

DOI: 10.19650/j.cnki.cjsi.J2514128

基于视觉语义约束的激光雷达大目标尺寸测量方法 *

何贊澤^{1,2}, 郭 猜¹, 郭隆強¹, 邓堡元¹, 王耀南¹

(1. 湖南大学电气与信息工程学院 长沙 410082; 2. 湖南大学深圳研究院 深圳 518000)

摘要:针对单帧视角数据难以反映大尺寸目标的整体轮廓从而导致尺寸测量受限的问题,以汽车为大目标测量研究对象提出了基于视觉语义约束的激光雷达大目标尺寸测量方法。该方法首先通过联合标定和时间戳最近邻匹配实现激光雷达、相机和惯性测量单元数据的时空同步;然后移动小车获取整个目标3个维度上的信息,利用同时定位与建图技术实现被测目标轮廓重建,并且在该模块通过基于地面的残差优化和回环检测框架提升算法精度;之后对点云去噪后使用地面分割算法分割地面点与非地面点,并采取直通滤波保证分割效果,与此同时使用目标检测算法获取图像中目标的类别和位置信息;其次通过自适应阈值点云聚类方法,将不同点云簇中心进行视觉投影,根据视觉目标检测结果定位目标对应点云;最后,设计了一种轮廓拟合算法完成目标点云的轮廓拟合,再利用三维框拟合算法实现目标的尺寸计算。实验结果表明,对于汽车这类尺寸较大的物体,提出的方法在有较多车辆的停车场,车辆长度的最大误差<1.97%、平均误差<0.82%,宽度的最大误差<3.26%、平均误差<2.08%,高度的最大误差<3.99%、平均误差<1.99%,具有良好的工程应用前景。

关键词: 尺寸测量;语义约束;轮廓重建;目标检测

中图分类号: TH721 TH701 文献标识码: A 国家标准学科分类代码: 420.40

A large target size measurement method of LiDAR based on visual semantic constraints

He Yunze^{1,2}, Guo Cai¹, Guo Longqiang¹, Deng Baoyuan¹, Wang Yaonan¹

(1. College of Electrical and Information Engineering, Hunan University, Changsha 410082, China;
2. Shenzhen Research Institute, Hunan University, Shenzhen 518000, China)

Abstract: To address the issue that single-frame perspective data struggle to depict the complete contour of large-scale targets, thereby limiting size measurement, this article proposed a large-target size measurement method of LiDAR based on visual semantic constraints, with automobiles as the research object for large-target measurement. Firstly, this method achieves spatio-temporal synchronization of LiDAR, camera, and inertial measurement unit data through joint calibration and timestamp nearest-neighbor matching. Subsequently, a mobile cart is used to acquire information about the entire target in three dimensions. The simultaneous localization and mapping technology is employed to reconstruct the contour of the measured target. In this module, the algorithm's accuracy is enhanced through a ground-based residual optimization and loop closure detection framework. After denoising the point cloud, a ground segmentation algorithm is used to separate ground points from non-ground points, and a pass-through filter is applied to ensure the segmentation effect. Meanwhile, a target detection algorithm is utilized to obtain the category and position information of the target in the image. Next, through an adaptive threshold point cloud clustering method, the centers of different point cloud clusters are visually projected, and the point cloud corresponding to the target is located according to the visual target detection results. Finally, a contour fitting algorithm is designed to complete the contour fitting of the target point cloud. Then, a three-dimensional box fitting algorithm is used to calculate the target's size. Experimental results show that for large-sized objects such as automobiles. In a parking lot with a large number of vehicles, the proposed method yields a maximum error and an average error of less than 1.97% and 0.82% respectively for vehicle length, less than 3.26% and 2.08% respectively for width, and less than 3.99% and 1.99% respectively for height, demonstrating promising prospects for engineering applications.

Keywords: dimensional measurement; semantic constraints; contour reconstruction; object detection

0 引言

随着激光雷达技术的不断发展和以及成本的降低,同时由于激光雷达分辨率高,能够获取目标的形状、大小、位置等信息,激光雷达在目标检测、自动驾驶和同时定位与建图(simultaneous localization and mapping, SLAM)等领域有着广泛的应用^[1-2]。而视觉传感器可以获取目标丰富的纹理和颜色信息,融合激光雷达数据和视觉数据,可以可靠感知周围环境信息^[3-4],对于推动各领域发展有着重要意义。

随着物体尺寸测量成为智能决策的关键数据源,激光雷达与视觉传感器构成非接触式测量的主流方案。相较可能损伤目标且适应性受限的接触式测量,非接触式方法凭借无损特性及高精度、高效率优势,逐步成为工业检测与自动驾驶等场景的首选方案。

在非接触式测量方法中,基于机器视觉方法、基于激光雷达方法和融合机器视觉与激光雷达方法是目前主流的方法。郭隆强等^[5]提出了基于视觉修正的激光雷达体积测量方法,通过视觉检测对目标点云进行修正,实现目标的尺寸测量,解决了点云的噪声和扰动等问题,但是对于较大尺寸目标尺寸测量效果不佳。杨亮等^[6]基于双目视觉设计了废料尺寸测量系统,通过提取废料的角点坐标,从而得到物料的尺寸。针对固定场景设计测量系统,实际应用时受到限制。吴文涛等^[7]提出了融合相机与激光雷达的目标检测与尺寸测量方法,融合激光雷达和相机数据对物体的二维边界框进行尺寸测量,然后对三维目标检测到的尺寸信息进行补充。但是对于较远或较小的物体尺寸测量精度较差。综上,视觉与激光雷达测量方案均依赖多传感器协同与固定安装配置,存在自主性瓶颈;视觉法虽具成本优势但易受光照干扰;激光雷达法则需多传感器静态部署实现数据融合。

针对自动驾驶等领域中大尺寸非规则物体(如车辆)的测量难题,现有三维检测算法如基于点一体素区域的卷积神经网络(point voxel region based convolutional neural, PV-RCNN)^[8]、更快的基于区域的卷积神经网络(faster region based convolutional neural network, FasterRCNN)^[9]、和点柱网络(point pillars, PointPillars)^[10]等虽能适配道路场景,但受限于单帧数据难以完整表征目标轮廓,导致尺寸测量精度受限。

针对于此,本研究提出了一种基于视觉语义约束的激光雷达尺寸测量方法,以无人车为载体,实现对汽车的尺寸测量。主要工作为首先对激光雷达、相机和惯性测量单元(inertial measurement unit, IMU)数据进行时空对齐;之后利用基于激光雷达的同步定位与地图构建

(LiDAR-SLAM)算法重建被测物体轮廓同时对图像进行目标检测,对点云去噪后利用目标检测结果作为约束条件对点云进行分割和聚类;最后设计了轮廓拟合算法对目标点云进行轮廓拟合并利用三维拟合算法得到目标尺寸。实验证明,本研究方法可以对汽车尺寸进行实时、稳定、高精度的自动化测量。对道路场景下的汽车尺寸测量具有重要意义。

1 基于视觉语义约束的激光雷达大目标尺寸测量算法框架设计

当下大多数目标体积测量方法通常都假设传感器位置固定或者是基于特定的场景,而且对于较大尺寸目标存在单帧数据难以完整反映目标尺寸轮廓导致体积测量出错的问题。针对于此,本研究提出了基于激光雷达的大目标体积测量算法。利用无人车上的激光雷达、相机和IMU采集的数据即可实现对汽车尺寸的实时测量,算法框架如图1所示。

首先对传感器数据进行时空对齐;然后对图像进行目标检测获取目标的种类与位置信息;在点云处理时,使无人车绕被测物体至少半圈,获取目标3个维度上完整的信息,利用LiDAR-SLAM算法重建被测目标轮廓,对点云去噪之后利用地面分割算法分割地面点与非地面点云,利用直通滤波后对点云聚类,然后将不同点云簇中心进行视觉投影,根据图像目标检测的定位结果和KD Tree(K-dimensional tree)对目标点云进行轮廓拟合;最后对轮廓进行偏航角和长宽高的拟合,并且对目标附近地面点云进行平面拟合得到高度上的地面补偿;最后得到目标尺寸和位置信息。以此解决了单帧数据难以反映大目标完整尺寸轮廓和点云扰动性带来的干扰。

1.1 数据时空对齐

本研究方法需要用到激光雷达、相机和IMU这3种传感器。但由于各传感器输出独立,所以需要对其进行时空对齐。本文利用最近邻匹配算法对数据进行时间软同步,利用文献[11-12]方法对数据进行空间上的对齐,从而对数据完成时空对齐。

1.2 LiDAR-SLAM 算法优化

本研究利用LiDAR-SLAM算法对被测物体进行轮廓重建,采用更快的激光雷达-惯性里程计^[13](faster light detection and ranging inertial odometry, Faster-LIO)作为轮廓重建的基准算法。该算法在Newer College Dataset^[14]上具有最快的速度,可以很好满足实时性要求,并且在绝对姿态误差(absolute pose error, APE)和相对位姿误差(relative pose error, RPE)两个指标上也具有比较理想的结果。

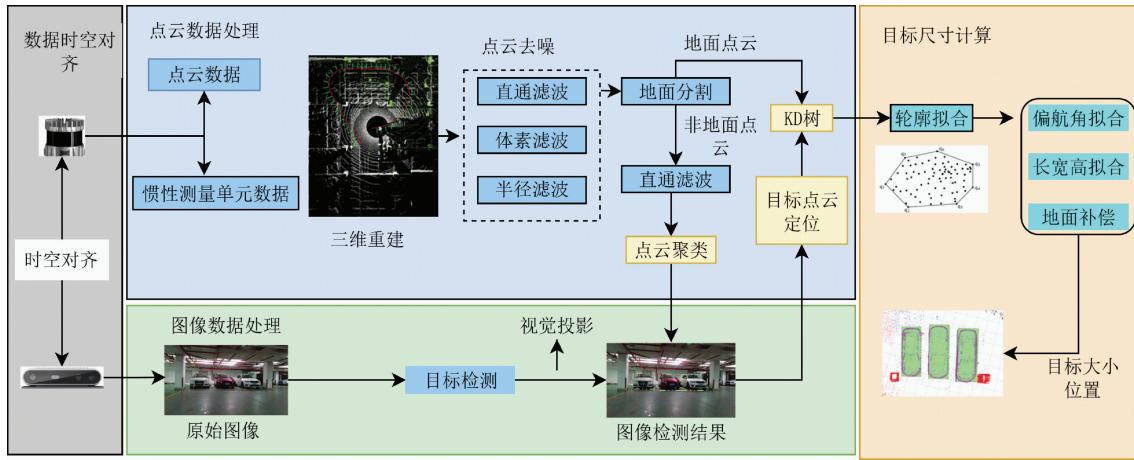


图 1 尺寸测量算法框架
Fig. 1 Dimension measurement algorithm framework

效果。以快速直接激光雷达慢性里程计 2 (fast direct LiDAR-inertial odometry 2, FAST-LIO2)^[15]为基础进行优化,简化了整体流程。由于无人车受限于平面运动,为减少漂移,本研究建立了地面观测模型和融合地面残差优化对 Faster-LIO 算法进行优化,增强其性能。为消除累计误差,提高定位精度以及建立全局一致的地图,本研究加入了回环检测框架(stable triangle descriptor, STD)^[16]。

1) 基于地平面约束的残差优化

从激光雷达数据中提取地面点,有助于估计激光雷达传感器相对于地面的相对方向,利用地平面的平面约束可以增强 LiDAR-SLAM 的性能,减少漂移。整个优化过程可以分为两步,建立地面观测模型和融合地面残差优化。

首先,从激光雷达输入中分割出地面,激光雷达地面分割旨在获得激光雷达相对于地面的旋转矩阵^c R_L ,利用 Patchwork++ 获取一组地面点并计算协方差矩阵,通过主成分分析法获取在 LiDAR 坐标系 L 下地平面的法向量 $\mathbf{n}_L = (a_L, b_L, c_L)^T$,地平面方程为:

$$a_Lx + b_Ly + c_Lz + d_L = 0 \quad (1)$$

设在地面坐标系 G 下的地平面法向量 $\mathbf{n}_G = (0, 0, 1)^T$,可知 LiDAR 坐标系 L 到地面坐标系 G 的旋转矩阵为:

$${}^c\mathbf{R}_L = \exp([\boldsymbol{\phi}_{GL}\mathbf{n}_{GL}]_{\times}) \quad (2)$$

式中: $\boldsymbol{\phi}_{GL} = \arccos(\mathbf{n}_G^T \mathbf{n}_L)$; $\mathbf{n}_{GL} = \mathbf{n}_G \times \mathbf{n}_L$; $[\cdot]_{\times}$ 表示斜对称矩阵。因为移动载体限制在 2D 空间运动,利用激光雷达坐标系 L 到世界坐标系 W 的旋转矩阵^w R_L 中 roll 和 pitch 与平移向量^w p_L 的 z 轴分量应该一致的约束,即在时间 k 得到地平面法向量,向量 \mathbf{n}_{L_k} 应该与地面坐标系 x-y 平面正交,平移分量^w p_{L_k} 的 z 轴分量相对于地面坐标系的位移应当为 0,可得观测模型为:

$$\begin{cases} [{}^c\mathbf{R}_L {}^w\mathbf{R}_{L_k}(\mathbf{n}_{L_k} + \boldsymbol{\epsilon}_{n_k})]_{1,2} = \mathbf{0} \\ \mathbf{e}_3 {}^c\mathbf{R}_L {}^w\mathbf{p}_{L_k} = \mathbf{0} \end{cases} \quad (3)$$

式中: $\boldsymbol{\epsilon}_{n_k}$ 表示 \mathbf{n}_{L_k} 的观测噪声; \mathbf{e}_3 表示单位向量; $[\cdot]_{1,2}$ 表示只取 3×1 向量的前两个元素。

Faster-LIO 中采用迭代误差状态卡尔曼滤波器 (iterated error state Kalman filter, IESKF) 进行状态估计,状态向量 \mathbf{x}_k 为:

$$\mathbf{x}_k = [{}^w\mathbf{R}_I^T, {}^w\mathbf{p}_I^T, {}^w\mathbf{v}_I^T, \mathbf{b}_g^T, \mathbf{b}_a^T, {}^w\mathbf{g}^T, {}^I\mathbf{R}_L^T, {}^I\mathbf{p}_L^T]^T \quad (4)$$

式中:^w \mathbf{R}_I^T 和^I \mathbf{R}_L^T 分别为 IMU 到世界坐标系和激光雷达到 IMU 坐标系的旋转矩阵;^w \mathbf{P}_I^T 和^I \mathbf{P}_L^T 分别为 IMU 到世界坐标系和激光雷达到 IMU 坐标系的平移向量;^w \mathbf{v}_I^T 为 IMU 到世界坐标系速度向量; \mathbf{b}_g^T 和 \mathbf{b}_a^T 分别表示角速度零偏向量和线加速度零偏向量;^w \mathbf{g}^T 为重力加速度。联合 IMU 的数据可得 IESKF 的状态预测 $\hat{\mathbf{x}}_{k+1}$ 和协方差矩阵 $\hat{\mathbf{P}}_{k+1}$ 为:

$$\begin{cases} \hat{\mathbf{x}}_{k+1} = \bar{\mathbf{x}} \boxplus (\Delta t \cdot f(\bar{\mathbf{x}}, \mathbf{u}_k, \mathbf{w}_k)) \\ \hat{\mathbf{P}}_{k+1} = \mathbf{F}_{\dot{\mathbf{x}}} \bar{\mathbf{P}}_k \mathbf{F}_{\dot{\mathbf{x}}}^T + \mathbf{F}_w \mathbf{Q}_k \mathbf{F}_w^T \end{cases} \quad (5)$$

式中:($\bar{\cdot}$)、($\dot{\cdot}$)、($\tilde{\cdot}$) 分别为最佳值、预测值、误差值; \boxplus 表示广义加法^[17]; \mathbf{u}_k 表示 IMU 输入; \mathbf{w}_k 表示 IMU 的噪声,实验中将 \mathbf{w}_k 置为 0; \mathbf{Q}_k 是 \mathbf{w}_k 的协方差矩阵, $\mathbf{F}_{\dot{\mathbf{x}}}$ 和 \mathbf{F}_w 是对 $\bar{\mathbf{x}}_k$ 和 \mathbf{w}_k 为 0 关于 $(\hat{\mathbf{x}}_{k+1} \boxplus \hat{\mathbf{x}}_{k+1})$ 的偏导数。融合状态量可得地平面的测量模型为:

$$\begin{aligned} \mathbf{0} &= h_{\pi}(\mathbf{x}_k, \mathbf{n}_{L_k} + \boldsymbol{\epsilon}_{n_k}) = \\ &\left([{}^c\mathbf{R}_L {}^I\mathbf{R}_{L_k}^{-1} {}^w\mathbf{R}_{L_k} {}^I\mathbf{R}_{L_k}(\mathbf{n}_{L_k} + \boldsymbol{\epsilon}_{n_k})]_{1,2} \right) \end{aligned} \quad (6)$$

再结合点的测量模型,即:

$$h_j(\mathbf{x}_k, {}^I\mathbf{p}_j + {}^L\mathbf{n}_j) = {}^G\mathbf{u}_j^T ({}^G\mathbf{T}_I {}^I\mathbf{T}_L ({}^I\mathbf{p}_j + {}^L\mathbf{n}_j) - {}^G\mathbf{q}_j) = \mathbf{0} \quad (7)$$

可得:

$$\begin{cases} h_j(\mathbf{x}_k, {}^L\mathbf{p}_j + {}^L\mathbf{n}_j) \approx \mathbf{z}_j + \mathbf{H}_j \tilde{\mathbf{x}}_k + \mathbf{v}_j, \mathbf{z}_j = h_j(\hat{\mathbf{x}}_k, {}^L\mathbf{p}_j + \mathbf{0}) \\ h_\pi(\mathbf{x}_k, \mathbf{n}_{L_k} + \mathbf{e}_{n_k}) \approx \mathbf{z}_\pi + \mathbf{H}_\pi \tilde{\mathbf{x}}_k + \mathbf{v}_\pi, \mathbf{z}_\pi = h_\pi(\hat{\mathbf{x}}_k, \mathbf{n}_{L_k} + \mathbf{0}) \end{cases} \quad (8)$$

式中: \mathbf{x}_k 表示状态量, 包含了 ${}^C\mathbf{T}_l$ 和 ${}^I\mathbf{T}_l$; ${}^L\mathbf{p}_j$ 和 ${}^L\mathbf{n}_j$ 分别代表扫描点和其对应的噪声; ${}^C\mathbf{u}_j^T$ 为对应平面法向量; ${}^C\mathbf{q}_j$ 为对应平面上的点; ${}^C\mathbf{T}_l$ 为 IMU 到世界坐标系的转换关系; ${}^I\mathbf{T}_l$ 为 LiDAR 到 IMU 坐标系的转换关系。 $\mathbf{v}_j \in N(\mathbf{0}, \mathbf{R}_j)$ 、 $\mathbf{v}_\pi \in N(\mathbf{0}, \mathbf{S}_k)$ 来源于测量噪声, 满足正态分布; \mathbf{H}_j 和 \mathbf{H}_π 的表达式如式(9)所示。

$$\begin{cases} \mathbf{H}_j = \frac{\partial h_j}{\partial \mathbf{x}_k} \Big|_{\mathbf{x}_k = \tilde{\mathbf{x}}_k} \\ \mathbf{H}_\pi = \frac{\partial h_\pi}{\partial \mathbf{x}_k} \Big|_{\mathbf{x}_k = \tilde{\mathbf{x}}_k} \end{cases} \quad (9)$$

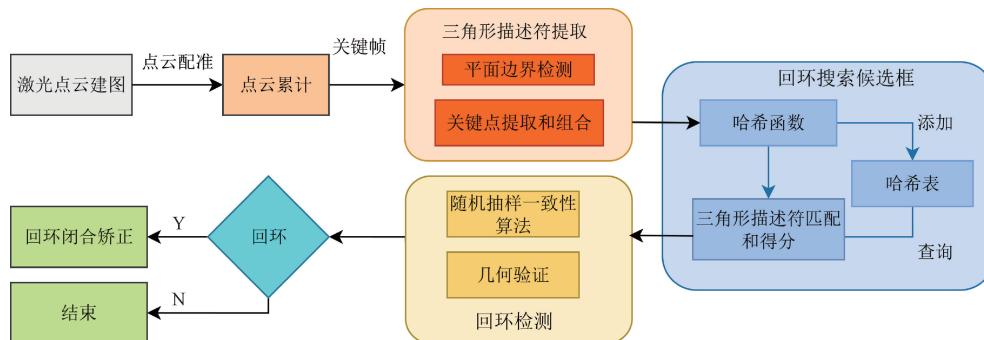


图 2 STD 回环检测框架
Fig. 2 STD loop closure detection framework

1.3 点云数据处理

1) 点云去噪

利用 LiDAR-SLAM 对环境建完图后, 目标周围的噪声点也被累计, 同时由于目标表面的激光折射导致更多的噪声点被统计, 这些虚假噪声点在很大程度上会影响后续检测效果, 因此需要将这些点去除, 本文采用直通滤波、体素滤波和半径滤波 3 种滤波方法联合对点云数据进行预处理。

体素滤波方法将点云划分成体素并利用体素中心进行下采样, 在很大程度上减少了数据量, 具有较高的效率; 直通滤波法是根据坐标轴范围进行滤波的简单方法, 对坐标轴某个范围内的点保留, 该方法简单快速, 对不在范围内的点可以有效去除, 但对于复杂形状的点云, 该方法会失效; 半径滤波根据某个点的半径阈值和点数阈值对点进行滤除, 对于离群点能有效滤除。

由于 LiDAR-SLAM 输出的是以第 1 帧为世界坐标系的点云, 首先利用直通滤波器对某个固定范围内的点进

行保留, 然后利用体素滤波器对点云下采样, 最后利用半径滤波器对明显离群点进行去除, 通过这 3 种滤波方法, 不仅能有效降低数据量还能有效去除噪声。具体效果如图 3 所示。

可得优化目标函数为:

$$\min_{\mathbf{x}_k^t} (\|\mathbf{x}_k \ominus \hat{\mathbf{x}}_k\|_{P_k}^2 + \sum_{j=1}^m \|\mathbf{z}_j^t + \mathbf{H}_j^t \tilde{\mathbf{x}}_k\|_{R_j}^2 + \|\mathbf{z}_\pi^t + \mathbf{H}_\pi^t \tilde{\mathbf{x}}_k\|_{S_k}^2) \quad (10)$$

式中: t 表示迭代次数; m 表示有效点云数量。目标函数利用 IESKF 进行求解。

2) 基于 STD 的回环检测

STD^[16] 回环检测算法创新性引入稳定三角形描述符(具旋转平移不变性), 通过 LiDAR-SLAM 配准点云生成关键帧并构建三角形拓扑, 采用哈希表实现描述符高速检索匹配性。STD 回环检测框架如图 2 所示。

其中, STD 提取过程为: 当给定点云的关键帧时, 先通过区域生长进行平面检测, 以边界点到平面的距离为依据提取特征点, 对邻域内的特征点进行组合形成三角形描述符。

2) 地面分割算法

在进行目标体积计算之前, 需要辨别地面点和非地面点, 避免地面点影响后面轮廓拟合部分。点云地面分割常用技术主要包含的类别为: 几何特征分析方法、聚类算法、统计模型拟合方法(比较经典的 Patchwork++^[18])以及机器学习与深度学习算法^[19-20]。针对实际场景中存在的端侧计算资源限制、地面分割精度要求严苛, 以及点云数据存在噪声干扰等技术挑战, 本研究选定 Patchwork++ 算法作为地面分割的实施方案。

该算法的核心在于通过多区域环形划分与动态自适应参数优化, 突破传统方法在点云密度不均、地形变化及实时性方面的局限。本文使用的 OS1-64 激光雷达随着探测距离的增加点云也更稀疏。该方法首先将点云划分为多个同心环形区域, 然后针对点云近密远疏的分

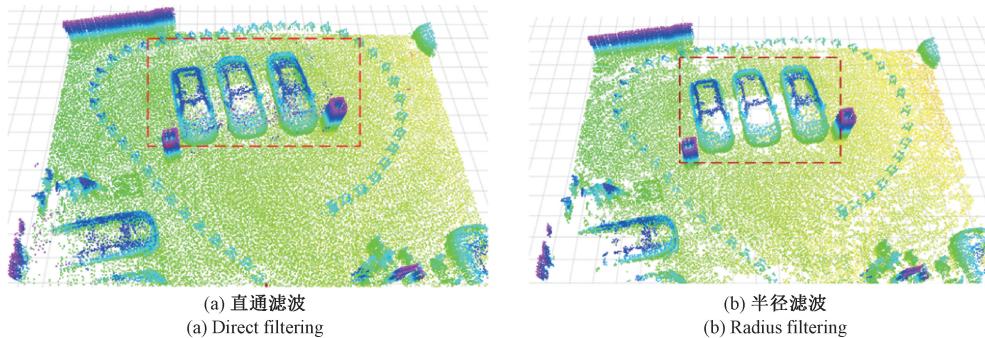


图 3 点云去噪效果

Fig. 3 Point cloud denoising effect

布特点,为每个区域独立配置高度差阈值、法向量约束等参数,实现局部精细化拟合;其次,结合时序信息动态更新地面模型,适应运动场景中的坡度变化与颠簸干扰,同时通过统计滤波与邻域一致性检验抑制噪声及动态物体残留点云;最后,利用区域间平滑约束与全局非线性优化,融合各分区结果并重构连续一致的地面曲面。

3) 目标点云分割

针对非地面点云分割问题,利用视觉投影与目标检测的结果相结合的方式^[5]提取目标点云,通过激光雷达与相机坐标系转换及内参计算实现点云视觉投影,结合目标检测结果筛选需聚类的点云,并采用基于密度的聚类(density-based spatial clustering of applications with noise,DBSCAN)^[21]完成分割。针对 LiDAR-SLAM 轮廓重建引入的噪声,基于雷达线束分布特性自动优化 DBSCAN 参数,即以 Z 轴间距作为参数选择依据^[5]。将领域半径设为距离的 2 倍,最小点数设为 10。

1.4 目标检测算法

不同于稀疏的激光雷达数据,图像数据含有丰富的语义信息,目标检测算法可以智能识别周围物体的语义信息,辅助定位不同目标点云位置,滤除大量背景噪声,可以避免人工设计不同目标点云特征。采用 YOLOv8s^[22]作为目标检测的算法,该目标检测算法是一种单阶段目标检测算法。YOLOv8 集成跨层连接与 Split-C3f 模块构建多层次特征提取网络,采用解耦检测头与 Anchor-Free 机制实现结构简化。为减少模型大小同时保证精度,在计算损失函数上,使用 SiLU(sigmoid linear unit)激活函数增强梯度流动和特征复用能力,在训练时,融合混合精度训练与自动超参数优化技术,结合动态标签分配策略和 Varifocal Loss 动态平衡样本权重,提高模型在不同数据集上的表现能力。

1.5 目标尺寸计算

1) 轮廓拟合

在对轮廓进行拟合时,需要将点云投影至 $x - o - y$ 平面,由于本文研究的对象为汽车,其形状为凸包,所以

需要对其表面的尖锐部分进行去除,轮廓的拟合过程为:

(1) 将目标点云投影至 $x - o - y$ 平面,使用 ConvexHull^[23]凸包算法找出点云的初始轮廓点;

(2) 遍历所有的轮廓点,建立八叉树寻找在半径为 r (默认值为 0.05 m)的球型区域内的点云并计算法向量,如果点云的法向量与轮廓的法向量夹角超过阈值(默认值 30°),则判断该点为凸起部分;

(3) 从目标点云中滤除所有被判断为凸起的点,并利用半径滤波对剩余点云进行滤波处理,然后重新利用 ConvexHull 凸包算法找出点云的轮廓点。其中 ConvexHull 凸包算法的处理流程为:首先进行极点的初始化,找到点集中 x 坐标最小的点 A 和最大的点 B ,因为这两个点必为凸包的顶点,并用直线 AB 将点集分为上下两部分(叉积计算),然后进行分治处理,对直线 AB 上方和下方的点集递归处理,处理过程为:

a) 找到离直线 AB 最远的点 C ,加入凸包点集;

b) 将点集分为 AC 左侧和 CB 左侧的子集,继续递归处理,当子集为空时停止递归。轮廓拟合完成后,执行偏航角、长宽高的拟合。

为方便观察各类目标的轮廓拟合效果,此处直接对聚类后的所有点云进行轮廓拟合,在实际使用中,只筛选出与视觉目标检测结果匹配成功的点云,然后对筛选出的点云进行轮廓拟合,对聚类后所有点云的轮廓拟合的具体效果如图 4 所示。

在本文实验中,视觉检测上输出 6 个汽车目标,点云数量分别为 [25 313, 29 624, 2 461, 14 625, 5 005, 1 714]。对 6 个目标直接进行三维框拟合的花费的时间为 4 158 ms,先进行轮廓拟合然后再三维框拟合花费的时间为 205 ms(其中轮廓拟合花费的时间为 186 ms、三维框拟合花费的时间为 19 ms),可看出进行轮廓拟合显著降低了后续计算负担。

2) 三维框拟合

在实际应用中,针对点云集中于物体前表面导致三维边界框拟合失准的问题,本研究的核心思想是将点云

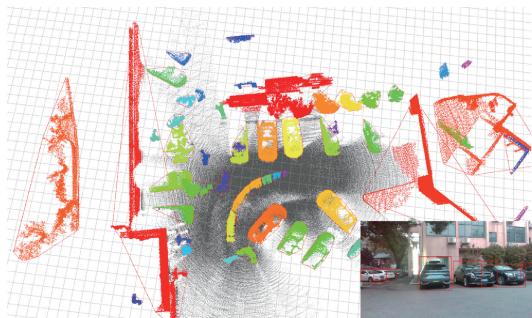


图4 轮廓拟合效果

Fig. 4 Contour fitting effect

投影到 $x-y$ 平面^[4],在忽略俯仰角和翻滚角的情况下,采用遍历搜索机制寻找最优偏航角,以点到边的平方误差 Varcriterion^[24]为损失函数确定目标长宽及中心点,同时依据地面点云计算高度。该方案兼具抗点云毛刺干扰与噪声鲁棒性优势。

2 实验验证与分析

2.1 实验平台搭建及数据集构建

为验证本章大尺寸测量算法对于汽车测量的精确性,以停车场的汽车作为实验对象。

实验平台使用移动小车作为载体,搭建激光雷达、相机、嵌入式平台等。搭建的硬件平台如图 5 所示。硬件平台的参数如表 1 所示。



图5 尺寸测量系统

Fig. 5 Dimension measurement system

其中激光雷达的型号为 OS1-64,激光雷达内置 IMU,其型号为 IAM-20680HT,相机的型号为 Intel RealSense D435i,有两个近红外摄像头和一个可见光摄像头,本研究只用可见光数据,嵌入式平台为 NVIDIA Jetson AGX

表1 硬件平台参数

Table 1 Hardware platform parameters

传感器	参数
	探测距离: 90 m
	测距精度: ± 2 cm
	水平视场角: 360°
OS1-64	垂直视场角: 45° ($\pm 22.5^\circ$)
	水平角分辨率: 0.35° (1 024)
	垂直角分辨率: 0.7° (64)
	扫描频率: 10 Hz
IAM-20680HT	陀螺仪量程(灵敏度): ± 250 dps (131 LSB/dps)
	加速度计量程(灵敏度): ± 2 g (16 384 LSB/g)
	输出频率: 100 Hz
Intel RealSense D435i	RGB 帧分辨率: 1 280×720 RGB 传感器 FOV (H×V): 69°×42° RGB 帧率: 30 fps

Xavier,配备移动电源、显示屏等辅助组件,移动电源为传感器、辅助组件和移动小车提供充足电源。

较低的相机分辨率会导致大目标边缘尺寸失真,进而导致外参标定及时间同步误差增大,影响检测结果,进而影响对目标的辅助定位,使目标尺寸检测失败,本研究选择了能确保目标边缘有足够像素的分辨率;较大的激光雷达角分辨率会导致点云稀疏,从而降低特征匹配精度,导致外参匹配失准,进一步影响多源数据空间对齐的准确性,同时还会直接影响到目标轮廓重建的精度和可靠性,导致点云分割及点云聚类出现较大偏差,降低目标尺寸测量精度;较小的角分辨率会导致数据量过大,增加计算负担;较大的 IMU 误差会导致数据截延迟和运动补偿错误,影响多源异构数据在时间和空间上的对齐,此外会导致建图出现严重的漂移,这些误差在点云数据处理过程中被传递,直接导致目标尺寸测量精度降低。

本研究构建了一个含有 6 类目标(武器箱、医疗箱、油桶、人、汽车、卡车)的图像数据集,含有 6 347 张图片,29 570 个标签,数据集的具体划分如表 2 所示,其中车辆作为本研究实验的主要检测目标。

本文算法在 NVIDIA Jetson AGX Xavier 嵌入式平台上基于 C++ 进行开发,在 ROS 框架下订阅相机与激光雷达驱动发布的话题,对两个传感器采集的数据进行实时处理,其软件环境具体如表 3 所示。

表 2 图像目标检测数据集

Table 2 Image object detection dataset

类别	类别名	训练集		验证集	
		图像数	标签数	图像数	标签数
汽车	Car	3 130	10 840	782	2 826
人	Person	2 574	6 114	616	1 447
卡车	Truck	821	1 870	204	480
医疗箱	Medicalbox	1 757	1 760	453	455
武器箱	Weaponbox	893	896	221	222
油桶	Oilbarrel	1 492	2 145	350	515
合计		5 078	23 625	1 269	5 945

2.2 消融实验

1) Faster-LIO 算法优化

为了验证所提的 LiDAR-SLAM 算法中加入的各模块

表 4 Faster-LIO 算法优化消融实验

Table 4 Ablation experiments on the Faster-LIO algorithm optimization

Ground_Seg	STD	ncd_00/cm		ncd_01/cm		时间/ms	
		绝对误差	相对误差	绝对误差	相对误差	合计	平均
✓		61.581 0	8.260 2	18.295 0	9.994 3	76	
✓	✓	15.522 6	7.911 4	15.985 6	9.963 6	88	
		15.788 7	7.900 9	14.367 7	9.963 4		123

在加入地面残差优化 Ground_Seg 后, 在 ncd_00 数据上, APE 指标下降 46 cm, 有较明显的提升, 在 ncd_01 数据上 APE 下降了 2.3 cm, 在此基础上加入 STD 回环检测, 在 ncd_00 数据上由于路径没有明显的回环, 反而造成了误差的略微提升, 在 ncd_01 数据上, 由于刚开始那段路径是个回路, 因此 STD 发挥了作用, APE 相较于 baseline 下降了 4 cm。

在速度上, 对于 Ground_Seg 部分, 需要额外花费时间在地面分割上, 是造成速度下降的主要因素, 对 STD 而言, 提取 STD 描述符在 10 ms 以内能够完成, 一旦检测到回环, 花费的时间平均为 123 ms。但需要注明地是, 当候选帧的数量提升, 花费的时间也会更多。如图 6 所示为 STD 回环检测效果。

2) 目标检测算法性能对比

目标检测算法性能使用 mAP (mean average precision) 作为精度指标, FPS (frame per second) 作为速率指标。基于本文自建的含有 6 类目标 (武器箱、医疗箱、油桶、人、汽车、卡车) 的图像数据集, 在服务器上对 YOLOv5s、CenterNet2、RT-DETR-1、YOLOv8s、YOLOv10s、YOLOv11s 模型进行训练测试, 其中服务器的 CPU 型号为 Intel i9-7920x, 48 核 64 GB 内存, GPU 为 Nvidia RTX3090ti, 模型皆在 Pytorch1.10 下训练与测试, 训练参

表 3 软件环境

Table 3 Software environment

软件名称	版本
CUDA	10.2
PCL	1.8.1
ROS	Melodic
OpenCV	4.1.1
TensorRT	7.1.3

对算法性能的提升作用, 在实验参数一致的条件下, 对算法进行了消融实验。

本研究共提出了 3 个消融实验: (1) 不加入任何模块, 用原始算法; (2) 加入基于地平面的残差优化; (3) 加入基于地平面的残差优化及 STD 回环检测模块。

如表 4 所示, 对于 RPE 指标而言, 优化前后只有细微的浮动, 并无明显区别。

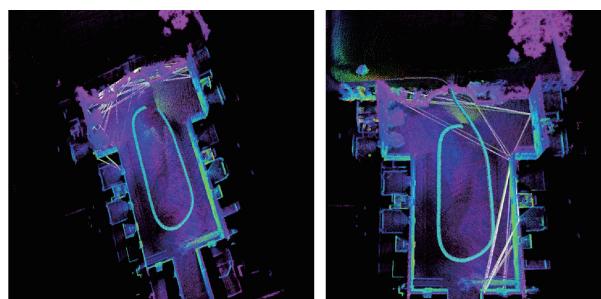


图 6 STD 回环检测效果

Fig. 6 Effectiveness of STD loop closure detection

数 batchsize 为 16、epochs 为 200。YOLOv8s 在 mAP50、mAP75、mAP50-95 这 3 个精度指标和 FPS 速度指标上均高于其余模型。其性能对比如表 5 所示。

YOLOv8s 在计算平台 Nvidia Jetson Xvaier AGX 上, mAP50 为 88.23%、mAP75 为 82.37%、mAP50-95 为 78.98%、FPS 为 29。

2.3 实验结果与分析

为验证针对大尺寸目标的尺寸测量算法的精确性, 在停车场拍摄了不同型号车辆数据进行实验, 采用最大误差和平均误差对车辆的长、宽、高尺寸进行评估。其实验结果如表 6 所示。

表5 目标检测网络性能对比
Table 5 Performance comparison of target detection networks

模型	mAP50/%	mAP75/%	mAP50-95/%	FPS	参数/M	GFLOPs _{imgsz=640}
YOLOv5s	89.11	82.80	77.53	67	9.11	23.8
CenterNet2	88.78	81.20	78.11	40	71.42	158.0
RT-DETR-l	85.83	75.84	70.60	33	32.00	103.5
YOLOv8s	90.69	84.16	80.62	74	11.13	28.4
YOLOv10s	90.22	83.69	80.00	57	8.04	24.5
YOLOv11s	90.50	83.93	80.25	61	9.42	21.3

表6 尺寸测量结果
Table 6 The result of dimensional measurement (mm)

车辆型号	长	测量值	宽	测量值	高	测量值
东风日产轩逸	4 631	4 600(-31)	1 760	1 706(-54)	1 503	1 443(-60)
奔驰C260	4 698	4 726(+28)	1 810	1 869(+59)	1 408	1 425(+17)
斯柯达SKODA	4 753	4 718(-35)	1 832	1 801(-31)	1 469	1 436(-33)
马自达3昂克塞拉	4 662	4 754(+92)	1 797	1 816(+19)	1 445	1 409(-36)
奥迪A3 35TFSI	4 351	4 357(+6)	1 815	1 791(-24)	1 458	1 430(-28)
平均误差	-	0.82%	-	2.08%	-	1.99%
最大误差	-	1.97%	-	3.26%	-	3.99%

对于汽车外廓尺寸测量,长度的最大误差和平均误差分别为1.97%和0.82%、宽度的最大误差和平均误差为3.26%和2.08%、高度的最大误差和平均误差为

3.99%和1.99%。可见本文方法对于汽车尺寸测量具有很好的效果,可以对较大尺寸目标进行精确的尺寸测量。实验中的部分结果图如图7、8、9所示。

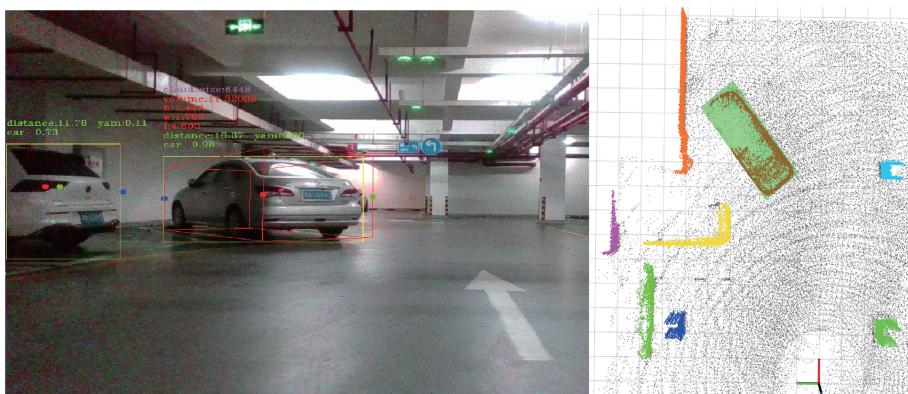


图7 场景1实验结果
Fig. 7 The experimental result of scene 1

值得注意的是,在测量过程中,需要保证对点云重建后能完整反映目标3个维度上的信息,否则即使成功拟合三维框,也无法保证测量精度,例如图8中车辆。在点云处理模块中,重建后的点云密度受建图过程中地图分辨率和环绕方式等因素的影响,并且需要对点云进行滤波处理,难免会丢失目标的部分细节信息,影响后续尺寸测量精度。在尺寸测量过程中,对于

目标点云而言,最重要的是关于目标偏航角的拟合,当偏航角测量有误差时,长宽的测量精度也随之受影响。但从图7~9效果图可以看出,本文方法不仅可以从重建点云中成功定位出目标点云,还可以对偏航角精确拟合,使目标的边框贴合目标点云。本文方法在经过多辆实车验证后,尺寸测量的最大误差<4.0%,平均误差<2.0%。



图 8 场景 2 实验结果

Fig. 8 The experimental result of scene 2

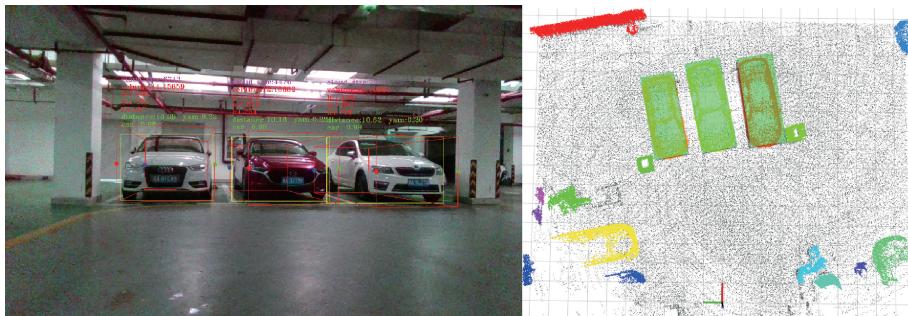


图 9 场景 3 实验结果

Fig. 9 The experimental result of scene 3

如图 10 所示为采用首帧点云和图像数据, 基于视觉修正的目标尺寸测量效果图, 目标与图 9 对应。可以明显看到对于大尺寸目标而言, 单帧数据的视角明显受限, 现有的基于视觉修正的尺寸测量方法也难以对该类目标进行准确的尺寸测量, 而如图 9 所示, 在本文的方法下, 像汽车这类大尺寸目标的尺寸可以被准确的测量, 解决了单帧数据视角受限的问题, 体现了本文方法在大尺寸目标尺寸测量中的有效性。

雨雾及雪天等较为恶劣的天气情况下, 激光的传播受到很大影响, 同时相机传感器也受到环境光的影响, 而本文算法依赖图像目标检测的结果及点云质量, 所以本文算法适用于光线充足和空气纯净的天气状况, 实验证明, 在合适的环境条件下, 本文方法可以对大尺寸目标进行精确的尺寸测量, 由于实验条件和设备受限, 未能在其他复杂条件下进行进一步的研究。

3 结 论

针对单帧数据难以完全反映大尺寸目标的完整轮廓而带来的尺寸测量效果差的问题, 本文首次提出了基于视觉语义约束的 LiDAR-SLAM 尺寸测量方法, 算法的主要创新点为利用优化后的 Faster-LIO 算法对点云重建, 设计了点云去噪流程以解决重建后点云数量庞大和累计噪声的问题。在尺寸测量模块, 设计了轮廓拟合方法以减少噪声和计算负担, 然后利用三维框拟合算法, 实现了对目标的尺寸测量。实验结果表明, 本研究算法在一定范围内能对汽车尺寸进行高精度实时测量, 可应用于道路上车辆尺寸的检测及自动驾驶等领域。

但本研究方法也还存在一定的局限性, 首先本研究只对汽车这类大目标的尺寸进行了测量, 其次, 在信息利用上, 只利用了第 1 帧的视觉信息, 忽略了小车建图过程中视觉检测部分的输出, 导致虽然可以同时检测多个目

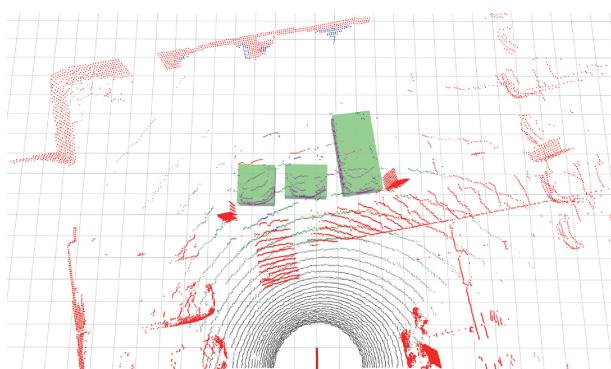


图 10 基于视觉修正的尺寸测量结果

Fig. 10 The result of size measurement based on visual correction

由于本文使用的是机械式激光雷达, 其原理不依赖于环境光, 所以能在白天和黑夜环境下正常工作, 但是在

標,但只局限於出現在首幀中的目標。未來工作將會驗證更多類型的大尺寸目標,同時在尺寸測量系統上加裝多個視角的相機,在建圖過程中同時檢測多個相機中的目標,最後再測量所有目標的尺寸。

參考文獻

- [1] 程亮,羅洲,何贊澤,等. 視覺和激光SLAM發展綜述與展望[J]. 中國測試, 2025, 51(6): 1-16.
CHENG L, LUO ZH, HE Y Z, et al. Review and prospect of visual and laser SLAM development [J]. Chinese Test, 2025, 51(6): 1-16.
- [2] 汤新华,代道文,陈熙源,等. 基於PointPillars的改進三維目標檢測算法[J]. 儀器儀表學報, 2024, 45(9): 260-269.
TANG X H, DAI D W, CHEN X Y, et al. Improved three-dimensional object detection algorithm based on PointPillars[J]. Chinese Journal of Scientific Instrument, 2024, 45(9): 260-269.
- [3] 張浩,左杭,劉寶華. 視覺與二維激光雷達的目標檢測方法[J]. 電子測量與儀器學報, 2022, 36(3): 79-86.
ZHANG H, ZUO H, LIU B H. Target detection method for visual and 2D laser radar[J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(3): 79-86.
- [4] 周志偉,周建江,王佳賓,等. 基於雷達和視覺多級信息融合的目標檢測網絡[J]. 電子測量技術, 2024, 47(24): 110-117.
ZHOU ZH W, ZHOU J J, WANG J B, et al. Target detection network based on multi level information fusion of radar and vision [J]. Electronic Measurement Technology, 2024, 47(24): 110-117.
- [5] 郭隆強,何贊澤,杜旭,等. 基於視覺修正的激光雷達體積測量方法[J]. 儀器儀表學報, 2023, 44(10): 48-59.
GUO L Q, HE Y Z, DU X, et al. LiDAR volume measurement method based on visual correction [J]. Chinese Journal of Scientific Instrument, 2023, 44(10): 48-59.
- [6] 楊亮,王琦. 基於雙目視覺的廢料尺寸測量系統設計與研究[J]. 大連交通大學學報, 2024, 45(5): 69-74.
YANG L, WANG Q. Design and research of waste size measurement system based on binocular vision [J]. Journal of Dalian Jiao Tong University, 2024, 45(5): 69-74.
- [7] 吳文濤,何贊澤,杜旭,等. 融合相機與激光雷達的目標檢測與尺寸測量[J]. 電子測量與儀器學報, 2023, 37(6): 169-177.
WU W T, HE Y Z, DU X, et al. Fusion camera and
- laser radar target detection and measurement[J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(6): 169-177.
- [8] TEUFEL S, GAMERDINGER J, VOLK G, et al. Collective PV-RCNN: A novel fusion technique using collective detections for enhanced local LiDAR-based perception[J]. 2023 IEEE 26th International Conference on Intelligent Transportation Systems, 2023: 1828-1834.
- [9] SHARMA, KUKREJA V. CPD: Faster RCNN-based dragon ball comic panel detection[J]. 2023 IEEE 12th International Conference on Communication Systems and Network Technologies, 2023: 786-790.
- [10] LANG A H, VORA S, CAESAR H, et al. PointPillars: Fast encoders for object detection from point clouds[C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 12689-12697.
- [11] KIM T, PAK G, KIM E. GRIL-Calib: Targetless ground robot IMU-LiDAR extrinsic calibration method using ground plane motion constraints[J]. IEEE Robotics and Automation Letters, 2024, 9(6): 5409-5416.
- [12] YU Y, FAN S, LI L, et al. Automatic targetless monocular camera and LiDAR external parameter calibration method for mobile robots [J]. Remote Sensing, 2023, 15(23): 5560.
- [13] BAI CH G, XIAO T, CHEN Y J, et al. Faster-LIO: Lightweight tightly coupled LiDAR-inertial odometry using parallel sparse incremental voxels [J]. IEEE Robotics and Automation Letters, 2022, 7(2): 4861-4868.
- [14] RAMEZANI M, WANG Y D, CAMURRI M, et al. The newer college dataset: Handheld LiDAR, inertial and vision with ground truth[C]. 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2020: 4353-4360.
- [15] XU W, CAI Y X, HE D J, et al. FAST-LIO2: Fast direct LiDAR-inertial odometry [J]. IEEE Transactions on Robotics, 2022, 38(4): 2053-2073.
- [16] YUAN CH J, LIN J R, ZOU Z H, et al. STD: Stable triangle descriptor for 3D place recognition [C]. 2023 IEEE International Conference on Robotics and Automation, 2023: 1897-1903.
- [17] XU W, CAI Y X, HE G J, et al. FAST-LIO2: Fast direct LiDAR-inertial odometry [J]. IEEE Transactions on Robotics, 2022, 38(4): 2053-2073.
- [18] LEE S, LIM H, MYUNG H. Patchwork++: Fast and robust ground segmentation solving partial under-segmentation using 3D point cloud[C]. 2022 IEEE/RSJ

International Conference on Intelligent Robots and Systems, 2022: 8618-8625.

- [19] ZHAO W SH, NING Y P, JIA X, et al. A rapid segmentation method of highway surface point cloud data based on a super voxel and improved region growing algorithm[J]. Applied Sciences, 2024, 14(7) : 2852.
- [20] ZHAO Y M. Multi-channel depth segmentation network based on 3D graph convolution algorithm and its application in point cloud segmentation[J]. Alexandria Engineering Journal, 2024, 109: 740-753.
- [21] KULKARNI O, BURHANPURWALA A. A survey of advancements in DBSCAN clustering algorithms for big data[C]. 2024 3rd International Conference on Power Electronics and IoT Applications in Renewable Energy and its Control, 2024: 106-111.
- [22] VARGHESE R, SAMBATH M. YOLOv8: A novel object detection algorithm with enhanced performance and robustness [C]. 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems, 2024: 1-6.
- [23] GAMBY A N, KATAJAINEN J. Convex-Hull algorithms: Implementation, testing, and experimentation[J]. Algorithms, 2018, 11(12):195.
- [24] ZHANG X, XU W D, DONG CH Y, et al. Efficient L-shape fitting for vehicle detection using laser scanners[C]. 2017 IEEE Intelligent Vehicles Symposium, 2017: 54-59.

作者简介



何贊澤,2006 年于西安交通大学获得学

士学位,2008 年于国防科技大学获得硕士学位,2012 年于国防科学技术大学获得博士学位,现任湖南大学教授,主要研究方向为嵌入式人工智能与边缘计算、红外热成像与机器视觉。

E-mail:yhe@vip.163.com

He Yunze received his B. Sc. degree from Xi'an Jiaotong University in 2006, his M. Sc. and Ph. D. degrees both from University of Defense Science and Technology in 2008 and 2012, respectively. He is currently a professor at Hunan University. His main research interests include embedded artificial intelligence and edge computing, infrared thermal imaging and machine vision.



邓堡元(通信作者),2018 年于湖南大
学获得学士学位,2024 年于湖南大学获得博
士学位,现为湖南大学助理研究员,主要研
究方向为成像与检测,探测与感知。

E-mail:dengbaoyuan@hnu.edu.com

Deng Baoyuan (Corresponding author) received his B. Sc. and Ph. D. degrees both from Hunan University in 2018 and 2024, respectively. He is currently an assistant research fellow at Hunan University. His main research interests include imaging and detection, as well as detection and perception.