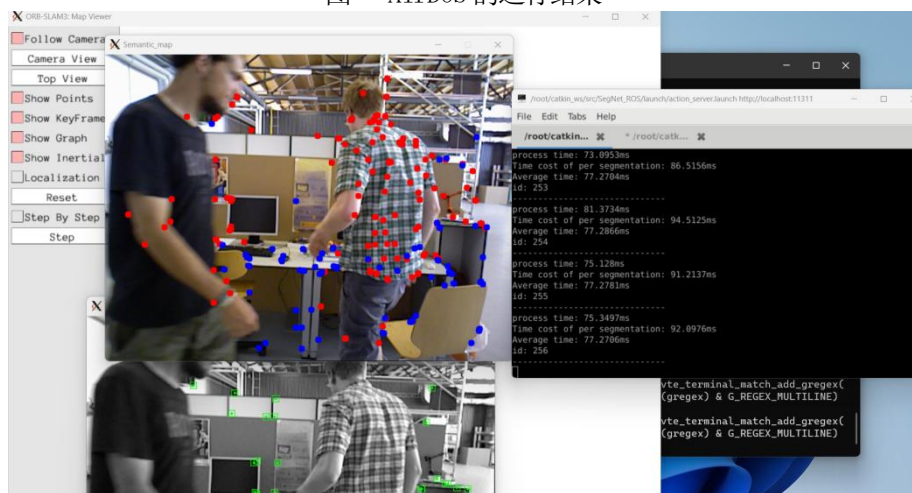
5. 工作积累和研究成果

(1) 研究生阶段的学习主要以 SLAM 为主, 学习了《视觉 SLAM 十四讲》并掌握了 ORB-SLAM2 以及 ORB-SLAM3 框架, 阅读了大量的与动态 SLAM 有关的文献。

(2) 在 WSL2 上测试了 TartanAir_shibuya 数据集测试了 AirDOS 的运行效果并使用 TUM 数据集测试室内动态环境下的 RDS-SLAM。



图一 AirDOS 的运行结果



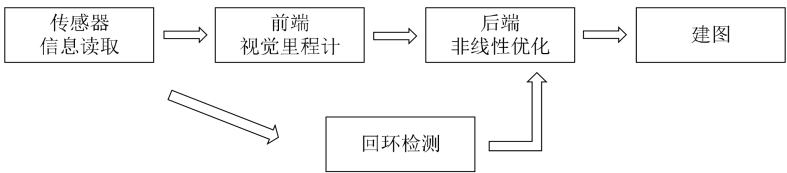
图二 RDS-SLAM 的运行结果

三、学位论文研究计划及预期目标

1. 拟采取的主要理论、研究方法、技术路线和实施方案（可续页）

(1) 主要理论

V-SLAM 的经典框架如下图所示：

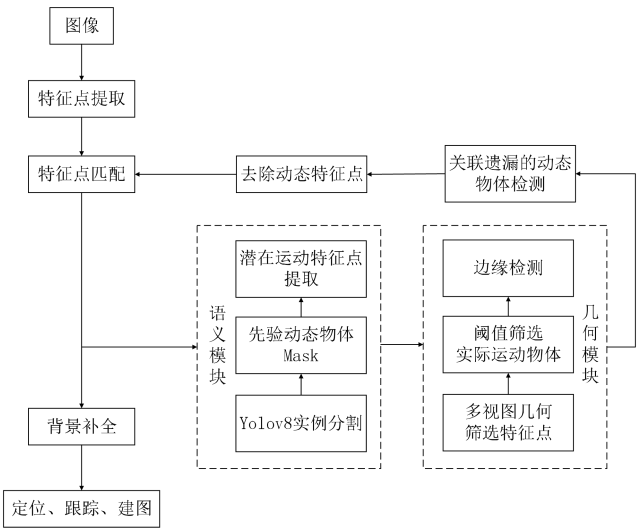


图三 V-SLAM 的框架

V-SLAM 利用摄像头获取环境信息，其中传感器信息读取主要是相机传感器，包括单目、双目、RGBD 等，有时还包括惯性传感器（IMU：陀螺仪+加速度传感器），对这些传感器的信息进行读取和预处理。通过前端进行特征提取、特征匹配和帧间跟踪，后端则负责进行回环检测和全局位姿优化，以完成精确的相机运动估计和地图构建。如图一所示在 V-SLAM 系统中，前端模块负责从连续的图像帧中提取特征点，并对这些特征点进行匹配和跟踪，以估计相机的运动。后端模块负责进行回环检测和全局位姿优化。回环检测的目标是识别相机曾经访问过的地点，以解决累积误差问题并提高定位的准确性。全局位姿优化的目标是通过最小化重投影误差来优化相机的位姿，并进一步提高地图的精度。这个过程通常使用非线性优化算法，如图优化或束优化，来求解最优的相机位姿。

(2) 研究方法

本课题主要采用文献研究法和实验法相结合的研究方法。首先收集并阅读近年来本领域有较大影响力的期刊或者会议发表的文献,对文献进行思考与总结。再充分调研理论知识和相关技术实现，结合自己的想法对各个模块进行改进，之后通过实验进行对比，确认方案的可靠性与有效性。



图四 总体技术路线

(3) 技术路线和实施方案

拟设计一个室内动态场景下的 V-SLAM 系统，总体技术路线如图四所示。图像首先通过语义模块进行 YOLOv8 实例分割获得包括人在内的先验动态物体的Mask 从而提取到处于该类别范围内的潜

在运动特征点，再提取出位于该类别的特征点并通过多视图几何筛选出真正处于运动状态的特征点。如果该 Mask 中的处于运动状态的特征点超过一定的阈值，则认为该 Mask 对应的先验动态物体确实处于运动状态。接下来再通过对该类别进行边缘检测，减少处于类别边缘的特征点运动状态的误判，这样就很好的提取正处于运动状态的先验动态物体上的动态特征点。之后再结合关联遗漏的动态物体检测算法判断与先验动态物体相接触的物体的运动状态，对动态特征点进行更完整的剔除从而避免将这些点进行错误匹配。最后通过基于静态信息融合的背景补全算法生成没有移动物体的图像，提高建图的质量。

2. 研究计划可行性，研究条件落实情况，可能存在的问题及解决办法（可续页）

（1）可行性分析

①YOLOv8 实例分割能力：YOLOv8 是一种高效的实例分割算法，能够实时地检测场景中的各种物体。在室内场景中，如家居、办公室等，通过使用 YOLOv8 可以准确地检测出物体的位置和类别，为后续的 V-SLAM 算法提供关键的输入信息。

②多视图几何筛选：多视图几何信息可以提供更准确的动态物体运动状态估计。通过对位于特定类别的特征点进行多视图几何筛选，可以准确地判断动态物体的运动状态，从而提取出真正处于运动状态的特征点。

③动态物体的处理：在室内动态场景中，有些不是先验动态物体、但与先验物体有关联的物体也有可能是动态的。如行走的人拿着的书、被人移动的椅子等。这些动态物体可能会对 V-SLAM 算法的相机位姿估计和场景重建造成干扰。通过关联遗漏的动态物体检测算法对之前遗漏的动态物体进行补充，从而更好地处理它们的影响。

（2）研究条件

①教研室拥有良好的师资力量，具有良好的科研环境、交流平台；实验室团队拥有丰富的论文投稿经验及专利申请经验。

②实验室设备丰富，具有 RK3568、RK3588 等开发板，Intel Realsense 深度摄像头 D354i、ORBEC Astra 深度摄像头、双目摄像头等设备。使用的数据集，如 KITTI、TUM、shibuya 等。

（3）可能存在的问题及解决办法

①YOLOv8 实例分割出现漏检情况

在提取先验动态物体时依赖于 YOLOv8 实例分割网络，但是可能会存在误检和漏检的问题，导致对先验动态物体的检测不准确或不完整。针对以上问题，可以考虑结合深度学习中的目标跟踪算法，如基于卡尔曼滤波器的目标跟踪算法，来提高先验动态物体的检测准确性和完整性。利用卡尔曼滤波器进行目标跟踪，估计物体的当前位置和速度，并预测未来的位置。通过融合 YOLOv8 检测结果和卡尔曼滤波器的跟踪结果，减少误检和漏检的问题。

①研究方法和研究进度的偏差

如果发现拟采取的研究方法和技术路线与实际效果不符，我将与导师进行进一步交流讨论。在与导师的沟通中，评估当前研究方案的可行性，并根据讨论结果适度调整研究方案，甚至考虑采用新的研究方法。这样可以确保研究能更加准确和有效地达到预期的目标。如果研究计划与实际研究进度不符，首先会进行客观评估，评估当前研究内容或实验进度的必要性和合理性。在评估的基础上，我会决定对当前研究内容的舍弃或者对研究周期进行适度缩减。这样可以确保在有限的时间内能够更好地完成研究任务，并确保研究的整体进展顺利。

3. 研究计划及预期成果		
研究计划	起止年月	完成内容
	2023.07-2023.12	调研室内动态场景下的 V-SLAM 算法
	2024.01-2024.03	构建基于 YOLOv8 和几何信息的动态 V-SLAM 算法，完成专利 1 篇
	2024.04-2024.06	构建室内动态场景下的 V-SLAM 算法，完成学术论文 1 篇
	2024.07-2024.09	优化算法，完成学术论文 1 篇
	2024.10-2024.12	验证实验效果，完成专利 1 篇
	2025.01-2025.06	撰写硕士学位论文
预期创新点及成果形式	(1) 预期创新点 ①基于 YOLOv8 和多视图几何的先验动态物体运动状态检测算法 ②关联遗漏的动态物体检测算法 ③研究基于静态信息融合的背景补全算法 (2) 成果形式 ①学术论文 在国内外期刊、国际会议上发表或录用论文 2 篇。 ②专利 申请发明专利 3 项。 ③研究报告 写相应的研究进度报告和总结报告。 ④大论文 撰写硕士学位论文一篇。	

四、开题报告审查意见

1.导师对学位论文选题和论文计划可行性意见，是否同意开题：

校内导师（组）签字：_____年 月 日

校外导师签字：_____年 月 日

2. 开题报告考评组意见

开题日期		开题地点	
考评专家			
考评成绩	合格____票	基本合格____票	不合格____票
结 论	☐ 通过 ☐ 原则通过 ☐ 不通过 通 过： 表决票均为合格 原则通过： 表决票中有 1 票为基本合格或不合格，其余为合格和基本合格 不 通 过： 表决票中有 2 票及以上为不合格		

考评组对学位论文的选题、研究计划及方案实施的可行性的意见和建议：

考评组签名：

_____年 月 日

3.学院意见：

负责人签名：

_____年 月 日