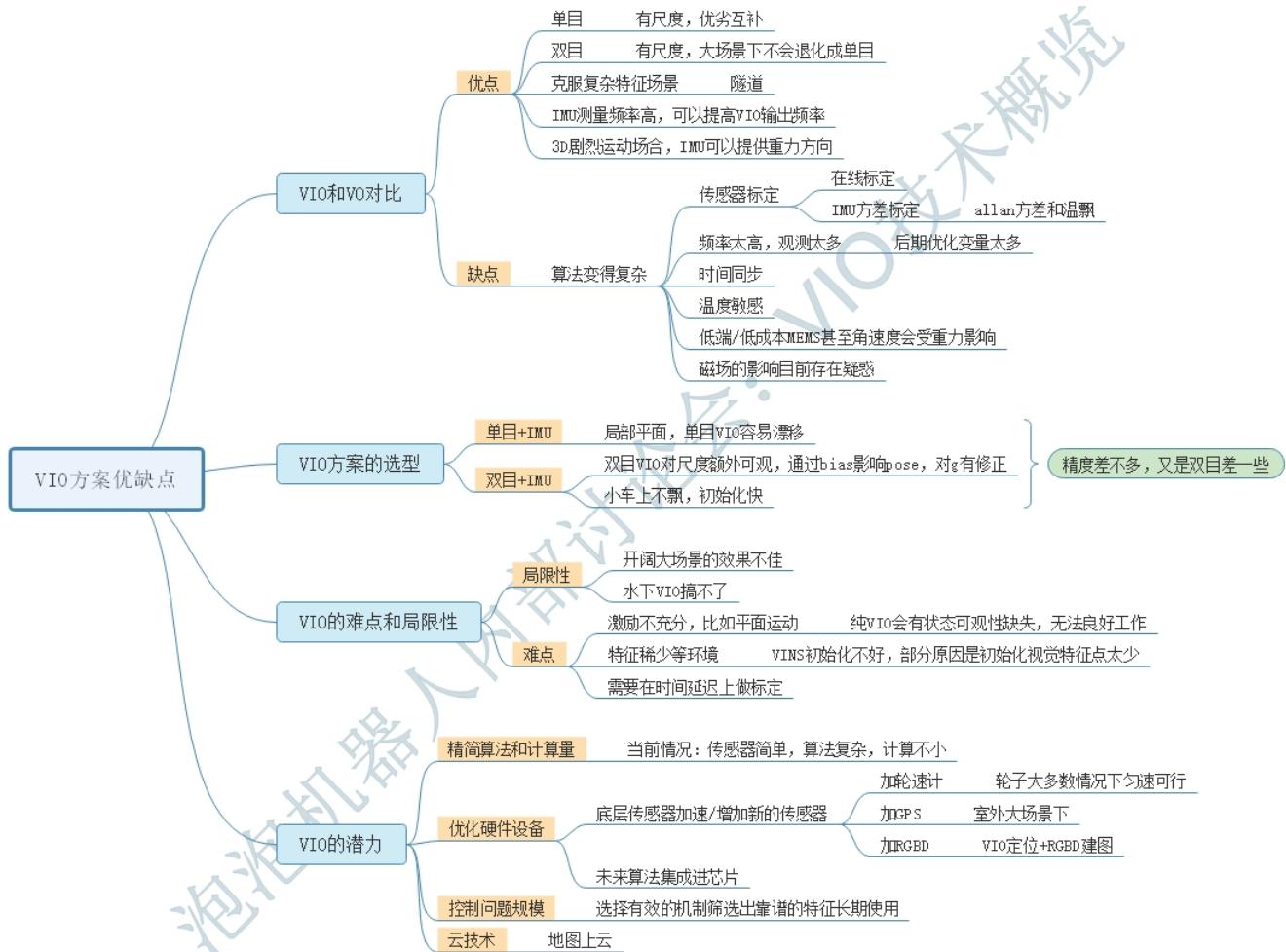


主题：VIO技术概览

思维导图



具体讨论内容

1. VIO相比VO（单目、双目和RGBD）好处和坏处在哪？

1.1 引入IMU的好处和潜在缺陷，例如引入IMU新增哪些可用信息，引入IMU不会增加哪些负担，IMU会受哪些因素影响，IMU会增加哪些负担

- 好处：
 - 和单目 最好的是有尺度 且优劣互补；
 - 和双目应该也算是有尺度，双目在大场景下会退化成单目；
 - IMU测量频率高 可以提高VIO的输出频率；
 - 对于3d剧烈运动的场合，imu可以提供重力方向；
 - 克服重复特征场景，比如隧道等；
- 坏处：
 - 算法变得复杂；
 - 传感器标定， 在线标定；
 - 时间同步， 对温度敏感；

- 低端mems甚至角速度会受重力影响；
- 因为频率太高 观测太多 后期优化变量太多
- 磁场的影响目前存在疑惑？？？？？(电容的不会受影响， mems的会受影响)

2. 一般vio方案用单目还是双目？

2.1 为何VI-ORB SLAM和VINS-Fusion中， 双目+IMU会比单目+IMU要差？

- 局部平面， 单目VIO容易漂移， 双目VIO对尺度额外可观， 然后传递到bias里面， 再传递到姿态里面， 于是对g有修正。
- 双目vio精度 和 单目vio单目差不多， 或有时候差一些。 双目的好处是 小车上不飘， 初始快。
- VINS-Fusion的双目+imu初始化其实用的就是双目初始化， g没做对齐？？？？
- IMU方差标定： allan方差和温飘

3. vio目前的难点、局限性？

3.1 难点可能是指目前难以解决的已知问题， 可能需要结合大家目前的研究方向来讨论， 因为研究方向通常是要解决的难点

3.2 局限性可能是指无法使用的场景， 可能需要结合camera或imu无法使用或者效果不好的场景讨论

- 局限：
 - 开阔大场景的效果不佳；
 - 水下VIO搞不了；
- 难点：
 - 激励不充分时（比如平面运动）， 纯vio会有状态可观性缺失， 无法良好工作；
 - 特征稀少和平面等环境， 虽然理论上imu可以支持， 但是实际上目前还是个痛点(VINS初始化不好， 部分原因是初始化时视觉特征点太少)；
 - 需要在时间延迟上做标定；

4. vio还有哪些潜力可以挖掘？ 有哪些可能提高当前VIO方法的策略？

4.1 轻量级： 计算平台， 速度与精度的折中， 能耗， 输出频率**

- 传感器简单， 算法复杂， 计算不小；
- 底层硬件加速， 或者可以有效的机制筛选出靠谱的特征长期使用， 控制优化问题规模， 或者增加新的传感器；
- 未来算法可能会进入芯片， 地图会上云；

4.2 与其他传感器结合（分应用讨论）： 轮式里程计？ GPS？ 结构光？ TOF？

- 加轮子， 轮子大多数情况下匀速可行；
- 场景大了就该上GPS；
- VIO+RGBD建图更给力；

主题： VIO标定

一、 VIO标定的目的

1. VIO标定的目的是什么？矫正误差和同步？

讨论内容：

1. 不止矫正误差吧，比如相机内参是计算的时候必须的数据啊；
2. 多传感器融合很重要的一个问题就是解决时空基准的统一，外参和IMU标定就是在解决这个问题；
3. 另外，对单个传感器的标定，主要是为了能够得到确定性误差的定量描述，和不确定性测量误差的统计特性，便于在算法中做处理；
4. 多传感器的时空统一简单点说是，时间统一就是同步，空间统一是坐标系关系是吧；
5. 我觉得既然叫标定肯定是系统里面的单个传感器内参和传感器相对关系都得标吧；
6. 各个相机与IMU的内参各个传感器之间的外参；
7. 还有各种传感器时间坐标系标定，就是所谓的时间偏置；
8. 以相机为例。左右目设计要求baseline 12cm，实际有安装误差。然后，镜头的sensor也有安装误差。这些误差需要确定。而每台相机，和每台相机安装误差不一样。有一类标定，就是标这个的。标定空间关系；
9. imu的标定和建模有关系，低成本的话，一般包括bias统计特性，非正交误差，标度因数，白噪声统计特性；
10. 然后，各传感器之间数据传输也有时间误差。这是标时间关系；
11. 标定就是传感器之间的空间误差和时间误差以及传感器本身的特性；
12. 还有些imu三轴角速度计由三个单轴角速度计拼成，有把这三个外参独立标定的；
13. 多个单轴角速度标定；
14. 传感器本身的误差；

相关文章： **Monocular Visual Inertial Odometry on a Mobile Device**

二、VIO标定现有方案对比

2.VIO标定现有方案有哪些？不论好坏请先罗列出来，开源的请给出地址，未开源请给出文章

讨论内容：

1. Kalibr，开源；
2. Vins；
3. Msckf；
4. 纯VIO的应该就这一个吧 但是我现在一般用camdocal标相机内参 然后用kalibr的外参；
5. imu_tk；

相关文章： **A robust and easy to implement method for IMU calibration without external equipments**

3. 目前哪些标定方案比较常用，他们之间的区别是什么？各自的优缺点是什么？

讨论内容：

1. 首先说一下我的体验 我一开始确实标不好 当然手法还有标定板都有很大关系；
2. 就大家用kalibr的时候注意打个好点的标定板，还有提前把IMU的参数标定好；
3. 我之所以不用这个功能（自标定），主要是目的是评估odometry结果，如果用他们标定还是要激励充分；
4. 用专门的标定工具，精度理论上是要更高的，毕竟有一个标定板作为强约束；
5. 而且kalibr对imu建模更好，建模细致；
6. 标定工具和设备的好坏都会影响标定的效果；

7. 我给大家一个参考 就是人家标定出来的误差是什么样子， kalibr:

```
Calibration results
=====
Reprojection error squared (cam0): mean 0.0258226320892, median 0.0147480323833, std: 0.050380835
Reprojection error squared (cam1): mean 0.0256774901704, median 0.0129226561016, std: 0.082312822
Gyro error squared (imu0): mean 0.0937716434596, median 0.0460440702065, std: 0.116936249
Accelerometer error squared (imu0): mean 0.170679362745, median 0.109385920057, std: 0.17249525343
```

参考论文：

A Benchmark Comparison of Monocular Visual-Inertial Odometry Algorithms for Flying Robots

2014-ICRA High-fidelity sensor modeling and selfcalibration in vision-aided inertial Navigation

Tutorial: IMU-camera calibration

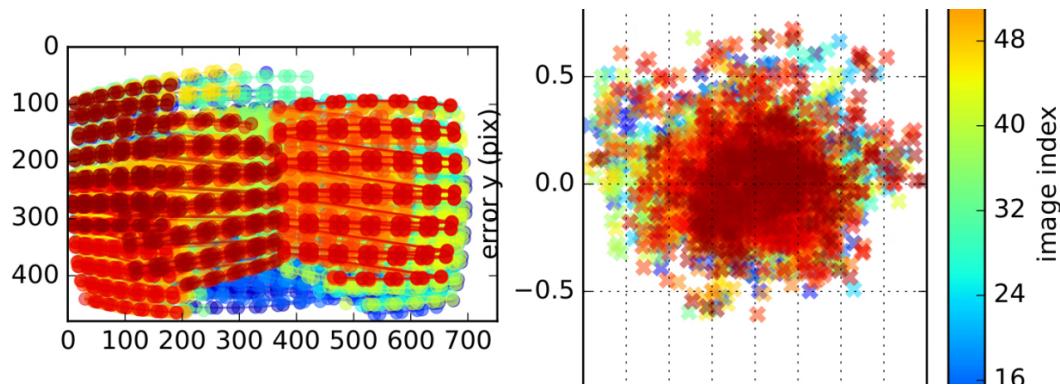
A video tutorial for the IMU-camera calibration can be found here:



4. 标定VIO需要哪些额外的工具？分方法举例说明

讨论内容：

1. 目标定就激励充分；
2. 选个环境纹理相对丰富的地方就ok吧；其实我刚刚发的那个截图里边的视频里边有；
3. 如果是kalibr xyz rpy6个自由度都最好激活；
4. 转台 是个工具；
5. 低精度imu转台完全没必要；
6. kalibr 打印一个好的标定版 保证标定版在一个平面上；
7. 标定的时候让标定板遍历到整个相机视野；
8. 遍历重叠的结果：



参考文献： [Extending kalibr: Calibrating the Extrinsic of Multiple IMUs and of Individual Axes](#)

5. 单目vio和双目vio哪个标定效果更好？两种标定方案的差别是什么？

讨论内容：

1. 双目VIO标定存在的问题是视野有限，无法全部覆盖，只能标定重叠部分，单目是可以利用视野里的所有内容；
2. 单目VIO的标定，实际上用了全图视野的内容；双目VIO因为只利用了视野重叠部分的内容，所以会有所偏差；
3. 也就是双目理论上确实标定比单目还是麻烦一点；
4. 标定工具不会管这些，他就用能用的信息，所以也没什么麻烦不麻烦的；
5. 传感器更多标定会更差一点；
6. 双目的话，两个相机都可以对imu的轨迹进行约束，这是双目的一个优势；构造误差项；

尝试：

1. 双目，能不能构造两个VIO,一个双目？组合一下？相当于两两组合标，标完搞个bundle，但是误差感觉还是会比单目差，因为两两标定也就是标单目；
2. 两两标定，再标个方差，融合怎么样；

三、VIO标定的分析工具：

6. 有哪些实用的Allan方差分析工具？

讨论内容：

1. 分析工具：matlab :psins工具箱中的avar函数；
2. 一个开源项目：<https://github.com/ethz-asl/kalibr/wiki/IMU-Noise-Model>；
3. 高文良：imu_utils，开源：https://github.com/gaowenliang/imu_utils；
4. kalibr_allan：https://github.com/rpng/kalibr_allan；
5. 如果编译不过去可以看一下我（晨博）提了一个merge，pull request里面，抱歉我的那个request是在他的依6. 赖库code utils里面，只是个很简单的头文件路径问题；
6. 严恭敏老师的书：

7.有的说用Allan方差的结果，再乘以10。这个10倍，也是个经验值。反以，感觉噪声项的标，有主观经验因素在里面；还是直接相信 allan方差的结果？

讨论内容：

1. 我是相信它的数量级，然后调系数；
2. Allan方差乘以10其实也是完全由Allan方差决定这个值的吧；
3. 比如标出来 $3e-5$ ，我实际可能用 $7e-5$ ；
4. 我自己的算法在EUROC上和用小觅，都是用的同一套参数；

四、VIO标定的实际应用

8.有哪些开源项目是关于单目+imu做slam的？

讨论内容：

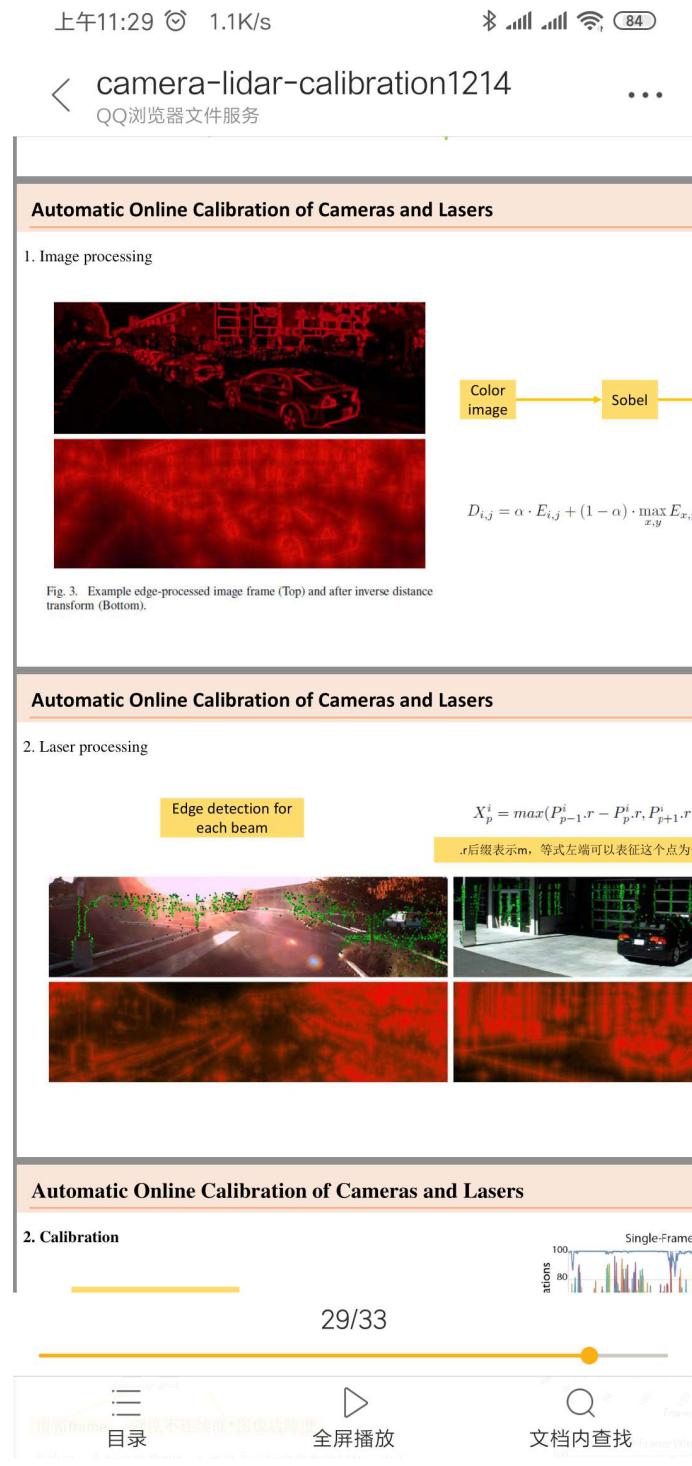
1. 知乎：<https://www.zhihu.com/question/53571648>；
2. vins、okvis、rovio、r-vio、msckf_mono；
3. TUM的VIO：<https://vision.in.tum.de/research/vslam/basalt>；

参考文章： **Basalt: Visual-Inertial Mapping with Non-Linear Factor Recovery**；

9.如何评估和每次开机自检外参好不好？

讨论内容：

1. 这个在相机和雷达标定论文有：



2. 而且其实VINS如果你只设置一个初值的话，也算是每次都标定一下吧；
3. 主要讲雷达和相机虽然在车体固连但是车跑起来的震动等等因素其实对外参是有影响到；
4. 改进方向：开机自检和评估外参无法做到，只能每次都检查，是个改进的点；

10. 怎么知道VIO，它已经飘了？还是在正常工作？

讨论内容：

1. 这个目前有工程经验方案，理论的没看到有研究的，可能可以归到故障检测里面，是一个研究点；
2. 其实要是系统崩了是可以有方法能检测到的，挺多算法都有system failure detection 但是如果说崩的因素归结到外参？？？嗯，我也觉得不会是外参导致崩；

3. 不知道如果把它归结到内参，从数学模型上不太好弄；
4. vins是有failure detection吧；
5. 可能要从固定传感器的工具上解决了；
6. 这个就像琛哥说的 可以从硬件固定入手；
7. imu 或者 camera 被撞歪了？硬件 或者外部的原因， imu 和 camera 不同步了？ 例如 转角的情况 imu 经常容易跑飞；
8. 位置变化的话是不是从模具上搞比较容易一点？如果是无人车这种的是否通过判断角度和之前差距太大？
9. 首先是知道自己偏了；
10. 然后怎么纠正过来 系统能容忍多大的误差；
11. 怎么保证系统在各种情况下的稳定运行；
12. 像无人车和平面机器人，本来应该只有yaw的，如果算出来其他方向的旋转也比较大，是不是就可以代表出问题了；
13. 我不知道现在的SLAM领域 是否引入一些 控制系统里面的容错的adaptive 的方法；

主题：传感器模型 + VIO分类

传感器模型

1、几种IMU的数学模型

1. 最简单的就是“测量值=真值+bias+noise”模式，这个模式下面bias一般建模成random walk，然后还有 bias建模成markov process的。在vio里不多见，gps/ins里比较常用
2. 测量值=真值+bias+noise 这个模式忽略了3轴加速度计和陀螺仪的非正交误差 标度因数误差
3. imu标定中有一个非常有名的方法叫多位置法 就是用来标定非正交和标度因数，非正交指的是实际的三轴并非完全正交，标度因数指测量值和实际值存在一个固定的比例关系，标度因数英文也是scale 不要和视觉里面的scale混淆
4. 非正交误差和标度因素都需要乘个矩阵吧？是的
5. 多位置法是将IMU六个面轮着朝下一遍吗？这个有各种 我以前做的时候好像是24个位置，不光有朝下的，还有斜着的
6. 低精度的陀螺仪还有一个误差源叫 g-sensitivity，这个误差源一般针对mems器件（低精度mems），mems也有高精度的一般高精度的器件不会考虑这个g-sen，高精度器件一般会离线标定好非正交和标度因数，不会在线估计这些东西
7. 是不是误差模型越复杂，效果越好？
不一定，估计的参数多了 滤波器不见得会更稳定

相关文章

For consumer grade devices, the influence of axis misalignment and incorrect measurement scaling as well as of linear accelerations on gyroscope measurements (“g-sensitivity”) can be modelled as

$$\varpi(t) = \mathbf{S}_\omega \mathbf{M}_{\omega A} \omega_{WB}(t) + \mathbf{A}_{\omega A} \mathbf{a}_{WA}(t) + \mathbf{b}_\omega(t) + \mathbf{\nu}_\omega \quad (9)$$

where \mathbf{S}_ω and \mathbf{M}_ω are defined analogously to \mathbf{S}_α and \mathbf{M}_α in (4), and \mathbf{A}_ω is a fully populated matrix. Despite presumably having different displacements from the IRA, only a single lever arm is considered in the calculation of ${}_A \mathbf{a}_{WA}(t)$. In general, the effect of linear accelerations on gyroscope measurements is small and insufficient to properly constrain the estimate of a spatial displacement. In this work, we assume the accelerometer and gyroscopes to be sufficiently close and employ the lever arm estimated for the accelerometers. In cases where an individual lever arm per accelerometer axis

Extending kalibr: Calibrating the Extrinsic of Multiple IMUs and of Individual Axes

The gyroscope measurements are modeled as:

$$\omega_m = \mathbf{T}_g {}^B \omega + \mathbf{T}_s {}^B \mathbf{a}_s + \mathbf{b}_g + \mathbf{n}_w$$

where \mathbf{T}_g and \mathbf{T}_s are 3×3 matrices, \mathbf{b}_g is the measurement bias, and \mathbf{n}_w the measurement noise. Similarly to the case of

High-fidelity Sensor Modeling and Self-Calibration in Vision-aided Inertial Navigation 李明阳

2、Mems指的是什么

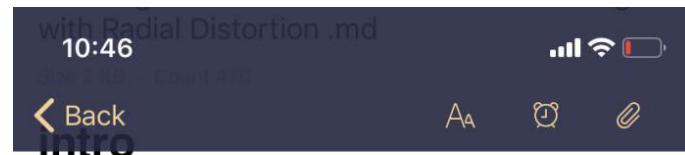
<https://baike.baidu.com/item/%E5%BE%AE%E6%9C%BA%E7%94%B5%E7%B3%BB%E7%BB%9F/1718090?fromtitle=MEMS&fromid=686299&fr=aladdin>

3、事件相机在SLAM中的数学模型

1. 没玩过事件相机，只是从资料看，推测它就是 vo
2. event我记得说是对于光暗变化剧烈的时候效果也不错
3. 事件相机，就是帧率高，低光照下也有特征点。没有imu也能跟踪的好。就是传感器好，而且还省电

4、鱼眼相机的高畸变在SLAM中是如何处理的？是畸变矫正还是特别的相机模型

1. 鱼眼用立方体投影模型比较好
2. 或者投影到一个平面，我觉得更好的思路还是用六面体把鱼眼拆成更多透视，而不是一个平面
3. 我觉得可以去看camodocal 就vins提供的那个，有很多相机模型



- 经向畸变随着wide fov 增加，造成一般的相机模型不再准确。
- 作者提出新的motion model来解决这种明显畸变情况下的feature tracking 问题。在可重复性和sfm进行了对比。效果好于当前的KLT tracking
- 宽的FOV相机，尤其是全景相机，在运动估计系统中更有优势，但是畸变也很大，而且畸变程度随着距离光心半径增加而增加
- image alignment techniques 应用在feature tracking上需要假设有运动模型的假设，从一般的translation model 到 affine motion model. 有论文分析发现设计合理的motion model对local feature tracking 的效果影响比较明显。但是传统的motion model 在经向畸变比较大的情况下并不适用。
- 目前KLT tracking已经在一些RD比较大的相机上使用。一些是直接在image registration 过程中忽略畸变；一些是采用预处理步骤处理畸变，去畸变这种方式的缺点是，在图像上进行插值操作是比较耗时的，另外因为增加了人工像素，也会造成不可靠。
- 论文主要聚焦于在有明显RD的图像上的feature tracking问题，论文的主要贡献是：
 - 1.融合feature motion和相机畸变的新的仿射模型
 - 2.在线估计RD



FOV model. This model is specifically designed for fisheye lenses and was first proposed by Fevernay and Faugeras [33]. It assumes that the distance of a given pixel in the image to the principal point is proportional to the angle between the optical axis and the ray connecting the optical center with the 3D point. This model assumes radially symmetric distortion, i.e.,

$$\gamma_F \left(\begin{bmatrix} u_u \\ v_u \end{bmatrix} \right) := \begin{bmatrix} u_d \\ v_d \end{bmatrix} := \frac{R(r_u)}{r_u} \begin{bmatrix} u_u \\ v_u \end{bmatrix}, \quad (2.6)$$

where $r_u := \sqrt{u_u^2 + v_u^2}$. The function R is characterized by only one parameter ω

$$R(r_u) := \frac{1}{\omega} \arctan \left(2r_u \tan \left(\frac{\omega}{2} \right) \right) \quad (2.7)$$

A useful property of this model is the existence of a closed-form inverse

$$\gamma_F^{-1} \left(\begin{bmatrix} u_d \\ v_d \end{bmatrix} \right) = \begin{bmatrix} u_u \\ v_u \end{bmatrix} = \frac{R^{-1}(r_d)}{r_d} \begin{bmatrix} u_d \\ v_d \end{bmatrix}, \quad (2.8)$$

with $r_d := \sqrt{u_d^2 + v_d^2}$ and

$$R^{-1}(r_d) := \frac{\tan(r_d \omega)}{2 \tan \frac{\omega}{2}}. \quad (2.9)$$

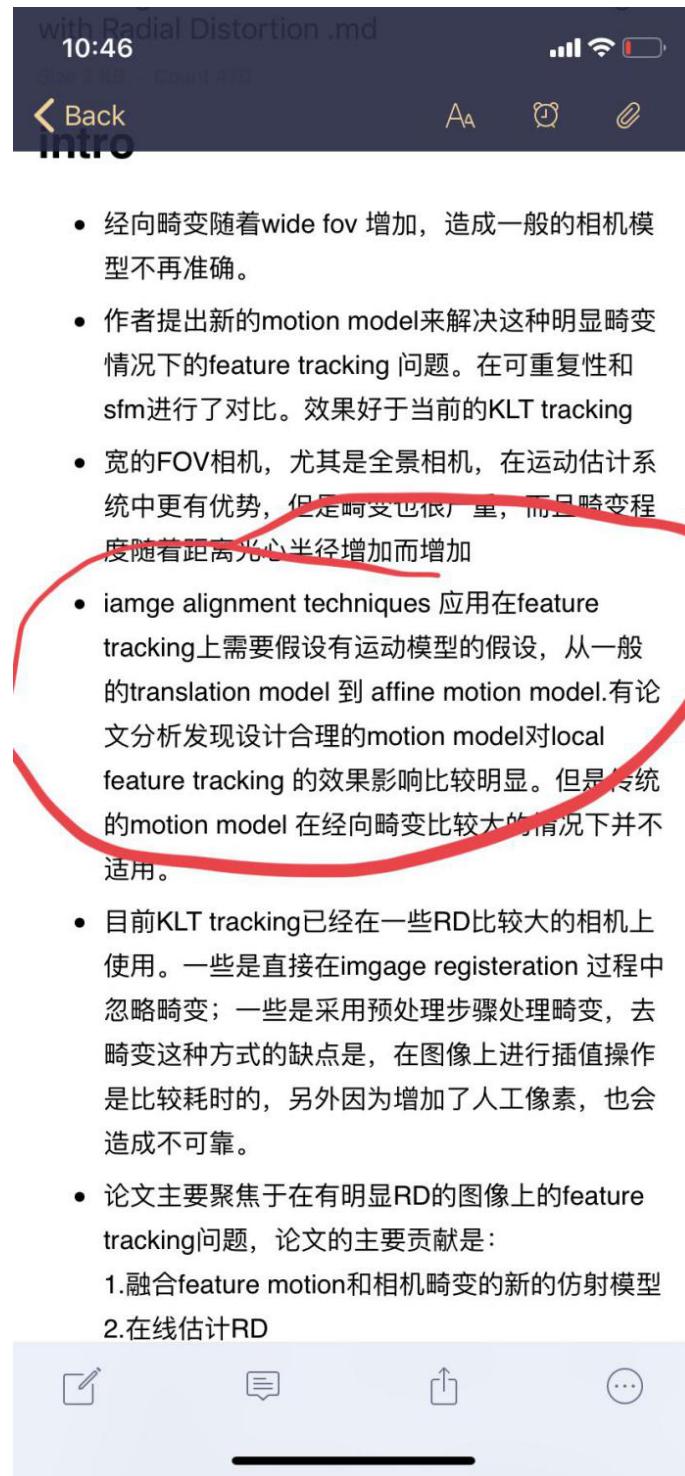
相关文章

Large-Scale Direct SLAM and 3D Reconstruction in Real-Time , PhD Thesis Jakob-Julian Engel

A Generic Camera Model and Calibration Method for Conventional, Wide-Angle, and Fish-Eye Lenses

5、畸变对feature tracking的影响

1. 有些论文提到做畸变矫正再tracking一个是增加计算负担，另外一个是矫正中使用差值方法引入人工像素，会对结果有影响，有的做法是修改klt (kanada lucas tracking) 的运动模型，之前假设是线性运动的，有的会使用考虑畸变的运动模型
2. klt跟踪的时候 有运动模型的嘛？不是直接迭代吗？



3. 还有这些运动模型的增加会不会引入额外计算量呢？会增加一些，大概15%这个论文提到，但是畸变大的情况下，效提升果比较明显。畸变小的情况没必要，所以要结合使用相机和参数设计motion model
4. 这里是物方还是像方的运动模型？就是是相机的运动还是像素的运动啊？应该是像素的，感觉是像方
5. 是指光流局部匹配时的模型吧？这个是说patch的模型？嗯，感觉是，我大概理解是之前是假设像素的patch是平的，现在不这么假设了，想当于对这个小patch做畸变矫正。畸变导致小patch形变需要设计模型来矫正这个东西再来算intensity error

6、卷帘相机的影响

1. 卷帘相机是不是可以理解，自身图像像素都不是同步的？是哒，卷帘门就是一行行或者一列列曝光的
2. 不能用于高速运动，自动驾驶好像都不用卷帘，主要是手机上

3. 那用卷帘相机有啥好处吗？便宜，现在全局快门也不贵

7、双目相机的数学模型

1. 不就是三角测量算深度么。然后，双目的图像时间要同步。
2. 标定，纠正使极线对齐
3. 自标定也要搞，主要是解决外参的微小形变。
4. 之前有看到一个说法，对极几何层面的双目标定，有优化极限，只能在对极几何误差上做到最优，但是这个结果不是最好的，所以有引入三视图做refine
5. 标定不好，匹配就不会好，深度图就不会好 极线对齐，就是双目不可能完全平行。通过标定，拿到偏移量(平移，旋转)，然后纠正图像。让图像变成平行的。
6. 其实我有一个问题 如果说是鱼眼或者广角的双目 是不是有另外一个什么约束 极线约束是针对透视投影的是吧
7. 我一直觉得把双目简化成一维 会不会增强了系统误差？有可能。不改极线，就做出深度图。太先进了，第一次听说。现在流程都是 标定，纠正，匹配，出视差图，出深度图
8. 包括s-msckf以及vins-fusion似乎都不校正？对，我就想说smsckf就是只用来初始化点 外参也不管直接当测量用
9. 其实我觉得多了一个同类型的约束 不见得多了就会精度有提升」 双目不仅是同类型约束吧 还有一个基线长度的新信息
10. 那倒是的 那我想想这个话应该怎么说，不过从重建的角度来讲 可能是你说的那个意思

8、双目相机应用中的关键问题

1. 标定
2. 同步也挺麻烦的
3. 在线标定 是个关键问题，主要解决长时间使用后，双目形变的问题。
4. 是不是缺乏一个精准的评判指标也可以认为是解决这个问题的困扰？是啊 我觉得这种东西很多时候只能从最终的定位或者重建结果来看是不是有效果，或者有一个比较好的仿真环境 可以开上帝视角
5. 小觅双目，板子后面在加强板。所以，双目外参的形变不会太大。如果，只是pcb板连接双目。双目长时间后，位移严重，影响效果。在线标定，解决的是外参的变化的问题。
6. 双目应用中的问题，如果用在建图上，就是 视角没有雷达广。雷达是360度。双目取决于FOV。肯定没有360度。
7. 我遇到的比较多的问题是， 双目跑vins飘——分析一下，还是用户对vio不了解。拿到相机，大部分对着白墙。或者，跑的时候翻腕。
8. 翻腕会导致双目飘？对， vins会这样，msckf没试过
9. 双目出深度图 都是用点吗，可以用线面做吗？有基于面匹配算深度的（面匹配，算力要大一些。还有，深度不连续的时候，有时候容易估计错，噪声比较多），目前deep learning做深度，效果也不错

VIO分类

1、在松耦合和紧耦合两类VIO中，有哪些相关工作

1. 紧耦合：vins, viorb, dsvio, svo2.0, msckf, iceba
2. 松耦合：松耦合利用信息不充分 已经没什么人研究了，soft slam里面有个松耦合的东西，用imu结果直接当旋转

2、VIO的分类除了紧耦合、松耦合，还有哪些类别

1. 直接法、间接法
2. 稀疏、稠密#

主题：初始化和预积分

1、VIO初始化的时候重力和加速度计的bias有没有估计准确一点的方法？

目前常用的初始化方法有VINS-Mono和VI-ORB两种，但差别不是特别大。除此之外，还有静止初始化，但是需要特定运动配合。特别的是，对于低成本的IMU，重力和bias初始化很难分离出来，所以也又在初始化不顾及bias的方法。但整体而言，在初始化方面需要进一步探讨。

参考文献

<https://blog.csdn.net/heyijia0327/article/details/82855443>
Visual-inertial monocular slam with map reuse

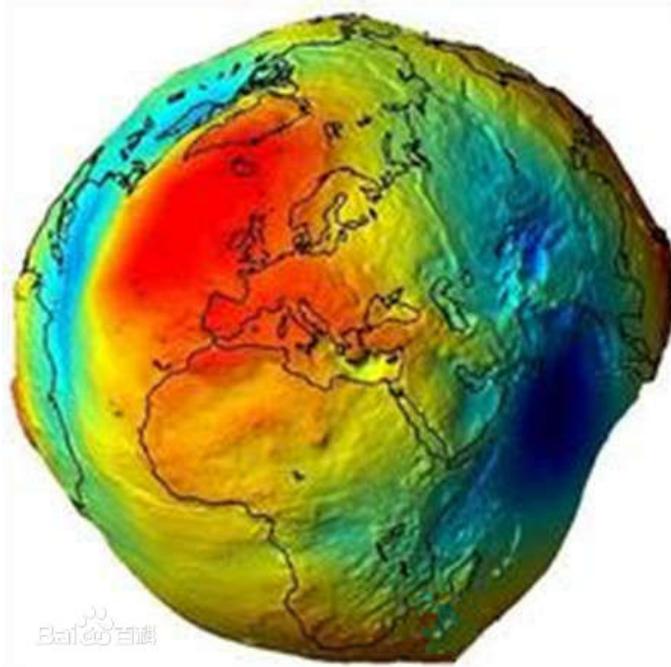
2、某些轴上的运动激励不够要怎么处理？

1、单纯使用IMU并不能解决此问题，通常需要使用其他传感器或者添加其他约束；其他传感器可以有轮速计、GPS等等；其他约束有高程约束等等。

参考文献

VINS on wheels

2、额外讨论，ins系统本身就存在高度通道发散的问题。原因是ins里面会严格根据高度和经纬度推算g：如果高度估计值比实际要高，那么估算出来的g会偏小，这样补偿后的加速度在高度方向上就会偏大，高度越发偏高。反之亦然。这会在高度通道形成一个正反馈，完全不稳定。（下图地球重力场）



3、额外讨论，在初始化时最好保证IMU各个轴上都有较大的运动，可以测量值幅值方差判断。

3、在尺度完全未知的情况下如何判断VIO初始化给出的尺度就一定是准确的？

一般尺度误差主要由ba (bias_acc) 引起，假如还是平面运动，天向的ba（由三轴关于天向的分量合成）就是不可观，那么这个方向的尺度就算不准也难以被发觉；并且即使在线运动激励不充分，有初始方差的限制，bias量也不会出现误差很大的情况；如果误差太大，则imu解算会与其他传感器对不上。

4、几种预积分方法的差异及区别？

1、常用的预积分有3套，第一套是12年最早的那一套 (lupton 2012)，用的是欧拉角来表示姿态；第二套是流型上的预积分 (forster)；第三套是vins里面的四元数预积分（沈老师搞出来的）。除此之外还有黄老师今年的工作，参考论文Closed-form Preintegration Methods for Graph-based Visual-Inertial Navigation。

参考文献

Visual-Inertial-Aided Navigation for High-Dynamic Motion in Built Environments Without Initial Conditions
 VINS-Mono: A Robust and Versatile Monocular Visual-Inertial State Estimator
 On-Manifold Preintegration for Real-Time Visual-Inertial Odometry

2、目前已知的三种方法的原理都差不多；但lupton基于欧拉角，存在万向节锁死问题；forster使用了李代数，沈老师使用了四元数，解决了欧拉角万向节锁死问题，并且还减少了cos和sin的计算量

主题：前端

前端所涉及的内容广泛，本次讨论主要关注特征点相关部分，其他内容在后续讨论中会进行补充。

1. 对于外点的处理目前有哪些方法，分哪几类？

在经典的slam框架中，前端对于特征点的处理，按照点对是否有对应关系可以粗略地分为两种，即有对应关系和无对应关系。

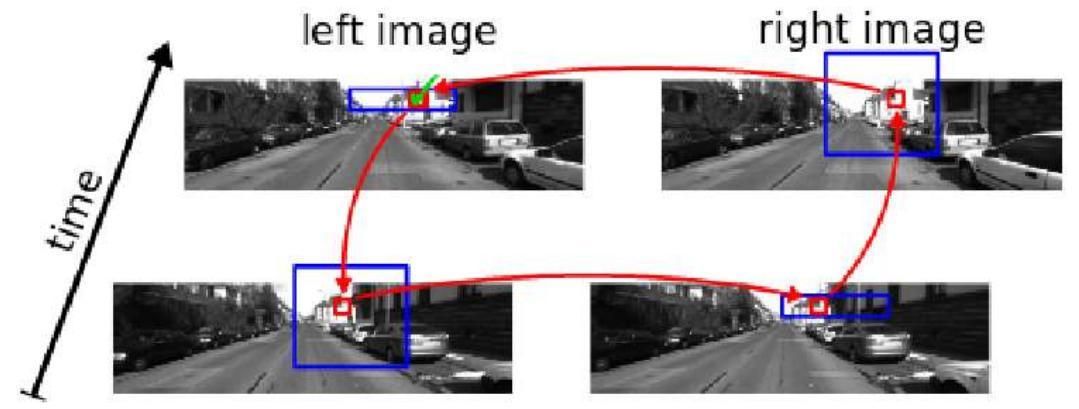
1.1 首先，对于具有对应关系的特征点来说，尽管对应关系存在，但对应关系的正确与否是未知的，外点即指那些具有错误对应关系的对应点，它们的存在会严重影响求解，因而有必要将外点剔除(outlier remove)，或选用对外点鲁棒的估计算法(robust estimation)。

- RANSAC类方法以及其变形。RANSAC是一种“假设检验”的思想，该简单却有力的经典思想于1981年提出，影响深远，后续的工作中有许多对传统RANSAC的变形。例如，在VIO里结合IMU以及运动模型，有2-ptRANSAC, 1-ptRANSAC。
- Hough变换。Hough变换的核心思想为“投票”，其实与RANSAC在某种意义上如出一辙。Hough目标是选取最高的投票数，RANSAC目标是选取最大一致集（允许精度范围内点个数最大），同时，经典的Hough变换提出之后，也确实有Randomized Hough变换延伸出来，其改进的地方即为投票的方式。这里还有一篇2013年CVPR文章可以借鉴，[Robust Feature Matching with Alternate Hough and Inverted Hough Transforms](#)，它的核心思路是内点的pose参数会聚合在一起，外点算出的pose则各不相同。

- M-estimation, 这里的“M”意味着最大似然估计，常用IRLS（迭代权重最小二乘法）来优化loss function。
- 全局类方法，其可以保证得到的解是全局最优解，而常见的RANSAC类方法不仅解是局部解，它的解还是随机的，不确定的。但往往由于全局类算法运行效率低，实用层面上还有一段距离。
- 确定性方法，这类方法不能保证得到的解是全局最优解，但可以保证无论运行多少次结果是唯一的，即消除了解的随机性。

前端完成对outliers的处理之后，仍不能保证剩下的都是inliers，可能有outliers进入后端。

- ekf方案中，在后端有一个基于卡方分布的gating test 也可以一定程度上剔除outliers。
- BA迭代过程中利用鲁棒核函数也可以剔除outliers。
- 测绘里还在BA的时候给残差加权重来剔除外点，这和核函数有点类似。
- 在circle match中，可以剔除outliers，下图是双目的示意图，若对单目来说，比较的就是前后两帧了。



1.2 关于无对应关系特征点，在激光SLAM中较常见，如无对应关系的pose求解问题。虽说在算法的使用过程中，不需要点对的对应关系，但在算法的求解过程中，寻找对应关系和求解pose是chicken-egg问题，只有找到对应关系才能相对准确求解pose，只有已知pose才能给出对应关系，如何找寻突破口迭代优化是这类算法的核心问题，因此这类算法本身也要涵盖对outliers一定程度上的处理能力。

2. 特征点提取方法中，关于识别性能和提取速度上，有没有能跟orb接近的，或者比他好的？利用深度学习来做特征点匹配，有没有比较成功的应用到SLAM系统中的？

这部分关于特征点提取的讨论内容，我们分为传统方法和基于深度学习方法。2.1 传统的特征点提取方法，由下图概括给出。在现有的开源的SLAM算法中，描述子提取特征点的方法还是orb，光流的方法还会由Harris 和Fast。这里提供一个比orb效率更高的特征点提取方法，[BFROST: Binary Features from Robust Orientation Segment Tests accelerated on the GPU](#)，本算法代码开源，比orb速度快约10倍左右，不过更适合并行化。

Traditional (slower, accurate):

1999 **Scale Invariant Feature Transform (Lowe)**

2006 **Speeded Up Robust Features (Bay, Tuytelaars, Van Gool)**

Binary (faster, real time, smartphone, performance):

2010 **Binary Robust Independent Elementary Features**

(Michael Calonder, et al.)

2011 **Oriented FAST and Rotated BRIEF (Ethan Rublee et al.)**

2011 **Binary Robust Invariant Scalable Keypoints**

(Leutenegger, Chli, Siegwart)

2012 **Fast Retina Keypoint (Alahi, Ortiz, Vandergheynst)**

2012 **Fast Retina Keypoint (Alahi, Ortiz, Vandergheynst)**

- SIFT: slow + good quality
- BRISK: fast + good quality
- FREAK \approx BRISK
- ORB: faster + littlebit less quality

2.2 基于深度学习方法，如开源的[GCNv2_SLAM](#)，是用深度学习提取特征描述子，替代orbslam中的orb描述子，网络训练好之后的使用阶段，特征点的提取可以做到实时，但是关于该开源方案的泛化能力还未尝试过。我们希望基于深度学习的方法能突破一些目前传统特征的瓶颈，比如白天夜晚光照变化大等恶劣情况。

3. 特征点在影像上的分布到底对VO的精度产生多大影响？有没有量化指标？

关于这点的讨论，转知乎上一讨论[ORB特征提取策略对ORB-SLAM2性能的影响](#)，作者在提取ORB特征时，尝试了OpenCV的实现和ORB-SLAM2的实现，特征点分布特点分别为“扎堆”和“均匀”，然后探讨特征点的分布对ORB-SLAM2的影响，结论简要概括为均匀化特征提取策略确实可以提高系统的精度和鲁棒性。

除此之外，为了提供系统的鲁棒性和精度，可以采用把translation和rotation分开解的策略，远处定rotation，近处定translation，例如ekf-slam。这种策略符合常理，比如人都是看远处来定方向的，可想而知远处的rotation即使有误差，也比近处的小，同时近处的特征点具有平移大的特点，分辨率也会更高。但为了区分远近，还是要先有一个depth map才行。

4. 深度学习网络来解pose是怎么编码到网络里的？

这里提供一篇CVPR2019单目深度估计的文章，[Digging Into Self-Supervised Monocular Depth Estimation](#)，一般求解pose，假设一个6个参数的矩阵，然后拟合出六个参数，即是用网络来拟合函数。

主题： VIO中的滑窗设计

本次讨论主要包含前端相关的两个问题和关于后端滑动窗口的四个问题。

前端

1. 基于描述子的光流意义大不大

光流追踪角点或特征点的时候加上周围一圈像素，像是ORB那样，保留周围的像素和被追踪像素的明亮关系，这样追踪的时候是否更鲁棒？注意并非光流跟踪整个patch，除了被追踪点，周围的信息保留的都是1和0的亮暗关系。

该方法具体可以参考以下论文：

H. Alismail, M. Kaess, B. Browning and S. Lucey, "Direct Visual Odometry in Low Light Using Binary Descriptors," in *IEEE Robotics and Automation Letters*, vol. 2, no. 2, pp. 444-451, April 2017.

可以第一帧提取描述子，光流跟踪的就是描述子，有时候也会使用特征点检测的点作为光流跟踪的初始点。但要注意而且光流跟的一些点可能根本不适合用于做描述子。用描述子初始化可以理解，但是如果光流角点+描述子，则就不一定好，因为特征点精度优于角点精度。

2. 直接法和特征法谁更具有优势？或者是两者可以互补？

直接法利用的信息更多，似乎更有潜力做的比间接法更好。但直接法对光照和旋转不太好，特征点的描述子对这两种情形有鲁棒性。注意到，弱纹理下肯定直接法比较好。

直接法少了提取特征的过程，可以提高帧率。关键是快速运动的时候，如何平衡直接法和间接法的帧率，从而解决快速运动模糊的问题。

直接法和间接法结合的工作可以参考以下论文：

S. H. Lee and J. Civera, "Loosely-coupled semi-direct monocular slam," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 399–406, April 2019. S. Li, T. Zhang, X. Gao, D. Wang, and Y. Xian, "Semi-direct monocular visual and visual-inertial SLAM with loop closure detection," *Robotics and Autonomous Systems*, vol. 112, pp. 201 – 210, 2019. X. Gao, R. Wang, N. Demmel, and D. Cremers, "LDSO: Direct sparse odometry with loop closure," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2018.

后端

后端所涉及的内容广泛，本次讨论主要关注滑窗相关部分，其他内容在后续讨论中会进行补充。

1. 为什么要引入滑窗？引入滑窗的优劣？滑窗marg出先验项作为先验信息的作用大不大？

引入滑窗的主要目的在于：减小状态维数，减小计算量。滑窗算是精度与速度的一个平衡，如果不使用滑窗的话优化时就要求解历史的所有信息。

使用滑窗是为了减少优化时间，损失了一点精确性。

比如DSO里面使用滑窗，是因为我们在使用的时候做优化时，随着时间的进行，会导致优化的问题越来越大。因此，DSO会维护一个大概七帧的Hessian阵以及两千个点的Hessian阵。边缘化的时候，会把不要的点和帧从海塞阵中去除，得到一个先验信息，把这部分先验存储到另外的帧或者点中。DSO这种用法给人感觉就是ORB-SLAM里面的局部优化，也就是local ba，一次只要局部共视图里的帧进行优化。

注意到DSO是用marg弄出了先验项，ORB则是在滑窗里固定一些节点来保持一致性，以此来作为先验信息。如果没有滑窗marg出的先验项的话，优化问题的解会在零空间上乱跑。

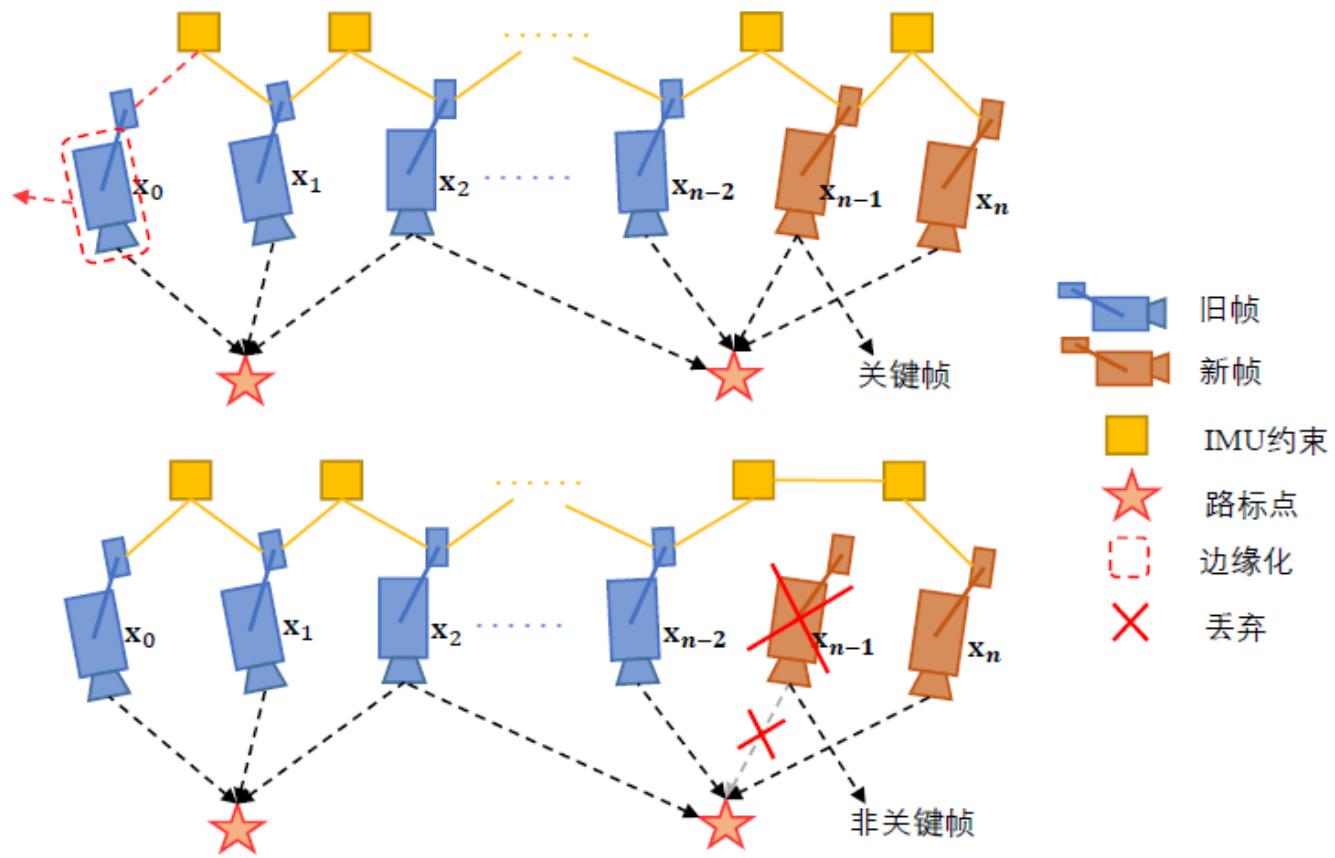
2. 滑窗这种算法思想都在哪些领域有应用，有哪些变种？

几项使用滑窗思想的经典工作比如：VINS，VI-ORB，MSCKF，MSCKF-VIO等。

滑窗的变种有时间滑，共视滑，空间距离等。各自的代表方法就是DSO和ORB。

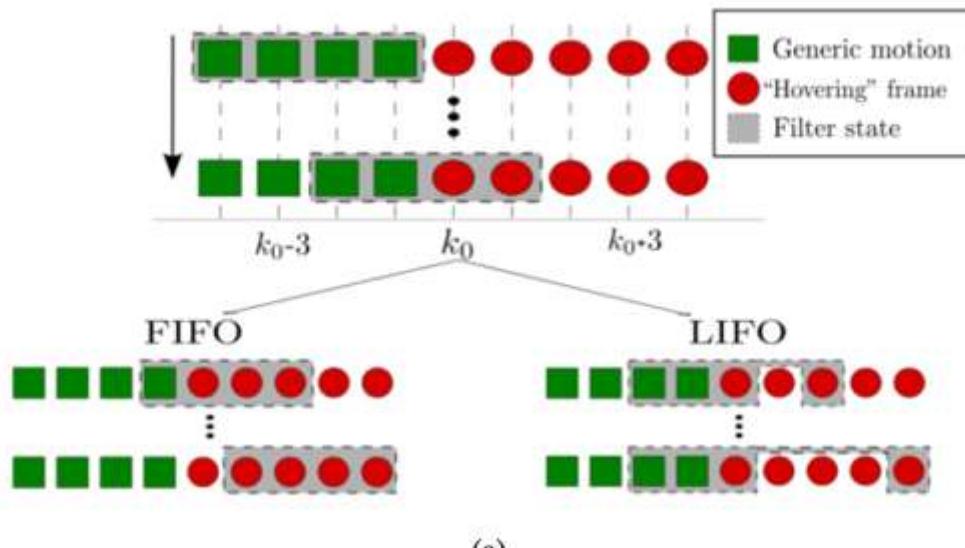
DSO不希望窗口里的帧离得太近，DSO的共视滑是每一帧的内点至少是5%，当然这个要求是比较容易满足的。所以它就再设置了一个空间距离上的，任意一个帧跟其他帧的空间距离合最近的剔除。从根本上，它就是不想帧间太近，而是可以在空间均匀分布。

以下是两幅滑窗的示意图：



而在下文工作中，主要思路在处理无人机悬停时候，根据如果是hovering状态，window是LIFO的；如果是正常状态window 是FIFO。判断是不是hover的标准是视差。

D. G. Kottas, K. Wu, and S. I. Roumeliotis, "Detecting and dealing with hovering maneuvers in vision-aided inertial navigation systems," in 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, Nov 2013, pp. 3172–3179.



3. 滑窗的参数设置（如窗口大小）有什么原则和规律？

原则是保证window里一定有足够的视差的pose。具体可参考：VINS， MSCKF-VIO。

4. 一般来说，滑窗大一点先验信息会好一些，但是速度会慢。有精度要求不能简单减少滑窗规模的情况下，如何优化？

首先明确的是未必滑窗大一些先验信息会好。先验信息倒确实会通过共视共约束关系来传递的，但是如果先验就是对差不多数量的节点有直接约束，那后面滑窗规模多大也不一定有很大作用。

第二个小问题很细节，要根据实际情况来调。

主题：未来SLAM讨论

1. 未来SLAM应该更加关注机器人对场景的理解，根据场景的动态变化规划自己的路径。

本次讨论中，机器人场景理解是计算机视觉研究一直研究的问题，目前是SLAM与深度学习结合比较多的部分。对于场景的理解是无线追求的，但是得考虑工程应用的必要性和制约性。SLAM现在更多被看成是解决工程问题，比如对场景的理解，这个理论的框架本质上不是属于SLAM这个范畴的，只是整个识别中的一个部分，然后应用到SLAM中。场景的理解还是得根据实际情况的条件而定。

此外还有一个问题就是对于场景理解的定义。这个定义本身比较模糊，讨论过程中由于对于场景理解的定义比较模糊，所以有许多小伙伴讨论起来有点没有准确目标。本人理解机器人场景理解是在运动过程中对周围场景的静态和动态的运动都能跟踪理解然后，做出路径规划和运动控制。现在SLAM中加入一些语义理解例如对于从添加语义标签的语义，来辅助机器人再环境中进行运动。技术流程可能是训练点云，图像，告诉机器人这是冰箱，然后机器人在原来地图上搜索吧

现在深度学习和slam结合的方向可以大致分为两块，松耦合和紧耦合。松耦合是对SLAM框架中一些组成部分进行修改，如前端的位姿估计，后端的建图部分进行修改，深度学习的结果可以作为SLAM框架继续运行的数据。紧耦合在场景理解中的使用深度学习算法构建深度SLAM框架。

2. 如果地图能够满足机器人导航的需求，地图的形式不再重要。

讨论中大家觉得地图的形式还是很重要，业界需要有统一的标准来共同打造解决方案。

地图是一种数据格式或者表示方式，应该有统一的标准。当前在智能车上面的高精度地图一般遵守国家上公认的两种协议，我国用opendrive的比较多，不过也可以是地图的形式不一样，但是抽象出来的对外接口有同样的标准。地图表达格式和标准是不是以后会根据传感器类型，或者有更高要求的地图形式表示而改变。

方宇提出拓扑地图这种形式比度量地图在机器人导航里面应该更有优势，不知道大家怎么看？感觉如果是为机器人导航而服务，可以建立拓扑地图，然后机器人在运动过程中逐渐去判断自己距离当前目标点的距离。王琛提出该是一个技术方向，但是自己计算距离会不会带来更大的计算量。几何度量地图也不一定就不能再减少体积，我觉得都是一些权衡。在拓扑地图里面融入尺度信息。现在我觉得一个工程问题抛出来说应该怎么做都是需要和实际应用结合的，看实际应用的限制和需求。

机器人大人导航我感觉根本不是规划出最优路径局部规划一下，而是局部规划。

3. 精度和鲁棒性应该是SLAM肯定要追求的目标，目前机器人还不够智能？

精度在技术上是永远要追求的，鲁棒性是永远要追求的。我的理解是够不够智能也是要看用在什么地方的，可能这个场景下已经足够智能了，只不过技术上的目标是希望它像人一样智能。智不智能主要看能不能适应场景。slam的建图过程和人记忆地图的过程有点像。有同学提出现在的gnn因子推理图可能是机器人智能化的趋势。gnn处于萌芽阶段，[介绍GNN因子图的博客 GNN综述文章 GNN综述文章2](#)

王琛：我觉得我需要反问一个问题，理解能力有标准定义吗？我们怎么知道它提高了？

这个问题还是对理解能力定义的标准，目前没有一个明确的标准。王琛建议，那就需要结合具体任务说明了，我同时认为不同种类的任务哪一个更复杂有时候无法区分。但是这些问题比较可能是各自制定的标准，到底哪个大哪个小也没有统一的标准，我认为一旦涉及到更，和提高，必须得有具体的定义和标准。

4. SLAM和深度学习都有哪些结合的点？这些点目前只是前沿探索，还是已经有实际落地应用了？

对于深度学习结合SLAM落地实际应用目前来说还是比较稀少，但是深度学习和SLAM结合的论文和代码还是有不少。例如cubeSLAM是一个深度学习与SLAM结合的比较好的，代码也开源，发表在TOR上的，TMU的CNN-DVO。

如果有传感器数据作为训练数据输入，和深度网络联合训练，然后再运行的时候输入传感器数据，这个思路可行吗？直接端到端进行训练就可以了。还有一个问题，同一个系统中同时运行多个任务的神经网络，三个及以上的时候有没有什么好的方法。不用考虑那么多，拉起直接跑，只要电脑性能好，想跑几个跑几个。有一个比较好的代码环境配置方法，你只用把一个代码搞好，封装成一个镜像。docker并不怎么耗资源，cuda有个docker定制版本的，只能在ubuntu上用，GitHub上能找到。

上周遗留问题讨论

1. 舒尔补解

[舒尔补解综述文章](#)

1) 什么是舒尔补解

舒尔补是margin的一种方法。舒尔补形式上我认为就是高斯消元，只不过我们手动高斯消元是将 $Ax=b$ 的A转换为上三角(或下三角)，然后一个个元素输出来。只不过舒尔补是先将A通过分块消元，如果将分块看成是 2×2 的矩阵，实际上也就是一个 2×2 矩阵的高斯消元。Sliding window filter with application to planetary landing

2)舒尔补解在高斯消元的时候有什么优势呢？

舒尔补我理解就是在边缘化的时候，为了保留被边缘化的重要信息，从而进行的降低求逆矩阵的维度，提高计算效率，高斯消元也会保存原有矩阵信息，不会改变 $Ax=b$ 的解。

3)舒尔补中分块求逆如果有不可逆的情况时候如何处理呢，实际问题中怎么解决不可逆问题呢？

不可逆，就是 rank deficient，SVD 分解，中间对角阵，element wise 倒数。舒尔补是作用在分块肯定有逆的矩阵上的。

舒尔补这个东西用在slam里，正如前面心亮说的，应该从概率方向来理解。协方差矩阵在去掉一些块的时候，正好对应信息矩阵做舒尔布。或者由多元高斯分布的概率密度做marg也能很清晰的看出来。 $\exp((x-u)^T A(x-u))$ 这个里面A就是信息矩阵，做完marg正好会剩下舒尔补的形式。

2.FEJ问题

因为信息矩阵中同一个状态有不同的展开点，导致零空间发生变化（可观性发生变化。于是引入fej 让相同状态的线性展开点一致 来修正这个变化。这个东西其实可以由非线性最小二乘的法方程来理解： $A\hat{x}=-d$ 。这个A其实就是信息矩阵， d 在A的零空间上无论发生什么样的变化 都不会影响这个等式的成立。也就是说在某些方向发生变化，不会影响最小二乘指标。这样很直观的就知道这些维度不可观了，A只可能是因为数值原因奇异。如果是严格的信息矩阵不可能奇异，因为协方差矩阵肯定是可逆的 因此他的逆——信息矩阵 也肯定是可以逆的。

A 是信息矩阵这个一直不太理解？信息矩阵不就是协方差矩阵的匿吗？有人认为不就是手动设定的吗？对角元非零且对称。 $J\Omega\omega J$ 这个东西就是信息矩阵该有的形式。

游振兴：感觉晨哥刚才那个从方程理解客观性挺有趣的，是不是可以这么理解， A 的秩对应可观的维度，其他不可观的对应零空间，所以无法求解优化。其实这个我认为和高斯消元等价。其实是从高斯分布的信息矢量与信息矩阵的表示形式推出来的。 $A=[1,0;0,0]$, $x=[0;k]$, k 随便怎么取，都不影响 $Ax=0$ 。这个方程对应SLAM优化的过程是怎么对应起来呢？就是迭代过程中的法方程，实际 $Hx=b$ 。 H 的质 (md找了半天没找到对的字)=可观状态维数。所以对不可观的维度比如yaw做优化其实是没意义的了是不是可以说，相对yaw还是有用的。

不可观的维度没法优化，不管怎么优化都对你的能量没有影响。但从物理意义上来说相对变化还是可观的？一般都是说不可观的维度是全局的，但从物理意义上来说 相对变化还是可观的。

如果信息矩阵精准的话 直接看零空间对应它的哪些维 然后咱们知道哪些维是哪些状态 就知道哪些状态不可观。我们在初始化的时候其实是固定了这些不可观的初值，比如yaw=0,全局位置等于0。感觉是不是可以从这个角度去看数学上有什么变化，如果是固定不可观（零空间维度上）的值，对最后的cost function收敛理论上是没有影响的。

王琛：但是有时候看不出是具体哪个维度不可观，比如尺度（随便举个例子），有时候可能是参数化的问题，如果某种参数化方法能把这些维度隔离开，那就又不一样。

笑晨：但是有时候看不出是具体哪个维度不可观，比如尺度（随便举个例子），有时候可能是参数化的问题，如果某种参数化方法能把这些维度隔离开，那就又不一样。这个东西反正我没想透，我经常这样想：比如在某一次SW里面，它已经有先验的yaw或者其他东西，如果零空间正常，这些东西应该就不会有不正常的变化（它们的变化只受相对约束的影响），但是如果人为加一个变化，在法方程上也反应不出来。[EKF论文](#)

主题：VIO中后端优化相关问题

1.SLAM中常用的优化方法有哪些？

1.1 [Gauss-Newton；LM；Dog-leg；Conjugate Gradient\(CG,共轭梯度法\)与Preconditional Conjugate Gradiem method\(PCG,预优共轭梯度法\)](#)

1.2 GN,LM,Dog-leg三种优化方法的分析与对比见[链接1](#) [链接2](#). CG与PCG方法的原理与实现见[链接](#).

1.3 开源VIO算法中VINS用到了Dog-leg. 项目[EIBA](#)确认用了PCG, [ICE-BA](#)也用了.

2.一阶与二阶优化方法的区别？

2.1 一阶与二阶的定义：

从 cost function开始说，构建error function 之后，就有了cost function. 我们对cost function进行泰勒展开

In this section, we will briefly explain the reason of poor convergence by discussing the approximation of the Hessian matrix, which serves as the normal matrix in both the GN solution and LM solution. The function $f(\mathbf{X})$ in (3) can be approximated using Taylor series expansion around the current guess \mathbf{X}_k , written as

$$f(\mathbf{X}) \approx f_k + \mathbf{g}_k(\mathbf{X} - \mathbf{X}_k) \quad (4)$$

here, $\mathbf{f}_k = f(\mathbf{X}_k)$ and $\mathbf{g}_k = \nabla f(\mathbf{X}_k) = 2\mathbf{J}_k^T \mathbf{r}$, where \mathbf{J}_k is the Jacobian matrix calculated using the first derivate of observations with respect to the unknown parameters \mathbf{X} .

To determine the descent direction and step size, we can compute the incremental values $\Delta \mathbf{X}_k$ by enforcing the first derivate of the function (4) as zero, i.e.,

$$(\mathbf{J}_k^T \mathbf{J}_k + \frac{\partial^2 \mathbf{r}}{\partial \mathbf{X}^2} \mathbf{r}) \Delta \mathbf{X}_k = -\mathbf{g}_k \quad (5)$$

Here $\mathbf{J}_k^T \mathbf{J}_k + \frac{\partial^2 \mathbf{r}}{\partial \mathbf{X}^2} \mathbf{r}$ is called a Hessian matrix, and then the solution is termed as Newton solution. In this letter, the first and the second items are called the *L-derivative matrix* and the *H-derivative matrix*, respectively. Many experimental results in public

其中泰勒展开的阶次则对应着一阶和二阶。

2.2 一阶是最速下降方法，有可能收敛慢，不一定准；二阶比较准，不容易奇异Hessien,较慢。GN方法是近似的二阶。

一阶，二阶，GN，LM方法的分析与速度对比见[链接](#).

2.3 优化问题几大要素

- costfunction的构建。
- 初值的选取（一般代数法结果作为初值，然后数值迭代）
- 迭代方向的确定
- 步长的确定（牛顿法是木有的，GN的步长默认为1）

2.4 深度学习优化方法总结见[链接](#).

3.什么时候使用线搜索或者置信域方法？

3.1 [线搜索与置信域](#)的解释与说明。

3.2 讨论结论来看，大家一般都可能用直接现成的库，其实内部核心只是配置相关参数而已。

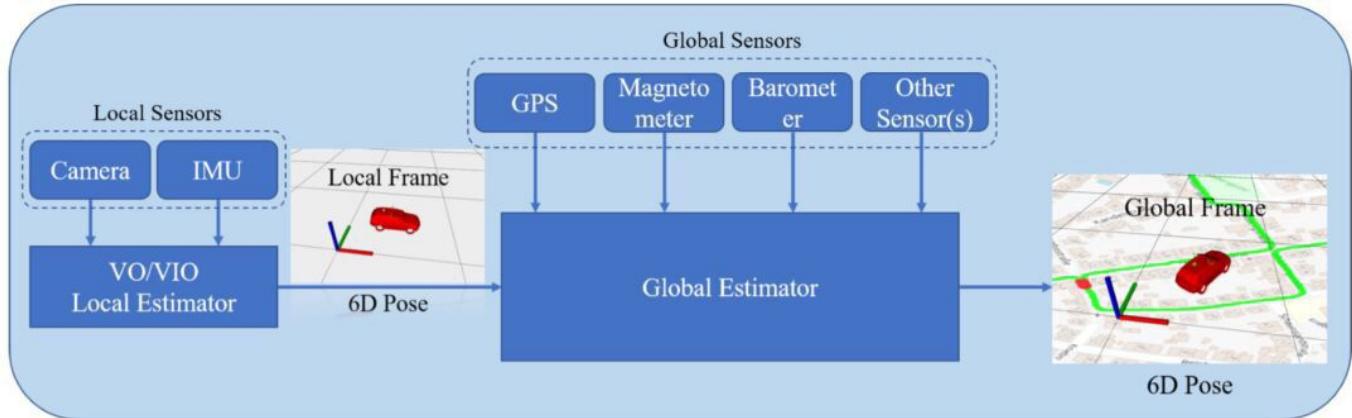
4.深度学习用于SLAM优化求解器的前景怎么样？

虽然也有很多Demo在做，譬如无监督的端到端的[unDeepVO](#).但是泛化能力一直是一个很难解决的问题，现在得到的反馈都是在数据集上确实能有很好的效果，但将模型运用到自己数据集上并不能得到很好的效果。而且端到端的问题是说有可能是一个黑盒，没有很好的办法去debug。

5.多传感器融合的约束对于后端优化的帮助有多大？（比如vins现在开源的接口添加了各种全局位姿传感器的目的是消除vo的累积误差）

5.1 BA (Bundle Adjustment) 又叫捆集调整，摄影测量学又叫光束法平差，测绘学又叫空三（空中三角测量）。

5.2 多传感器譬如IMU的引入能够减小累计误差，而且IMU与Vision是相辅相成的。[Vins-fusion](#)中IMU与视觉是紧耦合的，vio与GPS等等是松耦合的。



5.3 GPS以及回环都是可以减小全局误差的，另外讨论中提到了回环问题，其实也算是Pose Graph问题，它形象上讲是将误差匀一匀，但不一定是均分，使得整体误差最小，因为每个节点对总体误差的贡献程度不一样。

主题： VIO回环检测问题汇总

问题罗列

1. 目前比较实用的闭环检测方法都有哪些？还有类似DBoW2这样的方便使用的库吗？
2. 闭环检测中关键帧的选取机制是什么？未来的方向是什么？

3. 闭环检测可能带来误匹配的情况，有什么约束可以降低这种概率的发生？比如，在一个区域的不同位置，采集到的图像画面是非常相似的情况？
4. 在VIO中，IMU信息对闭环检测是否有帮助？
5. 闭环检测和重定位的关系？检测到回环后如何进行重定位和整体优化？pose graph是闭环优化比较经常采用的方法，那么还有哪些别的策略？
6. 基于深度学习的闭环检测的大概原理和流程是什么？目前有什么新的进展吗？
7. 如何提高闭环检测的实时性？深度学习的闭环检测有实际应用么？它相对一般的闭环检测的优缺点？
8. 大规模时空环境下，比如一年四季白天黑夜的城市环境下，地图如何维度闭环检测和更新？

泡泡讨论

1. 目前比较实用的闭环检测方法都有哪些？还有类似DBoW2这样的方便使用的库吗？关于图像的目前了解到的闭环检测方法有**FAB-MAP**和**RTAB-MAP**两种，其中FAB-MAP是第一个能够实时进行大规模场景的回环检测应用，但是实时性依然不佳，后来出现的DBoW2比较适合。RTAB-MAP通过缓存的方式构建STM、WM、LTM三个缓存工作栈，实现一个长时间的大规模室内场景闭环检测。其中STM\WM\LTM是Short Term Memory, Working Memory, Long Term Memory。只有Working Memory用于闭环检测，Short Term Memory的话是存放当前帧附近的图像，这部分图像由于视觉变化不大，所以不用于闭环检测，Long-Term Memory是用于长时间的闭环检测，这部分图像由于访问时间过长，并且访问次数过少，所以也减少匹配次数。闭环检测也是数据检索的过程，**DBoW**主要的作用就是特征向量的检索，本质上也就是**kd-tree**，类似的还有inverted multi-index。除此之外，还有facebook出品的faiss，包括了各种检索方法，还提供gpu加速功能。它的主要面向对象是超大规模检索。除此之前GPS可以判断是否在历史位姿附近，lidar可以通过直方图的方法实现回环检测。
2. 闭环检测中的关键帧选取机制一般是什么？未来还有哪些改进的可能呢？**关键帧的选取一般都是采用经验公式**，比如VINS通过Delta transformation判断匹配数超过一定阈值后，选取作为关键帧。ORB-SLAM中关键帧依据重叠度不超过90%来进行筛选。
3. 闭环检测可能带来误匹配的情况，有什么约束可以降低这种概率的发生？比如，在一个区域的不同位置，采集到的图像画面是非常相似的情况？**回环的目的，就是在历史轨迹中找到当前帧与历史帧中匹配分数比较高的一帧。共视图中的关键帧重叠度比较高，检索的结果也相对较高，所以容易造成误匹配**。假设有30个关键帧，后10个关键帧构成第30帧的共视图，那么也就是说这些帧都与第30帧有较高的重合度，换句话说就是共视特征点比较多。那么这些帧在bow匹配的时候，分数就会比较高。而历史帧，通常分数可能没有这么高，结果就是在bow返回的几个候选帧中，共视图相关的关键帧排序就会比较高。**所以在实际应用中，回环检测过程往往不采用共视图中的关键帧，而在闭环检测的过程中才将其取出进行整体优化，ORB-SLAM中还要求地图保持一致性。也有方法是通过每个关键帧所带的ID，直接去除共视图中的关键帧，防止回环候选帧与当前观测帧太近**
4. 在VIO中，IMU信息对闭环检测是否有帮助？IMU本身的自由度就已经决定了不能将其用于回环
5. 闭环检测和重定位的关系？检测到回环后如何进行重定位和整体优化？pose graph是闭环优化比较经常采用的方法，那么还有哪些别的策略？回环检测的关键帧都是历史帧，不在滑窗里，所以需要要进行全局优化。而闭环检测就是在全局的姿态图优化中，多加了一条边。现在也有通过训练多种特征（点、线）的BOW来进行回环操作。
6. 如何提高闭环检测的实时性？**闭环检测的实时性本身就是伪命题**，实际中闭环的主要应用场景是建图过程，而整个建图过程都是offline，另开线程来进行操作，所以没必要对实时性进行要求。
7. 基于深度学习的闭环检测的大概原理和流程是什么？目前有什么新的进展吗？深度学习得到的就是个描述向量，本身可以看成构建一个大词袋。**hfnet分了两个端，一个做全局的图像描述，一个提取局部**

特征，用全局找出候选匹配帧，再用局部特征做匹配验证求解位姿，但回环的目的是提供一条约束边，如果用深度学习来做检测，那么得到的仅仅位置识别的结果。并没有几何验证。深度学习通常直接通过暴力求解相似性来判断场景相似程度

8. 一年四季白天黑夜的城市环境下，地图如何维度闭环检测和更新？ seqSLAM。

主题：Lidar

1. lidar的分类以及研究方向，应用方向

- 机械旋转式 (velodyne64)
- mems半固态 (旋镜，16线32线那种)
- flash式 (固态雷达主要在做这个)，相控阵式
- 主要拿来做perception localization odometry
- 应用在自动驾驶，扫地机器人（二维Lidar），无人机测绘
- Lidar的应用场景是取决于应用场景，相机受限于FOV，单位时间内获取的周边环境有限，且对光照敏感，缺失尺度信息。无人机上2D lidar和3D lidar的选择我觉得也是取决于场景和指标限制，2D相比于3D轻一些，对载荷压力比较小，但是在需要大速度高机动的场景下，3D Lidar+点头舵机的配置是比较常见的

2. lidar的内参标定（多根射线之间的相对位姿关系），同型号的不同lidar之间的intensity标定

- 多条射线间的应该是厂家固定了；
- intensity标定比较复杂，分为按模型固定变量来标和数据驱动的方法来标，可以参考：Unsupervised Calibration for Multi-beam Lasers

3. lidar与相机，IMU之间的标定方法以及各自的优缺点

- lidar-imu 车画8字 imu有个轨迹 lidar有个轨迹 轨迹去对齐
-

4. 以64线雷达为例，一般处理前后两个样本时间里面的运动，需要什么硬件配置？如果是嵌入式设备呢？

- 一般处理都要先降采样的吧，当然硬件配置越高越好

5. 激光中的特征点如何计算？角点抽取除了利用曲率还有没有其他方式？具体地，车道线特征点提取方法，匹配方法？

- 至于特征就有好多啦，一般都基于统计、直方图，如点云的FPFH, SHOT特征
- 逐帧来的数据里面lidar是很稀疏的 所以大部分是有一条线上做文章
- 主要是曲率或者你叠了几帧 地图稠密了 地图里面找角点

- lidar一般是已经有了很好的位姿，把很多帧，很多批次的数据叠在一起，车道线上的点反射强度（也就是intensity）一般与旁边的路面不一样，用反射值提取车道线，不过那个已经是高精地图层面了

6. 多线激光生成的点云转化成图片有哪些方法？其中有哪些可以无损的还原出点云？

- 根据每条线变成图片的一行，64线就是64行pixel的图片
- 激光雷达点云转化成图片的方法我提名SqueezeSeg的，球体投影

7. 基于lidar的深度学习模型有哪些；lidar数据结合dl做目标检测（我记得kaiming he去年有相关文章）

- PointNet, PointNet++, PIXOR, VoxelNet, SqueezeSeg
- kitti 3d 检测榜 http://www.cvlibs.net/datasets/kitti/eval_object.php?obj_benchmark=3d

8. lidar和轮速计、imu的融合（扫地机器人？）

- imu其实正确做法应该就是vins那一套 套到lidar上
- LINS: A Lidar-Inerital State Estimator for Robust and Fast Navigation

9. 不回环，没有gps，如何利用lidar提高车辆的实时定位精度。

- 回环和定位没有关系 实时定位就是看你online的匹配精度。匹配可以加个motion model 或者用其他传感器给个初值

10. lidar在回环中的应用

- 用gps
- lidar可以对360°场景提取全局描述子，相比相机不容易受视角变化和天气、光照影响

11. lidar全局优化方法有哪些？