

数字图像处理大作业自选题目：无人机航拍场景的语义三维重建

小组成员

姓名	学号
陈煜	1701213988
王旭普	1701214012
王尧	1701214013

0. 任务背景及前瞻

较为成熟的室外场景数据集主要包括两种：街景与卫星

街景数据集一般由固定在车辆上的摄像机拍摄形成，卫星数据集顾名思义是卫星拍摄的图像，二者对于室外场景，尤其是建筑的刻画都不够完全，因此都不能直接拿来对建筑进行三维重建。

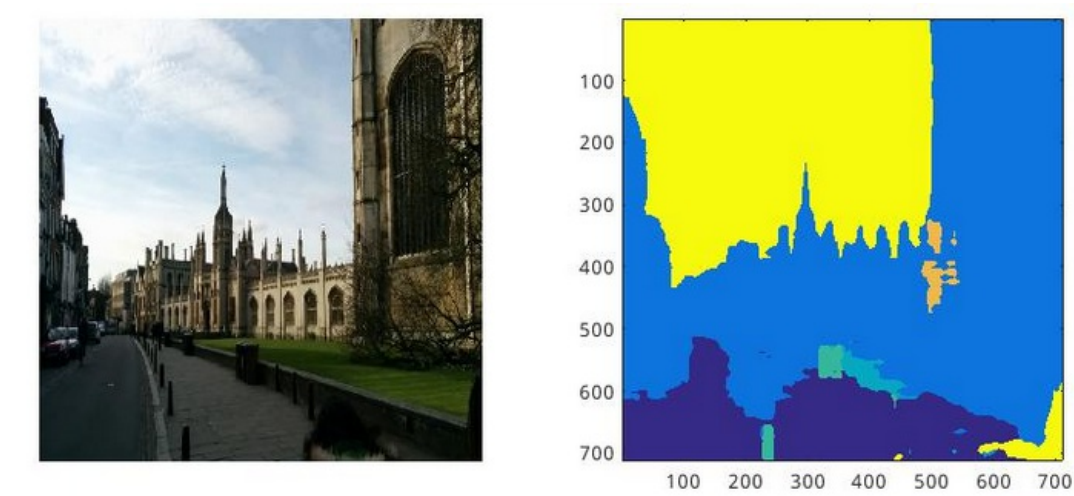


fig.1(a) 伦敦国王大学街拍数据集

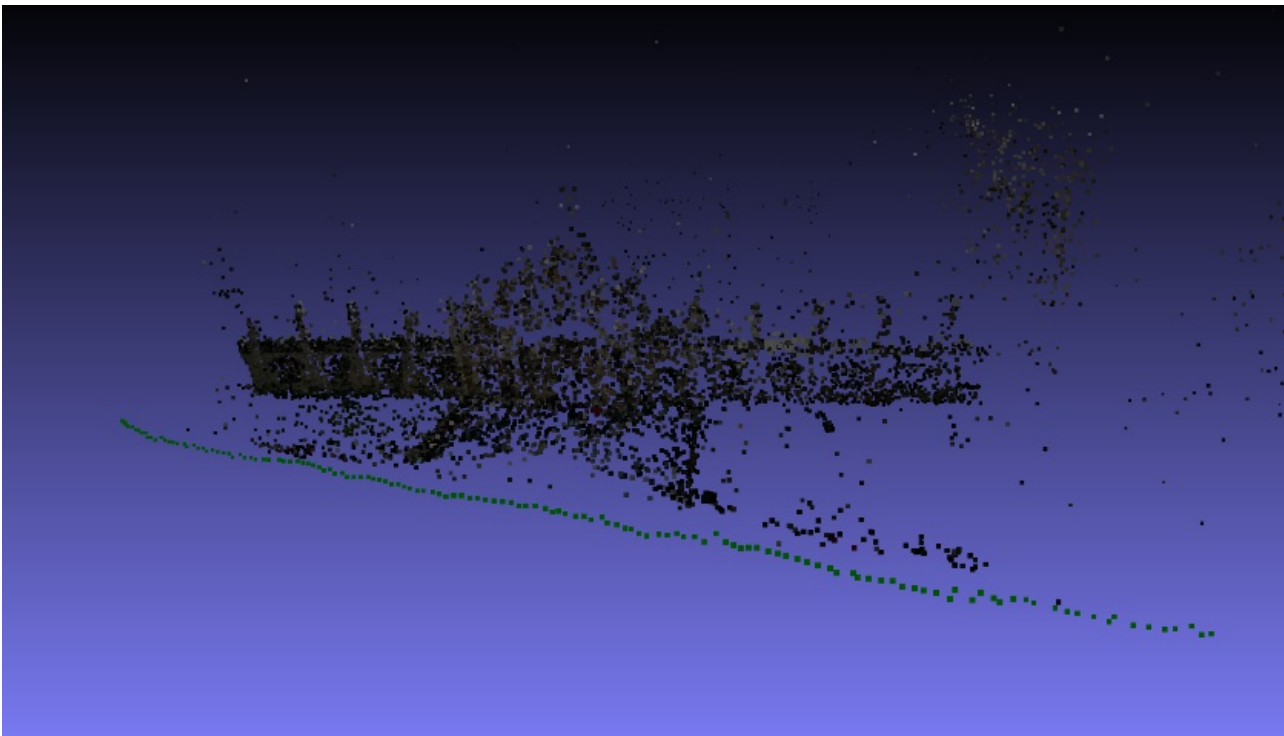


fig.1(b) 伦敦国王大学街拍场景重建出来的三维点云，建筑的顶面并没有被建出来



fig.2 INRIA卫星数据集，建筑的立面不在图像中

对于大规模场景三维重建任务，无人机航拍图像必然是最合适的。搭载在无人机上的相机可以拍摄建筑物的顶面与立面，且可以对目标场景快速获取大量数据。我们图形学与交互实验室先前采集的北大航拍影像有近千张，其分辨率较高(4000*3000)，且拍摄间隔较为紧凑。目前已经能分块跑出小规模场景的demo。

走通三维语义模型的构建，将有如下两个主要意义：

1. 三维语义模型重投影，可以得到带语义的二维仿真图像，当我们的模型构建足够准确，将节约大量的人工标注成本
2. 对于三维重建任务，语义信息也将优化匹配速度与结果，分类匹配是很自然的想法，不同类别的特征点必然是错匹配(如果语义标签足够准确)

但根据我们的调研，目前没有带有精细标注的无人机航拍数据集(拍摄高度40-150米)，因此决定在目前已有的航拍图像基础上进行扩充，建立第一个对外开放带有精细标注的无人机航拍数据集。我们分析得出影响无人机航拍图像语义分割性能的因素如下：

1. 场景内容
2. 拍摄高度
3. 光照(阴影)

为了构建泛化能力更强的网络，我们决定从这三方面扩充数据集，构建较为通用的无人机航拍数据集，并训练得到较好的分割模型。同时我们也将重构第三方库的代码，自行搭建SFM全流程语义模型。如有余力，我们也会对三维模型通过重投影的方式进行评估。

1. SFM系统构建

- UI: Qt/meshlab
- 后端: C++
- 主要依赖库: Eigen, Ceres, Open3D(OpenMVS)

任务目标:

- 实现稀疏点云重建算法Structure from Motion
- 结合图像语义信息生成语义点云
- 前端支持模型的放大、缩小、转动等功能
- 接入稠密点云重建到纹理重建算法(OpenMVS)
- 构建不少于2个模型, 每个场景使用图片数量不少于200张
- 基于当前的特征点匹配算法(SIFT、orb等), 设计算法提高特征点匹配速度以及匹配正确率

2. 神经网络模型训练

- 后端: matlab、python
- 主要依赖库: tensorflow

任务目标:

- 采集不少于1000张4000*3000的无人机航拍场景数据, 并完成不少于6类的语义标注(建筑、植被、地面、水面、汽车、自行车、行人)
- 训练一个在测试集上miou不低于70%的网络, 网络得到的预测结果用于优化SFM过程

3. 语义模型生成与仿真(进阶)

- UI: Qt/meshlab
- 后端: C++
- 主要依赖库: opencv

任务目标:

- 实现ray marching过程, 获得三维模型在原摄像机视点的图像
- 计算重投影误差, 评估模型的质量