

光电工程

Opto-Electronic Engineering

中文核心期刊 中国科技核心期刊
Scopus CSCD

基于图像信息约束的三维激光点云聚类方法

夏金泽, 孙浩铭, 胡盛辉, 梁冬泰

引用本文:

夏金泽, 孙浩铭, 胡盛辉, 等. 基于图像信息约束的三维激光点云聚类方法[J]. 光电工程, 2023, 50(2): 220148.

Xia J Z, Sun H M, Hu S H, et al. 3D laser point cloud clustering method based on image information constraints[J]. *Opto-Electron Eng*, 2023, 50(2): 220148.

<https://doi.org/10.12086/oe.2023.220148>

收稿日期: 2022-06-30; 修改日期: 2022-10-15; 录用日期: 2022-11-28

相关论文

融合空间掩膜预测与点云投影的多目标跟踪

陆康亮, 薛俊, 陶重霖

光电工程 2022, 49(9): 220024 doi: 10.12086/oe.2022.220024

基于点云与图像交叉融合的道路分割方法

张莹, 黄影平, 郭志阳, 张冲

光电工程 2021, 48(12): 210340 doi: 10.12086/oe.2021.210340

车载激光雷达点云数据地面滤波算法综述

黄思源, 刘利民, 董健, 傅雄军

光电工程 2020, 47(12): 190688 doi: 10.12086/oe.2020.190688

多场景下基于快速相机标定的柱面图像拼接方法

傅子秋, 张晓龙, 余成, 梁丹, 梁冬泰

光电工程 2020, 47(4): 190436 doi: 10.12086/oe.2020.190436

更多相关论文见光电期刊集群网站 

 | 光电工程
Opto-Electronic Engineering

<http://cn.ojournal.org/oe>



 OE_Journal



Website

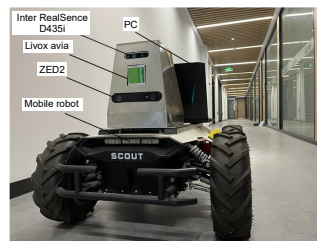


DOI: 10.12086/oe.2023.220148

基于图像信息约束的 三维激光点云聚类方法

夏金泽, 孙浩铭, 胡盛辉, 梁冬泰*

宁波大学机械工程与力学学院, 浙江 宁波 315000



摘要: 针对移动机器人在未知环境感知过程中对三维点云快速聚类分割的需求, 提出一种基于图像信息约束的三维激光点云聚类方法。首先通过点云预处理获取有效的三维环境信息, 采用 RANSAC 方法进行地面点云的分割剔除。其次传感器数据在完成时空配准后引入 YOLOv5 目标检测算法, 对三维点云 K-means 聚类算法进行改进, 利用二维图像目标物的检测框范围约束三维点云, 减少非目标物的干扰; 基于图像检测信息实现点云聚类算法的参数初始化; 采用类内异点剔除法优化聚类结果。最后搭建移动机器人硬件平台, 对箱体进行测试, 实验结果表明, 本文方法的聚类准确率和聚类时间分别为 86.96% 和 23 ms, 可用于移动机器人导航避障、自主搬运等领域。

关键词: 移动机器人; 激光雷达; 目标检测; 点云聚类

中图分类号: TP249

文献标志码: A

夏金泽, 孙浩铭, 胡盛辉, 等. 基于图像信息约束的三维激光点云聚类方法 [J]. 光电工程, 2023, 50(2): 220148

Xia J Z, Sun H M, Hu S H, et al. 3D laser point cloud clustering method based on image information constraints[J]. *Opto-Electron Eng*, 2023, 50(2): 220148

3D laser point cloud clustering method based on image information constraints

Xia Jinze, Sun Haoming, Hu Shenghui, Liang Dongtai*

School of Mechanical Engineering and Mechanics, Ningbo University, Ningbo, Zhejiang 315000, China

Abstract: Aiming at the requirement of fast clustering and segmentation of 3D point clouds for mobile robots in the process of perception of unknown environments, a 3D laser point cloud clustering method based on image information constraints is proposed. Firstly, the effective 3D environment information is obtained through point cloud preprocessing, and the RANSAC method is used to segment and eliminate the ground point cloud. Secondly, the sensor data is introduced into the YOLOv5 target detection algorithm after completing the spatiotemporal registration, and the K-means clustering algorithm of the 3D point cloud is improved. The detection frame range of the 2D image target is used to constrain the 3D point cloud and reduce the interference of non-target objects. The parameter initialization of the point cloud clustering algorithm is realized based on the image detection information. The clustering results are optimized by the intra-class outlier elimination method. Finally, the mobile robot hardware platform is built, and the box is tested. The experimental results show that the clustering accuracy and clustering time of the method in this paper are 86.96% and 23 ms, respectively, which can be used in mobile robot navigation

收稿日期: 2022-06-30; 修回日期: 2022-10-15; 录用日期: 2022-11-28

基金项目: 浙江省公益技术应用研究计划项目 (LGG21E050008); 宁波市科技创新 2025 重大专项 (2019B10100); 宁波市公益性科技计划项目 (2022S004)

*通信作者: 梁冬泰, liangdongtai@nbu.edu.cn。

版权所有©2023 中国科学院光电技术研究所

and obstacle avoidance, autonomous handling, and other fields.

Keywords: moving robot; LiDAR; target detection; point cloud clustering

1 引言

三维激光点云聚类识别是移动机器人感知环境的重要方式之一, 主要目的是为了获取目标物在三维空间中的语义、位置、尺寸以及姿态等信息。近年来, 许多学者不断加大在环境信息提取、分类方面的研究工作, 尤其是多传感器信息融合技术的深入研究, 能够有效弥补单个传感器获取信息的不足, 且有不少的研究成果^[1]。利用传感器数据对周边环境进行感知是实现移动机器人在未知环境下自主作业的首要任务, 如何快速有效从含有地面点云的复杂环境中提取目标物体信息具有十分重要的研究意义。

在传感器中, 相机有较好的环境细节信息提取能力, 适用于目标物分类。激光雷达能够探测目标位置、速度等特征, 具有分辨率高、抗干扰能力强、点云信息丰富的特点。目前, 处理三维激光点云的方法主要分为深度学习方法 and 传统聚类方法, 随着深度学习在图像领域的不断成熟, 逐渐发展了不少基于深度学习模型的方法来处理点云分类问题^[2], 如 Su 等提出了基于多视图的 MVCNN^[3] 卷积神经网络方法, 利用不同姿态下的虚拟相机捕获 2D 视图, 再通过卷积神经网络进行特征提取, 从而实现点云的识别。Maturana 等提出的 Voxnet^[4], 实现了非结构化的点云体素化为网格数据并应用于 3D 卷积神经网络。Qi 等提出了直接处理激光点云的 PointNet^[5]、PointNet++^[6] 网络算法, 有效解决了三维空间中点云无序性、旋转不变性等问题, 能较好地对点云进行特征提取。基于深度学习方法虽然泛化能力好、分割精度高, 但存在前期标注工作量庞大、硬件计算平台成本高等问题^[7]。点云在传统聚类方法^[8]中主要有基于密度的 DBSCAN 聚类算法、基于距离的欧氏聚类算法、基于划分的 K-means 聚类算法等。DBSCAN 算法^[7]能快速地对任意形状的三维点云进行聚类, 但内存资源消耗大。欧氏聚类算法^[9]对数据处理有良好的通用性, 但受距离阈值参数影响较大, 空间中点云存在近密远疏的特性, 易出现过欠聚类的情况。K-means 算法^[10]快速实用有效, 但由于存在 k 值的不确定性以及聚类中心点的随机性, 往往会导致聚类效果不佳。K-means++^[11]聚类是对 K-means 算法的改进, 对随机产生的初始聚类中心点添

加彼此分散的约束, 能提高一定的聚类效果。传统聚类算法虽然处理速度快, 但对聚类参数的依赖度高, 参数设定优劣直接影响该算法能否正确给出点云的语义等信息。

针对上述基于深度学习方法硬件成本高及传统点云聚类方法参数设定依赖度高等问题, 为满足三维激光点云快速聚类分割的需求, 提出了一种基于图像信息约束的三维激光点云聚类方法。首先考虑到激光点云数据量庞大, 采取感兴趣区域选取、体素栅格下采样滤波的点云预处理操作; 为防止地面点云被误检为有效数据, 结合 RANSAC 方法进行地面点云分割剔除; 其次引入基于卷积神经网络的 YOLOv5 目标检测算法, 在完成传感器时空配准后进行图像和点云的数据层融合, 进一步实现点云范围约束及聚类参数初始化, 有效解决传统三维点云 K-means 聚类参数难以选取的问题; 聚类结果采用类内异点剔除法进行优化, 并采用不同颜色的簇标记加以区分。最后搭建移动机器人硬件平台, 在室内环境下对箱体进行测试分析, 验证本文方法的有效性。整体算法流程如图 1 所示, 主要分为图像检测、点云预处理、地面分割、点云聚类四个模块。

2 点云预处理与地面分割

2.1 点云数据预处理

激光雷达传感器获取环境信息时, 对于庞大的点云数据采取感兴趣区域粗提取、体素栅格下采样操作, 实现点云有效数据获取及数据精简。三维目标识别考虑移动机器人正前方的数据区域, 因此本文选取的感兴趣区域以激光雷达为起始点框选出三维长方体 $8\text{ m} \times 4\text{ m} \times 3\text{ m}$ 的区域, 删掉区域外的点云, 初步给定点云处理范围。为加速后续算法处理速度, 采用体素栅格化方法, 其原理是生成若干个三维立方体, 储存在立方体中的点云由重心来替代, 保证点云结构不被破坏的情况下, 降低点云密度, 体素滤波参数根据实际精度与计算速度折中选取, 本文体素边长参数设定为 0.01 m 。考虑到单帧点云数据较为稀疏, 对于离群点及噪点未作处理, 点云数据预处理结果如图 2 所示。

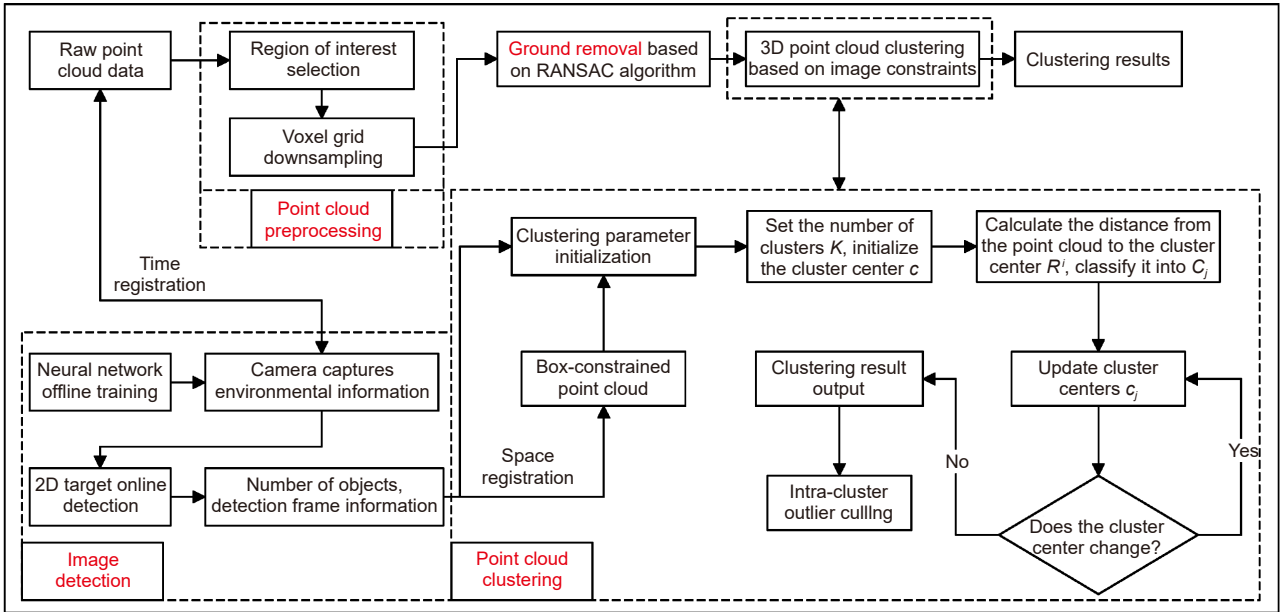


图 1 基于图像信息约束的三维激光点云聚类算法流程图

Fig. 1 Flow chart of 3D laser point cloud clustering algorithm constrained by image information

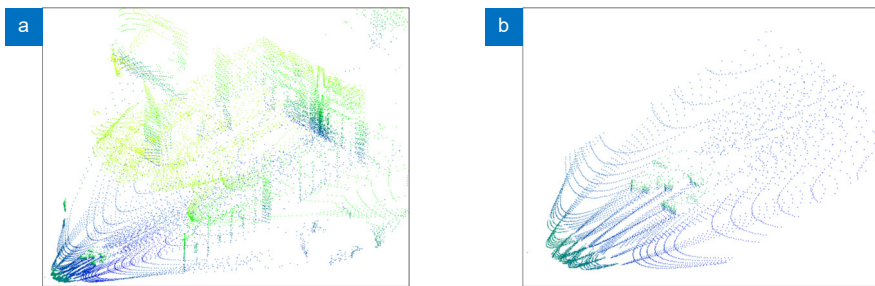


图 2 点云数据预处理。(a) 处理前; (b) 处理后

Fig. 2 Preprocessing of point cloud data. (a) Before processing; (b) After processing

2.2 基于 RANSAC 算法的地面分割

激光雷达点云数据经预处理后仍包含地面点云, 为排除地面点对目标物聚类的干扰, 需对地面点云检测剔除。RANSAC 算法^[12]利用随机采样的方式对三维点云数据拟合估计出平面数学模型参数, 速度快且效果较好。确定好 RANSAC 算法的点云数目 n 、迭代次数 i 以及距离误差阈值 $\Delta\tau$, 对于 RANSAC 拟合的平面模型设为

$$Ax + By + Cz + D = 0. \quad (1)$$

在当前点云数据 P 中随机 3 个点云 $P_1(x_1, y_1, z_1)$ 、 $P_2(x_2, y_2, z_2)$ 、 $P_3(x_3, y_3, z_3)$ 进行平面拟合, 由此可确定 A 、 B 、 C 、 D 四个参数:

$$\begin{cases} A = (y_2 - y_1)(z_3 - z_1) - (z_2 - z_1)(y_3 - y_1) \\ B = (z_2 - z_1)(x_3 - x_1) - (x_2 - x_1)(z_3 - z_1) \\ C = (x_2 - x_1)(y_3 - y_1) - (y_2 - y_1)(x_3 - x_1) \\ D = -(Ax_1 + By_1 + Cz_1) \end{cases} \quad (2)$$

计算空间中任意一点云 $P(x_i, y_i, z_i)$ 至拟合出的平面距离 D_i :

$$D_i = \frac{|Ax_i + By_i + Cz_i + D|}{\sqrt{A^2 + B^2 + C^2}}. \quad (3)$$

对于距离 D_i 小于预设的误差阈值 $\Delta\tau$, 判定为局内点 (满足拟合平面函数的点云), 统计当前有效局内点的数目。依次迭代, 直至满足最大迭代次数, 选取局内点最多的平面模型为最终结果, 实现地面点云的提取, 结果如图 3 所示。

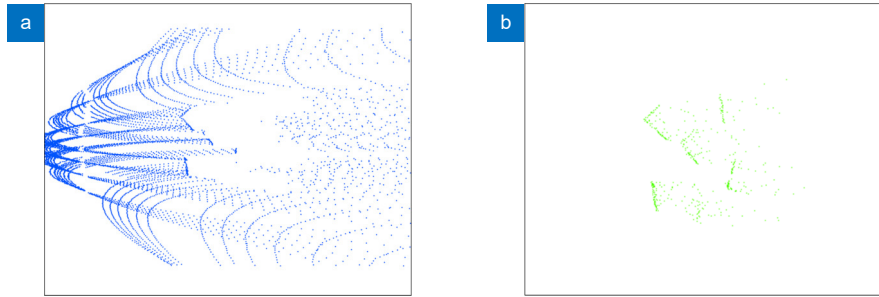


图3 地面分割。(a) 地面点; (b) 非地面点

Fig. 3 Ground segmentation. (a) Groud points; (b) Non-groud points

3 图像信息约束的三维激光点云聚类方法

3.1 传感器数据时空配准

在三维激光点云聚类过程中为添加图像信息约束的条件, 实现点云在图像的像素坐标系中完成数据融合, 必须完成激光雷达与相机的时空配准。首先构建两个不同传感器的空间映射关系, 设 $M(X_L^i, Y_L^i, Z_L^i)$ 是激光雷达坐标系下的某一点云, 则其在相机坐标系下对应坐标为 $M^*(X_C^i, Y_C^i, Z_C^i)$, 图像坐标系下对应坐标为 (x, y) , 像素坐标系下坐标为 (u, v) 。不同传感器坐标系的对应关系如图4所示, 其中 $O_L X_L Y_L Z_L$ 为激光雷达坐标系、 $O_C X_C Y_C Z_C$ 为相机坐标系、 oxy 为图像坐标系、 uv 为像素坐标系。

激光雷达坐标系与像素坐标系的空间映射关系为

$$Z_C^i \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f/d_x & 0 & u_0 & 0 \\ 0 & f/d_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_L^i \\ Y_L^i \\ Z_L^i \\ 1 \end{bmatrix}, \quad (4)$$

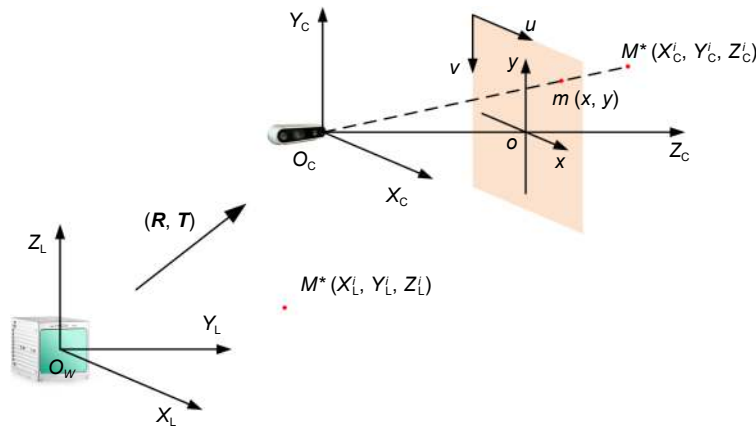


图4 传感器坐标系

Fig. 4 Sensor coordinate system

其中: 旋转矩阵 R 、平移矩阵 T 可由激光雷达与相机联合标定外参得到, f 为相机焦距, u_0 、 v_0 为像素平面原点至像平面中心的距离, 空间中的点云通过式(4)投影至像素坐标系。

相机内参及畸变系数获取参考张式标定法^[13], 激光雷达与相机的外参数确定参考无标定板像素级外参自标定方法^[14], 主要利用空间中存在大量的边缘特征, 对图像和点云共有的边缘特征进行检测与匹配, 添加优化后的图像及点云中边缘特征约束, 构建边缘特征外参计算的残差方程:

$$n_i^T (f(\pi_L^C T(L P_i + {}^L w_i))) - (q_i + {}^L w_i) = 0, \quad (5)$$

其中: ${}^L P_i$ 为激光点云检测到边缘点, n_i 为激光点云边缘线投影至图像上边缘特征的法向量, q_i 是激光特征投影至图像上对应的边缘点, ${}^L w_i$ 、 ${}^L w_i$ 分别表示服从零均值分布激光点云与图像的噪声, ${}^L T$ 是非线性方程迭代求解的对象。

在传感器数据融合时, 各个传感器有着不同测量精度和周期, 为保证测量结果反映同一状态下的客观世界, 必须保证多传感器的时间配准。本文采用时间

最近邻匹配方法, 以激光雷达时间作为基准, 匹配对应时刻下时间间隔最小的图像数据, 从而完成激光雷达与相机的时间戳对齐。

3.2 二维图像检测框尺寸优化

二维图像目标检测对于三维目标检测任务而言, 缺少深度信息, 且单一的纹理信息无法直接确定目标在三维空间中的位置、姿态等信息, 但能输出目标物的类别、数量等信息。YOLOv5 算法^[15] 是一种基于卷积神经网络的单阶段目标检测算法, 速度快、精度高, 相较于 YOLO 之前系列的检测模型, 性能有较大的提升。YOLOv5 整体网络结构如图 5 所示, 由输入端、主干网络、颈部、预测端四部分组成。输入端主要对 608×608 图像进行 Mosaic 数据增强、自适应锚框计算、自适应图片缩放的操作。主干网络使用 CSPDarkNet53 网络来提取图像丰富的特征, 主要由 Focus 结构、CSP 结构和 SPP 结构构成, Focus 结构通过对图像的切片操作加深图像的特征维度, CSP 结构可减少特征传输过程中的信息丢失和计算量, SPP 结构可以扩展网络可处理的图片类型。颈部采用

FPN+PAN 结构, 从不同的主干层对不同的检测层进行参数聚合。预测端部分包括预测边界框、损失函数计算和非极大值抑制, 能够有效消除多余的检测框, 筛选出正确的目标物位置。

目标检测算法输出的二维检测框信息为三维点云提供约束条件, 二维目标检测框的边界范围将直接影响后续的点云分割, 过小的检测框会导致提取的目标物点云缺失, 过大则会引入干扰点云, 因此对检测框作适当扩张处理, 避免丢失目标物点云。YOLOv5 边框预测公式为

$$\begin{cases} b_x = \sigma(t_x) + c_x \\ b_y = \sigma(t_y) + c_y \\ b_w = (1 + E_f)p_w e^{t_w} \\ b_h = (1 + E_f)p_h e^{t_h} \end{cases}, \quad (6)$$

其中: b_x 、 b_y 为预测框中心点坐标值, b_w 、 b_h 为预测框宽高值, E_f 为扩张因子, 设定为 0.05, $\sigma(t_x)$ 、 $\sigma(t_y)$ 为归一化后预测的坐标偏移值, t_w 、 t_h 为尺度缩放值。

完成预测框尺寸优化后, 对于检测框外的非目标物点云, 如图 6 所示, 通过检测框约束点云方法, 对

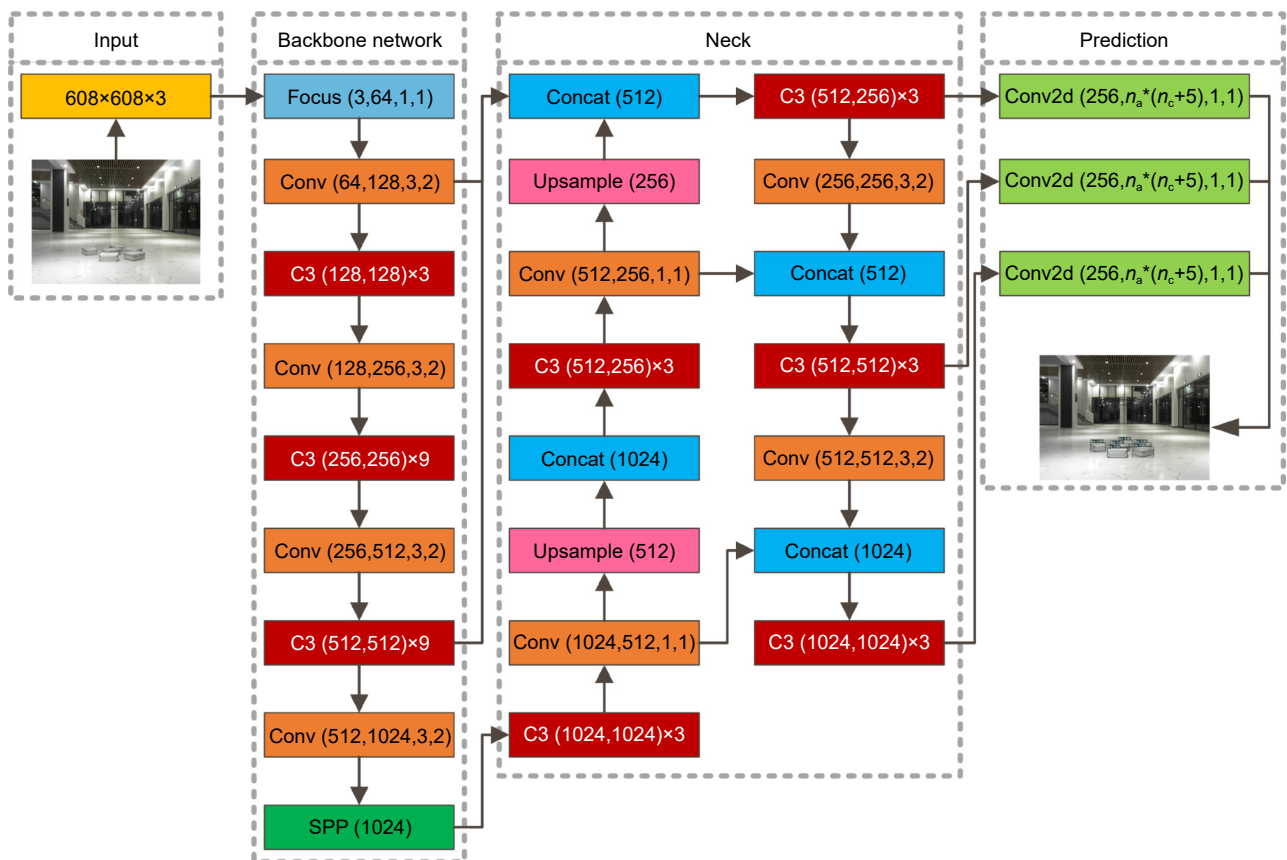


图 5 YOLOv5 网络结构图
Fig. 5 YOLOv5 network structure diagram

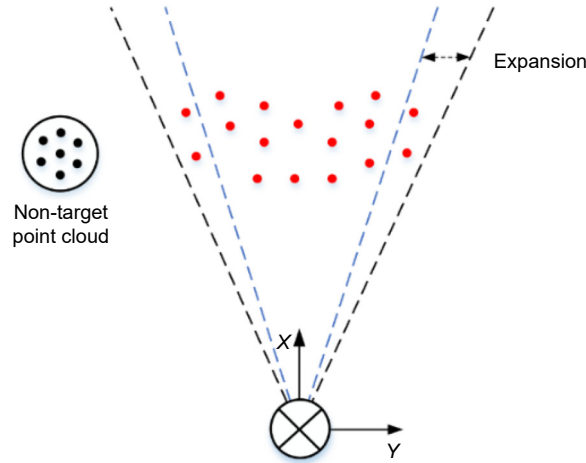


图 6 检测框约束点云示意图

Fig. 6 Schematic diagram of detection frame constraint point cloud

于每一点云在像素坐标系下判断是否位于检测框内, 对于检测框外的点云剔除, 实现对非目标物点云滤除。

3.3 三维点云 K-means 聚类改进

K-means 算法^[10]属于较为经典的非监督学习的聚类算法之一, 以欧几里得距离作为衡量数据相似度的指标, 聚类速度快, 常用于三维点云的聚类分割。核心思想是在未标记的点云数据中随机选取 K 个不重叠的点云对象作为迭代求解的初始聚类中心, 计算每个点云与聚类中心的距离, 根据就近原则分配给 K 个聚类中心, 每完成一次聚类过程, 将会更新聚类中心, 循环往复, 直至聚类中心值收敛不变化。

给定点云数据集 $P = \{p_1, p_2, \dots, p_n\}$, 初始化聚类中心 $c = \{c_1, c_2, \dots, c_k\}$, 计算点云至 K 个聚类中心的欧几里得距离 R^i :

$$R^i = \arg \min \|p^i - c_j\|^2, \quad (7)$$

其中: $p^i \in P$ 表示点云集合中第 i 个点云, $c_j \in c$ 表示第 j 个聚类中心。

将点云 p^i 划分至最近的 C_j 中, 聚类中心更新公式为

$$c_j = \frac{\sum_{x_i \in C_j} x_i}{|C_j|}, \quad (8)$$

其中: C_j 表示第 j 个类簇, $|C_j|$ 表示第 j 类簇中点云数量, x_i 表示第 j 个类簇中第 i 个点云对象, $1 \leq i \leq |C_j|$ 。

三维点云 K-means 聚类存在难以选定最佳聚类个数 K 、初始聚类中心敏感的问题, 取值优劣对聚类效果起着决定性作用, 对于每次迭代, 参数不同往往会导致不同的聚类结果。因此添加图像信息约束来解决

上述问题, 利用 YOLOv5 在检测过程中生成的目标物检测框数量来确定点云聚类算法的 K 值。对于初始聚类中心的设定, 设置缩放因子 S_f 对二维检测框采取缩框处理, 为保证框内点云属于所聚目标物, S_f 设置为 25%, 统计框内有效点云的质心 P_c , 将质心作为初始聚类中心, 完成聚类参数初始化后开展聚类迭代。质心在选取过程中, 点云尽可能靠近二维检测框中心, 同时与其最近点云在三维空间中的彼此距离阈值为 10 cm, 确保质心尽量位于实际目标物内部且分布均匀, 质心选取效果如图 7 所示。

对于聚类结果, 考虑到图像约束的检测框, 深度信息未知, 检测框在三维空间中以视锥的方式获取点云, 检测框框选的点云数目多于实际目标物点云数目, 后方其他物体的点云数据有几率被保留, 对聚类结果造成干扰。设定异点判定阈值, 判断当前点云与本类簇初始聚类中心 P_c 的空间欧氏距离是否大于设定阈值, 阈值大小由实际目标物尺寸决定, 本文阈值设定为 50 cm。通过此方法实现类簇内异点剔除, 优化聚类结果。

4 实验结果与分析

4.1 实验硬件平台搭建

本文实验所采用的移动机器人硬件平台如图 8 所示, 由 Scout 四轮差动移动机器人、Livox-Avia 激光雷达、Intel RealSense D435i 相机以及 PC 组成, 其中 PC 的 CPU 为 Intel i5-7300H, 显卡为 NVIDIA GeForce GTX 1050。数据处理端为 Ubuntu 18.04 的操作系统, 各个传感器数据基于 ROS 平台实现与 PC 端的通讯。

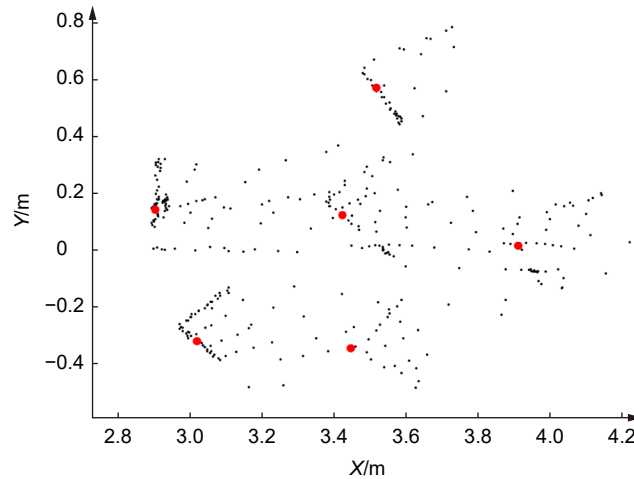


图 7 聚类质心选取图

Fig. 7 Cluster centroid selection graph

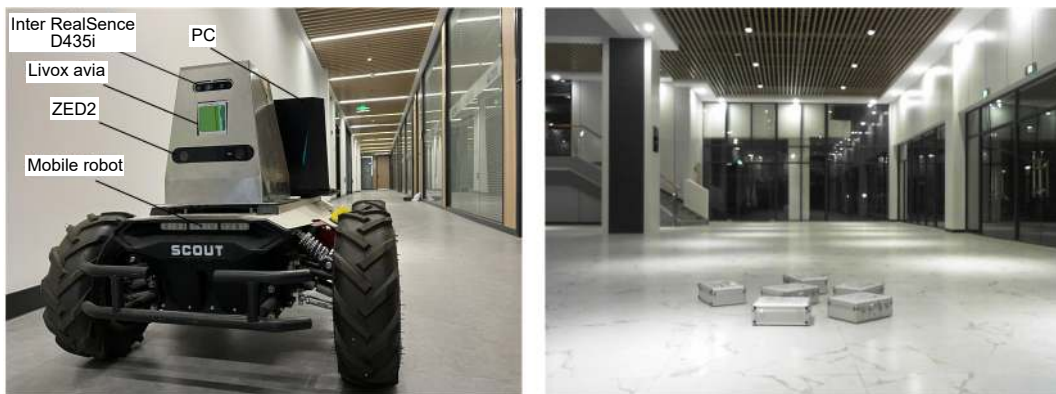


图 8 实验硬件平台及实验场景

Fig. 8 Experimental hardware platform and experimental scene

硬件系统中相机图像的输出频率为 30 Hz, 激光雷达点云的输出频率为 10 Hz, 以激光雷达点云的时间为基准, 完成传感器的时间戳对齐, 如图 9 所示。

4.2 内外参数标定实验

根据搭建的移动机器人硬件平台, 由于激光雷达与相机刚性连接, 相对位姿保持不变, 本文采用张氏标定法求解相机的内参矩阵 K 及对应畸变系数 D , 利用无标定板的像素级外参自标定法求解外参矩阵 T ,

表 1 中内参矩阵参数依次为 f_x 、 f_y 、 c_x 、 c_y , 畸变系数参数依次为 k_1 、 k_2 、 p_1 、 p_2 , 表 2 中外参矩阵依次为平移项 (x 、 y 、 z (mm)) 及旋转项 (roll、pitch、yaw(rad))。

为了验证标定系数的准确性, 将标定前与标定后的三维激光点云分别投影至对应图像信息中, 实现 RGB 信息融合, 标定效果如图 10 所示, 其中图 10(a) 为未标定状态的点云与像素投影效果, 图 10(b) 为标定后的点云与像素投影效果, 此参数基本实现了图像

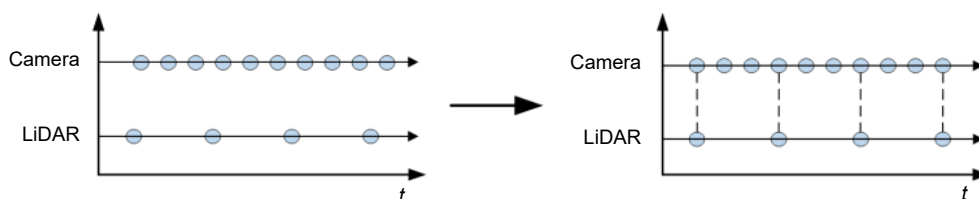


图 9 时间戳对齐

Fig. 9 Align timestamp

表 1 内参标定结果

Table 1 Calibration results of internal parameters

	f_x	f_y	c_x	c_y	k_1	k_2	ρ_1	ρ_2
K	657.58	660.12	296.12	246.35	—	—	—	—
D	—	—	—	—	0.238809	-0.643802	0.001786	-0.024125

表 2 外参标定结果

Table 2 Calibration results of external parameters

	x/mm	y/mm	z/mm	Roll/rad	Pitch/rad	Yaw/rad
T	59.94	52.76	-14.46	-1.540	0.031	-1.581

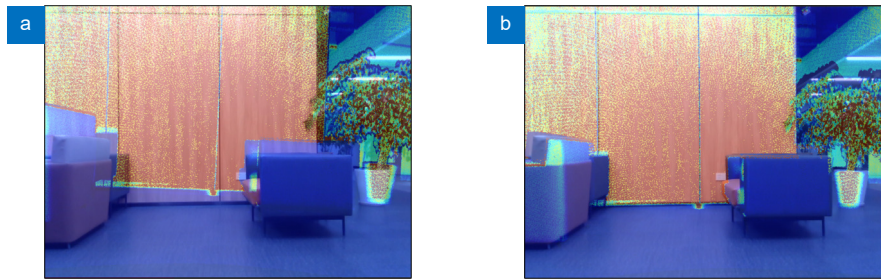


图 10 激光雷达与相机标定。(a) 标定前; (b) 标定后

Fig. 10 LiDAR and camera calibration. (a) Before calibration; (b) After calibration

点云的正确匹配, 可用于传感器的数据融合。

4.3 三维点云聚类实验

为验证本文方法的有效性, 在室内环境下, 实验设计场景如图 8 所示, 将箱体作为目标物检测对象, 主要设计了两个试验: 目标物分布疏密程度对聚类算法鲁棒性的影响、对 50 帧数据随机分布 368 个目标物进行聚类分割测试, 将本文方法与 Euclidean Clustering、DBSCAN、K-means、K-means++等算法进行对比分析。

目标物分布疏密程度对聚类算法有较大的影响。位置分布紧密的目标物容易被聚类成同一物体, 导致误检情况的发生。设置五组目标物不同的分布间距, 测试本文算法鲁棒性, 以 ω 表示当前帧是否聚类完全, α 表示单个目标物聚类准确度:

$$\alpha = \frac{T_p}{T_A}, \quad (9)$$

其中: T_p 表示正确分割点云数, T_A 表示实际目标物点云数, 设 α 大于 80% 表示单个目标物聚类成功, 若当前数据中所有目标物 α 均满足条件, 则 ω 聚类完全。

考虑到 K-means、K-means++算法聚类结果存在随机性, 统计完全聚类成功率 η 来测试算法性能:

$$\eta = \frac{I_S}{I_A}, \quad (10)$$

其中: I_S 表示聚类成功的次数, I_A 为测试次数设置为 50 次。

如表 3 结果所示, 本文方法在不同目标物分布间距下, 均能获得较好的聚类效果, 由于聚类参数给定相对较优且稳定, 不存在结果随机性问题, 聚类效果稳定。K-means 与 K-means++算法有时能聚类完全, 但存在随机性, K-means++完全聚类成功率优于 K-means 算法。Euclidean Clustering 算法及 DBSCAN 算法在分布间距过小会出现聚类不完全现象, 但 DBSCAN 算法效果略好于 Euclidean Clustering 算法。图 11 实际测试中分布间距为 10 cm 的多种聚类算法结果展示, DBSCAN 与 Euclidean Clustering 算法会滤除点云数量低于 10 的类簇, 会导致部分点云缺失, DBSCAN 及本文算法能聚类完全, 而其余传统算法也存在一定的过欠聚类情况。

表 4 是对随机分布 368 个目标物进行聚类分割测试数据, 每帧测试的数据集中存在 5~10 个目标物, 主要对比平均耗时、聚类迭代次数及聚类准确率。聚类准确率 C_A 为

$$C_A = \frac{N_C}{N_A}, \quad (11)$$

其中: N_C 是正确分割目标物数量, N_A 是实际目标物数量。

表 3 分布间距对算法影响

Table 3 Affects of distribution spacing on the algorithm

Distribution spacing/cm	My-method		K-means		K-means++		Euclidean Clustering		DBSCAN	
	ω	η	ω	η	ω	η	ω	η	ω	η
2	✓	—	✓	0.44	✓	0.72	✗	—	✗	—
5	✓	—	✓	0.46	✓	0.70	✗	—	✗	—
10	✓	—	✓	0.72	✓	0.88	✗	—	✓	—
15	✓	—	✓	0.62	✓	0.74	✓	—	✓	—
20	✓	—	✓	0.54	✓	0.70	✓	—	✓	—

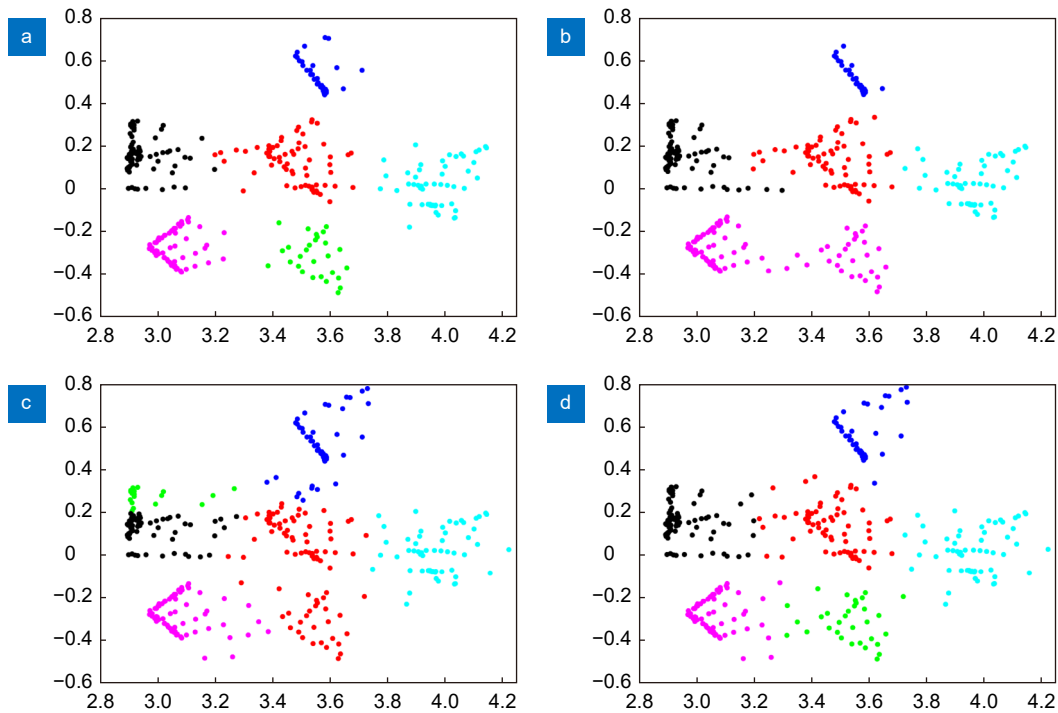


图 11 多种算法聚类结果。(a) DBSCAN; (b) Euclidean Clustering; (c) K-means++; (d) My-method
Fig. 11 Clustering results of multiple algorithms. (a) DBSCAN; (b) Euclidean Clustering; (c) K-means++; (d) My-method

表 4 多种算法性能对比

Table 4 Performance comparison of multiple algorithms

Algorithm	Number of correct divisions/number	Clustering accuracy/%	Average time spent/ms	Average number of iterations/number
DBSCAN	258	70.11	3.625	—
Euclidean Clustering	262	71.20	2.517	—
K-means	210	57.07	1.951	12
K-means++	222	60.33	3.373	10
My-method	320	86.96	1.106	6

由表 4 记录数据分析可知, 平均耗时指算法聚类迭代时间, K-means 算法平均消耗时间较短, 但聚类准确率最低, 且平均迭代次数达 12 次。K-means++ 算法优化了初始聚类参数, 聚类准确度得到提升且平均迭代次数降低, 但聚类效果随机性仍存在。Euclidean Clustering 及 DBSCAN 算法, 过欠聚类情

况时有发生, 聚类准确率分别为 71.20%、70.11%, 但 Euclidean Clustering 算法平均消耗时间快于 DBSCAN 算法。本文方法正确分割数目多、平均耗时最少, 聚类平均迭代次数达 6 次, 聚类准确率达 86.96%。如图 12 所示为本文方法各模块 50 帧的运行时间, 其中图像检测、点云预处理、地面分割、点云

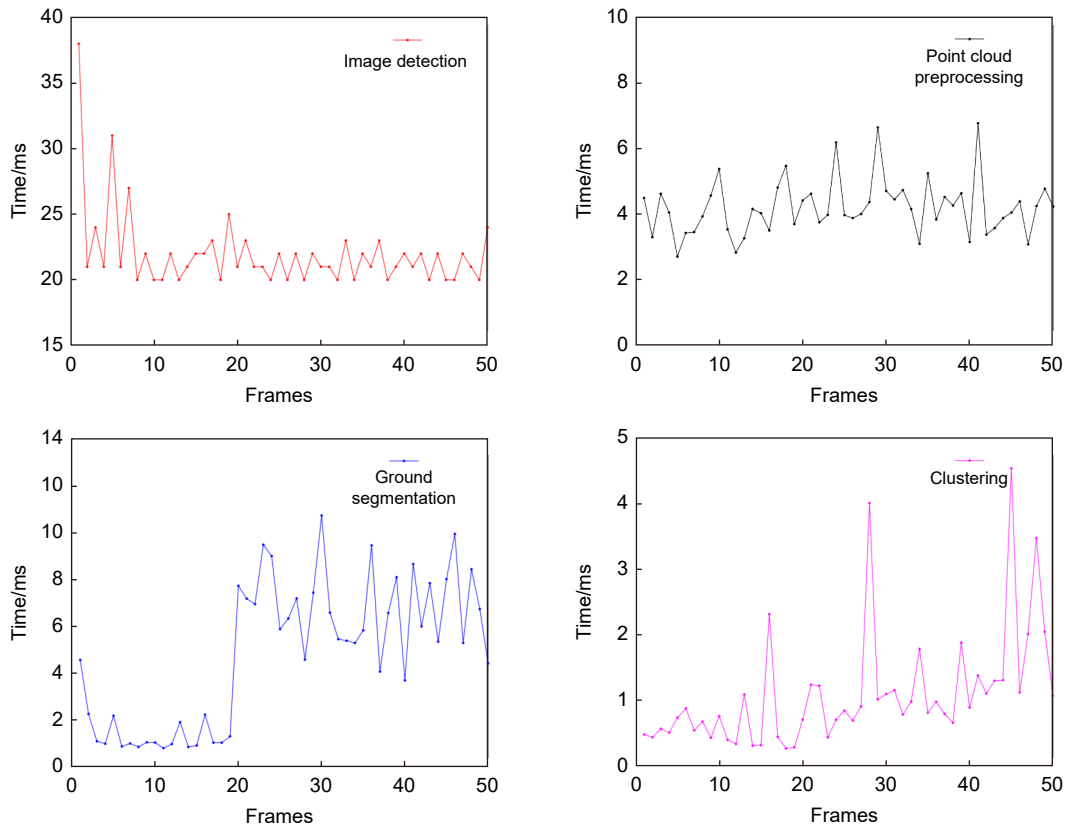


图 12 本文方法各模块运行时间

Fig. 12 Running time of each module of this method

聚类的平均耗时分别为 21.96 ms、4.17 ms、4.83 ms、1.11 ms, 本文方法整体平均总耗时约为 23 ms, 能够完成移动机器人对三维点云的实时聚类分割。

5 结论

本文提出一种基于图像信息约束的三维激光点云聚类方法。经点云预处理、基于 RANSAC 的地面点云剔除, 获得有效点云数据。引入 YOLOv5 目标检测算法, 利用图像二维目标检测框约束三维点云范围, 进一步排除其他物体对于目标物的干扰; 基于图像检测信息完成聚类参数初始化, 有效改善因聚类参数难以设定而导致聚类效果不佳的问题; 聚类结果采用类内异点剔除进行优化。本文对箱体进行测试, 实验结果表明, 本文方法的聚类准确率和聚类时间分别为 86.96% 和 23 ms, 对于较为密集的目标物能较好地进行聚类分割。本文提出的移动机器人三维点云聚类识别方法, 检测效果好、准确率高且具有良好的鲁棒性, 满足移动机器人在未知环境感知过程中对三维点云快速聚类分割的需求, 可用于移动机器人导航避障、自主搬运等领域。

实验中发现, 单帧点云数据视野有限, 实际还会存在物体相互遮挡的情况, 无法恢复识别物体的整体结构, 下一步将研究基于非重复性扫描激光雷达的高精度三维场景重建 (Slam), 从而获取更为丰富的物体三维点云信息, 进一步提升物体识别的准确率及鲁棒性。

参考文献

- [1] Ji C Y, Zhou J. Current situation of navigation technologies for agricultural machinery[J]. *Trans Chin Soc Agric Mach*, 2014, **45**(9): 44-54.
姬长英, 周俊. 农业机械导航技术发展分析[J]. *农业机械学报*, 2014, **45**(9): 44-54.
- [2] Zhao J Q, Zhou Y, He X, et al. Research progress analysis of point cloud segmentation based on deep learning[J]. *J Electron Inf Technol*, 2022, **44**(12): 4426-4440.
赵佳琦, 周勇, 何欣, 等. 基于深度学习的点云分割研究进展分析[J]. *电子与信息学报*, 2022, **44**(12): 4426-4440.
- [3] Su H, Maji S, Kalogerakis E, et al. Multi-view convolutional neural networks for 3D shape recognition[C]//2015 IEEE International Conference on Computer Vision, 2015: 945-953.
<https://doi.org/10.1109/ICCV.2015.114>.
- [4] Maturana D, Scherer S. VoxNet: a 3D convolutional neural network for real-time object recognition[C]//2015 IEEE/RSJ International Conference on Intelligent Robots and Systems

- (IROS), 2015: 922–928. <https://doi.org/10.1109/IROS.2015.7353481>.
- [5] Qi C R, Su H, Mo K C, et al. PointNet: deep learning on point sets for 3D classification and segmentation[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, 2017: 77–85. <https://doi.org/10.1109/CVPR.2017.16>.
- [6] Qi C R, Yi L, Su H, et al. PointNet++: deep hierarchical feature learning on point sets in a metric space[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017: 5105–5114.
- [7] Jin L S, He Y, Wang H H, et al. Point cloud segmentation algorithm based on adaptive threshold DBSCAN for roadside LiDAR[J]. *Automot Eng*, 2022, 44(7): 987–996.
金立生, 贺阳, 王欢欢, 等. 基于自适应阈值DBSCAN的路侧点云分割算法[J]. *汽车工程*, 2022, 44(7): 987–996.
- [8] Wang Z Y, Li Q Q, Zhang Z Y, et al. Research progress of unmanned vehicle point cloud clustering algorithm[J]. *World Sci-Tech R & D*, 2021, 43(3): 274–285.
王子洋, 李琼琼, 张子蕴, 等. 应用于无人驾驶车辆的点云聚类算法研究进展[J]. *世界科技研究与发展*, 2021, 43(3): 274–285.
- [9] Wei Y F, Liang D T, Liang D, et al. Visual identification and location algorithm for robot based on the multimodal information[J]. *Opto-Electron Eng*, 2018, 45(2): 170650.
魏玉锋, 梁冬泰, 梁丹, 等. 基于多模态信息的机器人视觉识别与定位研究[J]. *光电工程*, 2018, 45(2): 170650.
- [10] Yang J C, Zhao C. Survey on K-means clustering algorithm[J]. *Comput Eng Appl*, 2019, 55(23): 7–14, 63.
杨俊闯, 赵超. K-Means聚类算法研究综述[J]. *计算机工程与应用*, 2019, 55(23): 7–14, 63.
- [11] Li P T, Zhao Q Z, Tian W Z, et al. Point cloud simplification method combining K-means++ clustering with UAV LiDAR point cloud normal vectors[J]. *Remote Sens Land Resour*, 2020, 32(2): 103–110.
李沛婷, 赵庆展, 田文忠, 等. 结合无人机载LiDAR点云法向量的K-means++聚类精简[J]. *国土资源遥感*, 2020, 32(2): 103–110.
- [12] Liu Y K, Li Y Q, Liu H Y, et al. An improved RANSAC algorithm for point cloud segmentation of complex building roofs[J]. *J Geo-Inf Sci*, 2021, 23(8): 1497–1507.
刘亚坤, 李永强, 刘会云, 等. 基于改进RANSAC算法的复杂建筑物屋顶点云分割[J]. *地球信息科学学报*, 2021, 23(8): 1497–1507.
- [13] Pan J Y, Zhang S Q, Su Z L, et al. Measuring three-dimensional deformation of underwater propellers based on digital image correlation[J]. *Acta Opt Sin*, 2021, 41(12): 1212001.
潘济宇, 张水强, 苏志龙, 等. 基于数字图像相关的水下螺旋桨三维变形测量[J]. *光学学报*, 2021, 41(12): 1212001.
- [14] Yuan C J, Liu X Y, Hong X P, et al. Pixel-level extrinsic self calibration of high resolution LiDAR and camera in targetless environments[J]. *IEEE Robot Autom Lett*, 2021, 6(4): 7517–7524.
- [15] Li D G, Wang L, Li F. Review of typical object detection algorithms for deep learning[J]. *Comput Eng Appl*, 2021, 57(8): 10–25.
许德刚, 王露, 李凡. 深度学习的典型目标检测算法研究综述[J]. *计算机工程与应用*, 2021, 57(8): 10–25.

作者简介



夏金泽, 男, 硕士生, 主要从事移动机器人智能感知技术方面的研究。

E-mail: 1322214816@qq.com

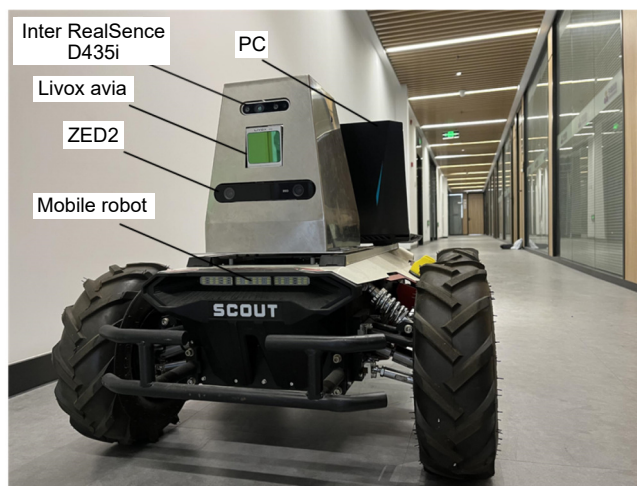


【通信作者】梁冬泰, 男, 博士, 宁波大学副教授, 硕士生导师, 主要从事机器人视觉感知及操控系统、工业视觉检测、机器人系统方面的研究。

E-mail: liangdongtai@nbu.edu.cn

3D laser point cloud clustering method based on image information constraints

Xia Jinze, Sun Haoming, Hu Shenghui, Liang Dongtai*



Experimental hardware platform

Overview: 3D laser point cloud clustering recognition is one of the important ways for mobile robots to perceive the environment. The main purpose is to obtain the semantics, position, size and attitude of the target in three-dimensional space. How to use sensor data to quickly and effectively extract target object information from complex environments containing ground point clouds is of great research significance.

This paper proposes a 3D laser point cloud clustering method based on image information constraints to meet the needs of fast clustering and segmentation of 3D point clouds for mobile robots in the process of perception of unknown environments. First of all, considering the huge amount of laser point cloud data, point cloud preprocessing operations of the region of interest selection and voxel grid down sampling filtering are adopted. To prevent the ground point cloud from being falsely detected as valid data, the ground point cloud is segmented and culled by the RANSAC method. In order to construct the spatial mapping relationship of the lidar and camera, the internal and external parameters are obtained by Zhang's calibration method and the pixel-level external parameter self-calibration method without a calibration board, and the calculated parameters enable spatial synchronization of point clouds and images. The time nearest neighbor matching method is adopted to complete the multi-sensor time registration based on the lidar timestamp. Secondly, the YOLOv5 target detection algorithm is introduced to improve the K-means clustering algorithm of the 3D point cloud. The detection frame range of the 2D image target is used to constrain the 3D point cloud to reduce the interference of non-target objects. The parameter initialization of the point cloud clustering algorithm is realized based on the image detection information, which effectively solves the problem of poor clustering effect caused by the difficulty in determining the initial parameter setting of the traditional 3D point cloud K-means clustering, and then uses the intra-class outlier elimination method to optimize the clustering results. Finally, we build a mobile robot hardware platform and test the box that are compared with DBSCAN, Euclidean Clustering, K-means, and K-means++ algorithms. In the case of densely arranged boxes, it has better detection robustness.

After testing with 50 frames of random data, the experimental results show that the clustering accuracy and clustering time of this method are 86.96% and 23 ms, respectively, which are better than other algorithms, and can be used in mobile robot navigation and obstacle avoidance, autonomous handling, and other fields.

Xia J Z, Sun H M, Hu S H, et al. 3D laser point cloud clustering method based on image information constraints[J]. *Opto-Electron Eng*, 2023, 50(2): 220148; DOI: [10.12086/oe.2023.220148](https://doi.org/10.12086/oe.2023.220148)

Foundation item: Public Welfare Project of Zhejiang Natural Science Foundation (LGG21E050008), the Ningbo 2025 Science and Technology Innovation Major Project (2019B10100), and Public Welfare Project of Ningbo Science and Technology Foundation (2022S004)
School of Mechanical Engineering and Mechanics, Ningbo University, Ningbo, Zhejiang 315000, China

* E-mail: liangdongtai@nbu.edu.cn