

DOI:10.3969/j.issn.1003-5060.2024.03.009

用于钢铁回收中的目标检测与尺寸估计算法

范彬彬^{1,2}, 秦训鹏^{1,2}, 吴强^{1,2}, 王哲^{1,2}, 毕玖璐^{1,2}

(1. 武汉理工大学汽车工程学院, 湖北 武汉 430070; 2. 现代汽车零部件技术湖北省重点实验室, 湖北 武汉 430070)

摘要:为解决钢铁回收过程中由于开放式复杂场景而无法准确地获得目标废钢点云进行尺寸估计,文章提出一种基于 Mask R-CNN 模型预测掩膜自适应颜色阈值的点云提取算法。针对 Mask R-CNN 模型边缘分割的不完整性,通过预测框对点云进行截取,利用自适应颜色滤波阈值对截取的点云进行 HSV 颜色空间滤波,将非目标点云离散化,最后通过欧式聚类对目标点云进行分割并进行尺寸估计。通过废旧汽车 B 柱模拟回收场景,获取汽车 B 柱点云,验证了该算法的有效性。

关键词:双目立体视觉;实例分割;点云分割;颜色滤波;钢铁回收

中图分类号:TP391 **文献标志码:**A **文章编号:**1003-5060(2024)03-0338-09

Object detection and size estimation algorithm for steel recycling

FAN Binbin^{1,2}, QIN Xunpeng^{1,2}, WU Qiang^{1,2}, WANG Zhe^{1,2}, BI Jiuju^{1,2}

(1. School of Automotive Engineering, Wuhan University of Technology, Wuhan 430070, China; 2. Hubei Key Laboratory of Advanced Technology for Automotive Components, Wuhan 430070, China)

Abstract: To solve the problem of inability to accurately obtain the target scrap point cloud for size estimation due to open and complex scenes in the process of steel recycling, a target point cloud extraction algorithm based on the prediction mask generated by the Mask R-CNN model and adaptive color thresholding technique is proposed. For the incomplete edge segmentation of the Mask R-CNN model, the point cloud is intercepted through the prediction frame, the adaptive color filter threshold is used to filter the intercepted point cloud in HSV color space to discretize the non-target point cloud, and finally the target point cloud is segmented through European clustering and dimensionally estimated. The effectiveness of the algorithm is verified by simulating the recovery scene of the B-pillar of the scrapped car and obtaining the point cloud of the B-pillar of the car.

Key words: binocular stereo vision; instance segmentation; point cloud segmentation; color filtering; steel recycling

0 引言

钢铁回收过程中,不同大小尺寸的废旧钢铁回收价格不同,并且废旧钢铁回收处理过程中,如果钢铁尺寸过大,会使龙门剪切机卡料,损坏器械,因此需要对回收的钢铁进行尺寸判别,依据尺寸大小定价,并根据尺寸大小确定废旧钢铁的处

理流程。双目视觉测量是依据人类视差成像原理,利用两台相机通过左、右两幅图像获取视差进而生成三维点云数据的测量方式^[1]。双目视觉测量虽然精度相比激光雷达点云精度存在差距^[2],但是凭借双目视觉的高效率、精度适宜、结构简单、成本低的特点,在工业生产、三维测量领域得到了广泛的应用。

收稿日期:2023-04-07;修回日期:2023-05-04

基金项目:中国博士后科学基金资助项目(2020M682498);湖北省技术创新重大专项资助项目(2019AAA075;2020BED010)

作者简介:范彬彬(1999—),男,江西赣州人,武汉理工大学硕士生;

秦训鹏(1962—),男,湖北黄冈人,博士,武汉理工大学教授,博士生导师。

目前,点云的目标分割方法可分为基于深度学习的点云分割和基于颜色、几何等特征的点云分割。基于深度学习的三维点云语义分割模型主流上有 Point Net、Point Net++、KD-Net、3D-MPA 等^[3-6]。这类方法对数据集有非常高的要求,应用的广泛性受到很大的限制。基于颜色、几何等特征点云分割没有数据集的要求,具有更大的应用面。较为常用的特征有边缘特征^[6-8]、表面特征^[9-10]、颜色特征^[11-12]等。其中,颜色特征是双目点云与激光雷达点云的一个显著区别,双目点云的颜色来源于图像颜色,因此双目点云结合颜色特征进行分割得到广泛应用。针对开放式场景的复杂情况,本文结合基于图像的目标检测和实例分割,针对图像实例分割像素精度不高的问题,提出通过实例分割掩膜自适应提取颜色阈值的方法对点云进行颜色滤波和欧式聚类,得到更加完整的实例点云。

钢铁回收领域的突出特点是开放式环境的复杂背景,往往在废旧钢铁中会混杂有其他非钢铁物体,因此本文采用 Mask R-CNN 模型对图像进行目标检测和实例分割,目标检测由于较高的识别精度,能够快速定位废旧钢铁的位置。由于实例分割掩膜的不完整性和低精确度,本文根据掩膜提取目标颜色阈值再对点云进行颜色滤波,以避免掩膜不完整造成的点云缺失,再通过有向包围盒(oriented bounding box, OBB)对钢铁点云进行尺寸估计。

1 算法描述

本文获取目标点云及尺寸估计的算法如图 1 所示。

该算法的具体步骤如下。

- 1) 对采集的双目图像进行立体校正,并对左视图进行目标检测与分割,获得目标所在的矩形区域及掩膜。
- 2) 对校正图像进行立体匹配与视差计算,并进行目标矩形区域裁剪。
- 3) 结合视差与标定参数进行计算,得到三维点云。
- 4) 根据 Mask R-CNN 模型得到的 mask 掩膜自适应获取 HSV 颜色空间阈值。
- 5) 根据颜色阈值对点云进行颜色滤波。
- 6) 通过欧式聚类的方法将目标点云分割出来。
- 7) 通过 OBB 对钢铁点云进行尺寸估计。

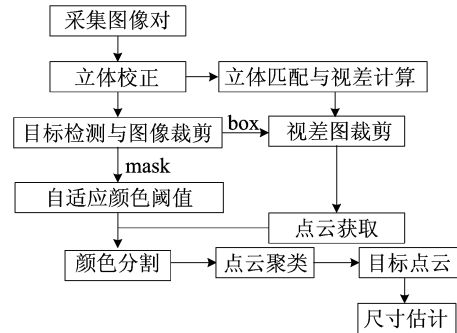


图 1 双目视觉目标点云分割流程

1.1 基于 Mask R-CNN 模型的实例分割算法

Mask R-CNN 模型^[13]是在 Fast R-CNN 基础上,在原本的分类与回归分支上扩展并添加语义分割分支,实现对图像的目标检测与实例分割。Mask R-CNN 是一种两阶段目标检测方法:第 1 阶段通过区域建议网络(region proposal network, RPN)生成一系列建议框;第 2 阶段通过生成的建议框与感兴趣区域对齐,再对特征层进行截取。最后将截取特征输入分类、回归和掩码 3 个分支进行不同的预测任务。本文所用 Mask R-CNN 模型以 ResNet101 为主干特征提取网络,在特征提取后利用压缩 2~5 次的特征层构建特征金字塔,实现多尺度特征融合。将提取到的 P2~P6 作为 RPN 的有效特征层进行建议框的提取,将 P2~P5 作为 3 个分支的有效特征层对建议框进行最终预测。通过 RPN 对有效特征层获取先验框调整参数并判断框内是否存在物体,随后对先验框进行调整和筛选后得到建议框。建议框是对图片有物体的区域进行初步筛选,再结合主干特征提取网络获得的多个公用特征层,经由 ROI Align 网络根据建议框对特征层进行截取,截取后的特征经卷积后输入到不同的检测分支对建议框进行调整和预测,并得到最终预测结果。Mask R-CNN 网络框架图如图 2 所示。

采用单独使用深度学习的方式对目标进行像素级的图像分割具有很大的限制。首先,对数据集的数量和质量要求较高,图像采集和打标签的工作量也非常庞大;其次,目前的实例分割模型对目标的边缘还无法做到精度很高的分割。然而,深度学习在目标检测领域能达到非常高的精度,因此本文采用 Mask R-CNN 模型作为点云算法的预处理模块,其功能包含目标检测和实例分割功能。目标检测得到的目标框将目标截取出来,减少计算量并突出目标,减少背景干扰;实例分割的边缘精度不高但是能够提取出目标的主体部分

作为自适应颜色阈值算法的输入。自适应颜色阈值算法对目标边缘不敏感,只需要获得目标的主

体颜色,因此不需要追求实例分割的高精度,这大大降低了对数据集的要求。

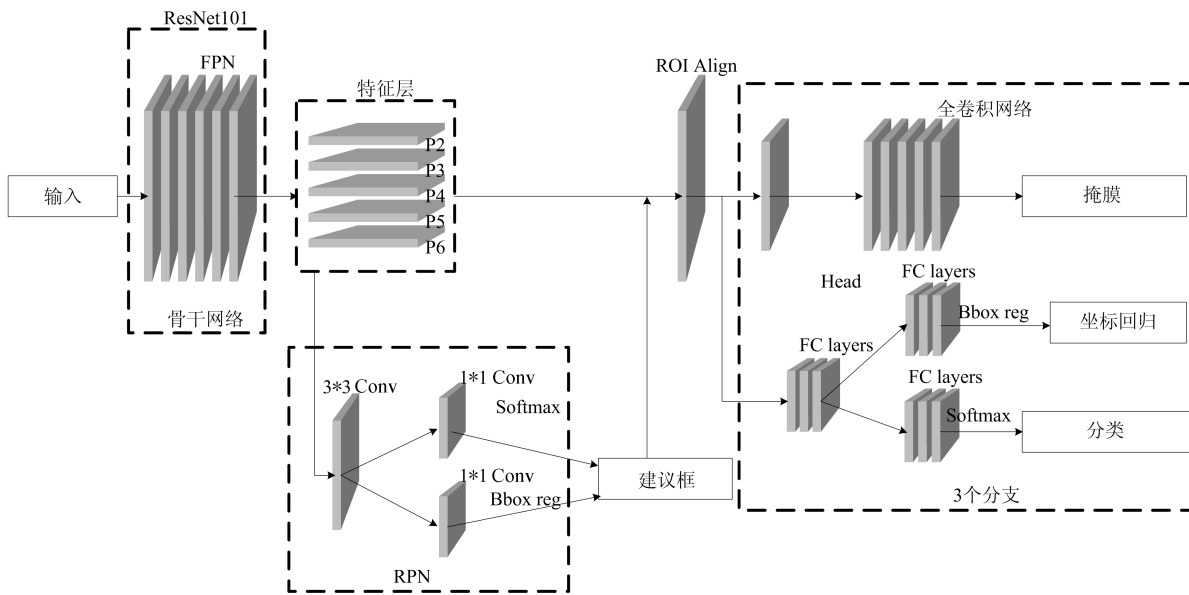


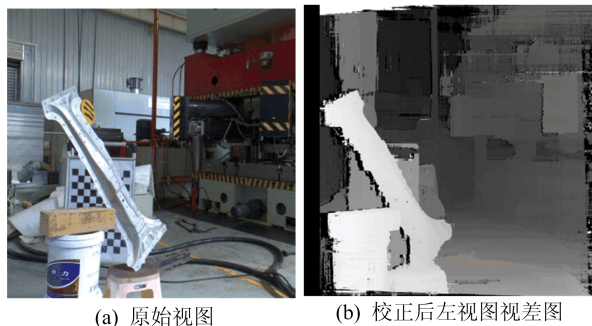
图 2 Mask R-CNN 网络框架图

1.2 基于双目视觉的立体重建

本文使用 open CV 库中半全局块匹配(semi-global block matching, SGBM)算法^[14]对校正图像进行立体匹配与视差计算,如图 3 所示。

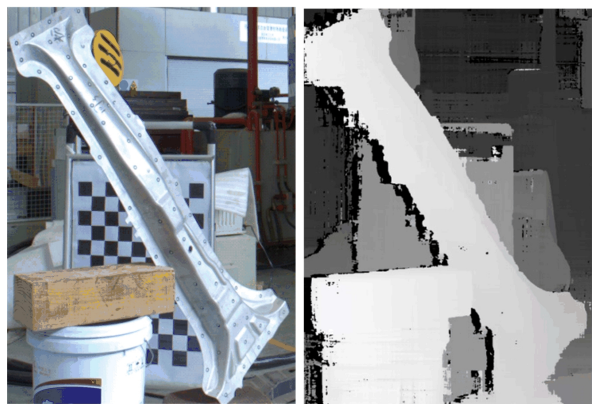
目标截取,根据识别出的目标框,截取图像获得更小的图片,减少计算量和背景干扰。

双目视觉利用双目视差原理计算物体实际坐标,如图 4 所示。



(a) 原始视图

(b) 校正后左视图视差图



(c) 目标左视图

(d) 目标视差图

图 3 图像裁剪

计算出的视差图如图 3b 所示,由于图像像素大小为 2 048×2 040,先对左视图和视差图进行

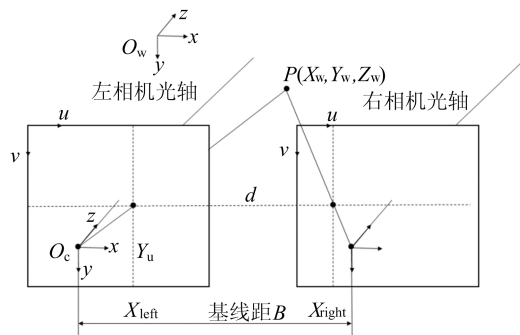


图 4 双目立体视觉示意图

计算公式如下:

$$\frac{B - (X_{left} - X_{right})}{B} = \frac{Z_w - f}{Z_c} \tag{1}$$

$$X_w = \frac{B}{d} X_{left} = \frac{B}{d} (u_1 - c_x) \tag{2}$$

$$Y_w = \frac{B}{d} Y_u = \frac{B}{d} (v_1 - c_y) \tag{3}$$

其中: B 为左右基线距离; X_{left} 、 X_{right} 、 Y_u 为相机坐标系下的图像坐标; u_1 、 v_1 为左相机图像的像素坐标; c_x 、 c_y 为光心坐标。通过式(1)~(3)可计算物体在左相机坐标系下的坐标。

1.3 基于颜色的点云滤波

本文选取的实验目标是汽车 B 柱,具有典型

的银白色金属光泽。针对场景中前景有多个干扰物体的情况,本文通过颜色的差异先对前景的物体进行处理,对目标进行颜色滤波,将非目标点云离散化,同时也减少点云数量,提高后续点云处理速度。

图像的颜色分割可通过不同的颜色空间进行区分,较为常用的有 RGB 颜色空间、HSV 颜色空间、YCbCr 颜色空间和 Lab 颜色空间。根据实际场景采集的图片,进行调参实验发现,HSV 颜色空间能够取得更理想的分离效果,能够在将前景中汽车 B 柱较完整保留下来的同时,将前景目标边缘的像素滤除,实验结果如图 5 所示,因此本文颜色滤波采用 HSV 颜色空间。

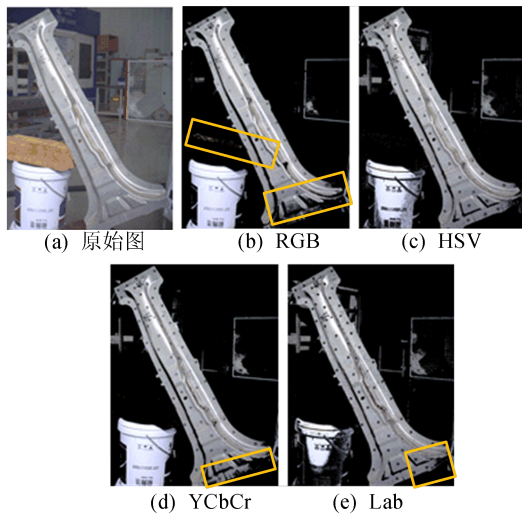


图 5 不同颜色空间下的颜色滤波图

通过掩膜提取目标的 RGB 三通道值,并转化到 HSV 颜色空间,再计算 HSV 三通道的像素分布。RGB 颜色空间到 HSV 颜色空间的转化公式如下:

$$H = \begin{cases} 0^\circ, & c_{\max} = c_{\min}; \\ \frac{60^\circ(G-B)}{c_{\max} - c_{\min}} + 0^\circ, & c_{\max} = R, G \geq B; \\ \frac{60^\circ(G-B)}{c_{\max} - c_{\min}} + 360^\circ, & c_{\max} = R, G < B; \\ \frac{60^\circ(B-R)}{c_{\max} - c_{\min}} + 120^\circ, & c_{\max} = G; \\ \frac{60^\circ(R-G)}{c_{\max} - c_{\min}} + 240^\circ, & c_{\max} = B \end{cases} \quad (4)$$

$$S = \begin{cases} 0, & c_{\max} = 0; \\ \frac{c_{\max} - c_{\min}}{c_{\max}} = 1 - \frac{c_{\min}}{c_{\max}}, & \text{其他} \end{cases} \quad (5)$$

$$V = c_{\max} \quad (6)$$

其中: H 、 S 、 V 分别为 HSV 颜色空间三通道的像

数值; R 、 G 、 B 分别为 RGB 颜色空间三通道的像素值; c_{\max} 、 c_{\min} 分别为当前像素点 RGB 三通道中的最大值和最小值。

根据实验发现,频率分布区间划分越多,阈值计算越准确,同时计算量也将加大,本文选择了 100 个分布区间。将 HSV 三通道的频率分布数列输入自适应颜色阈值选取算法中,得到目标物体的颜色分割阈值,算法伪代码如下。

Input: 颜色频率分布 a, b ; 颜色占比 $\text{scale}=0.9$

// a 存储频率, b 存储区间

Output: 颜色阈值 $\{\text{threshold}_{\min}, \text{threshold}_{\max}\}$

$x \leftarrow \max a$ // 将最高频率赋值给 x

$m, n \leftarrow \text{index}(\max a)$ // 左右起始阈值赋值

while $x < \text{scale}$ and $(m < a.\text{length}$ or $n > 0)$ do

if $m < a.\text{length}$ then // 最大值未越界

$m++$ // 最大阈值右移

$x \leftarrow x + a[m]$ // 频率和增加

end if

if $n > 0$ then // 最小值未越界

$n--$ // 最小值左移

$x \leftarrow x + a[n]$ // 频率和增加

end if

end while

$\text{threshold}_{\min} \leftarrow b[n]$

$\text{threshold}_{\max} \leftarrow b[m+1]$

return $\text{threshold}_{\min}, \text{threshold}_{\max}$

将获得的点云的 RGB 颜色空间转化到 HSV 颜色空间,并根据自适应 HSV 颜色阈值进行点云颜色滤波,使非目标点云离散而目标点云保持连通。

1.4 欧式点云聚类

通过上一步点云的颜色滤波,将非目标点云离散化而最大限度保留了目标点云的完整和连通。针对颜色滤波后的点云特点,本文结合欧式聚类的方法,将目标点云分割。欧式聚类利用点与点之间的欧式距离进行聚类,点与点之间的欧式距离小于设定阈值则视为一类。由于点云数据较大,为提高计算速度,本文在欧式聚类前先对点云进行随机下采样,同时采用构建 K-d 树的方法提高计算速度,算法伪代码如下。

Input: P - Cloud, tolerance, P - Sizemin, P - Sizemax

// 点云, 最大距离, 最小、最大点云数量

Output: 点云聚类结果 $\{\text{clusters}\}$

$\text{kdtree} \leftarrow \text{EstablishKdtree}(P\text{- Cloud})$

$P_Num \leftarrow P\text{- Cloud}.\text{length}$

for $\text{idx}=0 \rightarrow P_Num$ do

if $\text{processed}[\text{idx}] = 1$ then // 该点已处理

```

    continue
end if
seedq ← idx // 加入一个种子点
processed[idx] ← 1 // 标记该点已处理
sqidx ← 0
while sqidx < seedq.length do
    k, Ind ← kdtree(P_Could[seedq[sqidx]])
    // 获取种子点的邻近点数量及索引
    if k = 1 then
        sqidx ++
        continue
    end if
    for j = 0 → k do
        if Ind[j] = P_Num ||
            processed[Ind[j]] = 1 then
                continue
        // 已经处理的近邻点则跳出此次循环继续
    end if
    seedq ← Ind[j]
    processed[Ind[j]] = 1
    // 加入种子点, 并标记
end for
sqidx ++
end while
if P_Sizemax > seedq.length > P_Sizemin then
    clusters ← seedq
    // 聚类结果存入 clusters
end if
end for
return clusters

```

算法中最大距离参数的选择与图像大小和相机视野大小有关, 计算公式如下:

$$1.1\alpha F_s \leq t \leq 1.8\alpha F_s \quad (7)$$

$$F_s = \frac{\text{Fov}(L \text{ 或 } W)}{P(L \text{ 或 } W)} \quad (8)$$

其中: t 为最大距离阈值; α 为点云下采样倍率; F_s 为比例因子; Fov 为单边相机视野; P 为单边图像像素大小; L 、 W 分别为长和宽, Fov 和 P 的长或宽要相对应。

1.5 点云尺寸估计

钢铁回收过程中需要对废旧钢铁进行尺寸估计以进行定价和回收处理。针对点云的尺寸估计, 本文采用 OBB 的方式进行尺寸估计, 它沿着点云的主成分方向生成一个最小的可以随物体旋转的矩形包围盒, 能够更紧密地贴近物体, 该方法比较广泛地用于碰撞检测中。

点云的有向包围盒将通过对点云的主成分分析获得特征向量, 以这些特征向量作为包围盒的

主轴。点云的协方差矩阵 A 的特征向量表示有向包围盒的主轴方向。点云的中心为包围盒的中心, 以点云中心为包围盒的中心, 沿主轴方向计算 x 、 y 、 z 3 个方向的半长度, 即可求出点云的有向包围盒。具体计算公式如下:

$$\text{cov}(X_i, X_j) = E[(X_i - \mu_i)(X_j - \mu_j)] \quad (9)$$

$$A = \begin{bmatrix} \text{cov}(x, x) & \text{cov}(x, y) & \text{cov}(x, z) \\ \text{cov}(y, x) & \text{cov}(y, y) & \text{cov}(y, z) \\ \text{cov}(z, x) & \text{cov}(z, y) & \text{cov}(z, z) \end{bmatrix} \quad (10)$$

式(9)为协方差计算公式, 其中, μ 为均值。

2 实验结果与对比分析

2.1 实验平台

实验平台包括图像采集设备和算法运行环境, 图像采集设备采用 2 个相同型号的工业相机 (Basler acA2440-20gc) 和工业镜头 (C0828-5M) 组成双目立体视觉, 如图 6 所示, 最大图像像素大小为 2440×2048 , 工作距离为 2 m, 视野大小为 2.1×1.775 m。深度学习模型 Mask R-CNN 在工作站上完成 GPU 加速训练后将权重下载并移植到算法运行环境中。运行环境如下: 操作系统为 Windows 10; CPU 为 Intel(R) Core(TM) i5-7300HQ; IDE 为 VS Code; 深度学习框架为 Tensorflow 1.13.1 + keras 2.1.5。



图 6 双目相机

2.2 实验结果

将 400 张金属汽车零部件的图像按照 6 : 2 : 2 的比例划分训练集、验证集和测试集, 再对训练集进行数据增强到 3 000 张图像, 输入 Mask R-CNN 模型中进行训练, 训练 100 次迭代的损失曲线如图 7 所示, 从图 7 可以看出, 在迭代约 80 次后趋于收敛, 在 0.50~0.95 全部 IoU 阈值下, 回归框的平均精确率为 0.892, 平均召回率为 0.902; 掩膜

的平均精确率为 0.672,平均召回率为 0.900。

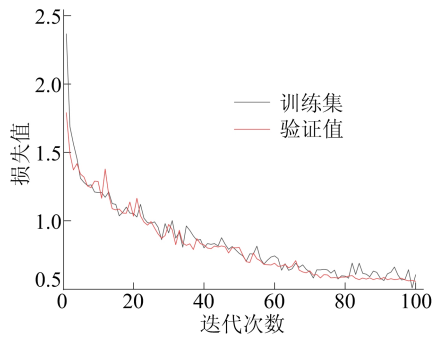


图 7 Mask R-CNN 模型训练迭代过程

校正后的图像对如图 8 所示,目标检测和实例分割的图像为校正后的左相机图像,识别结果如图 9 所示。为减小运算量,对左视图和视差图进行裁剪,如图 3 所示,并进行三维重建得到初始点云。为提高运行速度,将初始点云进行下采样,下采样倍率 $\alpha=10$ 。因此,依据式(7)~(8)取聚类阈值 $t=10$ 。通过颜色滤波与欧式聚类得到的聚类点云如图 10 所示。

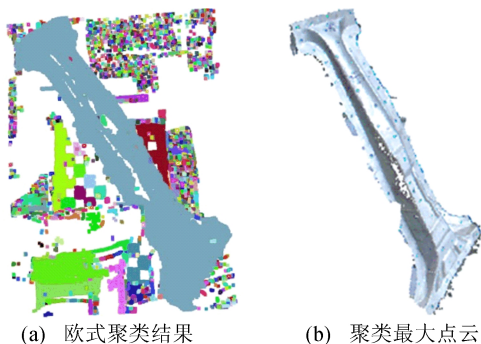


图 8 双目立体校正图像



(a) Mask R-CNN 检测图 (b) YOLOv5s-seg 检测图

图 9 汽车 B 柱目标检测图



(a) 欧式聚类结果 (b) 聚类最大点云

图 10 点云欧式聚类示意图

2.3 对比分析

为定量分析双目视觉目标点云分割算法的性能,参考文献[12,15-16]的滤波误差评价方法,将原始点云 O 分为目标点 O_T 和非目标点 O_F ,分割后的点云分为目标点 S_T 和非目标点 S_F ,滤除的点云包括目标点 D_T 和非目标点 D_F 。第 I 类误差为将非目标点判断成目标点的误差,即 $E_a = (S_F/D_F) \times 100\%$;第 II 类误差为总误差,将非目标点误判为目标点和将目标点误判为非目标点,即 $E_b = [(S_F + D_T)/O] \times 100\%$;定义目标的分割完整度为分割的目标点云与目标点云的比值,即 $I = (S_T/O_T) \times 100\%$ 。其中,真实的目标点云通过人为划分的方法来确定。

为验证本文算法的可行性,采用部分替换及与其他方法进行对比的方式进行对比实验。文献[17]提出了采用半径滤波进行预处理后采用欧式聚类的方法进行目标点云分割;文献[18]采用 Mask R-CNN 实例分割算法对桥梁进行目标检测与实例分割,根据分割掩膜对点云进行分割;基于此基础上,将 2 种算法结合,将 Mask R-CNN 模型与半径滤波和欧式聚类结合。

为验证目标检测模型对本文算法的影响,本文采用 YOLOv5s-seg 模型替换为 Mask R-CNN 模型后再进行颜色滤波与欧式聚类的方法,并通过对比进行验证本文算法。YOLOv5s-seg 是基于 YOLOv5 网络的轻量级实例分割模型。在相同的数据集训练后,0.50~0.95 全部 IoU 阈值下回归框的平均精确率为 0.664,掩膜的平均精确率为 0.673。

图像识别结果如图 9 所示。从图 9 可以看出;Mask R-CNN 模型的目标框识别准确率更高,在图像中能够将目标物体更加完整地包围;YOLOv5s-seg 则没能完整地包围目标物体,且掩膜中存在错误掩膜。

汽车 B 柱点云为本文实验对象,双目重建后原始点云中点数目为 107 142,其中:目标点数目为 25 416;需要滤除的点数目为 81 726。将本文算法与以上 4 种算法进行对比,结果见表 1 所列,效果图如图 11 所示。从表 1 和图 11 可以看出,本文算法对目标点云的分割误差比其他算法更小,点云完整度也更高。文献[17]利用半径滤波和欧式聚类的方法无法从前景的多个物体中将目标分离出来,虽然完整度很高,但是第 I 类误差和第 II 类误差均超过了 20%。

表 1 算法误差对比

| 算法序号 | 算法名称 | S_T | S_F | D_T | D_F | $E_a/\%$ | $E_b/\%$ | $I/\%$ |
|------|-----------------------|--------|--------|-------|--------|----------|----------|--------|
| 1 | 半径滤波+欧式聚类 | 25 402 | 22 286 | 14 | 59 440 | 37.493 | 20.813 | 99.944 |
| 2 | Mask R-CNN+半径滤波+欧式聚类 | 22 212 | 453 | 3 204 | 81 273 | 0.557 | 3.413 | 87.393 |
| 3 | Mask R-CNN+半径滤波 | 22 330 | 567 | 3 086 | 81 159 | 0.698 | 3.409 | 87.858 |
| 4 | 本文算法 | 22 908 | 500 | 2 508 | 81 226 | 0.615 | 2.807 | 90.132 |
| 5 | YOLOv5s-seg+半径滤波+欧式聚类 | 20 759 | 627 | 4 657 | 81 099 | 0.773 | 4.932 | 81.677 |

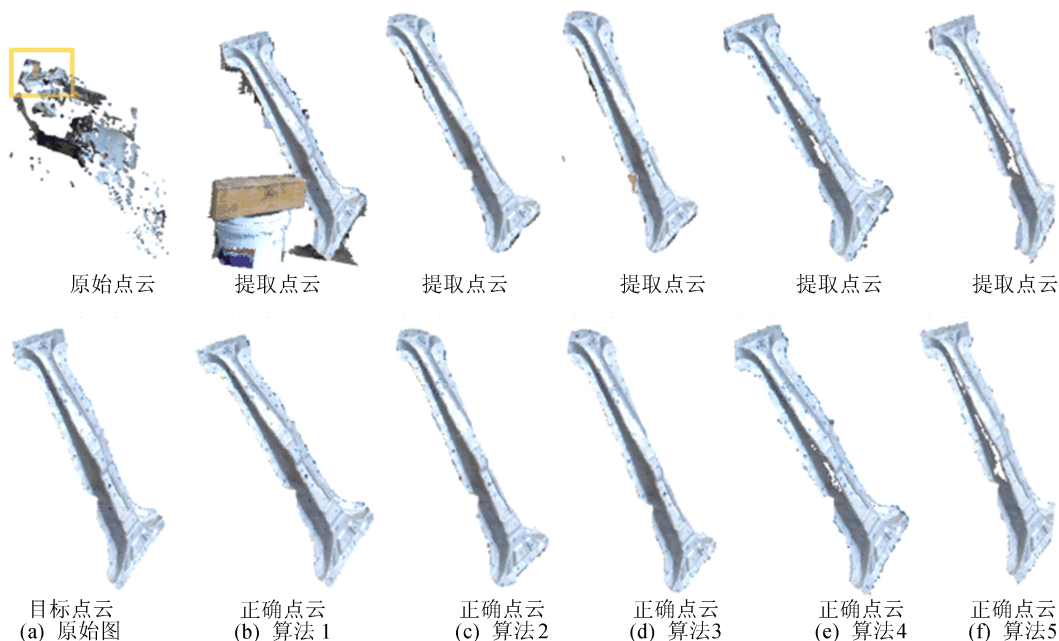


图 11 算法效果对比图

文献[18]利用 Mask R-CNN 模型结合图像分割点云,虽然误差减小,但是点云附近增加了很多噪点,并且由于图像像素分割不够准确,点云边缘还有很多非目标点云。在文献[18]基础上增加欧式聚类,可以将点云边缘附近的非目标点云滤除,减小误差,但从图 11 中可以发现,由于目标检测算法的掩膜分割不够准确,点云中的目标点丢失很多,导致点云缺失了一部分。将目标检测模块换成 YOLOv5s-seg 模型后,由于目标检测回归框的准确率较低,导致无法将目标完整地框选,因此会使保留的目标点 S_T 减少,滤除的目标点 D_T 增大,进而增大 2 种误差。同时,目标框内出现的错误掩膜将会影响颜色率波。由于 YOLOv5s-seg 模型识别的掩膜中出现了非目标物体,掩膜的颜色分布将会出现变化,颜色滤波的自适应阈值也会偏移,导致的结果如图 11f 所示,点云中出现了较大的空洞,导致点云的完整度降低。由以上分析可知,目标检测模型的回归框将影响最终点云的完整度,回归框应尽可能完整地包围目标;掩膜影响最终点云的分割误差和完整度,主要是由于错误掩膜的影响,掩膜轮廓精度对最终效果

无影响。

本文所提出的算法充分考虑到图像前景干扰物体多,深度学习算法数据集不够充分的情况,利用图像的颜色、几何、统计等特征,根据目标的颜色分布,设计了自适应颜色阈值选取算法,结合欧式聚类对目标进行点云分割,分割点云的完整度可达到 90% 以上,第 I、第 II 类误差分别为 0.615%、2.807%。综合来看,本文提出的算法效果最佳。根据图 11 算法效果对比可知,本文算法对于边缘点云的分割更加精确。

同时,在多个废旧钢铁汽车零部件上,本文算法均表现出较好的分割性能,具有一定的泛化能力。不同物体效果如图 12 所示,图 12 中:物体 a 为汽车 B 柱;物体 b 为汽车上的废旧零部件;物体 c 为一块方形废钢。通过本文算法进行双目点云自适应分割,得到的效果图能够较为清晰的表现出物体形貌。

根据实验统计,在本文的设备环境下,本文算法对图 12 的 3 个物体进行点云分割的运行时间统计结果见表 2 所列。

总运行时间包括目标检测时间、立体校正和

三维重构时间、颜色滤波时间、欧式聚类时间,其中颜色滤波和欧式聚类占用的时间较多,这主要与目标点云的数量有关,目标点云越多,耗费时间越多。从表 2 可以看出,点云分割误差也表现较好,说明本文算法具有一定的泛用性和实用价值。本文算法的应用对象为双目点云,因此点云的规模是可控制的,实验所用设备采集的最大图像像素为 $2\,448 \times 2\,040$,物体通常在整個图像上不会超过画面的 $1/4$ 。这种情况下,物体原始未处理

点云通常不会超过 1×10^6 个点,为提高运行效率需要在不影响计算准确度和点云形貌的基础上进行下采样,一般可以进行 $10:1$ 下采样,因此本文算法的点云处理规模通常为 10^5 以下。当图像中出现多物体时,采用多线程技术处理点云可以提高效率,本文算法处理一对图像的点云时间 30 s 左右。当采用性能更加优秀的计算机并对目标检测模型进行剪枝、知识蒸馏以及采用更高的下采样倍率,都将缩减运行时间。

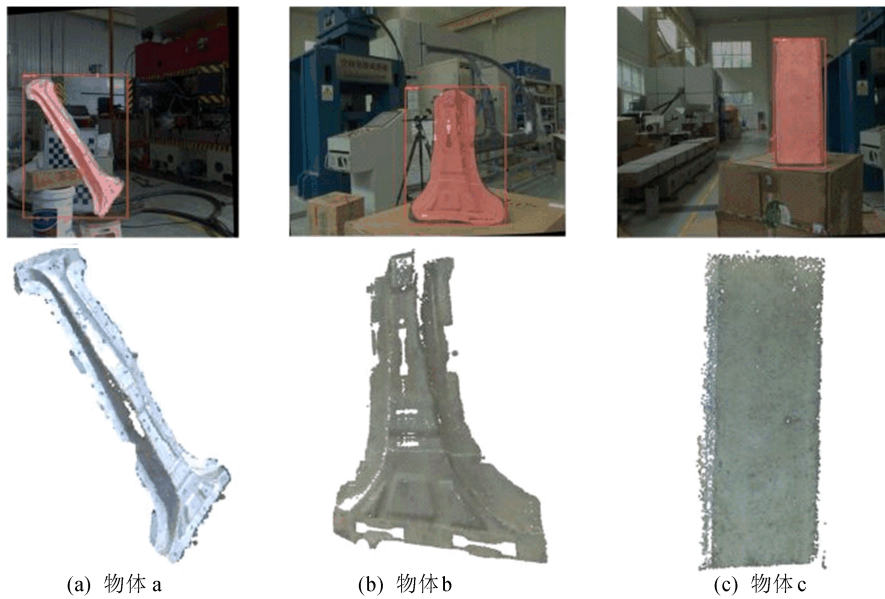


图 12 不同物体效果图

表 2 不同物体数据对比

| 物体 | 点云总数 | 保留点云数 | $E_a/\%$ | $E_b/\%$ | $I/\%$ | 颜色滤波耗时/s | 欧式聚类耗时/s | 总耗时/s |
|----|---------|--------|----------|----------|--------|----------|----------|-------|
| a | 107 142 | 22 908 | 0.615 | 2.807 | 90.132 | 7.2 | 8.9 | 21.6 |
| b | 96 105 | 42 495 | 1.042 | 6.815 | 88.128 | 12.1 | 11.3 | 30.5 |
| c | 50 493 | 37 843 | 0.206 | 3.987 | 94.998 | 8.3 | 10.2 | 24.7 |

根据 OBB 计算方法计算物体 a 点云的主方向和半长度大小,包围盒效果图如图 13 所示。

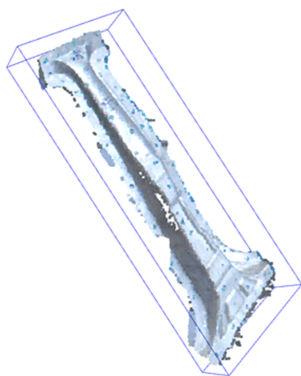


图 13 物体 a 的 OBB 效果

点云的长、宽、厚分别为 $1\,139.28$ 、 451.48 、 120.91 mm 。OBB 能够紧密地包裹住点云,较为准确地估计出点云的大小,实现了在钢铁回收过程中对废旧钢铁的尺寸评估,提高了钢铁回收效率和经济效益。

3 结 论

本文针对钢铁回收开放式环境下双目视觉点云数据,分析了场景和点云的几何、颜色、统计分布等特征信息,提出了一种针对钢铁回收开放式场景下的双目视觉钢铁点云分割和尺寸估计算法。首先通过基于 Mask R-CNN 的实例分割算法进行初步处理,将目标区域裁剪出来;然后通过

掩膜区域进行自适应 HSV 空间颜色阈值自适应选取,对裁剪区域进行三维重建生成点云;再对点云进行颜色滤波和欧式聚类,从而得到完整的目标点云;最后通过 OBB 对目标点云进行尺寸估计,得到点云的三维信息。

本文以废旧汽车 B 柱为实验目标,模拟回收中面临的开放式环境进行对比实验,实验结果表明,本文算法对汽车 B 柱点云的提取效果优于其他同类型算法。同时,对于点云边缘的完整性明显优于其他算法。算法还存在一些不足之处,如目标点云中存在少数点云的颜色与其他点云颜色不同,则该部分点云会被颜色滤波滤除,导致有一小部分目标点云缺失,点云完整度受到影响,因此进一步的工作是建立更完备的点云提取模型,以提高钢铁回收过程中钢铁点云的提取精度。

[参 考 文 献]

- [1] 杨春雨,顾振,张鑫,等. 基于深度学习的带式输送机煤流量双目视觉测量[J]. 仪器仪表学报, 2021, 41(8): 164-174.
- [2] 刘王超,娄鑫. 基于双目点云的三维物体检测准确性分析[J]. 中国科学院大学学报, 2022, 39(5): 677-683.
- [3] QI C R, SU H, MO K, et al. Pointnet: deep learning on point sets for 3rd classification and segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S. l.]: IEEE, 2017: 652-660.
- [4] QI C R, YI L, SU H, et al. Pointnet++: deep hierarchical feature learning on point sets in a metric space[J]. Advances in Neural Information Processing Systems, 2017, 30: 1-14.
- [5] KLOKOV R, LEMPITSKY V. Escape from cells: deep kd-networks for the recognition of 3d point cloud models[C]//Proceedings of the IEEE International Conference on Computer Vision. [S. l.]: IEEE, 2017: 863-872.
- [6] ENGELMANN F, BOKELOH M, FATHI A, et al. 3d-mpa: multi-proposal aggregation for 3d semantic instance segmentation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S. l.]: IEEE, 2020: 9031-9040.
- [7] NI H, LIN X, NING X, et al. Edge detection and feature line tracing in 3D-point clouds by analyzing geometric properties of neighborhoods[J]. Remote Sensing, 2016, 8(9): 710.
- [8] CHE E, OLSEN M J. Multi-scan segmentation of terrestrial laser scanning data based on normal variation analysis[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2018, 143: 233-248.
- [9] ADAMS R, BISCHOF L. Seeded region growing[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1994, 16(6): 641-647.
- [10] KANG C L, WANG F, ZONG M M, et al. Research on improved region growing point cloud algorithm[J]. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 2020, 42: 153-157.
- [11] 薛广顺,来智勇,张志毅,等. 基于双目立体视觉的复杂背景下的物体点云获取[J]. 计算机工程与设计, 2015, 36(5): 1390-1395.
- [12] 曹雄,林兆祥,宋沙磊,等. 基于颜色聚类的多光谱激光雷达点云去噪[J]. 激光与光电子学进展, 2021, 58(12): 499-506.
- [13] JOHNSON J W. Adapting mask-rcnn for automatic nucleus segmentation [EB/OL]. [2023-12-26]. <http://arxiv.org/pdf/1805.00500.pdf>.
- [14] GEHRIG S K, EBERLI F, MEYER T. A real-time low-power stereo vision engine using semi-global matching [C]//International Conference on Computer Vision Systems. [S. l.]: Springer, 2009: 134-143.
- [15] 朱俊锋,胡翔云,张祖勋,等. 多尺度点云噪声检测的密度分析法[J]. 测绘学报, 2015, 44(3): 282-291.
- [16] 谷帆,张长伦,郭志光,等. 提取桥梁点云的组合滤波算法[J]. 激光与光电子学进展, 2023, 60(2): 199-205.
- [17] 赖亦斌,陆声链,钱婷婷,等. 植物三维点云分割[J]. 应用科学学报, 2021, 39(4): 660-671.
- [18] BARRILE V, CANDELA G, FOTIA A. Point cloud segmentation using image processing techniques for structural analysis[J]. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, 2019: 181-193.

(责任编辑 李 凯)