

初始化

- 单目相机：单目相机需要两帧之间求相对位姿（分解矩阵H或F），然后三角测量出点云进行初始化。
- 双目相机：双目相机生成Frame对象时，会对左右图像提取特征，然后对左图像的每个特征点在右图像上进行立体匹配，然后就得到了左图中每个特征点的3D点。所以双目相机只需要一帧就能够进行初始化了。
- 深度相机：这个初始化更简单了，自带深度信息，也是一帧就初始化成功了。

note：不管是初始化还是追踪，右图像除了用来生成左图特征点的深度外，就没什么作用了

跟踪

- TrackWithMotionModel假设相机做恒速运动，根据前一帧的位姿加上恒速位移推算出当前帧的位姿，之后遍历前一帧的特征点，根据推算出的当前帧位姿，将特征点对应的3D地图点投影到当前帧，在投影点附近搜索特征点进行匹配，从而缩小特征点匹配搜索范围，加快匹配速度，这一过程通过SearchByProjection()函数实现。最终得到一组匹配好的3D地图点-2D特征点。然后进行Bundle Adjustment，最小化重投影误差来进行当前帧位姿优化，得到初始位姿
- TrackReferenceKeyFrame是在运动模型还未建立或者刚完成重定位时，通过BoW进行当前帧与参考关键帧特征点匹配，这一过程通过SearchByBoW()函数实现。得到一组3D地图点-2D特征点匹配后，与TrackWithMotionModel一样，优化3D-2D的重投影误差来得到初始位姿，这里在优化前将上一帧的位姿作为当前帧的初始位姿，可以使得优化收敛加快。
- Relocalization是在跟踪丢失时，将当前帧与候选关键帧通过BoW进行特征点匹配，之后通过EPnP算法估计位姿，和前两种方法一致，再通过最小化重投影误差对位姿进行优化。

note：单目相机与双目/深度相机的跟踪流程都是一样的，只不过搜索匹配时，阈值有所差别罢了。不同点在于优化位姿时，单目相机最小化的是3D-2D重投影误差；而双目/深度相机除了优化3D-2D重投影误差之外，还会优化mvRight，之所以优化该变量是因为双目/深度相机多了一个深度观察值。

以上三种方法都是基于两帧进行匹配，优化3D-2D重投影误差得到当前帧的初始位姿。这样得到的初始位姿不是可靠的，其只利用了两帧数据的信息，如果前一帧质量太差，得到的位姿可信度低。因此为了利用更多的信息，需要进一步将当前帧与局部地图进行匹配和优化，也就是TrackLocalMap()。

进一步优化位姿

TrackLocalMap的作用是使得当前帧能够获取更多的2D-3D点对应的，优化2D-3D重投影误差，从而得到更加稳健的重投影误差。

局部地图包含了局部关键帧（与当前帧帧共视关系高的帧）和局部3D点（局部关键帧所对应的3D点）。将局部地图的3D点（世界坐标系）根据之前追踪得到的位姿投影到当前帧上，然后在投影点的周围根据3D点的特征描述子进行匹配点的搜索，从而得到2D-3D匹配。最后根据这些2D-3D点进行位姿的优化