# 新近v-SLAM算法总结

主要摘录了从2018年至今的v-SLAM算法，目前总结算法共5种（不包括经典算法和未将代码发布在github上的内容）。

大多数视觉冲击系统可以根据两个关键属性进行分类。它们可以是直接的、间接的、密集的或稀疏的方法。间接方法使用中间表示，如基于图像补丁的局部特征或模板，而直接方法则使用原始像素强度。密集方法试图基于所有像素重建环境的完整3D模型，而稀疏方法仅使用图像点的子集来创建3D点。直接和密集方法的代表是密集跟踪和映射（DTAM）和大尺度直接撞击（LSD-Slam）。两者都试图基于原始像素强度最小化帧之间的光度误差项，以估计运动，这是计算上的要求。因此，一些方法结合了直接稀疏法，如直接稀疏里程法（DSO）。它只提供视觉里程计而不关闭环路，这是纠正大规模地图场景漂移的关键。对于回路闭合，必须依赖基于特征的方法。其他方法，如快速半直接单眼视觉里程计（SVO）使用直接的运动估计方法和间接的建立地图方法。直接方法通常需要原始像素强度可用，这需要在协作场景中进行大量数据交换。间接和稀疏的方法由于其图像抽象和稀疏的地图表示而引起了协同应用的关注。间接和稀疏系统的例子有monoSLAM和PTAM。最近的一个代表是ORB-SLAM2，它是一个全面的SLAM框架，能够实时并行跟踪、映射、循环关闭和重新定位。

目录

[新近v-SLAM算法总结 1](#_Toc10774)

[经典算法介绍： 3](#_Toc9622)

[RGB-D SLAM2： 3](#_Toc24846)

[ORB-SLAM2： 3](#_Toc21935)

[ElasticFusion： 4](#_Toc12256)

[新近算法： 5](#_Toc12675)

[CPL-Sync： 5](#_Toc4775)

[CubeSLAM： 6](#_Toc14592)

[Se2lam: 7](#_Toc13526)

[DRE-SLAM: 9](#_Toc20515)

[GCNv2-SLAM： 10](#_Toc32637)

[MVSNet 11](#_Toc9372)

[LAP-SLAM 13](#_Toc14055)

[Passive Initialization Method 14](#_Toc27003)

[SPM-SLAM 15](#_Toc2575)

[轻量级可扩展视觉惯性运动捕捉系统 16](#_Toc24917)

[CoSLAM 17](#_Toc19122)

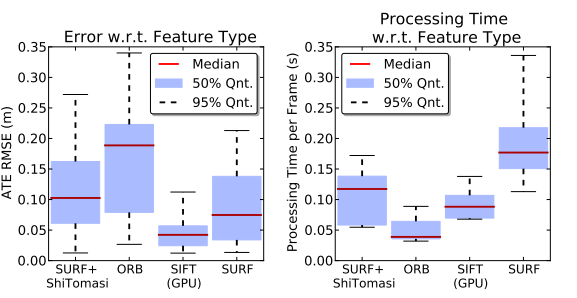
[Svl-SLAM 18](#_Toc12929)

[LLD 19](#_Toc31504)

## 经典算法介绍：

### RGB-D SLAM2：

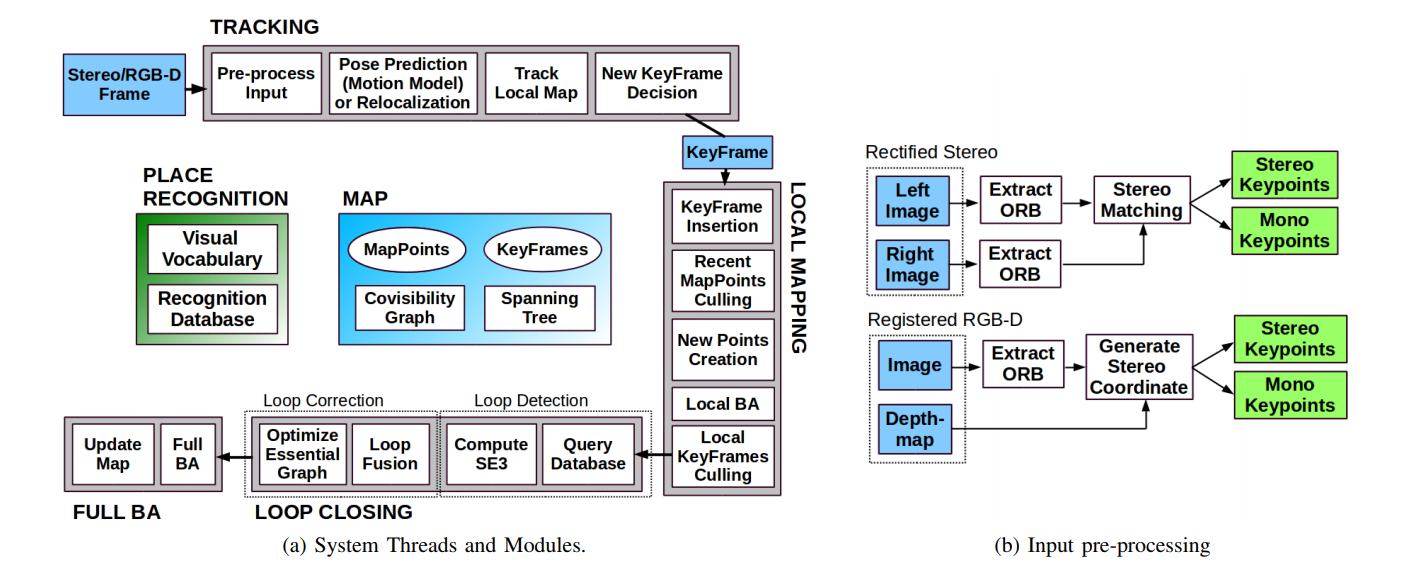
* 算法概述：14年由Endres提出的一种基于特征点提取的系统，用深度图像生成的点云通过ICP估计相机位姿，再依据相机位姿拼接多帧点云，并用TSDF模型表达重建结果。
* 技术特点：前端（RGB-D+SIFT+RANSAC+ICP+EMM（Environment Measurement Model）模型），后端（g2o+回环检测+八叉树地图octomap）。
* 技术缺陷：如果使用效果最好的SIFT算子提取特征点费时且对资源消耗极高，导致算法实时性不太好；图像采集慢；采集关键帧的频率很高，不太适合做长时间的SLAM。
* 应用范围：小范围（如室内）且对场景恢复程度要求高的情况。
* 算法效率：



可以选择不同算子，效率如图，有GPU时SIFT综合表现最好。 综合实时性、硬件成本和准确率来看，ORB较好。但用点云表达三维地图很耗费内存。

### ORB-SLAM2：

* 算法概述：15年提出的单目SLAM解决方案，并于16年更新了orb-slam2，在原来的单目基础上增加了双目和RGBD的接口。地图输出是稀疏特征点图。是目前为止效率和效果都不错的一套SLAM方案。
* 技术特点：前端（单目/双目/RGB-D图像+ORB算子）、后端（局部建图+回环检测）



ORB-SLAM2由三个平行的线程组成，跟踪，局部建图和回环检测。在一次回环检测后，会执行第四个线程，去执行BA优化。跟踪的线程在双目或者RGB-D输入之前进行，因此剩下的系统模块能够跟传感器模块独立运行。

代码轻量级并且能够在标准的CPU上面运行。可以稳定长时间运行。

* 技术缺陷：结果为稀疏点云，对地图的表达欠佳。
* 应用范围：已有地图且对地图表达精度要求不高的场景
* 算法效率：

在一台16G的RAM，Intel Core i7-4790的台式机运行，以低于传感器的帧率，对处理跟踪时间求平均。运行数据集5次，取中间值，来消除多线程系统的不确定性。

这个双目传感器有个小于54厘米的基线并且在在1392\*512像素上，以10Hz的采样速率进行采样，其中序列00,02,05,06，和09包含回环。我们的ORB-SLAM2能够检测出回环并且能够地图重用，除了09序列以外，09序列的回环只发生在尾端少数的几帧当中。在大多数序列当中比双目的LSD-SLAM要优秀很多，并且能够获得的相关误差低于1%。

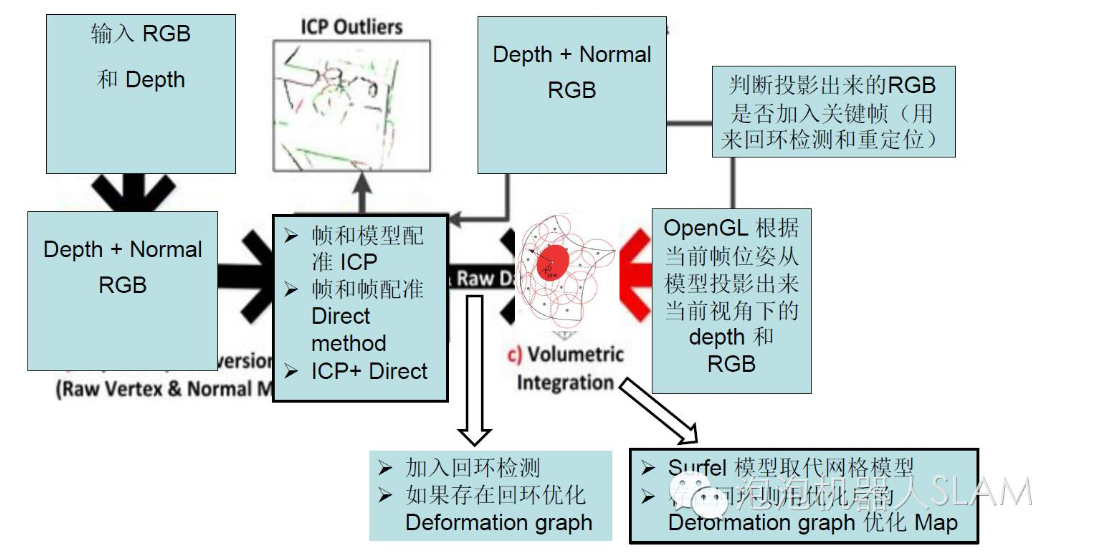


* 论文出处：

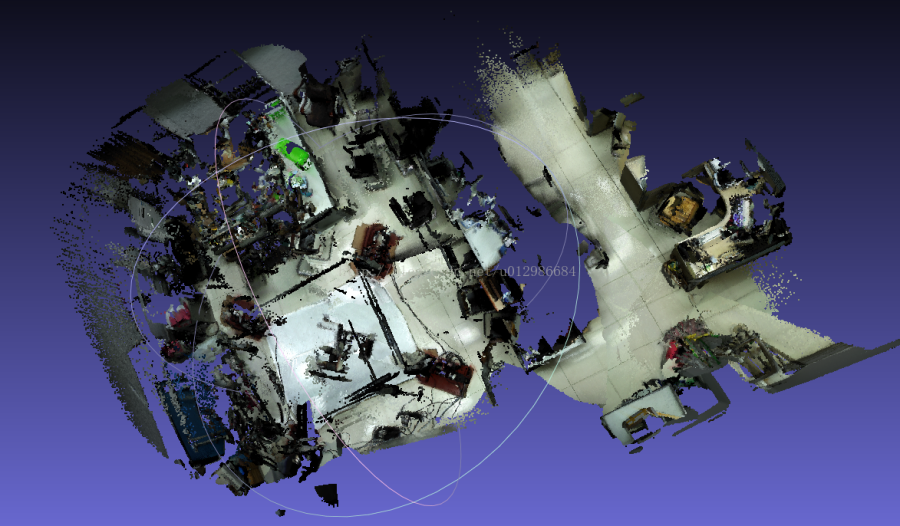
1. Mur-ArtalR, TardósJD (2016) ORB-SLAM2:an open-source SLAM system for monocular, stereo and RGB-D cameras. CoRR. abs/1610.06475
2. https://github.com/raulmur/ORB\_SLAM2

### ElasticFusion：

* 算法概述：15年提出的充分利用RGBD相机的信息，利用RGB的颜色一致性估计相机位姿，以及利用深度图像生成点云进行ICP来估计相机位姿，通过不断对比当前帧与之前的地图模型对比来提高相机位姿的估计精度，最后用surfel模型进行地图表达。
* 技术特点：

充分利用颜色与深度信息，在重建和定位的精度都是十分优秀的。

* 技术缺陷：这种方法忽略了非刚性形变地图的位姿和回环检测的性能。但由于在地图当中面元的数量影响计算的复杂程度，而且代码并未有效优化，只适合对房间大小的场景进行重建。
* 应用范围：小范围室内
* 算法效率：



由于只能适用于一个房间的小场景，虽然重建效果不错，但是缺少回环检测和内存占用大。

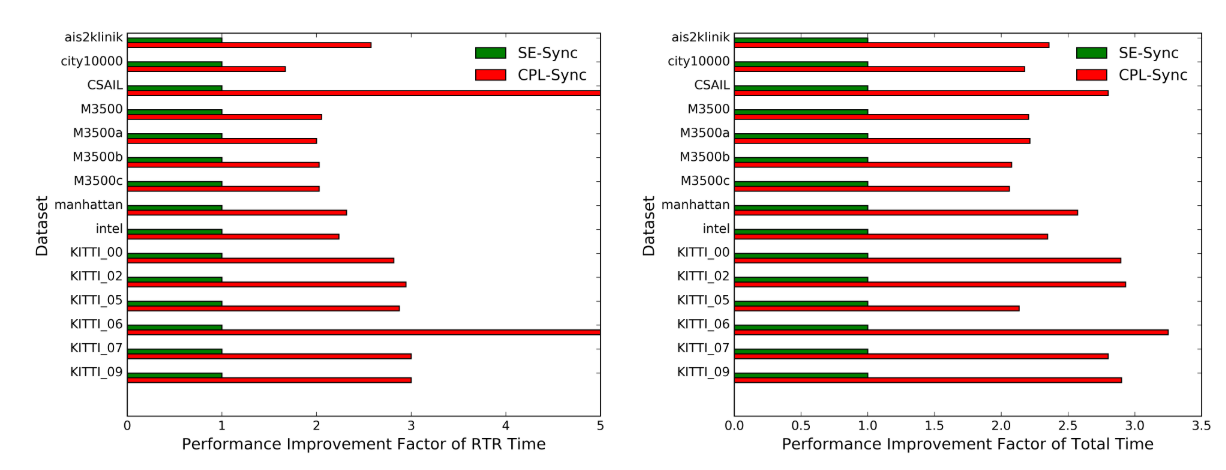
* 论文出处：

1. ElasticFusion: Dense SLAM Without A Pose Graph, T. Whelan, S. Leutenegger, R. F. Salas-Moreno, B. Glocker and A. J. Davison, RSS '15

## 新近算法：

### CPL-Sync：

* 算法概述：这并不是一套slam方案，只是新出的优化方法，CPL-Sync使用复数表示来解决平面姿态图优化（PGO）。
* 技术特点：使用复数表示在SE（2）上执行同步。
* 技术缺陷：
* 应用范围：
* 算法效率：



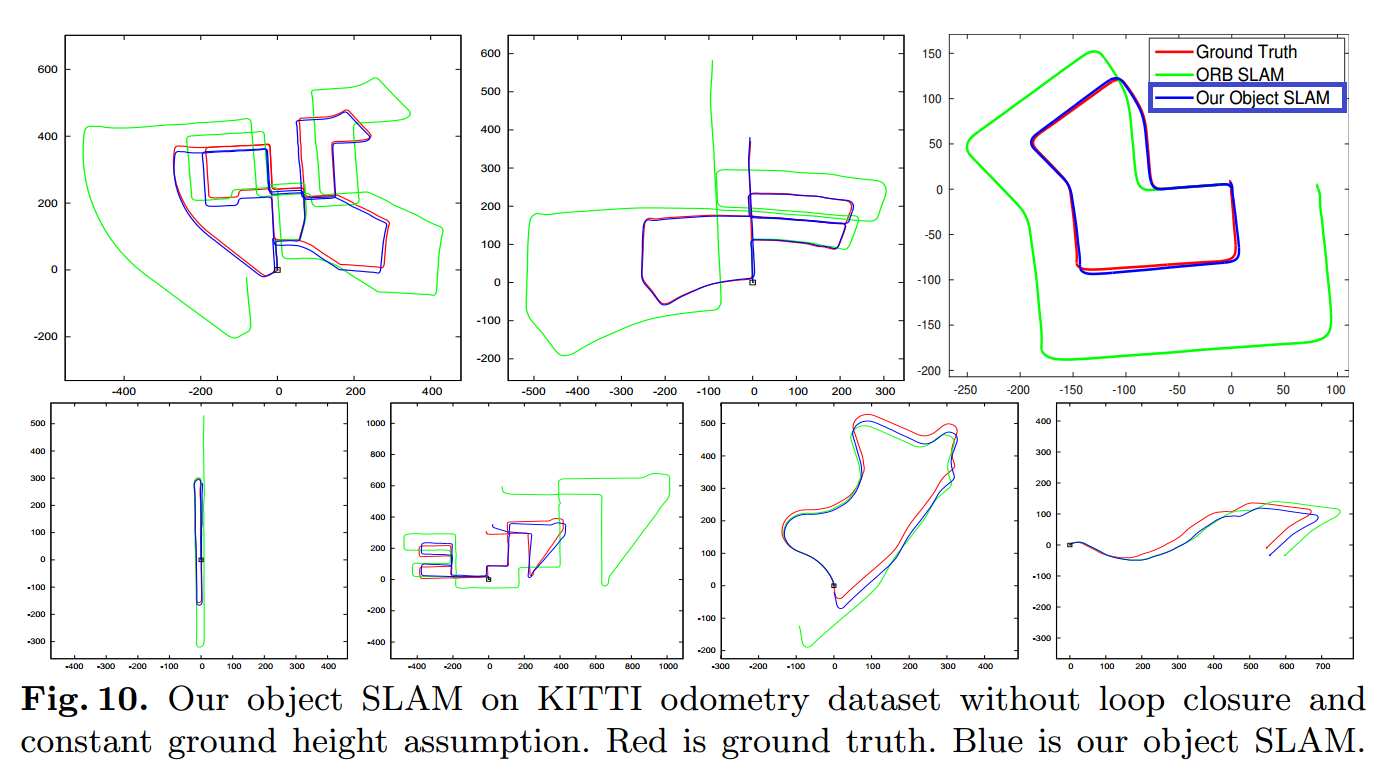
比现有的SE-Sync的效率提高显著

* 论文出处：<https://github.com/fantaosha/CPL-Sync>

### CubeSLAM：

* 算法概述：没有先验目标模型的单目三维目标检测与SLAM方法，并首次证明，语义目标检测与几何 SLAM 在一个统一框架中可以相互关联。
* 技术特点：

1. 提出一种目标级的 SLAM，在相机-目标-点之间产生测量，并提出新的目标关联方法，有效处理遮挡和动态情形；
2. 但需要提供初始高度。目标可以为点提供深度约束，为相机提供尺度约束，反过来，SLAM 也为 3D 目标检测提供初始化。

* 技术缺陷：受限制于目标提取精度，如果最近20帧中没有可见物体，会使用点云平面拟合，使用恒定地面高度假设来缩放相机位姿和局部地图；不适用于飞行器或手持相机，如果地面不可见，比如图 5 中的图像帧，会导致失败；
* 应用范围：室内或可控制实验区域
* 算法效率：

在没有回环检测的前提下，比orbslam的轨迹误差小

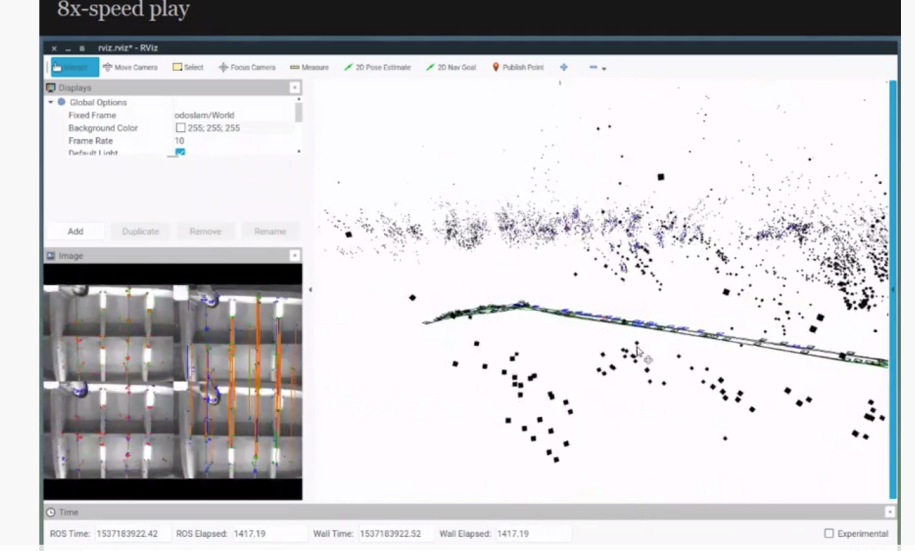
* 论文出处：

1. Yang S, Scherer S. CubeSLAM: Monocular 3D Object Detection and SLAM without Prior Models[J]. arXiv preprint arXiv:1806.00557, 2018.
2. http://www.wuxiaolang.cn/paper\_read\_cubeslam
3. 还有一篇后续的研究也是关于2D图像进行语义分割来帮助确定地图尺度和地图重建的：

Yang S, Scherer S. Monocular Object and Plane SLAM in Structured Environments[J]. arXiv preprint arXiv:1809.03415, 2018.

### Se2lam:

* 算法概述：首先，针对基于SE（2）的姿态估计问题，提出了一种新的SE（2）XYZ约束，该约束既包含图像特征测量，又包含非SE（2）运动扰动。它比随机约束更简单、更健壮，比以前的确定性SE（2）约束更符合现实世界。第二，提出了一种用于里程测量的预积分算法，它直接位于SE（2）上，与之前的SE（3）方法不同。
* 技术特点：

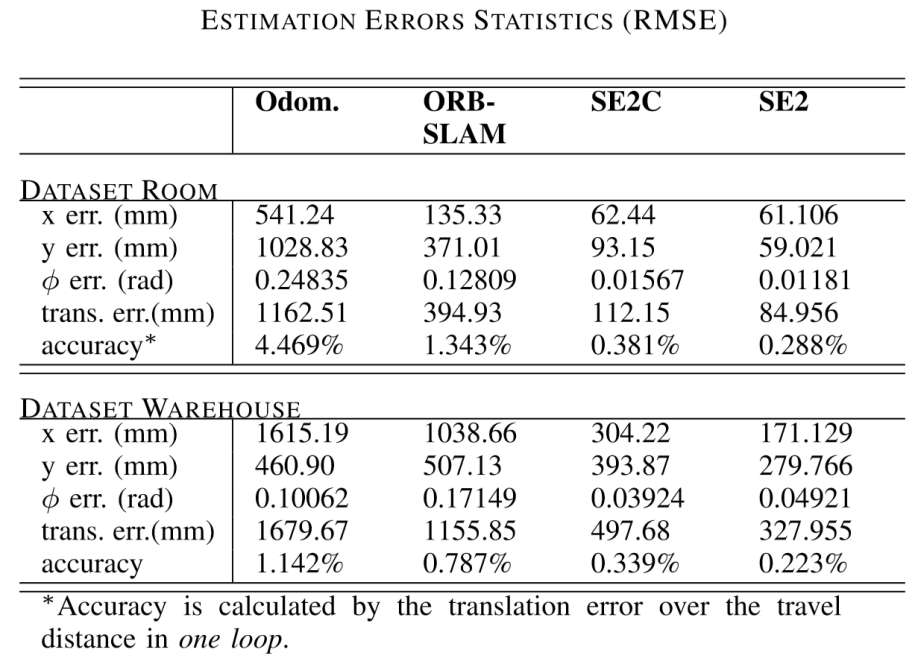


与传统的SE（3）-XYZ约束不同，这里的关键帧是用SE（2）姿态参数化的，并且引入了一个新的SE（2）-XYZ约束。





* 技术缺陷：只考虑SE（2）姿态，且恢复结果为稀疏点云
* 应用范围：针对地面导航，通常主要考虑SE（2）姿态，并且可以利用车辆在平面（或近平面）表面上移动的约束来辅助状态估计。
* 算法效率：



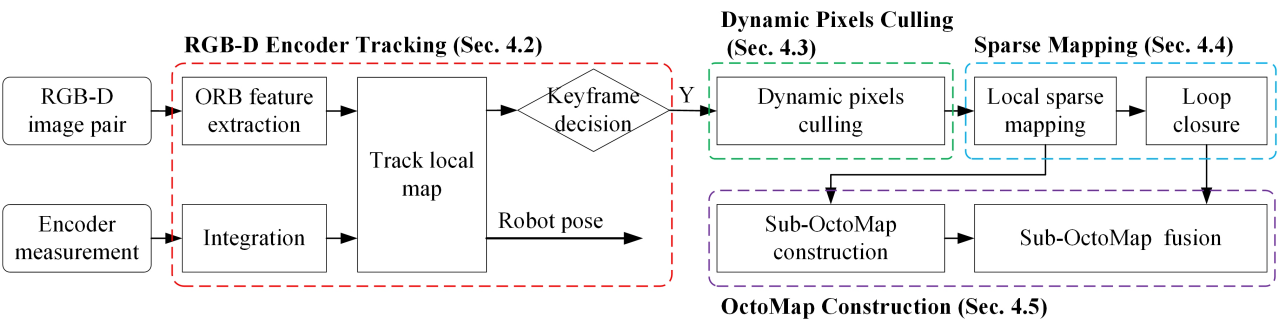
就RMSE而言，SE2测试结果在数据集室的准确率约为0.288%（0.085m/29.5m），在数据集的准确率约为0.223%（0.328m/147m），均优于SE2c测试结果（0.381%和0.339%）。与SE2C相比，SE2的一个更重要的改进是其鲁棒性。在SE2c的实验中，我们发现该估计对应用于SE（3）姿态的随机SE（2）约束的参数设置非常敏感。为了使估计系统正常工作，有时需要进行相当多的手动参数调整，这在实践中是不可取的。这可能是由于复杂系统数值优化固有的不稳定性。由于优化图是直接在SE（2）上建立起来的，这是一个更简单的线性系统的求解。

* 论文出处：

1. Fan Zheng, Yun-Hui Liu. "Visual-Odometric Localization and Mapping for Ground Vehicles Using SE(2)-XYZ Constraints". Proc. IEEE International Conference on Robotics and Automation (ICRA), 2019

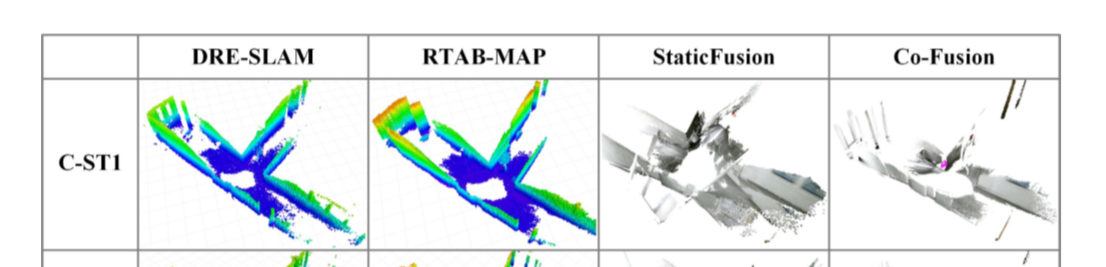
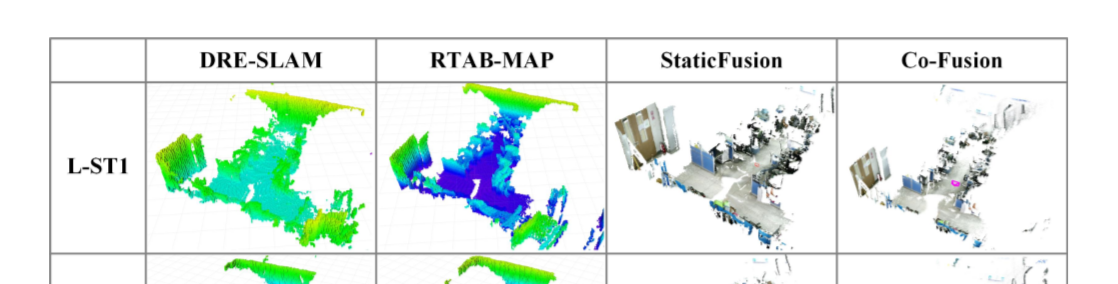
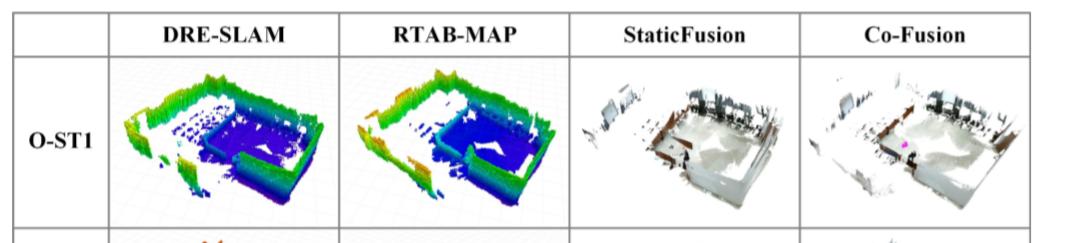
### DRE-SLAM:

* 算法概述：通过融合RGB-D摄像机和安装在差分驱动机器人上的两个编码器的信息，在动态场景中构建干净的静态OctoMap。首先使用关键帧构造局部子OctoMaps，然后将子OctoMaps融合成完整的OctoMap。这种基于子图的方法使OctoMap具有变形能力，并显着缩短了地图更新时间和内存成本。
* 技术特点：



首先使用关键帧构造局部子OctoMaps，然后将子OctoMaps融合成完整的OctoMap。这种基于子图的方法使OctoMap具有变形能力，并显着缩短了地图更新时间和内存成本。

* 技术缺陷：目前只在室内工业环境下测试完成，将增加与IMU系统的结合构建导航VINS
* 应用范围：平面级别的导航；存在车轮打滑的问题，以提高稳健性和准确性。将探索跟踪动态对象的功能。
* 算法效率：即使环境中存在动态对象，DRE-SLAM也能在所有序列中实现精确的姿态估计和令人印象深刻的静态地图。 这表明了系统的稳健性。 RTAB-MAP在大多数序列上也可靠地工作，因为它的局部姿态估计仅依赖于轮回测试。 只有环境和邻近检测是由动态对象共同构成的可见信息。

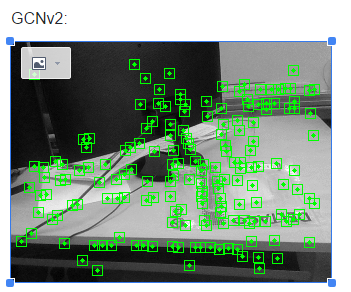
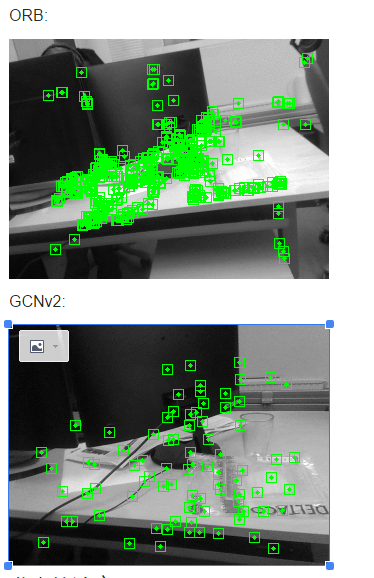


* 论文出处：

1. Dongsheng Yang, Shusheng Bi, DRE-SLAM: Dynamic RGB-D Encoder SLAM for a Differential-Drive Robot. Remote Sens. 2019, 11(4), 380;

### GCNv2-SLAM：

* 算法概述：GCNv2采用二进制描述符向量作为ORB特性，因此它可以在ORB-SLAM等系统中轻松替换ORB。属于orbslam在无人机嵌入方面的技术突破。
* 技术特点：显著提高了只能在桌面硬件上运行的gcn的计算效率。我们将展示如何使用gcnv2功能在嵌入式低功耗平台jetson tx2上运行ORB-SLAM的修改版本。



* 技术缺陷：稀疏点云，效率受制于网络和后台处理时间
* 应用范围：可用于无人机
* 算法效率：同orbslam
* 论文出处：

1. GCNv2: Efficient Correspondence Prediction for Real-Time SLAM, J. Tang, L. Ericson, J. Folkesson and P. Jensfelt, in arXiv:1902.11046, 2019
2. git clone https://github.com/jiexiong2016/GCNv2\_SLAM.git
3. Geometric Correspondence Network for Camera Motion Estimation, J. Tang, J. Folkesson and P. Jensfelt, RA-L and ICRA 2018

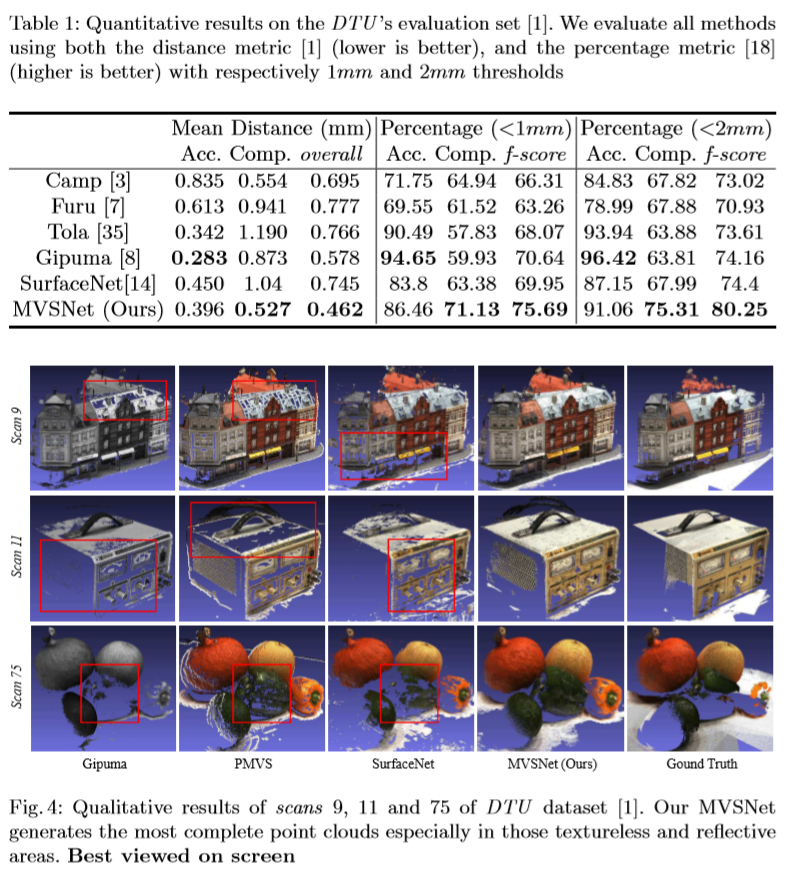
### MVSNet

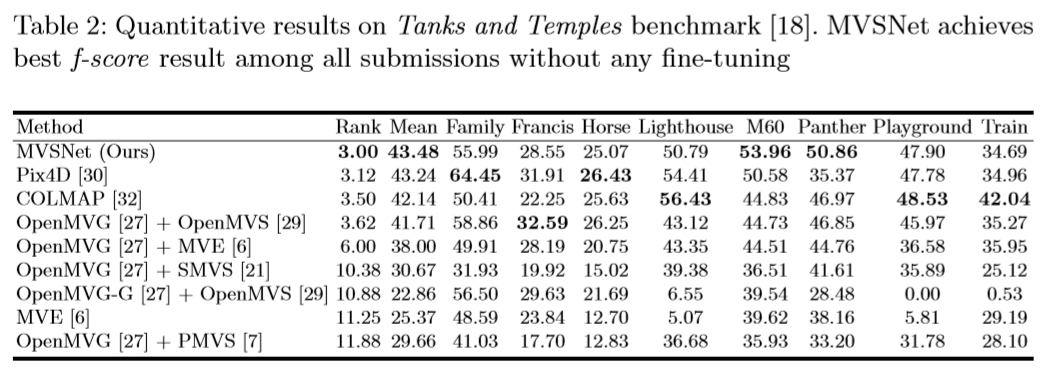
* 算法概述：2018年提出的高精度高效率的三维重建网络。
* 技术特点：

1. 通过可微单应映射（Differentiable Homography）操作将相机几何嵌入深度神经网络中。
2. 设计了一个基于方差的多视点匹配代价计算准则，将任意数量的像素特征向量映射为一个匹配代价向量，从而使得同一网络适用于任意数目的图像输入。

* 技术缺陷：内存占用大，不适合大场景
* 应用范围：
* 算法效率：

相较于传统的MVS算法，在保持了重建准确度的同时，极大地提高了模型的重建完整度。MVSNet在DTU数据集上获得了**远超**其他方法的结果；





Tanks and Temples benchmark：MVSNet在未重新微调训练的情况下，在室外场景Tanks and Temple榜单中排名第一（before April 18）。此外，MVSNet的运行速度也数倍快于之前的state-of-the-arts。

* 论文出处：

https://github.com/YoYo000/MVSNet

Mvsnet: Depth inference for unstructured multi-view stereo[C]//Proceedings of the European Conference on Computer Vision (ECCV). 2018: 767-783.

* 该团队在2019年引入循环神经网络，构建了R-MVSNet，依序地在深度方向通过GRU单元正则化2D的代价图，以减少内存消耗进而使其适用于大尺度的三维场景重建。

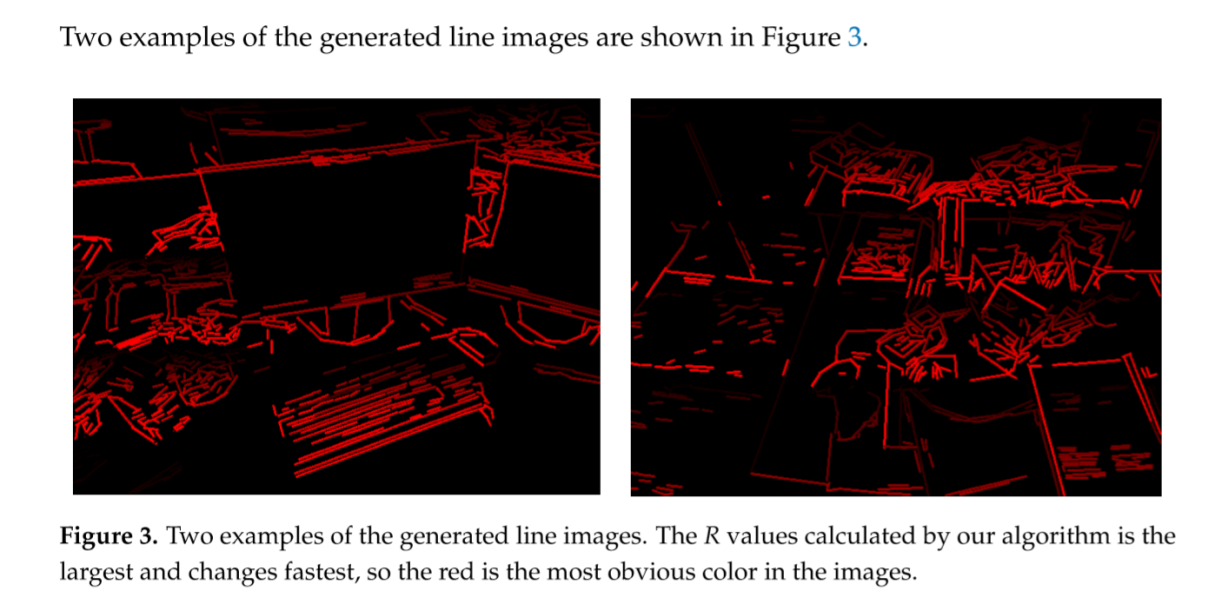
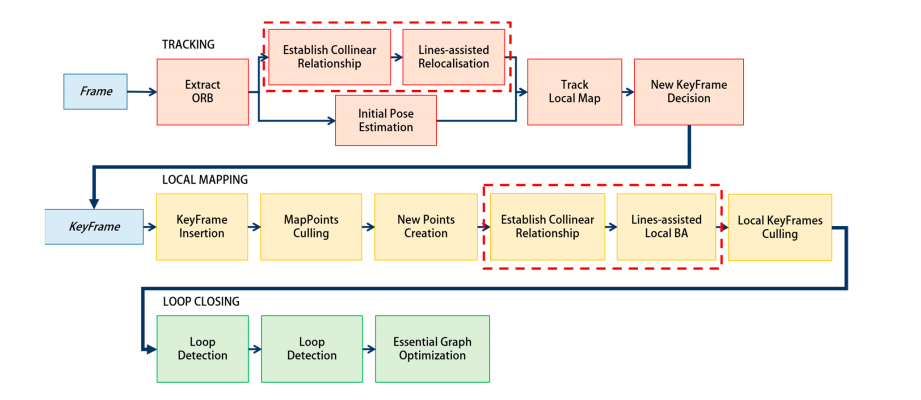
论文链接：

https://arxiv.org/abs/1902.10556

代码链接：

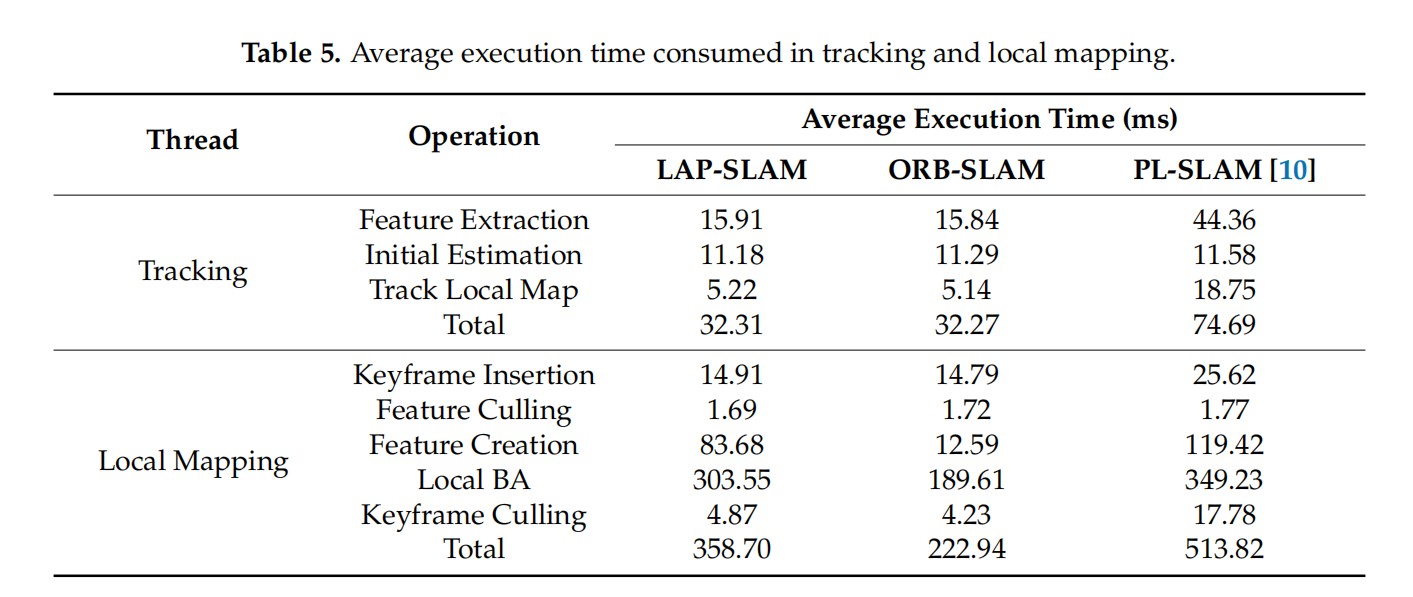
https://github.com/YoYo000/MVSNet

### LAP-SLAM

* 算法概述：基于线辅助点的单目slam，在orbslam基础上，修改了匹配原则，加快匹配速度
* 技术特点：

共线关系主要是基于不同帧之间的匹配点来建立的，它有两种功能：line-assisted bundle adjustment and line-assisted global.

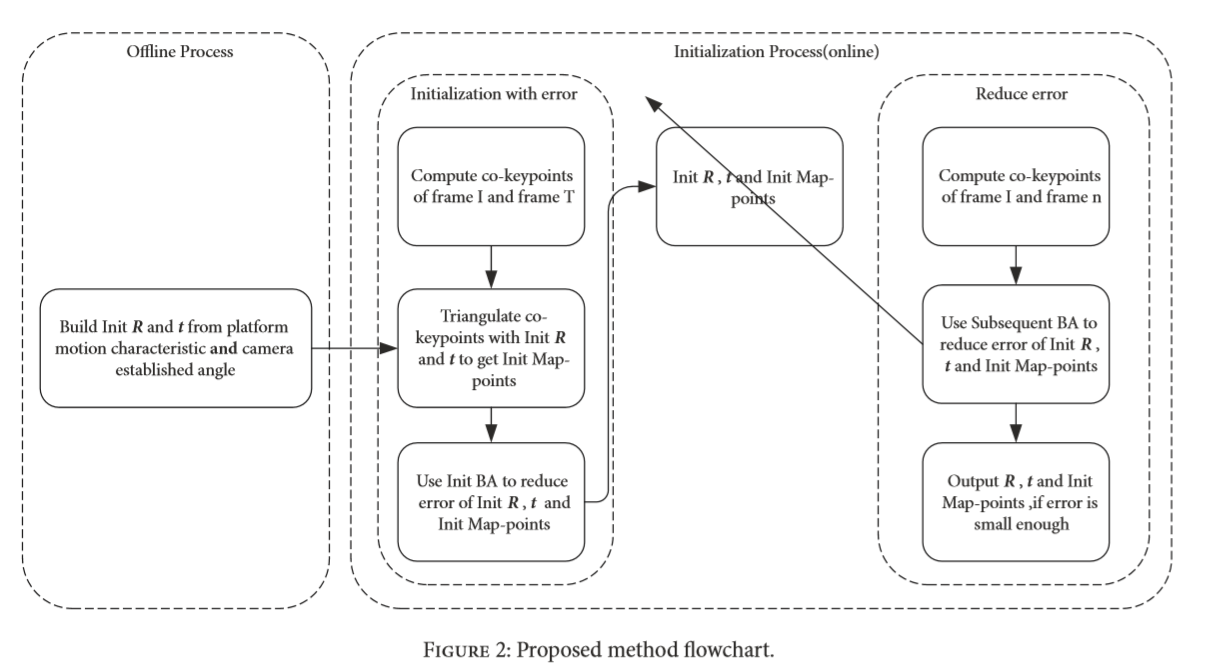
* 技术缺陷：在线条丰富的场景中效果好
* 应用范围：
* 算法效率：



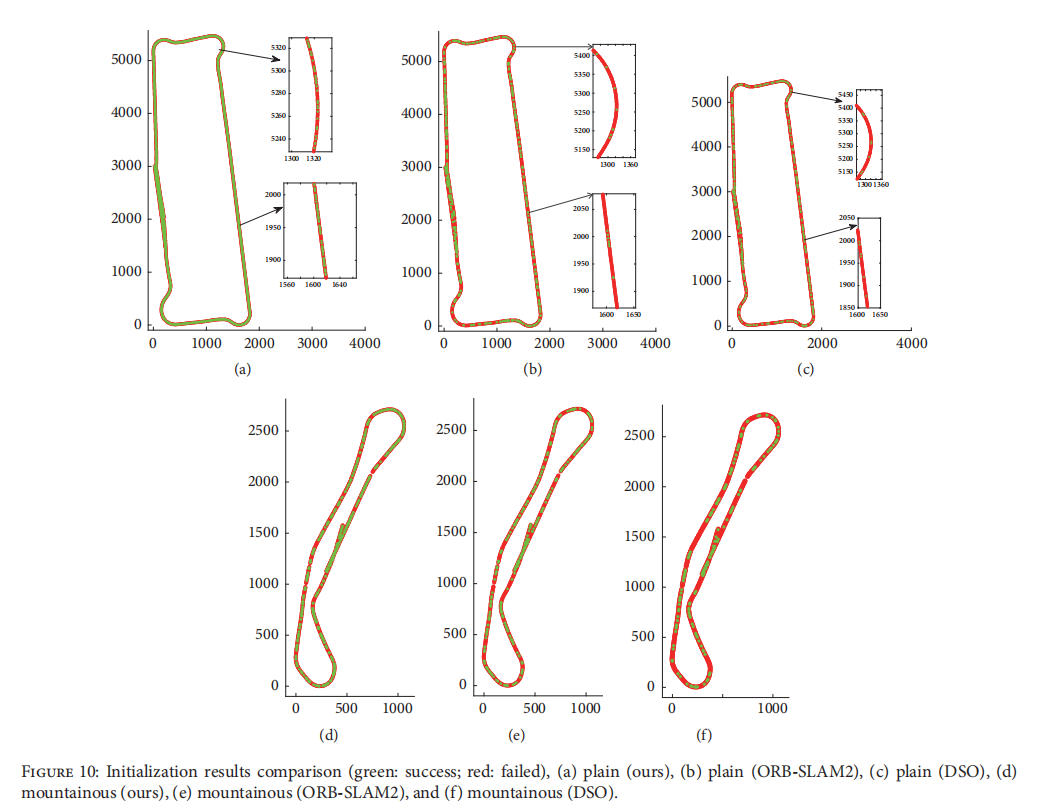
* 论文出处：<https://doi.org/10.3390/electronics8020243>

### Passive Initialization Method

* 算法概述：根据提出的平台运动的直线假设，从而推导出单眼SLAM的被动初始化方法。
* 技术特点：初始BA和随后的BA用于减少实际运动与所提出的假设之间的误差。 选择模拟固定翼飞机作为所提方法的测试平台。 结果表明，与ORB-SLAM2相比，单眼SLAM初始化的成功率大大提高。



* 技术缺陷：不能在以随机运动为特征的平台上不加区分地使用，例如人和动物。且该方法尚未在现实环境中进行测试，将在未来的工作中予以纠正。
* 应用范围：该方法仅在具有强运动特性的平台上有效
* 算法效率：



* 论文出处：https://doi.org/10.1155/2019/8176489

### SPM-SLAM

* 算法概述：IMU和编码器作里程计，视觉二维码作观测器，多传感器融合。
* 技术特点：二维码的边长已知，根据三角形相似原理则很容易求出特征点（二维码角点）的深度。一是在相机出现盲区时利用IMU和编码器进行短距离的位姿估计，二是利用多传感器信息融合在一定程度上提升机器人定位的精度。
* 技术缺陷：完全依赖二维码，对环境的依赖性较高，所以有必要加入一些其他的传感器，比如IMU（计算偏航角）和编码器（计算位移）。对于避障问题，在静态环境中可以利用轨迹规划，而在动态环境中可能还需要借助其他的避障传感器。
* 应用范围：可以布置二维码的室内
* 算法效率：



* 论文出处：https://doi.org/10.1016/j.patcog.2018.09.0

### 轻量级可扩展视觉惯性运动捕捉系统

* 算法概述：一个轻量级和可扩展的视觉惯性运动捕获系统。
* 技术特点：

1. 利用增量式平滑过的因子图，对视觉惯性进行了严格的联合优化。构造了一个基于Apriltag的标记因子。
2. 该系统对运动模糊、光照条件变化具有较强的鲁棒性，优于基于滤波的方法，更适合于大型环境（室内/室外）。
3. 我们提出了一种使用IMU测量和基准的中间故障恢复机制。

* 技术缺陷：
* 应用范围：可以布置标记的大型环境
* 算法效率：
* 论文出处：<https://doi.org/10.1007/s10514-019-09834-7>

### CoSLAM

* 算法概述：2019年提出的Co-VSLAM，使用深度和视图间编码方法扩展了我们之前的二进制特征编码框架。基于ORB-SLAM2的协作映射方案，其中多个SLAM映射是并行构建的，并且可以在检测到映射之间的重叠时合并。我们将两种方法结合到一个系统架构中，其中计算要求高的可视SLAM系统在中央处理节点上运行，并且只有视觉提示在客户端代理处被提取和压缩。我们在KITTI数据集上的编码效率，时序和绝对轨迹误差方面评估了我们的方法，
* 技术特点：

1. 扩展了现有的本地二进制特征编码框架[1]以包括来自RGB-D相机或立体声特征匹配的额外深度信息，以允许公制比例的视觉冲击。
2. 采用立体特征编码模式，利用空间相关性，完成了两种视点特征的编码框架。
3. 为了在协作场景中演示编码，我们实现了一个集中的可视化SLAM系统。

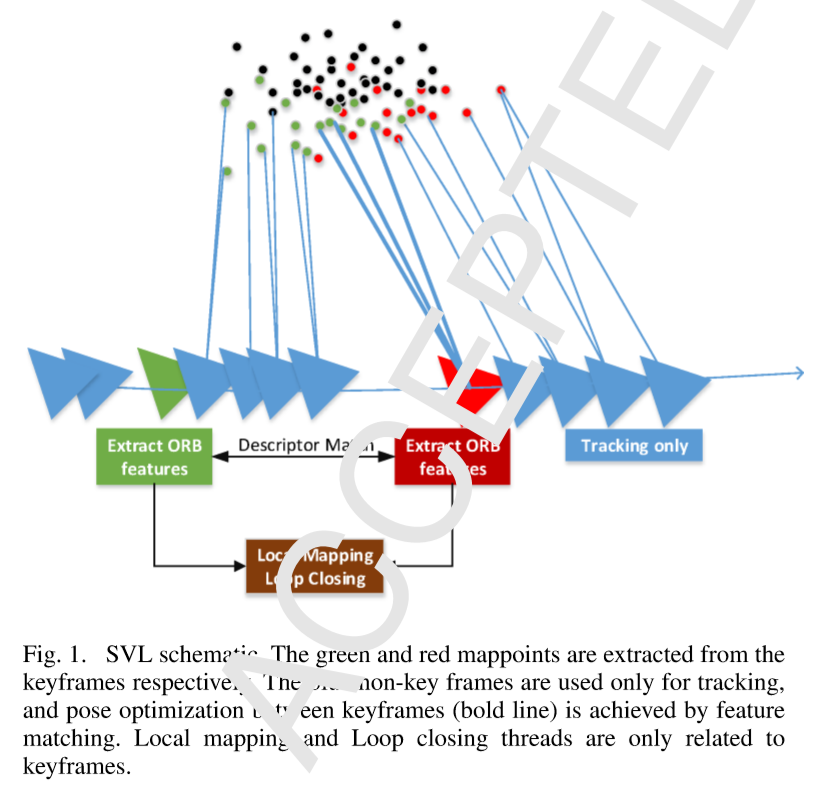
* 技术缺陷：还是基于orbslam的方法，受限制于数据传输
* 应用范围：硬件条件和网络条件合适的场景
* 算法效率：与独立映射相比，数据速率大幅降低70.8％，ATE降低53.7％ 。因此，解决Co-vSLAM设置中数据效率的需求来缩小差距。当从本地视觉测距仪与中央服务器交换关键帧时，也可以使用特征压缩。将这种轻量级里程计集成到客户端以用于本地控制循环留待将来工作。



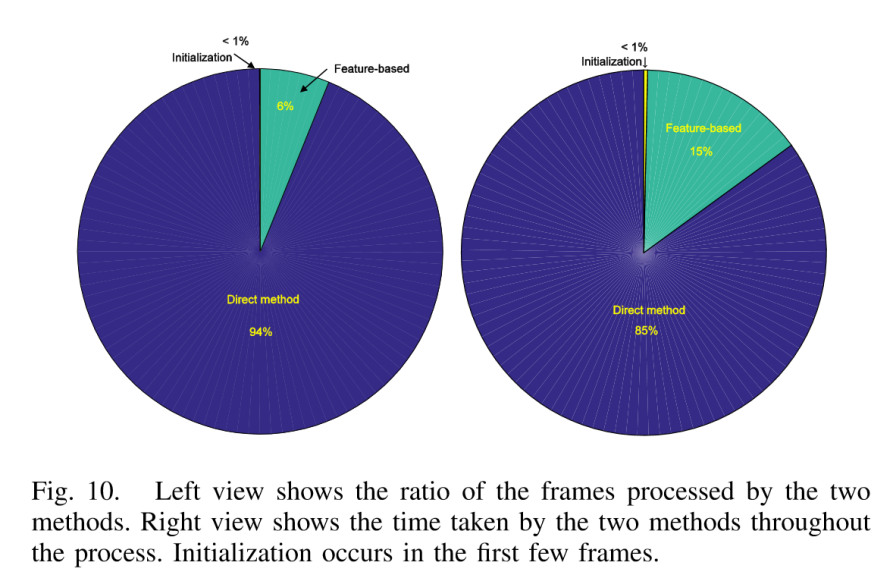
* 论文出处：[sci-hub.se/10.1109/lra.2018.2878920](https://sci-hub.se/10.1109/lra.2018.2878920)

### Svl-SLAM

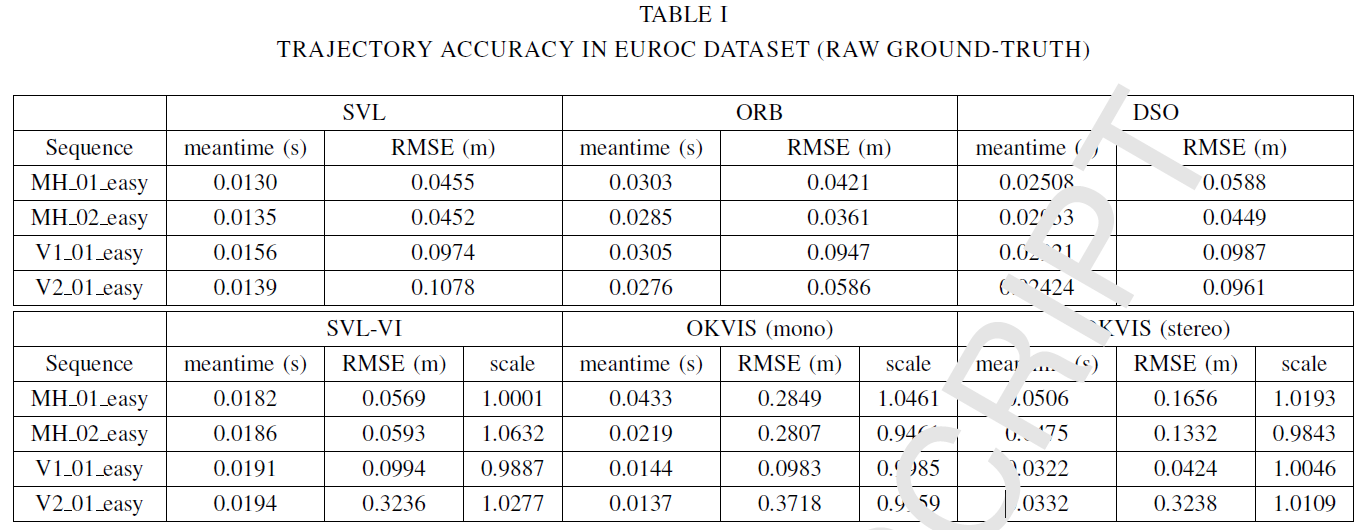
* 算法概述：semi-direct, visual and loop closure detection 的缩写，带有闭环检测的半直接SLAM系统
* 技术特点：系统根据场景的变化确定帧是否是关键帧，提取关键帧中的特征，并使用直接方法跟踪非关键帧中的这些特征。因此，在非关键帧中不再需要提取特征和匹配描述符。



SVL可以被认为是ORB和SVO的组合。 关键帧采用ORB方法，非关键帧采用SVO。 在实际实现中，改进了原始模型的方法，以完成两种方法的组合。



* 技术缺陷：
* 应用范围：
* 算法效率：



* 论文出处：

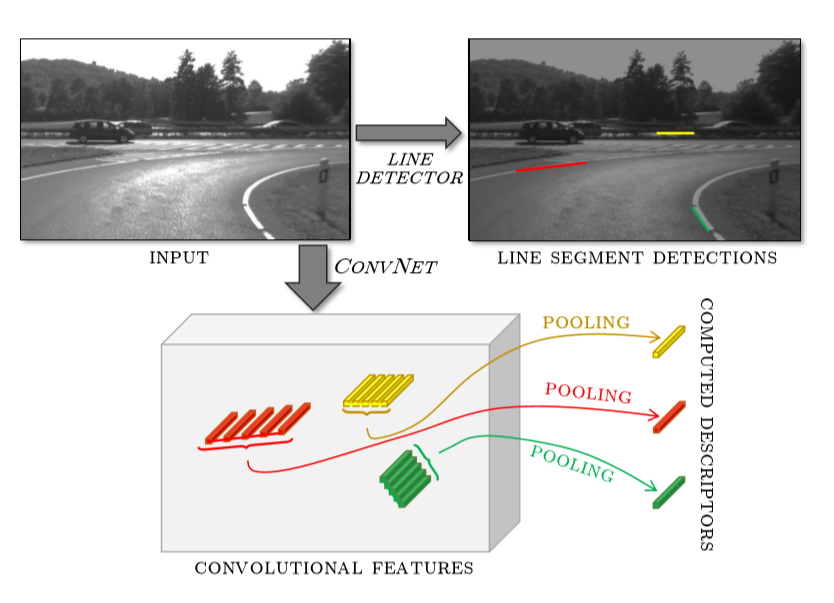
Semi-direct monocular visual and visual-inertial SLAM with loop closure detection.

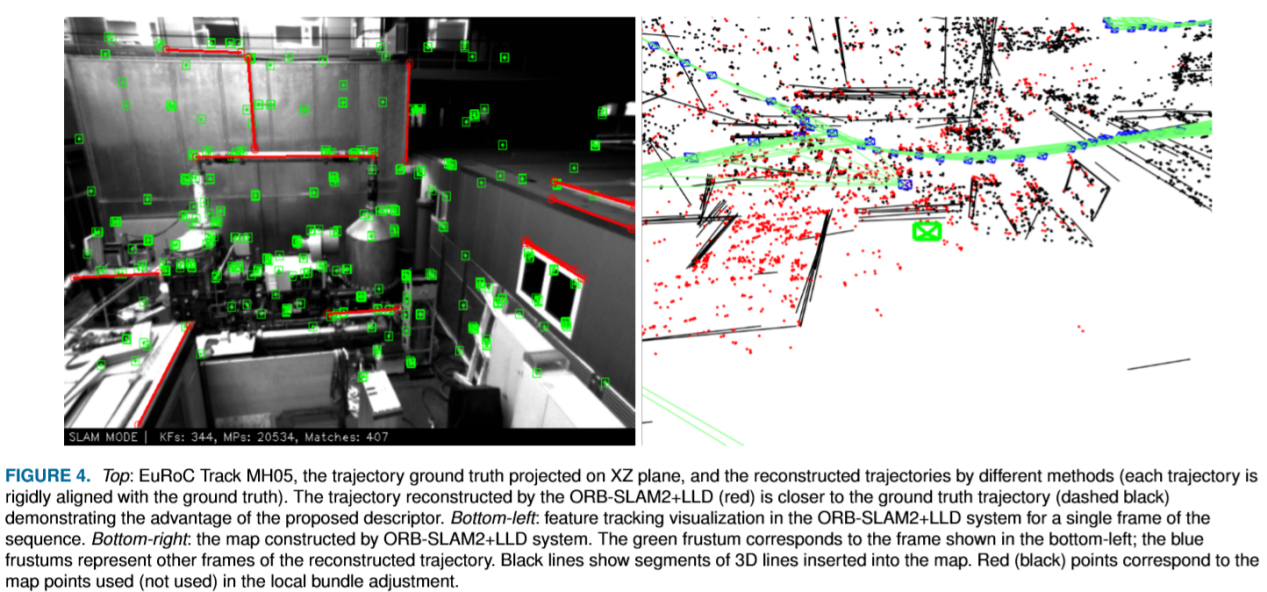
[sci-hub.se/10.1016/j.robot.2018.11.009](https://sci-hub.se/10.1016/j.robot.2018.11.009)

### LLD

算法概述：在ORB-SLAM2系统中添加线条。

技术特点：与手动设计的线段描述符不同，使用深而轻量级的完全卷积神经网络之上构建可学习的描述符

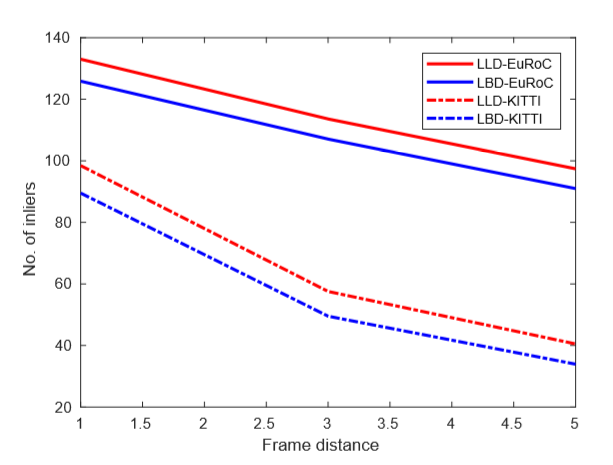


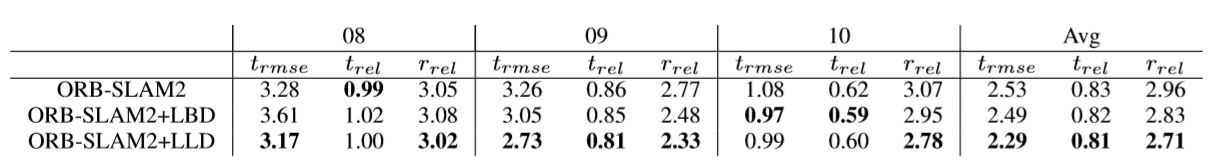


技术缺陷：

应用范围：

算法效率：





论文出处：https://ieeexplore.ieee.org/abstract/document/8651490