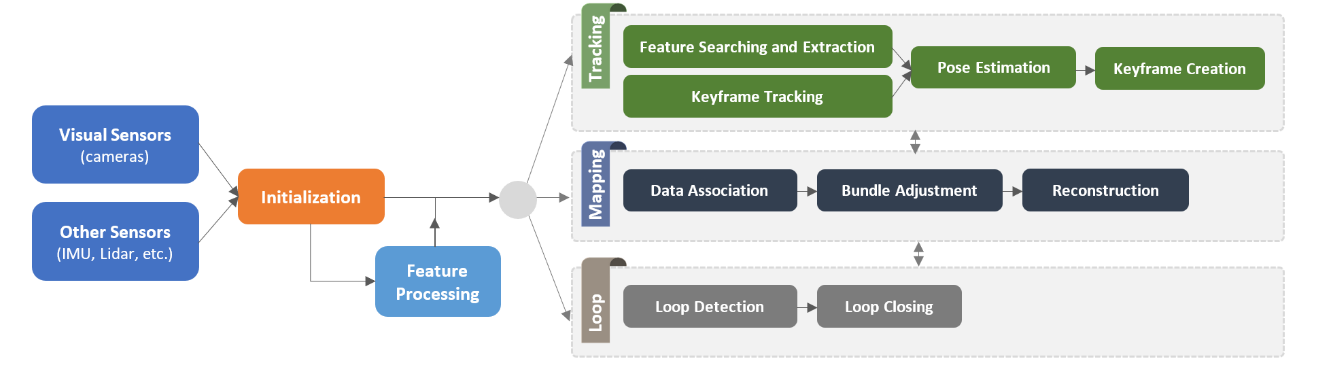
# 视觉slam综述

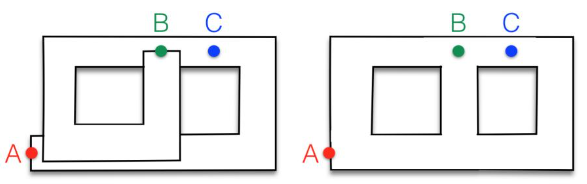
同步定位和建图（Simultaneous Localization and Mapping，SLAM）可以使用各种传感器（激光雷达传感器、视觉传感器、惯性测量单元等）从环境中收集数据。其中视觉传感器包括单目、立体、事件、RGB-D相机，使用视觉传感器的slam被称为vslam，具有硬件成本低、目标检测和跟踪更直观、可提供丰富的语义信息等优点。

数据预处理：传感器数据接收及同步

视觉里程计：位姿估计

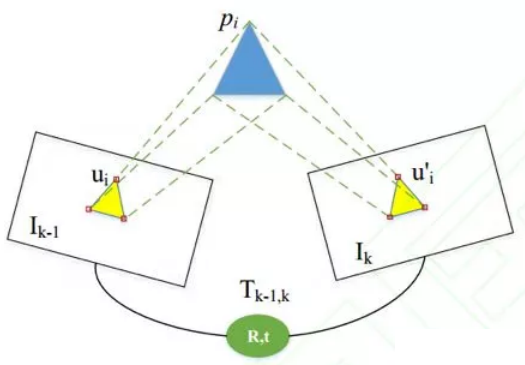
后端优化：接受不同时刻视觉里程计测量的相机位姿，以及闭环检测的信息，对它们进行优化，得到全局一致的轨迹和地图

闭环检测 ：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

建图：根据估计的轨迹，建立与任务要求对应的地图

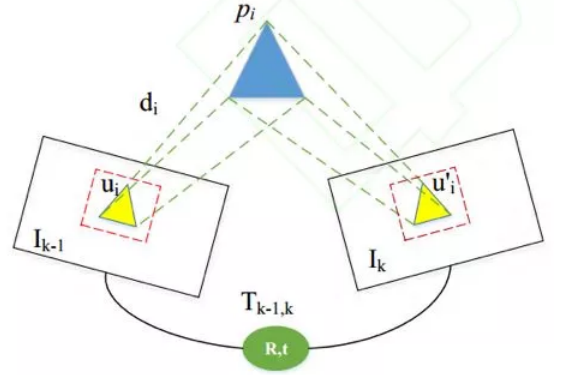
vlsam可分为直接法和间接法：

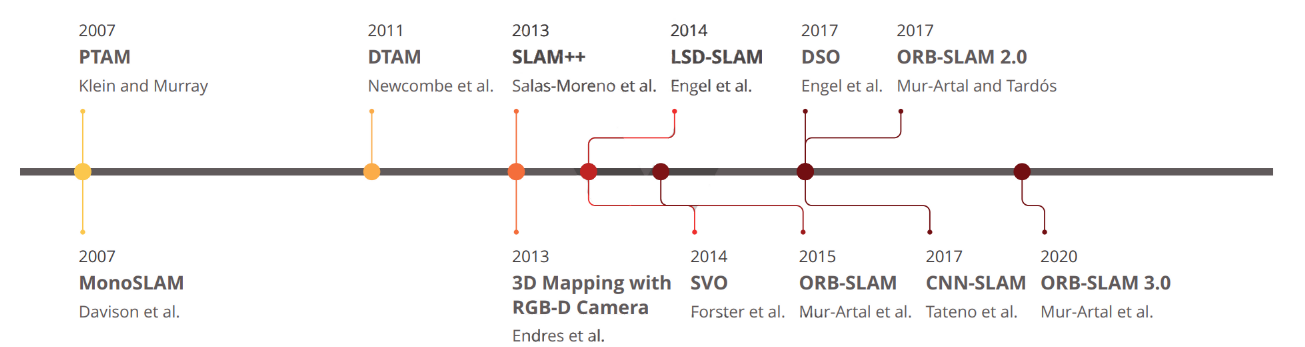
间接法首先通过对场景的预处理从纹理中提取特征点，再在连续帧中匹配特征点的描述符来跟踪特征点（也可提取稠密的光流计算光流向量）。间接法由于需要提取和匹配特征点，所以需要更多的算力，但它在光度测量上的鲁棒性相对较高。间接法的主要优化方向是减少几何误差：



其中ui为Ik-1中任意像素点，它投影到空间点的坐标为Pi， u′i是 Pi 投影到Ik上的坐标。之后利用特征点来估计周围物体的三维模型和相机运动轨迹。

直接法跳过预处理，直接从视觉传感器的数据估计相机运动（亮度颜色变化）。间接法在纹理较少的环境下表现较好且不需要提取特征。间接法的主要优化方向是减少光度误差：



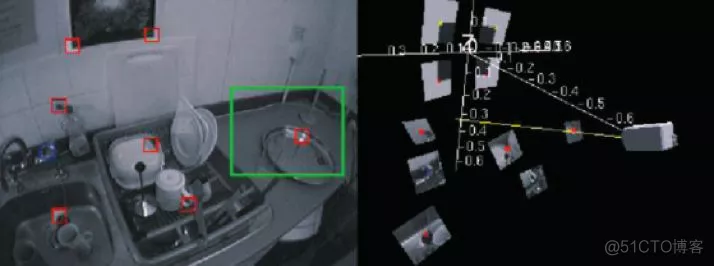


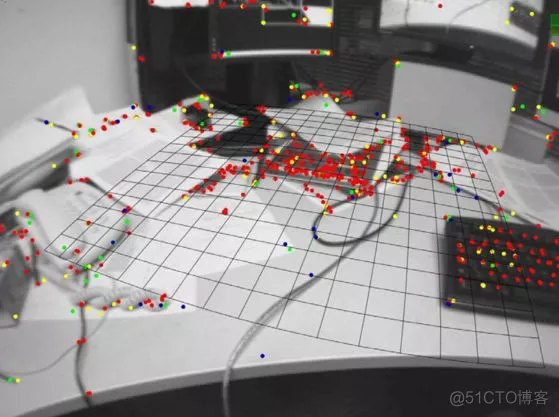
经典slam框架

MonoSLAM 2007

贡献：利用扩展卡尔曼滤波算法估计像机运动

不足：没有全局优化和循环闭合检测模块，重建的地图只包括地标

PTAM 2007

贡献：多线程并行处理提高了实时性

不足：设计复杂且第一阶段需要用户输入

DTAM 2011

贡献：可以稠密建图

不足：实时运行需要较高的运算成本

3D Mapping with RGB-D Camera 2013

贡献：可以低成本的实时执行

不足：无法在无特征环境中准确得到结果

SLAM++ 2013

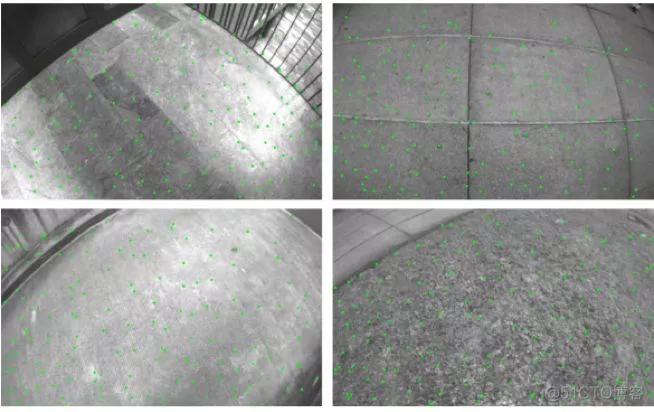
贡献：可以形成位姿图（节点表示位姿估计值，边表示两节点间的相对运动）

不足：需要提供物体3D模型

SVO 2014

贡献：结合了间接法和直接法

不足：使用了短期数据关联，无法进行循环闭合检测和全局优化

LSD-SLAM 2014

贡献：可以利用位姿估计模块重建大尺寸地图，并具有全局优化和闭环检测

不足：初始化阶段需要所有点在一个平面上

ORB-SLAM 2015 & ORB-SLAM2.0 2017 & ORB-SLAM3.0 2020

贡献：可以利用Oriented FAST and Rotated BRIEF特征进行高性能定位

不足：重建的地图比例未知，无法用于自主导航，不能在缺少纹理或纹理重复的环境中运行

CNN-SLAM

贡献：结合深度神经网络进行单目实时稠密建图

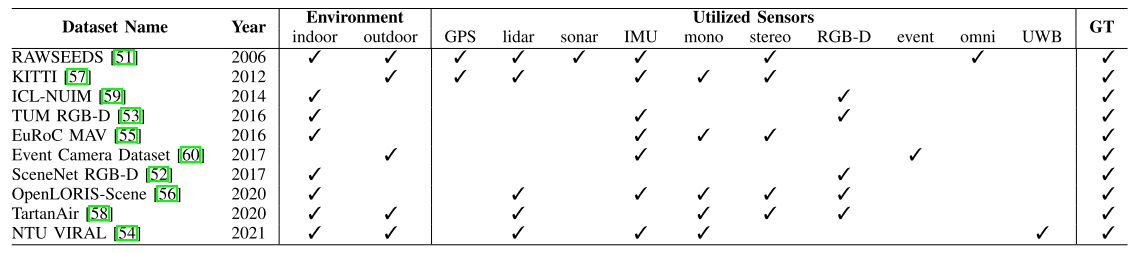
DSO

贡献：将直接方法和稀疏重建相结合，以提取图像块中的最高强度点

不足：没有回环检测

Vslam评估标准

传感器：视觉传感器 单目、立体、事件、RGB-D相机

其他传感器 激光雷达、IMU

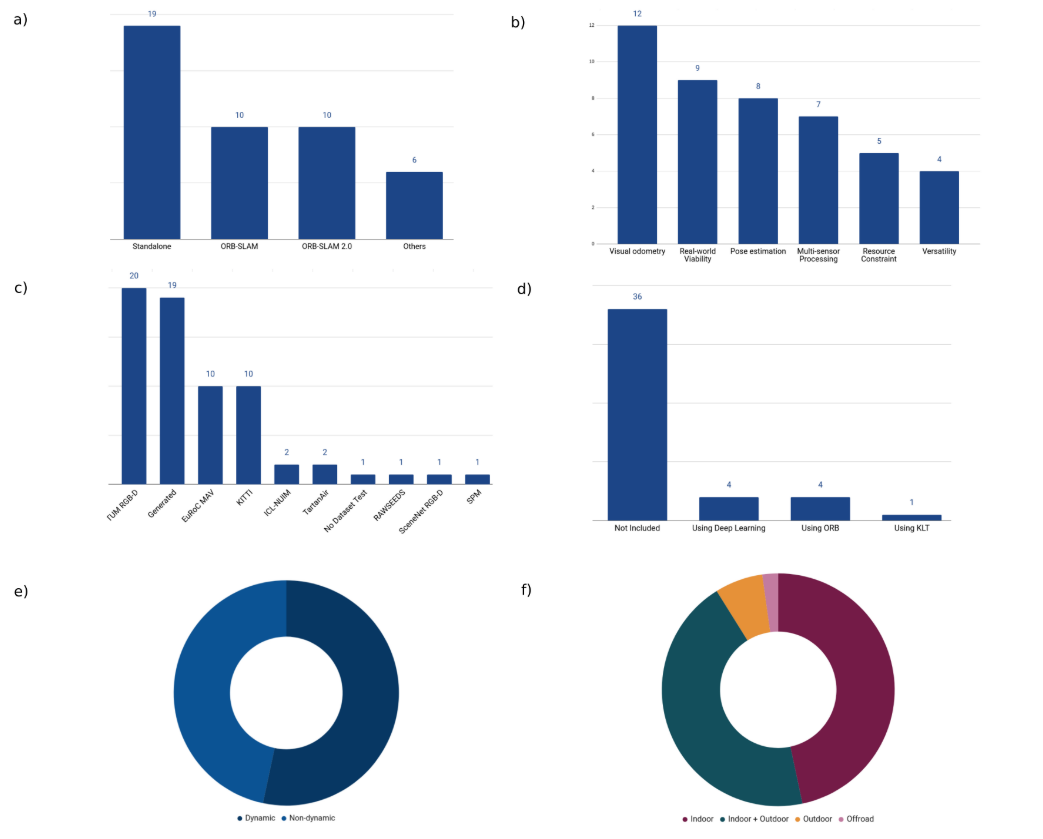
目标环境：静态、动态、室内、室外

特征提取：SIFT、SURF、FAST、BRIEF、ORB

数据集：

语义：机器人需要语义信息才能理解周围的场景并做出更优决策

Vslam研究方向

1. 多相机和多传感器
2. 位姿估计（使用点/线数据、使用额外特征、深度学习）
3. 动态环境
4. 资源限制（部署在边缘设备）
5. 通用性（便于二次开发或调用）
6. 视觉里程计（尽可能高精度的定位）
7. 大多数论文都独立完成整个SLAM框架，有部分基于ORB-SLAM和ORB-SLAM2框架
8. 研究方向排序：视觉里程计、动态环境、位姿估计、多传感器、资源限制、通用性
9. 大部分论文使用数据集TUM RGB-D、Generated、EuRoC MAV、KITTI
10. 大部分论文不包含语义信息
11. 超过一半的方法可以在有挑战性的环境下（动态、低照度、缺少纹理）工作
12. 大部分方法都是在“室内环境”或“室内和室外环境”中工作，小部分方法仅能工作在“室外环境”