

干货总结 | SLAM 面试常见问题及参考解答

之前我们分享过视觉SLAM找工作、面试经历，见《[2018年SLAM、三维视觉方向求职经验分享](#)》，《[经验分享 | SLAM、3D vision笔试面试问题](#)》。

从零开始学习SLAM知识星球里，会定期发布一些常见的SLAM问题引导大家讨论，并给出参考解答。以下列举几个已经发布的问题及回答。

1、视觉SLAM方法一般分为特征点法和直接法。请简述一下特征点法和直接法的概念，以及对应的优缺点。

特征点法，根据提取、匹配 特征点来估计相机运动，优化的是重投影误差，对光照变化不敏感，是比较成熟的方案。常见的开源方案 比如ORB-SLAM

优点：

- (1) 特征点本身对光照、运动、旋转比较不敏感，所以比较稳定
- (2) 相机运动较快（相对直接法来说）也能跟踪成功，鲁棒性好一些
- (3) 研究时间较久，方案比较成熟

缺点：

- (1) 关键点提取、描述子、匹配耗时长
- (2) 特征点丢失场景无法使用
- (3) 只能构建稀疏地图

直接法，根据相机的亮度信息估计相机的运动，可以不需要计算关键点和描述子，优化的是光度误差，根据使用像素数量可分为稀疏、半稠密、稠密三种。常见开源方案有SVO, LSD-SLAM

优点：

- (1) 速度快，可以省去计算特征点、描述子时间
- (2) 可以用在特征缺失的场合（比如白墙），特征点法在该情况下会急速变差
- (3) 可以构建半稠密乃至稠密地图

缺点：

- (1) 因为假设了灰度不变，所以易受光照变化影响
- (2) 要求相机运动较慢或采样频率较高（可以用图像金字塔改善）
- (3) 单个像素或像素块区分度不强，采用的是数量代替质量的策略

2、视觉SLAM常用的相机包括，单目，双目，RGB-D相机，请分别说说它们本身的优缺点、常用的相机型号等。

以下是我使用时的一些总结，可能有疏漏错误，欢迎补充指正。

单目相机：

常用型号：有非常多的种类可以选择

优点：

- 1、应用最广，成本可以做到非常低。
- 2、体积小，标定简单，硬件搭建也简单。
- 3、可以用于室内和室外（有适当光照条件下）。

缺点：

- 1、具有纯视觉传感器的通病：在光照变化较大，纹理特征缺失、快速运动导致模糊的情况下无法使用（睁眼睛）。
- 2、SLAM过程使用单目相机有尺度不确定性，需要专门初始化。
- 3、必须通过运动才能估计深度（帧间匹配三角化）

双目相机：

常用型号：Indemind，小觅，ZED等

优点：

- 1、相比于单目，在静止时就能够根据左右相机视差图计算深度。
- 2、可测量距离可以根据基线调节。基线距离越大，测量距离越远。
- 3、可以用于室内和室外（有适当光照条件下）。

缺点：

- 1、双目相机标定相对复杂
- 2、用视差计算深度比较消耗资源
- 3、具有纯视觉传感器的通病：在光照变化较大，纹理特征缺失、快速运动导致模糊的情况下无法使用（睁眼睛）。

RGB-D相机：

常用型号：Kinect系列、Realsense系列、Orbbec、Pico等

优点：

- 1、使用物理测距方法测量深度，所以避免了纯视觉传感器的通病，在没有光照的情况下、快速运动的情况下都可以测距。这是非常大的优势。
- 2、相对双目，输出帧率较高，更适合运动场景。
- 3、输出深度值比较准，结合RGB信息，容易实现手势识别、人体姿态估计等应用。

缺点：

- 1、测量范围窄，易受日光干扰，通常只能用于室内场景
- 2、在遇到透射材料、反光表面、黑色物体情况下表现不好，造成深度图缺失
- 3、通常分辨率无法做到很高，目前主流分辨率VGA（640x480）
- 4、标定比较复杂。

3、关键帧在SLAM里应用非常多，很多知名的开源算法都使用了关键帧。请你用自己的语言描述一下关键帧是什么？有什么用？如何选择关键帧？

关键帧目前是一种非常常用的方法，可以减少待优化的帧数，并且可以代表其附近的帧。可以理解为一个学校里有100个班级，每个班的班长就是一个关键帧，他可以代表他班里的人，那么如何选取关键帧呢？

选取的指标主要有：

(1) 距离上一关键帧的帧数是否足够多（时间）。比如我每隔固定帧数选择一个关键帧，这样编程简单但效果不好。比如运动很慢的时候，就会选择大量相似的关键帧，冗余，运动快的时候又丢失了很多重要的帧。

(2) 距离最近关键帧的距离是否足够远（空间）/运动

比如相邻帧我根据pose计算运动的相对大小，可以是位移也可以是旋转或者两个都考虑，运动足够大（超过一定阈值）就新建一个关键帧，这种方法比第一种好。但问题是如果对着同一个物体来回扫就会出现大量相似关键帧。

(3) 跟踪质量（主要根据跟踪过程中搜索到的点数和搜索的点数比例）/共视特征点

这种方法就是记录当前视角下的特征点数，或者视角，当相机离开当前场景时才会新建关键帧，避免了第2种方法的问题。缺点是比较复杂

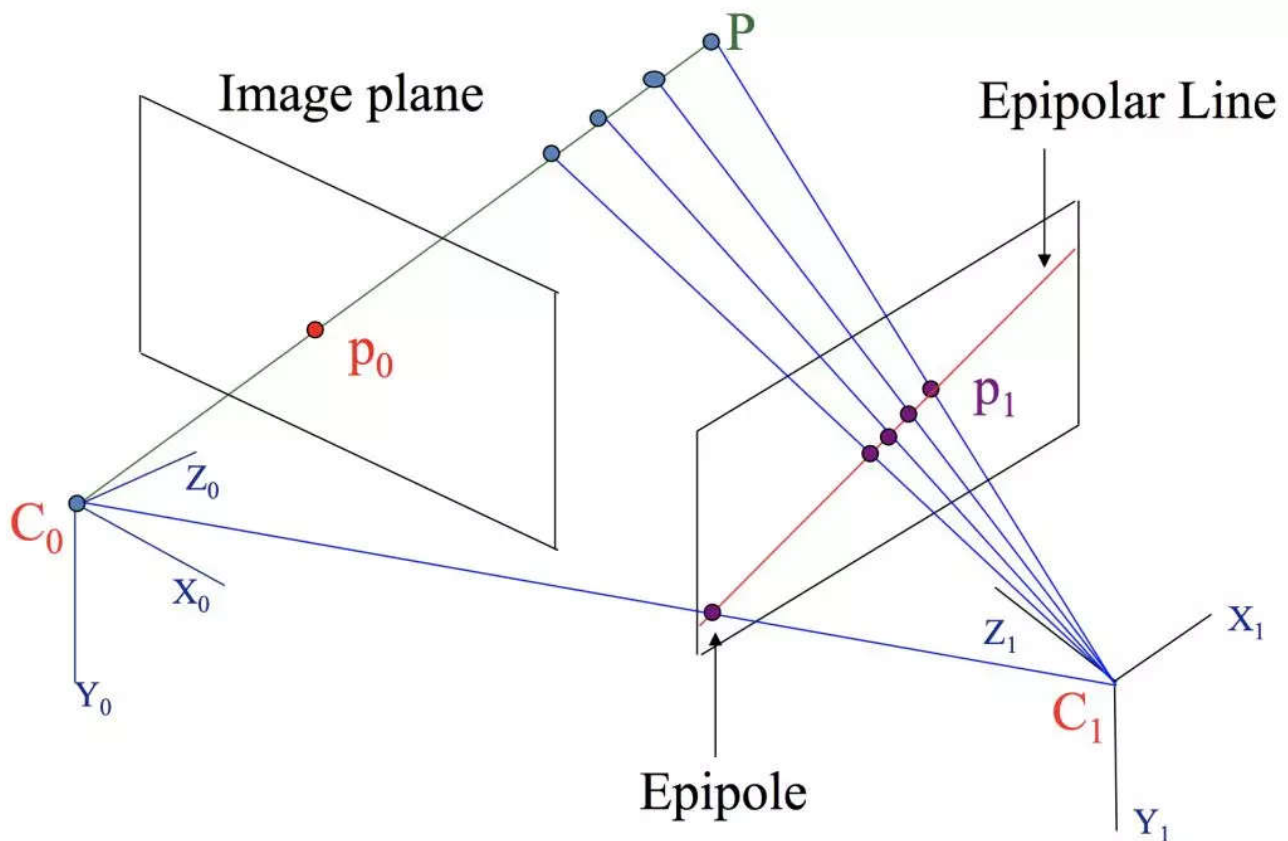
打个比方，关键帧相当于slam的骨架，是在局部一系列普通帧中选出一帧作为局部帧的代表，记录局部信息。举例来说，摄像头放在原处不动，普通帧还是要记录的，但关键帧因为总看到原场景，所以不会增加。

三角化需要一定程度的共视区域，所以普通帧每2帧之间会存在大量的信息冗余，如果所有帧全部参与计算，不仅浪费了算力，对内存也是极大的考验，这一点在前端vo递归处理方式中表现不明显，但在后端优化里是一个大问题，所以关键帧主要作用是面向后端优化的算力与精度的折中。此外，关键帧选择时还会对图片质量、特征点质量等进行考察，一定程度上也发挥了滤波的作用，防止无用的或错误的信息进入优化过程而破坏定位建图的准确性。

选择关键帧主要从关键帧自身和关键帧与其他关键帧的关系2方面来考虑。一方面，关键帧自身质量要好，例如不能是非常模糊的图像、特征点数量要充足、特征点分布要尽量均匀等等；另一方面，关键帧与其他关键帧之间的关系，需要和局部地图中的其他关键帧有少量的共视关系，但大部分特征点是新特征点，以达到既存在约束，又尽量少的信息冗余的效果，例如局部地图点投影到此帧的点数低于一个阈值或前一个关键帧的特征点在此帧里已经有90%观测不到等等。

在关键帧的运用上，我认为orb slam做的非常好，尤其是在回环检测中使用了以关键帧为代表的帧“簇”的概念，回环筛选中有一步将关键帧前后10帧为一组，计算组内总分，以最高分的组的0.75为阈值，滤除一些组，再在剩下的组内各自找最高分的一帧作为备选帧，这个方法非常好地诠释了“关键帧代表局部”的这个理念。

4、按照你的理解讲解一下什么是极线约束？这个约束能带来什么好处？



极线约束也叫对极约束。这个约束的意思就是说，假设相机在不同位置拍摄了两幅图像，如果一个空间点 P 在两幅图上分别有两个成像点，已知左图成像点为 p_1 ，那么右图成像点 p_2 一定在相对于 p_1 的极线上。

(以上过程面试的时候最好画图解释一下，见附图，面试官会感觉你很专业)

极线约束的好处：从上面的描述我们可以看到，我们在做特征点匹配时，左图成像点 p_1 的待匹配点 p_2 一定在相对于 p_1 的极线上，那么我们在做搜索时就可以在极线附近（考虑实际可能会有一点误差）进行搜索，相对暴力匹配极大减少待匹配的点的数量。

极线约束可以简洁的给出匹配点的空间位置关系，使得相机位姿估计问题变的简单。

5、SLAM后端一般有两种方法：滤波方法和非线性优化方法，这两种方法有什么优缺点？

6、单目视觉slam中尺寸漂移是怎么产生的？有什么解决办法

7、直接法估计相机位姿时，并不需要提取特征点，而是通过优化匹配点的像素值误差（也称光度误差）估计位姿，但也会面临快速运动，光照变化等的挑战，如果让你改善该问题，你会采用哪些方法来提高跟踪质量(精度，速度，鲁棒性等)？

8、什么是PnP算法？请用你的语言描述一下原理，它一般用在什么场景，解决什么问题？

...

此外，我们平时在SLAM的学习工作中也会遇到一些问题，我总结了一些常见的问题，也一并列在这里，并给出了答案（见知识星球）

9、我们知道相机的内参有 f_x, f_y, c_x, c_y , 畸变参数(只考虑 k_1, k_2)，相对世界坐标原点外参 T 。如果我们现在对相机拍摄的图片进行2倍的下采样，那么这些参数会如何变化？

10、我们知道双目相机两个相机光心的间距我们称之为 baseline。如果双目相机baseline比较大，我们称之为wide baseline.现在某代码中使用一个单目相机进行SLAM过程，在特征匹配时资料中提到了wide baseline，请问这个wide baseline怎么理解？

11、RGB-D相机我们知道可以直接输出 RGB + depth两张图比如我们常见的Kinect 是结构光原理，包括一个彩色相机，一个红外发射器，一个红外接收器。另外，Intel的Realsense系列RGB-D相机也非常常用，比如下面Realsense D415，官网说是Active IR stereo，也就是双目深度相机，这个双目和我们平时说的双目有何不同？为什么有如下四个孔？

12、我们在阅读文献或者代码中误差相关时，经常可以看到一个概念，叫逆深度（inverse depth）。也就是深度的倒数，那么同学们有没有想过，为什么使用逆深度误差而不是深度误差？

13、我们在看SLAM相关论文的时候，会遇到一个词“kidnap”，直译过来就是“绑架”，不了解的同学可能感觉怪怪的。你知道这个“绑架”是什么意思吗？可以用哪些方法解决这样的问题？

14、我们知道（不知道的话，去查一下十四讲）用g2o和ceres库都能用来进行BA优化，这两者在使用过程中有什么不同？

15、SLAM中回环检测（闭环检测）的目的是什么？简述一下SLAM中可以使用的回环检测方法？

16、SLAM中为什么要引入李群李代数？

17、为什么SLAM中常用LM算法而不是高斯牛顿求解优化问题？

18、讨论一下SLAM应用场景及落地的问题。大家觉得SLAM技术最适合的应用场景是什么？在哪个场景能够最快技术落地呢？

19、大家都是SLAM方向的研究者，不管是学生还是已经工作，以后都面临找（换）工作的问题，那么你知道哪些做SLAM技术的公司？

20、什么是ICP 算法？简述一下算法原理，SLAM中一般什么情况下会使用该算法？

.....

限于篇幅，只列举部分星球里已经发布的问题，以上所有问题《从零开始学习SLAM》知识星球里都有参考解答，并且会持续发布。

最近从零开始学习SLAM学习知识星球限时3天优惠，每天只需0.3元，加入3天内不满意全额退款。

关于星球的详细介绍请扫描下面二维码查看：