



物流场景稠密点云 融合与视觉定位

Department of Automation, Tsinghua University

>> 答辩学生: 自74 陈奕凡

>> 指导老师: 陈宝华





目录

CONTENTS

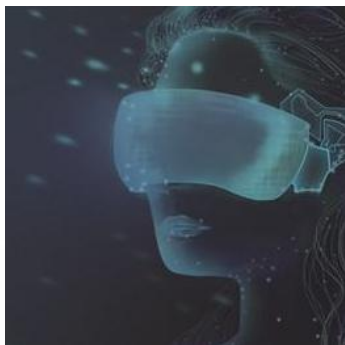
- 1 背景与意义
- 2 基于稠密点云融合的离线建图
- 3 基于2D-3D匹配的视觉定位
- 4 后期研究计划



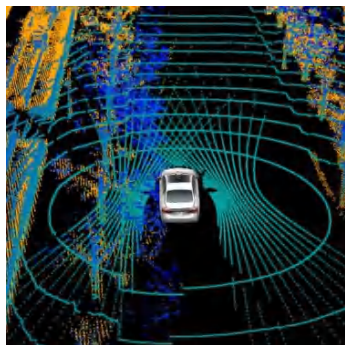
背景与意义

高精度三维点云地图构建与定位

● 应用场景



AR/VR

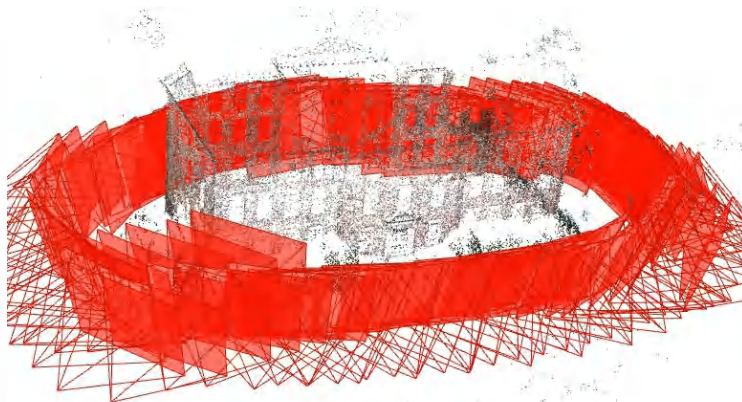


自动驾驶



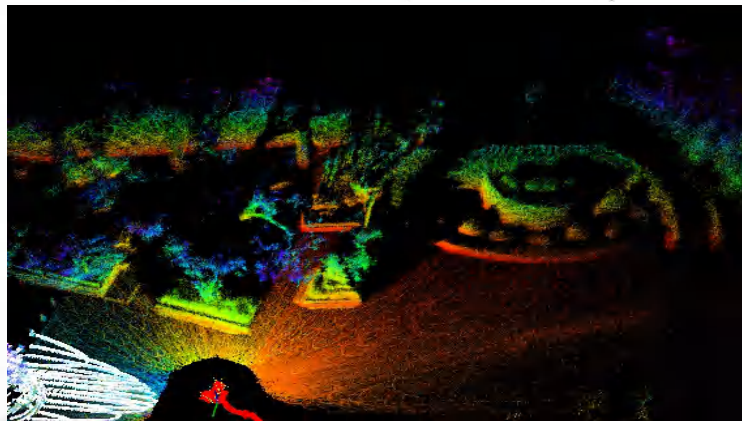
工业机器人

● 目前常用离线构图方法



Structure from motion (sfm)

- 基于单目相机图像
- 优点：数据采集方便
- 缺点：重建结果较为稀疏、没有绝对尺度



Lidar Odometry and Mapping (LOAM)

- 基于激光雷达
- 优点：数据准确、有绝对尺度
- 缺点：没有颜色纹理信息、运动激光雷达线数少、采样稀疏

[1] Schonberger J L , Frahm J M . Structure-from-Motion Revisited[C]// IEEE Conference on Computer Vision & Pattern Recognition. IEEE, 2016:4104-4113.

[2] Lin J , Zhang F . Loam_livox: A fast, robust, high-precision LiDAR odometry and mapping package for LiDARs of small FoV[J]. 2019.



背景与意义

□ 视觉与激光融合的建图与定位算法

● 数据采集设备



高精度三维激光扫描仪 Leica BLK360

- RGB点云, 信息丰富
- 静态扫描, 精度较高

采集代价高但精度大,
适合**离线建图**

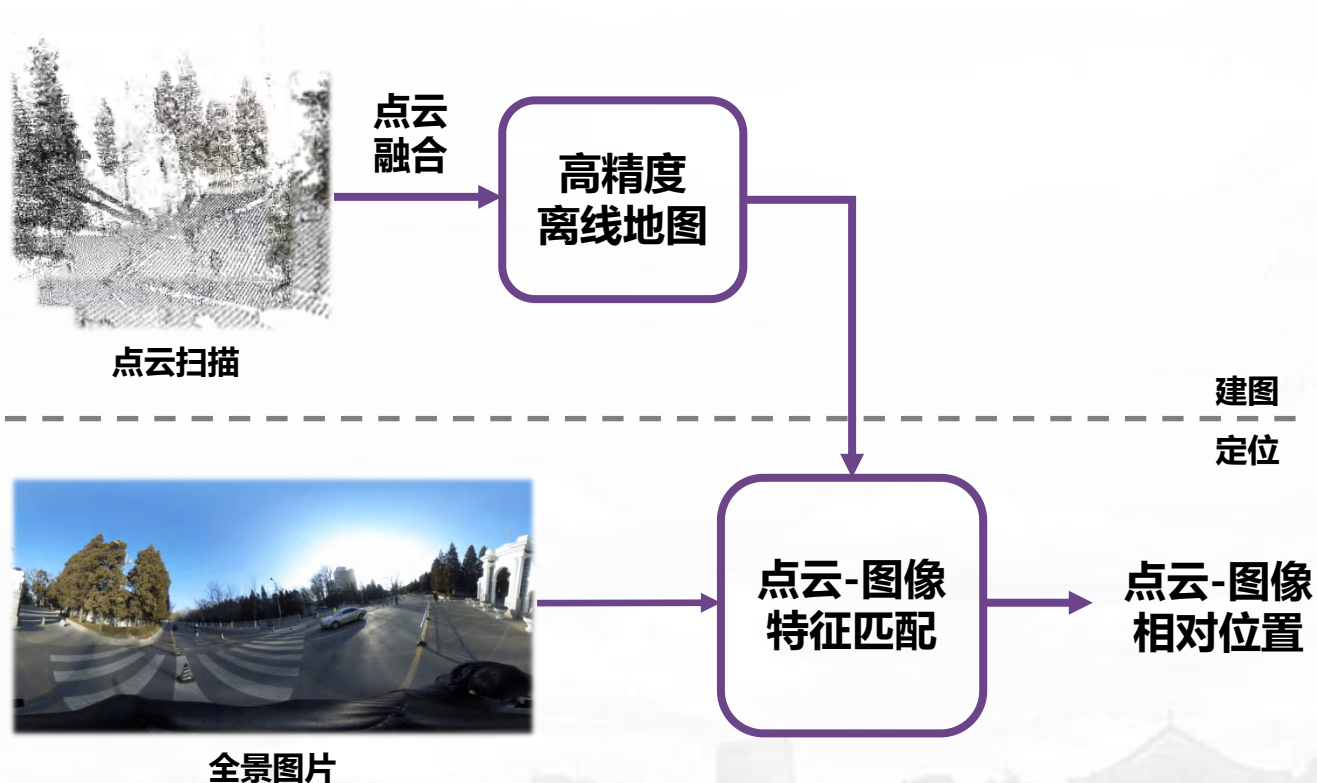


全景THETA相机

- 拍摄便捷, 即时成像
- 全景图像, 更多细节

2D数据但采集代价低,
适合**在线定位**

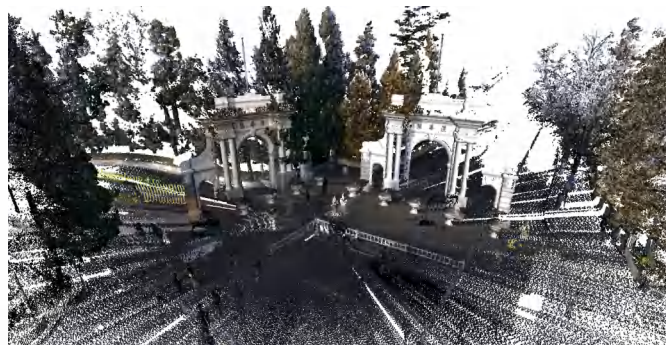
● 整体研究思路





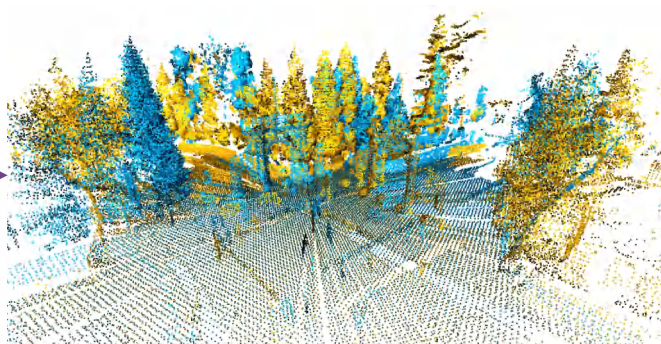
基于稠密点云融合的离线建图

点云融合步骤



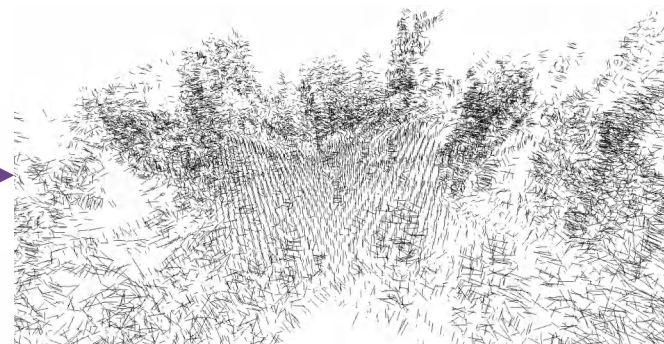
初始点云(2测站)

点云体素
降采样

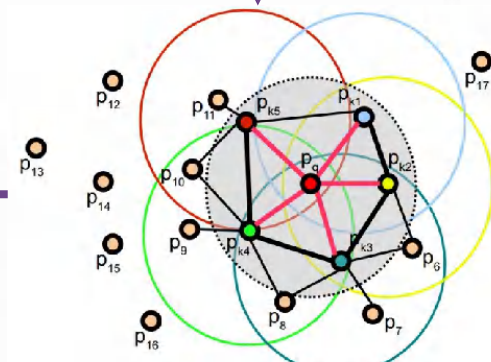


降采样结果 (去除颜色)

法线估计

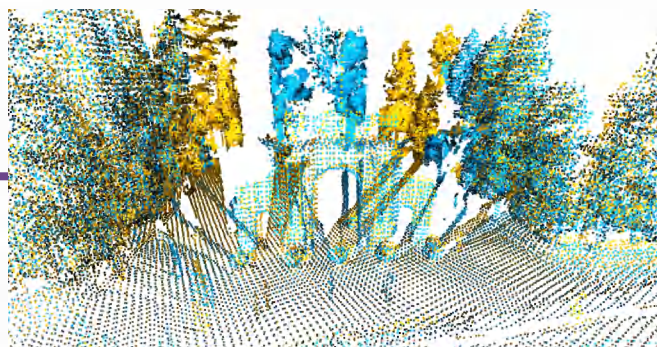


FPFH特征
估计



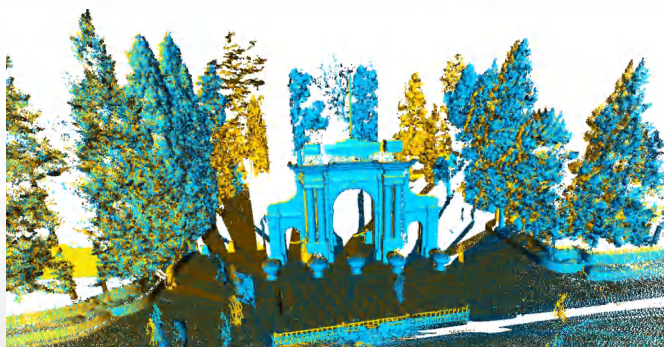
FPFH特征示意图

RANSAC
全局粗配准



粗配准结果

局部优化
Point-to-
plane ICP

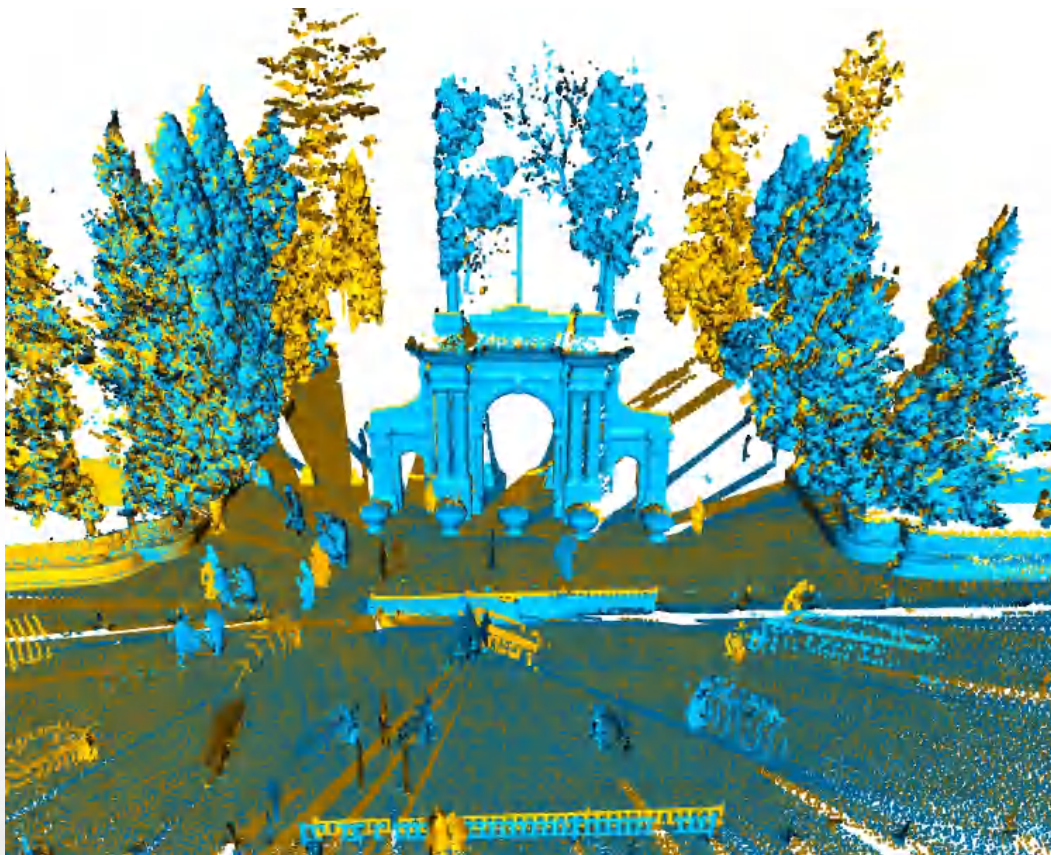


优化效果



基于稠密点云融合的离线建图

点云融合结果



评价指标

- **fitness** 融合分数
- **rmse** 匹配点的均方根误差

$$rmse = \sqrt{\frac{\sum_{i=1}^n ||a_i - b_i||}{n}}$$

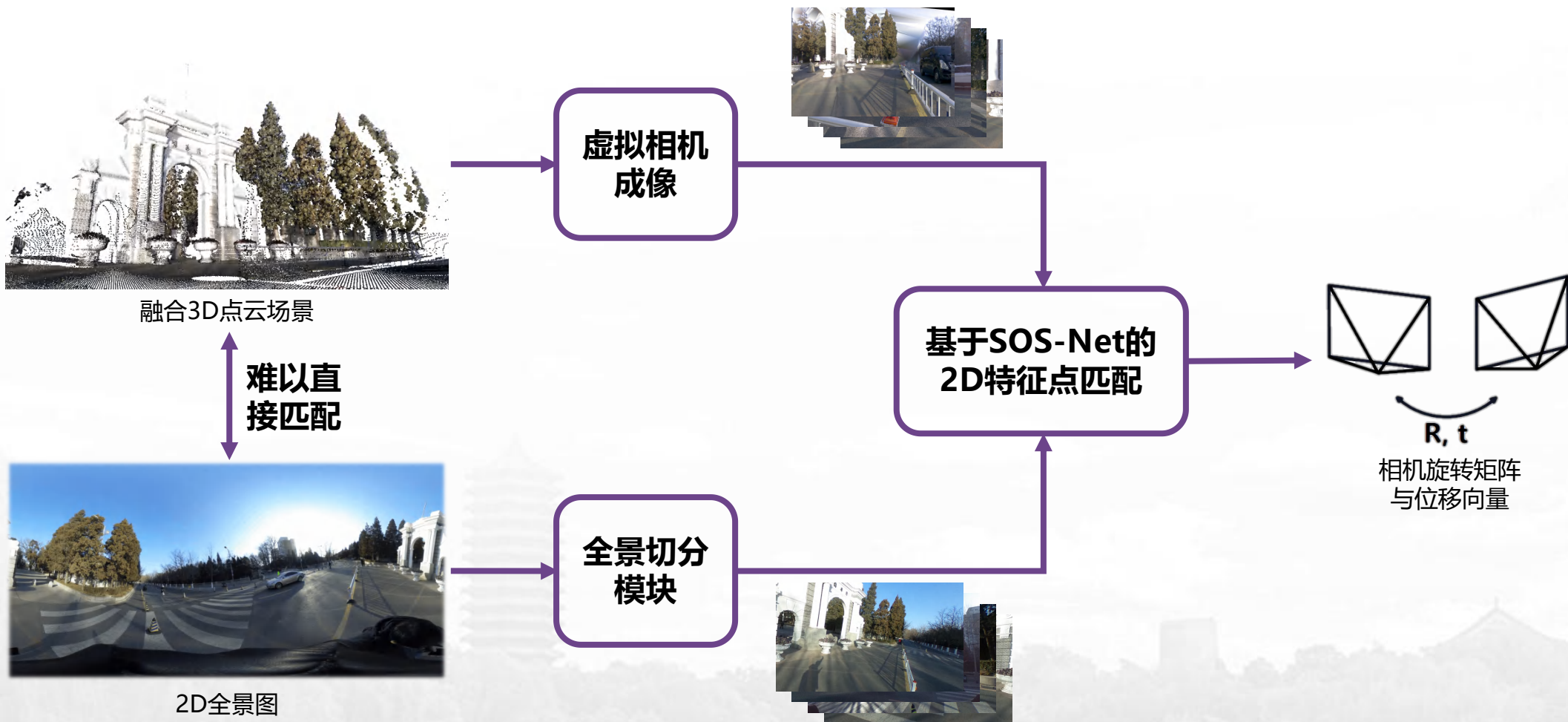
- **corr_size** 匹配的对数

	fitness	rmse	corr_size
RANSAC	0.0506	1.8394	808
Local Refinement	0.5752	0.0600	561193



基于2D-3D匹配的视觉定位

视觉定位算法流程





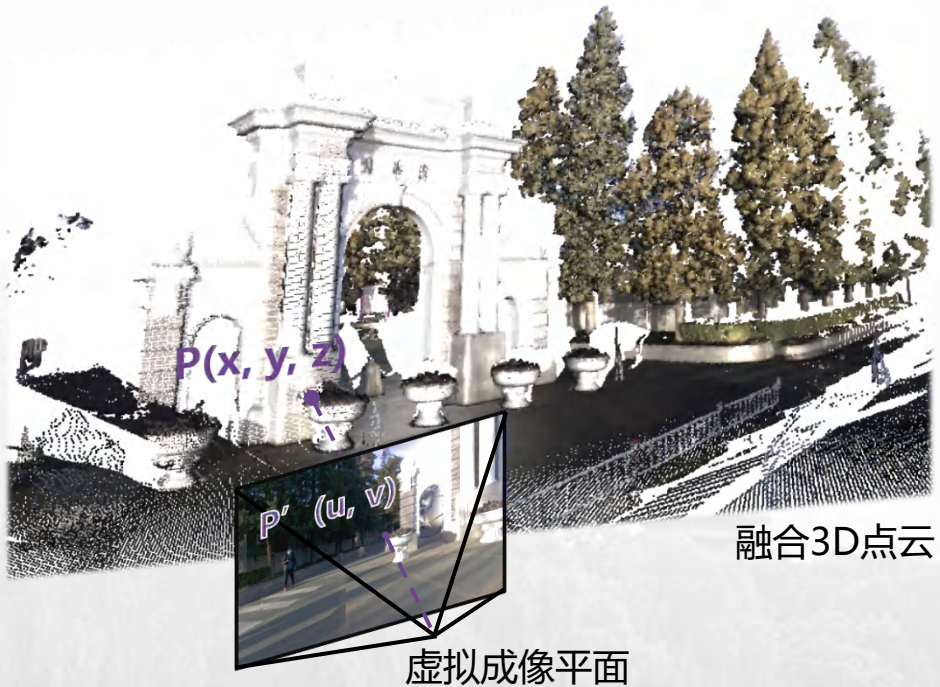
基于2D-3D匹配的视觉定位

□ 点云内虚拟相机成像

- 成像原理

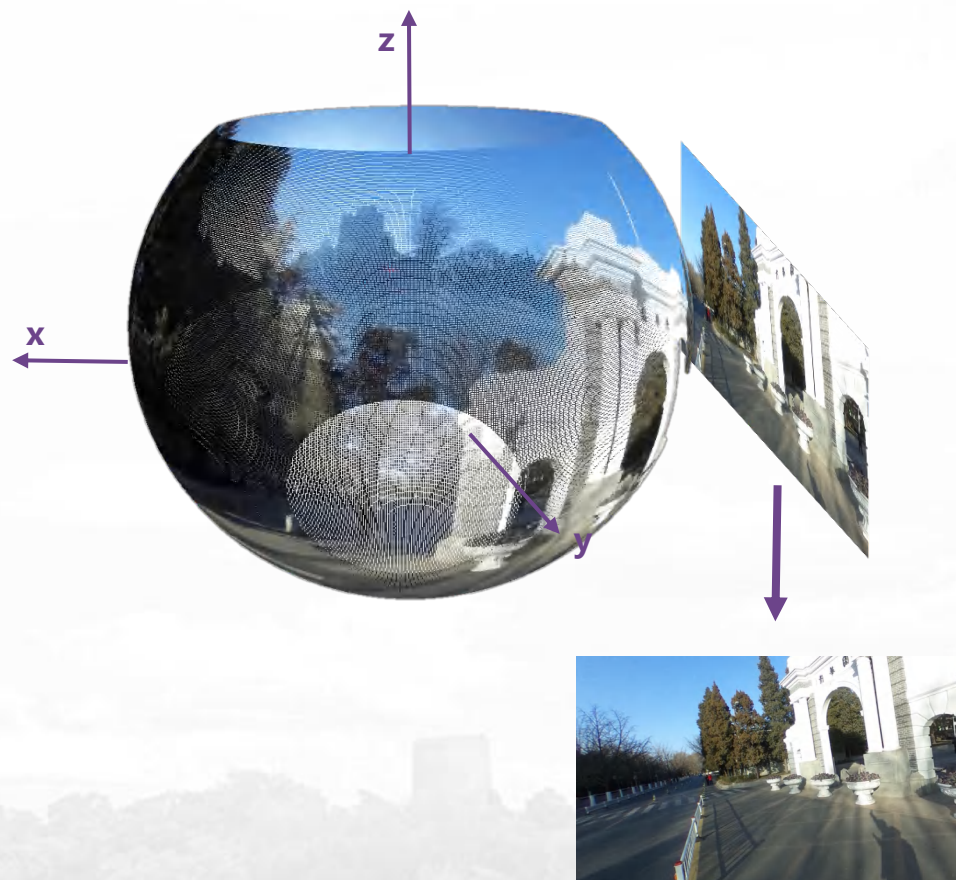
$$Z \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \triangleq KP$$

- 最近邻插值



□ 全景图像切分模块

- 成像原理示意图

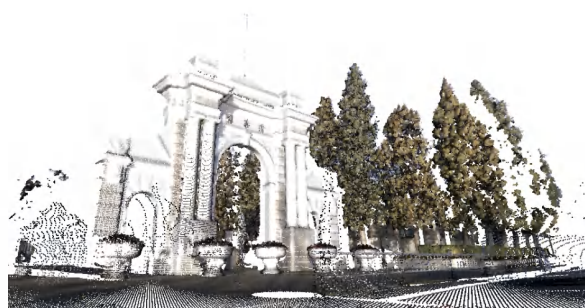




基于2D-3D匹配的视觉定位

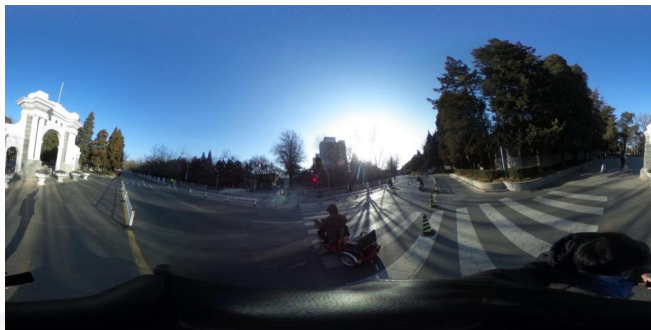
□ 点云内虚拟相机成像

- 实验结果
 - 水平视角1.5 rad
 - 垂直视角1 rad



□ 全景图像切分模块

- 实验结果
 - 水平视角 $\pi/2$ rad
 - 垂直视角1rad





基于2D-3D匹配的视觉定位

□ 基于深度学习的特征点匹配

- 虚拟成像结果与真实相机结果有差异
- 点云测站与全景相机之间位姿差距大
- 不同光照环境下更鲁棒



纹理边界不平滑



天空（远景）数据缺失

□ 基于SOS-Net的2D特征点匹配

- First Order Similativity

$$L_{FOS} = \frac{1}{N} \sum_{i=1}^N \max(0, t + d_i^{pos} - d_i^{neg})^2$$

- Second Order Similativity

$$R_{SOS} = \frac{1}{N} \sum_{i=1}^N \sqrt{\sum_{j \neq i}^N (d(x_i, x_j) - d(x_i^+, x_j^+))^2}$$

- Loss Function

$$L = L_{FOS} + R_{SOS}$$

□ SOS-Net训练网络

- L2-Net
- Pretrained model in github

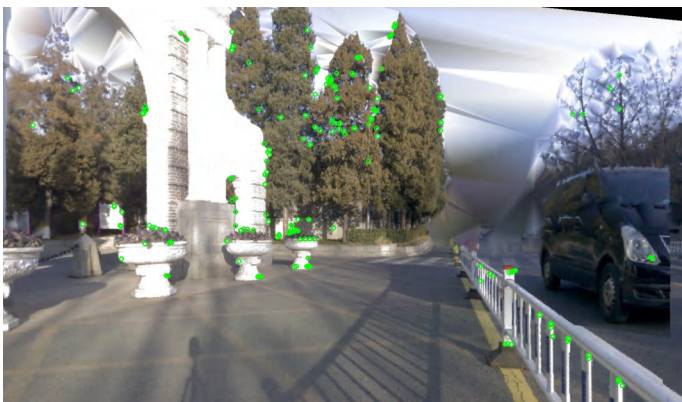




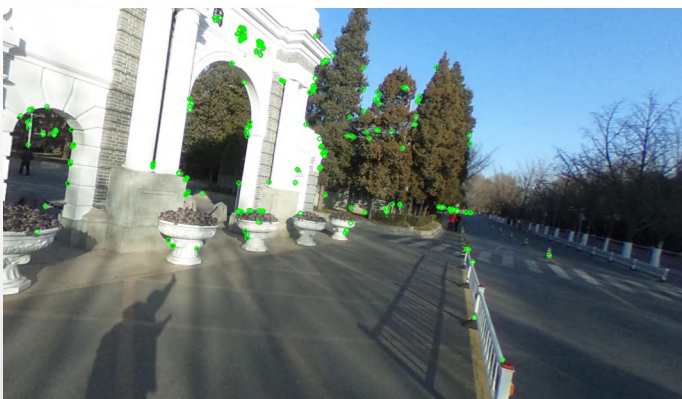
基于2D-3D匹配的视觉定位

□ 实验结果

● 特征点提取



img1: 点云虚拟成像图特征点提取



img2: 全景图像切分图特征点提取

● 特征点匹配



特征点匹配结果(19 matches)

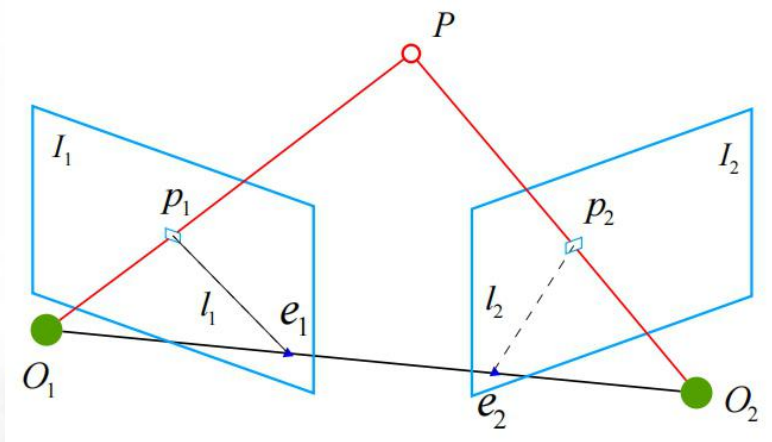


基于2D-3D匹配的视觉定位

□ 对极几何法求取相机旋转矩阵和位移向量

- 通过若干点对匹配估计相机外参
- 数学表示

$$p_2^T K^{-T} t^{\wedge} R K^{-1} p_1 = 0.$$



对极几何约束原理图

- 实验结果



```
R:
[[ 0.99858505  0.0509581  0.01520439]
 [-0.0511413  0.99862036  0.01191384]
 [-0.01457631 -0.01267455  0.99981343]]
t:
[[-0.7056037 ]
 [-0.26117016]
 [-0.65872116]]
```




后期研究计划

□ 将2D-3D匹配的思路应用在建图中

- 激光扫描仪采集成本高
- 纯基于视觉采集成本低

□ 视觉定位可视化

- 基于视频生成轨迹路线

□ 不同场景下重建和定位实验

- 校园环境：户外空旷场景（部分角度无法采集到有效数据）
- 仓储环境：室内重复场景（特征点少、重复场景多）

□ 论文撰写



谢谢大家 欢迎批评指正

Department of Automation, Tsinghua University

- >> 答辩学生: 自74 陈奕凡
- >> 指导老师: 陈宝华

