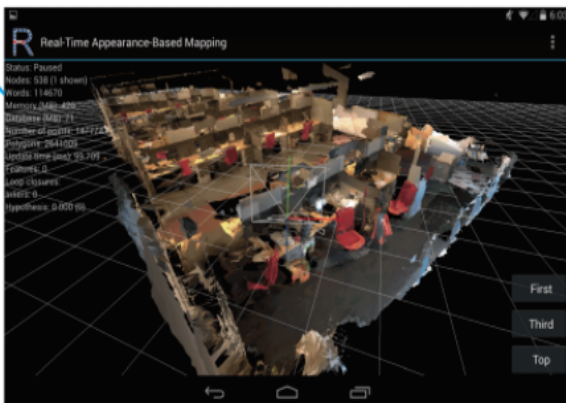
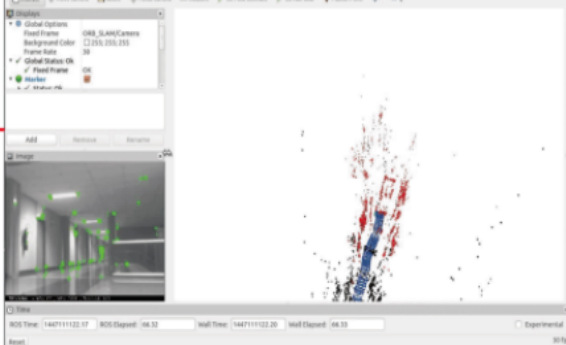
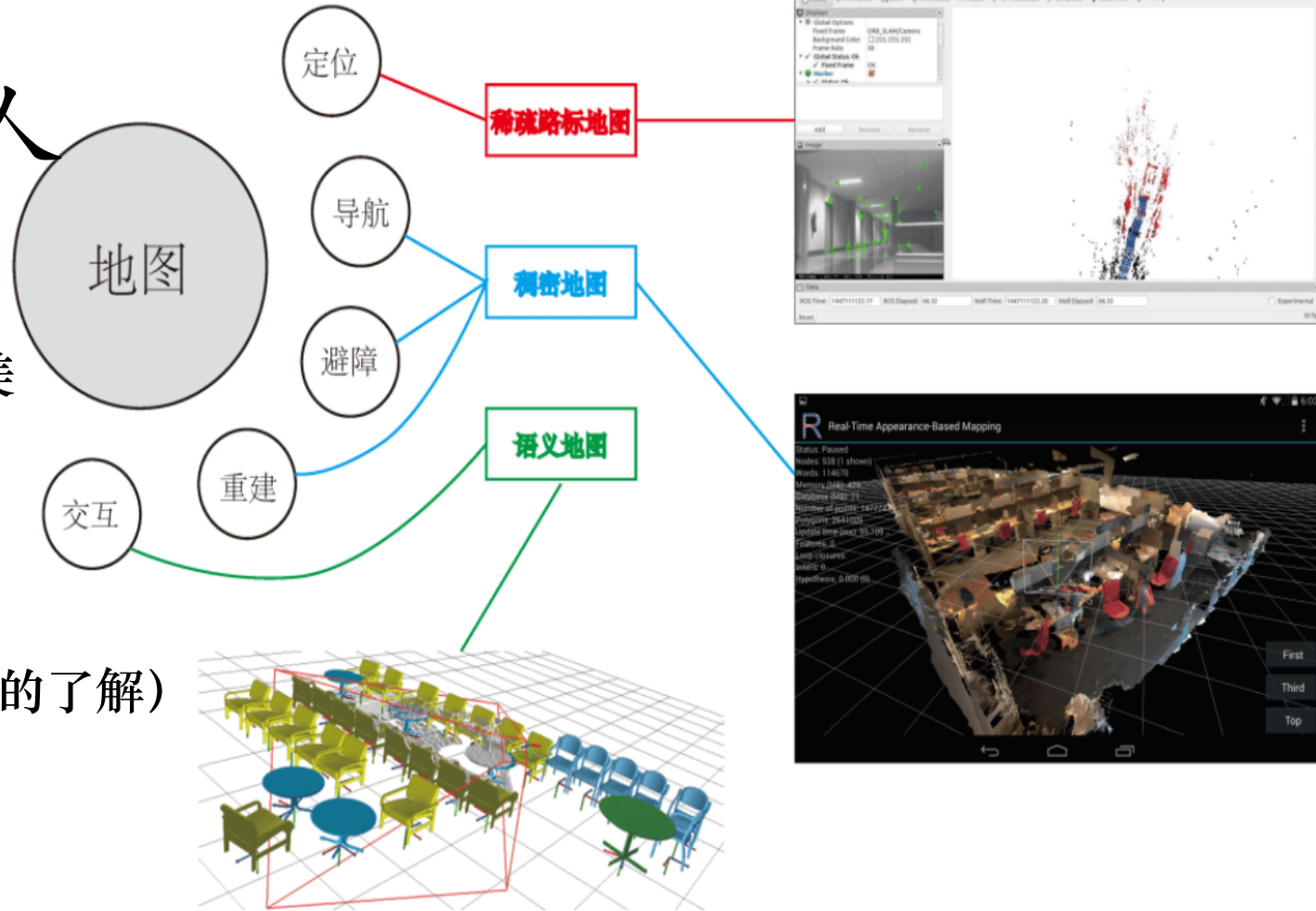


# 机器视觉

屠晓涵

2017年10月20日

# 背景：银行服务机器人



## 定位与建图：

机器人在陌生的环境中，比如说在一个布置精美的大厅里自主地运动，它至少需要知道两件事：

**定位：** 自己在哪里？（侧重于对自身的了解）

**建图：** 周围的环境是怎样的？（侧重于对外在的了解）

## 想法：

- 1、物体识别、分割往往只需考虑一个图，目前只拥有一个相机，如果把运动过程中的图片都带上物体标签，就能得到一个有标签的地图，根据有标签的地图，机器人就可以根据人的口令运动到任何目的标签处，我们的工作就更先进一点。
- 2、但物体识别和语义分割都需要大量的训练数据，要让机器人识别各个角度的物体，需要从不同的视角采集物体的数据，然后进行人工标定，这特别辛苦。而在SLAM中，由于我们可以估计相机的运动，可以自动计算物体在图像中的位置，节省人工标志的成本。如果我们采用深度学习方法，自动生成带有高质量标注的样本数据，将大大加速分类器的训练过程。

# 目的+问题（SLAM-同时定位与地图构建）

## 目的：

- 在没有环境先验信息的情况下，搭载传感器的机器人，要在运动时建立环境的模型，同时估计自己的运动，使自己具有自主运动能力。
- 采用最准确的方法实现机器人的实时定位与导航（“我在哪” “我要去哪” “我怎么去”）。

## 单目 SLAM 存在的问题：

- 1. 真实距离的确定，
- 2. 纹理特征较弱情况下稠密深度图像的获取，
- 3. 相机在纯旋转运动情况下立体匹配失效。

## 为解决这些问题，选择论文：

CNN-SLAM: Real-time dense monocular SLAM with learned depth prediction

会议CVPR: IEEE举办的计算机视觉和模式识别领域的顶级会议

# 经典SLAM框架

## 视觉里程计 Visual Odometry

定义：通过分析处理相邻图像序列来确定机器人的位置和姿态

## 后端优化

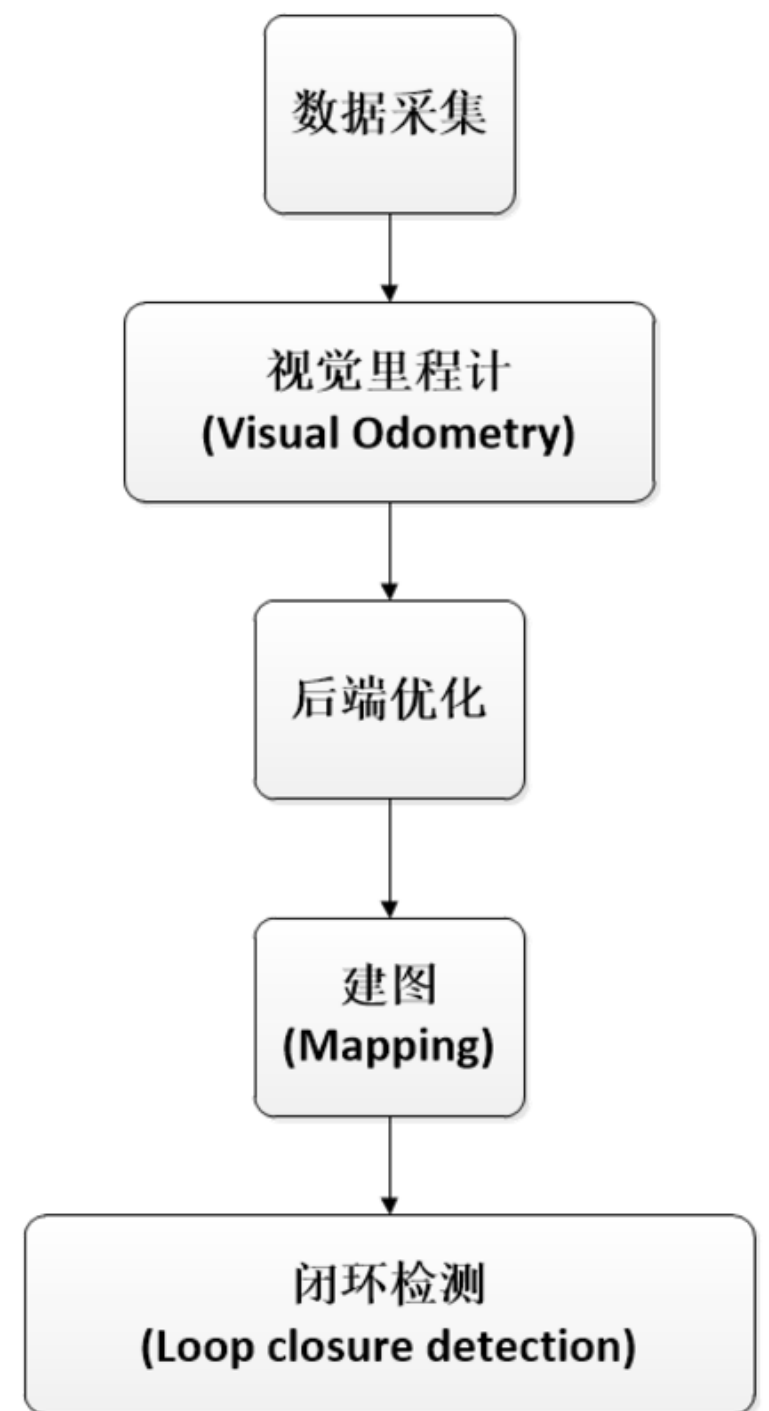
- 1、从带有噪声的数据中优化轨迹和地图（状态估计问题）
- 2、最大后验概率估计（MAP）
- 3、前期以 EKF 为代表，现在以图优化为代表

## 回环检测

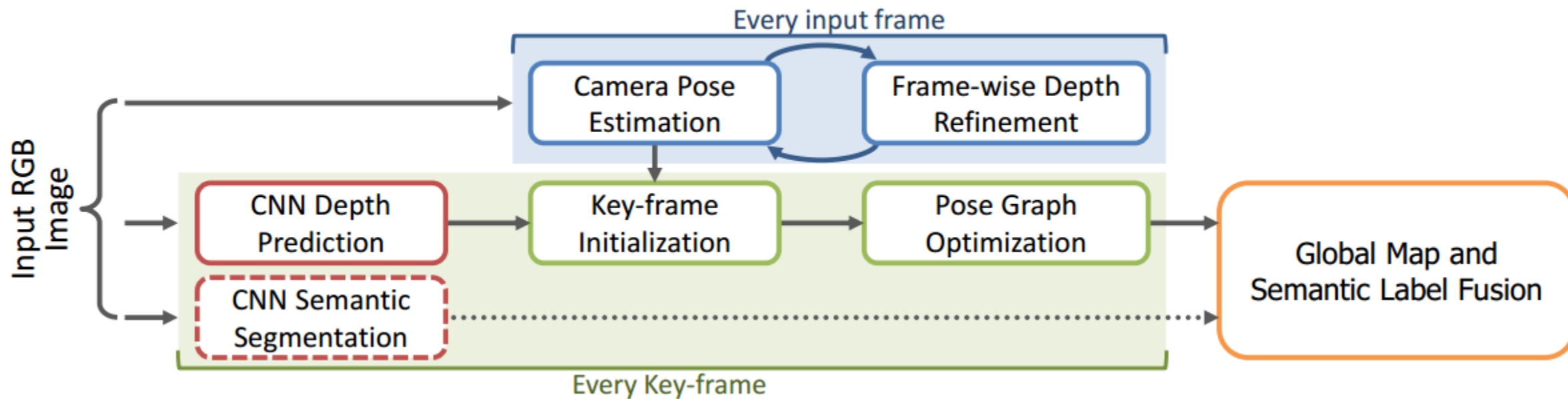
- 1、检测相机是否回到早先位置
- 2、识别曾经到过的场景
- 3、计算图像间的相似性
- 4、方法：词袋模型

## 建图

用于导航、规划、通讯、可视化、交互等



# 文章思路



该文章使用直接法估计相机姿态，使用CNN来估计Depth，以及做图像语义分割。然后将Geometry和semantic融合起来，生成具有语义信息的地图。

# 评估指标

- 数据集:

- ICL-NUIM数据集<sup>[1]</sup>和TUM RGB-D SLAM<sup>[2]</sup>数据集

- 定性评估:

- 纯旋转相机运动的鲁棒性

- 语义标签融合的准确性

- [1] A. Handa, T. Whelan, J. McDonald, and A. Davison. A benchmark for RGB-D visual odometry, 3D reconstruction and SLAM. In IEEE Intl. Conf. on Robotics and Automation, ICRA, Hong Kong, China, May 2014. 4, 5, 6, 7

- [2] J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers. A benchmark for the evaluation of RGB-D SLAM systems. In 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, pages 573–580, oct 2012. 6

# CNN-SLAM效果



Ground Truth



Ours



LSD-SLAM

Accuracy:

66.18%

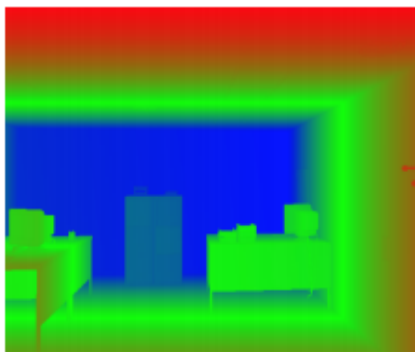
57.15%

11.91%

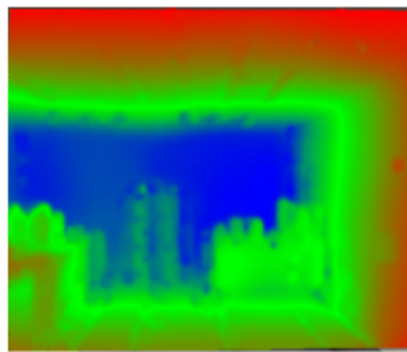
12.26%



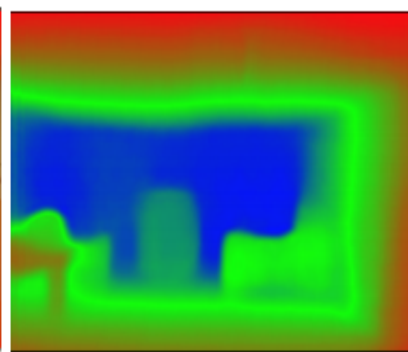
Color



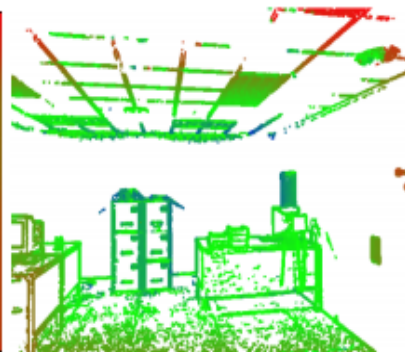
Ground Truth



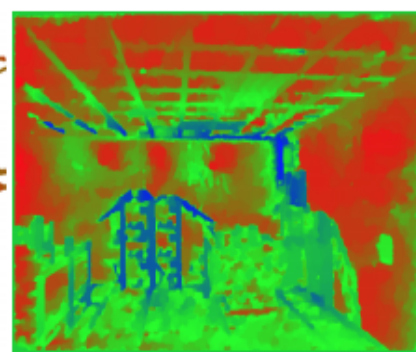
Ours



Raw Depth Prediction



LSD-SLAM



REMODE

文章在两个公开数据集上，对比了 LSD-SLAM、ORB-SLAM、REMODE 算法，验证了 CNN-SLAM 算法在：相机位姿精度、三维重建精度、相机旋转运动下算法稳定性方面的优势

# 下周工作

- 代码实现论文方法
  - 定位，使用SLAM+CNN来估计机器人位置。
  - 建图，使用CNN进行物体识别。
  - 将位置信息和物体识别融合起来，生成具有语义信息（标签）的地图。使机器人具有识别物体，找到特定物体的能力
- 高铁项目其他问题解决