

同濟大學

毕业设计(论文)任务书

课题名称 基于深度学习的语义 SLAM 研究
 与实现

副 标 题

学院 (系) 电子与信息工程学院
 计算机科学与技术系

专 业 计算机科学与技术

学生姓名 姚福飞 学 号 1452275

毕业设计(论文)起讫时间:
自 2020 年 2 月 17 日至 2020 年 5 月 22 日共 14 周

指 导 教 师 签 名 2020 年 2 月 15 日

教 学 院 长 (系 主 任) 签 名 _____ 年 ____ 月 ____ 日

一、毕业设计的课题背景

在机器人领域，智能服务机器人逐渐走上了行业的风口浪尖。我们的生活中也出现了越来越多移动机器人的身影。随着传感器技术、人工智能技术和计算技术等的不断提高，智能移动机器人逐渐的融入了人们的生产、生活和工作。无论是服务型移动机器人，比如餐厅里的机器人服务员、银行的机器人大堂经理，又或者是工厂里流水线上的移动装配机器人、物流行业的自动分拣机器人，都离不开 SLAM 实时定位与地图构建这一环节。如何让机器人清晰、高效的知道自己在哪里、周围有什么就成为了制约移动机器人工作能力和运行效率的一个重要条件。

目前，SLAM 技术已经广泛的应用于机器人定位导航、VR/AR、无人机、无人驾驶等领域。如 SLAM 可以辅助机器人执行路径规划、自主探索、导航等任务。智能家用扫地机器人都可以通过用 SLAM 算法高效绘制室内地图，智能分析和规划扫地环境。在 VR/AR 领域，SLAM 技术用于辅助增强视觉效果，构建视觉效果更为真实的地图，从而针对当前视角渲染虚拟物体的叠加效果，使之更真实没有违和感。在无人机领域，地图建模。SLAM 可以快速构建局部 3D 地图，并与地理信息系统 (GIS)、视觉对象识别技术相结合，可以辅助无人机识别路障并自动避障规划路径。在无人驾驶领域，SLAM 技术可以提供视觉里程计功能，并与 GPS 等其他定位方式相融合，从而满足无人驾驶精准定位的需求。

近年来由于深度学习 (Deep Learning) 的快速发展，尤其是全卷积神经网络 (Fully Convolutional Network, FCN) 的出现，实现了从图像像素到像素类别的变换，结合语义的 SLAM 技术得到了快速的发展。

构建语义地图都是一个大家都一致认同的发展方向，主要是因为目前视觉 SLAM 方案中所采用的图像特征的语义级别太低，造成特征的可区别性太弱。另外，采用当前方法构建出来的点云地图对不同的物体并未进行区分。这样的点云地图因为包含信息不足，再利用性十分有限。现阶段，语义 SLAM 的难点在于怎样设计误差函数，将深度学习的检测或者分割结果作为一个观测，融入 SLAM 的优化问题中一起联合优化，同时还要尽可能提高处理速度增强实时性。

在复杂环境下，怎样利用视频或时序图像中的语义信息，来增强移动机器人 SLAM 系统的实时定位和姿态确认一直是语义 SLAM 技术上一个亟待解决的问题。这一问题不仅涉及到复杂环境下的目标检测和语义识别，还涉及到因时序图像的帧冗余性和不确定性引发的语义分割标签的不稳定性。本选题是希望在这样的背景下，通过分析研究当前各类算法框架的特点，探索出一种基于图像语义的 SLAM 算法，以实现提高定位准确度的目的。

二、毕业设计的技术参数

本课题从现有的传感器和深度学习技术出发，探索使用深度学习实现语义 SLAM 的方法，使用深度学习进行图像语义分割并融合 SLAM 功能，实现一个性能优越的语义视觉 SLAM。

三、毕业设计应完成的具体工作

- (1) 查阅文献；
- (2) 撰写开题报告；
- (3) 文献分析和方案设计；
- (4) 实现基本 SLAM 功能；

- (5) 尝试使用深度学习实现图像语义分割;
- (6) 尝试融合 SLAM 和图像语义;
- (7) 实验数据结果分析;
- (8) 论文撰写和答辩;

四、毕业设计（论文）进度安排

序号	设计（论文）各阶段名称	时间安排（教学周）
1	查阅文献	2019.9~2020.1
2	撰写开题报告	（第1~2周）
3	文献分析和方案设计	（第3周）
4	实现基本 SLAM	（第4~6周）
5	尝试使用深度学习实现图像语义分割	（第7~8周）
6	尝试融合 SLAM 和图像语义	（第8~12周）
7	实验数据结果分析	（第13周）
8	论文撰写和答辩	（第14~15周）

同组学生姓名：

无

五、应收集的资料及主要参考文献

- [1] 高翔, 张涛等. 《视觉 SLAM 十四讲：从理论到实践》[M]. 北京:电子工业出版社. 2017:1-377.
- [2] Richard Hartley, Andrew Zisserman. 《Multiple View Geometry in Computer Vision》[M] Cambridge University Press. 2003.
- [3] Thrun S. Probabilistic robotics[M]. MIT Press, 2006
- [4] Kaiming He, Georgia Gkioxari, Piotr Dollar, Ross Girshick. Mask R-CNN[J]. arXiv:1703.06870v3 [cs.CV] 24 Jan 2018.
- [5] R. F. Salas-Moreno, R. A. Newcombe, H. Strasdat, P. H. J. Kelly, and A. J. Davison. SLAM++: Simultaneous localisation and mapping at the level of objects[J]. In 2013 IEEE Conference on Computer Vision and Pattern Recognition, pages 1352 - 1359, June 2013.
- [6] Jonathan Long, Evan Shelhamer, Trevor Darrell. Fully Convolutional

Networks for Semantic Segmentation[J]. arXiv:1411.4038v2 [cs.CV] 8 Mar 2015.

[7] Raul Mur-Artal, Juan D. Tardos. ORB-SLAM2: an Open-Source SLAM System for Monocular, Stereo and RGB-D Cameras. IEEE Transactions on Robotics[J]. 2017(33): 1255 – 1262.

[8] Michael Strecke, Jorg Stuckler. EM-Fusion: Dynamic Object-Level SLAM with Probabilistic Data Association[J]. arXiv:1904.11781v1 [cs.CV] 26 Apr 2019.

[9] Jurgen Sturm, Nikolas Engelhard, Felix Endres, Wolfram Burgard, Daniel Cremers. A Benchmark for the Evaluation of RGB-D SLAM Systems[J]. [2012 IEEE/RSJ International Conference on Intelligent Robots and Systems](#). 2012(13195110)

[10] Moritz Menze, Andreas Geiger. Object Scene Flow for Autonomous Vehicles[J]. [2015 IEEE Conference on Computer Vision and Pattern Recognition \(CVPR\)](#). 2015(15524438) DOI: [10.1109/CVPR.2015.7298925](https://doi.org/10.1109/CVPR.2015.7298925)