

# 无人艇同步定位与建图技术探究

于海深, 张会霞\*, 陶亚东

江苏海洋大学, 江苏 连云港

收稿日期: 2022年5月29日; 录用日期: 2022年6月21日; 发布日期: 2022年6月30日

## 摘要

同步定位与建图技术一直以来都是无人艇导航的一个热门研究方向, 通过使用相机来获得周围环境信息以及定位自身位置。本文罗列了近年来出色的视觉SLAM算法和视觉惯导SLAM算法, 系统的梳理了算法的优势与不足, 同时归纳了SLAM算法的最新研究进展。最后, 探讨了SLAM技术的发展趋势。

## 关键词

水面无人艇, 视觉SLAM算法, 视觉惯性SLAM算法

# Exploration of Simultaneous Localization and Mapping Technology for Surface Unmanned Vehicles

Haishen Yu, Huixia Zhang\*, Yadong Tao

School of Ocean Engineering, Lianyungang Jiangsu

Received: May 29<sup>th</sup>, 2022; accepted: Jun. 21<sup>st</sup>, 2022; published: Jun. 30<sup>th</sup>, 2022

## Abstract

Simultaneous localization and mapping technology has been a popular research direction for unmanned boat navigation, which uses cameras to obtain information about the surrounding environment and locate its own position. This paper lists the outstanding visual SLAM algorithms and visual inertial guidance SLAM algorithms in recent years, and systematically composes the advantages and shortcomings of the algorithm, and summarises the latest research progress of SLAM algorithms. Finally, the development trend of SLAM technology is discussed.

\*通讯作者。

## Keywords

Surface Unmanned Vehicle, Visual SLAM, Visual Inertial SLAM

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

随着人工智能的快速发展,在海洋科学等相关学科的研究中水面无人艇(Unmanned Surface Vehicle, USV)发挥着越来越重要的作用。无人艇不仅拥有自主运动控制和对复杂环境适应性强,而且可以进行目标识别和跟踪。在复杂的水面环境中,无人艇安全运行的关键是能否进行准确的定位与导航。解决这一问题的关键就是无人艇同步定位与建图(Simultaneous Localization and Mapping, SLAM)技术。视觉 SLAM 技术为无人艇在未知的水面环境下进行自主导航和环境探测提供了可行性方案[1]。

## 2. 视觉 SLAM 框架

经典视觉 SLAM 的框架如图 1 所示,由传感器数据、视觉里程计(Visual Odometry, VO)、后端优化(Optimization)、回环检测(Loop Closing)以及建图(Mapping)构成。

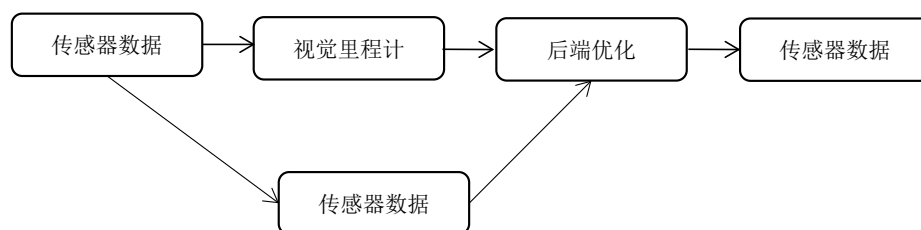


Figure 1. Classical visual SLAM flow

图 1. 经典视觉 SLAM 流程

### 1) 传感器数据

视觉 SLAM 中用于获取数据使用的相机主要分为:单目、双目、深度相机。如表 1 所示。

Table 1. Analysis of the characteristics of different vision sensors

表 1. 不同视觉传感器特点分析

| 相机 | 优点           | 缺点     |
|----|--------------|--------|
| 单目 | 成本低, 不受环境影响  | 尺度不确定  |
| 双目 | 通过左右视差得到深度信息 | 配置复杂   |
| 深度 | 可以直接获得深度信息   | 受光线影响大 |

### 2) 视觉里程计

在视觉 SLAM 中,前端是通过图像之间的位姿变换获得相机运动的过程。VO 的实现方法,按是否需要提取特征分为直接法前端和特征点法前端。

### 3) 后端优化

后端主要是将前端传输过来的结果进行优化, 利用扩展的卡尔曼滤波(extended Kalman filter, EKF)、粒子滤波(PF)等滤波理论或者优化理论进行树或图的优化, 最终获得较优的位姿估算。

### 4) 回环检测

回环检测是指根据传感器信息判断机器人在之前有没有访问过某个地方, 用于校正 VO 漂移产生的偏差, 从而获得信息一致的地图。

### 5) 建图

对于建图来说, 没有固定的算法和形式。侧重于定位时, 主要使用稀疏的路标地图; 侧重于真实场景建模时, 需要对 3D 地图进一步网格化。

## 3. 视觉 SLAM 算法

### 3.1. PTAM 算法

G. Klein 等在 2007 年提出了并行跟踪与建图技术(Parallel Tracking and Mapping, PTAM) [2], 首次区分了前端和后端, 后端优化使用非线性优化的方法。该技术将跟踪和建图分为两个独立的任务, 并在并行的线程中分别进行处理。在跟踪的线程中, 对运动的相机进行位姿估计; 在建图的线程中, 从先前观察到的视频帧中产生一个点图像的三维地图。

由于跟踪和建图分为两个独立的任务, 并在两个独立的线程上并行运行, 这样能够在不影响相机跟踪实时性的情况下, 提高相机跟踪算法的鲁棒性和精确性, 同时减少系统处理每帧图像的时间。但主要为了小场景设计的, 导致场景局限; 在相机快速运动时跟踪容易丢失。

针对使用特征点法的机器人定位容易丢失等问题, 刘贵涛等提出了一种基于多个双目相机紧耦合的视觉 SLAM 算法[3], 通过将多个多目相机的测量信息融合在一个统一的最小二乘法中来完成对自身的定位, 有效解决了局部弱纹理环境下机器人定位丢失的问题。

针对算法的实时性问题, 汪泽睿等在回环检测的过程中, 先通过计算图像之间相同单词个数来去除掉与新图像差异度较大的历史图像, 然后再进行图像查找的后续流程, 提高回环测算法的实时性[4]。

### 3.2. LSD-SLAM 算法

2014 年, Engel J 等提出的 LSD-SLAM (large-scale direct SLAM) [5], 该算法的提出标志着单目直接法 SLAM 的成功问世。LSD-SLAM 主要分为图像位姿跟踪、深度估计以及地图优化三大模块, 其中地图构建为该算法的主要模块; 该算法通过建立大尺度直接单目 SLAM 的框架, 提出一种用来直接估计关键帧之间相似变换、尺度感知的图像匹配算法, 在 CPU 上实现了半稠密场景的重建。

针对相机快速运动时跟踪容易丢失问题, 李攀伟等提出了一种结合特征关键子和直接灰度匹配方法前端视觉里程计信息融合方法, 使在快速运动的情况下跟踪目标不容易丢失, 以取得稳定的单目视觉定位效果[6]。王英格等通过将线特征加入到 LSD-SLAM 视觉里程计中, 提高了相机位姿估计的鲁棒性和准确性[7]。

### 3.3. ORBSLAM 算法

Mur-Artal 等提出了 ORB-SLAM 算法[8], 该算法基于 PTAM 的思想, 在 PTAM 的算法框架的基础上进行了优化和改进, 使用了 3 个独立并行的线程进行追踪、地图构建和回环优化[8]。ORBSLAM 整个算法过程基于 ORB (Oriented FAST and Rotated BRIEF)特征点, 具有较高的特征点利用率, 提高了系统的实时性。另外, ORBSLAM 具有地图自动初始化, 全局重定位功能, 并能实时对每帧图像进行回环检测

和优化,提高了系统的定位精度。经过 Mur-Artal 等人的不断优化,2016 年提出了 ORB-SLAM2 [9],它支持单目、双目和 RGB-D 三种相机。

该算法具有相对完善的功能和框架,可用于室内外场景以及大小不同的场景,有着优越的鲁棒性。但是,ORB-SLAM 使用了大量的特征点,计算这些特征点需要花费大量时间;另外构建的地图为稀疏地图,没有办法满足路径规划的需求。

针对 ORBSLAM 算法特征点问题,陶卓等采用 FAST 算法对特征点进行检测,同时利用点特征提取算子去除特征点中的伪角点,提高了特征点的质量[10]。赵宇琦等在原始 ORB 算法的基础上采用快速自适应图像金字塔,三维描述子质心法与二叉树特征管理算法等方法进行改进,改进后的算法提取的图像特征点鲁棒性更强[11]。秦绪佳等通过对提取周边点的方式和判断特征点条件进行了改进,使算法能够更快速准确的提取到图像上的特征点,减少了噪点的数目,提高了特征点的检测效率[12]。

## 4. 视觉惯导 SLAM 算法

纯视觉的 SLAM 在自身及周围环境发生剧烈变化的时候会受到较大的影响,而惯性导航测量单元(IMU, Inertial Measurement Unit)可以实时记录来自加速度计和陀螺仪的信息,在视觉信息失真的时候,可以通过短时积分以此来提供相对准确的自身运动信息。在一定程度上,相机和 IMU 这两个传感器能够互补,相机数据可以有效修正 IMU 的漂移,而 IMU 可以降低动态环境对相机的负面影响,视觉信息和惯导信息融合使用可以取得较为准确的定位结果。

### 4.1. MSCKF 算法

2007 年 Mourikis 等人提出以多状态约束卡尔曼滤波的 SLAM 算法(MSCKF, Multi-state Constraint Kalman Filter) [13] [14], MSCKF 采用滑动窗口(sliding window)的形式,通过连续时间段内的相机位姿状态建立非线性三角化约束,实现滤波更新和状态估计,大大提高了系统的鲁棒性和收敛性。

针对 MSCKF 的初始化阶段的缺陷,赵忠堂等采用 Sig-ma 滤波器和三焦点张量约束快速完成系统初始化,所生成的状态向量与 MSCKF 一致,从而实现了系统初始化阶段和后续导航之间的无缝衔接,提高了系统状态估计的鲁棒性和精度[15]。

针对传统 MSCKF 算法在光线不好的情况下造成滤波器估计准确度和稳定性降低,王磊等融合了交互式多模型(IMM)估计和多状态约束卡尔曼滤波(MSCKF)算法,提出了一种 IMM-MSCKF 算法,算法以 MSCKF 为模型匹配子滤波器,融合各子滤波器的输入、输出,大大提高了系统的鲁棒性和精确性[16]。

### 4.2. OKVIS 算法

2015 年,Stefan 等人提出了基于非线性优化的后端优化与图像中关键帧使用的视觉惯导 SLAM 算法(OKVIS, Open Keyframe-Based Visual-Inertial SLAM) [17],该算法是通过将视觉信息和惯导信息紧耦合的方式进行融合,后端通过优化的方法实现位姿准确估算。该方法利用视觉测量误差与惯导测量残差建立代价函数,同时对视觉信息和惯导信息进行优化,进而得到相对准确的位姿估算。同时优化视觉信息和惯导信息降低了系统的数据计算量,保证了算法的精确性和实时性。

针对视觉惯性里程计在处理图像信息的时候,特征点正确的匹配率较低问题,刘钊铭等通过结合图像的梯度信息提高实际纹理不同区域的描述子间的差异,并利用特征点的像素坐标信息,对相邻时刻采集到的图像间的待匹配特征点对进行筛选,提高了特征点正确匹配率[18]。

黄坤等对 OKVIS 算法进行改进,提出一种 MyVIO 跟踪注册算法[19]。把 OKVIS 算法使用的 BRISK 描述符算法替换为 MyFREAK 描述符算法,对图像数据进行处理,提升图像匹配的实时性和鲁棒性;还增加一种视觉惯性初始化过程,提升了算法整体的鲁棒性。

### 4.3. VINS-MONO 算法

香港科技大学 Tong Qin 等学者提出了一种基于优化的紧耦合视觉惯导 SLAM 算法(Monocular Visual-Inertial Systems, VINS-MONO) [20] [21], VINS-Mono 是 VINS 开源的单目视觉惯性 SLAM 方案, 是基于滑动窗口优化实现的 VIO, 使用 IMU 预积分构建紧耦合框架, 具有自动初始化、在线外参标定、重定位、闭环检测以及全局位姿图优化功能的一套完整的 SLAM 系统。该算法有效利用了视觉信息和 IMU 信息, 提高了算法的实时性, 在实际运行过程中具有良好的表现。

针对 VINS-MONO 的视觉导航系统初始化的性能不太理想, 徐晓苏等提出了一种基于 IMU 预积分封闭解的初始化算法, 该方法利用多约束的卡尔曼滤波器融合了 IMU 数据和相机数据, 提高了原来系统的定位准确度[22]。

针对 VINS-MONO 在水下环境中回环检测效果不太理想, 进而导致定位精度较差。赵恒飞等提出了一种基于几何约束和 ORB 字典的水下回环识别算法。将原来 VINS-MONO 使用的 FAST 角点替换成 ORB 特征点, 相对于 FAST 角点, ORB 特征点在水下环境中能提供大量正确的回环候选帧, 提高回环检测效果以及定位精度[23]。

## 5. 结论

应用视觉 SLAM 算法在水面环境中对无人船进行实时定位和建图, 纯视觉的 MONOSLAM、PTAM 和 ORB-SLAM/ORB-SLAM 算法仅使用高清相机来获取周围环境信息对无人船进行位姿估算, 但高清相机获取的图片信息中包括水面的动态区域部分, 会使图片中检测到的有效特征点不足的情况, 导致无法准确估计位姿信息, 因此纯视觉 SLAM 应用在水面环境下略有不足。相比于纯视觉 SLAM, 视觉惯导信息融合的 MSCKF、OKVIS 和 VINS-MONO 算法能够弥补纯视觉 SLAM 算法的不足。MSCKF 算法融合双目视觉和惯导信息进行位姿估算, 但在水面环境中要对图像进行再处理, 算法整体计算量太大, 实时性表现不好。OKVIS 算法没有闭环检测模块, 在大尺度环境下不能位姿校正, 因此不适合大尺度水面环境。VINS-MONO 算法具备相对完善的系统架构, 有较好的定位精度, 大尺度水面环境下有较稳定的表现。

综合近年来 SLAM 算法的研究成果, 如表 2 所示, 无人艇在 SLAM 领域正不断向鲁棒性、实时性和准确性方向发展, 未来可从以下方面展开探索与研究。

**Table 2.** Comparison table of visual SLAM algorithms

**表 2.** 视觉 SLAM 算法对比表

| 视觉 SLAM            | 类型       | 算法内容                            | 传感器      |
|--------------------|----------|---------------------------------|----------|
| PTAM               | 优化       | 后端首次使用非线性优化, 并行线程以及关键帧机制        | 单目       |
| LSD-SLAM           | 优化       | 将直接法用于半稠密的单目 SLAM 中, 实现半稠密场景的重建 | 单目       |
| ORB-SLAM/ORB-SLAM2 | 优化       | ORB 特征, 增加回环检测, 多线程             | 单目、双目、深度 |
| MSCKF              | 紧耦合和 EKF | 首次使用 IMU 和视觉融合, 后端滑动窗口法优化       | 双目 + IMU |
| OKVIS              | 紧耦合和优化   | 直接优化处理, 滑动窗口优化                  | 单目 + IMU |
| VINS-MONO          | 紧耦合和优化   | IMU 预积分, 滑动窗口优化                 | 单目 + IMU |



- 1) 考虑多种传感器融合, 如将激光雷达、GPS、相机以及 IMU 等多传感器进行融合, 实现对环境的感知。
- 2) 深度学习在传统的图像领域表现良好, 将 slam 和深度学习融合来完成一些更复杂的任务。

## 基金项目

2022 年研究生科研与实践创新计划项目支持

## 参考文献

- [1] Lee, D., Kim, D., Lee, S., Myung, H. and Choi, H.-T. (2013) Experiments on Localization of an AUV Using Graph-Based SLAM. 2013 10th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI), Jeju, 30 October-2 November 2013, 526-527. <https://doi.org/10.1109/URAI.2013.6677329>
- [2] Klein, G. and Murray, D. (2007) Parallel Tracking and Mapping for Small AR Workspaces. *Proceedings of the 6th IEEE/ACM International Symposium on Mixed and Augmented Reality*, Nara, 13-16 November 2007, 225-234. <https://doi.org/10.1109/ISMAR.2007.4538852>
- [3] 刘贵涛, 张雷, 徐方. 面向局部弱纹理环境的多双目视觉 SLAM[J]. 组合机床与自动化加工技术, 2022(3): 116-119+125.
- [4] 汪泽睿. 基于改进型词袋模型的回环检测算法研究[D]: [硕士学位论文]. 成都: 成都理工大学, 2020.
- [5] Engel, J., Schöps, T. and Cremers, D. (2014) LSD-SLAM: Large-Scale Direct Monocular SLAM. *European Conference on Computer Vision*, Zurich, 6-12 September 2014, 834-849. [https://doi.org/10.1007/978-3-319-10605-2\\_54](https://doi.org/10.1007/978-3-319-10605-2_54)
- [6] 李攀伟. 关于单目视觉实时定位与建图中的优化算法研究[D]: [硕士学位论文]. 深圳: 中国科学院大学(中国科学院深圳先进技术研究院), 2020.
- [7] 王英格. 视觉里程计在移动增强现实中的应用研究[D]: [硕士学位论文]. 北京: 北京工业大学, 2019.
- [8] Mur-Artal, R., Montiel, J.M.M. and Tardós, J.D. (2017) ORB-SLAM: A Versatile and Accurate Monocular SLAM System. *IEEE Transactions on Robotics*, **31**, 1147-1163. <https://doi.org/10.1109/TRO.2015.2463671>
- [9] Mur-Artal, R. and Tardós, J.D. (2017) ORB-SLAM2: an Open-Source SLAM System for Monocular, Stereo and RGB-D Cameras. *IEEE Transactions on Robotics*, **33**, 1255-1262. <https://doi.org/10.1109/TRO.2017.2705103>
- [10] 陶卓, 黄卫华, 章政, 姚艺, 何佳乐. 基于稀疏光流法的改进ORB特征匹配算法[J]. 组合机床与自动化加工技术, 2021(12): 1-5.
- [11] 赵宇琦, 连晓峰, 谭励, 刘栋. 一种均匀抗噪自适应的ORB特征提取算法[J]. 计算机仿真, 2022, 39(4): 321-327.
- [12] 秦绪佳, 陈国富, 王洋洋, 张美玉. 改进ORB特征点检测的图像拼接方法[J]. 小型微型计算机系统, 2022, 1(1): 98-102.
- [13] Mourikis, A.I. and Roumeliotis, S.I. (2007) A Multi-State Constraint Kalman Filter for Vision-Aided Inertial Navigation. *Proceedings of 2007 IEEE International Conference on Robotics & Automation*, Vol. 2, Rome, 10-14 April 2007, 3565-3572. <https://doi.org/10.1109/ROBOT.2007.364024>
- [14] Li, M. and Mourikis, A.I. (2013) High-Precision, Consistent EKF-Based Visual-Inertial Odometry. *International Journal of Robotics Research*, **32**, 690-711. <https://doi.org/10.1177/0278364913481251>
- [15] 赵中堂, 吴庆涛. 基于多状态卡尔曼滤波的双目视觉导航设计[J]. 兵器装备工程学报, 2020, 41(6): 169-173.
- [16] 王磊, 程向红, 李进, 王乐. 基于交互式多模型 MSCKF 的双目视觉/惯性里程计算法[J]. 中国惯性技术学报, 2021, 29(2): 221-228+236.
- [17] Leutenegger, S., Lynen, S., Bosse, M., Siegwart, R. and Furgale, P. (2015) Keyframe-Based Visual-Inertial Odometry Using Nonlinear Optimization. *International Journal of Robotics Research*, **34**, 314-334. <https://doi.org/10.1177/0278364914554813>
- [18] 刘铠铭. 基于SIFT特征的视觉惯性里程计研究与应用[D]: [硕士学位论文]. 西安: 西安科技大学, 2021.
- [19] 黄坤. 基于移动平台的增强现实跟踪注册技术研究[D]: [硕士学位论文]. 无锡: 江南大学, 2020.
- [20] Qin, T., Li, P. and Shen, S. (2018) VINS-Mono: A Robust and Versatile Monocular Visual-Inertial State Estimator. *IEEE Transactions on Robotics*, **34**, 1004-1020. <https://doi.org/10.1109/TRO.2018.2853729>
- [21] Qin, T. and Shen, S. (2018) Online Temporal Calibration for Monocular Visual-Inertial Systems. *International Conference on Intelligent Robots and Systems*, Madrid, 1-5 October 2018, 3662-3669. <https://doi.org/10.1109/IROS.2018.8593603>

- [22] 徐晓苏, 吴贤. 基于 IMU 预积分封闭解的单目视觉惯性里程计算法[J]. 中国惯性技术学, 2020, 28(4): 440-447.
- [23] 赵恒飞, 刘妹琴, 郑荣濠. 基于改进回环检测的 VINS-MONO 水下 SLAM 算法研究[D]: [硕士学位论文]. 杭州: 浙江大学, 2021.