

可变场景SLAM调研提纲

1 动态环境下的SLAM系统

动态环境下的SLAM，主要研究场景中有运动物体时，怎样进行稳定的SLAM求解。一般研究的核心问题在于对画面中运动部分和静止部分进行区分。

1.1 基于运动分割

这类方法最本质的思路是去除运动区域，构建静态地图进行定位。一般情况下将运动分割作为二分类问题，把图像分成静态区域和动态区域，其核心在于运动区域的提取。在静态区域占画面绝大部分的情况下，通过鲁邦的统计方式，比如RANSAC，可以结合几何和其它的约束条件滤去运动部分；一些方法利用了外部的传感器，比如IMU，来初始化相机位姿，以提高运动分割和相机位姿求解的鲁棒性；还有一些方法利用深度学习在数据的高阶特征提取上的优势，将运动分割作为分类问题进行学习。

1.2 基于运动物体跟踪

这类方法的核心是提取并跟踪各个运动物体。在上一问题中，我们知道，如果有准确的先验（位姿或者分割），那么运动分割和位姿求解这个联合优化问题的估计结果将会更加鲁邦。基于此，一些方法将这一问题作为特征对应的聚类问题，即找出不同运动模式的物体分别进行跟踪，利用跟踪的结果作为下一帧分割和位姿估计的先验，建立时序上的约束。一类方法将问题看做子空间聚类问题，其基本假设在于每一个独立的刚体运动可看做一个线性子空间；另一类方法将问题看做统计模型选择问题，其可看做是上一节运动分割的拓展，通过不断采样、建立内点集求解位姿、再在外点集进行下一次采样的方式，反复迭代找到 n 个运动模型。基于聚类得到的各个刚体运动区域，可以通过滤波的方式或三角化的方式进行跟踪，建立运动物体的时续联系

1.3 联合求解运动分割、位姿估计、和三维结构恢复

这一类问题考虑多个独立的刚体运动物体（multi-body SfM）或是局部刚体运动（non-rigid SfM, 主要包括一些dependent motion比如articulated motion或是piece-wise rigid motion）。这类方法主要基于因子化（factorization）来同时恢复运动和结构。 p 个特征点在 f 帧的所有二维轨迹可以看做 $2f \times p$ 的矩阵，这个观测矩阵可以因子化为 $2f \times 4$ 的运动矩阵和 $4 \times p$ 的结构矩阵。那么问题就转化为求解的就是 n 个独立运动的观测矩阵，每个观测矩阵可以分解为运动矩阵和结构矩阵相乘的形式，从而将非刚体问题转化为多个刚体SfM问题进行求解

2 长时变化环境下的定位

长时变化环境指的是环境在一定时间之后，场景可能会发生一点范围的变化，核心在于维护可动态调节的鲁棒地图以及在地图发生微调时进行定位。

2.1 物体为中心的方法

这类方法主要检测并删除或检测并调整地图中发生运动的物体。和基于运动分割的SLAM问题类似，一些方法基于静态世界假设，将场景中发生运动的区域剔除，只依赖静态地图进行定位，在方法上没有过多的差异。然而在实际应用中，静态背景区域或许在一些情况下只有很少的有用信息部分能被观测到，这限制了动态场景下定位的性能。

2.2 分别构建静态和动态物体地图

这类方法通过维护动态地图，运动区域的检测和静态地图的构建更加可靠，这类方法和基于dynamic object segmentation的SLAM方法相类似。但是，由于定位依赖静态地图，object-centric的局限性并没有被消除。

2.3 4D地图

这类方法建模了环境在时序上的变化情况，并且不考虑引起变化的原因。换句话说，地图被描述成状态随时间的转移过程。当然，由于加入了时间维的信息，地图的数据量和定位的计算复杂度是这类方法主要考虑的因素，其背后是long-term operation导致的scalability的问题。维护4D地图很好的保证了整个环境在时间上的consistency，同时，由于完整地反映了环境在不同时间的变化情况，4D地图能够很好地给出下一时刻环境的预测，为定位提供可靠的先验信息