

控制与决策

Control and Decision

基于视觉语义与激光点云交融构建的SLAM算法

佟国峰, 杨宇航, 彭浩, 孟祥政, 殷旗君

引用本文:

佟国峰, 杨宇航, 彭浩, 孟祥政, 殷旗君. 基于视觉语义与激光点云交融构建的SLAM算法[J]. *控制与决策*, 2024, 39(1): 103–111.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2022.0814>

您可能感兴趣的其他文章

Articles you may be interested in

一种基于多层语义特征的图像理解方法

An image understanding method based on multi-level semantic features

控制与决策. 2021, 36(12): 2881–2890 <https://doi.org/10.13195/j.kzyjc.2020.0927>

无人系统视觉SLAM技术发展现状简析

A survey of visual SLAM in unmanned systems

控制与决策. 2021, 36(3): 513–522 <https://doi.org/10.13195/j.kzyjc.2019.1149>

基于资源特征融合的复杂型面配合决策方法

Decision-making method for complex surface coordination based on resource feature fusion

控制与决策. 2021, 36(6): 1368–1376 <https://doi.org/10.13195/j.kzyjc.2019.1311>

基于DST融合多视图模糊推理赋值的三维目标检测

3D object detection based on DST fusion multi-view fuzzy reasoning assignment

控制与决策. 2021, 36(4): 867–875 <https://doi.org/10.13195/j.kzyjc.2019.0434>

融合稀疏编码与深度学习的草图特征表示

A feature representation of sketch based on fusion of sparse coding and deep learning

控制与决策. 2021, 36(3): 699–704 <https://doi.org/10.13195/j.kzyjc.2019.0941>

基于视觉语义与激光点云交融构建的SLAM算法

佟国峰[†], 杨宇航, 彭浩, 孟祥政, 殷旗君

(东北大学 信息科学与工程学院, 沈阳 110004)

摘要: 激光雷达作为同时定位与地图构建(SLAM)传感器之一,因精度高、性能稳定等特点而被广泛研究使用.但其获得的点云数据较稀疏,包含特征信息少,会导致误匹配、位姿估计误差大等问题,影响SLAM的定位和建图精度.对此,提出一种将视觉语义信息与激光点云数据融合的SLAM算法(VSIL-SLAM).首先,基于投影思想将聚类后的点云映射到语义检测框内,生成语义物体,解决原始激光点云特征稀疏问题;然后,在形状特征的基础上引入拓扑特征对语义物体进行表述,提出基于匹配的拓扑相似性度量方法,解决单一特征造成的误匹配问题,提高匹配准确度;最后,加入语义物体点到点的几何约束,基于几何特征和语义物体构建前端里程计,并完成后端回环检测和位姿图优化设计.实验结果表明,所提出算法在定位和建图效果上都有显著提高,改善了激光SLAM算法的性能.

关键词: 视觉语义信息; 点云聚类; 融合算法; 拓扑相似性度量; 前端里程计; 激光SLAM

中图分类号: TP391

文献标志码: A

DOI: 10.13195/j.kzyjc.2022.0814

引用格式: 佟国峰,杨宇航,彭浩,等.基于视觉语义与激光点云交融构建的SLAM算法[J].控制与决策,2024,39(1): 103-111.

SLAM algorithm based on fusion of visual semantics and laser point cloud

TONG Guo-feng[†], YANG Yu-hang, PENG Hao, MENG Xiang-zheng, YIN Qi-jun

(College of Information Science and Engineering, Northeastern University, Shenyang 110004, China)

Abstract: Lidar, as one of SLAM (simultaneous localization and mapping) sensors, has been widely studied and used due to its high precision and stable performance. However, the obtained point cloud data is sparse and contains little feature information, which could lead to some problems such as mismatching and pose estimation error, and affect the localization and mapping accuracy of SLAM. In view of the above problems, an SLAM algorithm (VSIL-SLAM) that integrates visual semantic information and laser point cloud data is proposed. Firstly, based on the projection idea, the post-clustering point cloud is mapped to the semantic detection frame to generate semantic objects and solve the problem of scarce features of the original laser point cloud. Then, on the basis of shape features, topological features are introduced to describe semantic objects, and a topological similarity measurement method based on matching is proposed to solve the problem of mismatching caused by single features and improve the matching accuracy. Finally, the front-end odometer is constructed based on geometric features and semantic objects by adding the geometric constraints of semantic objects point-to-point, and the back-end loopback detection and pose map optimization design are completed. Experimental results demonstrate that the proposed algorithm improves the performance of the laser SLAM algorithm in both localization and mapping.

Keywords: visual semantics; point cloud clustering; fusion algorithm; topological similarity measurement; front-end odometer; laser SLAM

0 引言

随着科技的不断进步,智能机器人被广泛应用于生活的各个领域,成为当今社会发展不可或缺的一部分.同时定位与地图构建(SLAM)技术作为智能机器人的核心支撑,在学术界和工业界已经备受关注.自

Smith等^[1]提出SLAM这一概念之后,SLAM算法经历了核心的3个发展阶段^[2],目前其技术已经逐渐成熟.根据选用的传感器不同,业内将SLAM分为视觉和激光两大类,其中视觉SLAM在语义感知方面有较好的表现^[3],但容易受光照强度影响.激光SLAM定

收稿日期: 2022-05-11; 录用日期: 2022-08-09.

基金项目: 国家重点研发计划项目(2019YFB1309905, 2020YFB1712802).

责任编辑: 冯俊娥.

[†]通讯作者. E-mail: tongguofeng@ise.neu.edu.cn.

*本文附带电子附录文件,可登录本刊官网该文“资源附件”区自行下载阅览.

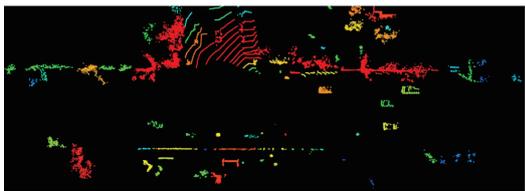
1.1.2 激光点云聚类

基于点的聚类方法包括划分聚类和层次聚类,精度主要受点云质量影响;基于体素的聚类方法主要受点云体素分辨率影响. 本文主要对激光点云进行聚类,选择基于点实现的聚类方法DBSCAN^[20],其核心思想是通过查找密度相连对象的最大集合来实现聚类.

为了降低该算法运行的复杂度,节省聚类时间,本文采用kdtree邻域搜索代替原始线性搜索方式,进行点云聚类. 选取KITTI数据集04序列中第251帧点云数据进行实验,设置DBSCAN聚类参数为 $r = 1.0$, $Minpts = 10$,实验结果如图4所示.



(a) 原始点云



(b) DBSCAN 聚类点云

图4 点云聚类结果

1.1.3 基于物体中心的融合算法

为了将上述得到的语义信息与点云聚类进行融合,本文提出一种基于物体几何中心的融合算法. 使用棋盘标定法获得相机外部参数,利用外参实现相机坐标系与激光雷达坐标系之间的转换,即

$$P_l = R_{lc}P_c + t_{lc}. \quad (1)$$

其中: P_c 为相机坐标系下点的坐标, P_l 为对应激光雷达坐标系下点的坐标, R_{lc} 、 t_{lc} 为相机外部参数.

相机的内部参数在出厂之后就已经固定,可以通过出厂商获得,也可以自行进行标定,利用相机内参实现相机坐标系与像素坐标系之间的转换,有

$$ZP_{uv} = Z \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = KP_c. \quad (2)$$

其中: K 为相机内参数矩阵, P_{uv} 为对应像素坐标系下点的坐标.

通过得到的内外参数,实现将激光点云数据投影到图像平面. 直接使用聚类点云进行投影处理,会造成较大误差,本文基于检测框中心与聚类点云中心近

似重合思想,提出基于物体中心实现的视觉语义与激光点云融合算法,具体流程如下.

1) 语义信息的提取是在相机视野内图像进行的,所以先对原始点云进行分割处理,再进行聚类,得到相机视野内的聚类点云. 点簇的中心为 O_i , 点簇标签为 C_i , 将两者进行组合可表示为 $\{(C_1, O_1), (C_2, O_2), \dots, (C_n, O_n)\}$.

2) 将聚类后点云投影到图像中,利用点云簇中心点与目标检测框中心点的距离以及点簇比值构建判别阈值,即

$$d_{ji} = 1 - 2 \times \frac{\text{dis}(O_i^{\text{fia}}, O_j^{\text{pro}})}{\sqrt{w_i^2 + h_i^2}}, \quad (3)$$

$$a_{ji} = 1 - \frac{n_{ji}}{N_j}, \quad (4)$$

$$p_{ji} = k_1 d_{ji} + k_2 a_{ji}. \quad (5)$$

其中: p_{ji} 为第 j 个激光点簇属于目标检测结果中第 i 个目标的判定阈值, d_{ji} 为距离相关性, w_i 、 h_i 分别代表第 i 个目标检测框的宽度和长度, $\text{dis}(O_i, O_j)$ 为两中心点在图像平面内的距离, a_{ji} 为映射结果相关性, n_{ji} 为第 j 个点簇落在第 i 个目标检测框的点数, N_j 为第 j 个激光点簇的总点数.

点簇与检测框中心点距离越近,即 d_{ji} 值越大时,两者关联性越强. 如果 d_{ji} 相近,则利用 a_{ji} 进行判断,选择投影后包含点数较多的检测框进行关联. 若不满足上述两个条件,则寻找新的目标检测框进行判断,最后将对应的语义标签赋给激光点簇.

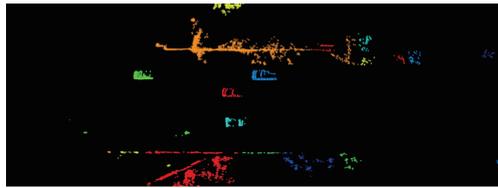
3) 保存所有得到语义标签的激光点簇,作为语义物体点云,其中每个点的信息包括 $(X_i^{C_j}, Y_i^{C_j}, Z_i^{C_j}, C_j, l_i, p_i)$, C_j 为点簇标签, l_i 为类别标签, p_i 为置信度.

利用该算法进行实验,将生成的语义物体点云投影到检测结果图像中,如图5所示. 结果表明,本文得到的语义物体点云消除了环境周围的噪声点及小物体点云,仅保留了对系统有利的大目标物体点云.

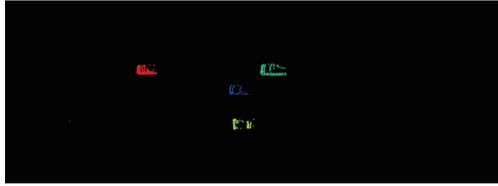


图5 语义物体点云在目标检测结果中的投影

如图6所示,对比融合语义信息前后的点云聚类情况,视觉语义信息与激光点云的融合相当于对聚类结果进行了深层次过滤处理,得到含有语义信息的点云物体,提高了准确率和鲁棒性,为后续SLAM系统的匹配、位姿估计等工作提供了有利条件.



(a) 融合前聚类点云



(b) 融合后聚类点云

图6 融合语义信息前后点云的聚类结果

1.2 联合拓扑特征进行数据关联

如果仅使用形状特征进行特征表述,则当场景中的物体相似性较高时,直方图会难以被区分,匹配误差较大.因此,本文在语义点云形状特征基础上通过引入周围物体的拓扑特征来提高可靠性.

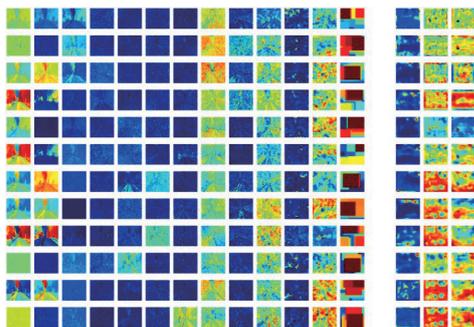
1.2.1 语义点云形状特征

本文采用点云和图像两个层次联合构建对象级的语义点云形状特征.点云形状特征可以通过点特征直方图(PFH)^[21]或视点特征直方图(VFH)^[22]进行表示,其中VFH描述子识别能力及效率较好,因此本文采用VFH表示语义物体点云层次的形状特征.

图像层次采用目标检测得到的特征图来表示.神经网络中主要通过卷积操作来提取特征,每个卷积得到的二维图像被称为特征图.如图7所示,图7(a)为需要检测的图像,图7(b)为卷积层得到的部分特征图.



(a) 被检测图像



(b) 卷积层得到的部分特征图

图7 网络层特征图

根据图2中YOLOv4-Tiny网络结构可以看出,其最后一个卷积层输出维度为 $26 \times 26 \times 255$,故特征图维度为255.将目标检测框中心点所在网格的所有特征值联合起来,用来表示该语义物体在图像层次的形状特征.

1.2.2 语义点云拓扑特征

基于图论思想,将语义物体视为图中一个节点,通过随机游走算法进行遍历,获得该语义物体的特征描述子.如图8所示,以中间红色外圈节点为例,设置游走次数和深度都为4,开始进行遍历,把经历过节点的语义信息按行进行存储,游走结束后,生成 4×4 的矩阵,便为当前语义点云的拓扑描述子.

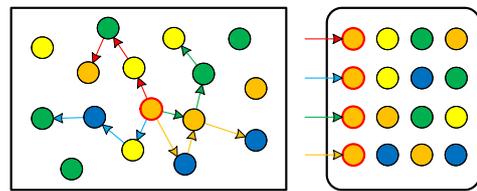


图8 拓扑特征表述

相比于传统拓扑特征的表示方法,上述形式不但能在匹配过程中减轻计算负担、节省时间,还能提高后续位姿估计的精度,对整个SLAM系统性能的提升非常有利.

1.2.3 语义点云数据关联算法

对于特征相似性,本文分别针对形状特征和拓扑特征单独进行度量.其中:形状特征通过相关性的方法进行度量,拓扑特征利用本文提出的拓扑特征相似性度量方法进行度量.

1) 形状特征.

基于皮尔逊相关系数得到直方图的相关性比较公式为

$$d(H_1, H_2) = \frac{\sum_I (H_1(I) - H'_1)(H_2(I) - H'_2)}{\sqrt{\sum_I (H_1(I) - H'_1)^2 (H_2(I) - H'_2)^2}}, \quad (6)$$

$$H'_i = \frac{1}{N} \sum_I H_i(I). \quad (7)$$

其中: H 代表直方图向量; N 代表直方图向量元素的数目; $d(H_1, H_2)$ 的绝对值介于0和1之间,越接近1,两直方图的相似性越高,反之亦然.

2) 拓扑特征.

对于拓扑特征,采用两节点描述子逐行匹配的方式进行相似性衡量,如图9所示.

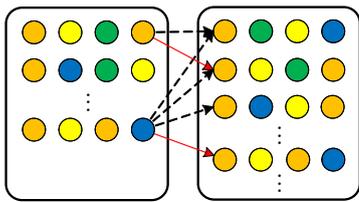


图9 拓扑特征相似性匹配

通过行与行之间的查找,可以找出互相对应的行描述子,最后得到对应的总行数,将其与两节点特征描述子的最大行数进行比值,有

$$d(n_1, n_2) = \frac{N(n_1 \cap n_2)}{N(\max(n_1, n_2))}. \quad (8)$$

其中: n 代表节点的拓扑描述子, $d(n_1, n_2)$ 为相关性指标. 通过 $d(n_1, n_2)$ 值的大小来衡量拓扑特征的相似程度.

将上述两种特征表述共同应用于语义物体的数据关联中,互相补充,能够避免因单一特征而导致的错误关联,提高实例物体关联的准确性.

1.3 联合位姿估计

激光SLAM算法中常用角点和平面点来构建位姿估计问题,为了提高前端位姿估计的准确度,本文扩展了基于语义物体点的位姿估计. 首先采用文献[5]的方法,通过当前帧 F_n 和参考帧 $F_{(n-1)}$ 构建角点到线以及平面点到平面之间的距离约束. 然后基于ICP点匹配^[23]的思想,从参考帧 $F_{(n-1)}$ 的语义物体中查找当前帧语义物体上点 i 的近邻点 j .

通过两点各自的坐标计算点点之间的距离来建立距离约束,有

$$d_0 = |X_{(n,i)}^L - X_{(n-1,j)}^L|. \quad (9)$$

由于基于几何特征构建的点线、点面距离以及基于语义物体构建的点点距离,其本质都是关于位姿的函数,可统一表示成

$$f(T) = d. \quad (10)$$

其中: T 代表位姿, d 代表距离.

针对式(10),利用非线性迭代方法进行位姿求解,实现基于几何特征和语义物体的位姿估计. 动态物体会影响激光里程计的精度,所以本文将当前和历史地图中语义物体的欧几里德距离与设定阈值进行比较,滤除动态目标,仅使用静态语义物体进行位姿估计.

为了说明语义拓扑特征对匹配精度的影响,在KITTI数据集上分别运行仅使用几何特征匹配的里程计算法和融合语义拓扑特征之后进行匹配的里程计算法,采用Layer-35网络层特征表示语义点云图像

层次的形状特征. 将评估轨迹与真实轨迹进行对比,结果如表1所示.

表1 融合语义拓扑特征前后里程计轨迹误差

序列	平均位置误差(%)/平均角度误差(deg/m)	
	融合拓扑特征前	融合拓扑特征后
01	2.892 0/0.007 9	2.846 9/0.007 3
04	1.585 6/0.006 4	1.585 6/0.006 4
09	2.204 1/0.008 4	1.865 1/0.007 5
10	2.964 3/0.009 8	2.640 5/0.008 4

由表1可见,除未识别到静态语义点云的04序列外,其余序列在融合语义拓扑特征之后,轨迹误差都有所下降,表明本文所提出的语义拓扑特征可以提高匹配精度,改善位姿估计的结果.

1.4 闭环检测及位姿图优化

后端采用基于位置关系实现的闭环检测,通过将当前关键帧与局部候选关键帧集合进行匹配的方式来确定闭环,缩小检测范围,减少计算量. 具体算法步骤如下.

step 1: 选取当前关键帧 P_i , 在历史帧中查询与其距离最近且时间间隔最长的帧作为回环帧 Q_l .

step 2: 在回环帧 Q_l 附近提取局部候选关键帧集合 $E\{Q_{l-m}, \dots, Q_{l+m}\}$.

step 3: 判断两关键帧内三维点数量是否分别大于阈值 N_{cur}, N_{pre} .

step 4: 执行ICP匹配,计算两关键帧得分情况,即

$$S(P_i, Q_j) = \sum_{k=0}^N (p_k - q_k)^2. \quad (11)$$

其中: P_i, Q_j 代表当前帧和候选帧, (p_k, q_k) 代表对应的最近点对.

利用点对的距离平方和作为得分,当检测到闭环时,将其约束加入到位姿图中进行全局优化. 图10为本文位姿图优化示意,主要包括激光里程计约束和回环检测约束,利用当前关键帧与附近局部关键帧集合进行 scan-to-map 匹配并构成里程计,根据当前帧和

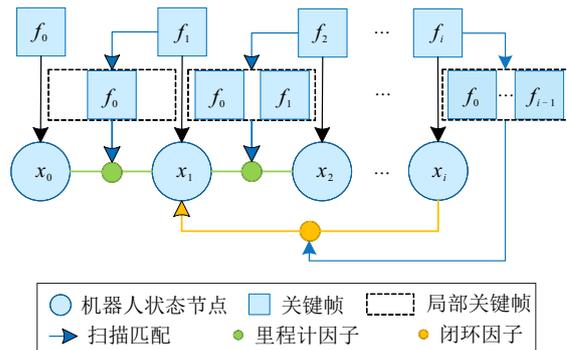


图10 位姿图优化

候选关键帧集合确定回环,执行优化,降低SLAM算法的累计误差。

2 系统概述

通过上述内容实现了融合语义信息的激光SLAM设计,包括系统前端里程计、后端闭环检测、位姿图优化。图11为完整SLAM算法的系统框架。

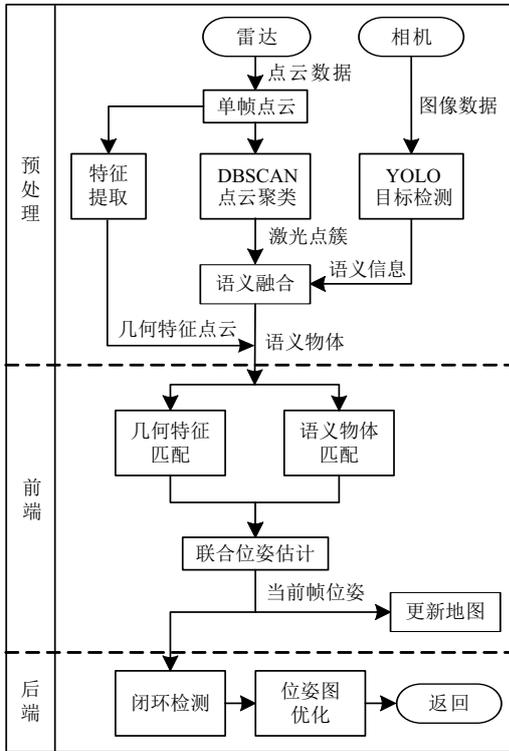


图11 本文SLAM系统框架

该框架主要分3部分。预处理:完成语义信息融合,提取语义物体;前端里程计:进行特征匹配,联合优化估计当前帧位姿,并对点云地图进行构建和更新;后端优化:检测闭环,利用位姿图进行全局优化。

3 实验结果与分析

选取KITTI数据集分别进行3组实验,验证本文算法性能的完整性及优越性。实验环境硬件配置如表2所示,软件配置包括64位Ubuntu18.04操作系统、

CUDA11.6、CUDNN8.3.0,使用GPU环节加速神经网络的计算。

表2 硬件配置

名称	参数
处理器	Intel(R) Core(TM) i5-10400F CPU @ 2.90 GHz
显卡	NVIDIA Corporation Device 1f06 (rev a1)
内存	24 G
硬盘	2 T

利用轨迹图像及精度作为评估指标,轨迹图像采用EVO进行评估,轨迹精度采用KITTI官网方法度量。从估计的位姿中按固定间隔 d 选取两帧,再从真实位姿中选出对应的两帧,分别计算位姿变换,即

$$\Delta T' = \begin{bmatrix} \Delta R' & \Delta t' \\ 0 & 1 \end{bmatrix}, \Delta T = \begin{bmatrix} \Delta R & \Delta t \\ 0 & 1 \end{bmatrix}. \quad (12)$$

根据估计位姿与真实位姿之间的变换矩阵,得到旋转误差和平移误差

$$\Delta \theta = \arccos\left(\frac{\text{tr}((\Delta R')^T \Delta R) - 1}{2}\right), \quad (13)$$

$$\Delta t = \sqrt{((\Delta R')^T (\Delta t - \Delta t'))^T (\Delta R')^T (\Delta t - \Delta t')}. \quad (14)$$

最后,对整个序列中连续间隔 d 的两帧位姿误差求和并取平均值,利用平均误差度量实验结果的准确性。

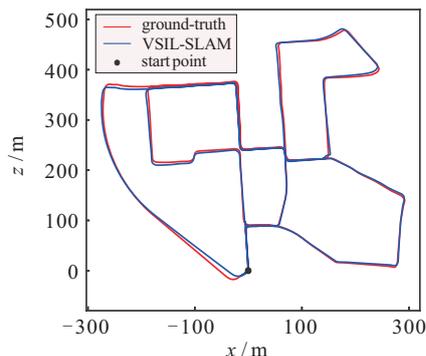
3.1 测试实验

本文分别选取KITTI数据集中的户外大场景00和小场景07序列对算法性能进行测试,利用本文提出的SLAM算法运行两个序列,输出点云地图,并保存轨迹信息,最后与KITTI真值轨迹数据进行对比。

图12为00序列运行结果。其中:平均位置误差为1.9101%,平均角度误差为0.0082 deg/m。图13为07序列运行结果,红色突出部分为运行过程中生成的静态语义物体。可以看出,本文算法能正确提取出场景中大量语义物体。提高定位和建图效果,输出轨迹与真实轨迹近似吻合。

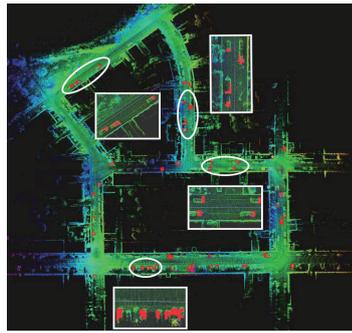


(a) 语义物体点云地图

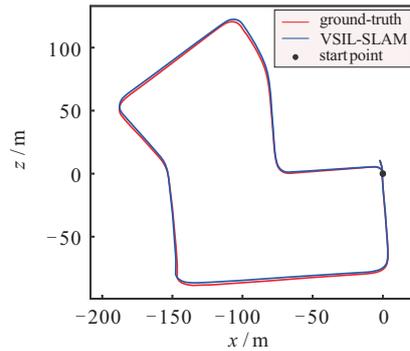


(b) 轨迹评估结果

图12 测试实验结果(00序列)



(a) 语义物体点云地图



(b) 轨迹评估结果

图13 测试实验结果(07序列)

除此之外,本文对该算法中生成语义物体部分进行了时间统计,取各阶段的平均运行时间,如表3所示.可以看出,本文所提出的算法在运行效率上也有良好的表现.

表3 语义环节平均运行时间 单位: s

语义环节	平均运行时间
目标检测	0.008 3
点云聚类	0.042 4
特征提取	0.006 7
数据关联	0.003 1
位姿估计	0.034 4

3.2 消融实验

利用融合语义前后的系统进行消融实验,验证本文算法中语义信息的有效性.结果如表4所示,可以看出,除了未识别静态语义点云的04序列外,其余序

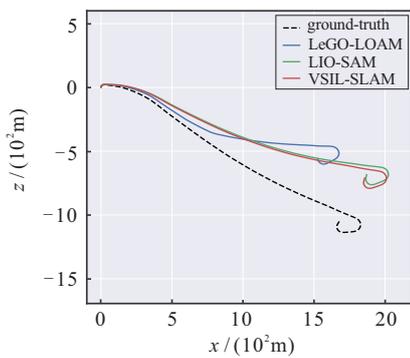
列融合语义后的结果都相对较好,从而表明语义信息的结合可以改善激光SLAM系统的性能.

表4 融合语义前后系统误差对比

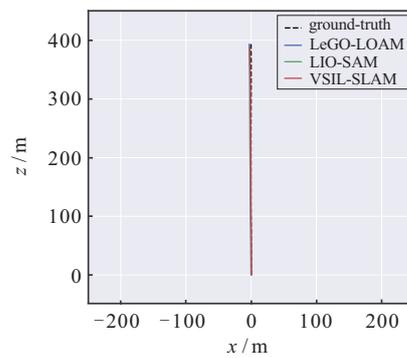
序列	平均位置误差(%)/平均角度误差(deg/m)	
	融合语义前	融合语义后
01	2.912 6/0.007 7	2.790 1/0.007 1
04	1.585 6/0.006 4	1.585 6/0.006 4
09	2.179 4/0.008 2	2.089 3/0.007 9
10	2.945 1/0.009 7	2.175 9/0.008 2

3.3 对比实验

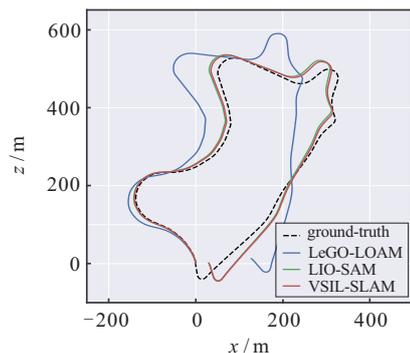
将本文提出的SLAM算法分别与目前开源的4种先进算法进行对比,包括2018年开源的轻量级LEGO-LOAM,以及2020年的扩展版本LIO-SAM.还对比了近两年新开源的激光强度SLAM算法,包括



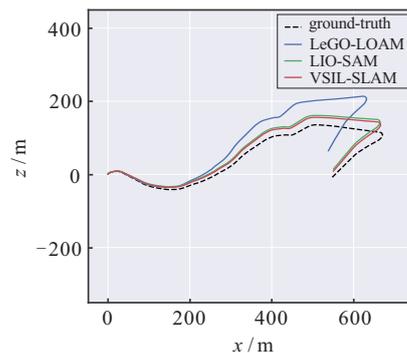
(a) 01序列



(b) 04序列



(c) 09序列



(d) 10序列

图14 与LEGO-LOAM和LIO-SAM输出轨迹对比

2020年基于强度信息,用ISC特征描述符实现闭环的ISC-LOAM^[24],以及2021年使用几何和强度信息实现激光里程计的Intensity-SLAM.

选取01、04、09、10四个序列进行实验,分别与LEGO-LOAM和LIO-SAM算法进行对比,图14给出了EVO评估的轨迹对比曲线.可以看出,本文算法在4个数据集中都有较好的表现,与LEGO-LOAM和LIO-SAM相比,轨迹与真实轨迹更加接近.

表5为对应的误差结果,可以看出,本文提出的算法在简单场景下误差相对较低,除了04序列相比于LIO-SAM差一点,其余表现均有所提升.

表5 与LEGO-LOAM和LIO-SAM轨迹误差对比

序列	平均位置误差(%)/平均角度误差(deg/m)		
	LEGO-LOAM	LIO-SAM	VSIL-SLAM
01	17.243 3/0.011 3	3.697 4/0.007 2	2.790 1/0.007 1
04	1.822 1/0.007 9	1.623 4/0.005 6	1.585 6/0.006 4
09	5.117 0/0.018 4	2.817 8/0.008 9	2.089 3/0.007 9
10	5.282 5/0.018 1	2.580 5/0.008 7	2.175 9/0.008 2

图15为轨迹对比曲线.在05序列中,本文算法相比于其他两种算法估计的轨迹更准确,对齐后与真实轨迹重合度更高,尤其在序列末端位置;在07序列中,本文算法与其他两种算法结果接近.

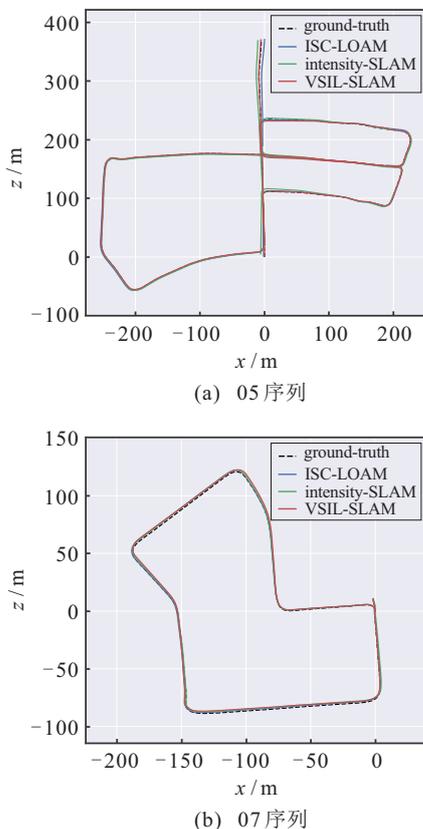


图15 与ISC-LOAM和Intensity-SLAM输出轨迹对比

为了更明显展示各算法精度的情况,同样对其平均位姿误差进行比较,如表6所示.显然,在平移误差上本文算法优于其他两种算法,在旋转误差上与其他两种算法接近.

表6 与ISC-LOAM和Intensity-SLAM轨迹误差对比

序列	平均位置误差(%)/平均角度误差(deg/m)		
	ISC-LOAM	Intensity-SLAM	VSIL-SLAM
05	1.139 3/0.004 7	1.231 2/0.004 9	1.028 6/0.005 1
07	1.377 3/0.007 7	1.387 4/0.007 3	0.969 1/0.007 6

通过上述3组实验,分别对本文提出的VSIL-SLAM算法性能及其运行效果进行了直观表述.实验结果表明,本文算法在大部分数据集上都优于目前开源的几种激光SLAM算法,整体位姿误差相对较低,定位和建图性能表现较好.

4 结论

本文提出了一种融合语义信息的激光SLAM算法,通过对激光点簇进行映射处理,实现与视觉语义信息的关联,得到语义物体.在VFH和网络层特征联合的形状特征基础上,引入拓扑特征,对语义物体点进行特征描述和匹配,构建了基于几何特征和语义物体的前端激光里程计,解决了激光SLAM算法中因点云特征信息不足而造成的匹配精度低、位姿估计误差大等问题.此外,通过加入后端闭环检测和位姿图优化部分,完成了整体SLAM设计.最后,通过大量对比实验验证了本文算法的优越性,达到了改善激光SLAM算法定位精度和建图效果的目的.

参考文献(References)

- [1] Smith R, Self M, Cheeseman P. Estimating uncertain spatial relationships in robotics[C]. Proceedings of 1987 IEEE International Conference on Robotics and Automation. Raleigh, 2003: 850.
- [2] Cadena C, Carlone L, Carrillo H, et al. Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age[J]. IEEE Transactions on Robotics, 2016, 32(6): 1309-1332.
- [3] Lovegrove S, Patron-Perez A, Sibley G. Spline fusion: A continuous-time representation for visual-inertial fusion with application to rolling shutter cameras[C]. Proceedings of the British Machine Vision Conference 2013. Bristol: British Machine Vision Association, 2013: 93.1-93.11.
- [4] Bresson G, Alsayed Z, Yu L, et al. Simultaneous localization and mapping: A survey of current trends in autonomous driving[J]. IEEE Transactions on Intelligent Vehicles, 2017, 2(3): 194-220.

- [5] Zhang J, Singh S. Low-drift and real-time lidar odometry and mapping[J]. *Autonomous Robots*, 2017, 41(2): 401-416.
- [6] Shan T X, Englot B. LeGO-LOAM: Lightweight and ground-optimized lidar odometry and mapping on variable terrain[C]. 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems. Madrid, 2019: 4758-4765.
- [7] Deschaud J E. IMLS-SLAM: Scan-to-model matching based on 3D data[C]. 2018 IEEE International Conference on Robotics and Automation (ICRA). New York: ACM, 2018: 2480-2485.
- [8] Ji X L, Zuo L, Zhang C H, et al. LLOAM: LiDAR odometry and mapping with loop-closure detection based correction[C]. 2019 IEEE International Conference on Mechatronics and Automation. Tianjin, 2019: 2475-2480.
- [9] Dubé R, Dugas D, Stumm E, et al. SegMatch: Segment based place recognition in 3D point clouds[C]. 2017 IEEE International Conference on Robotics and Automation. Singapore, 2017: 5266-5272.
- [10] Shan T X, Englot B, Meyers D, et al. LIO-SAM: Tightly-coupled lidar inertial odometry via smoothing and mapping[C]. 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems. Las Vegas, 2020: 5135-5142.
- [11] Wang H, Wang C, Xie L H. Intensity-SLAM: Intensity assisted localization and mapping for large scale environment[J]. *IEEE Robotics and Automation Letters*, 2021, 6(2): 1715-1721.
- [12] Lianos K N, Schönberger J L, Pollefeys M, et al. VSO: Visual semantic odometry[C]. *Computer Vision—ECCV 2018*. Cham: Springer International Publishing, 2018: 246-263.
- [13] Chen X, Milioto A, Palazzolo E, et al. SuMa: Efficient LiDAR-based semantic SLAM[C]. 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems. Macau, 2019: 4530-4537.
- [14] Milioto A, Vizzo I, Behley J, et al. RangeNet: Fast and accurate LiDAR semantic segmentation[C]. 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems. Macau, 2019: 4213-4220.
- [15] Wu Y M, Zhang Y Z, Zhu D L, et al. EAO-SLAM: Monocular semi-dense object SLAM based on ensemble data association[C]. 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems. Las Vegas, 2020: 4966-4973.
- [16] Wang J W, Rünz M, Agapito L. DSP-SLAM: Object oriented SLAM with deep shape priors[C]. 2021 International Conference on 3D Vision (3DV). London, 2021: 1362-1371.
- [17] Mur-Artal R, Tardós J D. ORB-SLAM2: An open-source SLAM system for monocular, stereo, and RGB-D cameras[J]. *IEEE Transactions on Robotics*, 2017, 33(5): 1255-1262.
- [18] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, 2016: 779-788.
- [19] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: Optimal speed and accuracy of object detection[J/OL]. 2020, arXiv: 2004.10934.
- [20] Ester M, Kriegel H P, Sander J, et al. A density-based algorithm for discovering clusters in large spatial databases with noise[C]. 1996 AAAI Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96). Portland: AAAI, 1996: 226-231.
- [21] Rusu R B, Blodow N, Marton Z C, et al. Aligning point cloud views using persistent feature histograms[C]. 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems. Nice, 2008: 3384-3391.
- [22] 朱德海. 点云库PCL学习教程[M]. 北京: 北京航空航天大学出版社, 2012: 288-291.
(Zhu D H. PCL learning course of point cloud library[M]. Beijing: Beijing University of Aeronautics & Astronautics Press, 2012: 288-291.)
- [23] Chen Y, Medioni G. Object modelling by registration of multiple range images[J]. *Image and Vision Computing*, 1992, 10(3): 145-155.
- [24] Wang H, Wang C, Xie L H. Intensity scan context: Coding intensity and geometry relations for loop closure detection[C]. 2020 IEEE International Conference on Robotics and Automation. Paris, 2020: 2095-2101.

作者简介

佟国峰(1973—), 男, 教授, 博士生导师, 从事SLAM、无人驾驶和无人机、测绘机器人及大尺度场景的高精度三维重建、工业智能感知与深度学习等研究, E-mail: tongguofeng@ise.neu.edu.cn;

杨宇航(1999—), 男, 硕士生, 从事激光SLAM、计算机视觉等研究, E-mail: yangyuhang1616@163.com;

彭浩(1990—), 男, 博士生, 从事点云处理、三维物体检测、计算机视觉等研究, E-mail: apengh@126.com;

孟祥政(1997—), 男, 博士生, 从事激光SLAM、点云处理等研究, E-mail: eminbogen@163.com;

殷旗君(1996—), 男, 硕士生, 从事激光SLAM、计算机视觉等研究, E-mail: yinqijun666@126.com.