

# Final Project 说明

高翔

2021 年 9 月 9 日

## 1 Final Project 说明

Final Project 可以由至多 3 人组队完成，最终提交完成的报告。组队过程可以由学员自主决定，也可以一人组队。由多人组队的，在提交报告时需说明分工情况。多人组队的情况下，对完成度要求会适当提高。

Final Project 有四周完成时间，提供若干道选题。队员可从中选择自己感兴趣的一题。Final Project 的题目是开放的，没有标准答案，最后按照报告的完成度和丰富程度进行打分。

### 选题 1: Kitti 的双目视觉里程计 （10 分）

Kitti 的 Odometry 提供了左右双目的视觉图像（以及激光，但本作业不使用激光数据），并提供了标定信息。它一共含有若干个 Sequence，其中一部分 Sequence 的真实轨迹是开放的，另一部分则是隐藏的，作为测试使用。在 Kitti 官网可以上传你对测试部分的轨迹估计，系统会计算与真实轨迹的差异，并给出评分。

现在我们已经介绍了所有关于视觉 SLAM 方面的内容，你可以基于已有算法，实现一个双目的视觉里程计了。Kitti 官网 odometry 分类下提供了双目相机的所有信息。请你根据已学知识，实现双目视觉里程计，然后运行任何一个 sequence，并与标准轨迹比较。



以下是实现过程中的一些提示。这是通常 SLAM 中的做法，你可以作为参考，但并不一定必须使用这样的结构。

1. 你可以先实现一个 Frame-by-Frame 的里程计，即仅估计当前图像与上一个图像之间的运动，然后把它们组成完整的相机轨迹。这样的工作方式应该是有效的，但误差会很快累积，导致轨迹发散。
2. 接下来，从轨迹中提取关键帧，对关键帧进行 bundle adjustment。由于 BA 过程重投影误差必定是下降的，可以有效的抑制误差累积。同时，你需要一个机制来

管理所有地图点和关键帧。

3. 最后，请测试你的方案在 kitti 上的表现，例如轨迹精度，运行时间等等。

## 选题 2：VIO 的 Bundle adjustment 推导 （10 分）

在 VIO 里，除了传统的 bundle adjustment 以外，我们还有 IMU 提供的测量量。请根据文献 [1]，推导并实现带有 IMU 测量的 Bundle Adjustment。以下是一些提示：

1. 首先，请推导 IMU 测量方程，即角速度和加速度分别测到的是什么量。这其中需要用到相机的运动学，请一并推导。
2. 第二，请说明按照图优化的思路，如何建立 VIO 的图模型？其中包括哪些节点和边？
3. 第三，说明各节点储存的量是什么，各边的误差如何计算，雅可比如何计算？
4. 最后，利用 Ceres、g2o 或 gtsam 等库实现这些运算。

你也可以参考其他文献来回答这些问题。请注意 VIO 中的数学符号比较多，请使用一致的符号。

## 选题 3：开源 SLAM 方案的评价与比较 （10 分）

在 SLAM 研究中，我们通常需要在各数据集上测试各个方案的性能情况。请对不少于三种开源 SLAM 方案（你可以在 github 上搜索），在不少于三个公开数据集上进行测试，并给出性能和精度和鲁棒性指标。其中，你需要说明：

- 评价了哪几个方案？使用了哪些数据集？
- 如何将各数据集的数据放到 slam 方案中？你书写了哪些运行实例？
- 性能、精度和鲁棒性指标是如何定义，如何计算的？例如整条轨迹的 ATE，RPE 是如何计算的，运行时间如何测量等等。
- 探讨参数对 SLAM 方案的影响，例如特征点的数量或运行过程中的阈值等等。

## 选题 4：手写 Bundle Adjustment （10 分）

港科大沈劭劼老师要求他的一年级研究生仅使用 Eigen 写出 Bundle Adjustment，所以我觉得这应该不难。请在只使用基础代数运算库的情况下实现完整的 Bundle Adjustment，并对比 g2o 或 ceres 之一，给出比较结果。

- 你不必实现一个通用的优化库。只需假设你仅有某几种节点或边即可。当然如果你愿意，写一个通用库然后测试曲线拟合之类的简单问题，也是很好的。
- 挑选一个 BA 数据集测试你的结果，包括精度、性能、内存占用等等。
- 在基本的 L-M 迭代基础上尝试一些其他的迭代方式，例如 Dog-Leg 迭代，或者 PCG 求解线性方程等等。完成度高的将有更高的评价。

### 选题 5：基于检测框的重建 （10 分）

物体检测是计算机视觉中一个常见的任务。通常，检测网络会输出所有图片中存在的物体，每个物体以一个 2D 外包框表示，每个框还带有物体标签、置信度等信息。这是一个传统的 2D 视觉问题，但是，通过 SLAM 的方式，如果我们在多个视角看到同一个物体的外包框，则可以推断该物体的 3D 位置。

请参照文献 [2] 和 [3]（或自行搜索对应文献），完成下列任务：

1. 推导基于对偶二次曲面的重建算法，说明其工作原理；
2. 使用 g2o 或 ceres 实现该算法，完成一个演示用例。检测框可使用任意检测网络完成。

## Bibliography

- [1] C. Forster, L. Carlone, F. Dellaert, and D. Scaramuzza, “On-manifold preintegration for real-time visual–inertial odometry,” *IEEE Transactions on Robotics*, vol. 33, no. 1, pp. 1–21, 2017.
- [2] N. Sünderhauf and M. Milford, “Dual quadrics from object detection boundingboxes as landmark representations in slam,” 2017.
- [3] L. Nicholson, M. Milford, and N. Sünderhauf, “Quadricslam: Dual quadrics from object detections as landmarks in object-oriented slam,” 2018.