

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： 202222090508

姓 名： 陈佐均

论 文 题 目： 基于 RGB-D 相机的物体级语义

SLAM 算法研究

校内指导教师： 王治国

校外指导教师： 梁熙

填 表 日 期： 2023 年 12 月 14 日

电子科技大学研究生院

一、学位论文研究内容

课题类型	<input type="checkbox"/> 应用基础研究 <input checked="" type="checkbox"/> 应用研究
课题来源	<input type="checkbox"/> 纵向 <input type="checkbox"/> 横向 <input checked="" type="checkbox"/> 自拟
学位论文研究内容	<p>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）</p> <p>1. 研究目标</p> <p>语义 SLAM(Semantic SLAM)是 SLAM(Simultaneous Localization and Mapping)技术的一种发展形式，它在传统 SLAM 的基础上引入了语义信息，即对环境中物体的语义理解。这使得机器不仅能够感知周围环境的几何结构，还能理解环境中不同物体的语义关系，从而提高对环境的高层次理解能力。语义 SLAM 中，使用目标识别算法通常会输出物体的 2D 检测框，然而这些框通常包含物体周围的前景和背景部分，而且物体在检测框中的占比可能较小，可能存在较多的噪声。这种情况可能对物体的整体形状等信息造成干扰，影响构建准确的物体表示。因此，研究从 2D 检测框中精确分割出物体与背景信息，以提高物体构建的精准性。这一过程的目标是有效消除噪声，使物体的表示更为准确和可靠。</p> <p>在物体被检测时，可能是在不同角度和外观条件下被观察到，导致形状、颜色等特征出现较大差异。这种情况下，检测到的物体与在语义地图上构建的 3D 物体可能被错误地判断为不同的实体从而导致算法对物体位姿的估计错误，同时重复构建同一个物体最终导致构建的语义地图失效，无法正常使用。研究物体的数据关联，以提高语义地图的可用性。通过强化物体数据之间的联系，可以更准确地将不同视角和外观下的物体关联为同一实体，从而改善语义地图的一致性和准确性。</p> <p>语义 SLAM 还需要构建包含语义信息的地图。也就是说地图中的区域不仅有几何属性，还具有语义标签，例如墙壁、椅子、桌子等。研究如何在语义地图中高质量地表示物体，能提供更为精确的位置和语义信息，为上层的应用提供更为强大的支持。通过提高对物体的高质量表示，可以增强机器人对环境的感知和理解能力，使其更准确地识别和定位各类物体，为智能机器人在不同任务和场景中的应用提供更可靠的基础数据。</p> <p>2. 研究内容</p> <p>语义 SLAM 不仅仅侧重于几何结构的感知，还注重对环境语义信息的理解。在实际场景中，算法识别到的 2D 检测框中包含的背景信息可能对物体信息的准确使用造成一定的干扰，影响不同观测角度下的同一物体的数据关联，导致物体建模不准确。通过有效地去除背景干扰并确保物体之间的正确关联，可以改善物体建模的准确性和一致性，从而提高整体系统对场景的理解和表示。主要的研究内容包括：</p> <p>1. 基于深度和灰度值的物体前景分割方法</p> <p>这一问题针对图像中检测到的物体，通过利用背景与物体的深度值、灰度值的不一致性，实现对物体的像素级层面分割。该方法的实质是检测图像中深度值或灰度值发生显著变化的像素点，将其视为物体的边界，从而实现对物体的精准分割。这种方</p>

法通过捕捉边缘信息的突变，有效地识别出物体的形状和轮廓，为后续的物体建模和识别提供了更为准确的基础。基于深度的物体前景分割

2. 基于静态特征点的数据关联方法

在语义 SLAM 中，解决静态物体的数据关联问题旨在对不同图像中的静态物体进行判断，确保同一物体的不同观测不会被孤立地重新构建。这一问题的核心在于通过算法分析和识别，对同一静态物体的观测点进行准确的关联融合，从而在现有地图的基础上对物体进行修改。此过程需要综合考虑几何和语义信息，以保障对静态物体的观测关联的准确性和鲁棒性，以提高系统对环境的感知和地图构建的精度。

3. 基于目标识别的物体准确重建方法

精准的物体重建有助于 SLAM 系统对于自身的定位，以及对环境的感知理解。通过目标识别所得的语义信息以及物体位置信息等，在语义地图上重建出准确的物体。其实质是断动态物体在先前帧中的状态，从而在整个时间序列中重建出静态物体的位置和形状。

3. 拟解决的关键性问题

1) 静态物体前景分割边界不准确问题

在目标检测输出的物体检测框中，背景信息作为一个潜在的干扰噪声，对于物体的重建会产生显著的影响。背景信息的存在可能导致物体的形状、轮廓和纹理等特征受到干扰，进而影响物体重建时的准确性和系统的鲁棒性。针对这一问题，计划研究更精确的前景分割方法，通过利用图像的深度、灰度等信息，辅助前景分割，以获取更精准的分割结果，提高物体重建的精度，减少背景干扰对系统性能的影响，为语义地图的构建提供更可靠的基础。

2) 静态物体数据关联不一致性问题

在语义 SLAM 中，静态物体的数据关联涉及将在不同时间获得的观测数据正确关联起来，以确保系统对环境的持续理解和地图构建。针对这一问题，除了考虑几何信息外，还需要考虑物体的语义信息。其中的主要问题是确保语义一致性。由于环境中可能存在语义相似的物体，如相似的家具或结构，系统需要能够正确地将观测到的语义标签与先前的数据关联起来，同时避免将不同物体错误地关联在一起。因此静态物体数据的有效关联能保证对环境的正确感知和地图构建的完整。

3) 语义地图静态物体重建准确性较差问题

传统 SLAM 系统主要侧重于几何信息，而在语义地图中准确重建静态物体则需要更深层次的语义理解。静态物体的精准重建要求系统能够区分不同物体的类别、准确标记其位置，并将其在语义地图中恢复出真实的形状和特征。针对这一问题，需要涉及到解决物体遮挡、视角变化和光照变化等方面的问题，以确保语义地图不仅包含准确的几何属性，还能反映物体的语义信息。因此，精准的重建物体能为机器人提供更全面、准确的环境理解。

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

同时定位与建图(Simultaneous Localization and Mapping, SLAM)近年来成为基础研究领域，因为它有望解决与自我探索型人工智能移动机器人领域相关的大多数问题^[1]。例如，无需对其所探索的环境有任何先验知识且无需任何人为干扰即可进行探索的能力、机器人在城市搜索救援、地下采矿、水下监视和行星探索等许多应用领域。SLAM 是实现机器人在未知环境中实时进行自身位置估计和地图构建的关键技术。其主要目标在于使机器人能够，动态感知周围环境、精确定位自身位置，并在这一过程中构建地图。SLAM 系统涵盖了多种类型，可以分为基于激光雷达的激光 SLAM、基于相机的视觉 SLAM(VSLAM)、基于惯性测量单元的惯性 SLAM 等多种类型。

随着技术的发展，移动机器人有了越来越复杂的应用场景，对环境的感知能力提出了更多的要求。而深度学习的出现与快速发展，Segnet^[2]、YoLo 系列^[3]等神经网络为 SLAM 系统提供了应对更复杂环境能力以及自身定位精度的提升。SLAM 结合语义分析算法，借助语义信息，我们可以将数据关联从传统的像素级别升级到物体级别，提高机器人对环境的感知能力，一方面可以先验性的判断部分动态目标进行处理，另一方面提高了 SLAM 算法在回环检测，位姿优化等部分的精确度，而且对场景的语义分析还有利于机器人的自主理解和人机交互等功能。然而目前的语义 SLAM 对于语义地图的构建并不完善，需要对静态物体的数据进行正确关联，重建一个更加精细的物体因此对于语义 SLAM 的检测物体的准确前景分割、静态物体的数据正确关联、语义地图上物体的精准重建、等工作是需要的。

2. 国内外研究现状和发展趋势

SLAM 技术是机器人领域中一个重要的研究方向，其目标是使移动机器人在未知环境中能够实时进行自身位置估计和地图构建。Cesar Cadena 等人将 SLAM 发展历程分为了三个阶段，经典时代(1986-2004)、算法时代(2004-2015)以及鲁棒感知时代(2015-现在)^[4]。在经典时代，引入了 SLAM 的主要概率公式，包括基于扩展卡尔曼滤波器、RaoBlackwellized 粒子滤波器和最大似然估计等方法；此外，这个时期还涵盖了与效率和强大的数据关联相关的基本挑战。Durrant-Whyte 和 TimBailey 的两项工作^[5,6]对经典时代的早期发展和主要公式结论进行了详细回顾，内容基本全面覆盖了整个经典时代的发展。接着是算法时代，Gamini Dissanayake 等人的工作^[7]回顾内容涵盖这个时期的一些发展，并提出了一些 SLAM 面临的一些挑战。目前，我们正处于鲁棒感知时代，其中涉及到一些新的挑战如，鲁棒性能、高层次理解、资源感知和任务感知、驱动推理。

传统的 SLAM 主要依赖于昂贵的激光雷达、IMU 等传感器进行环境感知。通过几何特征点的提取和匹配，传统 SLAM 系统能够实现机器人的定位和环境地图构建。然而，这些传感器通常较昂贵，限制了 SLAM 系统的广泛应用。随着技术的发展，视觉 SLAM(Visual SLAM)开始出现，利用相机等视觉传感器进行环境信息的获取，推动了 SLAM 技术的更广泛应用，并催生了多种经典算法。2013 年 Salas-Moreno 等人提出的 SLAM++^[8]是物体级

SLAM 领域的最早的代表作；2015 年，Mur-Artal 等人提出的 ORB-SLAM^[9] 是一种基于特征的单目 SLAM 系统，可以在小型和大型、室内和室外环境中实时运行，成为模块化 SLAM 领域的一项重要工作，很多后续出现的基于特征匹配的 SLAM 系统都是由 ORB-SLAM 发展而来。如 Mur-Artal 等人接下来提出的 ORB-SLAM2^[10]，在保持框架整体性的基础上，对一些细节进行了改进，使其能够适用于更多种类的相机，包括深度相机和双目相机，此外，跟踪线程中引入了预处理模块，最后有一个全局 BA(Bundle adjustment)提高系统的鲁棒性；ORB-SLAM3^[11]在此基础上耦合了惯导 IMU、加入了融合估计以及子地图功能。另一方面，Tong Qin 等人提出的 VINS^[12,13]系列是不同于 ORB-SLAM 的又一经典框架。支持多种视觉惯性传感器类型包括 IMU、GPS 等。在此算法基础上也提出了很多新算法，Seungwon Song 等人提出的 DynaVINS^[14]，设计了一个鲁棒 BA 来丢弃一些异常的特征。

SLAM 框架早期是建立在静态假设成立的基础上，即认为环境中的所有物体都是静态不动的，唯一移动的物体是传感器本身。这种假设导致在存在动态物体的环境中，位姿估计容易变得不准确，甚至在高度动态的环境中可能完全失效。为了解决这一问题，提出了动态场景下的 SLAM，即动态 SLAM。动态 SLAM 将环境中的物体分为了动态和静态两类来进行区分。在一些动态 SLAM 中，动态物体被剔除，不纳入位姿估计计算当中。例如，Daniela Esparza 等人提出的 STDyn-SLAM^[15]采用对极几何的方法，通过建立当前帧和上一帧光流，根据对极几何约束判断是否为动态点，从而将动态点剔除；Wu 等人提出的 YOLO-SLAM^[16]将物体检测与 RANSAC 算法相结合，以去除动态物体。另一方面，一些 SLAM 框架采用不同的策略，将动态物体位姿进行估计并纳入到优化中，一同用来估计相机位姿。

传统 SLAM 系统主要依赖于几何信息，在某些场景下可能限制了对环境的深度理解。随着深度学习技术的兴起，VSLAM 系统得到了显著的改进。深度学习方法被广泛用于图像特征提取、深度图的生成、对抗性训练，可以提高 SLAM 系统的鲁棒性等性能。深度学习为 VSLAM 系统获取更多的环境语义信息，增强对环境的高层次理解能力，从而更好的感知环境。在 SLAM 系统中加入语义信息，可以形成语义 SLAM(Semantic SLAM)。语义 SLAM 利用深度学习网络对物体进行分割，能更好的识别可能的动态物体，同时构建出包含语义信息的地图，在导航和环境交互等方面有更好的效果。2017 年 Martin R`unz 等人提出的 Co-Fusion^[17]利用 SharpMask^[18]将场景分割成不同的对象(使用运动或语义线索)，同时跟踪和重建真实的 3D 形状，并随时间推移改进物体在地图上的模型。2018 年 Martin R`unz 等人进行改进，提出了 Mask-Fusion^[19]，使用 MASK-RCNN^[20]网络对场景中的不同对象进行识别，并在 SLAM 线程之外添加了一个用于分割的语义线程，以提高系统的实时性。2022 年 Shuhong Cheng 等人提出的 SG-SLAM^[21]在 ORB-SLAM2 的基础框架上添加了两个新的并行线程：一个用于获取 2D 语义信息的对象检测线程和一个语义地图线程，然后利用语义信息和几何信息快速剔除动态点，使用静态特征点进行位姿估计，将静态物体构建到语义地图中。

语义 SLAM 代表了 SLAM 技术在理解和应用语义信息方面的进一步演进，为机器在真实世界中更智能、更有感知力地行动提供了新的可能性。

3. 理论研究和实际应用方面的意义和价值

在理论研究方面，语义 SLAM 的深度融合将传统 SLAM 与深度学习技术相结合，拓展了环境感知的深度和广度，为机器在复杂场景中的理解提供了更丰富的语义信息。其能够处

理动态环境,通过引入语义信息提高系统在现实场景中的鲁棒性,并实现高层次的环境理解。此外,语义 SLAM 涉及多传感器融合,如相机、激光雷达、IMU 等,为机器提供全面、多维度的感知数据。

在实际应用方面,语义 SLAM 为智能导航领域提供了更智能的环境感知和理解能力,使机器能够更安全、高效地导航。在自动驾驶中,语义 SLAM 通过对道路上的交通标识、行人、车辆等进行语义分析,提供了更精准的环境感知,提高了驾驶安全性。在增强现实应用中,语义 SLAM 为虚拟对象与真实世界的交互提供准确的定位和语义信息,提升了用户体验。此外,语义 SLAM 在环境监测领域也发挥着关键作用,帮助机器更全面地理解环境,适应各种复杂情况。

语义 SLAM 的发展不仅推动了 SLAM 技术的理论进步,也为各种智能系统提供了更强大、灵活的环境感知和理解能力,使得机器能够在复杂、动态的现实世界中更加智能和适应性强。

4. 主要参考文献

[1] A. R. Khairuddin, M. S. Talib and H. Haron, "Review on simultaneous localization and mapping (SLAM)," 2015 IEEE International Conference on Control System, Computing and Engineering (ICCSCE), Penang, Malaysia, 2015, pp. 85-90, doi: 10.1109/ICCSCE.2015.7482163.

[2] V. Badrinarayanan, A. Kendall and R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 12, pp. 2481-2495, 1 Dec. 2017, doi: 10.1109/TPAMI.2016.2644615.

[3] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 779-788, doi: 10.1109/CVPR.2016.91.

[4] C. Cadena et al., "Past, Present, and Future of Simultaneous Localization and Mapping: Toward the Robust-Perception Age," in IEEE Transactions on Robotics, vol. 32, no. 6, pp. 1309-1332, Dec. 2016, doi: 10.1109/TRO.2016.2624754.

[5] H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping: part I," in IEEE Robotics & Automation Magazine, vol. 13, no. 2, pp. 99-110, June 2006, doi: 10.1109/MRA.2006.1638022.

[6] T. Bailey and H. Durrant-Whyte, "Simultaneous localization and mapping (SLAM): part II," in IEEE Robotics & Automation Magazine, vol. 13, no. 3, pp. 108-117, Sept. 2006, doi: 10.1109/MRA.2006.1678144.

[7] G. Dissanayake, S. Huang, Z. Wang and R. Ranasinghe, "A review of recent developments in Simultaneous Localization and Mapping," 2011 6th International Conference on Industrial and Information Systems, Kandy, Sri Lanka, 2011, pp. 477-482, doi: 10.1109/ICIINFS.2011.6038117.

[8] R. F. Salas-Moreno, R. A. Newcombe, H. Strasdat, P. H. J. Kelly and A. J. Davison, "SLAM++: Simultaneous Localisation and Mapping at the Level of Objects," 2013 IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, 2013, pp. 1352-1359,

doi: 10.1109/CVPR.2013.178.

[9] R. Mur-Artal, J. M. M. Montiel and J. D. Tardós, "ORB-SLAM: A Versatile and Accurate Monocular SLAM System," in IEEE Transactions on Robotics, vol. 31, no. 5, pp. 1147-1163, Oct. 2015, doi: 10.1109/TRO.2015.2463671.

[10] Y. Diao, R. Cen, F. Xue and X. Su, "ORB-SLAM2S: A Fast ORB-SLAM2 System with Sparse Optical Flow Tracking," 2021 13th International Conference on Advanced Computational Intelligence (ICACI), Wanzhou, China, 2021, pp. 160-165, doi: 10.1109/ICACI52617.2021.9435915.

[11] C. Campos, R. Elvira, J. J. G. Rodríguez, J. M. M. Montiel and J. D. Tardós, "ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual-Inertial, and Multimap SLAM," in IEEE Transactions on Robotics, vol. 37, no. 6, pp. 1874-1890, Dec. 2021, doi: 10.1109/TRO.2021.3075644.

[12] T. Qin, P. Li and S. Shen, "VINS-Mono: A Robust and Versatile Monocular Visual-Inertial State Estimator," in IEEE Transactions on Robotics, vol. 34, no. 4, pp. 1004-1020, Aug. 2018, doi: 10.1109/TRO.2018.2853729.

[13] Qin, T., Pan, J., Cao, S., and Shen, S., "A General Optimization-based Framework for Local Odometry Estimation with Multiple Sensors", <i>arXiv e-prints</i>, 2019. doi:10.48550/arXiv.1901.03638.

[14] S. Song, H. Lim, A. J. Lee and H. Myung, "DynaVINS: A Visual-Inertial SLAM for Dynamic Environments," in IEEE Robotics and Automation Letters, vol. 7, no. 4, pp. 11523-11530, Oct. 2022, doi: 10.1109/LRA.2022.3203231.

[15] D. Esparza and G. Flores, "The STDyn-SLAM: A Stereo Vision and Semantic Segmentation Approach for VSLAM in Dynamic Outdoor Environments," in IEEE Access, vol. 10, pp. 18201-18209, 2022, doi: 10.1109/ACCESS.2022.3149885.

[16] Wu, W., Guo, L., Gao, H. et al. YOLO-SLAM: A semantic SLAM system towards dynamic environment with geometric constraint. Neural Comput & Applic 34, 6011–6026 (2022). <https://doi.org/10.1007/s00521-021-06764-3>

[17] M. Rünz and L. Agapito, "Co-fusion: Real-time segmentation, tracking and fusion of multiple objects," 2017 IEEE International Conference on Robotics and Automation (ICRA), Singapore, 2017, pp. 4471-4478, doi: 10.1109/ICRA.2017.7989518.

[18] Lin, TY. et al. (2014). Microsoft COCO: Common Objects in Context. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds) Computer Vision – ECCV 2014. ECCV 2014. Lecture Notes in Computer Science, vol 8693. Springer, Cham. https://doi.org/10.1007/978-3-319-10602-1_48

[19] M. Runz, M. Buffier and L. Agapito, "MaskFusion: Real-Time Recognition, Tracking and Reconstruction of Multiple Moving Objects," 2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), Munich, Germany, 2018, pp. 10-20, doi: 10.1109/ISMAR.2018.00024.

[20] He, K., Gkioxari, G., Dollár, P., and Girshick, R., "Mask R-CNN", <i>arXiv e-prints</i>, 2017. doi:10.48550/arXiv.1703.06870.

[21] S. Cheng, C. Sun, S. Zhang and D. Zhang, "SG-SLAM: A Real-Time RGB-D Visual SLAM Toward Dynamic Scenes With Semantic and Geometric Information," in IEEE Transactions on Instrumentation and Measurement, vol. 72, pp. 1-12, 2023, Art no. 7501012, doi: 10.1109/TIM.2022.3228006.

4. 已有的工作积累和研究成果

- 1) 研究生阶段，学习了 SLAM 的基础知识，在 SLAM 方向有一定基础。
- 2) 在自己的 PC 机上搭建环境，使用单目、双目以及深度相机测试过 ORB-SLAM2 和 ORB-SLAM3 框架进行稀疏建图，并修改代码使用数据集进行稠密建图的测试。

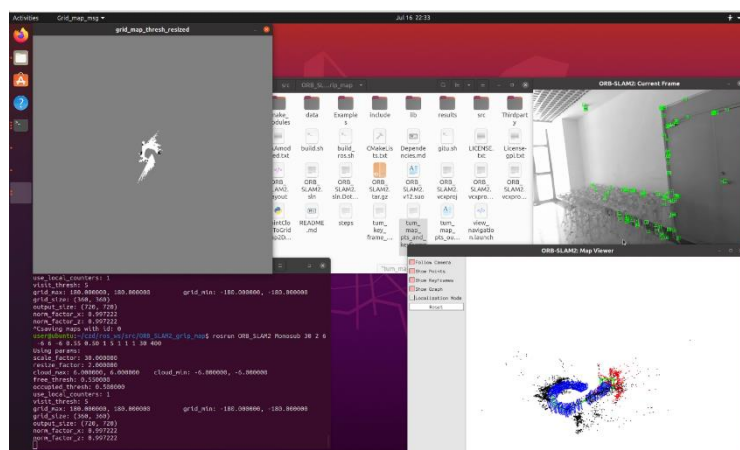


图 1 单目实时稀疏建图

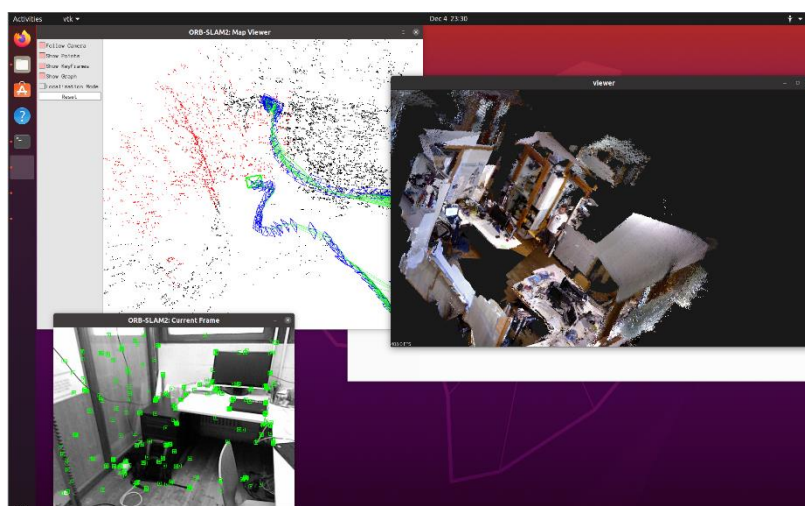


图 2 数据集构建稠密点云地图

- 3) 简单了解过一些目标识别的算法理论，在目标识别方面有一定的基础

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. 拟采取的主要理论

SLAM 系统通常包括五个模块：传感器数据、前端里程计、后端优化、回环检测以及建图。

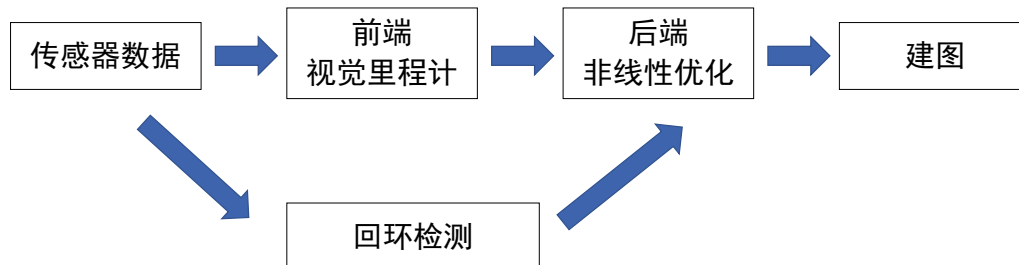


图 3 SLAM 系统流程

在 SLAM 中传感器有激光雷达、相机、惯性测量单元等，主要为整个 SLAM 系统获取环境数据并做一些预处理。前端视觉里程计在相邻图像之间追踪特征点或特征描述符的运动，从而推断相机的位姿变化。后端非线性优化后端不同时刻视觉里程计测量的相机位姿，以及回环检测的信息，以最小化误差，提高对机器人轨迹和环境地图的估计精度，得到全局一致的轨迹和地图。回环检测判断机器人是否到达过先前的位置。如果检测到回环，它会把信息提供给后端进行处理。建图部分根据估计的轨迹，建立与任务要求对应的地图。

2. 研究方法

1) 文献法

通过文献查询，尤其是关注相关领域具有显著影响力的会议和期刊上近期发表的文章，深入了解了语义 SLAM 的最新研究方向。在这个过程中，系统性地总结与语义 SLAM 相关的问题、技术等关键内容。通过交流和深入探讨，对这些文献中采用的方法进行了研究和分类，以更全面、深入的视角理解语义 SLAM 领域的最新发展趋势。这一过程不仅可以加深对当前研究方向的理解，也为进一步的研究和创新提供了有益的参考和启示。

2) 实验法

通过对文献中采用的方法进行复现，整理实验数据，并对比各种算法的实验结果，从数据上直观地评估这些算法。通过实践的过程，不仅可以加深对各种算法的理解，还可以提炼并形成对这些算法性能和特性的深刻认识。这样的实验过程不仅有助于积累经验，提升实际操作的能力，还为提出新的算法改进的思路奠定基础。动手复现算法，还能够更全面地理解算法的细节，从而为今后的研究提供更坚实的实验基础。

3. 技术路线和实施方案

本论文的总框架如图 4 所示：

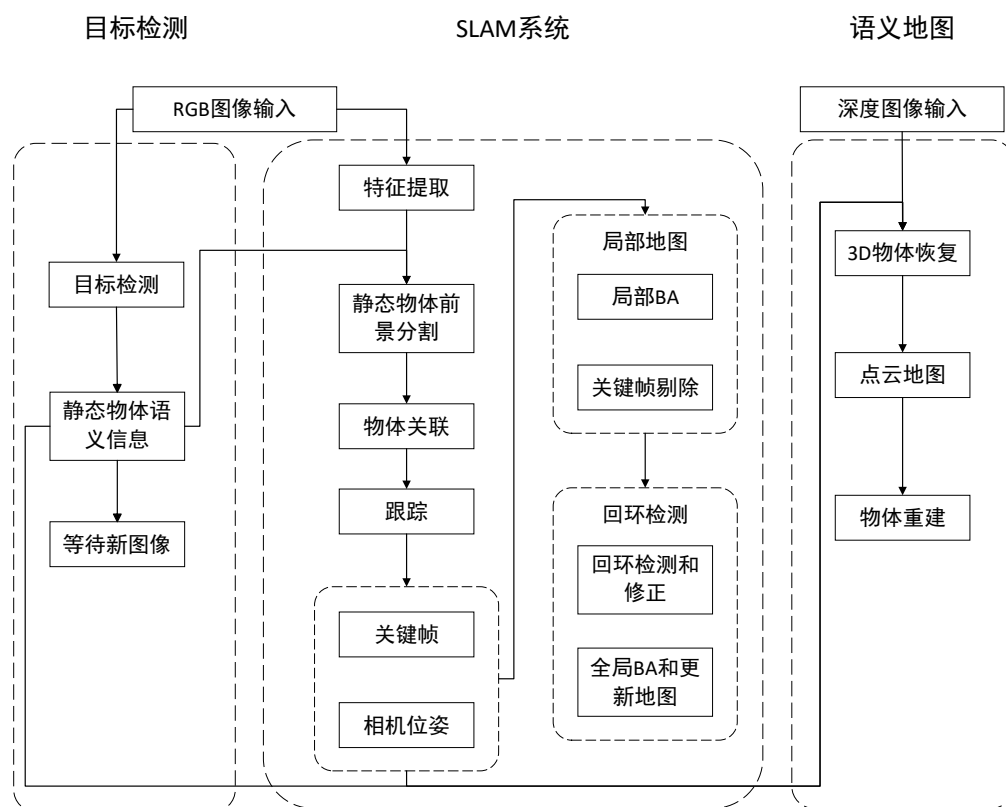


图 4 算法框架

室内静态场景下的语义 SLAM 算法研究需要关注静态物体的准确的前景信息静态物体的数据关联以及静态物体的准确重建等关键技术方面。主要需要研究物体的前景分割算法、静态特征点关联物体以及物体之间的关联、物体在地图上的精准表示等。

图 4 描述的语义 SLAM 的整体框架，主要分三个模块，目标检测模块、SLAM 系统模块以及语义地图模块。目标检测模块对物体进行分割，精确获取目标 2D 检测框，以及物体标签等语义信息。SLAM 框架系统接收目标检测部分的输入，获取物体的大致位置信息以及对应的语义信息，通过深度、灰度等信息精确分割出物体。语义地图模块将 2D 的物体正确关联到 3D 物体，并附上语义信息之后，将其重建在全局地图上。

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. 研究计划可行性

根据文献调研的结果，当前的 SLAM 系统已经成功将目标识别算法融入系统中，以提供更多的信息来对物体进行一定的重建、参与整体优化，从而辅助 SLAM 系统进行更准确的自身位姿估计。在实际应用中，对动态物体的判断表现出较好的效果。一些 SLAM 系统采用 MSK-RCNN 网络进行像素级掩码分割，还有使用诸如 YoLo 等框架进行目标识别，仅获取物体的 2D 检测框信息。

除了利用语义信息，一些文献中还探讨了使用几何信息的方法。通过一系列算法，如极线约束、运动不一致性、深度不一致性等，利用特征点的深度、位置信息来判断连续帧中匹配点是否发生了移动。综合利用语义和几何信息的方法为 SLAM 系统提供了更多判断动态物体的手段，为提高系统鲁棒性和环境理解能力提供了有效途径。

2. 研究条件

1) 深度摄像机 Astra Pro，是一个单目结构光相机，有一个 RGB 摄像头、一个 IR 摄像头以及一个深度处理芯片和麦克风。

2) rk-3588 开发板，集成了 4 个高性能的 CPU 核心和 4 个低功耗 CPU 核心，同时内置了一个高频率的 GPU 和一个 NPU 协处理器。

3) 使用的数据集，如 KITTI、TUM、shibuya 等。

3. 可能存在的问题及解决办法

1) 静态物体前景分割不准确问题

目标检测框中含有丰富的背景信息，特别是在目标与背景相似或有重叠的情境下，使用前景分割算法对物体进行背景分离可能导致物体边界分割的不够精确。直接使用目标检测的结果检测框时，就会受到背景的混淆，导致对物体的建模产生较大的误差。针对这个问题，可以利用边界上像素点深度值和灰度值的不一致性来进行物体边界的优化。通过分析深度和灰度在物体与背景过渡区域的变化，可以更精确地确定物体的边界。从而优化物体边界的分割结果，为后续的物体重建和场景理解提供更准确的信息。

2) 物体太近导致的数据关联错误问题

当多个物体距离非常近或者由于前后遮挡而导致它们的目标框重叠时，可能导致物体关联的错误，从而影响物体在优化、重建等过程中的精确度。针对这个问题，可以利用物体上的特征点前后匹配结果，借助刚性物体上特征点相对不改变的特性，以及物体运动的一致性条件，对物体进行更精确的分离和重建。

3. 研究计划及预期成果		
研究计划	起止年月	完成内容
	2023.12-2024.02	研究静态物体的前景分割算法
	2024.02-2024.04	研究静态物体关联方法
	2024.04-2024.06	研究物体在语义地图上的准确重建方法
	2024.06-2024.09	优化算法，实现更好的效果，发表学术论文一篇
	2024.09-2024.12	进行实验，获取数据，发表专利一篇
	2024.12-2025.03	完成毕业论文并准备毕业答辩
预期创新点及成果形式	<p>1. 预期创新点</p> <p>1) 提出基于深度和灰度值的静态物体前景分割方法</p> <p>2) 提出物体数据关联关系的方法</p> <p>3) 提出高质量构建 3D 语义物体的方法</p> <p>2. 成果形式</p> <p>(1) 学术论文 发表一篇高质量的学术论文。</p> <p>(2) 专利 申请专利不少于一篇。</p> <p>(3) 论文 撰写硕士学位论文一篇。</p>	

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：

同意开题

校内导师（组）签字：王治国

2023 年 12 月 25 日

校外导师签字：梁熙

2023 年 12 月 25 日

2. 开题报告考评组意见

开题日期	2023.12.21	开题地点	
考评专家	庄杰，游长江，朱学勇		
考评成绩	合格____票	基本合格____票	不合格____票
结 论	<div><input type="checkbox"/>通过</div> <div><input type="checkbox"/>原则通过</div> <div><input type="checkbox"/>不通过</div> <div>通过：表决票均为合格</div> <div>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</div> <div>不通过：表决票中有 2 票及以上为不合格</div>		

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