



北京化工大学

BEIJING UNIVERSITY OF CHEMICAL TECHNOLOGY

COMPUTING METHODS

---

## 自动化科学导论课程报告

---

李昊

### 目录

第 1 章 自动化大类各专业之间的相互关系	1
第 2 章 学习的总结，体会和感受	1
2.1 学习的总结	1
2.2 学习的体会	1
2.3 学习的感受	1
第 3 章 相关专业方向：同步定位与地图构建 (SLAM)	1
3.1 AGV 的导航 (Navigation)	1
3.2 SLAM 定义	2
3.3 V-SLAM 的框架与分类	2
3.4 经典 V-SLAM 方法	3
3.4.1 稀疏法:ORB-SLAM2	3
3.4.2 多传感器融合:VINS-Mono	4

3.4.3	blablabla . . . . .	5
3.5	基于深度学习的 V-SLAM 方法 . . . . .	6
	参考文献	6

创建日期：2020 年 3 月 25 日  
更新日期：2020 年 5 月 25 日

## 第 1 章 自动化大类各专业之间的相互关系

其实吧，说专业差距大，其实也没那么大。举图像识别做例子：大家都做图像识别，可能在研究生阶段区别比较大，但是对于本科来说差别真的几乎没有。如果一定要做区分的话

- 自动化：偏向系统集成，控制算法，感知算法
- 测控：就我们学校，MEMS 无疑
- 通信：各种雷达信号处理

## 第 2 章 学习的总结，体会和感受

### 2.1 学习的总结

自动化学科是一门多学科交叉的高技术学科。研究内容广，几乎无所不包。但是，控制、信息、系统是其核心。自动化学科知识体系中的核心知识是：控制与智能、对象和建模、系统和工程等三个知识领域，也是自动化学科与其他学科的最重大区别。

### 2.2 学习的体会

选自动化就对了！选自动化就对了！选自动化就对了！千万不要搞过程控制！千万不要搞过程控制！千万不要搞过程控制！更加坚定了自己选择自动化和进行移动机器人以及机器人感知相关方向研究的意志。

### 2.3 学习的感受

与其说选择专业，我觉得更应该选择研究方向后再决定专业，因为每个专业的研究方向都有很大的重叠部分。

## 第 3 章 相关专业方向：同步定位与地图构建 (SLAM)

### 3.1 AGV 的导航 (Navigation)

导航技术是 AGV 技术的研究核心，没有导航功能的 AGV 系统是很难想象的。举个例子，如果 AGV 系统想要在工厂等工作场所运输货物，它必须要在不触碰到任何障碍物的情况下运动到不同的地点。随着信息化技术的提升，CPU 算力的提升和传感器的升级，AGV 的柔性路径导航吸引了越来越多的研究者进行研究，产业化应用越来越

越广。柔性路径导航技术应用较多的是视觉与激光导航、惯性 (IMU) 导航 [1]、基于 WIFI/UWB 等无线信号定位 [2] 的导航技术等。机器人实现其自主导航的前提是机器人自身位姿的确定，出于对定位精度的要求以及现实环境的复杂化，目前大多数导航系统都是由上述两种或者多种导航方式的结合组合导航技术 [3]。

## 3.2 SLAM 定义

同步定位与地图构建 (SLAM 或 Simultaneous localization and mapping) 是一种概念：希望机器人从未知环境的未知地点出发，在运动过程中通过重复观测到的地图特征（比如，墙角，柱子等）定位自身位置和姿态，再根据自身位置增量式的构建地图，从而达到同时定位和地图构建的目的。

在误差和噪音条件下，定位和地图构建技术上的复杂度不支持两者同时获得连续的解。即时定位与地图构建 (SLAM) 是这样一个概念：把两方面的进程都捆绑在一个循环之中，以此支持双方在各自进程中都求得连续解；不同进程中相互迭代的反馈对双方的连续解有改进作用。

同步定位与地图构建 (SLAM) 被定义为以下问题：在建立新地图模型或者改进已知地图的同时，在该地图模型上定位机器人。实际上，这两个核心问题如果分开解决，将毫无意义；必须同时求解。即通过控制数据  $u_{1:t}$  和观测数据  $z_{1:t}$  来求解位姿和地图的联合概率分布

$$p(x_{1:t}, m | z_{1:t}, u_{1:t-1}) \quad (3.1)$$

## 3.3 V-SLAM 的框架与分类

视觉 SLAM 由前端（视觉里程计）、后端（位姿优化）、回环检测、建图 4 个部分组成。

- 视觉里程计 (Visual Odometry): 仅有视觉输入的姿态估计 [4]
- 后端优化 (Optimization): 后端接受不同时刻视觉里程计测量的相机位姿，以及闭环检测的信息，对它们进行优化，得到全局一致的轨迹和地图 [5]
- 闭环检测 (Loop Closing): 指机器人在地图构建过程中，通过视觉等传感器信息检测是否发生了轨迹闭环，即判断自身是否进入历史同一地点 [5]
- 建图 (Mapping): 根据估计的轨迹，建立与任务要求对应的地图 [5]

流程如下图所示：

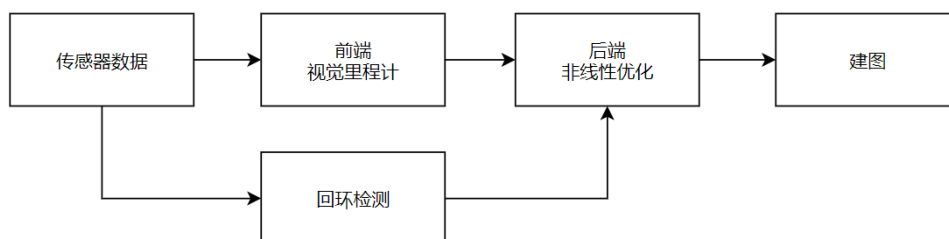


图 3.1: VSLAM 流程图 [5]

## 3.4 经典 V-SLAM 方法

### 3.4.1 稀疏法:ORB-SLAM2

该算法 [6] 是基于特征点的实时单目 SLAM 系统，在大规模的、小规模、室内室外的环境都可以运行。它主要有以下贡献：

- 1) 这是首个基于单目，双目和 RGB-D 相机的开源 SLAM 方案，这个方案包括，回环检测，地图重用和重定位。
- 2) 结果说明，BA 优化比 ICP 或者光度和深度误差最小方法的更加精确。
- 3) 通过匹配远处和近处的双目匹配的点和单目观测，双目的结果比直接使用双目系统更加精确。
- 4) 针对无法建图的情况，提出了一个轻量级的定位模式，能够更加有效的重用地图。

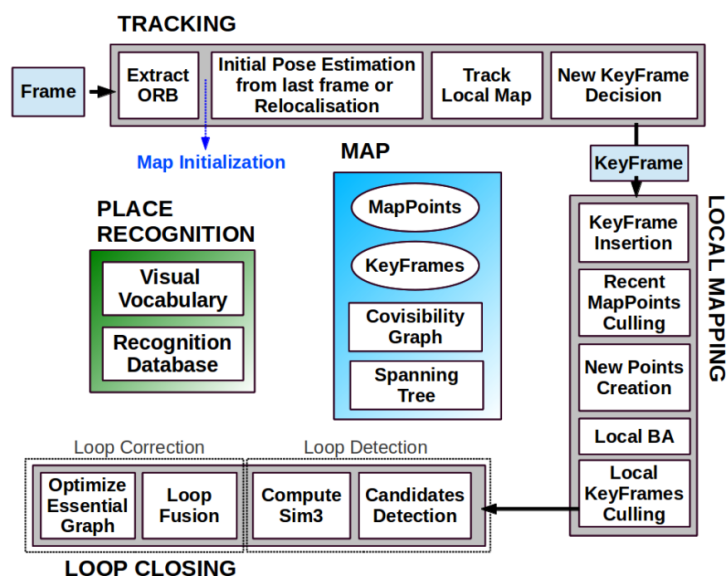


图 3.2: ORB-SLAM 主要分为三个线程进行，也就是如图所示的，分别是 Tracking、LocalMapping 和 LoopClosing [6]

### 1. 跟踪 (Tracking)

这一部分主要工作是从图像中提取 ORB 特征，根据上一帧进行姿态估计，或者进行通过全局重定位初始化位姿，然后跟踪已经重建的局部地图，优化位姿，再根据一些规则确定新的关键帧。

### 2. 建图 (LocalMapping)

这一部分主要完成局部地图构建。包括对关键帧的插入，验证最近生成的地图点并进行筛选，然后生成新的地图点，使用局部捆集调整 (Local BA)，最后再对插入的关键帧进行筛选，去除多余的关键帧。

$$\{\mathbf{R}, \mathbf{t}\} = \underset{\mathbf{R}, \mathbf{t}}{\operatorname{argmin}} \sum_{i \in \mathcal{K}} \rho \left( \left\| \mathbf{x}_{(\cdot)}^i - \pi_{(\cdot)} (\mathbf{R} \mathbf{X}^i + \mathbf{t}) \right\|_{\Sigma}^2 \right) \quad (3.2)$$

### 3. 闭环检测 (LoopClosing)

这一部分主要分为两个过程，分别是闭环探测和闭环校正。闭环检测先使用 WOB 进行探测，然后通过 Sim3 算法计算相似变换。闭环校正，主要是闭环融合和 Essential Graph 的图优化。

开源代码链接: [https://github.com/raulmur/ORB\\_SLAM2](https://github.com/raulmur/ORB_SLAM2)

## 3.4.2 多传感器融合:VINS-Mono

该算法 [7] 是一种鲁棒且通用的单目视觉惯性状态估计器。采用一种基于紧耦合、非线性优化的方法，通过融合预积分后的 IMU 测量值和特征观测值，获得高精度的视觉惯性里程计。结合紧耦合方法，回环检测模块能够以最小的计算代价实现重定位。处理视觉和惯性测量的最简单的方法是松耦合的传感器融合 [8,9]，其中 IMU 被视为一个独立的模块，用于辅助运动的视觉结构 (sfm) 获得的纯视觉位姿估计。融合通常由扩展卡尔曼滤波 (EKF) 完成，其中 IMU 用于状态传播，而视觉位姿用于更新。

### IMU 预积分 (pre-integration)

在实践中，IMU 通常以比摄像机更高的速率获取数据。不同的方法被提出来处理高速率的 IMU 测量值。最简单的方法是在基于 EKF 的方法中使用 IMU 进行状态传播 [11][13]。在图优化公式中，为了避免重复的 IMU 重复积分，提出了一种有效的方法，即 IMU 预积分 (IMU pre-integration)。通过一系列计算，得到下面的预积分估计值：

$$\begin{aligned} \hat{\alpha}_{i+1}^{b_k} &= \hat{\alpha}_i^{b_k} + \hat{\beta}_i^{b_k} \delta t + \frac{1}{2} \mathbf{R}(\hat{\gamma}_i^{b_k})(\hat{\mathbf{a}}_i - \mathbf{b}_{a_i}) \delta t^2 \\ \hat{\beta}_{i+1}^{b_k} &= \hat{\beta}_i^{b_k} + \mathbf{R}(\hat{\gamma}_i^{b_k})(\hat{\mathbf{a}}_i - \mathbf{b}_{a_i}) \delta t \\ \hat{\gamma}_{i+1}^{b_k} &= \hat{\gamma}_i^{b_k} \otimes \begin{bmatrix} 1 \\ \frac{1}{2}(\hat{\omega}_i - \mathbf{b}_{w_i}) \delta t \end{bmatrix} \end{aligned} \quad (3.3)$$

通过计算可以写下 IMU 测量模型所对应的协方差  $\mathbf{P}_{b_{k+1}}^{b_k}$  :

$$\begin{bmatrix} \hat{\alpha}_{b_{k+1}}^{b_k} \\ \hat{\beta}_{b_{k+1}}^{b_k} \\ \hat{\gamma}_{b_{k+1}}^{b_k} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_w^{b_k} (\mathbf{p}_{b_{k+1}}^w - \mathbf{p}_{b_k}^w + \frac{1}{2} \mathbf{g}^w \Delta t_k^2 - \mathbf{v}_{b_k}^w \Delta t_k) \\ \mathbf{R}_w^{b_k} (\mathbf{v}_{b_{k+1}}^w + \mathbf{g}^w \Delta t_k - \mathbf{v}_{b_k}^w) \\ \mathbf{q}_{b_k}^{w^{-1}} \otimes \mathbf{q}_{b_{k+1}}^w \\ \mathbf{b}_{ab_{k+1}} - \mathbf{b}_{ab_k} \\ \mathbf{b}_{wb_{k+1}} - \mathbf{b}_{wb_k} \end{bmatrix}. \quad (3.4)$$

### 估计器初始化

单目紧耦合 VIO 是一个高度非线性的系统。由于单目相机无法直接观测到尺度，因此，如果没有良好的初始值，很难直接将这两种测量结果融合在一起。当 IMU 测量结果被大偏置破坏时，情况就变得更加复杂了。事实上，初始化通常是单目 VINS 最脆弱的步骤。需要一个鲁棒的初始化过程以确保系统的适用性。通过对齐 IMU 预积分与纯视觉 SfM 结果，我们可以粗略地恢复尺度、重力、速度，甚至偏置。这足以引导非线性单目 VINS 估计器。

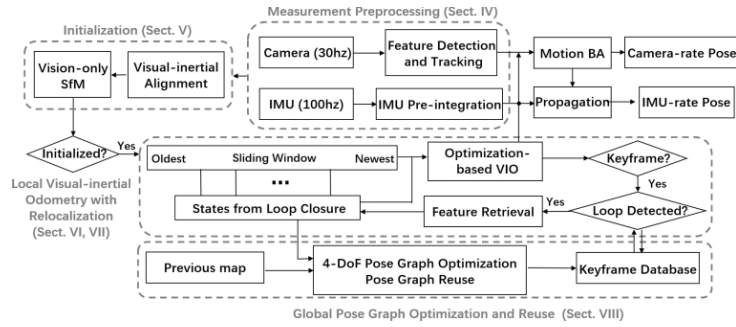


图 3.3: 总体架构图

开源代码链接: <https://github.com/HKUST-Aerial-Robotics/VINS-Mono>

### 3.4.3 blablabla

比较项目	视觉里程计	IMU 里程计
更新频率	低	高
长时间稳定性	高	低
场景依赖度	高	低
重定位能力	高	低
CPU 功耗	高	低

表 3.1: 视觉里程计和 IMU 的特性对比

比较项目	传统 SLAM 算法	基于深度视觉的 SLAM 算法
模型参数调整难易程度	数据少, 调参周期短	数据多, 调参周期长
魔性物理含义	直观意义明确	缺少直观意义
模型泛化能力	信息利用不充分, 泛化能力弱	信息利用充分, 泛化能力强
适应能力	迁移能力弱	迁移能力强
设计流程	特征设计与分类训练分离	同步完成特征设计和分类器训练

表 3.2: 传统 SLAM 算法与基于深度学习的 SLAM 算法对比 [10]

相关文献	方法	所用传感器	算法名称	发表年份
[11]	稀疏法 (特征点)	单目 SLAM	MonoSLAM	2014
[12]	半稠密法	单目 SLAM	LSD-SLAM	2014
[13]	稀疏法 (特征点)	单目 SLAM	ORB-SLAM	2015
[14]	半稠密法	单目 SLAM	SVO	2014
[6]	稀疏法 (特征点)	双目 SLAM	ORB-SLAM2	2017
[15]	稀疏法 (特征点)	双目 SLAM	S-PTAM	2015
[16]	稠密法	RGBD-SLAM	RGBD-SLAMv2	2014
[17]	稠密法	RGBD-SLAM	RTAB-SLAM	2019

表 3.3: 传统 SLAM 算法的总结

### 3.5 基于深度学习的 V-SLAM 方法

## 参考文献

- [1] O. Wongwirat and C. Chaiyarat, “A position tracking experiment of mobile robot with inertial measurement unit (imu),” in *ICCAS 2010*, pp. 304–308, 2010.
- [2] S. Krishnan, P. Sharma, Z. Guoping, and O. H. Woon, “A uwb based localization system for indoor robot navigation,” in *2007 IEEE International Conference on Ultra-Wideband*, pp. 77–82, 2007.
- [3] S. Ma, Y. Zhang, Y. Xu, B. Wang, J. Cheng, and Q. Zhao, “Indoor robot navigation by coupling imu, uwb, and encode,” in *2016 10th International Conference on Software, Knowledge, Information Management Applications (SKIMA)*, pp. 429–432, 2016.
- [4] D. Nister, O. Naroditsky, and J. Bergen, “Visual odometry,” in *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*, vol. 1, pp. I–I, 2004.



- [5] 高翔, 视觉 SLAM 十四讲. 05 2020.
- [6] Mur-Artal, Raúl, Tardós, and J. D., “ORB-SLAM2: an open-source SLAM system for monocular, stereo and RGB-D cameras,” *IEEE Transactions on Robotics*, vol. 33, no. 5, pp. 1255–1262, 2017.
- [7] T. Qin, P. Li, and S. Shen, “Vins-mono: A robust and versatile monocular visual-inertial state estimator,” *IEEE Transactions on Robotics*, vol. 34, no. 4, pp. 1004–1020, 2018.
- [8] S. Tsao and S. Jan, “Observability analysis and performance evaluation of ekf-based visual-inertial odometry with online intrinsic camera parameter calibration,” *IEEE Sensors Journal*, vol. 19, no. 7, pp. 2695–2703, 2019.
- [9] J. Jackson, K. Brink, B. Forsgren, D. Wheeler, and T. McLain, “Direct relative edge optimization, a robust alternative for pose graph optimization,” *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 1932–1939, 2019.
- [10] Z. Yang, L. Guoliang, T. Guohui, L. Yong, W. Ziren, Z. Wei, and L. Junwei, “A survey of visual slam based on deep learning,” *ROBOT*, vol. 39, no. 6, p. 889, 2017.
- [11] Russo, Ludovico, Rosa, Stefano, Bona, Basilio, Matteucci, and Matteo, “A ros implementation of the mono-slam algorithm,” vol. 4, pp. 339–351, 01 2014.
- [12] J. Engel, T. Schöps, and D. Cremers, “Lsd-slam: Large-scale direct monocular slam,” in *Computer Vision – ECCV 2014* (D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, eds.), (Cham), pp. 834–849, Springer International Publishing, 2014.
- [13] Mur-Artal, Raúl, Montiel, J. M. M., Tardós, and J. D., “ORB-SLAM: a versatile and accurate monocular SLAM system,” *IEEE Transactions on Robotics*, vol. 31, no. 5, pp. 1147–1163, 2015.
- [14] C. Forster, M. Pizzoli, and D. Scaramuzza, “Svo: Fast semi-direct monocular visual odometry,” in *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 15–22, 2014.
- [15] T. Pire, T. Fischer, J. Civera, P. De Cristóforis, and J. Jacobo berlles, “Stereo Parallel Tracking and Mapping for robot localization,” in *Proc. of the International Conference on Intelligent Robots and Systems (IROS)*, pp. 1373–1378, September 2015.

- [16] F. Endres, J. Hess, J. Sturm, D. Cremers, and W. Burgard, “3-d mapping with an rgb-d camera,” *IEEE Transactions on Robotics*, vol. 30, no. 1, pp. 177–187, 2014.
- [17] M. Labbé and F. Michaud, “Rtab-map as an open-source lidar and visual simultaneous localization and mapping library for large-scale and long-term online operation,” *Journal of Field Robotics*, vol. 36, no. 2, pp. 416–446, 2019.