

DOI:10.16356/j.1005-2615.2021.01.008

## 基于区域感知的机械零件位姿计算方法

李红卫<sup>1</sup>, 张煜奇<sup>2</sup>, 周军<sup>2</sup>, 汪俊<sup>2</sup>

(1. 航空工业西安飞机工业(集团)有限责任公司, 西安 710089; 2. 南京航空航天大学机电学院, 南京 210016)

**摘要:** 机械零件位姿识别的准确性与高效性是实现生产过程自动化与仓储物流智能化的关键性能之一。本文对一般机械零件的位姿计算方法进行了研究, 针对其形状种类多、结构特征明显等特点, 结合场景语义分割方法以及点对特征, 提出了一种基于区域感知的机械零件位姿计算方法, 计算机械零件的位姿。为验证方法的有效性, 使用多组数据进行了实验, 并和其他方法进行了对比, 试验结果表明该方法具有较高的位姿计算精度, 且可以有效地提高计算效率。

**关键词:** 位姿计算; 语义分割; 特征描述子; 机械零件

中图分类号: V19 文献标志码: A 文章编号: 1005-2615(2021)01-0078-07

## Pose Estimation for Mechanical Parts Based on Area Sensing

LI Hongwei<sup>1</sup>, ZHANG Yuqi<sup>2</sup>, ZHOU Jun<sup>2</sup>, WANG Jun<sup>2</sup>

(1. AVIC Xi'an Aircraft Industry (Group) Company Ltd, Xi'an 710089, China; 2. College of Mechanical and Electrical Engineering, Nanjing University of Aeronautics & Astronautics, Nanjing 210016, China)

**Abstract:** The accuracy and efficiency of the object pose estimation is one of the key performance to realize the automation of production process and the intellectualization of warehousing and logistics. Some general mechanical parts pose estimation methods are studied. This work proposes a pose recognition method based on area sensing for mechanical parts characterized by abundant varieties of shapes and obvious structural features for mechanical parts pose estimation. This is achieved by combining the state-of-the-art semantic segmentation method and the point pair feature. In order to verify the effectiveness of the method, we conduct vast experiments and comparisons using multiple sets of data. The experimental results show that the method has higher pose recognition accuracy and can effectively improve the recognition efficiency.

**Key words:** pose estimation; semantic segmentation; feature descriptors; mechanical parts

随着工业机器人在零件识别、抓取和现场巡检等制造环节、场景的大量应用<sup>[1]</sup>, 基于机器人对零件的位姿计算技术已经成为智能制造领域中的重要一环。飞机制造过程中大量的机械零件, 由于其形状种类多、尺寸大小不一等特点, 始终缺乏快速、有效的零件自动分拣手段, 使分拣环节难以与零件的自动配送系统形成高效的有机整体, 限制了飞机制造的自动化程度及生产效率的进一步提升。研究机械零件的位姿计算技术已日益迫切。

传统的机械零件位姿识别计算方案如 Delta 机器人识别方法、射频识别技术等, 由于准确率与效率较低、应用局限性大以及位姿计算精度低等原因, 很难满足现代飞机制造的需求。近年来, 国内外许多科研人员基于点对特征(Point pair feature, PPF)<sup>[2]</sup>对三维物体的位姿识别开展了大量的研究<sup>[3-7]</sup>。王化明等<sup>[8]</sup>提出了基于点对特征和局部参考系的位姿计算算法, 提高了目标物体识别的准确率以及效率。鲁荣荣等<sup>[9]</sup>提出一种增强型点对特

基金项目: 江苏省杰出青年基金(BK20190016)资助项目。

收稿日期: 2020-05-25; 修订日期: 2020-07-25

通信作者: 汪俊, 男, 教授, 博士生导师, E-mail: wjun@nuaa.edu.cn。

引用格式: 李红卫, 张煜奇, 周军, 等. 基于区域感知的机械零件位姿计算方法[J]. 南京航空航天大学学报, 2021, 53(1): 78-84. LI Hongwei, ZHANG Yuqi, ZHOU Jun, et al. Pose estimation for mechanical parts based on area sensing[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2021, 53(1): 78-84.

征用来识别目标物体,解决了原始算法中存在的内存浪费以及效率低下的问题。然而,在严重遮挡、杂乱堆放的场景中,如图1所展示的机械零件和机械臂抓取现场,现有的位姿识别方法面临较大的挑战,无法达到工业现场要求的识别效率以及准确率。飞机机械零件因表面无纹理或纹理少、几何形状种类繁多、大小不一,位姿计算的技术难度更大。



图1 机械零件及机械臂抓取现场

Fig.1 Mechanical parts and grabbing scene

本文针对上述问题,通过引入基于深度学习的机械零件语义分割方法,智能感知机械零件的空间区域信息,优化基于特征描述子的机械零件位姿计算算法,以实现基于机械臂的机械零件高效率精准抓取。具体地,本文首先构建一些常用的机械零件语义分割RGB数据集,训练深度学习分割网络,实

现基于RGB图像数据的机械零件语义分割;接着,结合深度数据及视觉传感器内参,计算二维图像与三维点云的映射关系,实现不同机械零件所在三维空间区域的分割;最后,基于分割的不同类别的零件,分别计算描述子特征,提高常用机械零件位姿计算的准确率及效率,最终通过实验对本文方法进行验证。

## 1 算法概述

本文首先基于机械臂视觉传感器,获取到待抓取机械零件的RGBD数据。为了提高直接利用特征描述子进行机械零件位姿计算的精度及效率,在进行机械零件位姿计算之前,本文提出了一种基于深度学习的机械零件三维空间区域感知方法,即利用深度学习对机械臂视觉信息进行分析,实现机械零件在二维空间的分类及分割。进一步地,结合传感器参数及深度信息,利用二维与三维的空间映射关系,实现不同机械零件所在三维空间区域感知。基于三维感知结果,本文针对不同类别的机械零件,分别计算PPF特征,缩小特征匹配的范围,提高机械零件位姿计算的精度及效率。最后基于计算结果,自动规划机械臂路径,支持机械零件的自动分拣抓取。整体的实验配置及方案如图2所示。

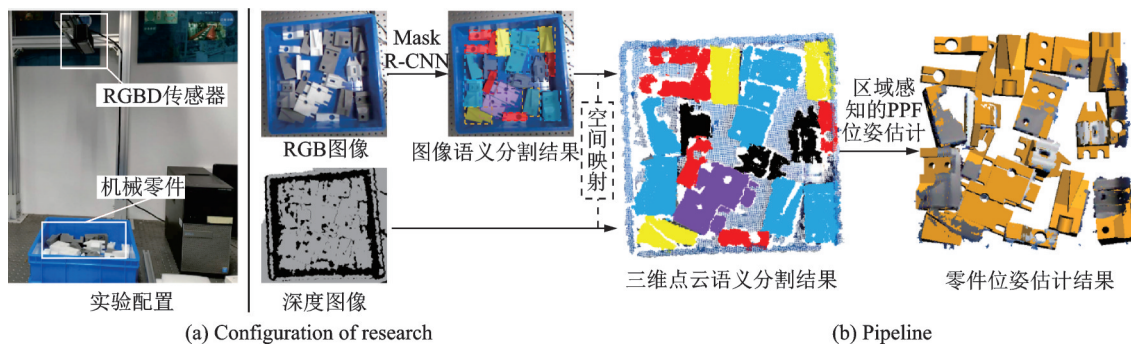


图2 实验配置及方案

Fig.2 Configuration and solution of this research

## 2 算法流程

本文首先利用深度学习框架Mask R-CNN<sup>[10]</sup>对机械臂视觉传感器获取的视觉信息进行智能分析;接着基于二维到三维的空间映射关系,实现机械零件的三维区域感知;最后利用PPF特征,实现机械零件的位姿计算。

### 2.1 基于Mask R-CNN的机械零件RGB图像语义分割

为了实现基于视觉信息的机械零件的分类与分割,本文采用Mask R-CNN深度学习框架。Mask R-CNN是Faster R-CNN<sup>[11]</sup>的扩展,在Faster R-CNN原有的用于对象分类和边界框识别的分支

上增加了一个用于预测对象掩码的并行分支。因此,Mask R-CNN与Faster R-CNN类似,同样由两个阶段所构成。在第1阶段,通过区域生成网络(Region proposal network, RPN)生成候选对象边界框;在第2阶段,除了进行预测对象类别和边界框识别外,Mask R-CNN还使用一个全卷积网络(Fully convolutional networks, FCN)为每个感兴趣区域(Region of interest, ROI)生成一个二进制掩码。基于以上两个阶段,形成一个小型的、灵活的通用对象实例分割框架,不仅可以检测图像中的目标,还可以为每个目标提供高质量的分割结果。然而,原始的Mask R-CNN框架是基于COCO开源训练数据集<sup>[12]</sup>,将其直接用到机械零件的语义分割任务

中,精度较差。因此,本文首先构建几类常用的机械零件语义分割RGB数据集,基于该数据集进一步优化模型参数,提高机械零件语义分类分割的精度。

### 2.1.1 数据集构建

本文使用的视觉传感器配置如图2(a)所示。规划机械臂路径之前,视觉传感器获取当前机械零件摆放场景的RGBD数据。本文选取10类比较常见的机械零件,具备通孔、倒角和棱边等机械零件的典型特征,且表面光滑无纹理。设计不同的摆放场景,共获取1000张RGB图像数据,并通过中心裁剪统一调整大小到1024像素×1024像素。此外,为了实现更加鲁棒的训练,本文对采集到的数据集进行扩充。具体来说,数据集主要通过3种方式进行扩充:随机旋转,水平/垂直翻转和模糊。随机旋转和翻转技术可以模拟传感器在不同的角度和不同的方向采集到的图像。图像模糊用于模拟相机抖动或无法对焦时的图像捕获。在此基础上,本文采用Label me软件对所有图像进行人工的语义标注及分割,作为训练Mask R-CNN的真实值。最终,本文数据集包括4000张机械零件无序摆放的RGB图像及对应的语义分类、分割真值。这一数据集被分成3个部分:80%的样本用于训练,10%的样本用于验证,而10%的样本用于测试。

### 2.1.2 框架细节

本文采用特征金字塔网络<sup>[13]</sup>(Feature pyramid network, FPN)和深度残差网络<sup>[14]</sup>(Deep residual network, ResNet101)作为Mask R-CNN的主干网络。输入视觉传感器获取到的RGB图像,基于ResNet101中Block1至Block5每个残差块的最后一层卷积计算得到自底而上的特征图。卷积层之间的最大池化操作使每一层生成的特征图大小分别为输入图像的1/2、1/4、1/8、1/16和1/32。接着,通过特征融合操作将高层特征映射逐层合并到低层特征映射,实现上下文特征融合。此外,本文是针对10类机械零件的语义分割,因此需要对模型关于分类的分支最后的全连接层进行修改,模型的结构框架如图3所示。

本文同时考虑提高机械零件定位精度、分类及分割准确度,定义如下优化方程

$$\text{Min}_{\theta} \left\{ \sum_{i=0}^N \left[ \sum_{j=0}^K \alpha \|l_{ij} - l_{ij}^{gt}\|_2^2 - \beta p_{ij}^{gt} \log(q_{ij}) + \sum_{k=0}^M \gamma \|f_{ik} - f_{ik}^{gt}\|_2^2 \right] \right\} \quad (1)$$

式中: $\theta$ 是Mask R-CNN模型参数。式(1)第1项计算机械零件的定位精度, $l$ 包含机械零件包围盒的位置及尺寸信息;第2项计算机械零件的分类准确

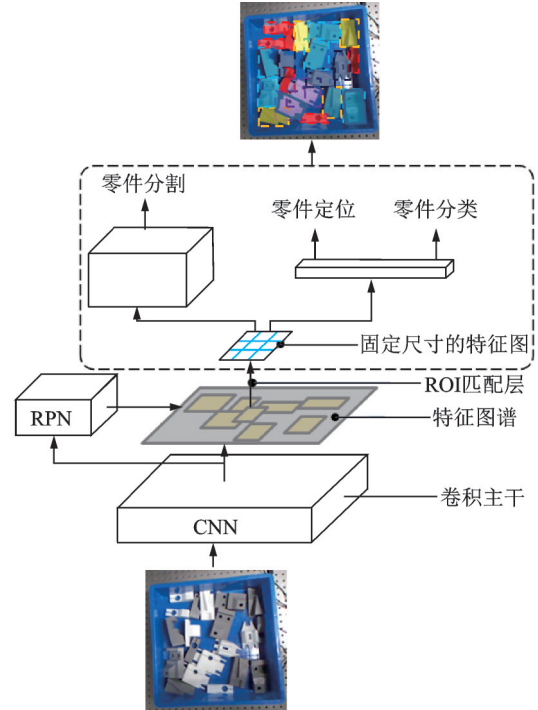


图3 基于Mask R-CNN的机械零件RGB图像语义分割模型结构框架

Fig.3 Mask R-CNN architecture for semantic segmentation of mechanical parts

度, $q$ 表示模型对单个机械零件输出的分类的概率;第3项计算机械零件分割的精度, $f$ 表示输出图像上的单个像素; $gt$ 表示这些信息来自本文数据集,包括:第*i*张RGB图像上,机械零件*j*真实包围盒的位置及尺寸 $l_{ij}^{gt}$ ,机械零件真实分类 $p_{ij}^{gt}$ ,及输入图像每个像素对应的分割真值 $f_{ik}^{gt}$ ,辅助优化,更新模型参数。此外, $\alpha$ 、 $\beta$ 和 $\gamma$ 调节以上3项的相对重要性,通过大量的实验验证,本文的权重设为 $\alpha=0.8$ 、 $\beta=1$ 和 $\gamma=0.5$ 。

### 2.1.3 训练实验配置

给定机械零件RGB图像和数据集中对应的语义分割真实结果,本文采用交并比(Intersection over union, IOU)策略来定义FPN锚框中的正样本和负样本。IOU表示预测边界盒与地面真实边界盒的交集与并集之比。本文将ROI得分大于0.7的锚框定为正,将IOU得分小于0.3的锚框定为负,这样可以丢弃大部分无效的区域提案。接着,在每个小批量迭代中,使用所有选择的锚点,基于式(1),计算分类和协调回归损失。本文共训练50000次迭代,将权值衰减和动量分别设置为0.0001和0.9。本文将初始学习速率设置为0.001。经过30000次迭代后,将学习率重新设置为0.0001。实验环境描述如下:深度学习框架Tensorflow 2.0, Windows 10, Python 2.7, Intel



Core I7-8700,带有 8 GB 内存的 GTX 1080 图形处理单元(GPU)。

## 2.2 基于区域感知结果的机械零件位姿计算

为了实现机械零件所在三维空间区域感知,本文基于以上机械零件 RGB 图像语义分割结果,结合视觉传感器的内参及采集到的深度信息,利用二维与三维的空间映射关系,构建附带语义信息的机械零件三维点云。

机械零件具有表面光滑无纹理、结构特征明显等特点,PPF 特征可有效利用机械零件的结构特点进行特征表达,因此,本文利用 PPF 特征对机械零件和场景进行描述。传统基于 PPF 的位姿计算方法对整个场景进行采样,在运行阶段同时计算并存储场景和模型的特征,存在以下问题:(1)特征计算量大,存储空间内存消耗大,需要较长的匹配时间,算法效率低。(2)除了待识别的机械零件,存在其他冗余数据,可能造成误匹配,影响位姿识别的精度。为了提高位姿计算的精度及效率,本文基于 PPF 特征和上述区域感知结果改进位姿识别算法。在离线阶段,计算待检测机械零件模型 PPF 特征,构建机械零件特征库;在检测阶段,计算由区域感知结果分割的机械零件点云数据的 PPF 特征,搜索特征库获得与场景机械零件上点对对应的模型点对。由这两组点对执行局部坐标对齐和投票决策,优化得票值高的位姿结果,实现机械零件的位姿计算。

### 2.2.1 PPF 特征描述

PPF 特征用于描述点对,它是一个由 4 个参数组成的向量,如图 4 所示。点云中存在两个点  $p_1, p_2$ ,构成点对,用  $\angle(a, b) \in [0, \pi]$  表示两个向量之间的夹角,则 PPF 特征可表示为

$$F(p_1, p_2) = (\|d\|_2, \angle(n_1, d), \angle(n_2, d), \angle(n_1, n_2)) \quad (2)$$

式中: $d$ 表示由点  $p_1$  指向  $p_2$  的向量; $n_1, n_2$  分别表示两个点所在表面的法向量。

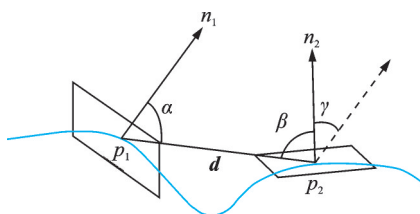


图 4 PPF 特征

Fig.4 Point pair feature

视觉传感器采集场景数据转化成点云,点云中每对点对计算一个 PPF 特征将导致存储空间的巨大消耗,计算量十分庞大,为了加快算法速度,本文首先对点云进行泊松采样,采样点构成点对计算

PPF 特征,随后利用 Hash 表结构对 PPF 特征进行编码,利用链表和数组结构存储信息,具有相同 Hash 值的点对将存储在同一个线性表中,构建机械零件特征库,如图 5 所示。

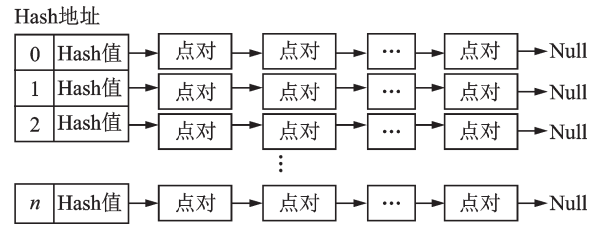


图 5 Hash 编码

Fig.5 Hash coding

### 2.2.2 局部坐标对齐

建立机械零件特征库后,基于区域感知结果分割场景中的机械零件,对于场景机械零件数据进行相同的泊松采样,计算采样点间的 PPF 特征,匹配 Hash 值。在这里,基于分割结果避免对场景中背景等其他数据的计算,一定程度上提高匹配的精度和效率。假设场景中存在机械零件点对  $S_i, S_j$ ,则模型机械零件上存在点对  $M_i, M_j$  与其匹配,为了获得一个统一的表示,本文建立局部坐标系,如图 6 所示,分别将两点对对齐到局部坐标原点,围绕法线旋转物体,完成场景与模型间的局部坐标对齐,具体公式为

$$R_s \cdot t_s \cdot S_i = R(\alpha) \cdot R_m \cdot t_m \cdot M_i \quad (3)$$

式中: $R_s, R_m, R(\alpha)$  为旋转变换; $t_s, t_m$  为平移变换。场景点对  $S_i, S_j$  和模型点对  $M_i, M_j$  存在对应关系,建立局部坐标系,通过刚体变换,分别将点  $S_i, M_i$  变换到坐标原点,并旋转其法线与  $x$  轴方向对齐。此时,两个点对之间存在错位,再通过旋转变换  $R(\alpha)$  旋转角度  $\alpha$  完成局部坐标对齐,最后,通过 5 个变换关系,求解模型空间到场景空间的转换关系,即可获得待求机械零件的一个位姿结果。

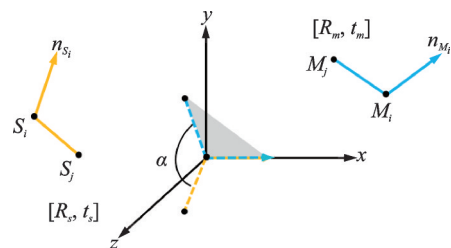


图 6 局部坐标变换

Fig.6 Local coordinate transformation

### 2.2.3 投票策略

研究目标是找到“最优”的位姿结果,使得经过位姿变换后的模型点落在场景中机械零件上的数量最大。本文采用一种投票策略对局部坐标进行投票,最终找到“最优”局部坐标。

首先,创建一个二维累加器,行数与模型采样点数量一致,列数与旋转角度采样步数一致,本文设置为30,即仅允许采样角度与正确角度之间的最大差值为 $12^\circ$ ,在保证识别精度的同时降低采样步数,以满足效率要求;接着选定场景机械零件上一点作为参考点,与其他场景点配对,计算PPF特征并从库中寻找匹配项,返回在模型上具有相似距离和方向的点对;随后对齐局部坐标,根据参考点和角度建立二维累加器的索引值,并计数投票;最后,所有采样点遍历上述阶段,得到一组二维累加器阵列,完成投票。

由此,获得了多个位姿结果,本文引入位姿优化策略,来提高位姿的精确性。首先,对于一个位姿,与已存在的位姿相比较,若旋转量和平移量小于给定的阈值,则将这个位姿归为一类,否则,创建新的位姿类。遍历所有位姿结果,即可得到位姿聚类结果,累加同一聚类内的位姿得票值,取峰值为正确位姿聚类结果。最后,均值化类内位姿,得到“最优”位姿结果。

### 3 结果分析

#### 3.1 合成数据

为了验证本文方法对于位姿识别的准确性,本文首先利用人工合成的机械零件数据搭建虚拟场景,采集RGBD数据,并对其中的机械零件位姿进行手动标注,获得真值位姿,以此作为评估精度的标准。此外,本文与传统的PPF位姿识别算法比较,分别评估两种方法对相同的合成数据集进行位姿计算的精度,并引入匹配时间来评判算法的效率,通过精度和效率两个指标评判两种算法的优劣。

图7表示两种方法的位姿识别结果,识别的位

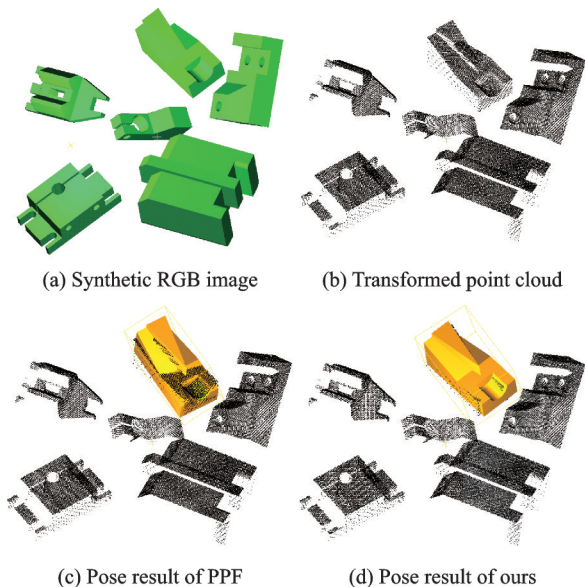


图7 合成数据的单实例识别

Fig.7 Single-instance recognition of synthetic data

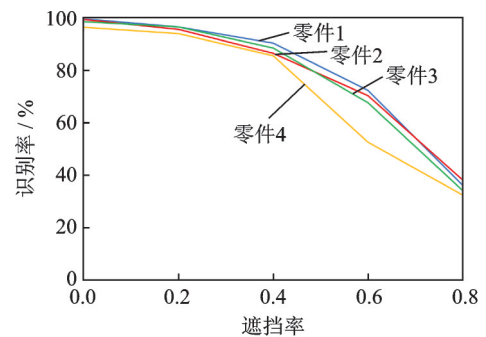
姿用3D包围盒表示,虽然两种方法都能计算目标物体的位姿,但本文方法优于PPF方法,原因在于本文方法进行三维点云语义分割,避免冗余数据特征影响位姿识别结果。表1给出两种方法位姿结果与真值位姿的平均误差,本文方法的识别结果平均误差平移量为0.52 mm,旋转角度为 $3.38^\circ$ ,相较于传统的PPF算法提高一倍,优于传统算法。从数据可以得出结论,本文的方法在位姿识别精度方面有所提高,而匹配速度明显高于PPF算法,原因是本文对待抓取的机械零件进行三维感知,语义分割出待识别的机械零件,而待识别的机械零件在整个场景数据中占比小,此方法只对相同类别的场景零件数据计算特征,避免大量冗余数据的计算。由此可见,本文的方法在提高位姿识别精度的同时,提高位姿识别的速度,在实际工程应用中更具有优势。

表1 两种方法的误差对比

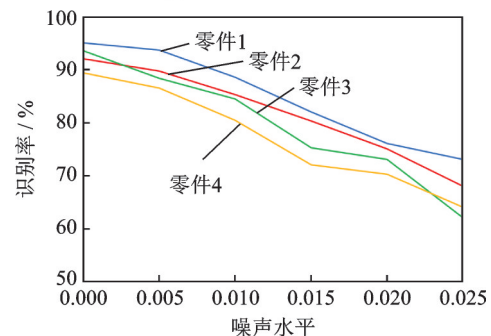
Table 1 Error comparison of the two method

评估指标	PPF	本文方法
旋转角度/rad	7.06	3.38
平移量/mm	1.11	0.52
匹配时间/s	69.6	12.9

此外,本文还测试算法对遮挡、噪声的影响,其中遮挡率及噪声比率定义及设置方式参考文献[15]。本文分别以20%、40%、60%、80%的遮挡率进行遮挡测试,添加0.5%、1.0%、1.5%、2.0%、2.5%的高斯噪声进行噪声测试,如图8所示,该算



(a) Influence of occlusion



(b) Influence of noise

图8 遮挡和噪声的影响

Fig.8 Influence of occlusion and noise

法在一般的遮挡和噪声水平下识别率能达到 80%,鲁棒性较好。

### 3.2 真实场景数据

为了测试本文方法在真实场景中的适用性,本文采集真实的机械零件数据进行位姿识别实验,为了采集测试数据,利用视觉传感器采集场景的 RGB-D 图像,作为本文方法的原始输入。图 9 展示实际场景中的位姿识别效果,其中,蓝框部分的位姿识别效果较好,模型数据和场景数据基本拟合,红框部分位姿识别效果较差,原因在于该机械零件处于数据采集框边缘,数据质量较差,特征被遮挡。本文利用机械臂和视觉传感器在实验室进行抓取实验。表 2 展示了本文方法进行的若干次标准件的抓取结果,可以观察到,位姿识别结果直接影响抓取结果,正确的位姿识别结果可以有效引导机械臂抓取,本文的方法对抓取的成功率较高。经过上述实验,表明该方法适用于实际场景。总体来说,本方法对真实的场景有效,能够识别出零件的位姿。



图 9 真实场景多实例识别

Fig.9 Multi-instance recognition of real data

表 2 抓取实验结果

Table 2 Result of crawling experiment

机械零件类别	实验次数	抓取成功次数	成功率/%
零件 1	200	187	93.5
零件 2	200	192	96.0
零件 3	200	173	86.5
零件 4	200	184	92.0

## 4 结 论

本文提出一种基于区域感知的机械零件位姿计算方法,并在大量的合成数据和实际场景中进

行了实验论证。结果表明,本文的位姿计算方法的平移误差在 1 mm 以内,旋转角度误差在 4° 以内,识别结果用于抓取任务中成功率可达 85%。针对实际项目需求,在场景中摆放 20 个机械零件,利用机械臂进行实际抓取实验并设定 200 次的抓取次数上限,通过本方法可在限定次数下将场景中的机械零件全部抓取成功,验证了本方法的实际应用价值,可以为实际工程应用提供有效支撑。

### 参考文献:

- [1] 田威,焦嘉琛,李波,等.航空航天制造机器人高精度作业装备与技术综述[J].南京航空航天大学学报,2020,52(3): 341-352.  
TIAN Wei, JIAO Jiachen, LI Bo, et al. High precision robot operation equipment and technology in aerospace manufacturing[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2020,52(3): 341-352.
- [2] DROST B, ULRICH M, NAVAB N, et al. Model globally, match locally: Efficient and robust 3D object recognition[C]//Proceedings of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2010: 998-1005.
- [3] 赵安安,吴志鹏,曹涛,等.基于局部不变性的航空 C 型梁配准方法[J].南京航空航天大学学报,2020,52(3): 430-437.  
ZHAO An'an, WU Zhipeng, CAO Tao, et al. Aeronautical C-beam registration based on local invariance [J]. Journal of Nanjing University of Aeronautics & Astronautics, 2020,52(3): 430-437.
- [4] 易杰,李蹊,易辉成,等.基于点对特征的三维点云匹配算法[J].传感器与微系统,2019(9): 115-117.  
YI Jie, LI Xi, YI Huicheng, et al. 3D point cloud matching algorithm based on point pair feature[J]. Transducer and Microsystem Technologies, 2019 (9): 115-117.
- [5] VIDAL J, LIN C Y, MARTÍ R. 6D pose estimation using an improved method based on point pair features [C]//Proceedings of the 4th International Conference on Control, Automation and Robotics (iccar). [S.l.]: IEEE, 2018: 405-409.
- [6] LIU D, SHOGO A, MIAO J, et al. Point pair feature-based pose estimation with multiple edge appearance models (PPF-MEAM) for robotic Bin picking [J]. Sensors, 2018, 18(8): 2719.
- [7] SAHIN C, GARCIA-HERNANDO G, SOCK J, et al. A review on object pose recovery: From 3D bounding box detectors to full 6D pose estimators[J]. Image and Vision Computing, 2020, 96: 103898.



- [8] 王化明, 郭交通, 陈作特. 基于点对特征和局部参考系的六维位姿估计算法[J]. 江苏大学学报(自然科学版), 2019, 40(6): 695-700.  
WANG Huaming, GUO Jiaotong, CHEN Zuote. 6D pose estimation based on point pair feature using local reference frame constraint[J]. Journal of Jiangsu University (Natural Science Edition), 2019, 40(6): 695-700.
- [9] 鲁荣荣, 朱枫, 吴清潇, 等. 基于增强型点对特征的三维目标识别方法[J]. 光学学报, 2019, 39(8): 245-254.  
LU Rongrong, ZHU Feng, WU Qingxiao, et al. Three-dimensional object recognition based on enhanced point pair features[J]. Acta Optica Sinica, 2019, 39(8): 245-254.
- [10] HE K, GEORGIA G, PIOTR D, et al. Mask R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision. [S. l.]: IEEE, 2017: 2961-2969.
- [11] GIRSHICK R. Fast R-CNN[C]// Proceedings of the IEEE International Conference on Computer Vision. [S. l.]: IEEE, 2015: 1440-1448.
- [12] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft COCO: Common objects in context[C]//Proceedings of European Conference on Computer Vision. Cham, Switzerland: Springer, 2014: 740-755.
- [13] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S. l.]: IEEE, 2017: 2117-2125.
- [14] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S. l.]: IEEE, 2016: 770-778.
- [15] JOHNSON A E, HEBERT M. Using spin images for efficient object recognition in cluttered 3D scenes[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 1999, 21(5): 433-449.

(编辑:陈珺)