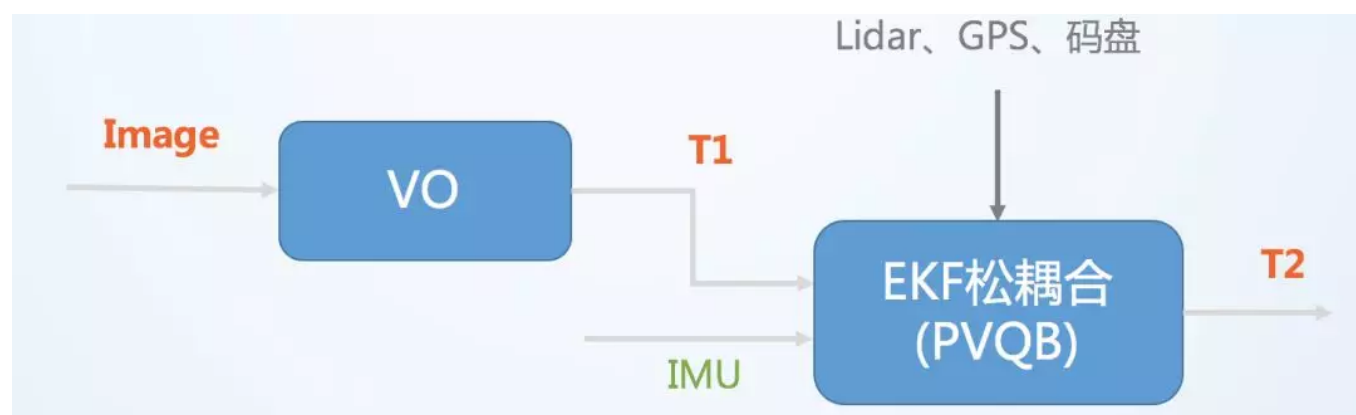


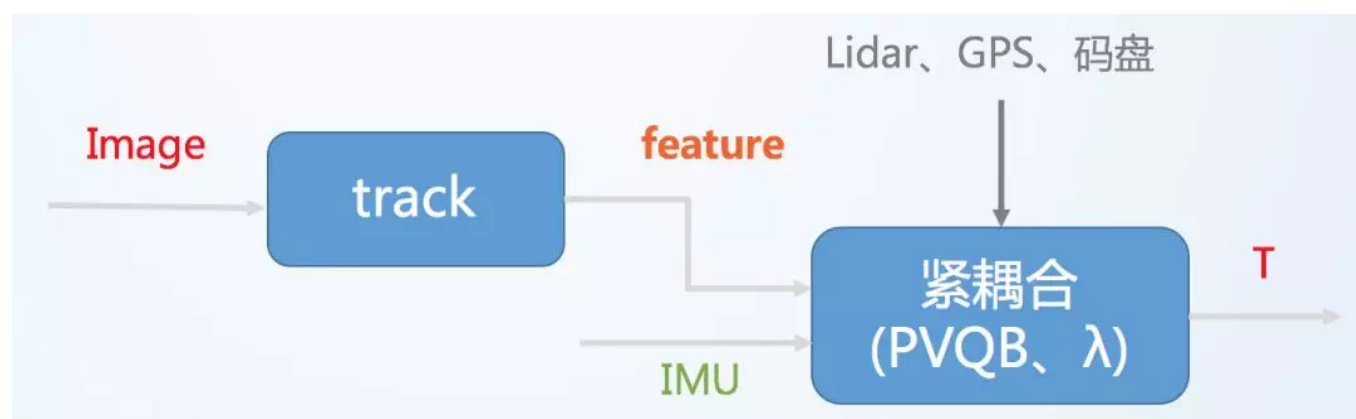
计算机视觉方向简介 | 视觉惯性里程计(VIO)

VIO-SLAM

Visual-Inertial Odometry (VIO) 即视觉惯性里程计，有时也叫视觉惯性系统 (VINS , visual-inertial system)，是融合相机和IMU数据实现SLAM的算法，根据融合框架的不同又分为**松耦合**和**紧耦合**。



其中VO (visual odometry) 指仅视觉的里程计，T表示位置和姿态。松耦合中视觉运动估计和惯导运动估计系统是两个独立的模块，将每个模块的输出结果进行融合。



紧耦合则是使用两个传感器的原始数据共同估计一组变量，传感器噪声也是相互影响的。紧耦合算法比较复杂，但充分利用了传感器数据，可以实现更好的效果，是目前研究的重点。

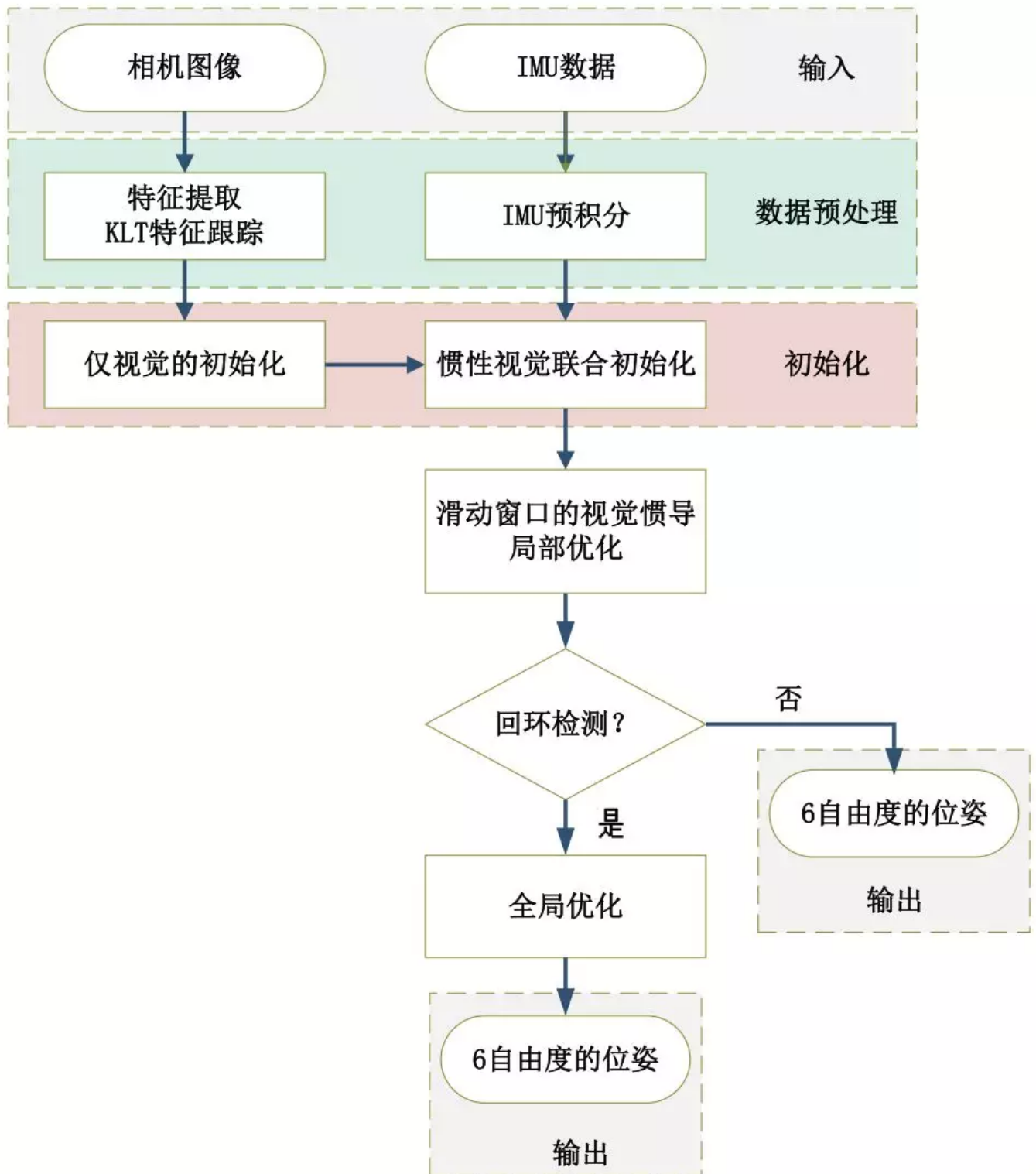
相机和IMU的缺点及互补性

Camera	单目相机无法解决尺度问题 在纹理少的区域无法工作 快速运动时图像模糊 动态场景会有错误匹配
IMU	纯IMU预测会有漂移 角度漂移相对时间 t 是线性的 位置漂移是 t^2 级别的

相机和IMU融合有很好的互补性。首先通过将IMU 估计的位姿序列和相机估计的位姿序列对齐可以估计出相机轨迹的真实尺度，而且IMU 可以很好地预测出图像帧的位姿以及上一时刻特征点在下帧图像的位置，提高特征跟踪算法匹配速度和应对快速旋转的算法鲁棒性，最后IMU 中加速度计提供的重力向量可以将估计的位置转为实际导航需要的世界坐标系中。

随着MEMS器件的快速发展，智能手机等移动终端可以便捷地获取IMU数据和摄像头拍摄数据，融合IMU 和视觉信息的VINS 算法可以很大程度地提高单目SLAM 算法性能，是一种低成本高性能的导航方案，在机器人、AR/VR 领域得到了很大的关注。

算法流程



整个流程图可以分解为五部分：数据预处理、初始化、局部非线性优化、回环检测和全局优化。

各个模块的主要作用是：

图像和IMU数据预处理：对于图像，提取特征点，利用KLT金字塔进行光流跟踪，为后面仅视觉初始化求解相机位姿做准备。对于IMU，将IMU数据进行预积分，得到当前时刻的位姿、速度、旋转角，同时计算在后端优化中将要用到的相邻帧间的预积分增量，及预积分的协方差矩阵和雅可比矩阵。

初始化：初始化中，首先进行仅视觉的初始化，解算出相机的相对位姿；然后再与IMU预积分进行对齐求解初始化参数。

局部非线性优化：对应流程图中滑动窗口的视觉惯导非线性优化，即将视觉约束、IMU约束放在一个大目标函数中进行优化，这里的局部优化也就是只优化当前帧及之前的n帧的窗口中的变量，局部非线性优化输出较为精确的位姿。

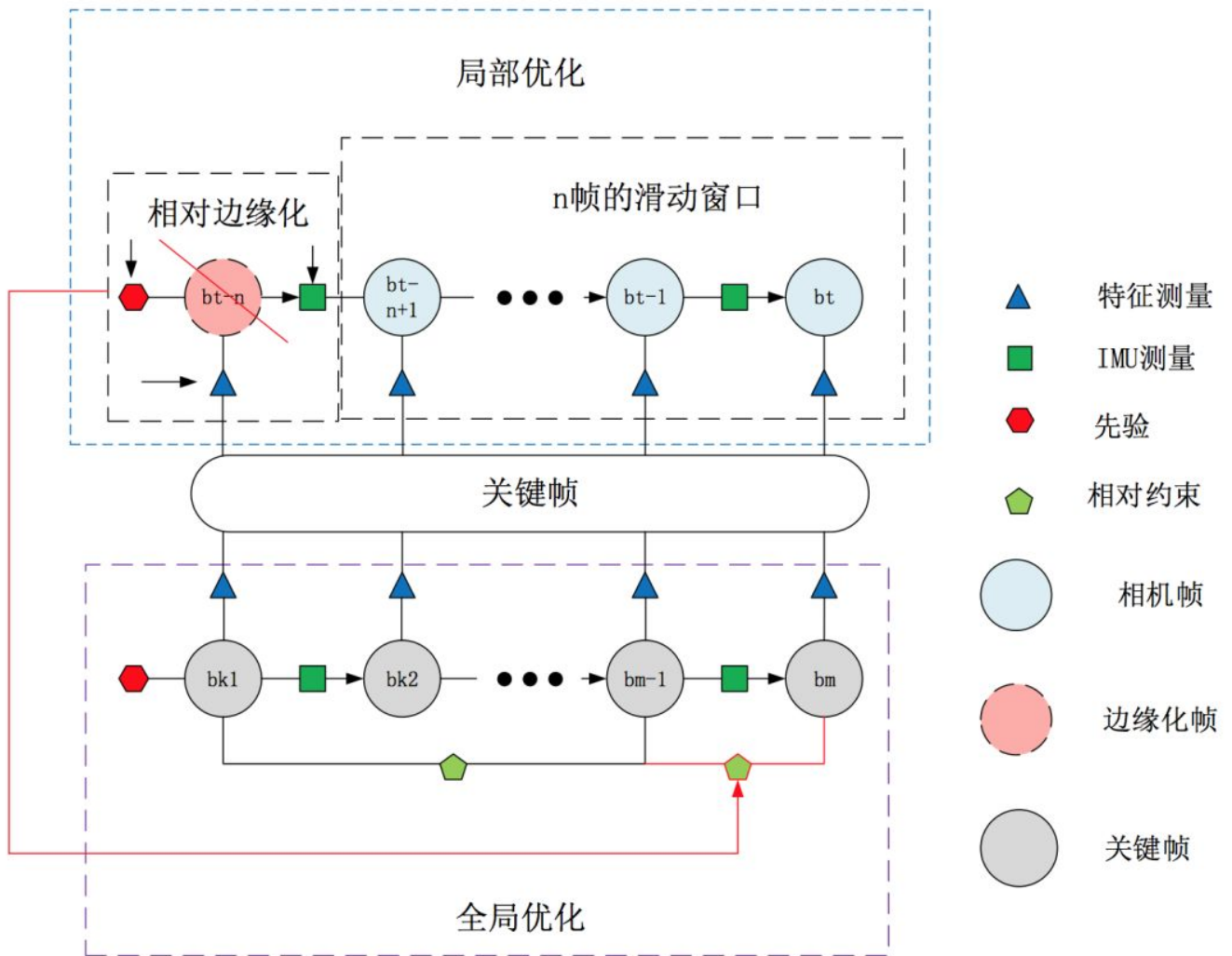
回环检测：回环检测是将前面检测的图像关键帧保存起来，当再回到原来经过的同一个地方，通过特征点的匹配关系，判断是否已经来过这里。前面提到的关键帧就是筛选出来的能够记下但又避免冗余的相机帧（关键帧的选择标准是当前帧和上一帧之间的位移超过一定阈值或匹配的特征点数小于一定阈值）。

全局优化：全局优化是在发生回环检测时，利用相机约束和IMU约束，再加上回环检测的约束，进行非线性优化。全局优化在局部优化的基础上进行，输出更为精确的位姿。

算法核心

局部优化会用到边缘化，仅用局部优化精度低，全局一致性差，但是速度快，IMU利用率高；仅用全局优化精度高，全局一致性好，但是速度慢，IMU利用率低；两者侧重点不同，所以将两者结合，可以优势互补。

因此小编设计实验采用局部优化和全局优化融合的方法。



局部优化是滑动窗口内相机帧的优化，全局优化是所有关键帧的优化，两者结合会产生边缘帧冲突的问题，因为局部优化会固定滑动窗口边缘帧，而全局优化发生回环检测的时候则会固定回环起点的帧。这里的改进就是采用相对的位姿边缘化，即边缘化以后的点是相对于它上一时刻关键帧的位姿而不是全局的位姿，这样局部优化边缘化相对位姿（关键帧），扔给全局优化整体优化。局部边缘化和全局边缘化的结合部分是关键帧。

相对边缘化可以具体解释为，相对边缘化的参考坐标系不再是世界坐标系，而是与当前帧共视且距离最近的一个关键帧的相机系（设为第 k_0 帧）。视觉约束可以表示为：

$$\varepsilon_{ij}(T_i, T_j, \rho_i) = \varepsilon_{ij}(T_{k_0}^{c_i}, T_{k_0}^{c_j}, \rho_i) = K(T_{k_0}^{c_j} T_{k_0}^{c_i-1} \frac{\bar{P}^{c_i}}{\rho_i}) - P_{uv}^{c_j}$$

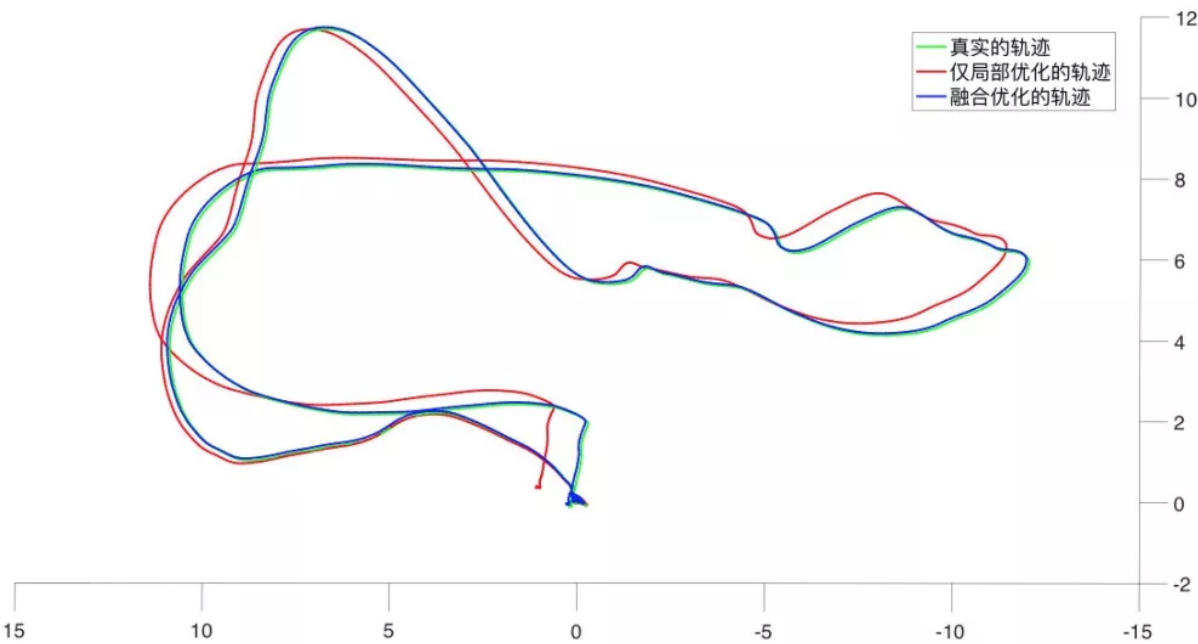
区别于绝对边缘化的视觉约束

$$\varepsilon_{ij}(T_i, T_j, \rho_i) = \varepsilon_{ij}(T_w^{c_i}, T_w^{c_j}, \rho_i) = K(T_w^{c_j} T_w^{c_i-1} \frac{\bar{P}^{c_i}}{\rho_i}) - P_{uv}^{c_j}$$

实验一：无人机数据集上的实验

数据集采用了欧盟机器人挑战数据集（EuRoC）。EuRoC 数据集使用 Asctec Firefly 六旋翼飞行器在仓库和房间采集数据，数据集中包括以20Hz采集的相机图像和200Hz的IMU数据，以及运动真值。

实验结果如下：

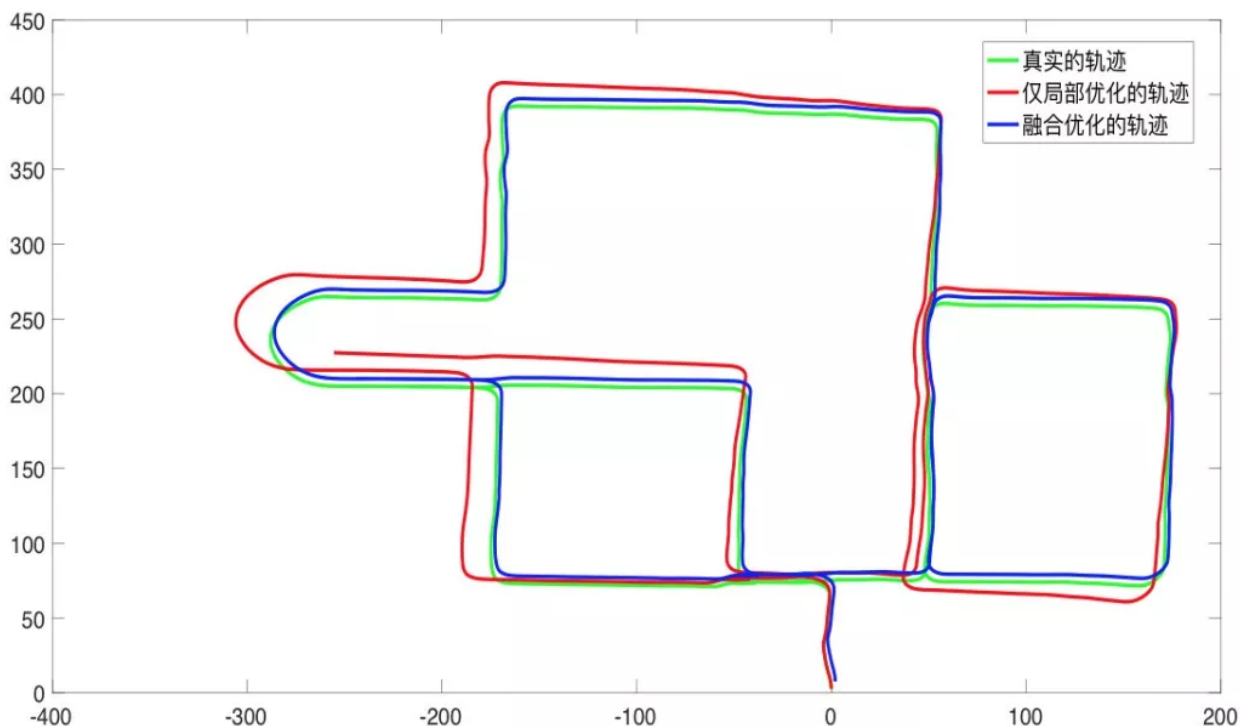


实验结果可见，融合优化的轨迹和真实轨迹很接近，而仅使用局部优化的定位结果误差不断累积。

实验二：车载数据上的实验

该车载数据是在北京市朝阳区某小区采集的，在数据采集阶段，车辆以5km/h 到 30km/h 的速度行驶，一共行驶2271m。

实验结果如下：



实验结果可见，融合优化后的定位结果明显优于仅使用局部优化的定位结果，融合优化中误差得到及时修正。

参考文献

1. J. Delmerico. A Benchmark Comparison of Monocular Visual-Inertial Odometry Algorithms for Flying Robots. 2018. ICRA
2. T. Qin. VINS-Mono: A robust and versatile monocular visual-inertial state estimator. arXiv preprint arXiv: 1708.03852, 2017.
3. N. Trawny. Indirect Kalman Filter for 3D Attitude Estimation. 2005.
4. Sola. Quaternion kinematics for error-state kalman filter. 2017.
5. K. Eickenhoff. Decoupled, Consistent Node Removal and Edge sparsification for graph-based SLAM. 2016.
6. G. Sibley. Sliding window filter with application to planetary landing. 2010.
7. S. Leutenegger. Keyframe-Based Visual-Inertial SLAM Using Nonlinear Optimization. 2015.

8. H. Liu. ICE-BA: Incremental, consistent and efficient bundle adjustment for visual-inertial slam. 2018. CVPR.
9. H. Liu. Robust keyframe-based dense SLAM with an RGB-D camera. 2017.