**2017-2018年度学习调研报告**

报告人：李昱霄

**一、概述**

SLAM，全称叫做Simultaneous Localization and Mapping，中文叫做同时定位与建图。而视觉slam是指仅仅利用图片这一外部信息来确定环境中机器人、车辆或者移动相机的位置，同时能够建立已探索区域的一种表示法（地图）。SLAM是移动机器人实现自主化的一项基本任务。

移动机器人的自主导航问题被分成三大块：定位，建图，路径规划。定位就是精确确定机器人在环境中的当前姿态。建图就是将环境的多个不完整观测集成到单个一致的模型中，路径规划就是在地图中确定一条通过环境的导航最优路径。

视觉slam可供选择的拓展项，1.与imu、激光雷达、超宽带等传统传感器所采集的信息进行信息融合研究及现在应用较为广泛的VIO系统。2. 增强现实的应用（虚拟的物体被放置在现实世界场景中）

**二、为什么以及怎么样研究slam**

机器人在陌生环境中主要需要解决三个问题：

　1.          自己在哪里？

　2.          这是什么地方？

　3.          怎么离开这个地方？

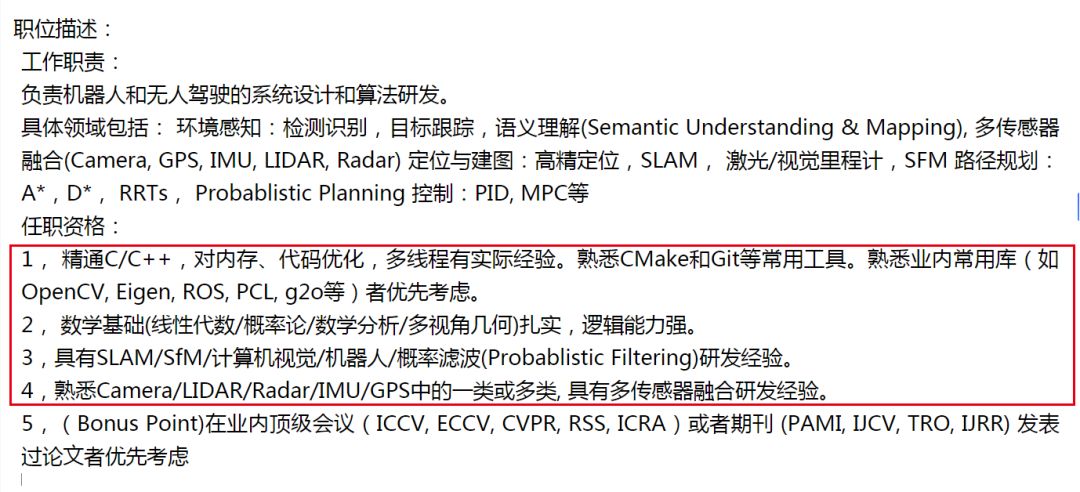
在SLAM理论中，第一个问题称为定位 (Localization)，第二个称为建图 (Mapping)，第三个则是随后的路径规划。

最近几年我们见证了机器人领域的快速发展，大疆的无人机已经很好地实现智能化飞行，谷歌的自动驾驶出行服务已经在局部范围进行试运营，波士顿动力的Atlas机器人一个漂亮的后空翻甚至让多数业内人士为止惊叹，AR/VR技术也在定义新的人机交互方式。这些都离不开稳定可靠的定位导航技术。

目前这一方向研究难度较大有以下几个方面：

1. 入门资料很少。虽然国内也有不少人在做，但这方面现在没有太好的入门教程。
2. 自己动手编程需要学习大量的先决知识。首先要会C和C++，网上很多代码还用了11标准的C++。第二要会用Linux。第三要会cmake，vim/emacs及一些编程工具。第四要会用openCV, PCL, Eigen等第三方库。只有学会了这些东西之后，才能真正上手编一个SLAM系统。如果你要跑实际机器人，还要会ROS。

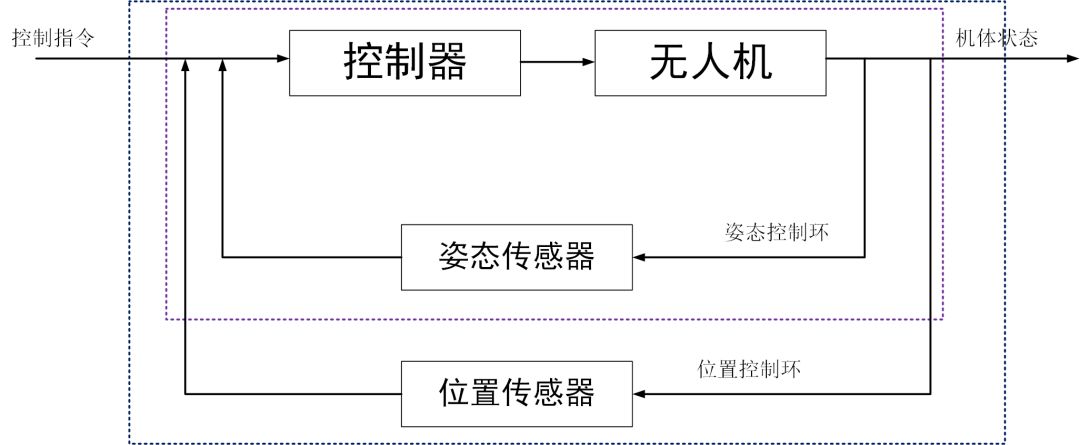
Slam技术目前可以广泛的应用于无人车，无人机，城市地图重建等热点AI领域，下面是一无人驾驶公司的招聘要求：



**三、研究方向的选择**

SLAM的输出是运动体状态+环境地图，所以研究SLAM会有两个不同的侧重点，侧重前者可以用于机器人定位导航（及制定一个机器人里程计并且对结果进行一个反馈纠偏，我们称之为前端，对应三维地图重建被称为后端），侧重后者则用于三维重建。（目前根据实验室条件建议将工作重心放在前段里程计的搭建和优化上，后端三维重建工作可作为日后的可扩展选项）。

以无人机的控制为例，总体控制流程图：



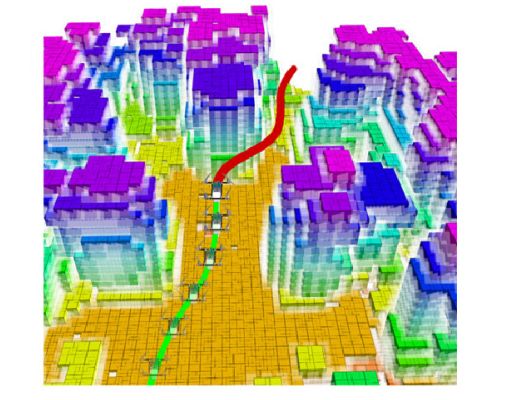
对无人机的运动控制是通过内环姿态控制和外环位置控制实现的，内环控制实现了机体在空间中的基本运动，外环位置控制则是实现路径规划、航点飞行的基础。为了对无人机姿态进行控制，需要有姿态量的反馈，这个是通过IMU/AHRS传感器进行姿态解算实现的；同样，位置控制需要位置量的反馈，可以实现位置测量和估计的传感器有GPS、UWB、LiDAR、Camera。GPS和UWB是利用外部信号实现的定位，LIDAR和Camera则是利用SLAM算法实现位置估计。同时，控制理论中，速度量可以加快机体响应速度，也是一个非常重要的量，复杂的姿态和位置控制环路里一般会加入速度（角速度、线速度）控制。所以通常我们也会考虑速度的测量估计。

我们做状态估计的目的就是得出机器人三组状态量（位置、速度、姿态），为机器人运动控制做状态反馈。尽管不同的传感器会有不同的特征，进行数据处理、运动估计的方法也不尽相同，但是目的是一致的，都是为了得到机器人的三组状态。

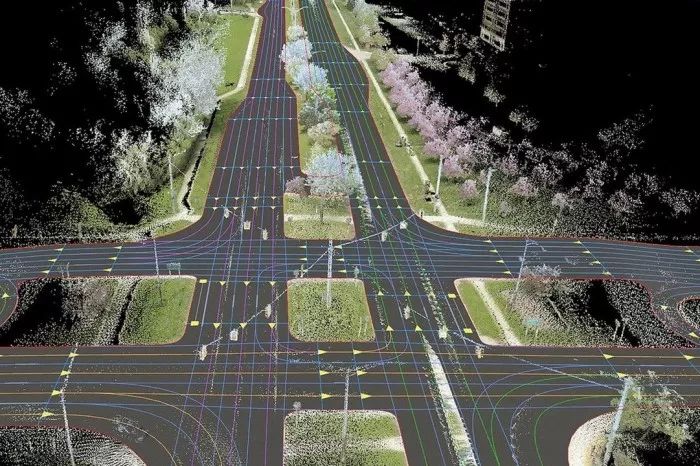
从机器人定位导航的层面上认识SLAM的地位，SLAM只是机器人定位导航的一种方法，不同传感器的不同特性使其适用于不同的场合，为了提高定位导航系统的可靠性稳定性，在实际中往往涉及多传感器融合。而SLAM方法进行定位导航的优势在于不需要外部先验信息，使用范围更广泛。

SLAM在机器人领域的应用

下图是港科沈绍劼老师团队利用VIO实现的无人机导航和单目稠密重建的工作



在自动驾驶里，一方面则是使用多传感器融合的方式提高定位系统的可靠性稳定性，这里面激光/视觉SLAM是其中一个重要的定位来源；另一方面在构建高精地图时，也会用到SLAM技术，高精地图被认为是实现自动驾驶的重要技术方案，使用地图采集车先构建出包含非常详细的环境细节的地图，在正常运行的自动驾驶车辆上则利用SLAM重定位技术从高精地图里获取定位信息



**四、设备**

Pc一台（已有），单目/双目摄像头一组（约1000元），数据集（互联网获得），imu（约100元）

**五、传统技术路线**

1.通过摄像头获得场景的图片信息。

2.对获得的图片进行特征点匹配（方法有SIFT，SURF,ORB等，可以用到很多**机器学习**的方法）。

3.通过对极几何方法获得一个初始位移信息（其中涉及到了相机的标定，本质矩阵，基本矩阵的求解等几何知识对**矩阵**的操作要熟练），对初始信息进行再优化（目前主流方法是Bundle Adjustment这里需要对**最优化**有所理解）。

4.通过解出的路径信息重建出无人机的路径。

——————————以下部分为后端———————————

5.对前端信息进行去噪，优化在这里会用到卡尔曼滤波知识。

6.对之前的所有信息进行回环检测（如果是在一个开放的环境中此步可省略）

7.三维地图重建

**六、补充文献**

目前国内slam技术较为前沿的公司及团队：香港科技大学沈邵劼团队（同时也是大疆无人机首席机器人学家），中科院自动化所模式识别国重胡占义团队，清华自动化高翔博士，百度无人车团队等。

**由于水平有限更进一步的技术细节可以看高翔博士的博客：**

<https://www.cnblogs.com/gaoxiang12/p/3695962.html>

**目前主流的开源框架：**

VIO目前实现比较好的有vinsmono,okvis,MSCKF。前两个是基于非线性优化的方案而且框架比较相似，后者是基于滤波优化的方案，也是Google Tango上使用的方法，MSCKF目前并没有开源，不过宾夕法尼亚的Kumar实验室18年有一个相似的工作，目前已经开源。此外还有ROVIO。值得注意的是，虽然在纯视觉SLAM中，学界已经公认基于非线性优化方法的SLAM方法效果要好于滤波的方法，但在VIO中，非线性优化和滤波方法目前还没有很明显的优劣之分。我的理解是结合相机和IMU两种传感器的信息，提供了对当前状态更多的观测，使得算法对历史观测之间约束信息的依赖降低，这也是为什么okvis、vinsmono采用滑动窗口法（global bundle adjustment和filter 的折中）也能取得很好效果的原因。

**开源数据集选择：**

飞行数据集，目前使用比较多的是EUROC数据集，给出了几种场景下飞行平台采集到的双目图像，IMU数据以及VICON提供的运动基准值。

无人车的数据集，最流行的当数KITTI。

定位误差有ATE和RPE两种评测指标，TUM给了评测估计值和基准值误差Python脚本，而且对两种误差的解释，TUM的评测脚本是假设估计轨迹和真实轨迹都是实际物理尺度的。ORB作者改写了一个自动缩放尺度的脚本，可以用来评测单目SLAM运动估计效果。

Euroc dataset: https://projects.asl.ethz.ch/datasets/doku.php?id=kmavvisualinertialdatasets

KITTI dateset: http://www.cvlibs.net/datasets/kitti/

https://vision.in.tum.de/data/datasets/rgbd-dataset/tools#evaluation

**工具介绍：**

C++是做SLAM的主流编程语言，需要重点掌握，也是很多公司非常重视的技能，C++学习可以参考北大的C++公开，配合着网站给出的作业一起搞。Pyhton的基本语法比较简单，网上介绍的资料比较多，Python中的Matplotlib可以很方便的替代Matlab 进行数据分析的事情。Linux 下的C++、Python开发环境推荐Clion和PyCahrm，两款软件是同一个公司出的，学生可以利用教育邮箱在官网上申请免费使用。

Cmake是目前主流的C++工程构建方法，可以参考《Cmake Practice》。

同时，需要对Linux系统有一些宏观了解，并掌握基本的命令，学习资源推荐一门公开课，讲解的非常清晰。

ROS的资料推荐官网的wiki，网上也有很多介绍的博客。

**七、个人计划**

在硕士期间完成至少一个完整的slam前段框架的搭建，对现有的著名开源框架如ORBslam、vinsmono完成搭建和运行，对其中的机器学习可优化模块进行进一步的优化完成论文写作，同时留下一个较好的可持续展开的baseline。