**RM: 精**

|  |  |
| --- | --- |
| 记录项 | 记录内容 |
| 文献名称 | LabelFusion A Pipeline for Generating Ground Truth Labels for Real RGBD Data of Cluttered Scenes |
| 文献作者 | Pat Marion\*, Peter R. Florence\*, Lucas Manuelli\* and Russ Tedrake  Computer Science and Artificial Intelligence Laboratory  Massachusetts Institute of Technology |
| 来源信息 | IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2018: 3235-3242。  DOI: 10.1109/ICRA.2018.8460950 |
| 关键词 | Pipelines，Three-dimensional displays，Robot sensing systems, Image segmentation, Cameras  , Image reconstruction |
| 针对性问题 | 人工标注数据集（图像分割、位姿估计）费时费力。 |
| 主要方法 | 基于已有的制作数据集的方法pipline ElasticFusion进行改进，提出自己的制作数据集的方法。 |
| 主要贡献 | 1. 主要贡献是，设计了一个开源的制作数据集的流程，只需要额外的RGBD相机就能制作数据集。极大地减少了杂乱场景人工每像素标注的时间。 2. 一个很大的带有标签的数据集。 3. 使用制作的数据集回答了训练DNN需要多少数据集以及数据集必须是什么质量才能获得较好性能的问题，对数据集的价值进行了经验量化。 |
| * Developing an open-source pipeline that reduces the amount of human annotation time needed to produce labeled RGBD datasets for training image segmentation neural networks. * Leveraging dense RGBD reconstruction to fuse together RGBD images taken from a variety of viewpoints. * Labeling with ICP-assisted fitting of object meshes. * Automatically rendering labels using projected object meshes. * Enabling the collection of over 1,000,000 labeled object instances in just a few days. * Providing a dataset and annotation pipeline that are available at labelfusion.csail.mit.edu. * Using the dataset to answer questions related to how much training data is required and of what quality the data must be to achieve high performance from a DNN architecture. |
| 核心算法记录（精读） | 1. 采集RGBD数据集。使用RGBD摄像机从多个视点收集场景的RGBD视频。 2. 利用ElasticFusion（现有的重建技术都可以替代，只要能够提供相机姿态就行）对场景进行三维密集重建。 3. 物体的网格生成。使用扫描仪扫描、建模等获得物体的网格。**（可能需要解决高反光物体和透明物体深度异常的问题）**。 4. 人工三次点击对网格进行姿态初始化，提供一个粗略的初始位姿，再基于ICP方法进行拟合，最终实现较为精准的对齐。 5. 对三维场景的标注结果进行渲染，得到每个场景每张图片的标注结果。 |
| 主要实验平台 | NVIDIA GTX1080 GPU |
| 实验数据集 | 制作的人工标注的数据集，包含分割、6D位姿  地址：labelfusion.csail.mit.edu |
| 性能指标 | IOU |
| 性能比较对象 | 主要的比较对象及其参考文献，如：  比较对象1描述，文中文献编号为[x1]  比较对象2描述，文中文献编号为[x2]  …… |
| 主要实验及目的（简记） | **实验一 ：数据生成pipline的评估：**   1. 对不同场景（严重杂乱且多物体、低光照、运功模糊、高相机与物体间距、25种不同环境）里的物体进行标注，效果都比较好。 2. 对数据收集时间、三维场景重建时间、人工三点标注时间和渲染时间做了分析，结果显示大部分时间花在渲染上面，人工标注的时间占所有时间的一小部分。   **实验二 ：** **讨论数据集的哪些因素会影响模型的泛化性。**   1. 首先研究严重遮挡和杂乱的多物体场景和单物体场景。作者分别使用18个单物体场景（a）、三个多物体场景（b）、a+b（c）、18个多物体场景（d）、和a+d（e）对模型进行训练，使用模型分别对单物体场景和多物体场景进行测试，最终得到每个网格的IOU图。      1. 其次，测试了不同环境的训练数据对模型的性能影响。实验结果显示随着背景数量增加，性能会随之提高。      1. 最后研究30hz是否必要，以及机械臂（手持相机）移动速度对性能的影响。结果显示随着有效传感器速率的增加，分割性能呈单调增长。但是对于较慢移动速度，在0.3hz以后回报率明显下降，而手持数据能够带来更高收益。 |
| 对文献的个人思考 | 1. ICP的方法应该还能优化。 2. 这个管线是开源的，能够用该管线来制作自己的数据集。所以后续优化管线是一个点，提出自己的数据集是一个点。 |
| 扩展阅读 |  |