

多种视觉SLAM方案对比

开源SLAM方案选型：

开源方案	链接地址	传感器	测试数据集	测试传感器
ORB_SLAM2	http://webdiis.unizar.es/~raulmur/orbslam/	mono/stereo/RGB-D	KITTI/TUM/EuRoC	Kinect V1/optor
RTAB MAP	http://webdiis.unizar.es/~raulmur/orbslam/	stereo/RGB-D	TUM	Kinect V1
RGBD SLAM V2	https://github.com/felixendres/rgbdslam_v2	RGB-D	TUM	Kinect V1
VI ORB SLAM2	https://github.com/jingpang/LearnVIO RB	ORB_SLAM2(mono) + IMU		
VINS MONO	https://github.com/HKUST-Aerial-Robotics/VINS-Mono	mono+IMU		
VINS FUSION	https://github.com/HKUST-Aerial-Robotics/VINS-Fusion	stereo/ stereo+IMU / mono+IMU	EuRoC	optor

数据集：

数据集	链接地址
EuRoC	https://projects.asl.ethz.ch/datasets/doku.php?id=kmavvisualinertialdatasets
TUM	https://vision.in.tum.de/data/datasets/rgbd-dataset/download
KITTI	http://www.cvlibs.net/datasets/kitti/eval_object.php

在SLAM研究中，我们通常需要在各数据集上测试各个方案的性能情况。如下主要对表1中6个视觉SLAM方案进行相关测试，测试其不同场景、要求下的精度和鲁棒性，并给出最终方案选型。

一、ORB-SLAM2

检测流程：Kinect生成地图【地图主要可见的有关键帧(包括相机的pose,相机的内参,ORB特征), 3D的地图点(空间中3D位置,法线方向,ORB的描述子),词袋向量,共视图等】→ 保存地图 → 加载地图和重定位

缺点：1、不能解决实际问题。它基于特征点法，建的图是稀疏的，只能满足定位需求，而无法提供导航、避障、交互等诸多功能。而且目前还没有开放存储和读取地图后重新定位的功能。

2、受光照和车辆、行人等动态物体的影响太大

后续工作：对前端数据进行轨迹的优化和回环，稠密建图

思考：1 如何利用视觉SLAM进行导航？ 激光SLAM的路线

2 导航的具体需求是什么？给个起点和终点，然后小车自主完成这个任务？这个路径是通过路径规划出来的？还是指定的路线，然后进行轨迹跟踪？

【自我备注：1、保存稀疏地图的MapPoint和KeyFrame成.bin格式；2、如果保存成pcd格式，那么关键帧与关键点之间的关系没有办法保存。。。即pcd保存的地图信息并不完整】

二、VINS MONO

实现原理：用紧耦合方法实现的，通过单目+IMU恢复出尺度

目标：AR

功能模块：

代码解读：

、RTAB-MAP

概述：与其他视觉slam方案不同在于提供了一个与时间和尺度无关的基于外观的定位与构图解决方案，优化了大型环境的在线闭环检测问题，主要是利用计算权重使得只利用有限数量的定位点参与闭环检测，但是若需要也可以访问全局的定位点。

流程：开始->特征提取和匹配->求两两帧的视觉里程计->优化php的结果->局部地图->运动状态估计->运动状态估计->位姿图->词袋模型->相似度计算->结束

特征点：使用词袋法创建图像的签名，基于opencv从图像中提取SURF特征来得到视觉单词。

建图方式：1.Octomap（3d占用栅格地图）

2.稠密点云地图

3.2d占用栅格地图

ROS节点：输入：1.TF（传感器相对于机器人底座的位置）
2.里程计
3.相机输入以及带有校准的信息

输出：1.Map Data和Graph

优点：1.适用于长期和大规模环境在线建图的要求
2.里程计鲁棒性较好好且低漂移
3.定位鲁棒性较好
4.地图生成开发实用且简便
5.提供软件包

缺点：1.方案偏向于商品化，二次开发难

参考：<https://blog.51cto.com/remyspot/1784914>

、RGBD-SLAM V2

概述：系统前端为视觉里程计，从每一帧的RGB提取特征，计算描述符，RANSAC+ICP计算两帧之间的运动估计，并提出了一个EMM（环境测量模型）判断运动估计是否可以接受，后端回环检测基于G2O的位姿图优化

特征点：1.SIFT by GPU(综合表现最好)

2.SURF by CPU(OPENCV)

3.ORB by CPU (OPENCV)（准确率，实时性更加好）

建图方式：八叉树地图octomap（利于导航；易于更新；存储方式比较省空间）

优点：1.二次开发较容易，可以在其基础上继续开发

缺点：1.实时性相对较差，相机要慢速运动

参考：<http://www.cnblogs.com/voyagee/p/7027076.html>

三、ORB_SLAM2、VI ORB SLAM2、VINS Mono对比实验

参考链接：<http://www.liuxiao.org/2018/02/vi-orb-slam2-%E4%B8%8E-vins-%E5%AF%B9%E6%AF%94%E5%AE%9E%E9%AA%8C/>

主要指标：均方根误差

实验结果：

1、ORB 与 VI ORB 对比

EuRoC	ORB Mono	ORB Stereo	VIORB Mono	VIORB Stereo	VIORB Mono Paper
	RMSE	RMSE	RMSE	RMSE	RMSE
V1_01_easy	0.015	0.087	0.094	0.09	0.023
V1_02_medium	0.020	0.065	0.118	0.067	0.027
V1_03_difficult	x	0.085	x	0.075	x
V2_01_easy	0.015	0.064	0.242	0.07	0.018
V2_02_medium	0.017	0.06	0.104	0.118	0.024
V2_03_difficult	x	x	x	x	0.047
MH_01_easy	0.070	0.038	0.092	0.072	0.068
MH_02_easy	0.066	0.048	0.089	0.054	0.073
MH_03_medium	0.071	0.038	0.091	0.053	0.071
MH_04_difficult	0.081	0.103	0.773	0.15	0.087
MH_05_difficult	0.060	0.056	0.286	0.099	0.060

由表可以看出：

- 1) 双目 VIO 比单目 VIO 效果要更好也更稳定；
- 2) 原始 VO 的优化已经比较彻底，IMU 误差项的加入给优化结果造成了更多的不稳定性，加入 IMU 的版本均不如原始双目版本的精度。

2、VI ORB 与 VINS Mono 对比（开启闭环）

EuRoC	VIORB Mono	VIORB Stereo	VIORB Mono Paper	VINS Mono
	RMSE	RMSE	RMSE	RMSE
MH_01_easy	0.092	0.072	0.068	0.06
MH_02_easy	0.089	0.054	0.073	0.066
MH_03_medium	0.091	0.053	0.071	0.135

MH_04_difficult	0.773	0.15	0.087	0.056
MH_05_difficult	0.286	0.099	0.060	0.058
V1_01_easy	0.094	0.09	0.023	0.098
V1_02_medium	0.118	0.067	0.027	0.061
V1_03_difficult	x	0.075	x	0.064
V2_01_easy	0.242	0.07	0.018	0.052
V2_02_medium	0.104	0.118	0.024	0.174
V2_03_difficult	x	x	0.047	0.140

由表可以看出：

- 1) VINS Mono 精度与 VIORB Mono Paper的单目精度比是明显不如的，但与 VIORB Mono/Stereo比精度要高一些；
- 2) VIORB 是不能完全跑过所有测试集的，特别是快速运动的 V2_03_difficult 测试集结尾阶段会跟丢；而vins mono采用的光流跟踪，不易跟丢。

3、VIORB 与 VINS Mono 对比（关闭闭环）

EuRoC		VIORB Mono Paper		VIORB Stereo		VINS Mono	
	No Loop	Loop	No Loop	Loop	No Loop	Loop	
	RMSE	RMSE	RMSE	RMSE	RMSE	RMSE	
V1_02_medium	0.028	0.027	0.07	0.065	0.105	0.066	
V2_02_medium	0.041	0.024	0.092	0.102	0.103	0.058	
MH_04_difficult	0.217	0.087	0.288	0.243	0.300	0.174	

由表可以看出：

VINS Mono 对闭环依赖较大，在关闭闭环后，VINS Mono 的性能下降较多。

6、实验小结

- 1) 在没有闭环情况下，VINS Mono 精度略低于 VIORB；而在有闭环情况下，二者精度差距不大。

- 2) 由于开源 VIO RB 版本并非官方实现，与官方版本有很多不同之处，没有办法测试真实官方 VIO RB 的性能；但基本可以看出对于 VIO RB SLAM2 框架，IMU 的引入主要是在快速运动时能够减少一些丢失，而精度上与 VO 相近或略有下降；
- 3) 以光流作为前端的 VINS Mono 比描述子作为前端的 VIO RB 具有更好的鲁棒性，在快速运动时更不容易丢失。因此在类似于无人机这样的场景，VINS 应该是比 VIO RB 更好的选择；而对于精度要求较高、场景更大而运动较慢的场景，VIO RB /ORB 仍然更合适；
- 4) 不论是否引入 IMU，双目对于精度和鲁棒性都会有一定的提升，VINS Fusion 相比于 vins mono，性能有所提升； {精度——慢速，双目；鲁棒——快速}
- 5) 除前端差别较大外（比如光流与ORB、滑窗与局部地图），虽然都是预积分，二者后端算法与实现亦有较大不同（李代数与四元数参数化不同、积分方式不同、g2o与ceres实现差别等），因此造成性能差异的原因较复杂。
- 5、VIO RB-SLAM初始化与VINS初始化对比(将vi orb-slam初始化方法移植到vins 中)

四、VIO：飞行机器人单目VIO算法测评

参考链接：<https://blog.csdn.net/u012348774/article/details/81414264>

结论：SVO+MSF是计算效率最高的算法，而VINS-Mono是状态估计精度最高的算法，ROVIO则处于两者之间

五、VINS FUSION

发布信息：港科大 2019年1月12号

版本：

- (1) 单目+imu
- (2) 纯双目
- (3) 双目+imu
- (4) 双目+imu+GPS

功能模块及函数接口：与vins mono相差不大，个别库的更新和优化

和vins mono相比：

优势：可以静止进行初始化；尺度信息不一定完全依靠IMU,不会造成尺度不可关的情况；鲁棒性上，双目明显优于单目

劣势：由于视觉误匹配等各种原因，双目的精度会比单目差一点。

传感器	单目	双目	深度
缺点	无法获取真实尺度，精度低	运算量大	不适合室外

【自我备注：使用vinsfusion闭环跑KITTI的纯双目数据集，跑出来的output结果有vio.txt位姿文件，但这个没有闭环，闭环的数据在单独的vio_loop.csv文件里，这个文件里的数据只有闭环帧，而且数据格式和真实轨迹不一样，不是12列。如何将闭环的vio_loop.csv数据融合到未闭环的vio.txt数据里，然后画出闭环后的整体轨迹估计？】

六、RTAB-MAP和RGBD-SLAM V2的对比

系统构建：RGBD SLAM2是一个非常全面优秀的系统，将SLAM领域的图像特征、优化、闭环检测、点云、octomap等技术融为一体，

RTAB Map是当前最优秀的RGBD SLAM

开发：RGBD SLAM2适合RGBD SLAM初学者，也可以在其基础上继续开发。

RTAB-MAP二次开发难度较高(著名的Google Tango（见如何评价Google 的 Project Tango和Google Project Tango 有哪些黑科技）就是使用RTAB Map做SLAM，)

实时性：RGBD SLAM2的缺点是其算法实时性不好，相机必须慢速运动，

RTABMAP 通过STM/WM/LTM的内存管理机制，减少图优化

和闭环检测中需要用到的结点数，保证实时性以及闭环检测的准确性，能够在超大场景中运行。

回环检测：RGBDSLAMV2:相机较快运动时，会出现很大的地图重叠。

RTABMAP:稳定性相对较好，但是也不稳定。

个人：实际运行中RGBDSLAMV2用点云表达，点云表达方式耗费内存，时间较长后，三维空间中的点云数量增多，画面出现明显的卡顿，但采集得结果细节表现得很好

RTABMAP实际运行过程中，画面比较流畅，但是是使用3d栅格占用地图表现的，细节表现得也不错。

七、RTAB-MAP和ORB-SLAM的对比

系统构建：RTAB-MAP：框架较全面，支持联合建图，建立的地图包括三维稠密点云和二维栅格地图（可以直接从ros获取二进制程序）

ORB-SLAM：只满足了定位的需求，无法提供导航，避障，交互等诸多功能。

代码可读性：RTAB-MAP：差，基本上封装了。

ORB-SLAM：可读性好。

应用：RTAB-MAP：适合作为slam应用

ORB-SLAM：适合研究使用

计算量：RTAB-MAP：内存管理方法计算量相对较小

ORB-SLAM：多线程结构，必须对每幅图像计算一遍orb特征，耗时和计算量大

回环检测：RTAB-MAP：回环检测只访问有限数量的定位点

ORB-SLAM：优秀的回环检测算法保证了ORB-

SLAM有效地防止累积误差，并且在丢失之后还能迅速找回，这一点许多现有的SLAM系统都不够完善。为此，ORB-SLAM在运行之前必须加载一个很大的ORB字典。

八、多种VIO算法对比：

一：视觉和IMU前端数据处理

对于视觉前端做法基本分为三种：

①特征点提取 (*feature detection*) + 根据描述子匹配特征点

orbslam, okvis

优点: 因为有描述子，因此对地图的维护很方便（包括重定位，闭环，全局优化）

尤其是对室内环境，当视觉上共视关系较多的时候，这种方法能很大提高定位精度和局部的稳定性。室内移动机器人建议使用这种方法

缺点: 每帧图像都要提取特征点数量的描述子，浪费时间; 在tracking过程中，如果运动过快（图像出现模糊）比较容易tracking failed, 而光流要好一些

②特征点提取+光流跟踪

vins, svo的初始化

优点: 简单高效，tracking要鲁棒一些

缺点: 不容易构建全局map，视觉约束只靠滑窗里面的关键帧。

比如vins的闭环和重定位是需要另外再提取特征点和描述子的；但后端做得好

③直接法

lsdslam, dso

优点: 在弱纹理下，鲁棒性好

缺点: 不容易维护全局地图，对光照影响较大，高精度地图中无法使用等。

对于IMU前端:

IMU前端基本都是使用预积分的方式，和积分的区别（预积分把重力考虑进去了）。imu预积分主要是在两帧图像之间更新imu的5个状态变量（ p, v, q, b_a, b_g ）的Jacobian和covariance。

jacobian是单位矩阵，主要是为了求一阶近似的 p, v, q ；covariance是零矩阵，主

要是为了求imu误差项的权重

IMU初值确定：一般结合vins mono 等的视觉信息来求

二：后端优化

滤波：不需要做边缘化，重定位难。因为不需要迭代，速度略快。

优化：需要边缘化（比较耗时间），重定位容易。迭代优化，速度略慢。

VINS-Mono与OKVIS类似，是基于非线性优化器优化一个滑窗内的关键帧，帧间通过鲁棒的角点关联。在系统初始化阶段，通过松耦合的方式融合多种传感器；在重定位阶段，则通过紧耦合的方式融合传感器。在优化之前，也通过IMU预积分减少计算量。除此至外，VINS还提供了基于4DoF的位姿图优化和回环检测。

九、实体测试

开源方案	传感器	实验结果分析
ORB_SLAM2	①Kinect V1	Kinect V1相机提取特征点比较稳定，在慢速运行时图像不易跑飞；但由于ORB_SLAM2基于Kinect生成的地图主要可见的有关键帧,3D的地图点,共视图等，与RTAB 和RGBD比较效果不明显。 整体来说在慢速运动下，ORB经过后期优化，和RTAB相差不大，RGBD主要用于三维重建。

	②optor	<p>optor是stereo+imu，在测试过程中非常不稳定，极易飞漂，且容易跟丢。</p> <p>产生原因：①摄像头问题，经多数卖家反映质量差；②optor相机漂移大，原因是imu和图像同步不好，初始化经常失败，导致运行时总是丢失位置</p> <p>改进方案：拟后期测试小觅双目或ZED相机，口碑不错</p>
RTAB MAP	Kinect V1	<p>画面用3D栅格占用地图表现，画面细节表现不错，相机运动较慢时，建图结果较稳定，但是第二次实验时不知原因出现建图失败的现象（不管相机怎么运动，地图都不会更新）</p>
RGBD SLAM V2	Kinect V1	<p>画面用点云图表现，建图容易出现重叠，说明回环检测稳定性较差，但画面细节表现不错，时间长了出现明显的卡顿</p> <p>原因分析：没有意识到kinect也要标定，可能效果因此打了折扣</p>
VINS FUSION	optor	<p>由于vins fusion为最新版本，运行非小觅相机时需自己写启动节点和订阅、话题影射，能力有限，后期再抽空实地测试。</p>

十、方案选型（个人观点，不喜勿喷）

出于应用考虑，前期拟选择。。。。，后期。。。。。