

杨老师好，师兄师姐们好，因为之前每次都有课这其实是我第一次来参加汇报，所以很多东西都不太了解，不知道形式内容是否恰当；然后是这篇文章实际学习时间只有两天半，所以我只能详细讲前面两个模块，我尽量把自己学到的都捋了一遍，肯定还有很多细节不太清楚也多少会有些错误的地方，还请及时指出谢谢。今天要讲的是之前杨老师推荐我先看的VINSMONO

**摘要：**

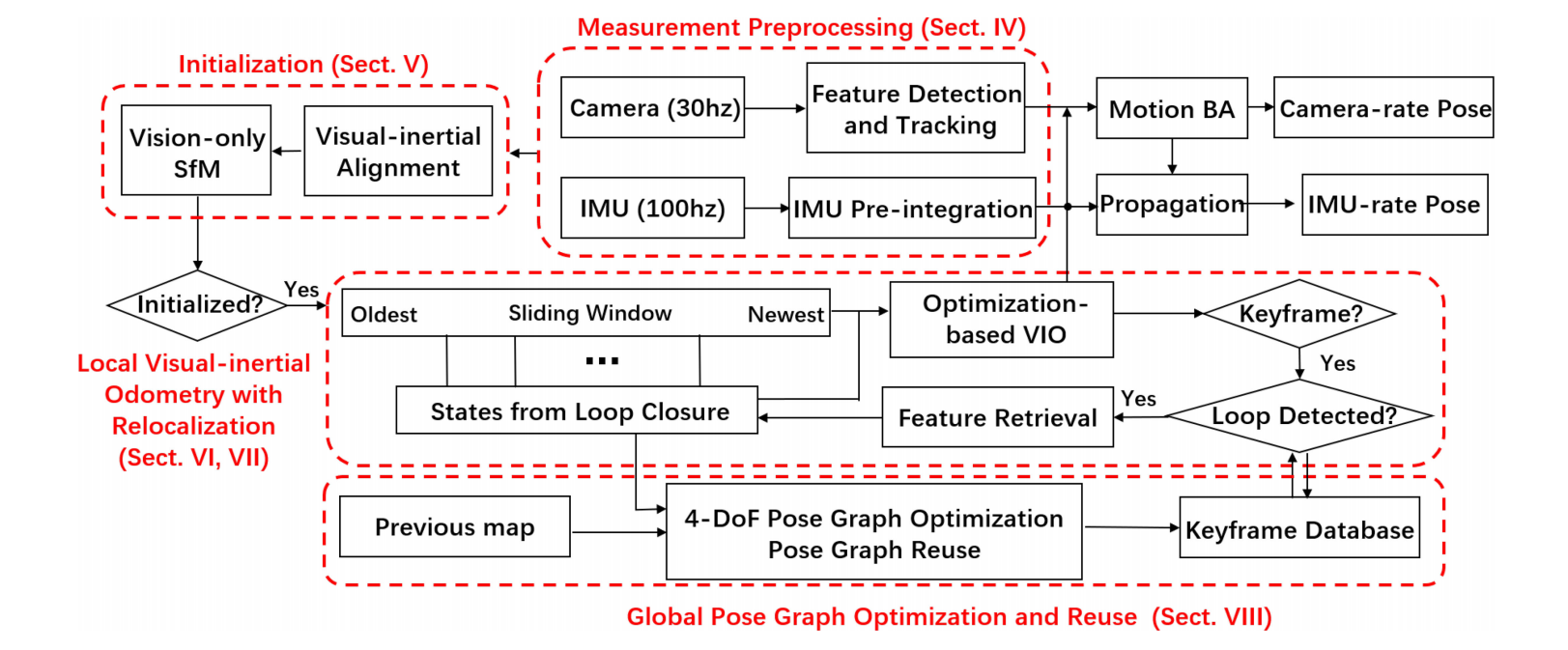
港科大17年提出的VINS系列可以说是VIO（VisualInertialOdometry）方向或是视觉SLAM传感器融合的必读经典算法，它主要提出了一个 单目摄像头和惯性测量单元（IMU）紧耦合、同时基于非线性优化的视觉惯性系统（**visual-inertial systems，**VINS）

这是一个有很强应用背景的方向，通常IMU可以通过加速度计和陀螺仪测量物体的线加速度和角速度，被认为和相机有明显互补性，首先1.IMU的测量量都存在明显的漂移(Drift)，使得积分两次得到的位姿数据非常不可靠。但是，对于短时间内的快速运动，IMU能够提供一些较好的估计。这正是相机的弱点，害怕快速运动。（害怕运动模糊）反过来，相机数据基本不会有偏移，可以有效地估计并修正IMU读数中的漂移，使得在慢速运动后的位姿估计依然有效。此外IMU能够感受到自己的运动信息，从某种程度上减轻动态物体的影响。

目前VIO的框架已经定型为两大类:松耦合(Loosely Coupled）和紧耦合(Tightly Coupled)。松耦合是指IMU和相机分别进行自身的运动估计，然后对其位姿估计结果进行融合（这也是这篇文章里面初始化模块用到的方法）。紧耦合是指把IMU的测量与相机的测量合并在一起，共同构建运动方程和观测方程，然后进行状态结果的估计（后面的紧耦合VIO是核心，体现在残差和代价函数的构建）

算法主要贡献：

1. 初始化：提出了一种能够从未知初始状态启动系统的初始化方法。（因为是单目所以这是必须的）
2. 紧耦合的，非线性优化的 里程计：通过融合预积分的IMU测量和视觉特征观测。
3. 回环检测和重定位：VIO结合回环检测模块，实现了低计算量的在线重定位，保持全局一致性。
4. 地图重用：系统能够高效地保存和加载地图，并通过四自由度全局位姿图优化将当前地图与先前地图合并，实现地图重用。（位姿图优化主要是用来回环检测重定位后的全局优化，也可以用来地图重用）



**数据预处理：**

IMU预积分理论：

我们会快速地推导一遍，理解为什么要用预积分以及其优势和后续会用到的一些关系式

IMU测量值建模：

图表, 图示, 示意图, 箱线图

描述已自动生成

t时刻体坐标系（IMU坐标系）下测得，a-线加速度矢量（xyz轴）=物体加速度+加速度计偏移+重力加速度（转到体坐标系）+噪声；w-角速度矢量=物体角速度+陀螺仪偏移+噪声；噪声通常设为高斯白噪声（均值为0的正态分布），偏移建模为random walk（什么东西，其一阶导服从高斯白噪声）

根据直接测量的物理量构建运动方程，考虑两个连续图像帧间隔时间下的PVQ（位置，速度，方向，世界坐标系）（基本的牛顿运动学方程）

文本, 信件

描述已自动生成

PV是三维向量，q是四元数，Ω矩阵用于将角速度转换为四元数的变化率，上面的积分项来源于四元数微分方程的推导

可以看出，IMU的k+1时刻的状态的传播需要前一时刻k在世界坐标系下的值。特别是在基于优化的算法中，每次调整位姿时当这些起始状态改变，我们都需要重新传播IMU的测量值，每次从头开始积分很浪费计算资源。于是提出预积分算法:

首先想到考虑增量，再想到增量与绝对位姿解耦（将参考坐标系从世界坐标系w调整为第k帧的体坐标系bk下）

将g提出来，因为世界坐标系下的g已知

文本, 信件

描述已自动生成

优化量——预积分项（pre-integration term）：考虑偏移给定，将k到k+1帧的IMU数据积分得到——α（位置增量），β（速度增量），γ（旋转增量）。将它们作为优化量，（与传统算法相比）bk的状态改变并不会对其产生影响（因为所积分间隔时间内的传感器数据不变），因此将其作为非线性优化变量（而非直接的优化位移速度方向，可以避免重复传递，减少计算量）。

R将t时刻测量值转换到起始时刻 k的体坐标系下。通过对陀螺仪测量数据进行积分获得：



以tk时刻的体坐标系为参考坐标系，所以初始化旋转矩阵为单位矩阵（不旋转）

文本

中度可信度描述已自动生成然后计算微小时间内的旋转矩阵，[·]×表示微小角位移矢量的反对称矩阵（同上），exp表示矩阵的指数映射，用于将李代数元素（反对称矩阵）映射到李群（旋转矩阵）。

罗德里格斯（Rodrigues） 公式提供了从旋转向量到旋转矩阵的映射

微小旋转时的近似计算（等价无穷小）

黑色的钟表

中度可信度描述已自动生成

文本

描述已自动生成

在每个小时间间隔内更新旋转矩阵，通过迭代乘法累积所有微小旋转，得到当前时刻到参考时刻的旋转矩阵

上面是在假设IMU的偏置已经确定的情况下，实际上偏置也是需要优化的变量，我们暂不加证明地给出下式，每次迭代优化时，偏置变化小则按上述用预积分项关于偏置的一阶近似修正优化量，变化大则用新估算的偏置重新算

文本, 信件

描述已自动生成

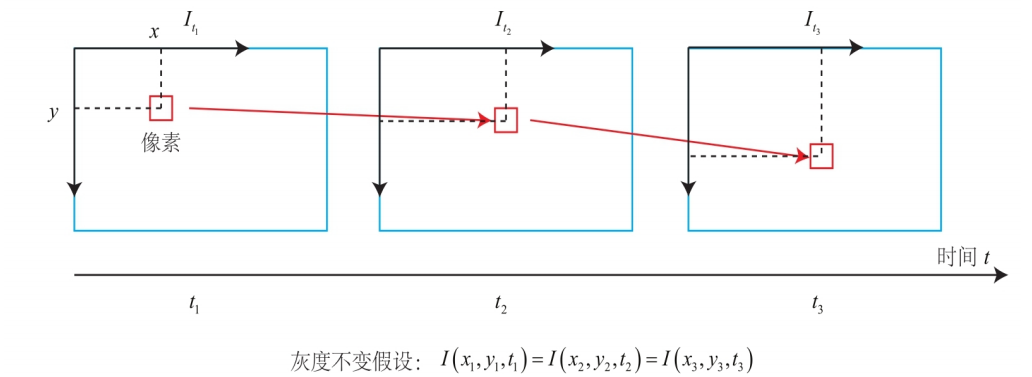
此外，离散时间的实现，协方差矩阵的传播因为在论文正文中没有提及，也暂不推导了

**视觉处理前端：**

1、对于每一幅新图像，用KLT稀疏光流算法Kanade-Lucas-Tomasi Sparse Optical Flow Algorithm对现有特征进行跟踪；

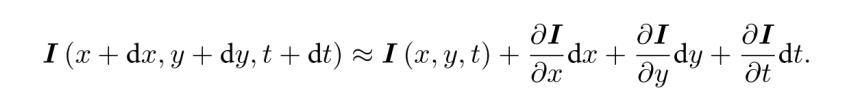
基于LK光流法（如下），结合了Tomasi提出的特征点选择标准（主要选择更稳定的角点或强纹理区域，计算Harris响应值并与阈值比较），用于在视频帧之间跟踪稀疏特征点，主要步骤包括特征点检测、光流计算和特征点跟踪

以LK光流法为例简单推导一下：



把图像看成关于位置与时间的函数，它的值域就是图像中像素的灰度。同时给出下面两个假设：在灰度不变假设下:同一个空间点的像素灰度值，在各个图像中是固定不变的。在LK光流中，我们假设某一个窗口内的像素具有相同的运动，并使用局部窗口计算。

考虑一像素在微小时间内的变化——泰勒展开保留一阶项



图片包含 图示

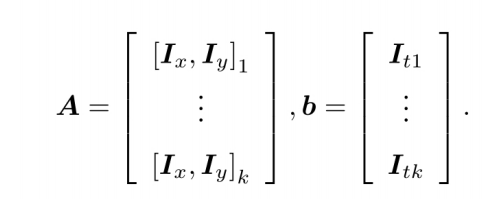
描述已自动生成因为灰度不变假设

图片包含 文本

描述已自动生成除以δt，可以看出分别是xy方向梯度和xy方向速度和对时间变化率，考虑一个局部窗口内的所有像素点（因为一个解不出来）我们构建如下**超定线性方程组**，我们通常采用最小二乘法来求解，这样就得到了描述特征点运动的光流向量

黑色的钟表

中度可信度描述已自动生成



2、检测新的角点特征以保证每个图像特征的最小数目(100-300，为什么最小是个范围？)；  
3、通过设置两个相邻特征之间像素的最小间隔来获得均匀的特征分布，避免挤在一起；  
4、二维点去畸，利用基本矩阵模型的RANSAC算法进行外点剔除，投影到单位球面。

基本矩阵（Fundamental Matrix, F）描述了两幅图像中的点对之间的几何关系。

两个x分别是两幅图像中的对应点像素坐标系下的齐次坐标

RANSAC（Random Sample Consensus）其基本思想是通过不断采样和验证，找到最符合模型的内点集合，从而剔除外点。

 随机采样：

* 从匹配的点对集合中随机选择一个最小子集（通常为8对点），用于估计基本矩阵。

 模型估计：

* 使用选定的点对计算基本矩阵 F。

 验证模型：

* 计算所有点对的代数误差，并判断哪些点对满足该基本矩阵模型（即误差在某个阈值内）。
* 满足模型的点对称为内点，不满足的点对称为外点。

 评估模型：

* 记录当前模型的内点数量和分布。如果当前模型的内点数量多于之前的模型，则更新最佳模型和内点集合。

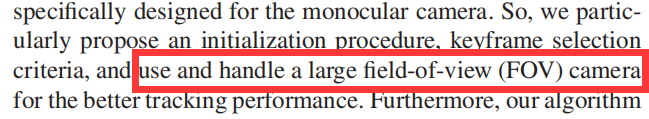
 迭代：

* 重复上述步骤，直到达到预定的迭代次数或内点比例达到一定标准。

 模型重估计：

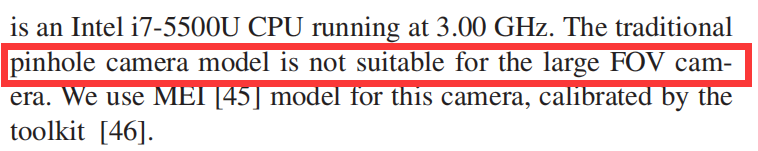
* 使用找到的最佳内点集合重新估计基本矩阵，以获得更加准确的结果，然后剔除外点。

5、对特征点进行去畸变矫正（？），然后剔除外点，然后投影到一个单位球面上(处理广角相机图像的非线性失真，更准确地反映特征点的三维空间位置，优化姿态估计的计算。)。广角相机（cata-fisheye camera）具有非常宽的视场角，会导致图像中出现显著的径向畸变，使得传统的基于平面的投影模型不再适用。提高算法兼容性



文本

描述已自动生成



而文章还提到的视觉处理前端对关键帧的两种选择将在之后进行详细讨论。

 **基于特征点数量的方法**

如果跟踪到的特征点降低到一定阈值，选择这帧作为关键帧，避免跟丢

 **基于视差变化的方法**

跟踪到的特征点的平均视差达到一定阈值时，选择作为关键帧，此外，因为旋转也会产生视差，但这样两帧间特征点不能进行三角化（后面初始化会说到），所以用陀螺仪的短时数据去补偿旋转带来的视差（陀螺仪误差大只在这里直接这么用，导致次优关键帧选择，不直接影响位姿估计）

**初始化：**

该算法的核心贡献之一

为什么要初始化：

 高度非线性：我们知道在同样是高度非线性优化问题的深度学习中，初始化参数的设置相当重要。而对于修正能力较有限的SLAM对初始参数设置的要求就更高，通常系统初始处于未知状态，需要一个准确的初始猜测（位姿、速度、重力向量和IMU偏差）来确保优化过程能够收敛到正确的解。

 缺乏直接距离测量：单目视觉没有尺度信息，IMU的测量又存在偏置误差，如果没有良好的初始值很难将这两种测量融合，因而初始化是VIO最脆弱的步骤，（这也是单目初始化较复杂且重要的主要原因，就是没有直接的深度信息）

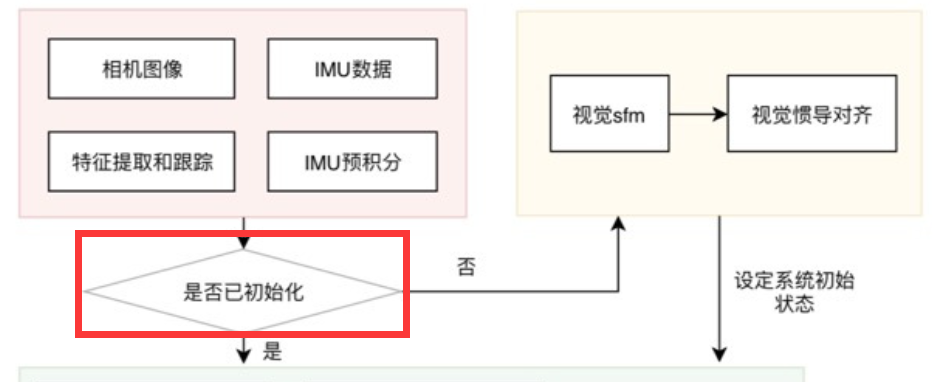
第一是深度初始化：仅通过单张图像单目视觉只能获得二维信息，损失了一维信息(深度)。所以需要动一下，也就是三角化（triangulation）才能重新获得损失的深度信息；

(2)但是，这个三角化恢复的深度信息，是个“伪深度”，它的尺度是随机的，不是真实的（利用对极几何原理从2D-2D恢复的位姿中旋转没问题，平移是一个up-to-scale的），所以就需要IMU来标定这个尺度；

(3)要想让IMU标定这个尺度，IMU也需要动一下，得到PVQ（位置，速度，方向）的P；

第二是(4)另外，IMU存在bias，视觉获得的旋转矩阵基本不存在bias，所以可以用视觉来标定IMU的旋转bias；

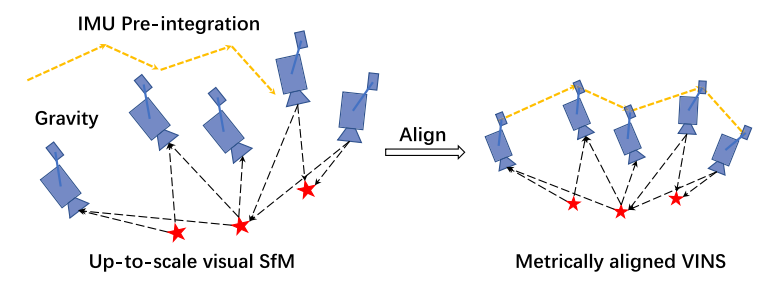
(5)还需要获得世界坐标系这个先验信息，借助重力加速度g来确定；



**初始化只进行一次**

**初始化步骤：**

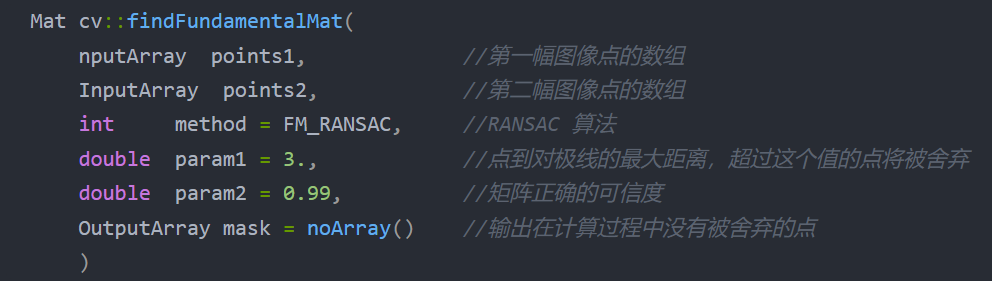
a vision-only SfM（structure from motion）：“up-to-scale”指的是相机位姿和特征点位置的几何结构的估计具有任意比例因子的缩放（也就是说可以得到相对位置和结构的位姿图（pose graph）），不包含绝对尺度信息（通过惯性测量单元IMU来引入绝对尺度信息）



* 1. 维持一个滑窗：限制计算量（10帧）
  2. 最新帧满足条件时（与滑窗内任意帧匹配到一定数量特征点（30+）和具有一定视差（20+））
  3. 五点法（应该是指利用对极几何原理求解至少需要五对点）计算这两帧相对位姿（以第一帧相机帧c0为参考系），在（用这两帧）任意设置尺度因子s三角化特征点得到深度，pnp算法估算滑窗内所有位姿，BA算法最小化所有特征点的重投影误差和

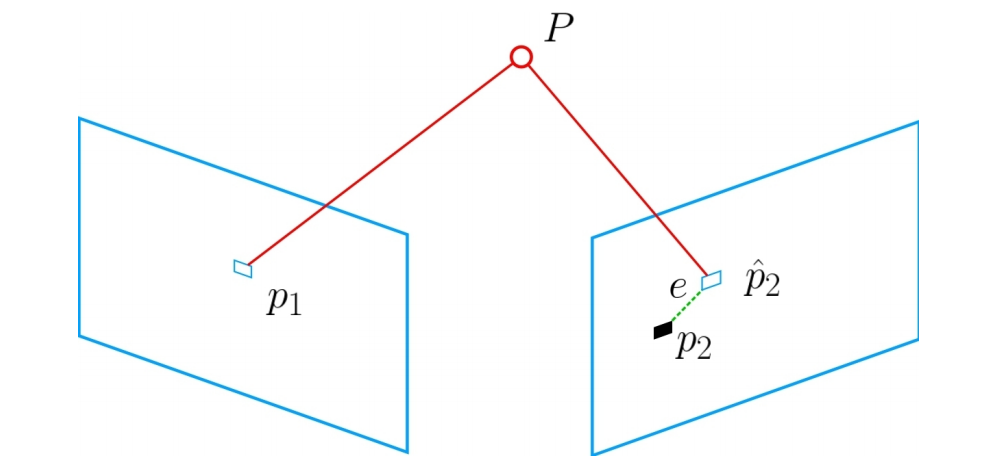
五点法通过从五对对应的图像点求解本质矩阵E来确定两台相机之间的相对旋转和位移。本质矩阵 E描述了两幅图像中的对应点在相机坐标系下的几何关系：

[t]×是平移向量 t的反对称矩阵形式，R 是相机间的旋转矩阵

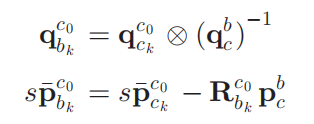


函数findFundamentalMat()既可以用来求基本矩阵F，也可以求本质矩阵E。求什么取决于你的参数points1和points2是像素坐标还是归一化平面坐标，基础矩阵 F 和本质矩阵 E 之间通过相机内参矩阵 K 相关联





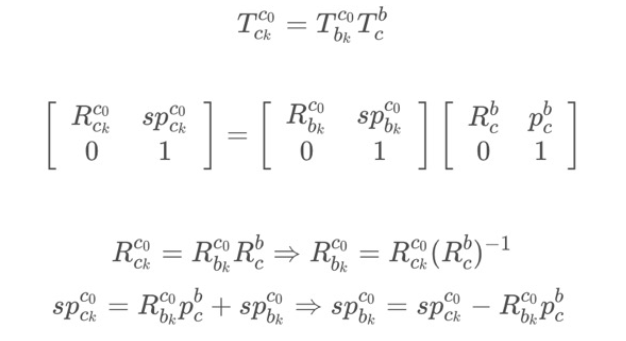
* 1. 已知相机到IMU外参——就是相机坐标系到IMU坐标系的静态四元数（即相机与IMU之间的固定旋转平移关系）



前面计算的是相机第k帧相对初始相机坐标系的位姿，现在已知外参，可以得到IMU坐标下第k帧位姿到c0的变换关系，换句话说，一系列bk帧的位姿也确定下来（c0为参考系）

上述等式是为了后面实现IMU和相机数据在初始相机坐标系下的一致性表达，使得姿态估计可以统一处理。

推导：

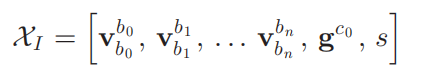


三角化利用相机之间的相对位姿和特征点匹配的视差恢复三维深度

PnP——已知3D空间点及其投影位置时，估计相机的位姿；在PnP中，这个Bundle Adjustment问题，是一个最小化重投影误差问题(Reprojection error)将相机位姿和空间点位置，都看成优化变量，放在一起优化。因为这两项理论体系都挺庞大，就不展开了

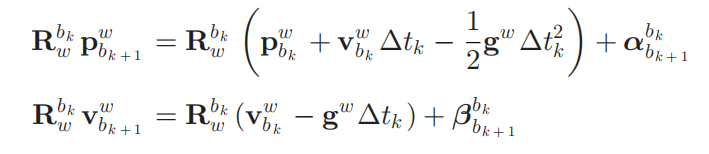
***Visual-Inertial Alignment***

* 1. 其他关键参量的估计：

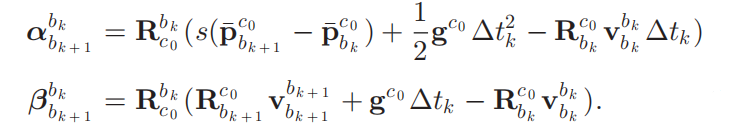


每帧的速度（体坐标系下），g（c0坐标系），尺度因子（米制）

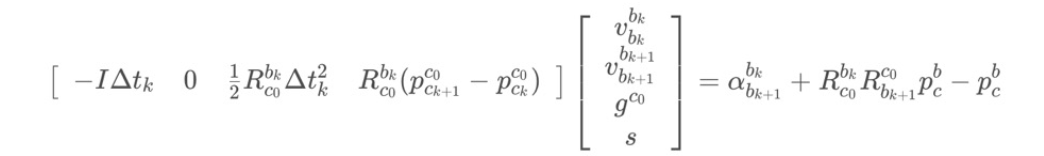
IMU预积分推导中（附录A）有等式：



把预积分提到左边，右边转换为c0坐标系下



考虑噪声的情况下等号左边减去等号右边就是残差，理想状态下这个残差是0，同时用前面bk到c0的关系替换（这就是IMU与SfM的直接联系），再化成 A x = b这种线性方程组的形式，用矩阵表示

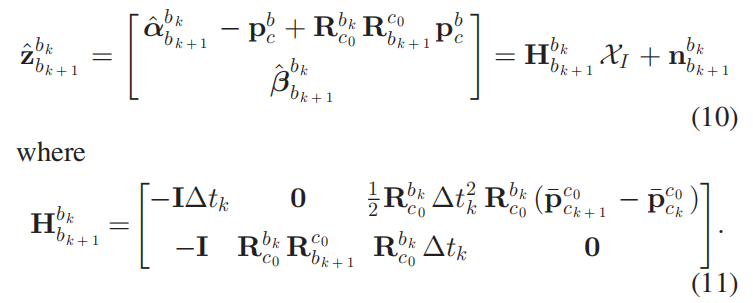


β同理：

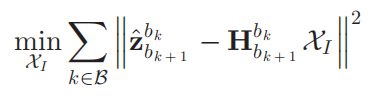
图片包含 图示

描述已自动生成

两个整合一下加上噪声构成实际的线性测量模型，就是论文的式子：

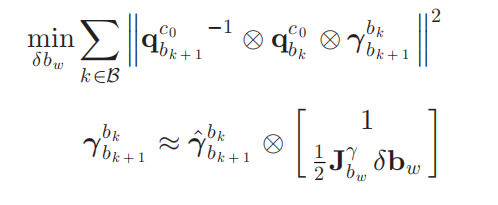


这个式子是一个典型的观测方程，H矩阵实现了状态到观测值的映射，然后z是IMU结合了视觉的观测值，融合了纯视觉和纯IMU的预测结果，同样采用LDLT分解解这个最小二乘问题得到最可能 的初始状态估计这里不像那种用高斯牛顿法迭求解，而更像是用直接法,也就是矩阵运算的方式来求待优化的状态量

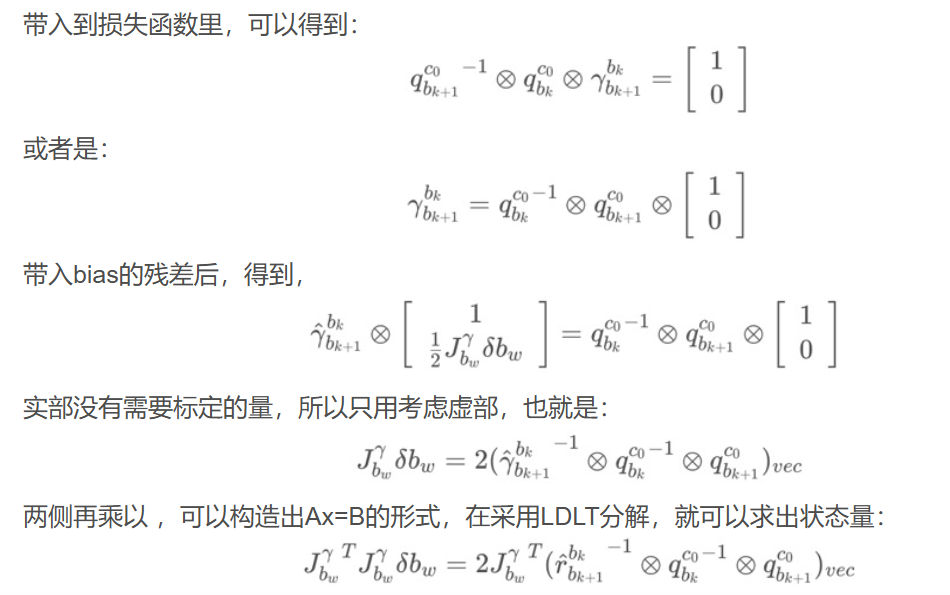


公式中只计算了相邻2帧，得到的SfM有很多帧，然后IMU的测量数据也有很多帧，上面的β是窗口内所有帧。

* 1. 陀螺仪测的角速度bias标定：



q是前面通过通过视觉sfm得到的，γ\_hat是纯IMU测量的（带偏差），目的是找到一个偏差，使上述代价函数最小（理想状态下，这三个数乘积应该是单位四元数，优化过程中每次bias微小变化时，γ线性地近似的通过第二个式子更新）



构造对称正定矩阵

带入求得的偏差重新计算所有帧之间的IMU预积分，这里只初始化陀螺仪的偏移可能是因为相对来说加速计的偏移比较小，which将在后面的VIO优化，然后刚好视觉的P不好，用IMU标定；IMU的Q不好，用视觉标定

约束重力再优化：

因为在上面把重力矢量当作3自由度来优化，但是重力模长已知，所以退化成2自由度。所以，在这里需要想一个办法，就是**怎么样用一个2自由度的表达式来表示3维的向量呢**？

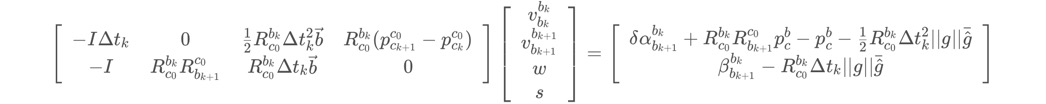
雷达图

描述已自动生成

在这里，采用球面坐标进行参数化，也就是用g的模长作为半径画一个半球，上图蓝色实线对应的是gc0的测量值的方向(也就是优化前的方向)，在这个交点上找到一个切平面，b1，b2是切平面的一组正交基底（论文中写的可任选），再这两个方向上分别加上微小扰动w1，w2



扰动后的重力，只剩下两个自由度，替换前面的g，再进行迭代优化，直到g收敛



我们可以发现这一部分做的工作和前面是相似的，但是却是在前面的基础上进一步做的工作。

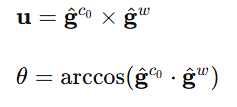
最后转到世界坐标系：

上面c0下的重力求出来了，然后这里假设的c0原点和世界坐标系原点重合，有关系：



所以我们然后可以用gc0和gw（化单位向量后）作叉乘得到一个旋转轴u；

重力方向是沿z轴，因此它们之间的夹角θ是能很快求出来的；

 其中，θ是旋转角度，u 是旋转轴的单位向量。

Rodrigues 公式提供了从旋转向量到旋转矩阵的映射

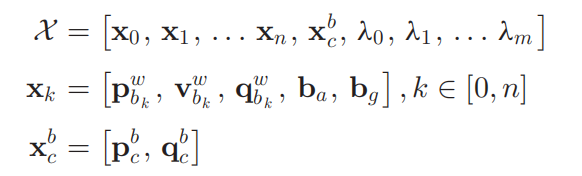
 

把所有 c0 坐标系下和b坐标系的变量都乘上相应的旋转矩阵旋转到 w系下

(3) 把相机平移和特征点尺度恢复到米制单位  
初始化大功告成！

**紧耦合单目VIO**

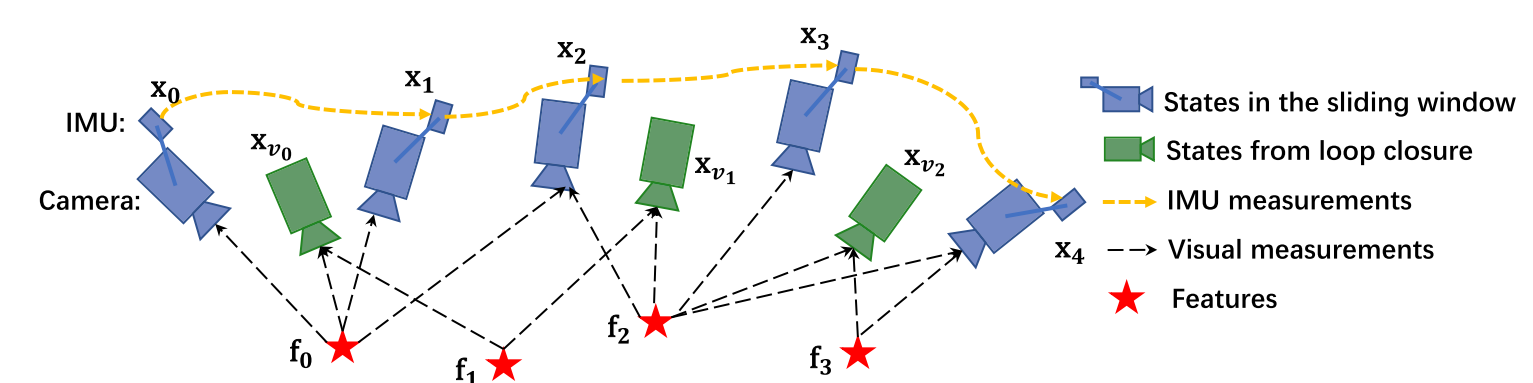
使用滑窗，紧耦合，来估计状态（滑窗内的状态向量如下）



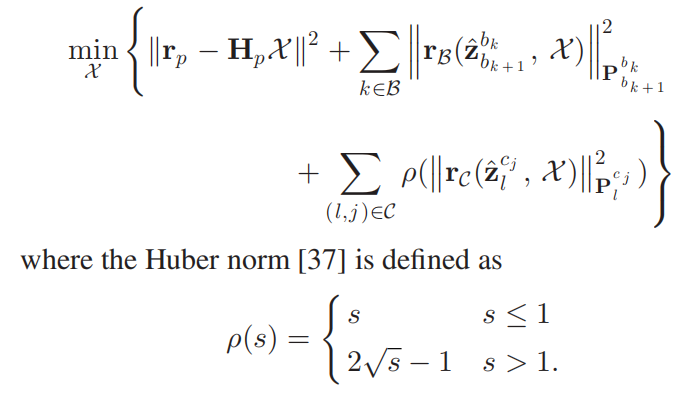
为什么用这些作为优化变量呢？

里程计要做的无非是估计当下的设备位姿状态和路标点的位置，对于设备的状态，相机有位置和姿态，IMU有PVQ和bias，这里首先将IMU的各种状态作为优化量，因为IMU和相机通过一个外参联系，而这个外参本身也不是精确的，所以也加进来优化。然后是路标点的位置信息，首先路标点通过光流法跟踪捕获其投影位置这个是确定的，经内参投影到归一化相机坐标系也是确定的，唯一不确定的就是深度信息，其次再是路标点从相机坐标系到世界坐标系的变换，这个就是前面提到的设备状态。

n是窗内关键帧总数，然后k是第k帧图像捕捉时刻，对于IMU有PVQ和两个偏差需要优化，此外前面初始化假定已知的相机到IMU的外参也是需要优化的变量（设计硬件系统时通过静态方法对相机和IMU的外参进行初步标定，为尽量减少误差保证高精度，这个外参也要优化），m是特征点总数（只包含观察过至少两次的特征点），λ是某个特征点到第一次观察到时的逆距离（以逆深度作为参数的原因，一是因为一些观测到的特征点深度值可能会非常大，难以进行优化；二是可以减少实际优化的参数变量（？）；三是逆深度更加服从高斯分布，提高数值稳定性。）



滑窗内有：1.相机（和IMU）位姿，回环测量，IMU测量，视觉测量，特征点位置

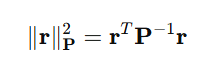


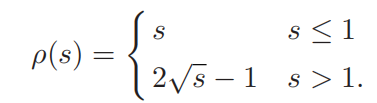
下面构建误差函数，最直观的无非是相机的重投影误差（估计的路标点位置经估计的状态重新投影，与实际匹配的特征点位置的偏差）和IMU的误差（根据估计的PVQ计算预积分值，与实际根据测量结果的预积分值的偏差），但随时间相机和IMU的约束项会越来越多，所以此处是维持一个窗口进行的优化，压缩历史信息（边缘化）而非直接丢弃，转换为新的约束——先验误差项，用来描述先前的估计信息和当前优化变量之间的误差，为后续优化提供基准并提高稳定性。

最小化先验之和（先验残差（来自于已经边缘化的先验数据，表示当前观测与模型先验的差距）- 先验残差关于状态的雅可比矩阵\*当前状态），（这里的 Hp 表示基于当前状态 X计算得到的残差。通过最小化这个式子，确保了新的状态估计不仅仅是基于当前观测，还考虑了先验信息，我们可以调整 X使得新的状态尽可能使残差最小）

以及最小化马哈拉诺比斯距离（协方差矩阵表示了测量的不确定性。通过协方差矩阵的逆（信息矩阵）对残差进行加权，加权残差的主要目的是考虑不同测量的置信度（不确定性），从而在优化过程中更加准确地调整状态估计，可以更好地体现置信度高的测量对状态估计的影响。）来获得最大后验估计，

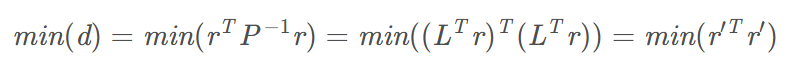
简单说，VINS-Mono的误差来源于三部分，分别是来源于视觉的**重投影误差**，来源于IMU预积分的**IMU误差**（PVBQ的运动误差）以及来源于滑动窗口维护的边缘化操作带来的**先验误差**。

r是残差向量，P是协方差矩阵

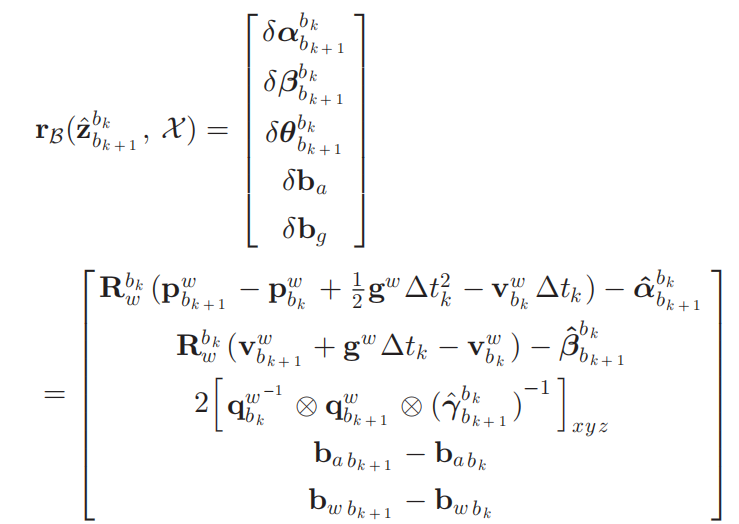
Huber范数

鲁棒核函数（robust kernel function），用于降低异常值（outliers）的影响

因为Ceres只接受最小二乘优化，即min(e^Te)，因此在代码中，需要把协方差矩阵做LLT分解，代码中对应为sqrt\_info，这样最后优化的



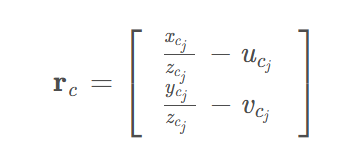
IMU的残差：可参考前面的IMU预积分得到的运动学方程



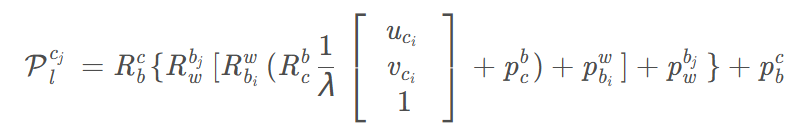
相对角位移那里取四元数的向量部分

因为使用广角相机，针孔相机模型不适用，不再将重投影误差定义在归一化相机平面，而是投影到一个单位球面上，其实这之间只有一个投影函数的差别，所以下面先用针孔相机模型进行推导然后再推广。

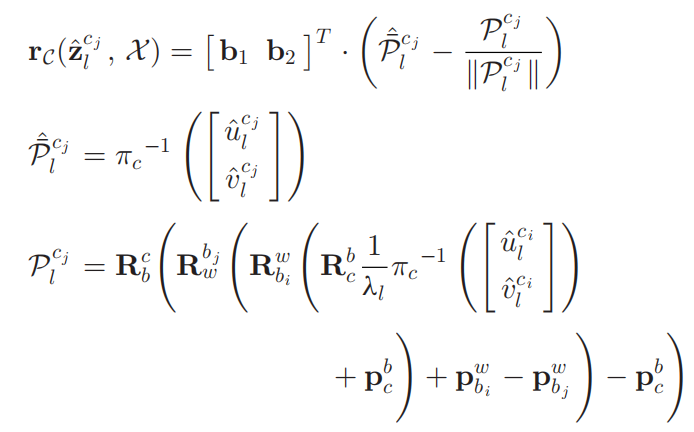
针孔相机模型下的视觉残差，考虑第L个特征点在i 和j时刻的情况。左边那个是模型根据预测的位姿数据和i时刻的观测重投影到相机归一化平面的位置，右边是j时刻的观测值（也是在相机归一化平面上）注意：这里的uv不是指代像素成像平面坐标，而是从像素平面投影到归一化相机平面的坐标



重投影方程，第L个特征在j时刻的位置（这里是i时刻的特征点根据估计的位姿信息转换过去的）



VINSMONO里面考虑了更一般的球面投影模型，视觉残差为：（这里的UV是像素坐标系下的）



π-1用相机内参将像素坐标转为一个单位向量，可以看作是相机内参

带把的是实际观测值；方差也定义在单位球面上

从上到下分别是：

视觉残差定义（观测值-模型预测值（重投影）），都投影到观测向量的切平面

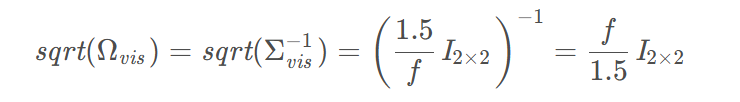
观测值从像素坐标转为单位向量

重投影方程，和针孔的一样，只是多了一个π投影函数

图示

描述已自动生成

VINS在代码中sqrt\_info取1.5个像素，对应到归一化相机平面还需除以焦距f，最后得到的信息矩阵（协方差的逆，同样表征测量不确定性）：（这里真正的信息矩阵其实是sqrt\_info的平方）



（优化过程的数学原理）IMU和视觉残差优化中的雅可比矩阵推导待看。

**边缘化：**

什么是边缘化？是指从优化问题中去除某些变量，同时保留它们对其他变量的影响。为什么要？用于压缩历史信息，优化计算和存储

**要边缘化哪些变量？**

边缘化变量的选择：

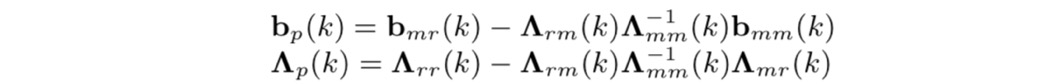
下一帧是关键帧：去掉最早一帧及其视觉惯性测量，转化为先验；

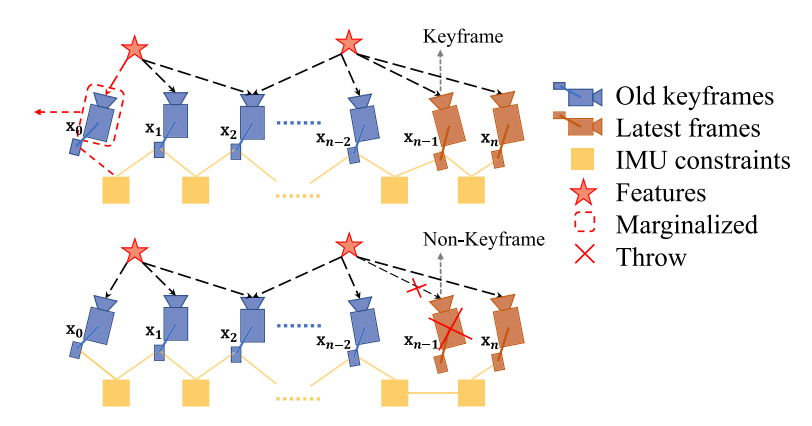
下一帧不是关键帧：去掉这一帧和它的视觉测量，保留IMU预积分测量

**怎么边缘化？**

利用Schur Complement：Schur补是一种用于简化线性系统或优化问题的数学工具。它通过消除某些变量将历史信息转化为当前的先验，将问题转化为一个较小的系统。

每有一个状态被边缘化，与其相关的测量就会被转化为一个先验，然后加到已有的先验上。





**重定位**

紧耦合，解决滑窗和边缘化带来的累计误差（因为只在窗内优化，而且边缘化导致早期的那些线性化的点被固定，导致次优估计）

回环检测：DBoW2词袋模型（基于几何外观，对所有特征点聚类得到词袋，判断一幅图的特征点都是词袋里的哪些词，获得一个独特的特征点分布图/表，加权处理成向量（如使用TF-IDF方法强调这幅图最有辨识度的部分）），除了前面VIO用的角点，这里还多检测了500个并计算BRIEF描述子，提高召回率；时间（设置窗口和阈值）和几何一致性检查后返回回环检测候选帧；保存描述子，舍弃原始图像

特征检索：检测到回环时，通过BRIEF描述子匹配找到对应关系，建立局部滑动窗口与回环候选帧之间的连接，然后使用两步几何外点拒绝法：1.基于基本矩阵的RANSAC（2D-2D）（同前面的视觉前端预处理）2.基于PNP的RANSAC（3D-2D），3D信息来自当前滑窗内帧（也只有窗口内的帧有特征点的3D信息）

紧耦合重定位：回环帧的位姿视为常数（来自位姿图或先前的里程计输出（如果是第一次重定位的话））代价函数多加上一项回环约束

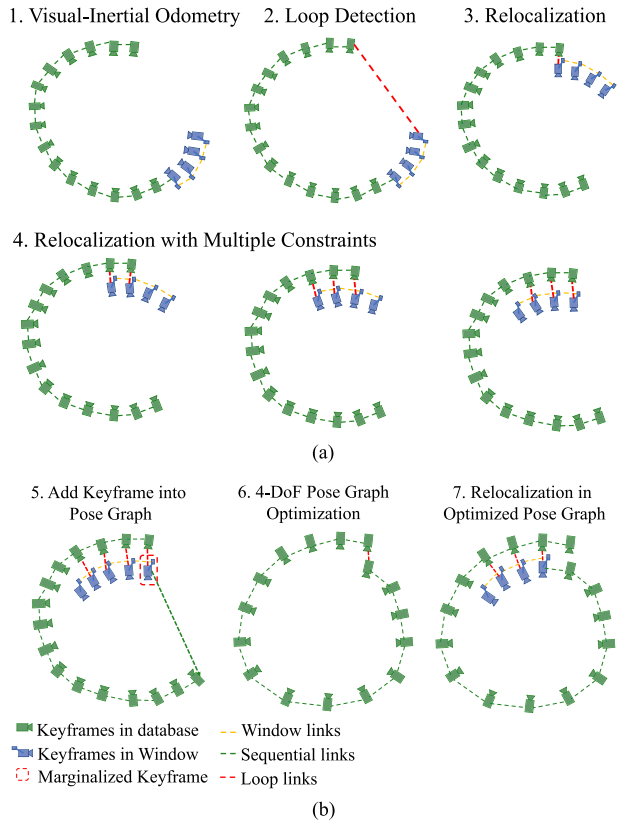
重定位过程使单目VIO维持的当前滑动窗口与过去的位姿图对齐。（重定位只优化调整当前窗口，此时没有实现全局一致性）

将所有回环帧的位姿作为常量（不只检测一个回环帧，多检测到几个后同时加进来优化，提高精度），利用所有IMU测量值、局部视觉测量和从回环中提取特征对应值，共同优化滑动窗口。

要求解的状态的维度保持不变（新加的pq是第v帧回环帧，为常数）

文本, 信件

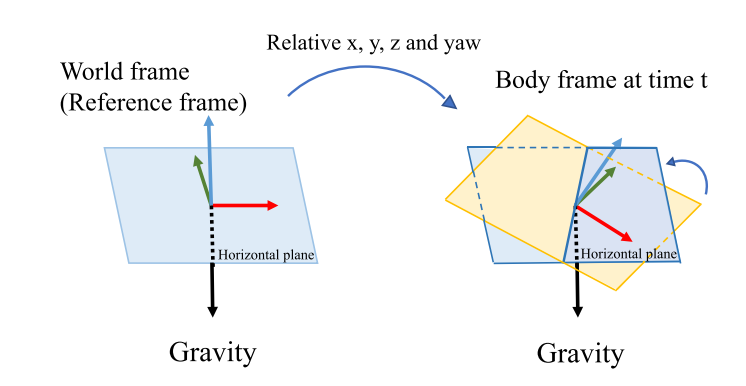
描述已自动生成



**全局位姿图优化**

确保过去位姿的一致性

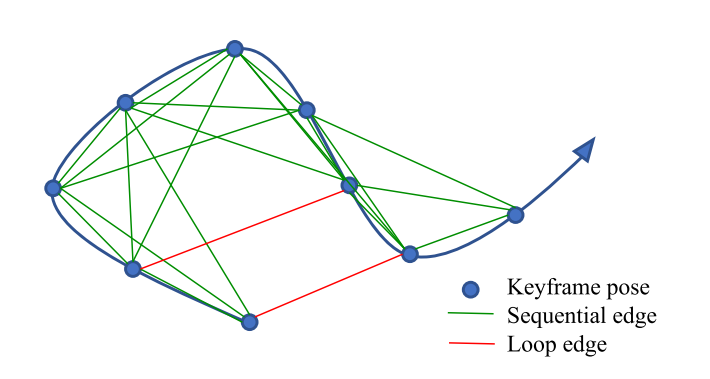
因为重力已知，所以可以知道绝对的roll和pitch角，然后就只剩下4个方向（自由度）存在累计的偏移误差，x, y, z, yaw角。翻滚，俯仰，偏航



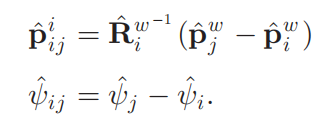
位姿图

VIO处理后的关键帧及其对应位姿作为顶点，

顶点间有两种边：序列边（时序上两帧间的那四个自由度变化，一个顶点对先前的几个点都有关系），回环边（检测到构成回环关系的两帧的四个自由度关系）



序列边：

来自VIO

当前帧是i，j是其中一个先前帧，然后这两个顶点间的序列边包含，两帧间的相对位移和相对yaw角变化，其位姿信息直接来自于VIO

回环帧的包含信息同上式，然后如果两帧确定具有回环关系，则其边包含的位姿信息来自于上面的重定位

**4DOF位姿图优化**

残差定义如下：

文本, 白板

描述已自动生成

Φ和θ是前面固定的roll和pitch角，来自于VIO（认为无偏差），上面的两项残差就是根据边包含的数据定义的，通过引入新的约束（回环）和先前的位姿，实现全局的一致性，最终代价函数定义如下：

电脑屏幕的照片上有字

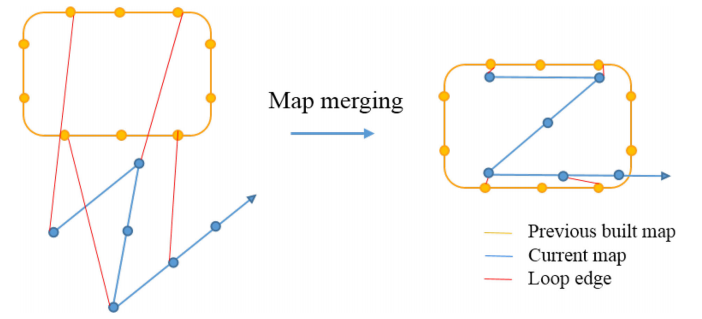
中度可信度描述已自动生成

右边那个是来自回环的残差，加上Huber范数进一步去减少错误回环的影响；然而序列没有加任何鲁棒核函数，因为其来自于VIO，已经有足够的外点拒绝措施了。

然后这里的全局位置图优化和重定位是两个独立的进程，这样可以在最新的位姿图没有优化完时重定位仍能使用上一个位姿图，

**地图融合：**

有先前地图，检测到回环，加入上面的全局位姿图优化，因为所有边都是相对变换所以很好加，如下，只需要固定黄色地图的第一个顶点，就能把蓝色拉到黄色对齐。



位姿图存储和加载

第i帧关键帧存储的信息



索引，绝对位姿，（如果有回环匹配：v是它回环匹配帧的索引，后面是回环边），D是特征点坐标和描述子

节点和回环边直接加载，再构建序列边（一个关键帧和它相邻的几帧），时间复杂度与位姿图大小呈线性相关，位姿图加载完成后立即进行优化，只有当窗口中有被边缘化的关键帧时才会把这帧加到位姿图（如有回环，其约束也会加进来）

重定位和位姿图优化在不同线程，在重定位过程中，建立在最新的位姿图上。