

基于 Mask R-CNN 实例分割及 FPFH 特征配对的喷涂工件识别方法

葛俊辉¹, 王健¹, 彭以平¹, 李婕瑄¹, 肖昌炎^{1**}, 刘勇^{2*}

¹湖南大学电气与信息工程学院, 湖南 长沙 410082;

²浙江同济科技职业学院, 浙江 杭州 311231

摘要 工件识别在柔性化机器人自动喷涂生产线中至关重要,它是机器人切换喷涂轨迹的重要依据。然而,实际应用中,由于喷涂工件尺寸和种类的多样性、表面的弱纹理性、多视异构件及相似件等情况的存在,准确且高效识别喷涂工件充满挑战性。为此,提出了一种二维(2D)实例分割与三维特征一致性配对的识别方法,即利用基于小样本训练的 Mask R-CNN 深度模型对 2D 工件分割及识别的快速性,再结合 fast point feature histogram (FPFH) 点云特征对局部细节的强区分性,实现对多视异构件及相似件由粗到精的准确识别。在精识别阶段,提出了一种基于 FPFH 特征配对的识别方法。该方法以 intrinsic shape signature 为工件的关键点,以 FPFH 为矢量特征,通过线性相关度配对 FPFH 特征,再以拓扑结构及空间变换关系的一致性为约束验证特征的匹配率,以此作为工件识别的评价标准。实验中,采用 34 种类别 1500 多个工件进行测试,所提方法的识别率高达 99.26%,单工件识别耗时低于 1500 ms。

关键词 机器视觉; 三维视觉感知; 工件识别; Mask R-CNN; fast point feature histogram 特征配对

中图分类号 TP391

文献标志码 A

doi: 10.3788/LOP202259.1415016

Recognition Method for Spray-Painted Workpieces Based on Mask R-CNN and Fast Point Feature Histogram Feature Pairing

Ge Junhui¹, Wang Jian¹, Peng Yiping¹, Li Jiexuan¹, Xiao Changyan^{1**}, Liu Yong^{2*}

¹College of Electrical and Information Engineering, Hunan University, Changsha 410082, Hunan, China;

²Zhejiang Tongji Vocational College of Science and Technology, Hangzhou 311231, Zhejiang, China

Abstract Workpiece recognition is critical for switching painting trajectory in a flexible robotic spray-painting production line. However, due to the wide variety of sprayed workpiece sizes and types, as well as the presence of poor surface texture, multiview dissimilar (MVD) components, and comparable parts, it is difficult to effectively and reliably identify sprayed workpieces in the real production line. In this study, a recognition approach is proposed based on two-dimensional (2D) instance segmentation and three-dimensional feature pairing. Specifically, the high efficiency of the Mask R-CNN learning model was used for 2D workpiece segmentation and coarse recognition based on small sample training; this was followed by the integration of the fast point feature histogram (FPFH) feature for fine recognition, with its strong discrimination of local details for accurately recognizing MVD and similar-topology workpieces. During the fine recognition stage, the intrinsic shape signature method was used as the key point of the workpiece and vectored using the FPFH feature. The extracted feature was then coarsely paired and verified with topological structure consistency and spatial transformation to obtain the pairing rate, which was used as the evaluation criterion to recognize the workpiece. In the experiment, more than 1500 workpieces of 34 categories are used for testing, and the recognition accuracy can reach 99.26% with a running time of less than 1500 ms for a single workpiece.

Key words machine vision; three-dimensional visual perception; workpiece recognition; Mask R-CNN; fast point feature histogram feature pairing

收稿日期: 2022-04-01; 修回日期: 2022-05-23; 录用日期: 2022-05-25

基金项目: 国家自然科学基金(62073128)

通信作者: *Z20420110204@zjtongji.edu.cn; **c.xiao@hnu.edu.cn

1 引言

随着以小批量、多品种、短周期为特点的柔性化机器人自动喷涂需求的不断增加,工件识别在自动喷涂环节变得越来越重要,它是机器人根据工件类别自动调整喷涂轨迹的重要依据。然而,在非结构化场景的机器人喷涂应用中,由于环境的复杂性、喷涂工件尺寸和种类的多样性、表面的弱纹理性、整体形状相似性、不同视角工件的局部异构性等各种非常规情况的存在,高效而准确识别喷涂流水线上在位运动的工件一直是一项充满挑战性的任务。针对非结构化应用场合的工件识别,目前存在很多方法。根据输入数据类型的不同,可分为基于二维(2D)图像的识别方法和基于三维(3D)数据的识别方法两类。

基于 2D 图像的识别方法通过对图像进行处理、特征分析、逻辑判断等实现分类识别。该方法主要分为传统方法和基于深度学习的方法^[1]。传统方法主要以基于局部特征的识别方法为主。该类方法的关键在于提取图像的局部特征,然后通过分类器进行识别。常见的特征有 Haar 小波^[2]、尺度不变特征变换(SIFT)^[3]、局部二进制模式(LBP)^[4]等。基于局部特征的识别方法能在杂乱和部分遮挡的场景中保持较好的鲁棒性,但是其对光照及环境噪声点比较敏感。随着深度学习的不断发展,基于深度学习的识别方法以其精度高、可扩展性强的优势逐渐取代大多数传统识别方法。根据模型结构的不同,基于深度学习的识别方法可分为两类:分步式法和端到端法^[1]。分步式法先得到图像候选区域再对其进行分类和边框回归,通过提取目标区域而后实现目标识别,具有代表性的有 SPPNet^[5]、Faster RCNN^[6]、RFCN^[7]等。端到端法基于回归的思想,在不需区域提名的情况下直接预测目标的边界框并实现目标识别。目前,端到端法主要有 YOLO^[8]、SSD^[9]及一些在它们基础上的改进算法等。虽然基于深度学习的识别方法在性能上表现优异,但其对硬件的算力有较高要求。此外,非结构化场景的不确定性对深度学习网络的泛化能力也是一个巨大考验。整体上,基于 2D 图像的识别方法已经取得了非常优异的成果,但基于 2D 空间的分析处理限制了其对 3D 目标的识别性能,尤其对于多视异构和结构相似的多面体工件。

基于 3D 数据的识别方法通过表征目标表面形貌的 3D 无序点集,使用 3D 数据处理技术实现目标的识别。根据识别原理的差异,可将其分为基于局部特征的方法、基于投票的方法及基于深度学习的方法^[10-11]。基于局部特征的目标识别方法主要包含 4 个过程,分别为特征提取、匹配、假设与验证:即在原点云中提取稀疏但区分性高的点集,并通过特征描述使其矢量化,成为区分容易、保持尺度及方向不变的特性^[12];然后基

于提取的特征,配对目标与模型,得到相对应的点对集合;接着,估算场景目标与模型可能的位姿关系,即变换假设^[13];最后,通过验证假设得到正确的变换关系,达到识别的目的^[14]。基于局部特征的方法在遮挡和杂乱的环境中表现较好,但当离线模型数量较多时,比较耗时。基于投票的目标识别方法基于模型与场景目标的固有属性直接匹配,生成有限的候选姿态集,然后利用先验条件构造奖罚函数投票每一个姿态,选出最优的变换矩阵。基于该原理的识别方法中,最经典的算法是 Drost 等^[15]提出的点对特征(PPF)算法,其利用了点对间最为朴素的特征(距离与法线夹角)构造出有 4 个参数的特征数组,然后结合哈希表进行穷举匹配,利用高效的投票方案得出最优的姿态估计,达到识别的目的。该类方法同样比较耗时,此外,在遮挡和杂乱场景中的表现一般。随着 3D 深度模型的不断发展,基于 3D 深度学习的识别方法逐渐成为研究热点,2017 年,斯坦福大学的 Charles 等^[16]提出了 PointNet 及其改进方法 PointNet++^[17]。之后,RotationNet^[18]等一些性能更好的深度模型相继被提出。相对于传统方法,基于 3D 深度模型的识别方法有诸多优势,如效率高、鲁棒性强等。然而,此类方法的模型训练通常需要大量的 3D 数据,而实际应用的数据并不容易获取。此外,3D 数据的打标也非常困难,并且十分耗时^[19]。

虽然目前已存在大量的识别方法,但是针对小批量、多种类喷涂工件在结构和场景上的复杂性,已有方法存在一些不足。为此,本文提出了一种融合 Mask R-CNN 模型与 fast point feature histogram(FPFH)特征配对的识别方法。该方法利用基于小样本训练的 Mask R-CNN 深度模型对 2D 工件分割及识别的快速性,再结合 FPFH 点云特征对局部细节的强区分性实现对多视异构件及相似件由粗到精的准确识别。

2 结合 Mask-RCNN 与 FPFH 特征配对的工件识别算法

所提基于 Mask-RCNN^[20-21]与 FPFH^[22]特征配对的工件识别算法通过 Mask-RCNN 的实例分割对 2D 工件进行初步识别,然后利用 FPFH 特征一致性对 3D 工件进行精准识别。这样,即可以通过 FPFH 特征对局部细节的强区分特性弥补 Mask R-CNN 模型对局部结构相似、视角异构的 3D 工件识别的不足,又可以通过 Mask R-CNN 粗识别缩小候选范围从而提高基于 FPFH 特征识别的效率。

2.1 算法框架

结合 Mask R-CNN 与 FPFH 特征配对的工件识别算法主要包含基于 Mask-RCNN 的实例分割和基于 3D 特征一致性对的工件识别两个过程,如图 1 所示。首先使用 RGB-D 相机进行数据采集,在 RGB 上进行 Mask R-CNN 实例分割,既可以对工件进行初步识别,又可从背景中分割出 2D 工件。基于 RGB-D 相机内部

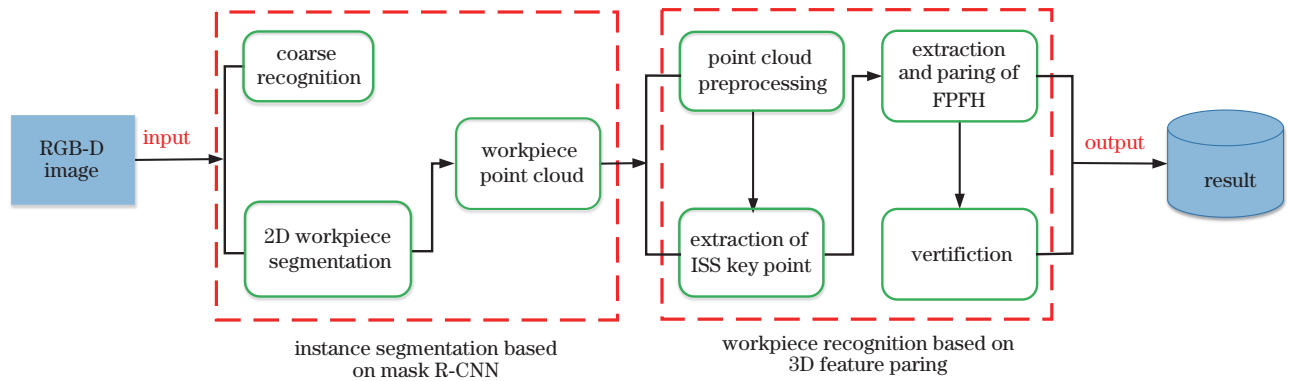


图 1 结合 Mask R-CNN 和 3D 特征配对的工件识别算法框架

Fig. 1 Algorithm framework of workpiece recognition combining Mask R-CNN and 3D feature paring

映射关系,可将分割的 2D 工件转换为 3D 点云。最后,利用 intrinsic shape signature (ISS) 关键点的重复性及 FPFH 特征的可区分性,进一步对点云工件进行精确识别。

2.2 基于 Mask R-CNN 的 2D 工件识别与 3D 分割

RGB-D 相机采集完工件数据后,在 2D 彩色图上采用 Mask R-CNN 模型对工件进行实例分割,一方面实现工件的初步识别,另一方面将 2D 工件从背景中分割出来,转换为 3D 点云后对工件进一步识别。

2.2.1 Mask R-CNN

Mask R-CNN 是在目标检测算法 Faster R-CNN^[6] 的基础上改进的实例分割算法。首先,输入图像经过主干网络的处理提取图像特征;然后基于 region proposal network (RPN) 生成候选感兴趣区域 (ROI);最后采用头部网络对候选 ROI 进行 Mask 表示 (即分割结果的掩膜)、位置坐标和类别预测。Mask R-CNN 中的很多超参数直接采用 Faster R-CNN 中已经训练好的值,所以简单直观、易于使用。此外,Mask R-CNN 能够达到 5 frame/s 的执行效率,在低耗时的同时还能

保持很高的分类、检测及分割的准确率,这些为机器人自动喷涂 2D 工件的识别及分割提供了有利先机。

2.2.2 基于 Mask R-CNN 的 2D 工件识别

本实验组直接应用 Mask R-CNN 进行工件的初步识别和 Mask 分割。首先,对网络进行训练,采用 Ge 等^[23] 的方法将相机安装在喷涂流水线运动工件的两边,连续采集工件的多视角数据流,每类工件标记 20 张,通过透视变换 (5 个不同) 和灰度空间增强 (3 个不同阈值) 的方式扩展到 300 张作为样本集。然后,对样本几何进行训练,训练中,采用权重衰减的优化策略,衰减参数为 0.0001。RPN 中锚框包含 5 种不同面积,3 种不同长宽比,其训练权重与 Mask R-CNN 共享。最后,对训练的模型进行测试。通过 RPN 产生 800 个候选 ROI,并对候选 ROI 进行类别分类、检测框回归、Mask 掩码分类。现场采集工件的实例分割结果如图 2(a) 所示,其中工件的类别 (整数) 和打分 (百分数) 被标记在定位框左上方,工件的定位结果用矩形框标注,分割结果采用带颜色 Mask 标记。

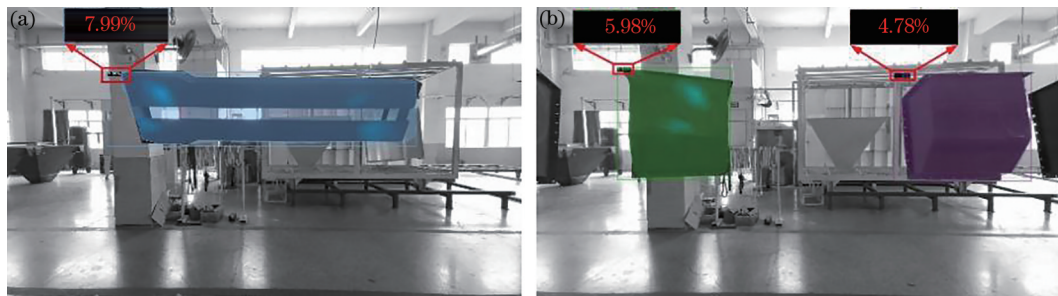


图 2 Mask R-CNN 实例分割结果。(a) 正确识别示例;(b) 错误识别示例

Fig. 2 Results of Mask R-CNN instance segmentation. (a) Example of correct identification; (b) example of misidentification

从图 2(a) 可以看出,工件被准确定位和识别,但是在分割上存在一定偏差,工件空洞部位的背景不能去除。此外,也存在误识别的情况,如图 2(b) 所示。由于 Mask R-CNN 在 2D 图像对 3D 工件进行识别,图像的几何变换及单视角约束不可避免降低了多视角异构工件的识别率,尤其当不同类别工件存在局部结构

相似时。图 2(b) 中,导爆罩 2 (类别 5) 因与导爆罩 1 (类别 4) 结构相似而被误识别。

综上所述,基于 Mask R-CNN 的实例分割在应用中虽然表现突出,但依然存在一些问题:1) 几何变换问题。深度网络本质上是不能对目标特征保持几何空间的不变性的,虽然网络的 Max pooling 层在某种程度上

对特征具有平移不变性,但是对旋转变换的特征不变性映射也是非常必要的^[23]。2)对于多尺度的小目标及局部结构相似的目标,尤其当物体表面纹理较弱时,Mask R-CNN提取的特征不够稳定,区分性不强,这直接影响实例分割的鲁棒性。3)此外,在非结构化场景中,环境的变化对Mask R-CNN的泛化能力也是一个考验。对此,本实验组以Mask R-CNN的分类结果作为初步识别结果,并将得分高于设定阈值 T_c 的结果作为候选类别输出。同时,根据实例分割的Mask结果,将从背景中分割出的2D工件转换为3D点云,进一步进行精确识别。

2.3 基于FPFH特征配对的3D工件识别算法

由第2.2节可知,当不同类别工件存在局部结构相似时,Mask R-CNN算法很难将这些工件区分开。而FPFH作为一种经典的特征描述算法,对局部细节具有很强的描述性和区分性,对于背景干扰及遮挡有较高的鲁棒性^[22],基于此优势,该算法被广泛用于局部到全局的配准和识别应用。虽然FPFH特征具有较强的区分性,但它通常与random sample consensus (RANSAC)算法组合的方式进行目标配准或者识别,该过程的时间复杂度非常高^[24-25],通常难以满足实际应用需求。

为了进一步识别Mask R-CNN难以区分的工件,同时保证算法的运行效率,本实验组采用FPFH特征配对的方式实现工件的准确快速识别,即在初步工件识别定位的基础上,利用FPFH特征对局部细节的强区分性,采用关键点检测、矢量化、特征配对的思路,以配对率为评价指标对粗分类工件进行精确识别。针对相似工件特征误配对的问题,提出了一种特征对分步验证算法,即以矢量化特征的线性相关度为特征初配对标准,再以拓扑结构及空间变换关系的一致性为约束,实现工件3D特征的准确配对。该算法主要包含点云预处理、ISS关键点提取、FPFH特征计算及配对、基于特征配对的工件识别等步骤。

2.3.1 点云预处理

数据采集设备为RGB-D相机,其获取的彩色图和深度图在图像坐标系可以实现一一对应^[26],因此可将2D彩色工件图转化为3D点云。因为Mask R-CNN不能分辨该工件空洞区域的背景,所以转换的工件点云含有噪点。本实验组通过条件滤波,即根据工件到相机的Z方向距离滤除背景。

此外,基于RGB-D相机获取的点云:一方面,受采集设备性能、工件表面特性的影响,点云数据集中存在噪点或离群点,尤其是特征区分度比较明显的边缘区域;另一方面,RGB-D相机采集的数据属于稠密点云,点云数量较大,而特征提取过程需要对每个点云进行计算,这样不仅非常耗时,还会引起计算机算力的不足。针对上述情况,需要对点云进行滤波和降采样预处理。考虑到点云中的噪点或离群点分布稀疏,且相

对孤立,所以采用统计滤波消除密度差异相对较大的无效点云,即以点云邻域内高斯分布的均值和方差判定离群点并滤除。然后,采用体素滤波对点云数据进行下采样处理,以稀释点云密度,提高算法的执行效率。

2.3.2 ISS关键点提取

虽然预处理后点云及噪点的数量得以降低,但点云不具有重复性及可靠性,这样势必影响基于点云所提取特性的有效性。因此,本实验组在点云中提取稀疏而区分度高的点云作为关键点,然后计算其矢量特征,这样不仅可以进一步减少点云数量,还能保证所提特征的强区分性和可靠性^[27]。

ISS是一种具有可区分性、可在异源性目标点云中重复检测的关键点。它通过与邻域几何信息建立联系,并利用特征值之间的关系来表征邻域点间的分布状态^[28],即以点云 p_i 与其邻域的欧氏距离反比为权重,构建协方差矩阵,并计算其特征值 $\lambda_i^1, \lambda_i^2, \lambda_i^3$,若它们满足以下不等式,则 p_i 为ISS关键点。

$$\begin{cases} \frac{\lambda_i^2}{\lambda_i^1} \leq \epsilon_1 \\ \frac{\lambda_i^3}{\lambda_i^2} \leq \epsilon_2 \end{cases}, \quad (1)$$

式中: ϵ_1, ϵ_2 为经验阈值。以邻域协方差矩阵特征值的关键点是具备几何意义的。特征值的大小实际上代表椭球轴的长度,而椭球的形态则是对邻域点分布状态的抽象总结。这样,ISS关键点多为邻域内特征稳定且显著的点,它们具有可区分性,可在不同视角甚至异源点云中重复检测。

2.3.3 FPFH特征提取及配对

工件识别的本质是特征的比对,即在线工件与模型库的特征匹配,匹配度最高的工件即为识别结果。对于在线场景工件,通过上文的实例分割已经初步获取其候选类别和分割的3D点云。假设工件点云为 P_w ,对其进行滤波预处理,然后检测ISS关键点并计算相应的33维FPFH特征向量 $f_i (0 < i \leq m)$,其中 m 表示ISS关键点的数量,这样,最终可得到 $33 \times m$ 的FPFH特征集合 F_w 。在计算FPFH特征时,应采用原点云作为关键点的邻域,以保证所提取特征更具有局部表现性和区分性。对于离线工件模型,采用与在线工件相同的方式,提取每个工件的FHFP特征 $F_m^j (0 < j \leq N)$,其中 N 表示工件的种类,建立离线特征库 F_m :

$$F_m = \{F_m^1, F_m^2, \dots, F_m^N\}. \quad (2)$$

接下来,配对在线工件与离线模型的FPFH特征。上文已经提到,已经初步确定在线工件的候选类别,所以只需与对应的工件模型比对应的FHFP特征即可,这样可以减少算法的执行时间。假设在线工件的FPFH特征集为 F_w ,候选类别对应的离线库中的

FPFH特征集为 $C_{ii}(0 < ii < kk)$, 分别估算 F_w 与 C_{ii} 间 FPFH 特征的对应关系。理论上, 在线与离线工件点云相同点对应的 FPFH 特征应该近似相同, 通常通过计算特征向量的欧氏距离来判定它们是否属于同一点, 然而受点云拓扑结构、稀疏程度、噪点等因素的影响, 该衡量标准存在较大的误差。针对该问题, 采用余弦相似度作为两个 FPFH 特征向量相似度的判断标准, 其定义如下: 在 n 维空间的向量 $a = (a_1, a_2, \dots, a_n)$ 和向量 $b = (b_1, b_2, \dots, b_n)$, 假设它们的空间夹角为 θ , 则余弦值为

$$\cos \theta = \frac{\sum_{i=1}^n a_i \times b_i}{\sqrt{\sum_{i=1}^n a_i^2 \times \sum_{i=1}^n b_i^2}} \quad (3)$$

余弦值越大表示 θ 越小, 则向量 a 和向量 b 越相似, 当余弦取值为 1 时, 表示 a 与 b 线性相关, 即二者的相似度最大。基于余弦相似度, 在 C_{ii} 中通过 KD 树最近邻查找法 (FLANN) 分别搜索与 F_w 中 f_i 配对的 FPFH 特征, 即可确定在线工件与候选类别工件间的 FPFH 特征配对关系。

2.3.4 基于 PPFH 特征配对率的工件识别

在提取并配对场景工件点云与离线模型点云的 FPFH 特征后, 理想情况下根据配对特征的数量即可识别工件。然而, 由于工件局部结构和特征的相似性, 大量误匹配的特征对会导致基于特征对数量的识别方式失效。对此, 很多学者采用 RANSAC 去除错误点对, 并在此基础上通过估算配准误差识别工件。该方法虽然取得不错的效果, 但是计算复杂度较高, 尤其当点集较大时, 识别过程非常耗时。针对此问题, 本实验组提出了一种基于仿射不变性与误差最小化的特征对验证算法, 该算法利用 4-points congruent sets (4PCS)^[29] 中 4 点集全等一致性的思想验证点集的空间拓扑结构, 以此计算场景工件与模型工件的相对变换矩阵, 再以配准误差为度量验证配对率, 达到识别工件的目的。

4PCS 算法基于 RANSAC 算法框架, 通过构建与匹配空间一致 4 点对的方式减少搜索过程, 从而加快特征对的匹配速度^[29]。该算法基于共面 4 点的空间仿射一致性搜索符合条件的对应点对, 使用 largest common pointset (LCP) 策略寻找符合条件的对应点对, 得到最优匹配结果。详细原理如下: 给定源点集 P 及目标点集 Q 。在 P 中选择共面 4 点 a, b, c, d , 由此构成共面的 4 点基 $B \subset P$ 。由 B 可确定两个点对, 拓扑信息如图 3 所示。该拓扑结构用 r_1, r_2 表示:

$$\begin{cases} r_1 = \frac{\|a - e\|}{\|a - b\|} \\ r_2 = \frac{\|c - e\|}{\|c - b\|} \end{cases} \quad (4)$$

根据仿射不变性原理, 比例因子 r_1 和 r_2 在点云旋转和平移变化中保持不变, 在 Q 中搜索共面四点 (a', b', c', d') , 使其构成的 4 点基 $B' \in Q$ 的空间拓扑结构与 B 相等, 则 B 与 B' 中对应元素满足全等一致性, 即 B 中 4 点与 B' 中 4 点一一对应。接着, 根据 B 和 B' 估算 P 和 Q 间的刚体变换矩阵, 以此计算二者间的对齐误差。然后基于上述过程迭代搜索其他全等一致性的 4 点基, 直到对齐误差最小使得点云 P 和 Q 足够接近。

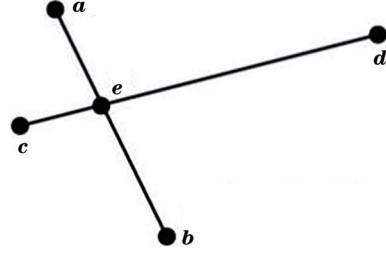


图 3 仿射一致性拓扑结构图

Fig. 3 Topological structure of affine consistency

4PCS 算法中, 输入的两个点云没有对应关系, 进而需要迭代搜索对应关系后才能实现配准和识别, 所以在计算上依然非常耗时。而场景工件点云和离线工件模型在提取 FPFH 特征后, 采用余弦相似度已经完成初步配对。此时只需以 4PCS 的仿射不变性作为约束, 通过一次点对遍历即可实现错误点对的快速剔除, 能够大大提高工件的识别效率。基于上述分析, 工件识别算法的具体实现如下: 1) 基于 4PCS 的全等一致性原则, 寻找在线工件特征集 F_w 与候选类别工件特征集 C_{ii} 中的全等 4 点基。在 F_w 中选择 4 个点构建 4 点基 B , 根据 F_w 和 C_{ii} 的 FPFH 特征配对关系, 在 C_{ii} 中快速的寻找对应的 4 点基 B' , 如果 B 和 B' 满足全等一致性, 则二者构成一个点对集合 M 。2) 根据 M 中的点对关系, 估算在线工件点云 P_w 与工件模板点云 P_m 间的刚性变换矩阵 T , 并根据 T 计算二者间的对齐误差 E_r 。

$$E_r = \frac{1}{N_p} \sum_{k=1}^{N_p} \|q_k - T p_k\|, \quad (5)$$

式中: p_i 表示 P_w 中的点; q_i 表示 P_m 中与 p_i 变换后最近的点; N_p 表示 P_w 的点云数量。3) 在 F_w 和 C_{ii} 建立其他全等一致性的 4 点基, 并依照上述步骤估算 T 并计算点云 P_w 和 P_m 的对齐误差。对齐误差最小的全等 4 点基所在的候选类别即为识别结果。

如果在线工件和候选类别工件的最小对齐误差大于给定阈值 T_c , 则与离线特征库其余工件进行 FPFH 特征配对, 按照上述过程识别工件, 算法具体实现的伪代码如算法 1 所示, 其中 Q_s 为工件的离线特征, $Q_s \subset C_{ii}$, K 表示粗识别后候选工件的数量。与基于 4PCS 识别算法相比, 该方法时间复杂度降到了 $O(n)$, 其中 n 为点云数量。

Algorithm 1: Recognition algorithm based on FPFH feature pairing

Input: FPFH feature F_w of workpiece online, FPFH feature $Q_s (s = 1, \dots, K)$ of candidate workpiece offline, error threshold T_e .

Output: Recognition result I_c

```

1  $t \leftarrow 10000$ ;
2  $c_1 \leftarrow \frac{\text{Num}(F_w)}{4}$ ;
3  $F'_w \leftarrow F_w$ ;
4 while  $i \leq K$  do
5   while  $j \leq c_1$  do
6      $\forall p_b \in F'_w (b = 1, 2, 3, 4)$ , construct four-point basis  $B$ , and determine the
       corresponding  $B'$  in  $Q_i$  based on the matching of  $F_w$  and  $Q_i$ ;
7     if  $B' = B$  then
8       Compute the transform matrix  $T$  between  $F_w$  and  $Q_i$ ;
9        $E_r \leftarrow \frac{1}{N} \sum_{k=1}^N \|q_k - p_k T\|$ ,  $q_k \in F_w, p_k \in Q_i$ ;
10      if  $E_r \leq t$  then
11         $t_j \leftarrow E_r$ ;
12      end
13       $F'_w \leftarrow F_w - p_b, (b = 1, 2, 3, 4)$ ;
14    else
15       $F'_w \leftarrow F_w - p_b, (b = 1)$ ;
16    end
17     $j \leftarrow j + 1$ ;
18  end
19   $i \leftarrow i + 1$ ;
20 end
21  $E_r \leftarrow \min(t_j), (j = 1, \dots, K)$ ;
22 if  $E_r \leq T_e$  then
23   return  $I_c$ , where  $c$  denotes the recognition result related to  $E_r$ ;
24 else
25   Compute the pairing relationship between  $F_w$  and  $Q'_s$ ;
26 end

```

3 实验结果与分析

为了验证所提算法的性能,从定性和定量两方面评估其对喷涂工件的识别效果。

3.1 实验条件

实验数据来自校企合作项目“面向机器人自动喷涂的 3D 工件在线识别系统开发”的实际喷涂生产线,共有 34 类 1500 多件待识别的喷涂工件。图 4 展示了部分工件的样图,工件的尺寸及整体结构存在较大差异。实验中,工件数据的采集选用 Microsoft 公司发布的 Azure Kinect RGB-D 相机,根据工件大小及现场对工作距离要求,相机的采集模式设置为 WFOV-1080。

实验中,算法实现主要采用 C/C++, 部分采用 Python 编程,深度学习部分基于 Tensorflow 完成。算法在主控计算机上测试,其配置如下: 3.5 GHz 4-核

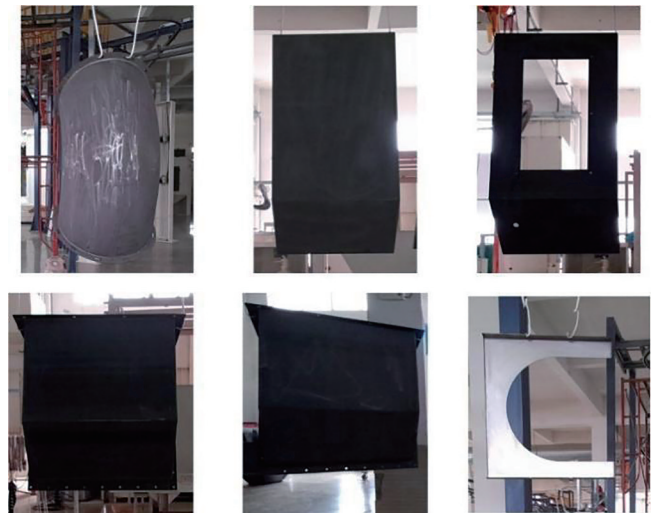


图 4 工件样本

Fig. 4 Workpiece samples

Xeon E-5v3 CPU、32 GB RAM、Geforce 1080Ti GPU, Windows 10 操作系统。

3.2 定性评估

为了定量评估算法性能,采用不同类别的工件展示识别算法的具体实现过程,并对结果进行分析。3种局部结构相似的工件的识别过程如表1所示。第2行为基于事件触发采集方式的工件采集结果(图片较大,只展示原图的部分裁剪图),每幅图中只含一个工件,以保证工件的单一性识别。第2行是基于Mask R-CNN实例分割的结果。虽然工件被精确定位和准确分割,但是由于Mask R-CNN在图像的几何变换及单

视角约束方面的缺陷,对视角异构和局部相似的3D工件经常存在误识别的情况,如表1第2行所示:导爆罩1(类别4)被识别为导爆罩2(类别5);上侧板1(类别1)与上侧板2(类别2)虽然被正确识别,但由于它们的形貌及结构非常相近,二者的预测分数不高,在某些视角下,如侧方位,二者因形貌一样而难以区分。对此,将Mask R-CNN的工件分类结果作为初步识别,基于其分割结果将2D工件转换为3D点云,基于点云特征进行精确识别,如表1第4行所示,根据RGB-D相机的内部映射关系,将分割的2D工件转换为相应的3D点云。

表1 工件识别过程及结果

Table 1 Process and result of recognition with different workpiece

Process	Detonating cover 1	Upper plate 1	Upper plate 2																																										
Original image																																													
Instance segmentation																																													
3D point cloud																																													
Filtering preprocessing																																													
Extraction of ISS key points																																													
Recognition based on feature pairs verification	<table border="0"> <tr> <td>Online Feature</td> <td>Offline Feature</td> <td>Error</td> <td>Result</td> </tr> <tr> <td>4</td> <td>5</td> <td>3.45</td> <td rowspan="3">0.89 4</td> </tr> <tr> <td></td> <td>4</td> <td>0.89</td> </tr> <tr> <td></td> <td>17</td> <td>10.54</td> </tr> </table>	Online Feature	Offline Feature	Error	Result	4	5	3.45	0.89 4		4	0.89		17	10.54	<table border="0"> <tr> <td>Online Feature</td> <td>Offline Feature</td> <td>Error</td> <td>Result</td> </tr> <tr> <td>1</td> <td>1</td> <td>0.68</td> <td rowspan="3">0.68 1</td> </tr> <tr> <td></td> <td>2</td> <td>2.35</td> </tr> <tr> <td></td> <td>4</td> <td>14.34</td> </tr> </table>	Online Feature	Offline Feature	Error	Result	1	1	0.68	0.68 1		2	2.35		4	14.34	<table border="0"> <tr> <td>Online Feature</td> <td>Offline Feature</td> <td>Error</td> <td>Result</td> </tr> <tr> <td>2</td> <td>2</td> <td>1.06</td> <td rowspan="3">1.06 2</td> </tr> <tr> <td></td> <td>1</td> <td>1.98</td> </tr> <tr> <td></td> <td>34</td> <td>15.67</td> </tr> </table>	Online Feature	Offline Feature	Error	Result	2	2	1.06	1.06 2		1	1.98		34	15.67
Online Feature	Offline Feature	Error	Result																																										
4	5	3.45	0.89 4																																										
	4	0.89																																											
	17	10.54																																											
Online Feature	Offline Feature	Error	Result																																										
1	1	0.68	0.68 1																																										
	2	2.35																																											
	4	14.34																																											
Online Feature	Offline Feature	Error	Result																																										
2	2	1.06	1.06 2																																										
	1	1.98																																											
	34	15.67																																											

然而,由于Mask R-CNN无法区分非实体工件的背景,转换的3D工件点云包含场景信息,如上侧板2

的背景也被转换为点云。此外,受RGB-D相机深度分辨率的限制,3D点云中存在诸多噪点。因此,识别前

需对工件点云进行滤波预处理。经过条件滤波、统计滤波以及体素滤波,结果如表 1 第 5 行所示。从表 1 第 5 行可以看出,背景点云及边缘噪点被去除,点云密度被稀疏化且均匀化。接下来,提取工件的 ISS 关键点,如表 1 第 6 行所示,被提取的具有显著性、重复性的 ISS 关键点(红色点标注)相对均匀地分布在工件点云上,这样不仅能够保持工件原有的整体形貌特征,还能够通过代替近万工件点云来降低运算量。然而,ISS 关键点在尺度空间及方向维度上不具备区分性,所以无法直接用于识别。采用 FPFH 特征对其进行描述,使其具备大小和方向后采用特征一致性实现工件识别,如表 1 第 7 行所示,待识别工件的 FPFH 特征与候选工件(实验中选用预测得分最高的三类作为候选类别)的离线 FPFH 特征配对后,通过 4PCS 思想进行仿射一致性验证并估算变换矩阵,对齐误差最小的工件类别即为识别结果。从表 1 第 8 行可以看出,工件被准确分类。

3.3 定量评估

基于现场采集的喷涂工件数据,从识别率和计算效率两方面定量评估算法性能,并与其他识别算法比较。

3.3.1 识别率

由于喷涂工件包含多个类别,每个类别具有多个相同的实例,所以喷涂工件的识别也是分类问题。因此,采用目标分类中的精度 P_r 、召回率 R_c 和综合评价 F_1 作为工件识别的评价指标,如图 5 所示:其中 N_{TP} 表示正相关预测为正的工件数,即正确识别的工件数; N_{FP} 表示负相关预测为正的工件数,即被误识别的工件数; N_{FN} 表示正相关预测为负的工件数,即错误识别的工件数; N_{TN} 表示负相关预测为负的工件数。基于此, P_r 、 R_c 和 F_1 的表达式如图 5 中公式所示,其中 P_r 表示正确预测的正相关占实际预测为正相关的比例, R_c 表示正确预测的正相关占实际正相关的比例。由于 P_r 和 R_c 只反映分类器性能的两个方面,单一依靠其中一个难以全面地评价分类器的性能,所以引入 F_1 综合评价指标平衡 P_r 和 R_c 的影响,较为全面地评价分类器性能。

基于上述评价指标,统计算法的识别率。对于在

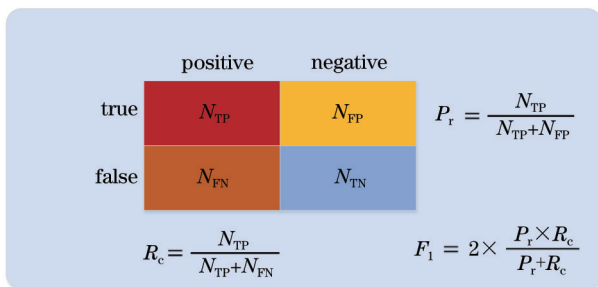


图 5 评价指标

Fig. 5 Evaluation index

线采集的喷涂工件,采用结合 Mask R-CNN 实例分割和 FPFH 特征配对算法对其进行识别,并将结果标注在输入图像的左上角,结果如表 2 所示。从表 2 可以看出,喷涂工件的识别率达到很高的水平,其中精度为 99.10%,召回率为 99.43%,综合指标达到 99.26%。然而,受工件表面材质的影响及消费级 RGB-D 相机深度精度的限制,采集的工件点云存在缺失的情况,进而引起工件的误识别。

表 2 识别结果

Table 2 Recognition results

unit: %

Parameter	P_r	R_c	F_1
Value	99.10	99.43	99.26

为了更深入评估算法的性能,对 8 类具有代表性的工件,包括结构相似类别不同的工件、类别接近尺寸不同的工件等,以混淆矩阵的方式统计识别结果,具体如图 6 所示,其中横轴代表真实的工件类别,纵轴代表工件的预测结果,各个元素的值表示对预测结果的统计。很明显,对角线被黑色填充,填充的元素即为不同种类工件的识别正确率。从图 6 可以看出,大部分工件的识别率为 100%,个别类别工件出现误识别的情况,如长侧板 2 及导爆罩 1。其主要原因在于工件表面色泽偏暗,RGB-D 采集的深度图受此影响出现空洞甚至缺失的情况,进而影响点云工件的特征提取及后续的识别结果。

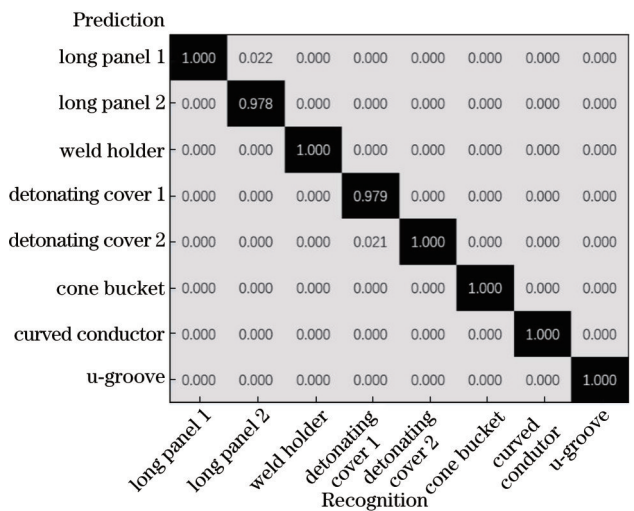


图 6 不同工件的识别结果

Fig. 6 Recognition results with different workpieces

为了验证所提算法的优势,与一些经典的识别算法进行了比较,包括传统的基于点云特征的识别算法和最近发布的基于 3D 深度学习的工件识别算法,如基于 SHOT 特征的识别算法^[30]及基于 PointNet++ 的识别算法^[17],其中 SHOT 是一种通过构建局部参考系将邻域点的空间位置和几何特征结合的点云特征,该特征因对噪点和密度变化具有较强鲁棒性而被广泛用于

3D 物体识别^[31]。PointNet++ 利用分层结构的思想,通过局部邻域的采样表征方式学习点云特征。基于 PointNet++ 提取的点云特征,可实现 3D 物体的识别。此外,所提算法是基于 Mask R-CNN 实例分割和 FPFH 特征而实现的,所以也分别与 Mask R-CNN 及基于 FPFH 特征的工件识别算法进行了对比。对比结果如图 7 所示。整体而言,所提算法识别率最高,经典的基于 FPFH 和 SHOT 点云特征的识别算法与所提算法相当,而基于 Mask R-CNN 和基于 PointNet++ 的识别算法效果偏弱。具体而言,Mask R-CNN 是基于在

2D 图像的识别,对于多视异构的 3D 工件存在单视角局限。PointNet++ 算法虽然对点云有较好的特征学习能力,但是对环境变化及噪点很敏感,实际应用中识别率受到影响。基于 FPFH 和 SHOT 点云特征传统识别方法对噪点、遮挡、密度变化等具有较强的鲁棒性,所以识别率较高。但是,此类算法的运算效率和工件类别数量呈正相关,所以通常比较耗时。相比而言,所提算法通过 Mask R-CNN 的初识别缩小候选工件类别,然后利用 FPFH 特征较强的稳定性及区分性进一步精确识别 3D 点云工件,最终达到稳定高效的识别性能。

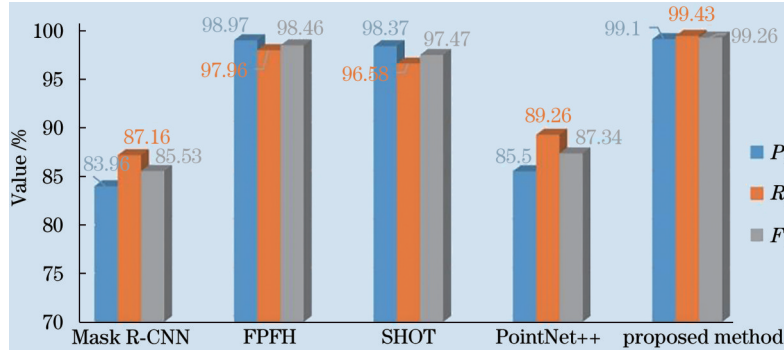


图 7 识别方法对比

Fig. 7 Comparison with different methods

3.3.2 计算效率

喷涂工件的识别是面向于机器人自动喷涂生产线的,实际的喷涂作业对实时性有一定的要求,所以测试算法的计算效率也是算法性能评估的一项重要指标。采用导爆罩 1、焊接架、上侧板 2 等 3 种类别和尺寸差异巨大的工件进行实验,分步测试所提算法的耗时。具体情况如表 3 所示,其中 Num 代表点云数量。从表

3 可以看出,尺寸不同的工件识别耗时存在差异,工件的点云数量越多,其识别时间越久。主要原因在于精确识别阶段,ISS 关键点的提取和 FPFH 特征的计算需要处理各单个点云,点云数量越多,处理的过程越耗时,这也是精确识别相对于初识别更加耗时的原因。整体而言,工件识别时间不超过 1500 ms,可以满足生产线对喷涂工件在线识别的需求。

表 3 算法耗时

Table 3 Running time

Workpiece	Dimension / (m×m×m)	Num / 10 ⁴	Coarse /ms	Fine /ms	Time /ms
Detonating cover 1	0.85×0.82×0.50	11.90	315	887	1202
Weld holder	2.40×0.65×0.03	17.65	332	1064	1396
Upper plate 2	0.75×0.75×1.35	5.92	310	524	834

4 结 论

针对小批量多种类自动喷涂生产线对工件识别的需求及对多视异构件及相似件识别不准确的难题,提出了一种基于 Mask R-CNN 实例分割与 FPFH 特征配对的识别方法。该方法先基于小样本训练的 Mask R-CNN 深度模型快速对工件类别进行初步判定,然后采用一种 FPFH 特征配对的识别方法实现点云工件的精确识别。整体上,基于 Mask R-CNN 的粗识别因减少工件的候选类别而提高基于 FPFH 特征精识别的效率,而 FPFH 特征对局部细节的强区分性又弥补了 Mask R-CNN 对多视角异构件和相似件识别的不足。实验结果表明,所提算法对 30 多种喷涂工件的识别达

到较高水平,其中精度、召回率、 F_1 分别为 99.10%、99.43%、99.26%,单工件识别耗时低于 1500 ms,可以满足机器人自动喷涂生产线在线识别的需求。

参 考 文 献

- [1] Zou Z X, Shi Z W, Guo Y H, et al. Object detection in 20 years: a survey[EB/OL]. (2019-05-13)[2021-05-06]. <https://arxiv.org/abs/1905.05055>.
- [2] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features[C]//Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, December 8-14, 2001, Kauai, HI, USA. New York: IEEE Press, 2001.
- [3] Lowe D G. Object recognition from local scale-invariant

- features[EB/OL]. (1999-09)[2021-02-01]. <https://www.cs.ubc.ca/~lowe/papers/iccv99.pdf>.
- [4] Ojala T, Pietikainen M, Maenpaa T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(7): 971-987.
- [5] He K M, Zhang X Y, Ren S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(9): 1904-1916.
- [6] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137-1149.
- [7] Dai J F, Li Y, He K M, et al. R-FCN: object detection via region-based fully convolutional networks[EB/OL]. (2016-05-20)[2021-05-06]. <https://arxiv.org/abs/1605.06409>.
- [8] 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, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 779-788.
- [9] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector[M]//Leibe B, Matas J, Sebe N, et al. *Computer vision-ECCV 2016. Lecture notes in computer science*. Cham: Springer, 2016, 9905: 21-37.
- [10] Guo Y L, Bennamoun M, Sohel F, et al. 3D object recognition in cluttered scenes with local surface features: a survey[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, 36(11): 2270-2287.
- [11] Hodaň T, Michel F, Brachmann E, et al. BOP: benchmark for 6D object pose estimation[M]//Ferrari V, Hebert M, Sminchisescu C, et al. *Computer vision-ECCV 2018. Lecture notes in computer science*. Cham: Springer, 2018, 11214: 19-35.
- [12] Salti S, Tombari F, di Stefano L. A performance evaluation of 3D keypoint detectors[C]//2011 International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission, May 16-19, 2011, Hangzhou, China. New York: IEEE Press, 2011: 236-243.
- [13] Taati B, Greenspan M. Local shape descriptor selection for object recognition in range data[J]. *Computer Vision and Image Understanding*, 2011, 115(5): 681-694.
- [14] 郭裕兰. 点云局部特征描述与 3D 目标重建识别技术研究[D]. 长沙: 国防科学技术大学, 2015.
Guo Y L. Local feature description 3D modeling and object recognition in point-clouds[D]. Changsha: National University of Defense Technology, 2015.
- [15] Drost B, Ulrich M, Navab N, et al. Model globally, match locally: efficient and robust 3D object recognition [C]//2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 13-18, 2010, San Francisco, CA, USA. New York: IEEE Press, 2010: 998-1005.
- [16] Charles R Q, Hao S, 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, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 77-85.
- [17] Charles R Q, Yi L, Su H, et al. Pointnet++: deep hierarchical feature learning on point sets in a metric space [C]//Advances in Neural Information Processing Systems 30, December 4-9, 2017, Long Beach, CA, USA. [S.l.: s.n.], 2017: 5099-5108.
- [18] Kanezaki A, Matsushita Y, Nishida Y. RotationNet: joint object categorization and pose estimation using multiviews from unsupervised viewpoints[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 5010-5019.
- [19] 刘通, 高思洁, 聂为之. 基于多模态信息融合的多目标检测算法[J]. *激光与光电子学进展*, 2022, 59(8): 0815002.
Liu T, Gao S J, Nie W Z. Multitarget detection algorithm based on multimodal information fusion[J]. *Laser & Optoelectronics Progress*, 2022, 59(8): 0815002.
- [20] He K M, Gkioxari G, Dollár P, et al. Mask R-CNN [C]//2017 IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2980-2988.
- [21] 叶雅欣, 王佳盛, 吴烽云, 等. 消毒机器人目标识别定位与包围盒优化[J]. *激光与光电子学进展*, 2022, 59(4): 0415008.
Ye Y X, Wang J S, Wu F Y, et al. Target recognition and localization, bounding box optimization of disinfection robot[J]. *Laser & Optoelectronics Progress*, 2022, 59(4): 0415008.
- [22] Rusu R B, Blodow N, Beetz M. Fast point feature histograms (FPFH) for 3D registration[C]//2009 IEEE International Conference on Robotics and Automation, May 12-17, 2009, Kobe, Japan. New York: IEEE Press, 2009: 3212-3217.
- [23] Ge J H, Li J X, Peng Y P, et al. Online 3-D modeling of complex workpieces for the robotic spray painting with low-cost RGB-D cameras[J]. *IEEE Transactions on Instrumentation and Measurement*, 2021, 70: 5011013.
- [24] Lenc K, Vedaldi A. Understanding image representations by measuring their equivariance and equivalence[J]. *International Journal of Computer Vision*, 2019, 127(5): 456-476.
- [25] 赵明富, 曹利波, 宋涛, 等. 3D 点云配准中 FPFH 邻域半径自主选取算法[J]. *激光与光电子学进展*, 2021, 58(6): 0610002.
Zhao M F, Cao L B, Song T, et al. Independent method for selecting radius of FPFH neighborhood in 3D point cloud registration[J]. *Laser & Optoelectronics Progress*, 2021, 58(6): 0610002.
- [26] Charles R Q, Liu W, Wu C X, et al. Frustum PointNets for 3D object detection from RGB-D data[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 918-927.

- [27] Theiler P W, Wegner J D, Schindler K. Markerless point cloud registration with keypoint-based 4-points congruent sets[J]. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2013, II-5/W2: 283-288.
- [28] Zhong Y. Intrinsic shape signatures: a shape descriptor for 3D object recognition[C]//2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops, September 27-October 4, 2009, Kyoto, Japan. New York: IEEE Press, 2009: 689-696.
- [29] Aiger D, Mitra N J, Cohen-Or D. 4-points congruent sets for robust pairwise surface registration[C]//ACM SIGGRAPH 2008 papers on-SIGGRAPH '08, August 11-15, 2008, Los Angeles, California. New York: ACM Press, 2008: 1-10.
- [30] Salti S, Tombari F, di Stefano L. SHOT: Unique signatures of histograms for surface and texture description [J]. Computer Vision and Image Understanding, 2014, 125: 251-264.
- [31] 李宇翔, 郭际明, 潘尚毅, 等. 一种基于 ISS-SHOT 特征 的点云配准算法[J]. 测绘通报, 2020(4): 21-26.
Li Y X, Guo J M, Pan S Y, et al. A point cloud registration algorithm based on ISS-SHOT features[J]. Bulletin of Surveying and Mapping, 2020(4): 21-26.