

硕士学位论文

基于多传感器融合的室内移动机器人  
定位算法研究

**RESEARCH OF INDOOR MOBILE  
ROBOT LOCALIZATION ALGORITHM  
BASED ON MULTI-SENSOR FUSION**

哈尔滨工业大学

2018 年 6 月

国内图书分类号：TP242.2  
国际图书分类号：681.5

学校代码：10213  
密级：公开

## 工程硕士学位论文

# 基于多传感器融合的室内移动机器人 定位算法研究

硕 士 研 究 生：李巍  
导 师：孙立宁 教授  
申 请 学 位：工程硕士  
学 科：机械工程  
所 在 单 位：机电工程学院  
答 辩 日 期：2018 年 6 月  
授 予 学 位 单 位：哈尔滨工业大学

Classified Index: TP242.2

U.D.C: 681.5

Dissertation for the Master Degree in Engineering

**RESEARCH OF INDOOR MOBILE ROBOT  
LOCALIZATION ALGORITHM BASED ON  
MULTI-SENSOR FUSION**

<b>Candidate:</b>	Li Wei
<b>Supervisor:</b>	Prof. Sun Lining
<b>Academic Degree Applied for:</b>	Master of Engineering
<b>Speciality:</b>	Mechanical Engineering
<b>Affiliation:</b>	School of Mechatronics Engineering
<b>Date of Defence:</b>	June, 2018
<b>Degree-Conferring-Institution:</b>	Harbin Institute of Technology

## 摘 要

随着老龄化社会的到来，青年劳动力的减少，发展家庭服务型机器人已成为迫在眉睫的需求。智能扫地机器人作为服务机器人中第一个能够落地的产品，已经吸引了大批的研究者和工程技术人员研究相关的技术，这就要求智能扫地机器人能够实时估计自己的位姿信息，同时进行路径规划完成全覆盖清扫，从而使得 EKF 数据融合、视觉 SLAM 和紧耦合算法成为研究中的关键技术。本文对智能扫地机器人室内定位算法涉及的相关技术进行研究，主要工作如下：

首先，研究了光流计定位原理，针对单个传感器无法满足所有的应用场景问题，改进了一种基于扩展卡尔曼滤波器对多传感器数据进行融合的方法。该方法建立了移动机器人的系统模型和观测模型，然后设计了适用于扫地机器人平台的扩展卡尔曼滤波器。通过实验，验证了基于 EKF 融合较里程计+IMU 融合（或光流计+IMU 融合）能够在一定程度上克服尺度不一致的问题，提高室内移动机器人航位推算定位的精度和鲁棒性。

其次，研究了 ORB-SLAM2 算法框架和三线程工作原理，针对目前 ORB-SLAM2 中加载 DBoW2 词汇库耗时长和占用大量系统内存等问题，改进了一种对 DBoW2 词汇库进行优化的方法；同时训练得到了针对家庭场景下的词汇库。另一方面，针对三维空间 6-DoF 下优化结果误差大、实时性差等问题，改进了一种平面移动机器人 3-DoF SLAM 优化算法。通过实验，验证了所提出的算法较优化前有显著的性能提升。

再次，针对目前主流的预积分算法误差大、精度低等问题，提出了一种融合多传感器的惯导预积分算法；同时利用松耦合算法估计了初始化视觉尺度  $s$ 。针对松耦合算法的定位精度低等问题，提出了一种基于多传感器融合的紧耦合算法，并推导了系统模型和观测模型。

最后，完成了 Create2 平台的硬件搭建和相应的软件设计，然后在不同材质表面进行了定位对比实验、室内小场景和大场景下两种算法的定位精度对比实验、以及针对本文提出的紧耦合算法进行了 VISLAM 系统优化的对比实验。实验结果表明本文所提出的基于多传感器融合的紧耦合算法能够适用于大部分的室内场景，且受表面材质、纹理特征、光照等条件的影响小。

**关键词：**扫地机器人；EKF；VISLAM；多传感器融合；紧耦合

## Abstract

With the advent of the aging society and the reduction of youth labor force, the development of family service robots has become an imminent need. Intelligent sweeping robot as the first landing product service robot, has attracted a large number of researchers and Engineers research related technologies. This requires intelligent sweeping robot real-time pose estimation to their own information, at the same time to complete the full coverage path planning for cleaning, so that the multi-sensor fusion , visual SLAM and tight coupling algorithm has become the key technology in the research. Research on the relevant technology of intelligent sweeping robot indoor positioning algorithm are involved, the main works are as follows:

First of all, the principle of optical flow meter positioning is studied. Aiming at the problem that a single sensor can not meet all the application scene problems, a method of multi-sensor data fusion based on extended Calman filter is proposed. This method establish the system model and observation model of the mobile robot, and then designs the extended Calman filter for the sweeping robot platform. Through experiments, it is verified that the dead reckoning algorithm based on EKF data fusion has higher accuracy and robustness than traditional dead reckoning algorithm.

Secondly, the framework of ORB-SLAM2 algorithm and the principle of three thread works are studied. In view of the problems of loading DBoW2 vocabulary in ORB-SLAM2 and taking up a large amount of system memory, a method of optimizing the DBoW2 vocabulary library is proposed, and the vocabulary library for home scene is trained. On the other hand, a 3-DoF SLAM optimization algorithm for planar mobile robot is proposed to solve the problem of large error and poor real-time performance in three-dimensional space 6-DoF. Experiments show that the proposed algorithm performs better than before.

Thirdly, in order to solve the problems of large error and low precision of the current preintegration algorithm, a multi-sensor fusion inertial navigation preintegration algorithm is proposed, and the initialization visual scale  $s$  is estimated by the loosely coupled algorithm. A compact tight coupling algorithm

based on multi-sensor fusion is proposed, and the system model and observation model are derived in order to solve the problem of low accuracy of the loosely coupled algorithm.

Finally, the hardware construction and the corresponding software design of the Create2 platform are completed. Then the positioning precision comparison experiments of two algorithms are carried out on the different material surface, and the comparison experiments of the VISLAM system optimization are carried out according to the tight coupling algorithm proposed in this paper. The experimental results show that this proposed tight coupling algorithm based on multi-sensor fusion can be applied to most of the indoor scenes, which are affected by the surface materials, texture features, illumination condition of small.

**Keywords:** Intelligent Sweeping Robot, EKF, VISLAM, Multi-Sensor Fusion, Tight Coupling Algorithm

## 目 录

摘 要 .....	I
ABSTRACT.....	II
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 智能扫地机器人研究现状 .....	2
1.2.2 SLAM 技术研究现状 .....	4
1.2.3 多传感器融合技术研究现状 .....	6
1.2.4 紧耦合的视觉惯导 SLAM 技术研究现状 .....	7
1.3 本文的主要研究内容 .....	9
第 2 章 基于 EKF 数据融合的定位算法研究 .....	11
2.1 引言 .....	11
2.2 光流计定位原理研究 .....	11
2.2.1 光流计简介 .....	11
2.2.2 光流计测试实验 .....	13
2.3 移动机器人系统模型和观测模型的建立 .....	14
2.3.1 移动机器人系统模型的建立 .....	14
2.3.2 传感器观测模型的建立 .....	16
2.4 扩展卡尔曼滤波器设计 .....	18
2.5 三种定位方案实验对比 .....	19
2.5.1 融合模型的建立 .....	19
2.5.2 实验结果 .....	20
2.6 本章小结 .....	25
第 3 章 视觉 SLAM 算法优化研究 .....	26
3.1 引言 .....	26
3.2 ORB-SLAM2 算法研究 .....	26
3.3 DBoW2 词汇库优化研究 .....	28
3.3.1 DBoW2 词汇库优化 .....	28
3.3.2 DBoW2 优化实验 .....	30
3.4 基于平面 3-DOF VSLAM 算法优化研究 .....	32
3.4.1 平面 3-DOF VSLAM 算法优化 .....	32
3.4.2 平面 3-DOF VSLAM 算法仿真 .....	35

3.5 本章小结 .....	36
第 4 章 基于多传感器融合的紧耦合算法研究 .....	38
4.1 引言 .....	38
4.2 VIO 系统架构 .....	38
4.2.1 系统描述 .....	38
4.2.2 惯导预积分 .....	39
4.3 视觉惯导初始化状态估计 .....	40
4.3.1 陀螺仪零偏估计 .....	41
4.3.2 视觉尺度估计 .....	41
4.4 紧耦合的视觉惯导融合算法 .....	42
4.4.1 滑动窗口模型 .....	42
4.4.2 相机测量模型 .....	44
4.4.3 里程计和光流计测量模型 .....	45
4.4.4 IMU 测量模型 .....	45
4.5 本章小结 .....	46
第 5 章 移动机器人平台搭建及实验研究 .....	47
5.1 引言 .....	47
5.2 移动机器人平台搭建 .....	47
5.2.1 Create2 硬件平台搭建 .....	47
5.2.2 STM32F407 平台软件设计 .....	49
5.3 Create2 平台实体实验 .....	50
5.3.1 不同材质表面对比实验 .....	50
5.3.2 两种场景下的定位实验 .....	52
5.3.3 VISLAM 系统优化对比实验 .....	57
5.4 本章小结 .....	59
结 论 .....	61
参考文献 .....	62
哈尔滨工业大学学位论文原创性声明和使用权限 .....	67
致 谢 .....	68



# 第 1 章 绪论

## 1.1 课题研究背景及意义

近年来，在机器视觉、多传感器融合、人工智能等学科的研究成果显著，大量的研究者和工程技术人员将其用于无人驾驶汽车、机器人、四旋翼无人机上<sup>[1]</sup>，用于提升用户体验。机器人技术已经发展到第三代智能型机器人阶段，第三代智能机器人能够按照主人的指令或控制自主完成一些简单的功能，例如：陪伴小孩学习、搬运小物品、自主清扫、用于安防监控等，因此第三代智能型机器人成为产业界和学术界的研究热点领域。如图 1-1 所示机器人三代的发展。

另一方面，随着老龄化社会的到来，青年劳动力的减少，人们对于家庭服务型机器人的需求日益增加<sup>[2]</sup>，同时也促使了家庭服务型机器人得到了快速发展。从小孩陪伴机器人到家庭安防监控机器人，从智能扫地机器人到炒菜机器人，从机器人宠物到智能管家机器人。家庭服务机器人将逐渐融入我们的生活，帮我们做一些简单枯燥的家务，例如：清扫、安防监控等。比尔盖茨曾预言，机器人将与人类共同生活。

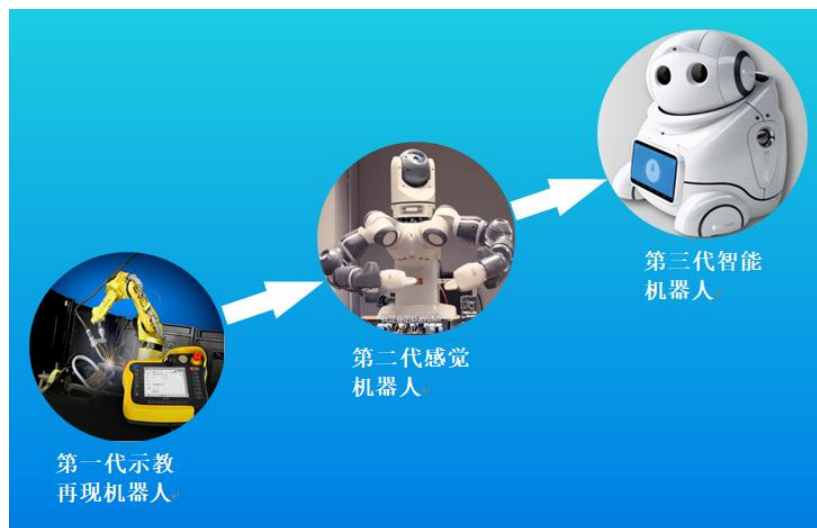


图 1-1 机器人三代

目前低成本的室内移动机器人一般由航位推算定位算法提供定位信息（里程计+IMU），航位推算定位算法存在一个很大的问题就是误差会随着时间的增加而累积。因此在对定位要求比较高的移动机器人上，航位推算定位算法不能满足应用的需求。而目前一些处于实验室阶段的室内移动机器人通过搭载高精

度的激光雷达、陀螺仪，可以提供准确的定位信息，但这对于低成本的室内移动机器人应用并不现实。

综合以上在低成本室内移动机器人上无法满足精确定位的需求，因此本课题着眼于满足低成本、高精度定位需求的室内移动机器人，基于多传感器融合技术研究室内移动机器人定位算法。

目前在室内环境下能够提供全局定位信息的低成本传感器只有摄像头（低成本激光雷达的价格也是摄像头的 10 倍以上），因此本文选取摄像头作为提供全局定位信息的传感器。在传统的航位推算定位算法中一般由里程计和 IMU 提供输入信息，经融合得到位姿信息。但里程计在不同材质表面上的表现性能不一致，甚至出现打滑现象，这对航位推算定位算法来说是致命的缺点。本文通过在移动机器人上引入一种全新的传感器——光流计，来解决在不同材质表面尺度不一致的问题。因此，基于多传感器的室内移动机器人定位算法研究是本课题的关键问题。

本文以智能扫地机器人为载体，通过融合里程计、光流计、IMU 和摄像头信息，给出移动机器人精确的位姿信息。本文研究的意义是基于融合多种低成本传感器研究室内移动机器人通用定位技术，为家庭服务型机器人自主移动提供准确的定位输出信息。

## 1.2 国内外研究现状

### 1.2.1 智能扫地机器人研究现状

随着老龄化成为人们日常讨论的话题，而青年群体忙于工作。大量的清扫体力劳动任务随之而至，劳动力缺乏成为一个很严肃的问题。在科技和技术的驱动下，各种以服务为导向的机器人被提出<sup>[3-6]</sup>。智能扫地机器人主要是帮助主人减小清扫灰尘、碎屑和拖地的任务量。

1997 年，世界上第一台扫地机器人<sup>[7]</sup>诞生于伊莱克斯三叶虫公司，如图 1-2 所示。它只能通过随机碰撞来进行清扫，配备有最基本的传感器，定位方面的传感器有里程计和 IMU。这种扫地机清扫效率极低，覆盖率低，且很难清扫干净。

2015 年美国 iRobot 公司推出其最新款清扫机器人 Roomba980，如图 1-3 所示。第一次在扫地机器人上应用单目视觉 SLAM 技术取代了昂贵的激光雷达传感器，第一次实现了扫地机器人规划式清扫而不是随机碰撞式，Roomba980 的发布对 VSLAM 的发展具有里程碑式的意义。Roomba980 在定位方面配备的

传感器有里程计、IMU、摄像头、光流计、以及检测打滑信号的传感器。Roomba980 第一次使得扫地机器人能够像“人”一样进行清扫，体现了家庭服务机器人的智能性。



图 1-2 世界上第一台扫地机器人



图 1-3 Roomba980 清扫机器人

2018 年科沃斯发布了其最新款的扫地机器人地宝 DJ35，如图 1-4 所示。DJ35 同样搭载了 VSLAM 视觉导航技术，并且在手机端能够实时查看清扫路径以及构建的地图。DJ35 在定位方面配备的传感器有里程计、IMU、摄像头，较 Roomba980 来说较少，但在定位方面的效果不逊于 Roomba980。路径规划方面采用“工字型”规划清扫，清扫效率较高，覆盖率高。从技术层面分析，DJ35 采用顶视 VSLAM 方案，顶视 VSLAM 较前视所需的计算资源更少、定位效果更好。

2018 年小米发布了其最新款的扫地机器人 2 代，如图 1-5 所示。小米目前在售的扫地机器人都是基于激光雷达传感器方案，相比于视觉方案的扫地机器人成本较高，同样采用规划式清扫，覆盖率高。可以在手机端实时查看清扫路径以及构建的地图。



图 1-4 地宝 DJ35 扫地机器人



图 1-5 小米扫地机器人 2 代

目前基于激光方案的 SLAM 比较成熟，在扫地机器人、AGV 小车、送餐机器人等都有应用，且定位效果较好；而基于视觉方案的 SLAM 商业应用还比

较少,2018 年是第一年将视觉 SLAM 技术比较成熟地应用于机器人上用于提供定位信息。另一方面,摄像头较激光雷达提供的信息量多且大,大量的视觉信息可用于后续人工智能方面的应用需求。因此,基于视觉方案的智能扫地机器人将往家庭服务型机器人方向发展。

### 1.2.2 SLAM 技术研究现状

Smith 和 Cheeseman 认为 SLAM<sup>[8]</sup>是解决移动机器人自主移动的关键技术,该算法于 1987 年第一次被提出来。传统的 SLAM 算法采用扩展卡尔曼滤波器<sup>[9]</sup>(EKF)来估计机器人的位姿信息。由于 EKF 对于非线性问题有较好的处理效果,因此 EKF 算法在机器人 SLAM 领域得到了广泛的应用。近年来,随着以图像为核心的机器视觉技术的进步以及芯片处理能力的快速提高,产生了大量以摄像头作为传感器的 VSLAM 算法。

目前,对于 SLAM 的研究主要集中在两种主流的算法上:基于滤波的 SLAM 算法和基于优化的 SLAM 算法。

#### (1) 基于滤波的 SLAM 算法

早期的 SLAM 算法是以机器人状态和 3D 特征点,通过对测量到的含噪声数据进行加权平均处理,并以概率的方法进行表示,该过程称为滤波。EKF 作为滤波算法的代表,被认为是 SLAM 算法的首选。在 EKF-SLAM<sup>[10]</sup>中,机器人位姿和特征点间维持着一个大的协方差矩阵用来表示不确定性。随着特征点数量的增加,EKF-SLAM 的计算复杂度成二次方增长,因此在实时性要求比较高的应用中存在局限性。

Davison<sup>[11]</sup>于 2003 年首次将 EKF 引入到实时单目 SLAM(MonoSLAM)中。在此之前,最先进的 SLAM 系统只能离线运行,实时性意味着需要在精度和计算复杂度间做出权衡。在 MonoSLAM 中,经过多重筛选的特征点被加入到地图中,加入的特征点用于计算机器人的位姿信息,此外被移除的特征点将不再被系统使用。虽然 MonoSLAM 能够实时运行,但由于计算复杂度随特征点成二次方增长,使得其仍然局限在小区域内应用。此外, MonoSLAM 用特征向量和初始未知深度估计参数来检测特征点,视觉测量中的小误差能够导致深度的巨大非线性误差,导致滤波器的不一致性。为了解决这个问题,Montiel<sup>[12]</sup>等提出了一种改进的一致性逆深度表示法,通过逆深度法表示的特征点能够允许更高精度的不确定性,因此高度的线性化能够表示更远距离的特征点。图 1-6 给出了典型的基于 EKF 的视觉 SLAM 流程图。

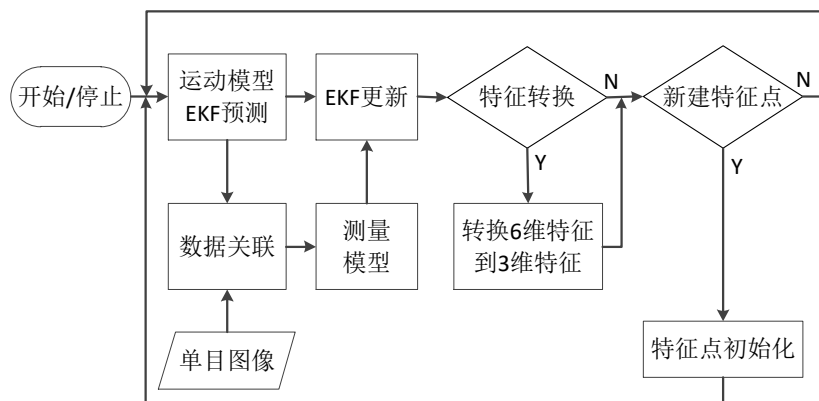


图 1-6 基于 EKF 的视觉 SLAM 流程图

## (2) 基于优化的 SLAM 算法

为了能够在单目 SLAM 系统上获得更好的实时性，Klein 和 Murray<sup>[13]</sup>等于 2007 年提出了 PTAM 算法。基于优化的 SLAM 算法是目前应用最广泛的算法，而 PTAM 就是应用了该算法。在搜索地图时，基于优化的 SLAM 算法维持一组稀疏的关键帧、特征点和相机位姿。该算法的一个优点是保留以前的信息用于当前使用，而不是像 EKF-SLAM 中以前的特征点和相机位姿被边缘化。虽然最初的 PTAM 仍然局限于一个相对较小的区域，但相比于 EKF-SLAM 显著增加了可用的特征点数。PTAM 将建图线程从跟踪线程中分离出来并行处理，这种解耦操作对于研究应用图优化的建图任务很有必要。在图优化中，每个关键帧可以表示为具有局部欧几里得约束的节点。对于实时大规模 SLAM，已经提出了多种用于图优化的方法<sup>[14,15]</sup>。PTAM 的双线程系统框图如图 1-7 所示。

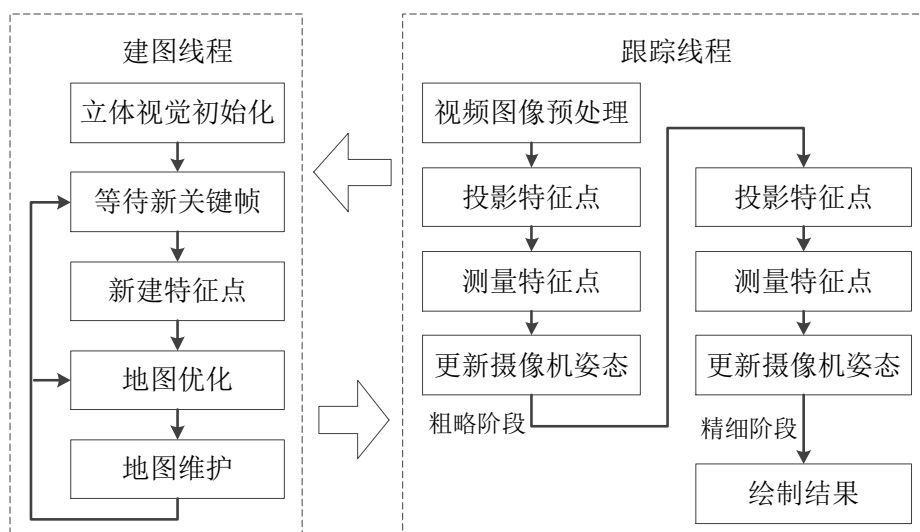


图 1-7 PTAM 双线程系统框图

Mur-Artal 和 Tardos<sup>[16]</sup>等于 2016 年提出了 ORB-SLAM2, ORB-SLAM2 可以应用于单目、双目和 RGB-D 相机。该系统主要由 3 个线程组成: 1) 跟踪线程用于寻找与本地地图特征匹配的重投影误差最小的图像帧; 2) 建图线程用于管理本地地图和执行局部 BA 优化; 3) 闭环检测用来检测大回环, 和通过位姿图修正累积误差, 该线程在位姿图优化后启动第四个线程来执行全局 BA, 以计算最佳的运动估计。该系统还嵌入了基于 DBOW2<sup>[17]</sup>的位置识别功能, 用于在跟踪丢失、已创建地图的环境中重新初始化、或者回环检测时进行重定位。该系统维持一个共视图, 用于连接观察到的公共点的任意两个关键帧和连接所有关键帧的最小生成树。

浙江大学章国峰课题组于 2016 年提出了 RKSLAM<sup>[18]</sup>, RKSLAM 主要为了解决在强旋转下快速运动的鲁棒性问题, 该系统首先提出了一种基于多单应性的鲁棒、高速的特征追踪方法。基于此, 又提出了一种实时局部地图的扩展方案, 对所观察到的特征点进行无延迟的三角化操作。针对 AR 领域的应用, 提出了一种基于滑动窗口的相机姿态优化框架, 该框架通过真实或模拟的 IMU 数据进行运动先验约束。

清华大学高翔博士于 2015 年提出了一种基于平面点特征的 SLAM 算法<sup>[19]</sup>。由于目前在机器人估计位姿和测绘研究中, RGB-D 相机是很受欢迎的传感器, 而基于 RGB-D 相机的 SLAM 系统采用迭代最近点 (ICP) 算法<sup>[20]</sup>来对齐点特征点, 通过测量的深度数据获取其空间位置。然而, 由于视觉特征往往位于真实物体的边缘, 用于计算特征的深度数据常常受到噪声的影响, 因此提出了该 SLAM 算法。平面特征在一定程度上能够实现实时性应用, 并且提高 ICP 算法的准确性和鲁棒性。

### 1.2.3 多传感器融合技术研究现状

多传感器数据融合<sup>[21-23]</sup>是一种能够组合来自多个输入信息以形成统一的数据输出的技术。数据融合在传感器网络、机器人、图像处理、智能系统设计等领域有着广泛的应用, 数据融合是综合分析多传感器的原始数据, 并结合其差异性与互补性, 来获得稳定、可靠的输出数据, 优点主要是增强数据的可靠性和可用性。近年来随着传感器技术的飞速发展, 国内外学者运用多传感器技术结合路径规划技术来实现移动机器人的高智能化、高自由度的运动。

Lynen.S 和 Achtelik.M<sup>[24]</sup>等提出了通用的多传感器扩展卡尔曼滤波器 (MSF-EKF) 框架, 在理论上能够处理不限数量、不限类型传感器的延时、相

对和绝对的测量信息，同时能够进行传感器自校准。**MSF-EKF** 的模块化允许在运行过程中处理增加或丢失的传感器数据，同时采用迭代 **EKF**<sup>[25]</sup> (**IEKF**) 更新状态缓冲区，以允许对预测进行有效的重新线性化，来获得绝对状态和相对状态的最佳线性化点。

**Quan.M** 和 **Piao.S**<sup>[26]</sup>等提出了采用紧耦合方案融合相机、里程计和 **IMU** 数据，以得到准确的机器人位姿信息，首先分别对里程计和 **IMU** 的原始数据进行预积分，然后将预积分结果与图像帧一起放入优化模型中得到优化后的机器人位姿。

**K.Paul** 和 **Kejian.W**<sup>[27]</sup>等第一次提出了基于双目 **VINS** 的紧耦合方案应用于实时的移动处理器上，详细地比较了单目、双目和多目 **VINS** 在不同场景和不同运动参数下的运动估计，并提供了从多目过渡到双目时信息矩阵增益的理论解释。

#### 1.2.4 紧耦合的视觉惯导 SLAM 技术研究现状

随着 **MEMS** 技术的进步使得惯性测量单元 (**IMU**) 小体积、低成本和低功耗成为可能。因此，基于 **IMU** 的惯性导航系统 (**INS**) 已经被广泛应用于智能汽车、四旋翼和机器人的运动状态估计中。然而由于误差累积问题，仅利用 **IMU** 很难得到可靠的结果。为了解决该问题，一些 **INS** 依靠 **GPS** 信号周期性地校正 **IMU** 的累积误差。但是在一些没有 **GPS** 信号的环境中 (例如：室内、地下、隧道等)，这些方法就失去了作用。另外，**GPS** 信号质量也会受环境干扰。

相机是校正 **IMU** 的另一种选择，相机低成本、小体积，能够提供丰富的周围环境信息。通过跟踪多帧图像的特征点，可以估计相机的运动轨迹<sup>[28]</sup>。相机和 **IMU** 是两个互补的传感器，**IMU** 在快速运动下具有较低的测量不确定性，相机能够在慢速运动下精确跟踪特征点。另一方面，单目 **SLAM** 没有尺度信息，需要依靠 **IMU** 来给出真实的尺度信息<sup>[29-31]</sup>。基于上述原因，结合单目相机和 **IMU** 进行传感器融合来估计机器人自身位姿信息<sup>[32-36]</sup>是必然的选择。

根据直接或间接地融合传感器的测量信息，视觉惯导里程计 (**VIO**) 分为两种主流的方法：松耦合和紧耦合。松耦合方法<sup>[31,37]</sup>是分别估计图像序列和 **IMU** 数据帧的相对运动，并融合两种数据的估计结果以获得机器人最优的位姿信息。紧耦合方法<sup>[38,39]</sup>是使用一个估计器通过融合原始相机和 **IMU** 测量数据来得到最优估计。相比于松耦合方法，紧耦合方法一般能够得到更加精确和鲁棒性的结果。紧耦合 **VIO** 方法<sup>[40]</sup>可以通过测量模型中线性化的数量来分类，基于

EKF 的方法仅在更新步骤中线性化一次，而非线性优化的方法在优化步骤中线性化多次。

基于滤波的方法<sup>[41,42]</sup>融合 IMU 测量信息用于预测阶段，然后用视觉测量信息来更新当前状态。由于 3D 特征点用向量形式表示，因此 EKF 方法的计算复杂度随特征点数量成二次方增长。为了解决该问题，Mourikis 和 Roumeliotis<sup>[33]</sup>提出了多状态约束卡尔曼滤波器（MSCKF），能够从状态向量中提取特征点坐标，这种方法的缺点是用于更新状态的特征点可能会移出相机视野，即意味着不是所有的特征点都会使用到。此外，线性化误差会使得滤波器不一致。

基于优化的方法通过最小化 IMU 测量残差和视觉重投影残差的联合非线性代价函数来求得最优估计。因此相对于滤波的方法，基于优化的方法<sup>[40]</sup>能够重复在不同特征点处线性化状态向量以获得更高的精度。IMU 的测量是通过在图像帧间进行预积分得到。然而，正确的 IMU 预积分与第一帧中 IMU 的初始状态密切相关。当状态估计发生改变时，需要重新计算所有 IMU 测量的预积分。

IMU 预积分广泛地应用于基于优化的 VIO 算法<sup>[43,44]</sup>上。Lupton 和 Sukkarieh<sup>[45]</sup>等于 2012 年第一次提出了高效地计算预积分而不是单纯的积分计算，作者提出了在局部参考坐标系中使用 IMU 测量数据动态的离散积分，以避免在每一次优化步骤中重新积分。虽然此方法解决了计算复杂度问题，但由于使用欧拉角作为旋转表示存在奇异问题。为了提高该算法的稳定性，Forster 和 Carlone<sup>[40]</sup>等于 2017 年引入了流形表示，在  $SO(3)$ <sup>[46]</sup>流形上给出了无奇异的方向表示，通过使用旋转矩阵来替代使用欧拉角进行 IMU 预积分，将 IMU 预积分引入到一个有效的图优化算法（VINS）中。Liu<sup>[47]</sup>等提出了连续的 IMU 预积分方法。尽管基于优化的方法已经达到了高精度，但是将越来越多的特征点加入到优化中时所需要的计算资源也急剧增加。

Leutenegger 和 Lynen<sup>[39]</sup>等于 2015 年提出了基于关键帧的 VINS 方法（OKVIS），即在非线性优化中使用了一组无序的过去相机位姿和一系列最近的惯导状态，再结合 IMU 测量数据用于精确的位姿估计。这些惯性因素采取状态预测的形式，每当惯性状态的线性化点开始改变时，就需要重新计算 IMU 预积分。这在惯导测量处理上表现出低效性，并且大量的惯导数据不可用。

香港科技大学沈劭劼课题组于 2015 年介绍了连续预积分方法<sup>[43]</sup>，但在没有提供闭环形式解的情况下，它们仍然离散地采样测量数据。这使得从时间连续性角度来看，距预积分理论的理论完整性还有很大的差距。同样在 2017 年又提出了一种基于智能手机的单目视觉惯导 SLAM 系统<sup>[48-50]</sup>。该系统是针对于增强现实（AR）领域的应用，但缺乏直接的距离度量测量和相机的快速运动，对



AR 设备的定位具有重大的挑战。因此，提出了一种基于紧耦合、图优化的单目视觉惯性状态估计方法，用于在复杂的室内和室外环境中进行鲁棒性的相机定位。该方法不需要任何人为的标记，并且能够通过单目相机的设置来恢复测量尺度。整个系统能够在线初始化，而不依赖于对环境的任何前提条件。并开发了一个轻量级的闭环检测模块，该检测模块与状态估计紧密集成，以消除漂移。

### 1.3 本文的主要研究内容

本文的目标是研制一款能够在室内环境下提供精确定位信息的智能扫地机器人，提高扫地机器人的清扫效率、覆盖率以及智能性，提升用户体验。本文采用低成本的摄像头、里程计、光流计和 IMU 进行多传感器数据融合，研究针对家庭环境下的移动机器人定位算法，主要涉及如下几个关键技术：EKF 数据融合、视觉 SLAM、预积分算法、紧耦合算法。本文主要研究内容包括：

#### （1）基于 EKF 数据融合的定位算法研究

对光流计定位原理进行研究，针对单个传感器无法满足所有的应用场景问题，改进一种基于扩展卡尔曼滤波器对多传感器数据进行融合的方法。首先建立移动机器人的系统模型，和里程计、光流计和 IMU 的观测模型，根据系统模型和观测模型设计扩展卡尔曼滤波器。最后通过对比实验，验证所提出算法的可行性。

#### （2）视觉 SLAM 算法优化研究

对 ORB-SLAM2 算法框架和三线程工作原理进行研究，针对目前 ORB-SLAM2 中加载 DBoW2 词汇库耗时长和占用大量系统内存等问题，改进一种对 DBoW2 进行优化的方法；同时训练得到针对家庭场景下的词汇库。另一方面，针对三维空间 6-DoF 下优化结果误差大、实时性差等问题，改进一种平面移动机器人 3-DoF SLAM 优化算法。最后通过仿真实验，对算法的可行性进行验证。

#### （3）基于多传感器融合的紧耦合算法研究

针对目前主流的预积分算法误差大、精度低等问题，提出一种融合里程计、光流计和 IMU 的惯导预积分算法。同时根据惯导预积分测量值和视觉测量值利用松耦合算法估计初始化视觉尺度  $s$ 。针对松耦合算法的定位精度低等问题，提出一种基于多传感器融合的紧耦合算法，并推导系统模型和观测模型。

#### （4）移动机器人平台搭建及实验研究

首先进行 Create2 实验平台的搭建及相应的软件设计，然后在不同材质表

面进行定位对比实验、室内小场景和大场景下的定位精度对比实验，以及针对基于多传感器融合的紧耦合算法进行 VISLAM 系统优化对比实验，对算法的鲁棒性和可行性进行验证，并对扫地机器人实际的应用场景进行定位精度测试。

## 第 2 章 基于 EKF 数据融合的定位算法研究

### 2.1 引言

通常移动机器人上会结合多种传感器（摄像头、激光雷达、里程计、IMU、超声波测距传感器<sup>[51]</sup>、红外测距传感器<sup>[52]</sup>、WIFI 定位<sup>[53]</sup>、UWB 定位<sup>[54]</sup>等）进行数据融合以获得更精确的位姿估计。在过去的几十年里，90%以上的多传感器数据融合算法都是基于 EKF 进行融合，EKF 对于非线性系统模型具有较好的估计效果，且是实时估计。

本章首先引入一种用于移动机器人定位的全新传感器——光流计，并研究光流计的定位原理。然后针对引入的光流计，设计一种基于里程计、IMU 和光流计进行航位推算定位的卡尔曼滤波器，用于实时估计移动机器人的位姿，该方案在一定程度上能够解决传统的里程计+IMU 定位效果较差的问题。

### 2.2 光流计定位原理研究

#### 2.2.1 光流计简介

空间物体投影在相机成像平面上的像素运动速度被称为光流，表示了物体的运动关系与其图像的灰度变化有关。因此光流计可以通过一系列图像帧来获得其在 X、Y 方向的运动信息。目前市面上主流的光流计有 Avago 公司研发的 ADNS-3080（如图 2-1 所示），和 PX4 四旋翼上用的 PX4FLOW<sup>[55]</sup>（如图 2-2 所示）。本文选取 ADNS-3080 光流计作为课题研究的传感器。



图 2-1 ADNS-3080 光流计



图 2-2 PX4FLOW 光流计

ADNS-3080 是一种传统的鼠标芯片传感器，它包含一个低分辨率相机和一个 DSP 处理器，用来估计图像帧中特征点的相对位移。ADNS-3080 的技术参数如表 2-1 所示，图像帧速率是光流计相机每秒拍摄的图片数量，图像帧速率越高意味着能够检测到更小的运动，进而识别更快的速度。

表 2-1 ADNS-3080 技术参数

项目	内容
图像帧速率	2000 - 6469 FPS
图像大小	30×30 Pixels
分辨率	400 或 1600 CPI
通信方式	SPI

正常情况下，ADNS-3080 能够给出  $\Delta X$ 、 $\Delta Y$ 、 $Squal$  三个值， $\Delta X$  和  $\Delta Y$  表示光流计下面平坦表面的位移信息， $Squal$  表示图像帧特征丰富情况。ADNS-3080 返回的  $\Delta X$  和  $\Delta Y$  为前后两帧图像间表面特征的平均位移，为了将上述距离转化为实际的位移信息，需要知道传感器感光器件到表面的高度  $h$ 。如果两个同样的 ADNS-3080 处于不同的高度但移动相同的距离，高度低的则认为表面特征移动速度较快。如示意图 2-3 所示，左图中小球穿过了整个相机视野，而右图中小球则穿过了 1/2 的相机视野。

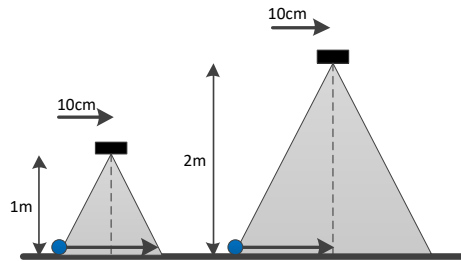


图 2-3 高度对光流计的影响

将传感器的测量值转化为 X、Y 方向的位移，取决于视场角（FOV）、相机的分辨率和传感器下单个像素移动距离所对应的尺度  $Scalar$ ：

$$Dist_{X/Y} = \frac{\Delta X / Y}{R_{OF} * Scalar} * h * 2 * \tan(FOV / 2) \quad (2-1)$$

式中  $R_{OF}$  为相机的分辨率。

ADNS-3080 同样能给出表面特征丰富的信息，4 倍的  $Squal$  值表示当前图像帧中较好的特征点数量。由于光流计类似于一个相机，需要环境有较好的照明条件，因此光流计还需要配置一个红外补光灯（红外补光灯安装在镜头的侧下方）。图 2-4 为实验平台 Create2 上光流计的安装情况，图 2-5 为 iRobot980 上光流计的安装情况。iRobot980 是第一次将光流计作为定位传感器引入到扫地机器人上，光流计更普遍的应用是在四旋翼无人机上进行定高悬停。

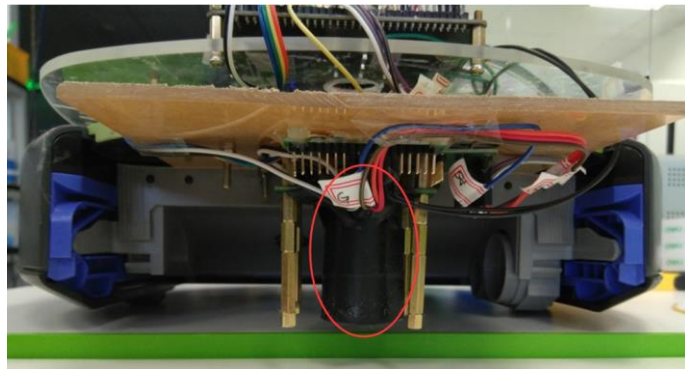


图 2-4 Create2 实验平台上光流计的安裝



图 2-5 iRobot980 上光流计的安裝

### 2.2.2 光流计测试实验

为了验证光流计能够适应各种不同材质的表面，故选取一些有代表性的表面进行测试，测试用例是光流计分别沿其自身的 X、Y 正方向移动 1m 的距离，然后累加  $\Delta X$  和  $\Delta Y$  的读数，得到 SumX 和 SumY。本次实验选取以下 6 种不同材质的表面进行实验，如图 2-6 所示。

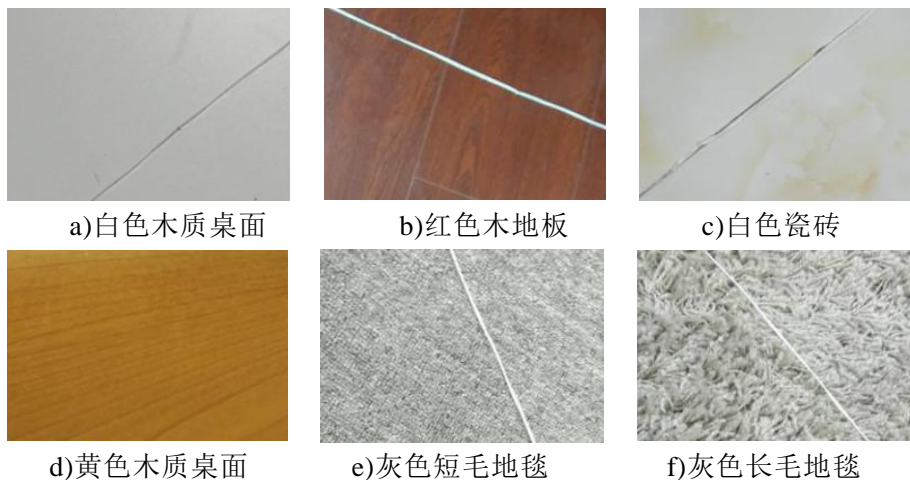


图 2-6 ADNS-3080 在不同材质表面测试

ADNS-3080 在上述 6 种材质表面上的测试结果如表 2-2 所示，表中位移累加的均值为 2406.25，由此可得出以下结论：1) ADNS-3080 在特征比较丰富的表面上效果较好（如表面 b/e/f）；2) ADNS-3080 在反光材质表面上效果没有亚光材质好；3) ADNS-3080 的定位误差最大为 4.52%，符合短距离高精度定位的需求（适用于大部分材质表面）。4) ADNS-3080 沿 Y 方向运动的效果较 X 方向要好，因此安装时 ADNS-3080 的 Y 方向尽量与差速驱动的移动机器人前进方向一致；5) ADNS-3080 光流计对室内环境下主流的材质表面都可以应用，因此验证了 ADNS-3080 光流计可以满足室内移动机器人的定位需求。

表 2-2 ADNS-3080 在不同材质表面的测试结果

不同材质表面	移动方向	位移累加和	Squal 均值	误差
(a)	X 轴正向	2316	37	-3.89%
	Y 轴正向	2435	43	+1.18%
(b)	X 轴正向	2345	58	-2.61%
	Y 轴正向	2471	62	+2.62%
(c)	X 轴正向	2302	29	-4.52%
	Y 轴正向	2359	32	-2.00%
(d)	X 轴正向	2348	55	-2.48%
	Y 轴正向	2452	57	+1.86%
(e)	X 轴正向	2448	110	+1.71%
	Y 轴正向	2466	112	+2.42%
(f)	X 轴正向	2461	97	+2.22%
	Y 轴正向	2472	101	+2.66%

## 2.3 移动机器人系统模型和观测模型的建立

### 2.3.1 移动机器人系统模型的建立

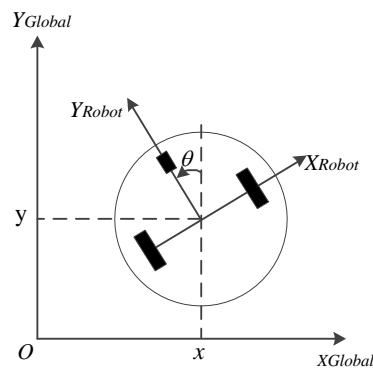



图 2-7 机器人坐标系与全局参考坐标系间的关系

为了确定机器人在全局参考坐标系下的位姿信息，需要建立机器人坐标系与全局参考坐标系间的关系，如图 2-7 所示。

$X_{Global}$ 、 $Y_{Global}$ 、 $X_{Robot}$  和  $Y_{Robot}$  分别为全局参考坐标系和机器人坐标系的坐标轴。机器人在全局参考坐标系下的位姿可以用点  $P(x, y)$  和角度  $\theta$  表示，因此机器人的位姿可以用 3 个变量  $x, y, \theta$  完全确定，方向向量  $\alpha$  和旋转矩阵  $R(\theta)$  用来表示机器人在全局参考坐标系下的位姿和方位。

$$\alpha_{Global} = \begin{bmatrix} x \\ y \\ \theta \end{bmatrix} \quad (2-2)$$



$$R(\theta) = \begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2-3)$$

其中

$$\dot{\alpha}_{Robot} = R(\theta) \dot{\alpha}_{Global} \quad (2-4)$$

本文采用的 Create2 实验平台由两轮差速驱动，且两轮中心距为  $L$ 。因此移动机器人的运动由其线速度 ( $v$ ) 和角速度 ( $\omega$ ) 来控制。

$$\dot{\alpha}_{Global} = \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{bmatrix} = f(v, \omega) \quad (2-5)$$

$$\dot{\alpha}_{Global} = R(\theta)^{-1} \begin{bmatrix} v \\ 0 \\ \omega \end{bmatrix} \quad (2-6)$$

因此，移动机器人的运动可以由如下的运动学模型表示为

$$\dot{\alpha}_{Global} = \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} v \sin \theta \\ v \cos \theta \\ \omega \end{bmatrix} \quad (2-7)$$

其中  $v$  是沿机器人坐标系  $Y_{Robot}$  轴的线速度， $\omega$  是机器人坐标系相对全局参考坐标系的角速度。定义系统状态变量  $X = [x, y, \theta, v, \omega]$ ，则移动机器人的运动学模型可以表示如下：

$$\dot{X} = \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \\ \dot{v} \\ \dot{\omega} \end{bmatrix} = \begin{bmatrix} v \sin \theta \\ v \cos \theta \\ \omega \\ 0 \\ 0 \end{bmatrix} \quad (2-8)$$

为了将式 (2-8) 应用于扩展卡尔曼滤波器, 需要对式 (2-8) 进行离散化处理。因此, 以采样周期  $\Delta t$  对系统进行离散化, 则扩展卡尔曼滤波器的系统模型如下:

$$X_k = \begin{bmatrix} x_k \\ y_k \\ \theta_k \\ v_k \\ \omega_k \end{bmatrix} = \begin{bmatrix} x_{k-1} + v_{k-1} \Delta t \sin \theta_{k-1} \\ y_{k-1} + v_{k-1} \Delta t \cos \theta_{k-1} \\ \theta_{k-1} + \omega_{k-1} \Delta t \\ v_{k-1} \\ \omega_{k-1} \end{bmatrix} \quad (2-9)$$

### 2.3.2 传感器观测模型的建立

观测模型描述了传感器实际测量到的信号与系统状态变量间的关系。里程计测量左右轮的位移增量, 光流计测量其自身相对全局参考坐标系的位移增量, IMU 测量沿 Z 轴的旋转角速度。

#### (1) 里程计测量模型

差速驱动的移动机器人左右两轮均配备有编码器, 编码器通过累加一个采样周期内的脉冲数来计算车轮的平动速度。

$$V_R = \Delta C_R / \Delta t = \frac{2v + \omega L}{2} \quad (2-10)$$

$$V_L = \Delta C_L / \Delta t = \frac{2v - \omega L}{2} \quad (2-11)$$

式中  $V_L$ 、 $V_R$  分别为左右轮的平动速度,  $\Delta t$  为采样周期,  $\Delta C$  为车轮在一个采样周期内的脉冲增量,  $v$ 、 $\omega$  分别为机器人中心的线速度和角速度。

因此, 里程计的观测模型为

$$h_{encoder} = \begin{bmatrix} V_R \\ V_L \end{bmatrix} = \begin{bmatrix} v + \omega L / 2 \\ v - \omega L / 2 \end{bmatrix} \quad (2-12)$$

对应的观测模型雅克比矩阵为



$$H_{encoder} = \begin{bmatrix} 0 & 0 & 0 & 1 & L/2 \\ 0 & 0 & 0 & 1 & -L/2 \end{bmatrix} \quad (2-13)$$

即

$$Z_k = H_{encoder} X_k + v_k \quad (2-14)$$

## (2) 光流计测量模型

差速驱动的移动机器人上光流计安装位置一般都远离机器人中心位置，光流计测量一个采样周期内沿其自身坐标轴的位移增量。

$$V_x = \Delta x / \Delta t = \omega d \quad (2-15)$$

$$V_y = \Delta y / \Delta t = v \quad (2-16)$$

式中  $\Delta x$ 、 $\Delta y$  分别为沿其坐标系 X、Y 轴的位移增量， $v$ 、 $\omega$  分别为机器人中心的线速度和角速度， $d$  为光流计距机器人中心的距离，光流计在 Create2 实验平台上的安装位置如图 2-8 所示。

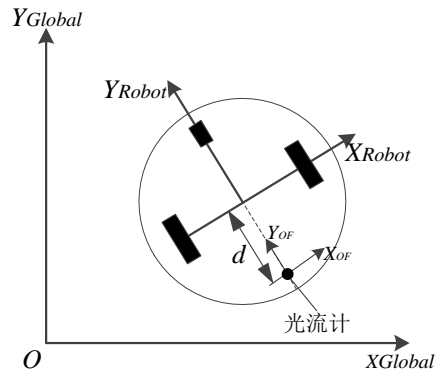


图 2-8 光流计在 Create2 实验平台上的安装位置

因此，光流计的观测模型为

$$h_{optical\_flow} = \begin{bmatrix} V_x \\ V_y \end{bmatrix} = \begin{bmatrix} \omega d \\ v \end{bmatrix} \quad (2-17)$$

对应的观测模型雅克比矩阵为

$$H_{optical\_flow} = \begin{bmatrix} 0 & 0 & 0 & 0 & d \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad (2-18)$$

即

$$Z_k = H_{optical\_flow} X_k + v_k \quad (2-19)$$

## (3) IMU 测量模型

对于室内移动机器人，陀螺仪只需测量移动机器人 Z 轴的角速度即可。

因此，IMU 的角速度观测模型为

$$h_{IMU} = [\omega_{IMU}] = \omega \quad (2-20)$$

对应的观测模型雅克比矩阵为

$$H_{IMU} = [0 \ 0 \ 0 \ 0 \ 1] \quad (2-21)$$

即

$$Z_k = H_{IMU} X_k + v_k \quad (2-22)$$

通过对角速度积分，得到机器人的方位角  $\theta$ 。

因此，IMU 的角度观测模型为

$$h_{heading} = \int \omega_{IMU} dt = \theta \quad (2-23)$$

对应的观测模型雅克比矩阵为

$$H_{heading} = [0 \ 0 \ 1 \ 0 \ 0] \quad (2-24)$$

即

$$Z_k = H_{heading} X_k + v_k \quad (2-25)$$

## 2.4 扩展卡尔曼滤波器设计

扩展卡尔曼滤波器是使用系统动态模型和一系列随时间变化的观测模型，以及包含统计噪声和其他不确定性，用来产生对未知变量估计的算法。通过使用贝叶斯推导和每一个时间帧上进行联合概率密度估计，通常 EKF 比只依靠单个测量数据具有更高的精度和稳定性。因此，选取 EKF 对多种定位传感器数据进行融合，输出更加准确的位姿信息。没有直接使用 KF 而用 EKF，是由于移动机器人的运动学模型存在非线性问题。应用 EKF 需要提供系统模型和测量模型，系统模型和测量模型分别表示如下：

$$X_k = f(X_{k-1}) + w_{k-1} \quad (2-26)$$

$$Z_k = h(X_k) + v_k \quad (2-27)$$

其中， $f$  和  $h$  分别表示非线性系统模型和测量模型，动态系统噪声  $w_{k-1}$  和测量噪声  $v_k$  均为零均值高斯噪声，且其协方差矩阵为： $w_k \sim N(0, Q_k)$ ， $v_k \sim N(0, R_k)$ 。

首先需要对非线性函数  $f$  和  $h$  进行线性化处理，因此求得函数  $f$  和  $h$  在  $X_k$  处的雅克比矩阵，求雅克比矩阵的公式如下：

$$F_{k-1} = \frac{\partial f(X)}{\partial X} \Big|_{\hat{X}_{k-1}} \quad (2-28)$$

$$F_k = \begin{bmatrix} 1 & 0 & V_k \Delta t \cos \theta_{k-1} & \Delta t \sin \theta_{k-1} & 0 \\ 0 & 1 & -V_k \Delta t \sin \theta_{k-1} & \Delta t \cos \theta_{k-1} & 0 \\ 0 & 0 & 1 & 0 & \Delta t \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (2-29)$$

预测过程：

$$\hat{X}_{k-} = f(\hat{X}_{k-1}^+, u_k) \quad (2-30)$$

$$P_k^- = F_k P_{k-1} F_k^T + Q_k \quad (2-31)$$

更新过程：

$$K_k = P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1} \quad (2-32)$$

$$\hat{X}_k^+ = \hat{X}_{k-} + K_k (z_k - h(\hat{X}_{k-})) \quad (2-33)$$

$$P_k^+ = P_k^- - K_k H_k P_k^- \quad (2-34)$$

上式中  $F_k$  为上一状态向量  $X_{k-1}$  对应的状态转移矩阵， $u_k$  为控制输入向量， $w_k$  是假定服从零均值协方差为  $Q$  的多元正太分布中的过程噪声， $H_k$  是将实际状态空间映射到观测空间的观测状态矩阵， $v_k$  是假定服从零均值协方差为  $R$  的高斯白噪声的观测噪声。

## 2.5 三种定位方案实验对比

里程计和光流计均为增量式测距传感器，且普遍应用航位推算定位算法进行定位。航位推算定位算法特别适用于短时短距离定位且精度很高。单一传感器一般无法独立完成定位功能，需要融合多种传感器来进行定位，多种传感器融合能够提高定位精度。本文对比了三种融合方案的定位结果：①里程计+IMU 融合；②光流计+IMU 融合；③里程计+光流计+IMU 进行 EKF 融合。

### 2.5.1 融合模型的建立

#### (1) 里程计+IMU 融合模型

基于里程计的航位推算定位算法：

$$X(k) = X(0) + \sum_{i=0}^{k-1} D(i) \cos(\theta_k) \text{Encoder}_{\text{scale}} \quad (2-35)$$

$$Y(k) = Y(0) + \sum_{i=0}^{k-1} D(i) \sin(\theta_k) \text{Encoder}_{\text{scale}} \quad (2-36)$$

$$\theta(k) = \theta(0) + \sum_{i=0}^{k-1} \Delta\theta(i) \quad (2-37)$$

式中  $X(k)$ 、 $Y(k)$  为机器人在  $k$  时刻的位置坐标， $\theta(k)$  为机器人在  $k$  时刻的角度值， $\Delta\theta(i)$  为一个采样周期内的角度增量， $D(i)$  为一个采样周期内机器人中心的位移增量（由里程计测得）， $\text{Encoder}_{\text{scale}}$  为里程计的尺度因子。

## (2) 光流计+IMU 融合模型

基于光流计的航位推算定位算法：

$$X_{OF}(k) = X_{OF}(0) + \sum_{i=0}^{k-1} (\Delta x \cos(\theta_k) - \Delta y \sin(\theta_k)) \text{OF}_{\text{scale}} \quad (2-38)$$

$$Y_{OF}(k) = Y_{OF}(0) + \sum_{i=0}^{k-1} (\Delta x \sin(\theta_k) + \Delta y \cos(\theta_k)) \text{OF}_{\text{scale}} \quad (2-39)$$

$$\theta(k) = \theta(0) + \sum_{i=0}^{k-1} \Delta\theta(i) \quad (2-40)$$

$$X_{\text{Robot}}(k) = X_{OF}(k) - d \sin(\theta_k) \quad (2-41)$$

$$Y_{\text{Robot}}(k) = Y_{OF}(k) + d \cos(\theta_k) \quad (2-42)$$

式中  $X_{OF}(k)$ 、 $Y_{OF}(k)$  为光流计中心在  $k$  时刻的位置坐标， $X_{\text{Robot}}(k)$ 、 $Y_{\text{Robot}}(k)$  为机器人中心在  $k$  时刻的位置坐标， $\theta(k)$  为机器人在  $k$  时刻的角度值， $\Delta x$ 、 $\Delta y$  分别为光流计在一个采样周期内沿自身坐标系  $X$ 、 $Y$  轴的位移增量， $\text{OF}_{\text{scale}}$  为光流计的尺度因子。

## 2.5.2 实验结果

本次实验在实验室地面上布置 12 个标记点，每两个点之间的距离均为 1m，如图 2-9 所示。通过手动控制扫地机器人沿如下路径运动，然后分析三种融合方案的定位结果与理论标记点的误差大小。路径一：沿 A-B-C-D-E-F-G-H-I-J-K-L 运动 3 圈（长工字型）；路径二：沿 A-L-K-B-C-J-I-D-E-H-G-F 运动 2 圈（短工字型）。

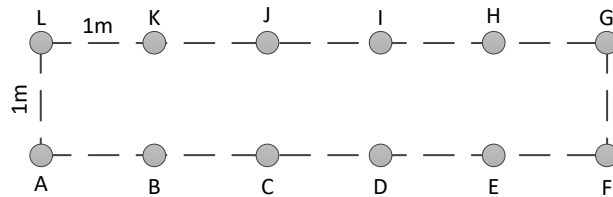


图 2-9 实验测试点布置

路径一实验结果：三种融合方案轨迹图如图 2-10 所示，12 个标记点的坐标记录及误差分析如表 2-3、2-4 所示。

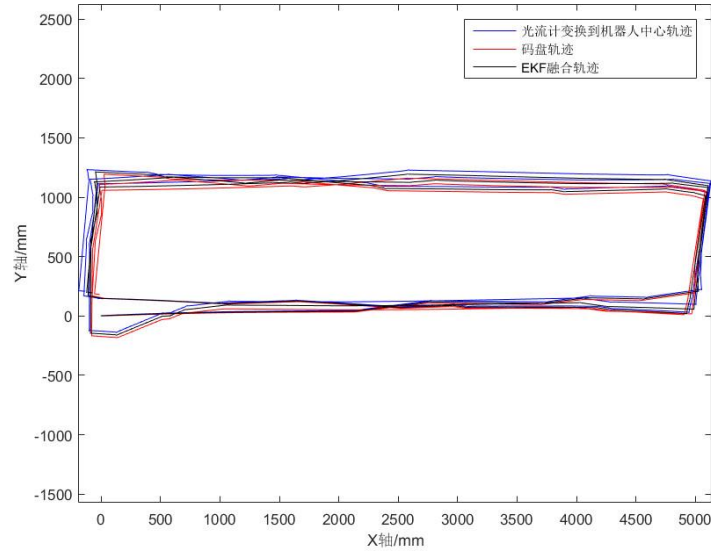


图 2-10 三种融合方案轨迹图（路径一）

表 2-3 12 个标记点坐标记录（光流计+IMU 融合）

单位：mm

次数	标记点 A (0,0)	标记点 B (1000,0)	标记点 C (2000,0)	标记点 D (3000,0)	标记点 E (4000,0)	标记点 F (5000,0)
1	(0,3)	(984,34)	(2062,46)	(3052,123)	(4045,158)	(5052,224)
2	(-101,-122)	(937,108)	(1947,119)	(2953,130)	(3985,150)	(5007,97)
3	(-147,171)	(966,114)	(1901,117)	(2855,92)	(3924,84)	(4946,31)
均值	(-62,13)	(929,85)	(1970,94)	(2953,115)	(3984,130)	(5002,118)
误差	63	111	98	124	132	118

标记点 G (5000,1000)	标记点 H (4000,1000)	标记点 I (3000,1000)	标记点 J (2000,1000)	标记点 K (1000,1000)	标记点 L (0,1000)
(5134,1137)	(4100,1195)	(3100,1218)	(2048,1174)	(1016,1137)	(-25,1113)
(5123,1116)	(4067,1145)	(3049,1165)	(2015,1159)	(988,1161)	(-97,1151)
(5107,1036)	(4017,1072)	(2997,1089)	(2008,1138)	(943,1183)	(-70,1230)
(5121,1096)	(4061,1137)	(3049,1158)	(2024,1157)	(982,1161)	(-63.6,1165)
155	150	165	159	162	177

表 2-4 12 个标记点坐标记录（里程计+IMU 融合）

单位：mm

次数	标记点 A (0,0)	标记点 B (1000,0)	标记点 C (2000,0)	标记点 D (3000,0)	标记点 E (4000,0)	标记点 F (5000,0)
1	(0,0)	(980,26)	(2060,33)	(3024,103)	(4002,132)	(4996,193)
2	(-69,-168)	(910,46)	(1903,53)	(2920,57)	(3962,72)	(4970,17)
3	(-63,167)	(927,107)	(1926,105)	(2867,76)	(3910,63)	(1904,10)
均值	(-33,0)	(939,60)	(1963,63)	(2937,79)	(3958,89)	(4957,73)
误差	33	85	73	101	98	85

标记点 G (5000,1000)	标记点 H (4000,1000)	标记点 I (3000,1000)	标记点 J (2000,1000)	标记点 K (1000,1000)	标记点 L (0,1000)
(5073,1058)	(4057,1122)	(3065,1148)	(2037,1111)	(1044,1078)	(9,1058)
(5081,1049)	(4039,1083)	(3024,1105)	(2035,1104)	(1042,1112)	(-15,1107)
(5054,990)	(3998,1027)	(3019,1047)	(2067,1097)	(1027,1145)	(46,1192)
(5069,1033)	(4031,1078)	(3036,1100)	(2046,1104)	(1038,1112)	(13,1119)
76	84	106	114	118	120

路径二实验结果：三种融合方案轨迹图如图 2-11 所示，12 个标记点的坐标记录及误差分析如表 2-5、2-6 所示。

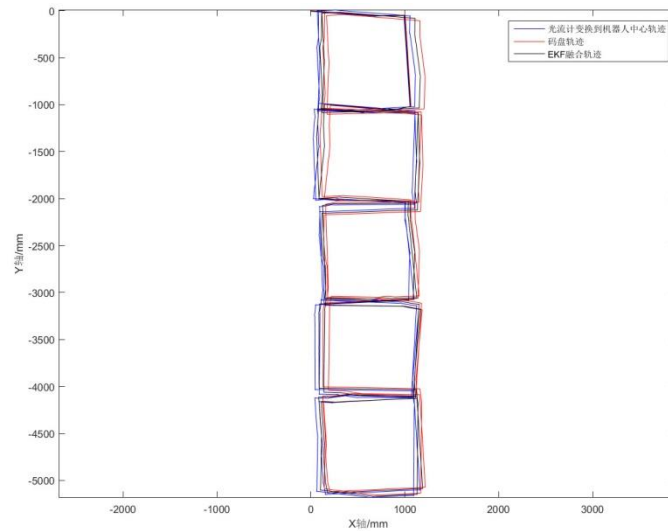


图 2-11 三种融合方案轨迹图（路径二）

表 2-5 12 个标记点坐标记录（光流计+IMU 融合）						单位：mm
次数	标记点 A (0,0)	标记点 L (1000,0)	标记点 K (1000,-1000)	标记点 B (0,-1000)	标记点 C (0,-2000)	标记点 J (1000,-2000)
1	(-1,-7)	(992,-53)	(1055,-1020)	(37,-1046)	(53,-2016)	(1007,-2046)
2	(67,-3)	(1054,-56)	(1108,-1068)	(72,-1067)	(96,-2140)	(1102,-2103)
3	NAN	NAN	(1104,-1021)	(84,-987)	(28,-2001)	(1002,-2029)
4	(83,11)	NAN	(1109,-1094)	(74,-1063)	(93,-2088)	(1093,-2061)
均值	(50,0)	(1023,-54)	(1094,-1051)	(67,-1041)	(67,-2061)	(1051,-2060)
误差	50	59	107	78	91	79

标记点 I (1000,-3000)	标记点 D (0,-3000)	标记点 E (0,-4000)	标记点 H (1000,-4000)	标记点 G (1000,-5000)	标记点 F (0,-5000)
(1099,-3073)	(111,-3079)	(89,-4082)	(1099,-4123)	(1161,-5100)	(155,-5148)
(1148,-3179)	(140,-3135)	(87,-4163)	(1101,-4127)	NAN	NAN
(1039,-3082)	(50,-3132)	(45,-4034)	(1081,-4047)	(1094,-5159)	(62,-5115)
(1124,-3142)	(115,-3081)	(42,-4125)	(1073,-4120)	NAN	NAN
(1103,-3119)	(104,-3107)	(66,-4101)	(1088,-4104)	(1128,-5130)	(109,-5132)
157	149	121	137	182	171

表 2-6 12 个标记点坐标记录（里程计+IMU 融合）						单位：mm
次数	标记点 A (0,0)	标记点 L (1000,0)	标记点 K (1000,-1000)	标记点 B (0,-1000)	标记点 C (0,-2000)	标记点 J (1000,-2000)
1	(0,0)	(1003,-48)	(1066,-1015)	(110,-1050)	(120,-1979)	(1065,-2013)
2	(166,-52)	(1160,-107)	(1178,-1112)	(171,-1103)	(142,-2175)	(1167,-2138)
3	NAN	NAN	(1205,-1045)	(210,-1007)	(143,-1994)	(1108,-2027)
4	(147,2)	NAN	(1176,-1077)	(140,-1038)	(157,-2061)	(1153,-2036)
均值	(105,-17)	(1081,-78)	(1156,-1062)	(158,-1049)	(141,-2052)	(1123,-2054)
误差	106	112	168	165	150	134

标记点 I (1000,-3000)	标记点 D (0,-3000)	标记点 E (0,-4000)	标记点 H (1000,-4000)	标记点 G (1000,-5000)	标记点 F (0,-5000)
(1155,-3035)	(216,-3041)	(187,-4038)	(1156,-4086)	(1218,-5071)	(203,-5120)

(1185,-3182)	(180,-3133)	(133,-4160)	(1132,-4126)	NAN	NAN
(1141,-3062)	(141,-3112)	(129,-4003)	(1158,-4024)	(1170,-5133)	(146,-5090)
(1178,-3115)	(178,-3048)	(124,-4103)	(1116,-4102)	NAN	NAN
(1165,-3098)	(179,-3083)	(143,-4076)	(1141,-4085)	(1194,-5102)	(174,-5105)
192	197	162	164	219	203

注：表格中的 NAN 表示没有数据，因为扫地机器人没有经过该测试点。

由上述原始数据绘制误差随距离变化曲线，路径一误差变化曲线如图 2-12 所示，路径二误差变化曲线如图 2-13 所示。

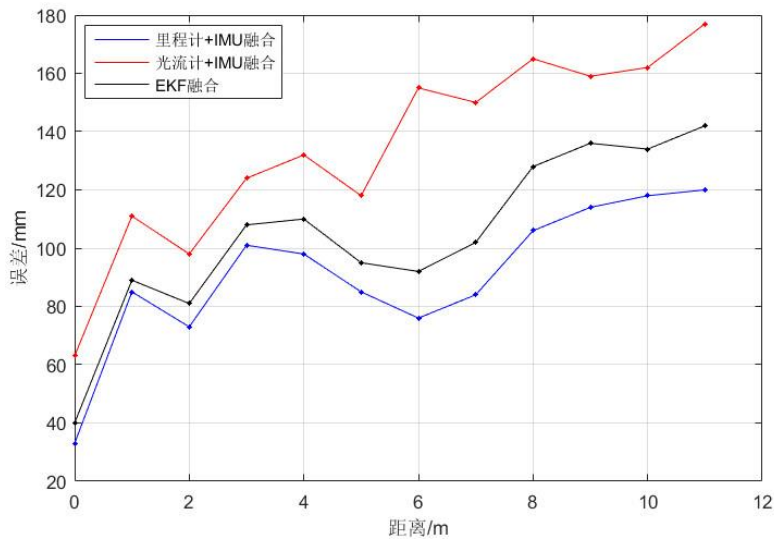


图 2-12 三种融合方案误差变化曲线图（路径一）

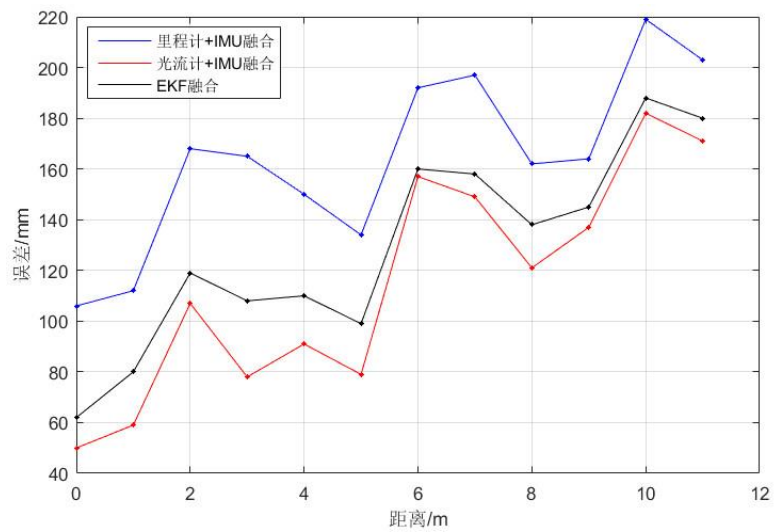


图 2-13 三种融合方案误差变化曲线图（路径二）



通过分析上述曲线可知,光流计+IMU 融合和里程计+IMU 融合的误差均是随着距离的增加而累积,且同一标记点的偏差也随着时间的增加而累积。在实验路径一中里程计+IMU 融合的累积误差增长率要低于光流计+IMU 融合;相反在实验路径二中光流计+IMU 融合的累积误差增长率要低于里程计+IMU 融合。因此,在里程计不打滑的情况下光流计和里程计的相对定位精度相当,且光流计的相对定位精度符合 2.2.2 小节中的指标(最大 4.52%)。由图 2-10 和图 2-11 可知,光流计和里程计均存在尺度不一致的问题,因此里程计+光流计+IMU 进行 EKF 融合可以适当地减弱尺度变化的影响,提高航位推算定位的精度和鲁棒性,但终究由于上述传感器均为相对定位传感器无法彻底解决累积误差的问题。

## 2.6 本章小结

本章首先深入研究了光流计的基本测量原理,并在 6 种不同材质表面上验证了光流计的定位效果,证实了 ADNS-3080 光流计可用于室内移动机器人的位姿估计,并基于实验结论给出了光流计在移动机器人上的合理安装要求。然后基于扫地机器人的实际结构建立了移动机器人的运动学模型和系统模型,其次根据三种传感器(里程计、光流计和 IMU)的不同测量原理,分别建立了测量模型。再次根据上述的系统模型和观测模型,设计了适用于 Create2 实验平台的扩展卡尔曼滤波器。最后通过实验,验证了基于三种传感器(里程计、光流计和 IMU)的 EKF 融合较里程计+IMU 融合(或光流计+IMU 融合)能够在一定程度上克服尺度不一致的问题,提高室内移动机器人的航位推算定位的精度和鲁棒性。

## 第 3 章 视觉 SLAM 算法优化研究

### 3.1 引言

目前比较优秀的开源 VSLAM 系统有 MSCKF<sup>[33]</sup> (基于特征点的滤波方案)、OKVIS<sup>[39]</sup> (基于特征点的优化方案)、ROVIO<sup>[42]</sup> (基于特征点的滤波方案)、SVO<sup>[56]</sup> (基于直接法的优化方案)、ORB-SLAM2<sup>[16]</sup> (基于特征点的优化方案) 等, ORB-SLAM2 作为目前最优秀的 SLAM 框架之一, 大多数公司基于 ORB-SLAM2 框架开发自己的 VSLAM 系统, 用在实际的产品上进行落地。开源的算法只适合做 DEMO, 跑跑数据集还可以, 但商业化过程中会遇到很多的实际问题。比如针对移动机器人的应用场景, 6-DOF SLAM 不仅需要的优化时间长, 而且额外未约束的自由度导致优化结果出现跳动现象 (误差大); 另一方面 ORB-SLAM2 中加载词典文件耗时长、且占用大量的系统内存。

本章首先深入研究 ORB-SLAM2 的系统架构, 并基于 ORB-SLAM2 开发本课题的 VSLAM 算法。然后针对商业化过程中可能遇到的实际问题, 提出相应的优化方案: 1) 对 ORB-SLAM2 中的 DBoW2 词汇库进行优化, 并针对扫地机器人实际的应用场景, 训练家庭场景下的词汇库; 2) 针对扫地机器人在平面内的运动特性, 研究平面内的 3-DOF SLAM 系统。

### 3.2 ORB-SLAM2 算法研究

ORB-SLAM2 是一个基于 ORB 特征点可实时用于单目、双目和 RGB-D 相机的完整 SLAM 系统, 可应用于室内外的大小场景。ORB-SLAM2 包含 3 个线程: 1) 跟踪线程用于寻找与本地地图特征匹配的重投影误差最小的图像帧; 2) 建图线程用于管理本地地图和执行局部 BA 优化; 3) 闭环检测线程用来检测大回环, 通过位姿图修正累积误差, 该线程在位姿图优化后启动第四个线程来执行全局 BA 优化, 以计算最佳的运动估计。该系统还嵌入了位置识别模块 DBoW2, 可用于初始化、跟踪丢失和重定位场合。DBoW2 更大的作用在于进行图像帧间匹配时, 通过使用词汇库对描述子进行分类, 然后按相应的类别进行比较, 去替换高维的特征描述子间的比对, 可以大大提高系统帧间匹配的实时性。ORB-SLAM2 的系统架构如图 3-1 所示, ORB-SLAM2 的三个线程如图 3-2 所示。

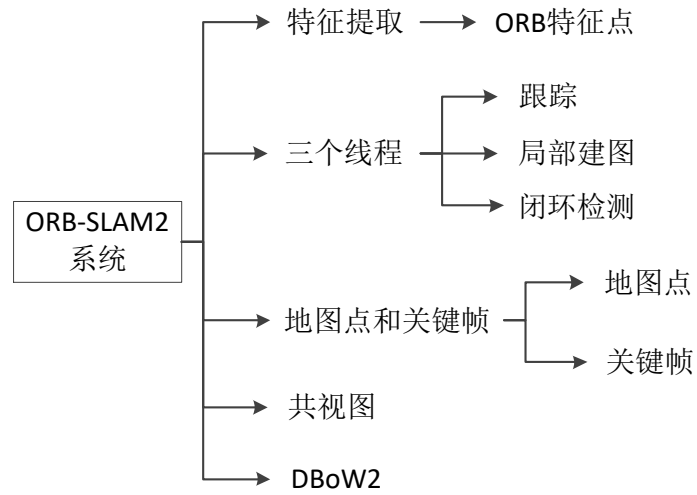


图 3-1 ORB-SLAM2 系统架构

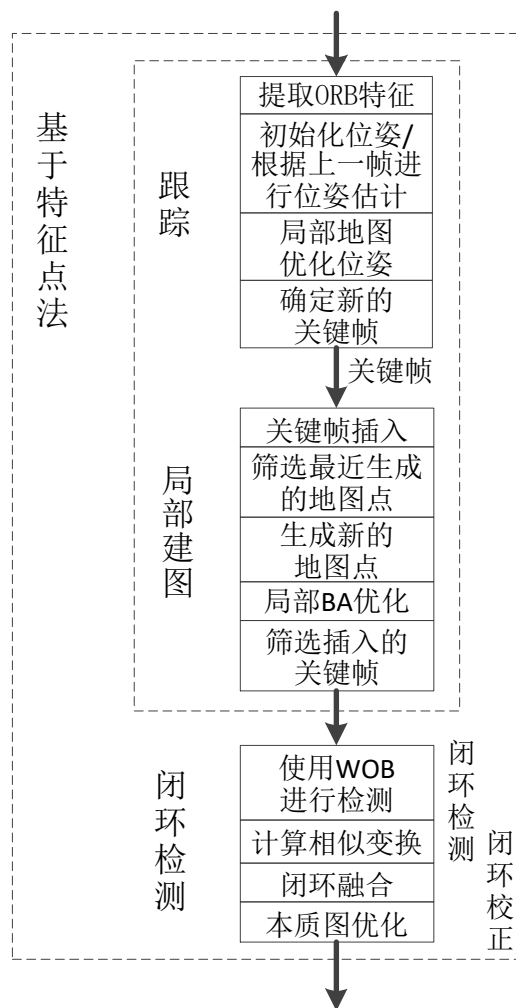


图 3-2 ORB-SLAM2 的三个线程

ORB-SLAM2 系统由 3 个线程构成：

(1) 跟踪线程

跟踪线程主要是从当前帧中提取 ORB 特征点，并与上一帧利用 PnP 进行帧间匹配，通过最小化重投影误差得到帧间运动估计。然后根据当前状态（帧间运动是否大于阈值、是否长时间未加入关键帧、局部地图是否空闲）判断是否将此帧作为关键帧加入到地图中。

(2) 局部建图线程

局部建图线程主要是插入新的关键帧，并对新关键帧中的地图点进行筛选（Culling），剔除新加入的坏的地图点，将该帧中与已有关键帧中共视程度高的地图点进行三角化生成新的地图点，之后进行局部 BA 优化。最后再对关键帧进行筛选，删除掉有过多共视关键点的旧关键帧。

(3) 回环检测线程

回环检测线程主要是判断当前关键帧是否与已有关键帧构成回环。首先判断当前关键帧与上一次回环检测是否大于阈值，若满足条件，则计算当前关键帧与相邻关键帧的相关度得分（BOW），若得分过高则判断为候选帧。紧接着找出与候选帧有公共词汇的非相邻关键帧，求解 Sim3 并进行优化，确定候选帧的位置，并对已有的关键帧和地图点进行融合和更新。

### 3.3 DBoW2 词汇库优化研究

#### 3.3.1 DBoW2 词汇库优化

DBoW2 是由 Lopez<sup>[17]</sup>等人开发的用于相似度匹配的词汇库，由于其在 ORB-SLAM2 回环检测中的惊人效果，得到了广大 SLAM 研究者的追捧。

BoW（Bag of Words，词袋模型）是一种使用视觉词汇将图像转换成低维向量的技术，使用这种技术能够管理大量的图像集。DBoW2 采用树状层级结构将处于相邻位置的图像特征聚集在一起，进而创建词汇库。DBoW2 词汇库生成过程如图 3-3 所示。

这棵树的节点数为

$$1 + K + \dots + K^L = (K^{L+1} - 1) / (K - 1) \quad (3-1)$$

总共有  $K^L$  个叶节点，每个叶节点代表一类，用每一类中所有特征的平均特征代表一个词汇（word）。词汇库中的每一个词汇都被赋予一个权重，有 TF-IDF、BINARY、IDF 和 TF 等权重可供选择，DBoW2 选用 TF-IDF 权重。

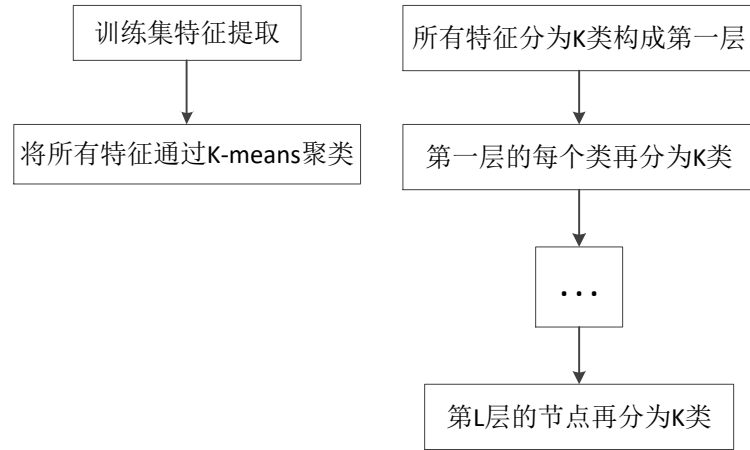


图 3-3 DBoW2 词汇库生成过程

TF 代表词频 (Term Frequency)，表示词汇  $T$  在文档  $D$  中出现的次数，IDF 代表逆文档频率 (Inverse Document Frequency)，表示词汇是否在所有文档中都是常见的或稀少的，比如如果包含词汇  $T$  的文档越少，则 IDF 值越大。

树结构中的每个单词 TF 和 IDF 分别定义如下：

$$TF_k(t, d) = \frac{f_{t,d}}{\max\{f_{t',d} : t' \in d\}} \quad (3-2)$$

$$IDF_k(t, D) = \log \frac{N}{|\{d \in D : t \in d\}|} \quad (3-3)$$

$$TFIDF_k(t, d, D) = TF_k(t, d) * IDF_k(t, D) \quad (3-4)$$

式中  $f_{t,d}$  表示处于叶节点的特征数量， $\max\{f_{t',d} : t' \in d\}$  表示所有的特征数量， $N$  为所有的图像数量， $\{d \in D : t \in d\}$  表示处于叶节点的图像数量。

词汇库可以通过训练大量的图片集得到，词汇的 TF 越高，说明词汇在该幅图片中出现的频率越高；词汇的 IDF 越高，说明该词汇本身具有较好的区分性。将二者结合起来，即是图片的 BoW 描述。

得到了词汇库后，即可计算任意两帧图像间的相似度得分，分两步进行：

第一步：将图像用特征向量  $v$  进行表示，图像将通过特征向量  $v$  在词汇库中搜索其相邻的叶节点，所有叶节点上的权重和构成了该图像的 BoW 向量  $V$ 。

第二步：根据两帧图像的 BoW 向量，计算其相似度得分

$$s(V_1, V_2) = 1 - \frac{1}{2} \left| \frac{V_1}{|V_1|} - \frac{V_2}{|V_2|} \right| \quad (3-5)$$

由式(3-5)可知,两张图像的相似度越高得分就越高,图像与其自身图像的相似度得分为1,两张完全不相关的图像相似度得分为0。

### 3.3.2 DBoW2 优化实验

目前 ORB-SLAM2 的词汇库是基于 DBoW2 库,用于在**帧间匹配和回环检测**过程中**对暴力匹配算法进行加速**。即将比较两个图像帧的相似度转化为比较两个向量的相似度,所以在 SLAM 的初始化阶段需要将整个词汇库读入到内存中。原 ORB-SLAM2 词汇库格式为 TXT 文档格式,在将其读入到内存中将耗费大量的时间,且浪费大量宝贵的内存空间,而且在对匹配帧计算相似度得分时耗费更多的时间,严重影响到整个 SLAM 系统的实时性。

故将 ORB-SLAM2 中的词汇库 DBoW2 进行优化,升级为 DBoW3。另一方面,ORB-SLAM2 作者提供的词汇库是公用的词汇库,是对多种场景下的图片集训练得到,这就导致词汇库中包含大量冗余的词汇。

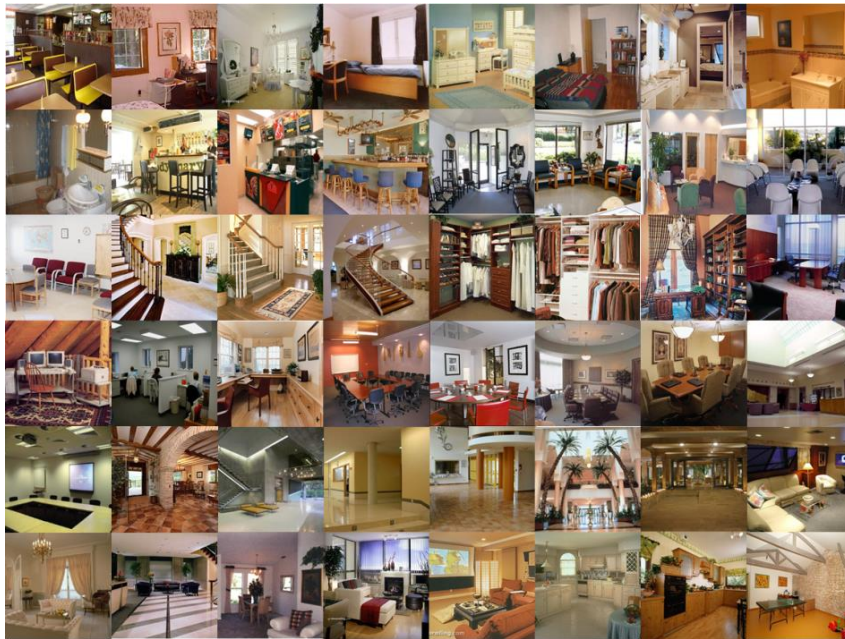


图 3-4 词汇库训练图片集

因此,本文**首先训练针对自己特有场景的词汇库**,由于本文的研究场景为家庭环境,故对大量家庭环境下的图片集进行训练,训练图片集如图 3-4 所示(这里只给出 48 张样例图片,而实际上需要上万张图片),得到针对家庭环境的词汇库。下面在笔记本和嵌入式 R16 平台上分别测试载入词汇库耗时以及载入前后内存占用量对比,实验结果如表 3-1、3-2 所示。

表 3-1 词汇库载入时间对比测试

运行平台	ORB-SLAM2/s	SKB-SLAM/s
Core i7 7700 四核 2.8GHz	7.32	0.45
R16 A7 四核 1GHz	无法测试	1.52

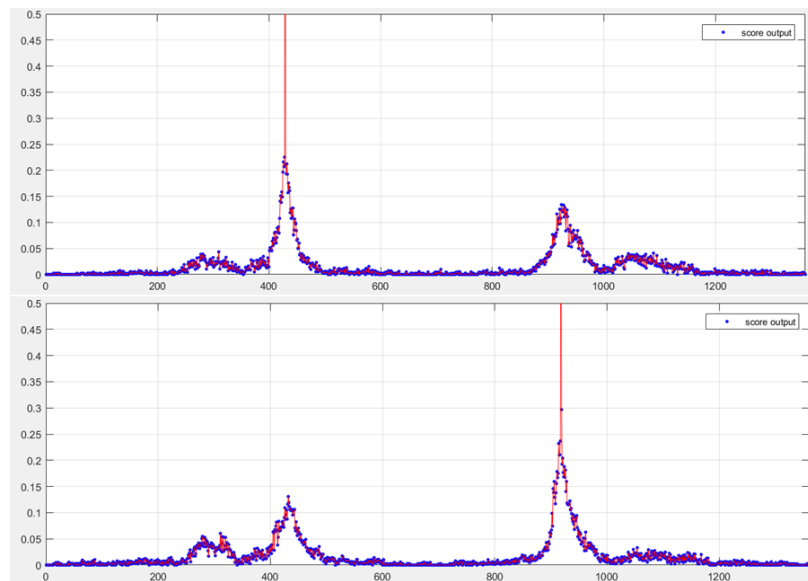
表 3-2 词汇库载入前后内存占用量对比测试

运行平台	ORB-SLAM2/Mb			SKB-SLAM/Mb		
	载入前	载入后	增加量	载入前	载入后	增加量
Core i7 7700 四核 2.8GHz	434.59	859.152	424.562	434.59	627.465	192.875
R16 A7 四核 1GHz	无法测试			137.34	313.999	176.659

注:ORB-SLAM2 加载原版词汇库文件,SKB-SLAM 加载针对家庭环境训练的词汇库文件。SKB-SLAM 为本文的 SLAM 系统。

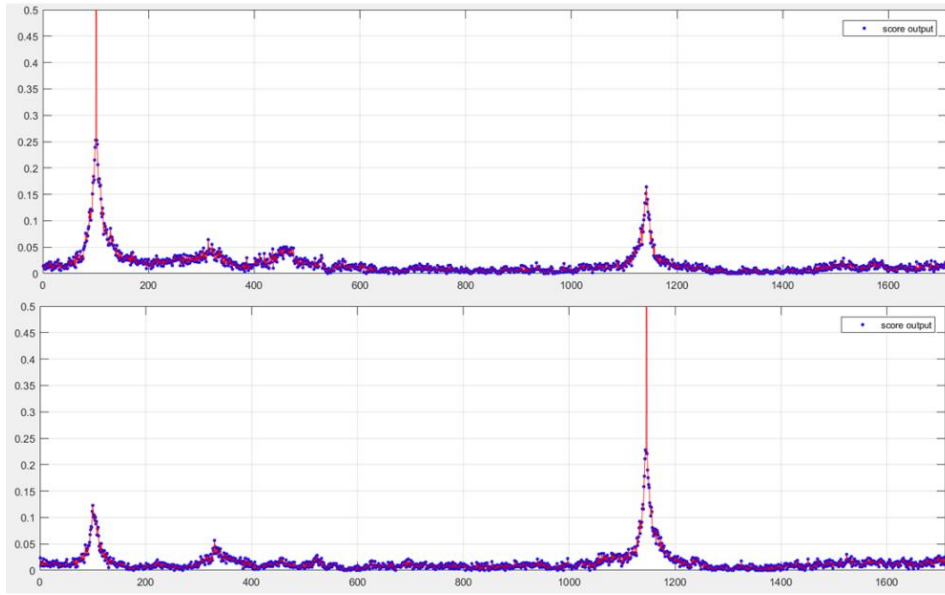
从测试结果可以看出,在笔记本上 SKB-SLAM 加载词汇库的时间相比 ORB-SLAM2 减少了 16.267 倍。在笔记本上 SKB-SLAM 加载词汇库的内存占用量相比 ORB-SLAM2 减小了 231.687Mb。由于 ORB-SLAM2 无法在嵌入式平台上运行,故无法测试。因此本文针对家庭环境训练的词汇库,有效地减小了 SLAM 系统初始化阶段加载词汇库的时间和节省了 SLAM 系统的内存空间,优化结果达到了本文的预期目标。

另一方面,针对家庭环境训练得到的词汇库进行回环检测测试,验证生成的词汇库是否能用于帧间匹配以及回环检测。下面将对两组家庭环境下的数据集,计算每一帧与所有帧的相似度得分,相似度得分实验结果如图 3-5 所示。



a) 家庭环境数据集一





b) 家庭环境数据集二

图 3-5 训练的词汇库用于回环检测测试

从测试结果可以看出，图(a)中数据集在第 430 帧和 920 帧处出现了回环，此时的相似度得分为 0.17；图(b)中数据集在 100 帧和 1140 帧处出现了回环，此时的相似度得分为 0.14，而在非回环的情况下相似度得分都远小于此值，形成了鲜明的对比。而图(a)和图(b)出现回环时的实际图像序列也处于相似的场景下。因此证实了本文针对家庭场景训练的词汇库是可以用于帧间匹配和回环检测。

### 3.4 基于平面 3-DOF VSLAM 算法优化研究

#### 3.4.1 平面 3-DOF VSLAM 算法优化

目前开源的主流 VSLAM 系统（MSCKF、ROVIO、SVO 和 ORB-SLAM2 等）都是针对三维空间 6-DOF 下的优化，没有专门针对平面移动机器人 3-DOF 下的优化。由于是三维空间内的 6-DOF，所以在 BA 优化时：一方面由于未约束的自由度导致优化结果出现跳动现象（误差大），另一方面由于局部 BA 和全局 BA 优化都需要频繁调用从空间点到像素点的变换矩阵  $T_{wc}$ ，SLAM 的实时性会降低。

因此本文基于 ORB-SLAM2 框架对平面内 3-DOF SLAM 系统进行研究，图优化示意图如图 3-6 所示。



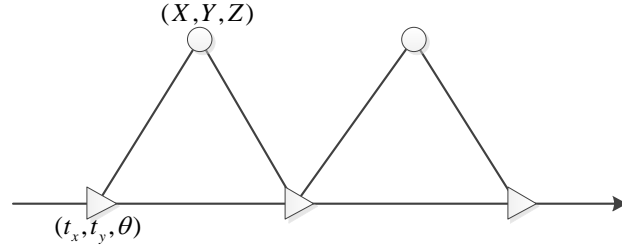


图 3-6 图优化示意图

图中圆圈表示 XYZ 上的顶点  $(X, Y, Z)$ ，三角形表示 SE2 上的顶点  $(t_x, t_y, \theta)$ ，连接圆圈和三角形的边为优化误差项。 $(X, Y, Z)$  表示空间 3D 点坐标， $(t_x, t_y, \theta)$  表示平面内移动机器人位姿。

在高斯分布的假设下，给出相机的观测方程：

$$z_{k,j} = h(x_k, y_j) + v_{k,j} \quad (3-6)$$

式 (3-6) 假设了噪声项  $v_k \sim N(0, Q_{k,j})$ ，所以观测数据的条件概率为

$$P(z_{j,k} | x_k, y_j) = N(h(x_k, y_j), Q_{k,j}) \quad (3-7)$$

$P(z_{j,k} | x_k, y_j)$  仍然是一个高斯分布，为了求得使它最大化的  $x_k, y_j$ ，通过最小化负对数来求  $P(z_{j,k} | x_k, y_j)$  的最大似然估计。

对一般的高维高斯分布  $x \sim N(u, \Sigma)$ ，则其概率密度函数展开形式为

$$P(x) = \frac{1}{\sqrt{(2\pi)^N \det(\Sigma)}} \exp\left(-\frac{1}{2}(x-u)^T \Sigma^{-1}(x-u)\right) \quad (3-8)$$

对式 (3-8) 两侧取负对数，则

$$-\ln(P(x)) = \frac{1}{2} \ln((2\pi)^N \det(\Sigma)) + \frac{1}{2}(x-u)^T \Sigma^{-1}(x-u) \quad (3-9)$$

对  $P(z_{j,k} | x_k, y_j)$  求最大化等同于对式 (3-9) 两侧求最小化。在最小化式 (3-9) 的  $x$  时，第一项与  $x$  无关，故可以省略。于是，只需要最小化式 (3-9) 右侧的二次型，即可得到观测方程的最大似然估计。于是由式 (3-6) 得到最小化形式为

$$x^* = \arg \min((z_{k,j} - h(x_k, y_j))^T Q_{k,j}^{-1}(z_{k,j} - h(x_k, y_j))) \quad (3-10)$$

因此，对于相机运动，定义观测数据与估计值之间的误差：

$$e_{k,j} = z_{k,j} - h(x_k, y_j) \quad (3-11)$$

对于平面移动机器人，上述 (3-6) 式的观测方程可以写成

$$z_{k,j} = C(RX_k + t) + v_{k,j} \quad (3-12)$$

$$C = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (3-13)$$

式 (3-12) 中  $C$  为相机的内参矩阵,  $f_x$ 、 $f_y$  分别为相机像素平面  $X$ 、 $Y$  方向的焦距 (理论上  $f_x = f_y$ ),  $c_x$ 、 $c_y$  为光心在像素平面上的坐标 (理论上光心位于像素平面中心)。

$$R = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3-14)$$

式 (3-12) 中  $R$  为相机相对于全局参考坐标系的旋转矩阵,  $\theta$  为对应的航向角。

$$t = \begin{bmatrix} t_x \\ t_y \\ 0 \end{bmatrix} \quad (3-15)$$

式 (3-12) 中  $t$  为相机相对于全局参考坐标系的平移向量。因此仅由  $R$  和  $t$  即可表示相机在全局参考坐标系下的位姿。

将式 (3-13) ~ (3-15) 代入式 (3-12) 中, 得相机观测方程表达式:

$$\begin{bmatrix} ZC_x + f_x(X_k \cos \theta - Y_k \sin \theta + t_x) + v_{k,j} \\ ZC_y + f_y(Y_k \cos \theta + X_k \sin \theta + t_y) + v_{k,j} \\ Z_k + v_{k,j} \end{bmatrix} \quad (3-16)$$

将式 (3-11) 在  $X$ 、 $Y$ 、 $Z$  方向上展开得

$$\begin{cases} e_{x,k,j} = u_{k,j} - [C_x + f_x(X_k \cos \theta - Y_k \sin \theta + t_x) / Z_k] \\ e_{y,k,j} = v_{k,j} - [C_y + f_y(Y_k \cos \theta + X_k \sin \theta + t_y) / Z_k] \\ e_z = 0 \end{cases} \quad (3-17)$$

则误差项  $e_{k,j}$  对顶点  $((X, Y, Z)$ 、 $(t_x, t_y, \theta)$ ) 的雅克比矩阵为

$$\frac{\partial e_{k,j}}{\partial X} = \begin{bmatrix} \frac{\partial e_{x,k,j}}{\partial X} & \frac{\partial e_{x,k,j}}{\partial Y} & \frac{\partial e_{x,k,j}}{\partial Z} \\ \frac{\partial e_{y,k,j}}{\partial X} & \frac{\partial e_{y,k,j}}{\partial Y} & \frac{\partial e_{y,k,j}}{\partial Z} \end{bmatrix} \quad (3-18)$$

$$\frac{\partial e_{k,j}}{\partial \xi} = \begin{bmatrix} \frac{\partial e_{x,k,j}}{\partial t_x} & \frac{\partial e_{x,k,j}}{\partial t_y} & \frac{\partial e_{x,k,j}}{\partial \theta} \\ \frac{\partial e_{y,k,j}}{\partial t_x} & \frac{\partial e_{y,k,j}}{\partial t_y} & \frac{\partial e_{y,k,j}}{\partial \theta} \end{bmatrix} \quad (3-19)$$

在开源的 6-DoF SLAM 系统中，上述观测方程中的旋转矩阵一般为

$$R = \begin{bmatrix} c\phi c\theta & c\phi s\theta s\varphi - s\phi c\varphi & c\phi s\theta c\varphi - s\phi s\varphi \\ s\phi c\theta & s\phi s\theta s\varphi + c\phi c\varphi & s\phi s\theta c\varphi + c\phi s\varphi \\ -s\theta & c\theta s\varphi & c\theta c\varphi \end{bmatrix} \quad (3-20)$$

式中  $s\theta = \sin\theta$ ， $c\phi = \cos\phi$ 。

可知该旋转矩阵 R 比 3-DOF 中的旋转矩阵 R 计算复杂度要大得多。

### 3.4.2 平面 3-DOF VSLAM 算法仿真

通过对一组模拟数据应用上述的 3-DOF 和 6-DOF SLAM 系统进行对比，以验证本文提出的 3-DOF SLAM 系统具有更高的精度以及帧率。模拟数据（共 481 帧）假设机器人在一个球内的平面上运动，其运动轨迹是  $15\text{m} \times 15\text{m}$  的方形，球的内表面是通过一张 360 度全景照片覆盖而成。模拟数据部分截图如图 3-7 所示，3-DoF 和 6-DoF 优化时间对比如表 3-3 所示，3-DoF 和 6-DoF 优化机器人轨迹对比如图 3-8 所示。

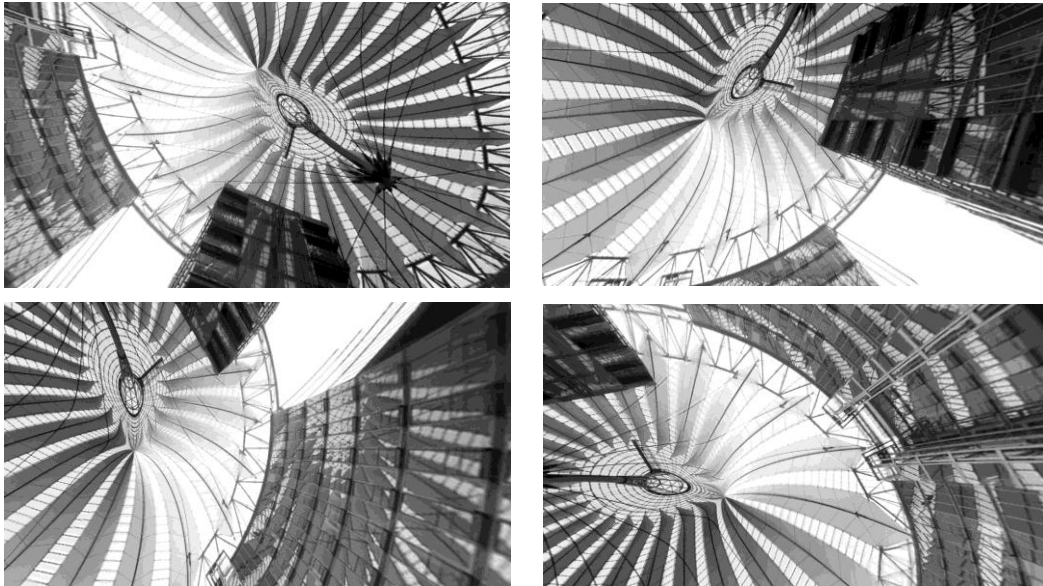


图 3-7 模拟数据部分截图

表 3-3 3-DoF 和 6-DoF 优化时间对比

3-DoF 优化结果						6-DoF 优化结果				
次数	1	2	3	4	5	1	2	3	4	5
耗时/s	37.21	37.33	37.54	36.97	37.04	42.79	42.62	42.51	42.89	43.87
均值/s	37.2238					42.9401				
每帧耗时/ms	77.3884					89.2725				
帧率/Hz	12.922					11.202				

从表 3-3 可知，在笔记本（Core i7 7700 四核 2.8GHz）上运行 6-DoF 优化时平均每帧耗时 89.2725ms，帧率 11.202Hz，而运行 3-DoF 优化时平均每帧耗时 77.3884ms，帧率 12.922Hz，故 3-DoF 优化下的耗时比 6-DoF 优化平均每帧减小 11.8841ms，帧率提高 15.35%。

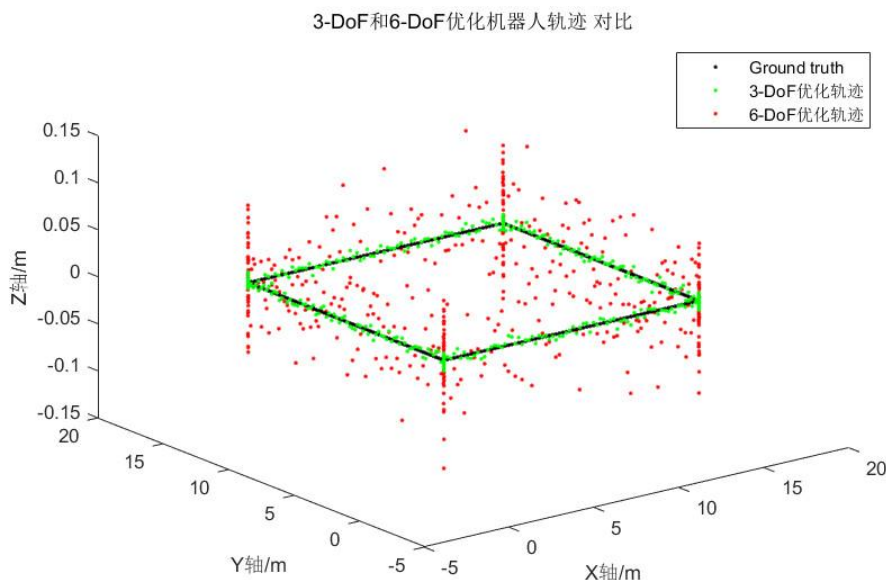


图 3-8 3-DoF 和 6-DoF 优化机器人轨迹对比

从图 3-8 可知，6-DoF 优化下的最大误差为 10.6cm，而 3-DoF 优化下的最大误差为 1.2cm，故 3-DoF 优化下的精度比 6-DoF 优化提高 8.83 倍。因此，证实了本文所提出的 3-DoF SLAM 系统具有更高的精度和帧率。

### 3.5 本章小结

本章首先研究了 ORB-SLAM2 的系统架构和三线程的工作原理，针对目前 ORB-SLAM2 中 DBoW2 词汇库存在加载时间长和占用大量系统内存等问题，对 DBoW2 进行了优化。同时训练得到了家庭场景下的词汇库，并对训练得到

的词汇库进行了实验测试，实验证实了本文的 **SKB-SLAM** 系统加载词汇库耗时相比 **ORB-SLAM2** 减少了 16.267 倍，内存占用量减小了 231.687Mb。同时相似度得分实验证实了本文针对家庭场景训练的词汇库是可以用于帧间匹配和回环检测。另一方面，针对三维空间 6-DoF 下的优化结果误差大以及实时性差等问题，对平面移动机器人 3-DoF SLAM 进行了优化研究，仿真实验结果证实了 3-DoF 优化比 6-DoF 优化的帧率提高了 15.35%，精度提高了 8.83 倍。

## 第 4 章 基于多传感器融合的紧耦合算法研究

### 4.1 引言

目前主流的 VIO（视觉惯导里程计）算法都是结合相机和 IMU 进行运动估计，而 IMU 用来做预积分计算两帧图像间的位姿增量。消费级 IMU 的数据噪声大和零偏随着时间而漂移，因此仅基于消费级 IMU 的 VIO 算法误差大、精度低、实时性差。

本章首先通过融合多传感器提出惯导预积分算法，通过融合相机、里程计、光流计和 IMU 为移动机器人提供准确和鲁棒的运动估计，然后根据视觉测量和惯导预积分测量通过松耦合算法估计视觉尺度  $s$ 。提出基于多传感器融合的紧耦合算法，实时估计准确、鲁棒的机器人位姿。

### 4.2 VIO 系统架构

#### 4.2.1 系统描述

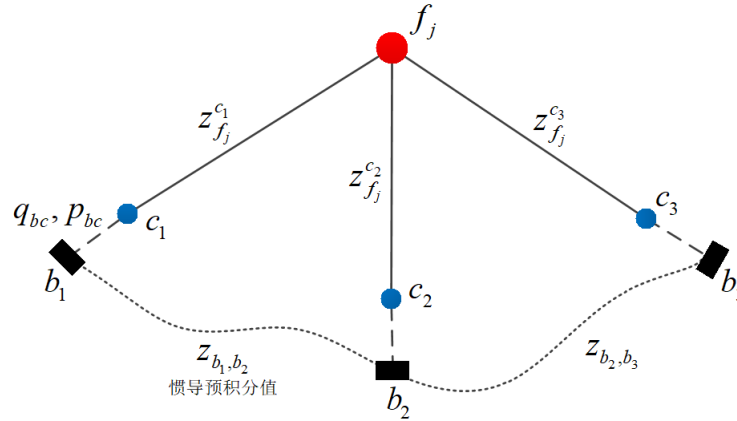


图 4-1 VIO 系统示意图

如图 4-1 所示为 VIO 系统示意图， $c_i$  表示  $i$  时刻的图像帧，且以第一帧图像坐标系作为世界坐标系  $(\cdot)_w$ ， $b_i$  表示  $i$  时刻的惯导数据帧。 $(\cdot)^c$  表示图像帧  $c$  中的向量  $(\cdot)$ 。四元数  $q_{yx}$  表示从图像帧  $x$  变换到图像帧  $y$  的旋转向量，对应的旋转矩阵为  $R_{yx}$ ； $p_{yx}$  表示从图像帧  $x$  变换到图像帧  $y$  的平移向量。四元数  $q_{bc}$  和向量  $p_{bc}$  表示从相机坐标系到机器人坐标系的变换矩阵。 $f_j$  表示地图中第  $j$  个特征

点， $z$  代表观测值， $z_{f_j}^{c_i}$  表示第  $i$  帧图像观测到的第  $j$  个特征点， $z_{b_i b_j}$  代表两个关键帧间的惯导预积分值。

#### 4.2.2 惯导预积分

里程计和光流计提供位移增量信息，IMU测量的角速度  $\omega$  相对于惯性系而言，IMU的角速度测量值  $\hat{\omega}$  容易受到零偏和噪声的影响：

$$\hat{\omega}^b = \omega^b + b_g^b + n_g^b \quad (4-1)$$

式 (4-1) 中  $b_g^b$ 、 $n_g^b$  分别为陀螺仪的零偏和噪声。因此机器人的运动学方程可以由下式给出

$$\begin{aligned} \dot{p}_{wb_i} &= v_t^w \\ \dot{v}_t^w &= a^{wb_i} = \Delta p_{wo_i} / \Delta t^2 \\ \dot{q}_{wb_i} &= q_{wb_i} \otimes \begin{bmatrix} 0 \\ \frac{1}{2} \omega^{b_i} \end{bmatrix} \end{aligned} \quad (4-2)$$

式 (4-2) 中  $p_{wo_i}$  为光流计（或里程计）得到的位移增量，符号  $\otimes$  表示四元数相乘。

通过IMU、里程计和光流计获取的数据，以及机器人在  $i$  时刻的状态量  $p_{wb_i}$ 、 $v_i^w$  和  $q_{wb_i}$ ，对式 (4-2) 积分即可得到机器人在  $j$  时刻的状态量

$$\begin{aligned} p_{wb_j} &= p_{wb_i} + v_i^w \Delta t + R_{wb_i} \alpha_{b_i b_j} \\ v_j^w &= v_i^w + R_{wb_i} \beta_{b_i b_j} \\ q_{wb_j} &= q_{wb_i} \otimes q_{b_i b_j} \end{aligned} \quad (4-3)$$

式中

$$\begin{aligned} \alpha_{b_i b_j} &= \iint_{t \in [i, j]} (R_{b_i b_t} a^{b_t}) \delta t^2 \\ \beta_{b_i b_j} &= \int_{t \in [i, j]} (R_{b_i b_t} \omega^{b_t}) \delta t \\ q_{b_i b_j} &= \int_{t \in [i, j]} q_{b_i b_t} \otimes \begin{bmatrix} 0 \\ \frac{1}{2} \omega^{b_t} \end{bmatrix} \delta t \end{aligned} \quad (4-4)$$

$z_{b_i b_j} = [\alpha_{b_i b_j}, \beta_{b_i b_j}, q_{b_i b_j}]$  被称为预积分测量值，在机器人状态未知的情况下能够直接计算该值。在连续运动中，将预积分测量值当做常量来处理。

预积分模型（公式（4-4））是在连续时间下推导且忽略了零偏和噪声，而实际上惯导数据是对离散点进行采样，因此本文以当前时刻的采样值作为两帧间的测量值进行积分（上一时刻到当前时刻的增量作为当前的积分值）。

实际上零偏会缓慢的变化，因此用随机游走噪声来对零偏进行建模

$$b_g^{b_{k+1}} = b_g^{b_k} + n_{b_g} \delta t \quad (4-5)$$

式（4-5）中的高斯白噪声服从分布  $n_{b_g} \in N(0, \sigma_{b_g}^2)$ 。在零偏值缓慢变化的情况下，用上一次的近似值去更新  $\hat{\alpha}_{b_i b_j}, \hat{\beta}_{b_i b_j}, \hat{q}_{b_i b_j}$ ，而不是迭代计算预积分测量值。

$$\begin{aligned} \hat{\alpha}_{b_i b_j} &\leftarrow \hat{\alpha}_{b_i b_j} + J_{b_g^i}^{\alpha} \delta b_g^{b_i} \\ \hat{\beta}_{b_i b_j} &\leftarrow \hat{\beta}_{b_i b_j} + J_{b_g^i}^{\beta} \delta b_g^{b_i} \\ \hat{q}_{b_i b_j} &\leftarrow \hat{q}_{b_i b_j} \otimes \begin{bmatrix} 1 \\ \frac{1}{2} J_{b_g^i}^q \delta b_g^{b_i} \end{bmatrix} \end{aligned} \quad (4-6)$$

式中  $J_{b_g^i}^{\alpha} = \frac{\partial \alpha_{b_i b_j}}{\partial \delta b_g^{b_i}}$ 、 $J_{b_g^i}^{\beta} = \frac{\partial \beta_{b_i b_j}}{\partial \delta b_g^{b_i}}$ 、 $J_{b_g^i}^q = \frac{\partial q_{b_i b_j}}{\partial \delta b_g^{b_i}}$  是零偏值在  $i$  时刻预积分测量值的

雅克比矩阵。

### 4.3 视觉惯导初始化状态估计

由于单目 SLAM 系统没有绝对的尺度信息，因此大部分研究者结合 IMU 来获取绝对尺度，但通过 IMU 得到的尺度误差大、精度低，直接导致定位精度低。另一方面，由于单目 SLAM 系统非线性优化阶段严重依赖初始化状态：IMU 零偏、视觉尺度等信息。因此本文提出通过融合里程计、光流计和 IMU 为单目 SLAM 系统提供精确、可靠的尺度信息，确保 VISLAM 系统的定位结果精度高。

初始化状态估计采用松耦合的方式，首先一个线程单独利用视觉测量值来得到少量的关键帧，同时另一个线程计算这些关键帧间的惯导预积分值。在上述过程中，当一个新的关键帧被创建时，松耦合算法通过迭代更新陀螺仪零偏和视觉尺度，直到尺度的估计值接近于真实值将触发紧耦合融合。初始化状态估计算法流程如图 4-2 所示。



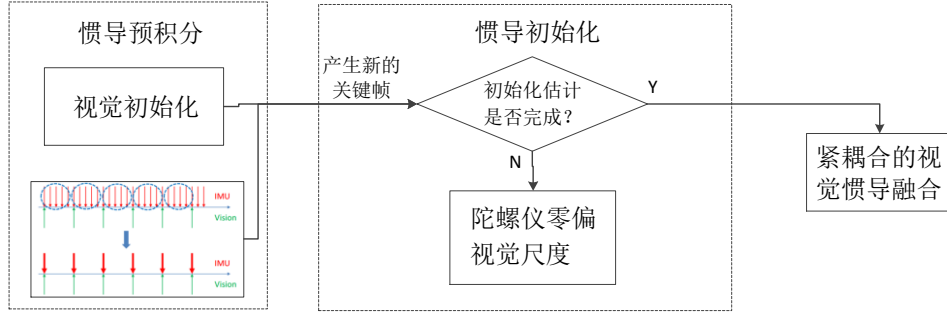


图 4-2 初始化状态估计算法流程

在松耦合估计视觉惯导初始化状态中，首先估计陀螺仪零偏，然后在不考虑加速度计零偏的情况下估计视觉尺度。

#### 4.3.1 陀螺仪零偏估计

考虑两个连续的视觉关键帧，通过视觉 SLAM 得到两个关键帧的方位分别为  $q_{wb_i}, q_{wb_j}$ ，以及惯导预积分的测量值  $q_{b_i b_j}$ 。通过最小化视觉旋转与惯导预积分旋转的残差来估计陀螺仪零偏：

$$\arg \min_{b_g} \sum_{i=1}^{N-1} \left\| \text{Log}((q_{b_i b_j} \text{Exp}(J_q^s b_g))^T q_{b_i w} q_{wb_j}) \right\|^2 \quad (4-7)$$

式（4-7）中  $N$  为视觉关键帧数， $J_q^s$  表示陀螺仪零偏变化的一阶近似。通过高斯牛顿或 Levenberg-Marquard 法求解上式即可得到陀螺仪零偏估计值  $b_g$ 。陀螺仪零偏在刚开始时被设置为 0，而在每一次迭代中都可以根据估计的零偏值  $b_g$  来更新预积分测量值  $\hat{\alpha}_{b_i b_j}, \hat{\beta}_{b_i b_j}, \hat{q}_{b_i b_j}$ 。

#### 4.3.2 视觉尺度估计

首先假设视觉尺度为  $s$ ，并将相机轨迹从相机坐标系变换到机器人坐标系：

$$p_{wb} = s p_{cb} + R_{wc} p_{cb} \quad (4-8)$$

将式（4-8）代入式（4-3）中得，表示两个连续关键帧间的位姿关系

$$\left[ \Delta p_{i,j} - R_{wb}^i (R_{wc}^j - R_{wc}^i) p_{cb} \right] = \begin{bmatrix} -R_{wb}^i \Delta t_{i,j} & R_{wb}^i (p_{wc}^j - p_{wc}^i) \end{bmatrix} \begin{bmatrix} v_i \\ s \end{bmatrix} \quad (4-9)$$

如果将式（4-9）进行累加，将有  $N-1$  个  $v_i$  需要求解，必将耗费大量的计算资源。故本文只选取连续的 3 帧（依次为  $i, j, k$ ）进行累加

$$\begin{aligned}
 \hat{z}_{i,j,k} &= [(R_{wc}^i - R_{wc}^j) p_{cb} \Delta t_{j,k} - (R_{wc}^j - R_{wc}^k) p_{cb} \Delta t_{i,j} - R_{wb}^j \Delta p_{j,k} \Delta t_{i,j} \\
 &\quad - R_{wb}^i \Delta v_{i,j} \Delta t_{i,j} \Delta t_{j,k} + R_{wb}^i \Delta p_{i,j} \Delta t_{j,k}] s \\
 &= [(p_{wc}^j - p_{wc}^i) \Delta t_{j,k} - (p_{wc}^k - p_{wc}^j) \Delta t_{i,j}] s \\
 &= H_{i,j,k} s
 \end{aligned} \tag{4-10}$$

式(4-10)中  $p_{wc}^{(\cdot)}$  和  $R_{wc}^{(\cdot)}$  从视觉 SLAM 中获取,  $\Delta p_{(\cdot)}$  和  $\Delta v_{(\cdot)}$  从惯导预积分中获取,  $\Delta t_{(\cdot)}$  为连续两关键帧间的时间间隔。

通过对式(4-10)中每三个关键帧连续积分, 即可将该问题转化为最小二乘问题求解, 因此得到尺度的求解公式

$$\min_s \sum_{i=1}^{N-2} \|\hat{z}_{i,j,k} - H_{i,j,k} s\|^2 \tag{4-11}$$

#### 4.4 紧耦合的视觉惯导融合算法

在VSLAM图优化中, 通过在像素平面内最小化重投影误差来优化相机位姿和 3D 地图点, 如图 4-3(a)所示。而在VISLAM图优化中, 通过紧耦合图优化模型来优化所有的状态变量, 如图 4-3(b)所示。

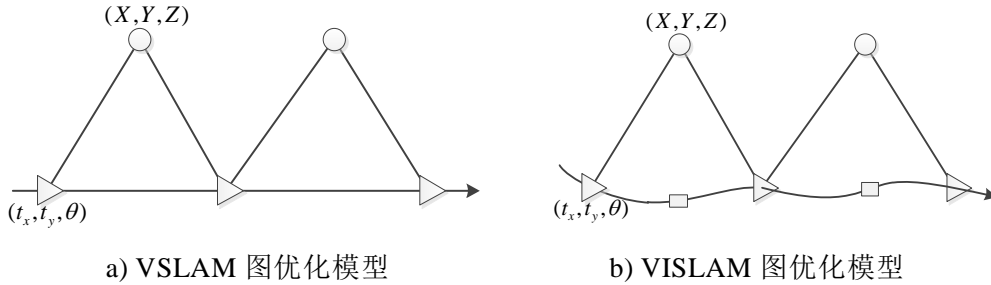


图 4-3 图优化模型

图中圆圈表示 3D 特征点, 三角形表示相机位姿, 方形表示惯导预积分值, 连接圆圈和三角形的边为优化误差项。

##### 4.4.1 滑动窗口模型

为了提高计算效率和定位精度, 本文以滑动窗口模型来实现图优化, 因此定义  $i$  时刻滑动窗口优化的状态变量:

$$\begin{aligned}
 \chi &= [x_n, x_{n+1}, \dots, x_{n+N}, \lambda_m, \lambda_{m+1}, \dots, \lambda_{m+M}] \\
 x_i &= [p_{wb_i}, q_{wb_i}, v_i^w, b_i^b]
 \end{aligned} \tag{4-12}$$

式中  $x_i$  表示全局参考坐标系下机器人的位置、方位、速度，以及IMU的零偏。使用逆深度  $\lambda_k$  来参数化表示第  $k$  个特征点。 $n$ 、 $m$  分别表示机器人状态量和 3D 特征点的起始下标， $N$ 、 $M$  表示滑动窗口中关键帧数和 3D 特征点数。

在滑动窗口中通过最小化所有测量模型的残差和来优化状态变量：

$$\min_{\chi} \rho(\|r_p - J_p \chi\|_{\Sigma_p}^2) + \sum_{i \in B} \rho(\|r_b(z_{b_i b_{i+1}}, \chi)\|_{\Sigma_{b_i b_{i+1}}}^2) + \sum_{i \in E} \rho(\|r_e(z_{e_i e_{i+1}}, \chi)\|_{\Sigma_{e_i e_{i+1}}}^2) \\ + \sum_{i \in O} \rho(\|r_o(z_{o_i o_{i+1}}, \chi)\|_{\Sigma_{o_i o_{i+1}}}^2) + \sum_{(i,j) \in F} \rho(\|r_f(z_{f_j}^{c_i}, \chi)\|_{\Sigma_{f_j}^{c_i}}^2) \quad (4-13)$$

式中  $r_b(z_{b_i b_{i+1}}, \chi)$ 、 $r_e(z_{e_i e_{i+1}}, \chi)$ 、 $r_o(z_{o_i o_{i+1}}, \chi)$  分别为IMU、里程计、光流计在状态  $x_i, x_{i+1}$  间的测量残差， $B$ 、 $E$ 、 $O$  分别为IMU、里程计、光流计在滑动窗口中惯导预积分的测量值； $r_f(z_{f_j}^{c_i}, \chi)$  为 3D 特征点的重投影误差， $F$  为相机拍摄图片中的特征点数。 $r_p$  为在滑动窗口中用于计算的先验信息， $J_p$  为从上一次优化结果 Hessian 矩阵中得到的雅克比矩阵。 $\rho$  为Cauchy核用于剔除外点。

在图优化中一般选用高斯牛顿或Levenberg-Marquard算法进行优化（LM法在一定程度上修正了GN法奇异的问题，鲁棒性更好，但收敛速度可能比GN慢）。因此本文选用Levenberg-Marquard算法进行非线性优化，状态估计  $\chi$  能够通过初始估计  $\chi_0$  迭代得到：

$$\chi'_{t+1} = \chi_t \oplus \delta\chi \quad (4-14)$$

式中  $\oplus$  符号用于对增量进行更新，位置、速度、逆深度和零偏的更新操作可以定义为：

$$p' = p + \delta p, \quad v' = v + \delta v, \quad \lambda' = \lambda + \delta \lambda, \quad b' = b + \delta b \quad (4-15)$$

然而由于方位是以四元数进行表示，因此方位  $q$  的更新比较复杂。针对平面移动机器人在 3-DoF 下的方位更新，本文采用扰动模型  $\delta\theta \in R^3$  作为旋转增量表示。因此方位  $q$  的更新通过四元数相乘即可：

$$q' = q \otimes \delta q, \quad \delta q = \begin{bmatrix} 1 \\ \frac{1}{2} \delta\theta \end{bmatrix} \quad (4-16)$$

同理可以将四元数写成旋转矩阵的形式。

因此，图优化中增量  $\Delta\chi$  可以表示为如下形式

$$\delta\chi = [\delta x_n, \delta x_{n+1}, \dots, \delta x_{n+N}, \delta \lambda_n, \delta \lambda_{n+1}, \dots, \delta \lambda_{n+M}]^T \\ \delta x_i = [\delta p, \delta\theta, \delta v, \delta b_g^{b_i}]^T, \quad i \in [n, n+N] \quad (4-17)$$

在每一次的迭代中，增量  $\Delta\chi$  可以通过下式求得

$$(H_p + H_b + H_e + H_o + H_f)\delta\chi = (b_p + b_b + b_e + b_o + b_f) \quad (4-18)$$

对于残差  $r(\cdot)$ ，定义

$$H(\cdot) = J_{(\cdot)}^T \Sigma_{(\cdot)}^{-1} J_{(\cdot)}, \quad b(\cdot) = -J_{(\cdot)}^T \Sigma_{(\cdot)}^{-1} r(\cdot) \quad (4-19)$$

式中  $J(\cdot)$  为残差  $r(\cdot)$  相对于  $\delta\chi$  的雅克比矩阵， $\Sigma_{(\cdot)}$  为测量值的协方差矩阵。

#### 4.4.2 相机测量模型

重投影误差是计算 3D 特征点在图像平面内投影与观测像素点的距离，本文将相机看做针孔相机模型进行处理，首先已知图像帧  $c_j$  中第  $k$  个特征点的观测值  $z_{f_k}^{c_j} = [u_{f_k}^{c_j}, v_{f_k}^{c_j}, 1]^T$ ，则重投影误差定义为

$$r_f(z_{f_k}^{c_j}, \chi) = \begin{bmatrix} \frac{x^{c_j}}{z^{c_j}} - u_{f_k}^{c_j} \\ \frac{y^{c_j}}{z^{c_j}} - v_{f_k}^{c_j} \end{bmatrix} \quad (4-20)$$

式中  $[x^{c_j}, y^{c_j}, z^{c_j}]^T$  定义如下

$$f_k^{c_j} = \begin{bmatrix} x^{c_j} \\ y^{c_j} \\ z^{c_j} \end{bmatrix} = R_{bc}^T (R_{wb_j}^T (R_{wb_i} ((R_{bc} \frac{1}{\lambda_k} \begin{bmatrix} u_{f_k}^{c_i} \\ v_{f_k}^{c_i} \\ 1 \end{bmatrix} + p_{bc}) + p_{wb_i}) - p_{wb_j}) - p_{bc}) \quad (4-21)$$

式中  $[u_{f_k}^{c_i}, v_{f_k}^{c_i}, 1]^T$  为图像帧  $c_i$  中第  $k$  个特征点的观测值，上式将图像帧  $c_i$  中第  $k$  个特征点的观测值重投影到图像帧  $c_j$  中第  $k$  个特征点的观测值。

为了使得式 (4-20) 重投影误差最小化，因此需要优化帧  $b_i, b_j$  中的位置和方位，以及特征点的逆深度  $\lambda$ 。求得相应的雅克比矩阵为

$$J_f = \frac{\partial r_f}{\partial f^{c_j}} \begin{bmatrix} \frac{\partial f^{c_j}}{\partial x_i} & \frac{\partial f^{c_j}}{\partial x_j} & \frac{\partial f^{c_j}}{\partial \delta\lambda} \end{bmatrix} \quad (4-22)$$

式中

$$\frac{\partial r_f}{\partial f^{c_j}} = \begin{bmatrix} \frac{1}{z^{c_j}} & 0 & -\frac{x^{c_j}}{(z^{c_j})^2} \\ 0 & \frac{1}{z^{c_j}} & -\frac{y^{c_j}}{(z^{c_j})^2} \end{bmatrix} \quad (4-23)$$

$$\begin{aligned}
 \frac{\partial f^{c_j}}{\partial x_i} &= \begin{bmatrix} R_{bc}^T R_{wb_j}^T & -R_{bc}^T R_{wb_j}^T R_{wb_i} [f^{b_i}] & 0 & 0 \end{bmatrix}_{3 \times 12} \\
 \frac{\partial f^{c_j}}{\partial x_j} &= \begin{bmatrix} -R_{bc}^T R_{wb_j}^T & R_{bc}^T [f^{b_j}] & 0 & 0 \end{bmatrix}_{3 \times 12} \\
 \frac{\partial f^{c_j}}{\partial \delta \lambda} &= -\frac{1}{\lambda} R_{bc}^T R_{wb_j}^T R_{wb_i} R_{bc} f_{c_i}
 \end{aligned} \tag{4-24}$$

式中  $f^{b_i}$  为式 (4-21) 中的  $[u_{f_k}^{c_i}, v_{f_k}^{c_i}, 1]^T$ 。

#### 4.4.3 里程计和光流计测量模型

里程计测量左右轮的位移量，故可间接得到机器人中心的位移量（假设轮轴位于机器人中轴线上）；而光流计测量两帧数据间的位移增量，故也可间接得到机器人中心的位移增量，因此里程计和光流计的测量残差相同。故里程计和光流计的测量残差定义如下

$$r_e(z_{b_i b_j}, \chi) = \begin{bmatrix} \Delta R_{b_i b_j} \\ \Delta p_{b_i b_j} \end{bmatrix} = \begin{bmatrix} R_{wb_i} R_{wb_j}^T \text{Exp}(\delta \theta_{b_i b_j}) \\ -R_{wb_i} R_{wb_j}^T p_{wb_j} + p_{wb_i} + \delta p_{b_i b_j} \end{bmatrix} \tag{4-25}$$

在图优化中，需要优化帧  $b_i, b_j$  下的状态  $x_i, x_j$ ，求得相应的雅克比矩阵为

$$J_e = \begin{bmatrix} \frac{\partial r_e}{\partial \delta x_i} & \frac{\partial r_e}{\partial \delta x_j} \end{bmatrix} \tag{4-26}$$

式中

$$\begin{aligned}
 \frac{\partial r_e}{\partial \delta x_i} &= \begin{bmatrix} 0 & (q_{wb_i}^{-1} \otimes q_{wb_i}) \text{Exp}(\delta \theta_{b_i b_j}) & 0 & 0 \\ I & 0 & v_i^w & 0 \end{bmatrix}_{6 \times 12} \\
 \frac{\partial r_e}{\partial \delta x_j} &= \begin{bmatrix} 0 & (\hat{q}_{b_i b_j}^{-1} \otimes q_{wb_i}^{-1}) \text{Exp}(\delta \theta_{b_i b_j}) & 0 & 0 \\ -R_{wb_i} R_{wb_j}^T & 0 & -R_{wb_i} R_{wb_j}^T v_j^w & 0 \end{bmatrix}_{6 \times 12}
 \end{aligned} \tag{4-27}$$

#### 4.4.4 IMU 测量模型

在惯导预积分 4.2.2 节中已经推导了连续两个关键帧间的惯导预积分公式，故IMU测量残差可定义为

$$r_b(z_{b_i b_j}, \chi) = \begin{bmatrix} r_\theta \\ r_{bg} \end{bmatrix} = \begin{bmatrix} 2[\hat{q}_{b_i b_j} \otimes (q_{b_i w} \otimes q_{wb_j})] \\ b_g^{b_j} - b_g^{b_i} \end{bmatrix} \tag{4-28}$$

在图优化中，需要优化帧  $b_i, b_j$  下的状态  $x_i, x_j$ ，求得相应的雅克比矩阵为

$$J_b = \begin{bmatrix} \frac{\partial r_b}{\partial \delta x_i} & \frac{\partial r_b}{\partial \delta x_j} \end{bmatrix} \quad (4-29)$$

式中

$$\begin{aligned} \frac{\partial r_b}{\partial \delta x_i} &= \begin{bmatrix} 0 & -[q_{wb_i}^{-1} \otimes q_{wb_i}]_L [\hat{q}_{b_i b_j}]_R & 0 & J_{b_i}^{r_\theta} \\ 0 & 0 & 0 & -I \end{bmatrix}_{6 \times 12} \\ \frac{\partial r_b}{\partial \delta x_j} &= \begin{bmatrix} 0 & -[\hat{q}_{b_i b_j}^{-1} \otimes q_{wb_i}^{-1} \otimes q_{wb_j}]_L & 0 & 0 \\ 0 & 0 & 0 & I \end{bmatrix}_{6 \times 12} \end{aligned} \quad (4-30)$$

式 (4-30) 中的  $J_{b_i}^{r_\theta}$  定义如下

$$J_{b_i}^{r_\theta} = \frac{\partial r_\theta}{\partial \delta b_{b_i}^{b_i}} = -[q_{wb_i}^{-1} \otimes q_{wb_i} \otimes q_{b_i b_j}] J_{b_i}^{q_{b_i}} \quad (4-31)$$

## 4.5 本章小结

本章首先针对目前主流的预积分算法（仅利用IMU）误差大、精度低等不足，通过融合多传感器提出了惯导预积分算法，能够准确测量相邻关键帧间的位姿增量，以及为移动机器人提供准确和鲁棒性的位姿估计。得到准确的预积分测量值后，通过松耦合算法估计视觉尺度s，进一步提高了VISLAM系统的定位精度和鲁棒性。最后提出了基于多传感器融合的紧耦合算法，并推导了系统模型，相机、里程计、光流计和IMU的观测模型，能够实时估计准确、鲁棒的机器人位姿。

## 第 5 章 移动机器人平台搭建及实验研究

### 5.1 引言

在前面的章节中，分别对基于 EKF 数据融合的定位算法、DBoW 词汇库优化、平面 3-DoF VSLAM 算法优化和基于多传感器融合的紧耦合算法进行了研究，并给予了一定的实验验证。本章将整合以上算法，在移动机器人平台上实现室内精确定位的功能。

本章首先介绍移动机器人平台搭建及相应的软件设计，然后利用实验平台在室内不同环境下、不同条件下进行定位实验，用以测试本文提出的定位算法的实际可用性。

### 5.2 移动机器人平台搭建

#### 5.2.1 Create2 硬件平台搭建

本课题的移动机器人平台采用 iRobot 公司的 Create2 扫地机器人通用开发平台（本文中 Create2 只作为行走执行机构），Create2 作为一款玩家级别的扫地机器人平台，对用户开放编程接口。Create2 扫地机器人主要性能参数如表 5-1 所示。

表 5-1 Create2 扫地机器人主要性能参数

参数	数值
整机尺寸	直径 348.5mm，高 92.25mm
轮轴长	235mm
车轮直径	72mm
编码器脉冲数	508.8/转
串口波特率	115200Bd/s
车轮速度范围	-500 ~ 500mm/s

通过在 Create2 平台上加装 IMU 模块、光流传感器、蓝牙模块、STM32F407 控制器、摄像头模组、以及嵌入式 R16 平台，以此来满足室内移动机器人对定位的要求，各硬件的具体信息如表 5-2 所示。里程计信息通过直接读取 Create2 左右编码器脉冲获得。移动机器人硬件平台结构图如图 5-1 所示。

表 5-2 实验平台硬件具体信息

硬件名称	型号	具体功能
摄像头	OV7251	采样图像信息
IMU	Razor IMU-6DoF	采样 Create2 角速度和加速度
光流计	ADNS-3080	采样 Create2 位移增量信息
蓝牙模块	HC-05	进行数据透穿
下位机主控	STM32F407 开发板 主频：168MHz	负责下位机传感器数据采样和 Create2 运动控制
上位机主控	R16 开发板 A7 四核 1GHz	接收摄像头和 STM32F407 下位机的采样数据
PC 端	笔记本电脑	用于 Create2 实时轨迹显示

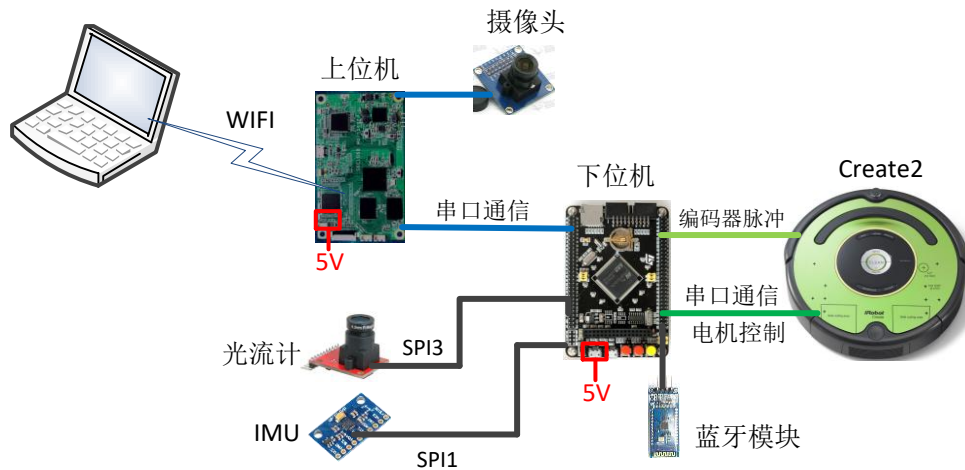


图 5-1 移动机器人硬件平台结构图

整个系统的工作流程：STM32F407 控制器作为移动机器人平台的下位机主控，负责采集 IMU、光流计的原始读数和编码器的脉冲信号（通过对 Z 轴角速度积分得到航向角 Yaw，将光流计增量信息进行积分，将码盘脉冲信号转化为实际的距离信息并积分）。由于 Create2 采用 3 线编码器（+5v，GND，A 相），故仅通过编码器无法判断车轮的实际转动方向，因此通过增加蓝牙模块，一方面用于实时控制 Create2 的行走状态，另一方面 STM32F407 根据获取的指令信息和码盘脉冲信号得到具有方向的脉冲信息。STM32F407 根据解码后的蓝牙指令通过串口控制 Create2 电机的运动状态。STM32F407 将所有的采集数据通过



串口发送给 R16 端进行融合。R16 端通过 WiFi 将移动机器人当前坐标和状态发送给 PC 端进行实时轨迹显示。Create2 实验平台实物图如图 5-2 所示。

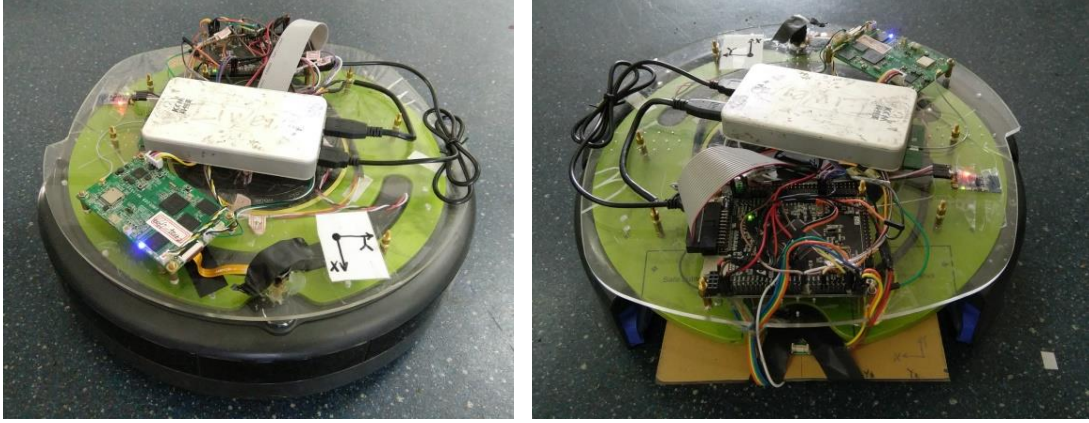


图 5-2 移动机器人平台实物图

### 5.2.2 STM32F407 平台软件设计

STM32F407 在整个系统中充当着数据采集以及控制 Create2 运动的角色，通过串口通信实时控制 Create2 的运动；通过硬同步在 3 个不相关的传感器间建立了联系。采用 uCOS-II 实时操作系统，并以系统时钟 SYSCLK（中断周期 1ms）作为多传感器数据采集的触发信号，以实现 3 个传感器数据采集的硬同步要求，硬同步采集数据的流程图如图 5-3 所示。

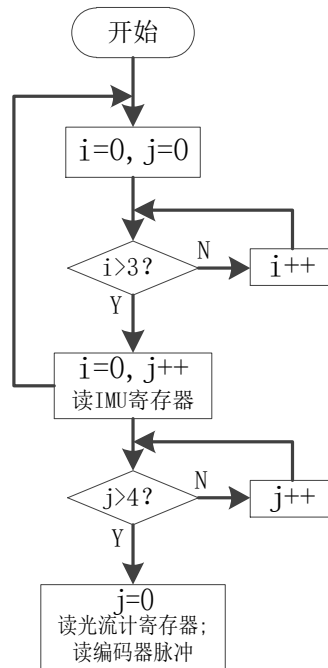
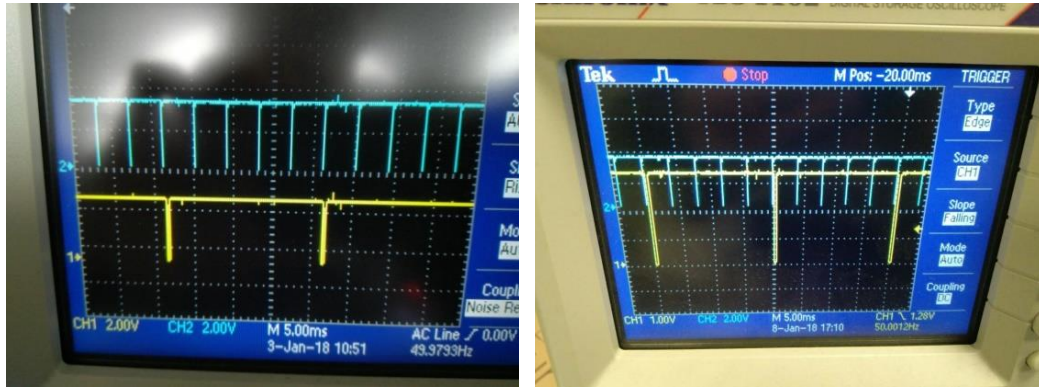


图 5-3 硬同步采集里程计、IMU 和光流计传感器数据

STM32F407 硬同步采样传感器数据的主要流程为：

- (1) STM32F407 控制器上电初始化，系统时钟 SYSCLK 以 1ms 为周期开始中断；
- (2) 当系统时钟中断 4 次后，计数值 i 置 0，j 加 1，并采样一次 IMU 数据；
- (3) 紧接着，i 继续累加；
- (4) 当计数值 j 达到 5 时，将计数值 j 置 0，并采样一次编码器和光流计数据；

因此，硬同步后的效果是 IMU 4ms 采样一次，光流计和编码器 20ms 采样一次，且 IMU 与光流计（或编码器）的采样时刻相对应。



（注：左图为 IMU 与光流计硬同步波形，右图为 IMU 与里程计硬同步波形）

图 5-4 三传感器硬同步采样波形图

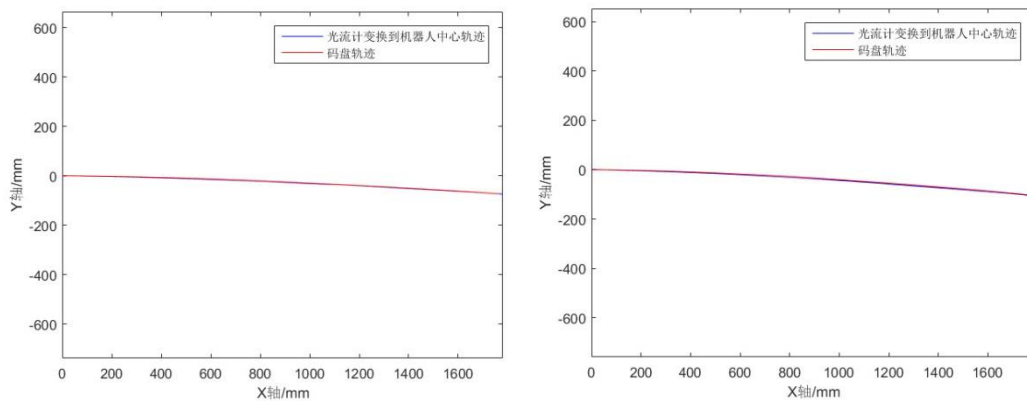
## 5.3 Create2 平台实体实验

### 5.3.1 不同材质表面对比实验

为了对比光流计和里程计在不同材质表面的实际表现性能，本实验将在两种不同的材质表面（短毛地毯和长毛地毯，由于在 2.5.2 小节已对比过在地砖上的实验，故省略此处地砖的对比实验）上进行实验，分别记录下 Create2 的实际位移、光流计融合和里程计融合的轨迹位移（本实验以机器人前进方向的距离作为记录值）。实验要求均为手动控制 Create2 前进指定的距离。

#### (1) 短毛地毯对比实验

在短毛地毯上 Create2 的融合轨迹如图 5-5 所示，坐标点记录及误差分析如表 5-3 所示。



(注：左图为第一次实验，右图为第二次实验)

图 5-5 短毛地毯融合轨迹对比实验

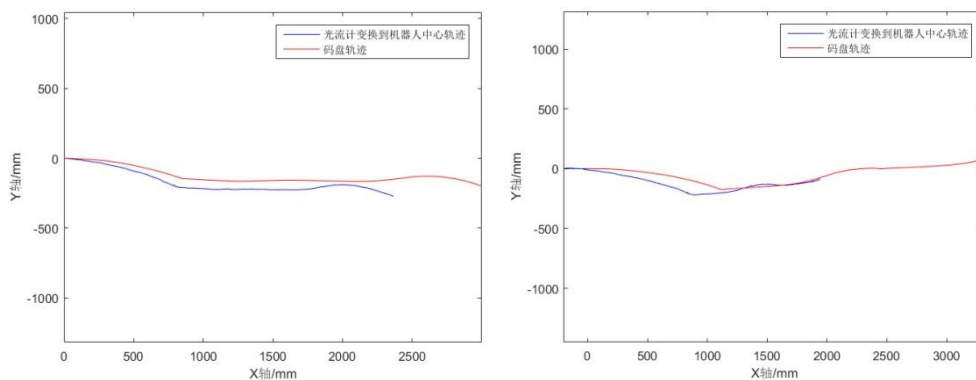
表 5-3 短毛地毯融合坐标及误差

单位：mm

实验次数	实际位移	光流计融合	误差	里程计融合	误差
第一次	1720	1779	59 (3.4%)	1788	68 (4.0%)
第二次	1750	1764	14 (0.8%)	1786	36 (2.1%)

## (2) 长毛地毯对比实验

在长毛地毯上 Create2 的融合轨迹如图 5-6 所示，坐标点记录及误差分析如表 5-4 所示。



(注：左图为第一次实验，右图为第二次实验)

图 5-6 长毛地毯融合轨迹对比实验

表 5-4 长毛地毯融合坐标及误差

单位：mm

实验次数	实际位移	光流计融合	误差	里程计融合	误差
第一次	1800	1943	143 (7.9%)	2999	1199 (66.6%)
第二次	1800	1866	66 (3.7%)	3310	1510 (83.9%)

从表 5-3 和表 5-4 的对比结果中可看出，在短毛地毯上光流计和里程计的定位精度相当（均小于 2.2.2 小节中的指标 4.52%）；在长毛地毯上光流计的定位误差最大为 7.9%，而里程计的定位误差最大达到了 83.9%。另一方面，图 5-5 和图 5-6 中的轨迹也符合 Create2 的实际轨迹形状。因此，实验结果证实了在不同材质表面上光流计都能够提供较为准确的定位信息，而里程计在长毛地毯（打滑情况下）上几乎处于失效状态，无法提供定位信息。从另一方面说明了相对定位的传感器在不同材质表面上，尺度信息是不一样的，只不过光流计的尺度信息更稳定。

### 5.3.2 两种场景下的定位实验

本文选定两种实验场景：室内小场景和室内大场景，如图 5-7 所示，用来验证本文提出的基于多传感器融合的紧耦合算法和 VISLAM 系统重定位功能的可行性。室内小场景和室内大场景主要用来模拟扫地机器人在家庭场景下工作过程中可能遇到的场景，图 5-7(a)为通过手动控制 Create2 在走廊环境下创建的一个小场景，用以模拟家庭小场景，大小为  $2\text{m} \times 4\text{m}$ ，扫地机器人除了在家场景工作外也常用于走廊、厂房等室内场景工作，而大场景对于扫地机器人的定位要求更苛刻，图 5-7(b)所示是一个场景特征相似的长走廊环境，大小为  $8\text{m} \times 42\text{m}$ ，主要用来测试本文 VISLAM 系统的重定位功能。本对比实验包括两大部分：分别为基于里程计的松耦合算法和基于多传感器融合的紧耦合算法的定位精度对比实验。



a) 室内小场景





b) 室内大场景

图 5-7 实验场景

### (1) 小场景下的定位实验

选取 Create2 平台起始点作为定位标记点，如图 5-8 所示，用来记录本实验的定位坐标。

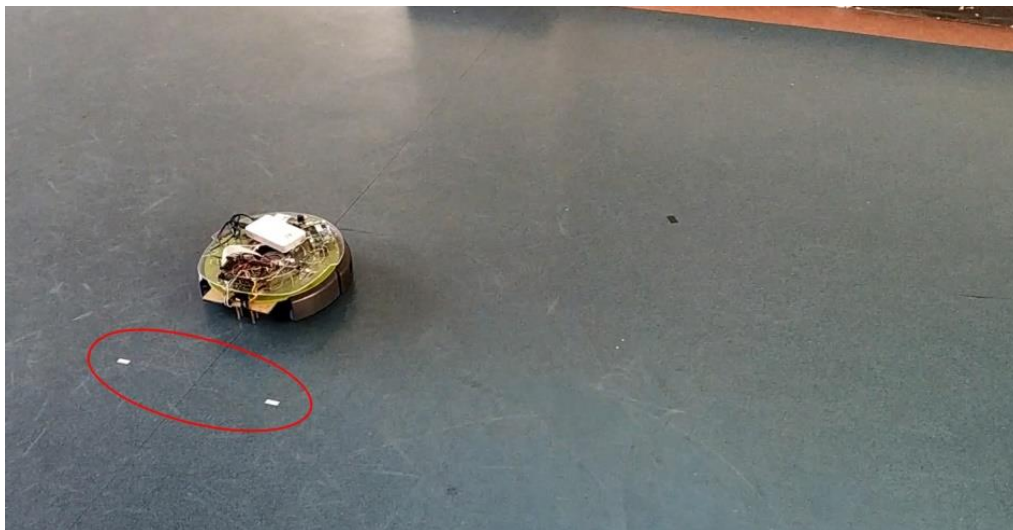


图 5-8 定位标记点

实验流程为 Create2 平台首先以紧耦合算法进行构建地图并记录回到起始点处的坐标，然后遮住摄像头进行纯积分版定位，紧接着模拟车轮打滑进行定位，最终以重定位修正光流计和里程计的累计误差。Create2 平台每次回到定位

标记点处的轨迹截图如图 5-9（基于里程计的松耦合算法）和图 5-10 所示（基于多传感器融合的紧耦合算法）。定位标记点坐标记录如表 5-5 所示。

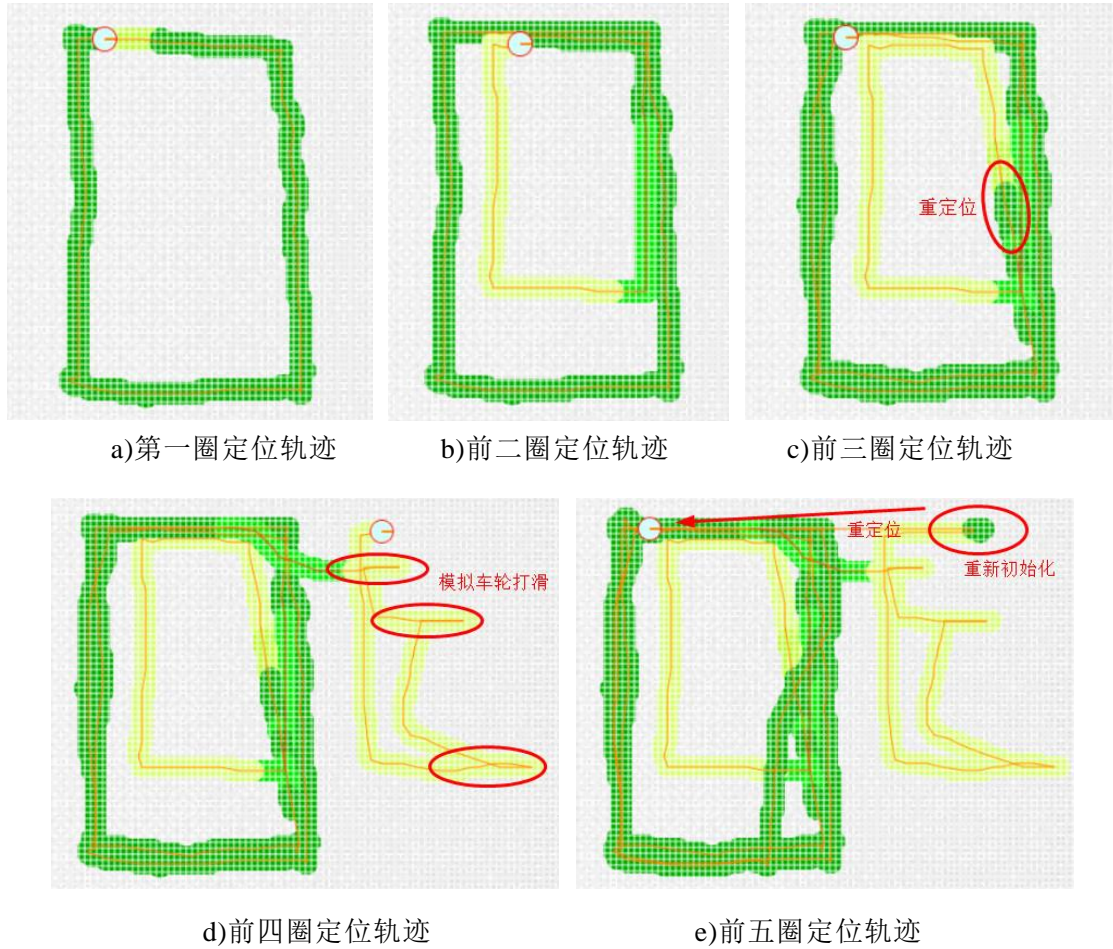
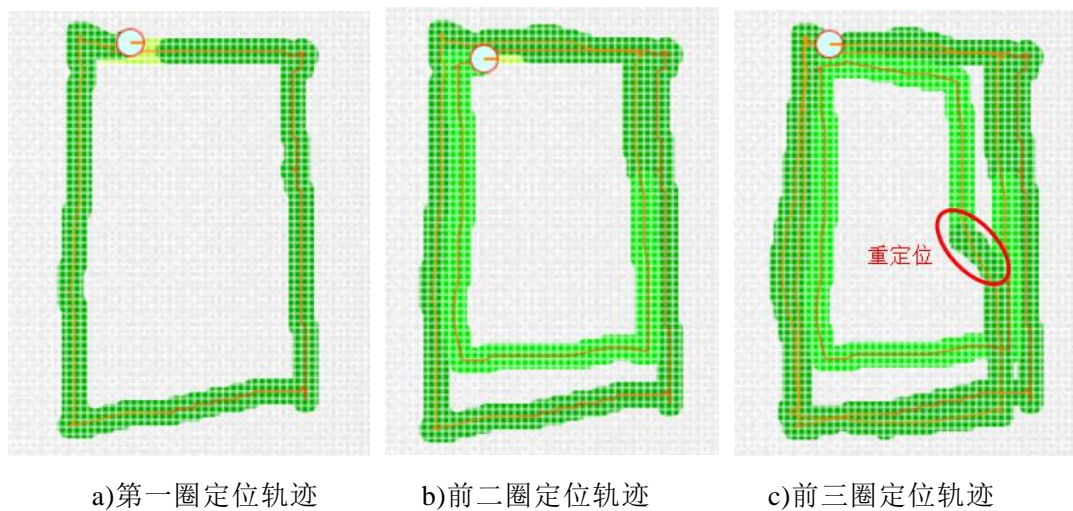


图 5-9 基于里程计的松耦合算法定位轨迹



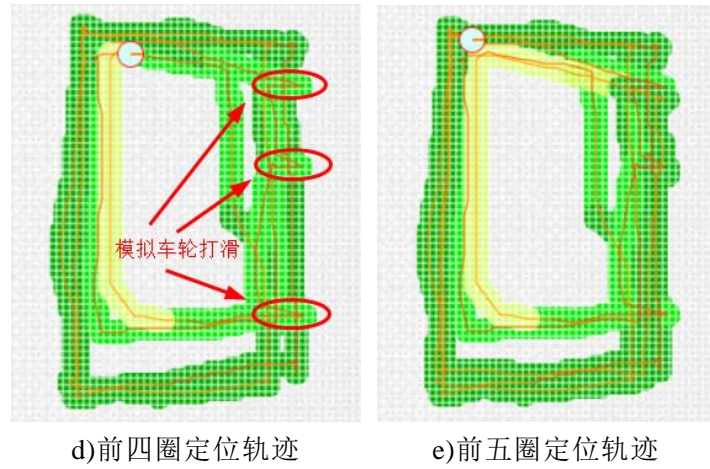


图 5-10 基于多传感器融合的紧耦合算法定位轨迹

通过对比图 5-9 和图 5-10 的(a)图可知,松耦合和紧耦合算法构建的地图精度相当,形状符合 Create2 的实际运动轨迹;

图 5-9 和图 5-10 的(b)图中遮住摄像头,对比可知,在进行纯积分版定位中由于光流计和里程计都存在尺度不一致问题,故定位轨迹都有一定程度的变形,而且里程计的变形要大于光流计;

图 5-9 和图 5-10 的(c)图中重新打开摄像头,对比可知,打开摄像头后都立刻进行了重定位,修正了里程计或光流计的累积误差;

图 5-9 和图 5-10 的(d)图中遮住摄像头模拟车轮打滑,对比可知,车轮在打滑过程中里程计的定位轨迹仍在累积,已经超出了实际的范围且轨迹变形异常,而光流计的定位轨迹符合 Create2 实际的运动轨迹,且形状与图 5-10(b)图中的轨迹几乎重合;

图 5-9 和图 5-10 的(e)图中重新打开摄像头,由于基于里程计的轨迹定位误差大,导致先进行了错误的重新初始化,然后重定位回正确的位置,而基于多传感器融合的轨迹误差小,直接重定位正确。

表 5-5 定位标记点坐标记录及误差分析

单位: mm

圈数	基于里程计的松耦合算法	基于多传感器融合的紧耦合算法
起始点	(0,0) → 0	(0,0) → 0
第一圈	(-22,-6) → 22.8	(-16,9) → 18.4
第二圈	(468,-127) → 484.9	(80,-95) → 124.2
第三圈	(-23,-11) → 25.5	(-14,17) → 22
第四圈	(2889,-46) → 2889.4	(210,-137) → 250.7
第五圈	(-11,-6) → 12.5	(-9,12) → 15



说明：由于实际控制 Create2 不可能精确回到定位标记点，因此上表中的坐标均为减去 Create2 的实际偏差后得到。

从表 5-5 定位结果可知，在正常 VISLAM 系统定位结果中（第一/三/五圈），基于多传感器融合的紧耦合算法的定位精度要优于基于里程计的松耦合算法，且其最大定位偏差分别为 22mm 和 25.5mm；在纯积分版的定位结果中（第二圈），基于里程计的松耦合算法定位偏差是基于多传感器融合的紧耦合算法的 4 倍左右；而在车轮打滑的定位结果中（第四圈），该定位偏差几乎达到了 10 倍多，基于里程计的松耦合算法几乎处于失效状态，相反基于多传感器融合的紧耦合算法定位偏差为 250.7mm，在可接受的范围内。

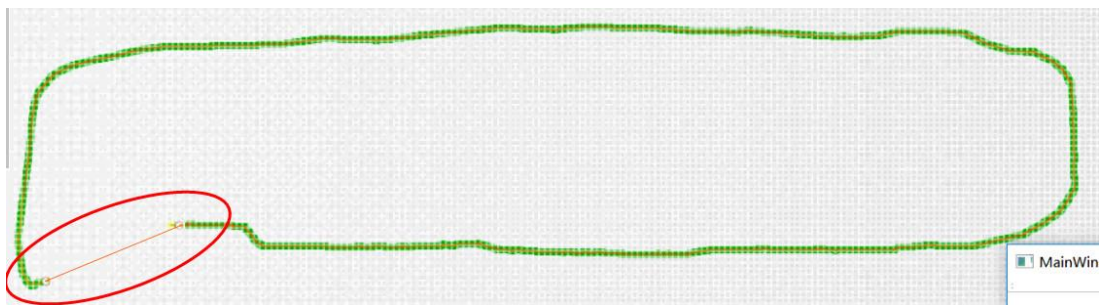
因此，通过上述小场景下的定位实验可知，本文提出的基于多传感器融合的紧耦合算法无论在视觉丢失以及车轮打滑的极端情况下，都能够给出比较满意的定位结果，满足扫地机器人的定位需求，解决了传统定位算法在失效情况下定位出错的问题，提升了用户体验。

## （2）大场景下的重定位实验

为了验证基于里程计的松耦合算法和基于多传感器融合的紧耦合算法的重定位效果，本实验在大型走廊场景下进行建图重定位，建图重定位过程如图 5-11 所示。



a) 基于里程计的松耦合算法重定位



b) 基于多传感器融合的紧耦合算法重定位

图 5-11 大场景下的建图重定位过程

图 5-11(a)和图 5-11(b)都出现了由于里程计或光流计的累积误差导致地图无法闭合的问题，而且图 5-11(a)中地图已经出现了严重的偏差，且构建的地图



无法用于后续的路径规划或导航等，相反图 5-11(b)中地图偏差为 2.5m，在一定程度上是可以接受的。由于重定位功能的存在使得两种情况下的定位点都能够找到起始点，使得地图闭合。因此，经过实验验证，本文所提出的基于多传感器融合的紧耦合算法在室内大场景下也具有较好的鲁棒性，能够满足扫地机器人的应用需求。

### 5.3.3 VISLAM 系统优化对比实验

#### (1) 初始化状态估计实验

本实验主要验证 4.3 小节提出的视觉惯导初始化状态估计算法的优化效果，由于本文结合里程计和光流计来进行初始化，较传统的仅采用 IMU 进行初始化能够得到更准确的尺度信息以及更优的收敛时间。尺度收敛时间及尺度值记录如表 5-6 所示。（注：当估计得到了一个尺度，且在后续的一个周期里估计的尺度变化不超过 5%，则认为得到了稳定的尺度，其对应的时间即为尺度收敛时间）

表 5-6 尺度收敛时间及尺度值

实验次数	收敛时间 (s)	尺度值
1	2.5	49.6
2	2.9	54.3
3	2.0	55.7
4	3.1	50.9
均值	2.625	52.625
均方根误差 (RMSE)	0.421	2.469

注：由于本文没有研究传统的 IMU 初始化方法，故无法得到对比实验数据。

从上表可看出，收敛时间均值为 2.625，RMSE 为 0.421，尺度均值为 52.625，RMSE 为 2.469，因此本文提出的视觉惯导初始化状态估计算法能够得到较稳定的、收敛快的尺度信息。

#### (2) 降帧率对比实验

前面的章节已经叙述或实验证实了由于里程计在不同材质表面上尺度不一致或者打滑导致机器人位姿估计不准确，那么 VISLAM 系统在 g2o 优化时所需要的时间或者优化得到正确位姿的成功率也不一样。因此，本实验将对基于里程计的松耦合算法和基于多传感器融合的紧耦合算法在不同帧率下的跟踪成功率。

由于前面的优化工作使得目前的 VISLAM 系统跟踪成功率都很高, 为了使实验结果更具差异化, 因此增加实验难度主动控制 Create2 平台沿实验室轮廓走一圈 (模拟扫地机器人真实的工作状况) 进行建图实验。本实验在两种算法下各进行 5FPS、10FPS 和 15FPS 的对比实验, 基于里程计的松耦合算法在不同帧率下的建图轨迹如图 5-12 所示, 基于多传感器融合的紧耦合算法在不同帧率下的建图轨迹如图 5-13 所示。具体跟踪成功率数据见表 5-7 所示, 并根据表 5-7 拟合跟踪成功率随帧率的变化曲线如图 5-14 所示。

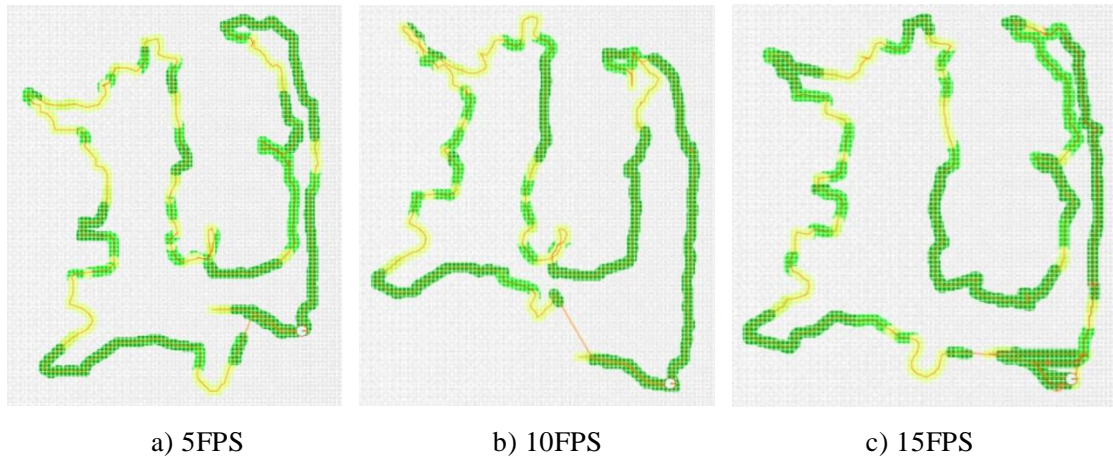


图 5-12 基于里程计的松耦合算法在不同帧率下的建图轨迹

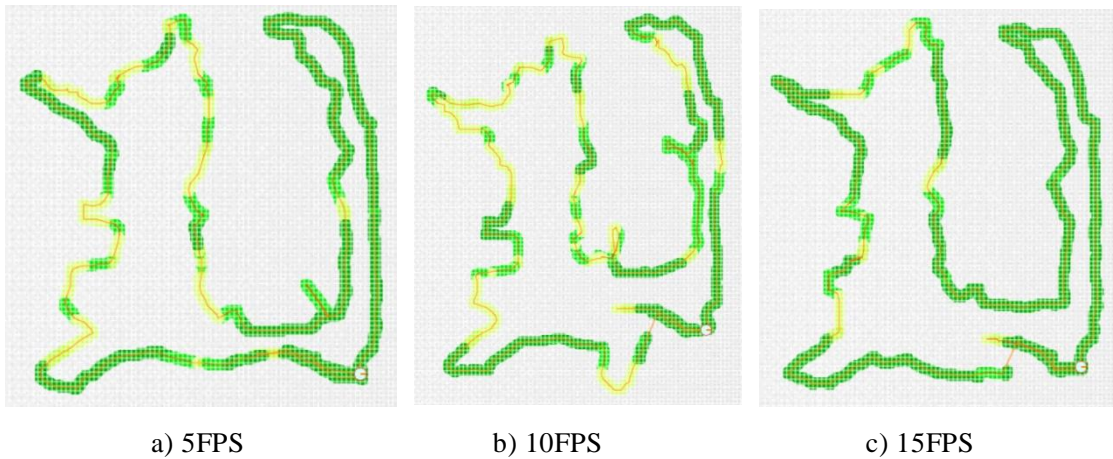


图 5-13 基于多传感器融合的紧耦合算法在不同帧率下的建图轨迹

从图 5-12 和图 5-13 可知, 随着帧率的增加, 跟踪成功率 (绿色线) 也越来越高。分析上述 6 种情况下的建图轮廓, 图 5-13 中 3 种轮廓图与真实实验室的轮廓一致, 地图的尺度更接近真实的尺度, 再一次说明了基于多传感器融合的紧耦合算法定位的累积误差要小于基于里程计的松耦合算法。

表 5-7 跟踪成功率数据记录

帧率	基于里程计的松耦合算法	基于多传感器融合的紧耦合算法
5FPS	6108/15564=39.2%	7293/12860=56.7%
10FPS	6834/14923=45.8%	9315/12751=73.1%
15FPS	7710/16191=47.6%	9693/12305=78.8%

注：6108/15564=39.2%表示在全部的 15564 跟踪帧中有 6108 帧跟踪成功。

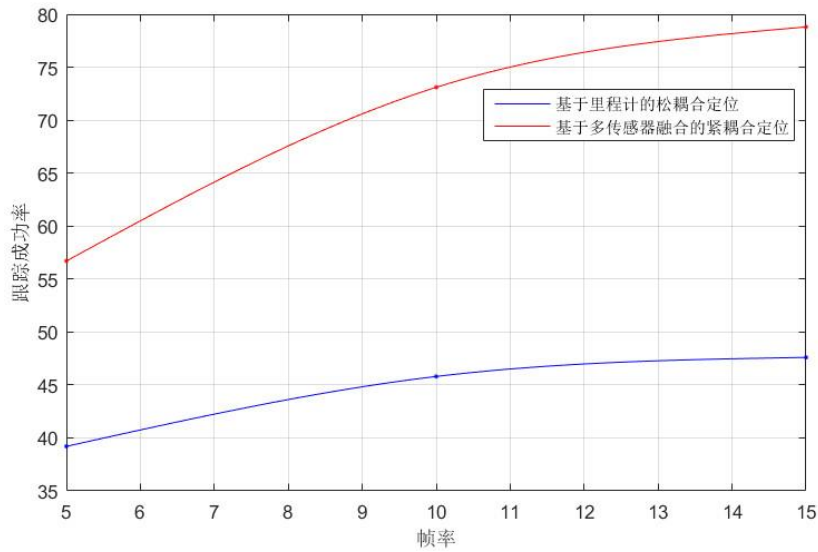


图 5-14 跟踪成功率随帧率的变化曲线

根据图 5-14 可知，在帧率为[5,15]范围内时，基于多传感器融合的紧耦合算法的跟踪成功率始终大于基于里程计的松耦合算法，因此在保证基于里程计的松耦合算法跟踪成功率的基础上，可以适当降低基于多传感器融合的紧耦合算法的帧率，节省处理器的计算资源。（说明：在扫地机器人正常工作过程中，上述两种算法的跟踪成功率都可以达到 80%以上（15FPS），此处只是为了增加难度，使得实验结果对比更明显）

## 5.4 本章小结

本章首先完成了 Create2 平台的硬件搭建和相应的软件设计，然后在不同材质表面进行了定位对比实验，进行了室内小场景和大场景下两种算法的定位精度对比实验，以及针对基于多传感器融合的紧耦合算法进行了 VISLAM 系统优化的对比实验。实验结果表明本文所提出的基于多传感器融合的紧耦合算法对于提升 VISLAM 系统的鲁棒性、地图的完整性都有很大的帮助，也证实了本

文所设计的扫地机器人能够适用于大部分的室内场景，且受表面材质、纹理丰富、光照等条件的影响极小。

## 结 论

扫地机器人作为一款已经或即将进入千家万户的服务机器人，在体力劳动方面为人们提供了极大的便利。研制具有更高清扫效率、覆盖率和智能型的扫地机器人是当前学术界和产业界研究的热点。本文主要研究了家庭场景下移动机器人的定位算法，主要研究成果如下：

(1) 改进了一种基于扩展卡尔曼滤波器对多传感器数据进行融合的方法，解决了单个传感器无法满足所有应用场景的问题。该方法建立了移动机器人的系统模型和观测模型，设计了适用于扫地机器人平台的扩展卡尔曼滤波器。验证了基于 EKF 融合较里程计+IMU 融合（或光流计+IMU 融合）能够在一定程度上克服尺度不一致的问题，提高室内移动机器人航位推算定位的精度和鲁棒性。

(2) 改进了一种对 DBoW2 词汇库进行优化的方法，解决了目前 ORB-SLAM2 中加载 DBoW2 词汇库耗时长和占用大量系统内存等问题，同时训练得到了针对家庭场景下的词汇库。另一方面，改进了一种平面移动机器人 3-DoF SLAM 优化算法，解决了三维空间 6-DoF 下优化结果误差大、实时性差等问题。验证了所提出的算法较优化前有显著的性能提升。

(3) 提出了一种融合多传感器的惯导预积分算法，解决了目前主流的预积分算法误差大、精度低等问题，同时利用松耦合算法估计了初始化视觉尺度  $s$ 。另一方面，提出了一种基于多传感器融合的紧耦合算法，解决了松耦合算法定位精度低等问题。

(4) 完成了 Create2 平台的硬件搭建和相应的软件设计，然后在不同材质表面进行了定位对比实验，进行了室内小场景和大场景下两种算法的定位精度对比实验，以及针对本文提出的紧耦合算法进行了 VISLAM 系统优化的对比实验。实验结果表明本文所提出的基于多传感器融合的紧耦合算法能够适用于大部分的室内场景，且受表面材质、纹理特征、光照等条件的影响小。

本文虽达到了预期目标，但仍有一些未解决的问题：

(1) 大量家庭场景下的测试。由于本文只在实验室或少量的家庭场景下测试过，还不能完全说明算法能够在家庭环境下鲁棒、稳定的长时间工作。

(2) VSLAM 算法方面。由于目前摄像头是顶视的方案，还无法完全发挥摄像头的优势（识别等应用），因此后续可以针对鲁棒、稳定的 VSLAM 算法进行前视方案的研究。

## 参考文献

- [1] 陈玉琨, 汤晓鸥. 人工智能基础[M]. 华东师范大学出版社. 第 1 版, 2018.
- [2] 中国老龄产业协会. 中国老龄产业发展报告[R]. 2018.
- [3] Cho M Y, Jeong Y S. Human gesture recognition performance evaluation for service robots[C]//Advanced Communication Technology (ICACT), 2017 19th International Conference on. IEEE, 2017: 847-851.
- [4] Zhaohui Z, Xuesong M, Xu B, et al. Development of an intelligent interaction service robot using ROS[C]//Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), 2016 IEEE. IEEE, 2016: 1738-1742.
- [5] Kuo C H, Chou H C, Tasi S Y. Pneumatic sensor: A complete coverage improvement approach for robotic cleaners[J]. IEEE Transactions on Instrumentation and Measurement, 2011, 60(4): 1237-1256.
- [6] Luo R C, Lai C C. Multisensor fusion-based concurrent environment mapping and moving object detection for intelligent service robotics[J]. IEEE transactions on industrial electronics, 2014, 61(8): 4043-4051.
- [7] 宋学滨. 聊聊扫地机器人的发展史 [EB/OL]. <https://www.iznb.cn/article/17834>, 2015-10-28.
- [8] Smith R C, Cheeseman P. On the representation and estimation of spatial uncertainty[J]. The international journal of Robotics Research, 1986, 5(4): 56-68.
- [9] Kalman R E. A new approach to linear filtering and prediction problems[J]. Journal of basic Engineering, 1960, 82(1): 35-45.
- [10] Jensfelt P, Kragic D, Folkesson J, et al. A framework for vision based bearing only 3D SLAM[C]//Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on. IEEE, 2006: 1944-1950.
- [11] 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.
- [12] Civera J, Davison A J, Montiel J M M. Inverse depth parametrization for monocular SLAM[J]. IEEE transactions on robotics, 2008, 24(5): 932-945.
- [13] Klein G, Murray D. Parallel tracking and mapping for small AR workspaces[C]//Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium on. IEEE, 2007: 225-234.
- [14] Kümmerle R, Grisetti G, Strasdat H, et al. g2o: A general framework for graph

- optimization[C]//Robotics and Automation (ICRA), 2011 IEEE International Conference on. IEEE, 2011: 3607-3613.
- [15] Strasdat H, Davison A J, Montiel J M M, et al. Double window optimisation for constant time visual SLAM[C]//Computer Vision (ICCV), 2011 IEEE International Conference on. IEEE, 2011: 2352-2359.
- [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] Gálvez-López D, Tardos J D. Bags of binary words for fast place recognition in image sequences[J]. IEEE Transactions on Robotics, 2012, 28(5): 1188-1197.
- [18] Liu H, Zhang G, Bao H. Robust keyframe-based monocular SLAM for augmented reality[C]//Mixed and Augmented Reality (ISMAR), 2016 IEEE International Symposium on. IEEE, 2016: 1-10.
- [19] Gao X, Zhang T. Robust RGB-D simultaneous localization and mapping using planar point features[J]. Robotics and Autonomous Systems, 2015, 72: 1-14.
- [20] Besl P J, McKay N D. Method for registration of 3-D shapes[C]//Sensor Fusion IV: Control Paradigms and Data Structures. International Society for Optics and Photonics, 1992, 1611: 586-607.
- [21] Khaleghi B, Khamis A, Karray F O, et al. Multisensor data fusion: A review of the state-of-the-art[J]. Information Fusion, 2013, 14(1): 28-44.
- [22] 张文. 基于多传感器融合的室内机器人自主导航方法研究[D]. 中国科学技术大学, 2017.
- [23] 吴显. 基于多传感器信息融合的移动机器人定位方法研究[D]. 北京交通大学, 2016.
- [24] Lynen S, Achtelik M W, Weiss S, et al. A robust and modular multi-sensor fusion approach applied to mav navigation[C]//Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on. IEEE, 2013: 3923-3929.
- [25] 李良群, 姬红兵, 罗军辉. 迭代扩展卡尔曼粒子滤波器[J]. 西安电子科技大学学报, 2007.
- [26] Quan M, Piao S, Tan M, et al. Tightly-coupled Monocular Visual-odometric SLAM using Wheels and a MEMS Gyroscope[J]. arXiv preprint arXiv:1804.04854, 2018.
- [27] Paul M K, Wu K, Hesch J A, et al. A comparative analysis of tightly-coupled monocular, binocular, and stereo VINS[C]//Robotics and Automation (ICRA), 2017 IEEE International Conference on. IEEE, 2017: 165-172.
- [28] Hartley R, Zisserman A. Multiple view geometry in computer vision[M]. Cambridge university press, 2003.

- 
- [29] Corke P, Lobo J, Dias J. An introduction to inertial and visual sensing[J]. 2007.
- [30] Hu J S, Tseng C Y, Chen M Y, et al. IMU-assisted monocular visual odometry including the human walking model for wearable applications[C]//Robotics and Automation (ICRA), 2013 IEEE International Conference on. IEEE, 2013: 2894-2899.
- [31] Weiss S, Siegwart R. Real-time metric state estimation for modular vision-inertial systems[C]//Robotics and Automation (ICRA), 2011 IEEE International Conference on. IEEE, 2011: 4531-4537.
- [32] Piniés P, Lupton T, Sukkariéh S, et al. Inertial aiding of inverse depth SLAM using a monocular camera[C]//Robotics and Automation, 2007 IEEE International Conference on. IEEE, 2007: 2797-2802.
- [33] Mourikis A I, Roumeliotis S I. A multi-state constraint Kalman filter for vision-aided inertial navigation[C]//Robotics and automation, 2007 IEEE international conference on. IEEE, 2007: 3565-3572.
- [34] Nilsson J O, Zachariah D, Jansson M, et al. Realtime implementation of visual-aided inertial navigation using epipolar constraints[C]//Position Location and Navigation Symposium (PLANS), 2012 IEEE/ION. IEEE, 2012: 711-718.
- [35] Asadi E, Bottasso C L. Tightly-coupled vision-aided inertial navigation via trifocal constraints[C]//Robotics and Biomimetics (ROBIO), 2012 IEEE International Conference on. IEEE, 2012: 85-90.
- [36] Indelman V, Gurfil P, Rivlin E, et al. Real-time vision-aided localization and navigation based on three-view geometry[J]. IEEE Transactions on Aerospace and Electronic Systems, 2012, 48(3): 2239-2259.
- [37] Kneip L, Weiss S, Siegwart R. Deterministic initialization of metric state estimation filters for loosely-coupled monocular vision-inertial systems[C]//Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on. IEEE, 2011: 2235-2241.
- [38] Bloesch M, Burri M, Omari S, et al. Iterated extended Kalman filter based visual-inertial odometry using direct photometric feedback[J]. The International Journal of Robotics Research, 2017, 36(10): 1053-1072.
- [39] Leutenegger S, Lynen S, Bosse M, et al. Keyframe-based visual-inertial odometry using nonlinear optimization[J]. The International Journal of Robotics Research, 2015, 34(3): 314-334.
- [40] Forster C, Carlone L, Dellaert F, et al. On-Manifold Preintegration for Real-Time Visual-Inertial Odometry[J]. IEEE Transactions on Robotics, 2017, 33(1): 1-21.



- 
- [41] Jones E S, Soatto S. Visual-inertial navigation, mapping and localization: A scalable real-time causal approach[J]. The International Journal of Robotics Research, 2011, 30(4): 407-430.
- [42] Bloesch M, Omari S, Hutter M, et al. Robust visual inertial odometry using a direct EKF-based approach[C]//Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on. IEEE, 2015: 298-304.
- [43] Shen S, Michael N, Kumar V. Tightly-coupled monocular visual-inertial fusion for autonomous flight of rotorcraft MAVs[C]//Robotics and Automation (ICRA), 2015 IEEE International Conference on. IEEE, 2015: 5303-5310.
- [44] Mur-Artal R, Tardós J D. Visual-inertial monocular SLAM with map reuse[J]. IEEE Robotics and Automation Letters, 2017, 2(2): 796-803.
- [45] Lupton T, Sukkarieh S. Visual-inertial-aided navigation for high-dynamic motion in built environments without initial conditions[J]. IEEE Transactions on Robotics, 2012, 28(1): 61-76.
- [46] Barfoot, Timothy D. State Estimation for Robotics. Cambridge University Press, 2017.
- [47] Liu Y, Chen Z, Zheng W, et al. Monocular Visual-Inertial SLAM: Continuous Preintegration and Reliable Initialization[J]. Sensors, 2017, 17(11): 2613.
- [48] Li P, Qin T, Hu B, et al. Monocular visual-inertial state estimation for mobile augmented reality[C]//2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR). IEEE, 2017: 11-21.
- [49] Qin T, Shen S. Robust initialization of monocular visual-inertial estimation on aerial robots[C]//Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Vancouver, BC, Canada. 2017: 24-28.
- [50] Yang Z, Shen S. Monocular visual-inertial state estimation with online initialization and camera-imu extrinsic calibration[J]. IEEE Transactions on Automation Science and Engineering, 2017, 14(1): 39-51.
- [51] Feng-Ji Z, Hai-Jiao G, Abe K. A mobile robot localization using ultrasonic sensors in indoor environment[C]//Robot and Human Communication, 1997. RO-MAN'97. Proceedings., 6th IEEE International Workshop on. IEEE, 1997: 52-57.
- [52] Mueller T J. Infrared location system: U.S. Patent 5,641,963[P]. 1997-6-24.
- [53] Shin B J, Lee K W, Choi S H, et al. Indoor WiFi positioning system for Android-based smartphone[C]//Information and Communication Technology Convergence (ICTC), 2010 International Conference on. IEEE, 2010: 319-320.
- [54] 丁锐, 钱志鸿, 王雪. 基于 TOA 和 DOA 联合估计的 UWB 定位方法[J]. 电子与信息学报, 2010, 32(2): 313-317.

- [55] Px4flow developer guide[EB/OL]. <https://pixhawk.org/dev/px4flow>. 2016-03-16.
- [56] Forster C, Zhang Z, Gassner M, et al. Svo: Semidirect visual odometry for monocular and multicamera systems[J]. IEEE Transactions on Robotics, 2017, 33(2): 249-265.

## 哈尔滨工业大学学位论文原创性声明和使用权限

### 学位论文原创性声明

本人郑重声明：此处所提交的学位论文《基于多传感器融合的室内移动机器人定位算法研究》，是本人在导师指导下，在哈尔滨工业大学攻读学位期间独立进行研究工作所取得的成果，且学位论文中除已标注引用文献的部分外不包含他人完成或已发表的研究成果。对本学位论文的研究工作做出重要贡献的个人和集体，均已在文中以明确方式注明。

作者签名： 李巍 日期：2018年6月27日

### 学位论文使用权限

学位论文是研究生在哈尔滨工业大学攻读学位期间完成的成果，知识产权归属哈尔滨工业大学。学位论文的使用权限如下：

(1) 学校可以采用影印、缩印或其他复制手段保存研究生上交的学位论文，并向国家图书馆报送学位论文；(2) 学校可以将学位论文部分或全部内容编入有关数据库进行检索和提供相应阅览服务；(3) 研究生毕业后发表与此学位论文研究成果相关的学术论文和其他成果时，应征得导师同意，且第一署名单位为哈尔滨工业大学。

保密论文在保密期内遵守有关保密规定，解密后适用于此使用权限规定。本人知悉学位论文的使用权限，并将遵守有关规定。

作者签名： 李巍 日期：2018年6月27日

导师签名： 孙立宁 日期：2018年6月27日

## 致 谢

“实践出真知”只有在实践中，才能真正得到真理。只有坐在同一条椅子上，才能缩短、拉近心与心的距离。只有面对面的交流，才能挖掘内心的真实想法。只有处于同一个屋檐下、身处同一个岗位上才能感受到真实的生活。

在哈尔滨工业大学学习的时间里，一直秉承着“规格严格，功夫到家”的校训，时刻牢记着自己是哈工大人的使命。在科研道路上，时常以实事求是、求真务实的心态砥砺前行，为此感谢哈工大提供了一个学习和进步的平台。

首先要感谢我的导师孙立宁教授，孙老师渊博的学识，开阔的眼界，严谨的科研态度时刻影响着我。感谢祝宇虹副教授，在我的研究生科研道路上给予了极大地支持，从课题选题、方案论证，到中期答辩，以及后期的毕业论文撰写都给予了悉心的指导和建议。还要感谢实验室孔民秀副教授在课题方面的指导和生活方面的关怀。

感谢实验室博士季晨师兄，上届的袁源强、赵兴平、宋博涵、荆鑫、徐鑫炎、陈松赞和李明师兄在开题过程中宝贵的指导意见，保证了课题后续的顺利开展。

感谢实验室同届的夏壮、韩博、白玉昊和宇文亮在科研学习中能够相互取长补短、互帮互助的精神，在日常生活中也给予我很大的帮助和支持。

感谢实验室师弟童宁、田宇、牛草源、沈锦华、徐胜润在实验过程中的帮助与支持。希望他们今后在科研道路上能够取得更好的成绩。

感谢我的父母，你们是我一生中最大的财富，你们的支持使我能够扬帆远航、追逐梦想。

最后，再次向在我求学路上所有关心和帮助过我的师长、同学和朋友们，表示最诚挚的感谢！

李 巍

2018年6月于哈工大机器人所