

视觉SLAM关键方法总结

最近在做基于激光信息的机器人行人跟踪发现如果单独利用激光信息很难完成机器人对行人的识别、跟踪等功能，因此考虑与视觉融合的方法，这样便可以充分利用激光雷达提供的精确位置信息及视觉提供的丰富纹理、颜色等场景信息。以下是最近调研视觉SLAM中的实现方法的总结，包括三方面内容：姿态计算、闭环检测、BA优化。

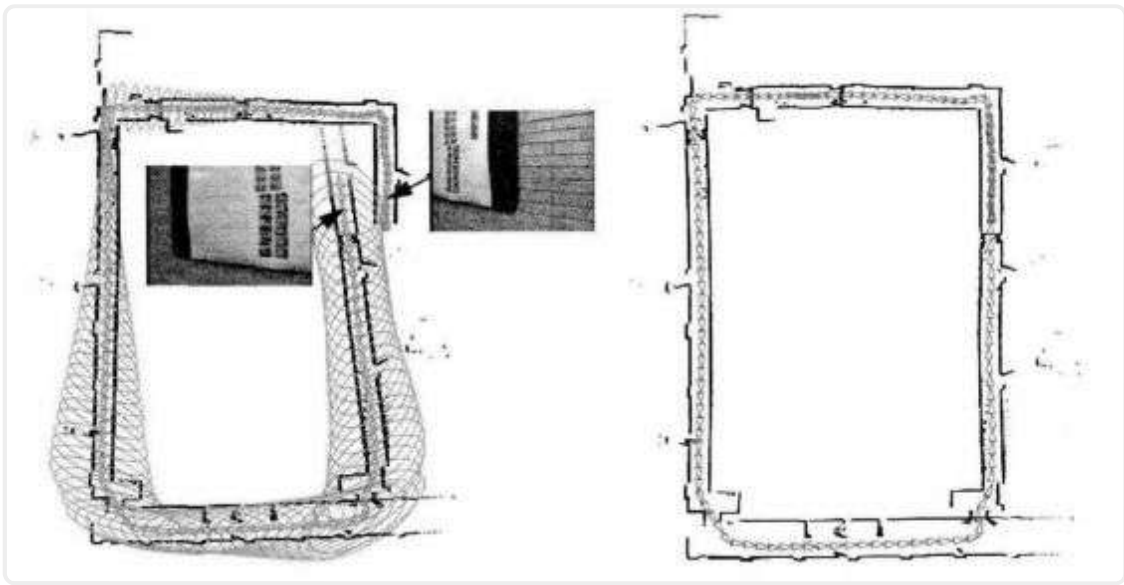
姿态计算

一、通过提取图像的特征描述子，如ORB、SURF和SIFT等特征描述子，然后通过RANSAC算法进行图像匹配去除匹配点中的外点，再将二维点对映射到三维之后，便可以利用PnP或ICP算法计算相机位姿。基于特征提取的位姿计算算法对场景有一定的要求，在无纹理场景会出现位姿计算失败的情形。

二、直接图像匹配方法：直接图像匹配并不对图片进行特征提取，核心思想是在旋转坐标系下，基于相机一致性的假设，在相机的刚体变换已知的情况下，利用相机变换矩阵将目标图片投影到当前图片上，其像素之间的差异应该最小，将姿态计算转换为加权最小二乘问题。直接图像匹配算法的计算效率很高，不依赖GPU，具有很高的理论和商用价值。

闭环检测

闭环检测算法指的是通过检测算法检测出之前访问过的场景。如图1所示，当机器人在移动过程中，特别是在探索大面积的场景时，由于模型的不确定性以及设备的噪声，不确定性会逐渐增长。通过引入闭环检测技术，识别出历史访问过的场景以增加位姿之间的约束，可以很好的减少这种不确定性。



闭环检测前后对比(a) 累计误差导致的构图错误, (b)闭环约束后的地图

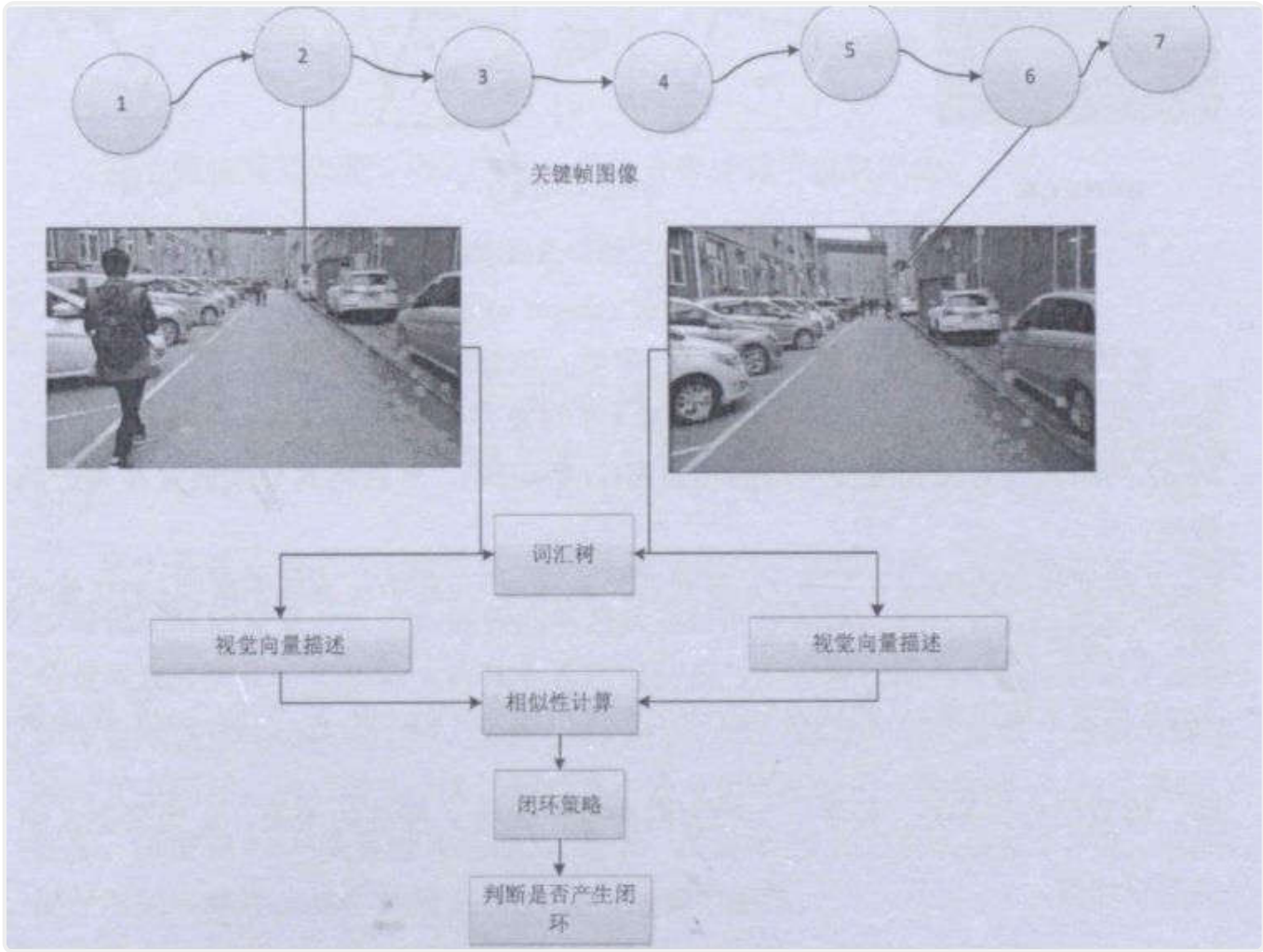
闭环检测方法有：

一、最简单的闭环检测算法是将新检测出来的关键帧和过去所有的关键帧一一进行比较，虽然这种方法能比较好的检测出当前场景是否在之前出现过，但是在大规模场景下，机器人往往有成千上万个关键帧，这种方法检测效率及其底下，不能再实际场景中使用。

二、通过将图像中的特征和整个地图中的路标对应起来，然后建立一个所有路标的数据库，通过使用路标分类器来加快场景识别速度。比如讲所有路标构建成一棵KD-Tree，当新添加一帧关键帧时，将关键帧在KD-Tree中进行检索，从中检索之前出现过的场景。

三、基于视觉词袋的闭环检测方法，这种方法通过将特征描述子抽象成词汇，通过TF-IDF方法识别出现过的场景。如图2所示，是闭环检测中基于词袋模型流程图。使用视觉词袋的方法

效率很高，可用在大规模地图的创建上。



BA优化

一、问题阐述：同时对三维点位置和相机参数进行非线性优化。

观测点

三维点坐标 $X_i = (X_i, Y_i, Z_i)^T$

投影点

相机参数 $C_j = (f_j, k_{ij}, k_{2j}, R_j, t_j)$

$\chi_{ij}=1$ 表示第 i 个点在第 j 个相机中可见

$\theta \in R^{9m \times 3n}$ 高维空间的非线性优化

$$g(\theta) = \frac{1}{2} \sum_i^n \sum_j^m \chi_{ij} \| \hat{u}_{ij} - u_{ij}(C_j, X_i) \|^2 = \frac{1}{2} \sum_i^n \sum_j^m \chi_{ij} e_{ij}^2, \quad \theta = (C_1, \dots, C_m, X_1, \dots, X_n)$$

The diagram shows a 3D point X being projected onto two camera planes. The first camera, O_1 , has a principal point c_1 and a point x_1 on its plane. The second camera, O_2 , has a principal point c_2 and a point x_2 on its plane. The projection of X onto the first plane is x_1 , and the projection onto the second plane is x_2 .

BA原理

二、LM法的原理与优势：

原理：是一种“信赖域”的方法，当收敛速度较快时，增大信赖域使算法趋向于高斯牛顿法；当收敛速度较慢时，减小信赖域使算法趋向于最速下降法。优势：速度快；可以在距离初始值较远处得到最优解。

SLAM优化算法对比

算 法	缺点	优点
K F/ E K F	假设噪声为高斯分布，在高维状态空间时计算效率较低，不适合大规模场景的地图构建	原理简单，小场景下收敛性好。
P F	会不可避免出现粒子退化现象，每个粒子需包含整张地图信息导致不适合大场景地图构建	将运动方程表示为一组粒子，通过蒙特卡罗方法求解，能够处理非线性非高斯噪声。
E M	计算量大，不能用于大规模场景	有效解决了数据关联问题
图 优 化	对闭环检测算法的要求严格	出现多种图优化框架，能够有效解决滤波器算法的缺陷，能用于大规模场景的地图创建

推荐阅读

[从零开始一起学习SLAM | 为什么要学SLAM?](#)

[从零开始一起学习SLAM | SLAM有什么用?](#)

[从零开始一起学习SLAM | C++新特性要不要学?](#)

[高翔：谈谈语义SLAM/地图](#)