

基于立体视觉与特征匹配的点云目标识别算法

王晓艳¹, 徐高魁^{2*}

(1. 西南林业大学 数理学院, 云南 昆明 650224;

2. 昆明铁道职业技术学院, 云南 昆明 650224)

摘要: 准确提取三维点云数据中待测目标的点云集合是三维点云目标识别技术的一个关键问题,也是近年来目标识别领域从二维向三维拓展的一个重要挑战,其主要难点在于快速寻找离散点云之间的相关函数关系。结合立体视觉与特征匹配构建了可以表征不同视场条件下的目标点云约束的机制,通过采用立体视觉作为约束条件完成了对原有特征匹配算法的优化。设计了基于立体视觉的估计算法,通过训练学习获得了不同选取比例条件下的识别规则。实验采用 ARIES 激光雷达采集点云,并通过 MATLAB 选取三种典型目标状态。当目标区分度高时,优化前后的目标识别率都在 98% 以上;当目标区分度低时,优化后对目标边界的限定条件可以很好地提高识别概率。采用优化的点云数据位置偏差量可达到 0.55 mm,相比未优化的 0.74 mm 提高了 0.19 mm。同时,优化后算法的收敛时间曲线要优于未优化的,3000 点以上的收敛时间均值约为 8.33 s,优于未优化的 12.76 s。综上所述,优化后的算法具有更好的识别效率。

关键词: 目标识别; 点云数据处理; 立体视觉; 特征匹配

中图分类号: TP391.41 **文献标志码:** A **DOI:** 10.3788/IRLA20210596

Point cloud target recognition algorithm based on stereo vision and feature matching

Wang Xiaoyan¹, Xu Gaokui^{2*}

(1. College of Mathematics and Physics, Southwest Forestry University, Kunming 650224, China;

2. Kunming Railway Vocational and Technical College, Kunming 650224, China)

Abstract: Accurately extracting the point cloud collection of the target to be measured in the three-dimensional point cloud data is a key issue of the three-dimensional point cloud target recognition technology, and it is also an important challenge in the field of target recognition from 2D to 3D in recent years. The main difficulty is to quickly find the correlation function relationship between discrete point clouds. Combining stereo vision and feature matching, a constraint mechanism of target point cloud that can characterize different field of view conditions is constructed, and the original feature matching algorithm is optimized by using stereo vision as the constraint condition. An estimation algorithm based on stereo vision is designed, and recognition rules under different selection ratios are obtained through training and learning. The experiment uses ARIES Lidar to collect point clouds, and selects three typical target states through MATLAB. When the target discrimination is high, the target recognition rate before and after optimization is above 98%. When the target discrimination is low, the

收稿日期:2021-08-24; 修订日期:2021-09-14

基金项目:国家自然科学基金(31760181); 云南省教育厅基金(2020J0404); 云南省重大科技专项计划项目(202002AD080002, 2019ZE005)

作者简介:王晓艳,女,讲师,硕士,主要研究方向为激光通信、光电检测等。

通讯作者:徐高魁,男,讲师,硕士,主要研究方向为信号处理、铁道信号分析等。

restriction conditions of the target boundary after optimization can improve the recognition probability. The position deviation of the optimized point cloud data can reach 0.55 mm, which is 0.19 mm higher than 0.74 mm before optimization. At the same time, the convergence time curve of the optimized algorithm is better than before optimization. The average convergence time above 3 000 points is about 8.33 s, which is better than 12.76 s before optimization. In summary, the optimized algorithm has better recognition efficiency.

Key words: target recognition; point cloud data processing; stereo vision; feature matching

0 引言

目标识别技术广泛应用于生产生活的各个领域,随着领域需求的提高,针对点云数据的目标识别成为了目标识别技术中的一项研究热点^[1-2]。由于点云是离散点构成的,用于二维图像的目标识别算法不再适用,研究点云目标识别算法具有实际意义。

随着激光雷达的应用越发普遍,三维目标的检测也成为了重要的发展趋势,例如,汽车自动避障、产品线残缺剔除等。Choi 等^[3]提出按照特征点对应的方式寻找目标匹配关系,可以提高识别精度,还采用了叠加颜色信息的方式提高识别能力。Mousavian 等^[4]曾将深度神经网络与几何约束进行融合,从而将其应用于自动驾驶过程中场景目标的识别。Engelcke 等^[5]采用 Vote 算法完成数据堆叠解算,增强对三维数据特征的提取能力。Yan Y 等^[6]通过卷积计算了点云目标数据之间的联系,提出了一种目标识别计算的思路。杨思远^[7]提出了一种路侧结构特征的分析方法,将其应用于车辆目标识别,根据不同车辆的点云特征分析车辆种类,具有较好的效果。Guo YL 等^[8]利用独特局部特征将目标与杂波进行分类识别,用相邻三角形隐式构造的方法提高了面型的重建效果。张哲等^[9]设计了一种点特征直方图的快速获取方法,可以更好地区分目标与背景点云之间的强度表征。

综上所述,在三维点云数据中直接提取目标信息是目前三维目标识别的一个重要方向。目前已有文献的研究重点主要是对点云滤波算法和点云数据分类算法的改进,虽然从测试效果上可以提高识别精度,但是普适性差,处理速度较慢^[10]。为了解决该问题,文中提出了一种基于立体视觉与特征匹配的点云目标识别算法。采用立体视觉作为待识别目标的几何约束,只要有待测目标的数学模型,则该算法具有很好的适用性;同时,采用特征匹配完成点云数据映射,可以压缩数据运算周期,提高算法收敛速度。

1 立体视觉测量模型构建

1.1 定义基准位置

为了表达点云数据中目标的三维信息,通过建立立体视觉测量模型对点云的边界进行限定与分析。由于点云获取主要是通过激光雷达扫描获得,所以为了便于数据匹配,采用球形网络作为三维数据节点的构建方式。将目标数模的中心定义为三维坐标系中原点位置 $o(x, y, z)$ 。 x 轴的方向在激光入射方向的垂直平面上,从原点向水平方向延伸; y 轴的方向在激光入射方向的垂直平面上,从原点向竖直方向延伸; z 轴为激光入射方向。原点到最远点的距离值为 r ,作为圆形区域的范围。

由此可以得到视点的最佳距离估计^[11],为了同时满足三个方向上覆盖率与重建精度,则坐标系与拟合球模型存在最佳距离估计 D_{best} ,其可表示为:

$$D_{best} = \min \left[\sqrt{(x_i - x_o)^2 + (y_i - y_o)^2 + (z_i - z_o)^2} \right] \quad (1)$$

式中: i 和 o 分别表示测试点云坐标系和拟合球模型的坐标系; x_i, y_i, z_i 和 x_o, y_o, z_o 分别表示对应坐标系中的三个坐标轴方向。故坐标对其后所有位置之间的最小距离就是最佳距离。

1.2 视点位姿判定

由于三维点云中待测目标涉及位姿状态问题^[12],所以需要对目标点云数模与测试坐标系点云进行对齐,其中基线中心计算关系可表示为:

$$\begin{cases} X_i = (a + D_{best}) \cos(\alpha_i) \cos(\beta_i) \\ Y_i = (b + D_{best}) \sin(\beta_i) \\ Z_i = (c + D_{best}) \sin(\alpha_i) \cos(\beta_i) \end{cases} \quad (2)$$

式中: X_i, Y_i, Z_i 为基线中心坐标值; a, b, c 分别表示在 x, y, z 轴上最佳距离的修正参数; α_i, β_i 分别表示对应第 i 个点的角度偏量, $\alpha_i = \alpha_{i-1} + \Delta\alpha, \beta_i = \beta_{i-1} + \Delta\beta$, 其中 $\Delta\alpha$ 和 $\Delta\beta$ 表示测试点与测试源之间的角度差异。

1.3 估计算法设计

立体视觉的本质是通过目标的三维立体结构去辅助构建待测目标的三维结构。在识别过程中,因为目标是具有固定外形结构的,所以可以将待测目标的数字模型作为模板,对点云数据进行分析。对于立体视觉的影响最明显的是测试角度,将不同测试角度投影得到的点云重叠分布特性作为目标识别依据,可以有效提升目标识别概率。估计算法步骤如下:

(1) 选定待测目标存在的点云区域三个坐标轴的区间分别是 $(x_m, x_{ml}), (y_n, y_{nl}), (z_b, z_{ll})$, 则其中任意一点 P 可以表示为 (p_x, p_y, p_z) , 当从激光雷达位置向该方向投影时,可以得到任意两个测试点之间在该角度 θ 上的投影距离值 D_i ;

(2) 因为立体视觉过程中是三维到二维的投影,故在相应的投影角度 θ 下,如果存在目标,则目标边缘位置的点云堆积最为明显,即投影距离值 D_i 与绝对距离 D_j 之间的变化最大。遍历计算所有点的 D_i/D_j 值,以一定比例选取 ($n\%$) 其中的最小值就能得到目标点云区域。如果不存在目标,则该运算结果无法符合预设的边界范围值,从而判断目标不存在;

(3) 根据不同的 θ 值,设置待测目标不同的选取比率 $n\%$,通过训练集完成对 $n\%$ 的训练,得到最优 $n\%$,从而使不同 θ 值具有合理的选取比率;

(4) 对选取测试点云区间中的点云数据进行视觉估计,将多种形状的组合进行排列分类,符合立体视觉预设逻辑的进行保留,不符合的剔除。由此可以剔除设定区域中的杂散点,通过该操作对目标点云集合进行滤波处理,进而获得较好的点云识别效果。

2 识别算法优化

2.1 特征匹配

由于点云数据是三维的,所以在完成立体视觉计算前,首先需要选定三维到二维的映射方向,因为获取点云的激光雷达方向是已知的,同时,该方向又是真实的投影方向,所以选取激光雷达方向作为投影方向。

对于点云集合进行特征提取,特征组中包括目标大小 f_1 , 目标与激光雷达的距离 f_2 , 目标长宽高对应的几何尺寸 f_3, f_4, f_5 , 长宽比 f_6 , 宽高比 f_7 , 反射强度比 f_8 。由这些特征值可以构成任意点的特征匹配矩阵,

通过计算该矩阵的特征值^[13]得到点云之间特征匹配关系,从而确定目标是否存在及存在目标的类型。设协方差 A 可由特征矩阵 C 求解,则:

$$A = (r - 1)^{-1} C^T C = \begin{pmatrix} \text{cov}(x, x) & \text{cov}(x, y) & \text{cov}(x, z) \\ \text{cov}(y, x) & \text{cov}(y, y) & \text{cov}(y, z) \\ \text{cov}(z, x) & \text{cov}(z, y) & \text{cov}(z, z) \end{pmatrix} \quad (3)$$

式中:协方差矩阵 A 是对称的,所以可通过特征分解得到特征值。

对于不同的特征值而言,不同的测试点云之间有相似的情况^[14]。为了区分这之间的区别,采用相似度函数 r 进行区分,有:

$$r(f_p, f_q) = \frac{n \sum f_p f_q - \sum f_p \sum f_q}{\sqrt{n \sum f_p^2 - (\sum f_p)^2} \cdot \sqrt{n \sum f_q^2 - (\sum f_q)^2}} \quad (4)$$

式中: P, Q 表示两个测试点; f_p 和 f_q 为两个点对应的特征集合; n 为需要计算的特征参数个数。

2.2 基于立体视觉的特征匹配

为了提高点云数据中目标识别的精度,将立体视觉作为边界条件导入特征匹配算法中,从而限定特征匹配过程中点云的选取范围,起到在点云源头进行滤波的效果。

只有符合待测目标边界条件的点云集才进入点云特征匹配算法,这样就可以降低由于特征数据中非特征量对解算精度的影响。将单纯的符合特征的点云匹配算法转化成符合立体视觉要求范围内的特征识别,则其实现步骤为:

(1) 对拟检测目标种类进行点云数模库构建,形成不同类型目标对应的点云集合,同时按照相应集合进行不同投影角度 θ 的分区;

(2) 在不同投影角度 θ 分区中,完成对合适的选取比率 $n\%$ 进行计算,计算过程中通过网路训练的方式实现,从而得到根据类别划分的选取比率集合 N ;

(3) 对测试点云数据集进行区域选取,然后遍历计算 D_i/D_j 值,并通过比较 D_i/D_j 值与选取比率 $n\%$ 之间的关系,对区域内点云进行筛选;

(4) 识别目标后,采用多种形状的相应点云集合去分析目标特征尺寸特征数,获得特征匹配集合 F ;

(5) 通过公式 (2) 计算特征匹配的协方差矩阵,从

而解算特征值,用于判断目标的准确参数及位姿信息;

(6) 通过相似度函数 $r(f_p, f_Q)$ 的对比计算,消除伪目标点。综上所述,算法对应的程序流程如图 1 所示。

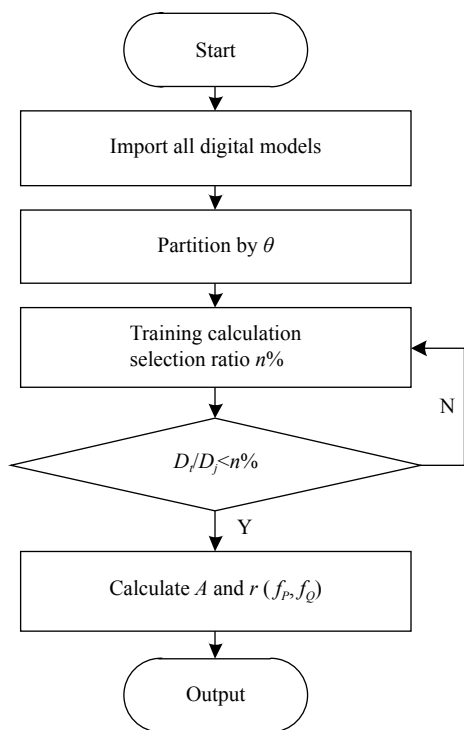


图 1 算法流程图

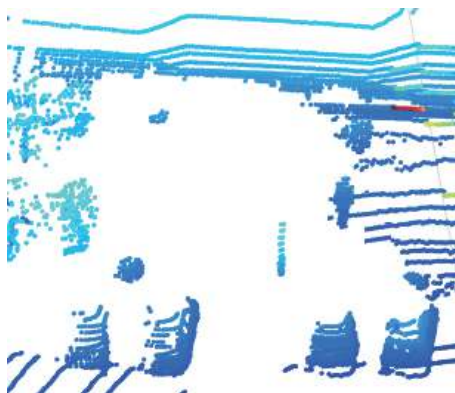
Fig.1 Flow chart of algorithm

3 实验

3.1 点云数据获取

实验采用 ARIES 激光雷达^[15] 采集待测区域的点云数据,处理系统为 Windows 10,主要测试对象为停放路边的车辆,对不同姿态和类型的车辆进行检测与识别。首先,通过激光雷达完成点云数据采集,经过滤波后获得点云由 MATLAB 显示如图 2 所示。实验测试数据已经经过基于时间带滤波处理,即符合测试距离的测试点被保留,其余由于杂散光、远背景噪声等产生的点云直接被消除。测试实验选取的场景是从笔者学校同一位置不同角度得到的,其测试目标也是一致的。

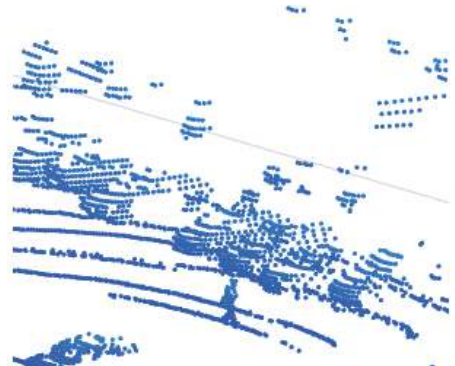
如图 2 所示,为了测试算法对不同点云情况中目标的识别效果,选取了三种典型的测试条件(无遮挡、部分、多目标重叠)。图 2(a)为目标分离度高的无遮挡状态,目标与背景具有一定距离,具有很好的



(a) 情况 1
(a) Situation 1



(b) 情况 2
(b) Situation 2



(c) 情况 3
(c) Situation 3

图 2 三种情况的点云数据

Fig.2 Point cloud data of three situations

区分度;图 2(b)为部分目标存在部位遮挡现象,有少量目标不易识别;图 2(c)为多目标存在重复遮挡的现象,目标均属于不易识别的状态。

3.2 识别概率分析

为了验证文中算法的目标识别能力,将文中算法与未加入立体视觉作为边界条件的点云识别算法进行对比。测试点云为 3.1 节中激光雷达采集得到的三

种典型点云分布。

由测试获得的识别概率可知,不同的测试算法结果不尽相同。对于第一种情况,由于目标分离度高,三种方式获得目标点云都是较为完整的,所以识别精度都在 90% 以上。同时,相比对人的识别精度略低于汽车,因为无论哪种汽车的车型是固定的,其数模边界清晰,所以其识别概率更高。相比第一种情况,第二种情况存在部分遮挡,从实验结果可以看出,识别概率明显降低,三个目标的平均识别率仅有 71.2%,当采用文中算法进行改进后,识别概率提升至 81.1%,但相比无遮挡时识别效果变差。分析认为部分遮挡虽然仍能够提取目标部分特征,但由于整体性判断依旧缺失,导致类别分析能力大大减弱。对于第三种情况而言,多目标重叠使有效信息进一步减小,如果不采用边界条件限制,仅通过特征匹配解算的识别概率均值仅为 48.1%,加入边界条件限制后系统识别概率均值达 71.4%。检出率与识别概率如表 1 所示。

表 1 检出率与识别概率

Tab.1 Detection and recognition probability

Situation	P_a	P_c	P_t
AQ1: B C	92.2%	91.3%	91.6%
	98.3%	98.2%	98.6%
	98.4%	98.4%	98.7%
AQ2: B C	67.5%	66.4%	73.6%
	75.6%	74.2%	84.5%
	70.7%	71.5%	85.4%
AQ3: B C	58.2%	52.3%	69.2%
	61.7%	46.2%	73.3%
	60.4%	45.7%	71.7%

注: Q1、Q2和Q3分别表示三种测试条件; A、B和C分别表示目标人、I型目标车辆和II型目标车辆; P_a 、 P_c 和 P_t 分别表示激光雷达内置算法、传统特征匹配算法和文中算法的识别概率

3.3 位置精度识别

实验中处理的是相同的点云数据,但是采用立体视觉进行优化会形成不同的目标点云集,这些目标点云集会影响到目标存在及类别的判断,其中对种类判断的影响主要是由于保留的点云之间特征值发生了变化,为了量化分析其对特征值的影响,选取了目标数字模型点云中 300 个点,与实际测试点中滤波后最近距离的点进行偏差计算,从而观察算法在引入立体

视觉后对最优匹配点的保留程度,位置精度结果如图 3 所示。

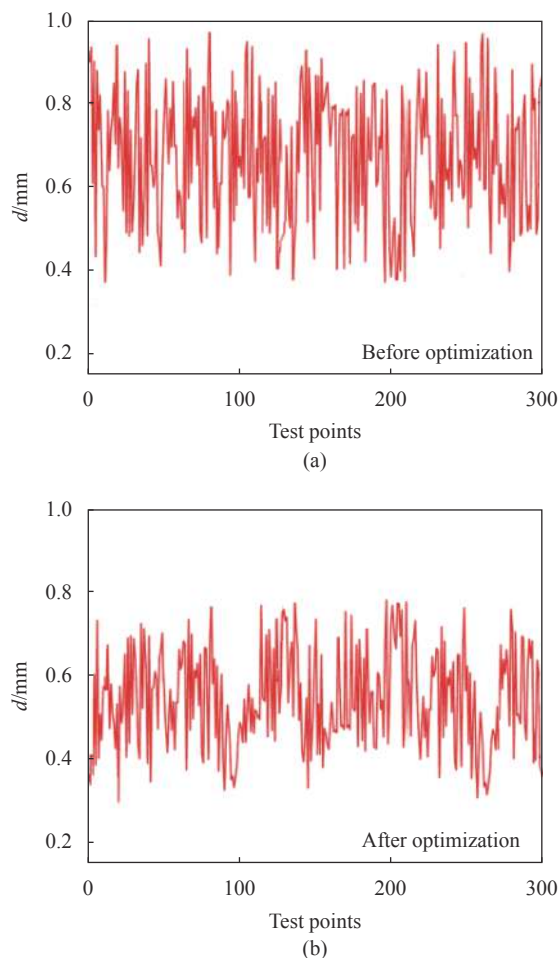


图 3 采用立体视觉优化前后对比

Fig.3 Comparison before and after using stereo vision optimization

图 3 中 d 表示所选取点与数模中相应位置点的绝对距离。由图 3(a) 和 3(b) 对比可知,在采用立体视觉对特征匹配算法进行边界限制与优化前,随机选取的 300 个测试点与数模点云中相应点的位置偏差平均值为 0.74 mm,最大偏差值为 0.38 mm;而采用优化后,位置偏差平均值为 0.55 mm,最大偏差值为 0.21 mm。对比两种方法可知位置偏差平均值提高了 0.19 mm。由此可见,采用算法优化后保留点目标集合与目标数模点云位置更匹配,也印证了优化后目标识别概率会得到提升。

3.4 收敛时间对比

在提高目标识别概率的基础上,还需要对算法的处理速度进行分析,从而分析算法的实用性。反映算

法处理速度的技术参数主要是算法的收敛时间,对比引入立体视觉边界条件前后的数据处理时间就能分析算法的时效性。当目标尺寸不变时,探测视场内点云数据总量的增加会导致处理时间的增长,所以对点云数据总量递增变化时的收敛时间进行分析就能区分算法改进前后的时效性,优化前后的收敛时间如图 4 所示。该测试数据曲线将三种不同条件下的测试数据进行均值后得到。

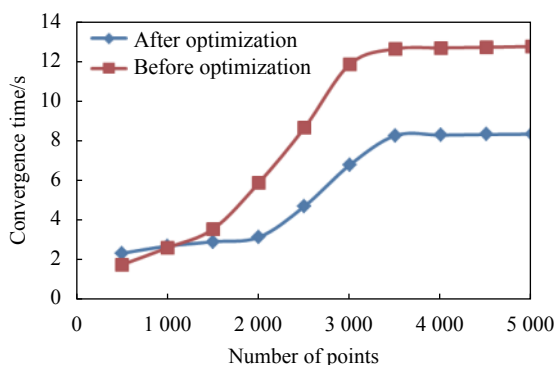


图 4 算法优化前后的收敛时间对比

Fig.4 Comparison of convergence time before and after algorithm optimization

从图 4 结果可以看出,随着测试数据量的增大,采用立体视觉优化后的收敛时间明显减小。在点云总量小于 1000 时,采用立体视觉优化由于同样需要进行数据预处理,所以此时相比传统方式更费时,但随着数据量的不断增大,预处理分类的优势逐渐显现。并且在实际测试的情况中往往数据量在 3000 点以上(由激光雷达测试区域和采样频率决定),所以,将立体视觉作为优化的边界条件可以提高大多数情况下的点云数据处理速度。优化前后的收敛时间分别趋近于 12.76 s 和 8.33 s,由计算可知,优化后收敛周期约为优化前的 65% 左右。但相比之下优化后大约在 3500 点以后才出现较为平稳的收敛曲线,相比优化前在 3000 点左右略慢一些,即优化后对数据总量的依赖性更强,但考虑测试过程中的点云数据总量而言,该差异对算法完成目标识别基本没有影响。

4 结 论

文中针对三维点云目标识别精度低、速度慢的问题,提出了一种基于立体视觉与特征匹配的点云目标

识别算法。该算法在立体视觉的约束条件下,对原有特征匹配算法进行了优化。通过训练不同角度下匹配合适的选取比率提高了目标的检出能力。实验对三种典型状态进行了分析,从识别精度、位置偏差以及收敛时间几个方面进行了对比,验证了采用立体视觉优化后,特征匹配算法得到了较好的改进,使在三维目标识别方面具有更高的适用性。

参考文献:

- [1] Ren S, He K, Girshick R, et al. Faster R-CNN: Real-time target detection in regional planning network [J]. *IEEE Pair Model Analysis and Machinery Information*, 2017, 39(6): 1137-1149.
- [2] Qiao Y J, Jia B M, Jiang J G, et al. Networking method of multi-view stereo-vision measurement network [J]. *Infrared and Laser Engineering*, 2020, 49(7): 20190492. (in Chinese)
- [3] Choi C, Taguchi Y, Tuzel O, et al. Voting-based pose estimation for robotic assembly using a 3D sensor [C]//IEEE International Conference on Robotics & Automation. IEEE, 2013: 1724-1731.
- [4] Mousaviana A, Anguelov D, Flynn J, et al. 3D bounding box estimation using deep learning and geometry [C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 5632-5640.
- [5] Engelcke M, Rao D, Wang D Z, et al. Vote3-deep: Fast object detection in 3D point clouds using efficient convolutional neural networks [C]//Proceedings of 2017 IEEE International Conference on Robotics and Automation. May 29-June 03, 2017, Singapore. New York: IEEE, 2017: 1355-1361.
- [6] Yan Y, Mao Y, Li B. Second: Sparsely embedded convolutional detection [J]. *Sensors*, 2018, 18(10): 3337.
- [7] Yang S Y, Zheng J Y. Vehicle target classification algorithm based on roadside 3D LiDAR [J]. *Transducer and Microsystem Technologies*, 2020, 39(7): 123-126. (in Chinese)
- [8] Guo Y L, Sohel F, Bennamoun M, et al. A novel local surface feature for 3D object recognition under clutter and occlusion [J]. *Information Sciences*, 2015, 293: 196-213.
- [9] Zhang Z, Xu H L, Yi H. A fast point cloud registration algorithm based on key point selection [J]. *Laser and Optoelectronics Progress*, 2017, 54(12): 121002. (in Chinese)
- [10] Yang J H, Liu W, Liu Y, et al. Calibration of binocular vision measurement system [J]. *Optics and Precision Engineering*, 2016, 24(2): 300-308. (in Chinese)

- [11] Gao H, Cheng B, Wang J, et al. Object classification using CNN —based fusion of vision and LiDAR in autonomous vehicle environment [J]. *IEEE Transactions on Industrial Informatics*, 2018, 14(9): 4224-4231.
- [12] Yoshoka M, Suganuma N, Yoneda K, et al. Object classification for autonomous vehicles using omni-directional LiDAR [J]. *The Japan Society of Mechanical Engineers*, 2016, 2016(25): 2307-2316.
- [13] Diyi L, Shogo A, Jiaqi M, et al. Point pair feature-based pose estimation with multiple edge appearance models (ppfmeam) for robotic bin picking [J]. *Sensors*, 2018, 18(8): 2719-2738.
- [14] Aliniya Z, Mirroshandel S A. A novel combinatorial merge-split approach for automatic clustering using imperialist competitive algorithm [J]. *Expert Systems with Applications*, 2018, 117(5): 243-266.
- [15] Lu Ch Q, Yang M F, Wu Y P, et al. Research on pose measurement and ground object recognition technology based on C-TOF imaging [J]. *Infrared and Laser Engineering*, 2020, 49(1): 0113005. (in Chinese)