论文:

Dubé R, Dugas D, Stumm E, et al. **SegMatch: Segment based loop-closure for 3D point clouds**[J]. arXiv preprint arXiv:1609.07720, 2016.

作者来自 Ethz asl 实验室

相关 YouTube 视频: https://www.youtube.com/watch?v=iddCgYbgpjE

相关开源代码: https://github.com/ethz-asl/segmatch

一. 背景介绍

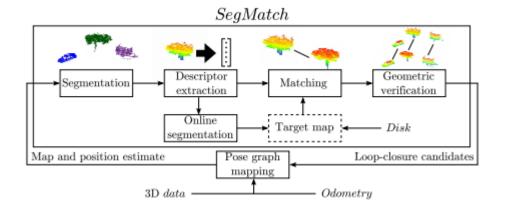
SLAM 里面很重要也很具有挑战性的一环是闭环检测。闭环检测可以使移动机器人真正地理解环境的结构[1],而不是简单的里程累积。机器人的重定位(relocalization)也可以看作是闭环检测的一个特例,比如机器人被绑架到已有地图之外,回到原有地图则需要有闭环检测的功能。

移动机器人利用视觉的 Bag of Words[2]可以检测闭环,2D 激光也可以通过 ICP 等 scan matching 算法来强行匹配二维点云数据。但是 3D 激光点云数据量更大,几何信息量更丰富,如果直接 ICP 等算法来检测闭环计算量会很大,也不准确。该论文里提出目前在 3D 点云里主要有三类方法:

- a. 基于新进来的点云(简称新点云, new scan)的局部特征。类似于图像的 SURF 特征点等等, 激光点云也可以提取出 Gestalt 特征[3]等等。通过特征 点匹配与投票等方法, 可以大致检测到机器人已经进入形式过的区域。
- b. 基于新点云的全局特征。比如,记录每簇点云在全局坐标系下的直方图特征,比如高度、密度、角度等。当地图上某段点云与新点云很相似时,再通过 ICP 等匹配算法计算位姿关系。
- c. 基于平面或物体特征。此类方法可以利用平面等物体特征来检测闭环,但 有时对环境的单一性要求比较高。

二. SegMatch 算法

SegMatch 算法可以分为四个模块: 分割(Segmentation)、特征提取(Descriptor extraction)、匹配(Matching)以及几何验证(Geometric Verification)。如论文中所示:



- 1. Segmentation 该算法需要将点云分割,论文中采用的是 Douillard 等人[4]的方法。该方法首先将地面提取出来并去除,然后将地面以上的物体用欧氏距离聚类可得到分割的物体。
- **2. Feature extraction** 对于一簇点云,目前没有 gold-standard 方法来提取特征。 SegMatch 利用了很多特征。文中只提到了两种特征,其中一种是基于特征向量[5],可以通过计算来描述点云的平面化、散射性等等。
- 3. Matching 获得了众多物体的特征之后,很难通过调整参数的方法来判断两 簇点云描述的是同一个物体。论文采用机器学习的方法来训练分类器,还在 特征空间内使用 kd 树来建立搜索树,以提高效率。具体采用的算法是随机森林,随机森林作为一种传统的机器学习算法有很多优势,这里就不一一赘述 了。当两簇点云的特征给定后,计算特征向量的差放入随机森林,如果最后 得到的分数大于一定阈值则认为是同一个物体。
- **4. Geometric Verification** 得到的匹配物体通过 RANSAC 来剔除不正确的匹配。 当匹配得到的物体数大于一定阈值可以认为检测到了闭环。分割中心可以用 来估计位姿变换。

增量式分割

提取分割物体的同时也在建立布满分割物体的地图。为了滤除噪声,论文提出只保留一定范围距离内的点云,此外如果探测到两个物体距离很近可以认为他们是同一个物体,这样就避免了重复的探测。在检测到闭环之后,已有地图上的所有物体位姿都得到了更新。

作者在 KITTI 数据集上做测试,效果要优于基于特征点的闭环检测方法。在做实物测试时,用 I7-4900MQCPU 的检测速度是 1Hz。

三. 个人想法

SegMatch 方法是 16 年 9 月发布的,应该是最新的基于 3D 激光点云来检测闭环的方法。个人认为该方法其实是糅合了基于特征点和物体的检测方法,但是有以下两个问题:

- 1. 从 Ethz-asl 实验室发布的视频以及 Wiki 来看,提取出来的物体基本是三类: 树、汽车以及墙。如果环境发生了缓变(比如汽车开走了),那么闭 环检测就会有可能受到影响。而墙的长度有可能大于激光的探测范围而 被当作两个物体,分割聚类的中心也估计不准,可能对闭环检测产生影响。
- 2. 通过机器学习训练分类器来匹配物体是一个很好的 idea,但是投入实物使用要采集大量样本,而且实物运行的时候会影响闭环检测的速度。如果环境比较单一(比如停车场或者农田),其实对物体模型有很容易的估计,可以不用随机森林分类器而使问题简单化。

此外,该论文大量篇幅在论述实验方法与结果,而对 SegMatch 具体方法描述 很淡,带有明显的 Ethz-asl 的风格。

相关文献:

- [1] Cadena C, Carlone L, Carrillo H, et al. Simultaneous Localization And Mapping: Present, Future, and the Robust-Perception Age[J]. 2016, 32(6).
- [2] Filliat D. A visual bag of words method for interactive qualitative localization and mapping[C]//Robotics and Automation, 2007 IEEE International Conference on. IEEE, 2007: 3921-3926.
- [3] Bosse M, Zlot R. Place recognition using keypoint voting in large 3D lidar datasets[C]//Robotics and Automation (ICRA), 2013 IEEE International Conference on. IEEE, 2013: 2677-2684.
- [4] Douillard B, Underwood J, Kuntz N, et al. On the segmentation of 3D LIDAR point clouds[C]//Robotics and Automation (ICRA), 2011 IEEE International Conference on. IEEE, 2011: 2798-2805.
- [5] Weinmann M, Jutzi B, Mallet C. Semantic 3D scene interpretation: a framework combining optimal neighborhood size selection with relevant features[J]. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2014, 2(3): 181.
- [6] http://www.cvlibs.net/datasets/kitti/