

[自控·检测]

DOI:10.3969/j.issn.1005-2895.2021.04.013

基于 Yolov4-Tiny 与 RANSAC 算法的 管纱识别抓取系统

李乐乐, 张团善, 马浩然, 张越

(西安工程大学机电工程学院, 陕西西安 710048)

摘要:针对络筒机依赖人工上纱的非自动化问题,课题组提出一种基于 Yolov4-Tiny 目标检测模型的识别抓取系统,实现了管纱的检测、定位及抓取任务。首先利用深度相机获取 RGB-D 图像,通过训练深度学习网络 Yolov4-Tiny,生成预测结果;然后利用预测框信息将原始点云进行裁剪,采用快速点特征直方图与 RANSAC 算法进行配准;最后利用手眼标定的方法将深度图像坐标信息转换到机械臂坐标系下完成抓取。实验结果表明:系统管纱平均抓取成功率达到 65%,在非结构化的环境中具有良好的抓取结果,满足管纱抓取的的实际生产需求。

关键词:机器视觉;RGB-D 图像;RANSAC 算法;Yolov4-Tiny 模型;点云配准

中图分类号:TS103.8;TH164 文献标志码:A 文章编号:1005-2895(2021)04-0068-06

Yarn Recognition and Grasping System Based on Yolov4-Tiny and Ransac Algorithm

LI Lele, ZHANG Tuanshan, MA Haoran, ZHANG Yue

(School of Mechanical and Electrical Engineering, Xi'an Polytechnic University, Xi'an 710048, China)

Abstract: To solve the non-automatic problem that the winding machine relies on manual yarn loading, a recognition and grasping system based on Yolov4-Tiny object detection model was proposed to realize the detection, positioning and grasping tasks of the cop. Firstly, the depth cameras were used to obtain RGB-D images, and the forecast results were generated through training Yolov4-Tiny deep learning network. Then the original point cloud was clipped with the prediction information, and the fast point feature histogram and RANSAC algorithm were used for registration. Finally the method of hand-eye calibration was used to convert the depth image coordinates information to mechanical arm coordinate system to complete the capture. The experimental results show that the average successful grasping rate of the proposed system reaches 65%, which has good grasping results in the unstructured environment and meets the actual production requirements of grasping cop.

Keywords: machine vision; RGB-D images; Ransac algorithm; Yolov4-Tiny model; point cloud registration

传统的纺织行业是一个典型的劳动密集型产业^[1],既耗时又耗力。管纱是织造过程中的重要原料,其上纱的速度关系着整个织造流程的效率,但目前自动络筒机的上纱工作绝大多数依旧依赖人工进行分拣和放置。伴随着纺织行业生产向智能化、自动化转变,机器人代替手工进行全自动化作业已成为一种不可避免的趋势^[2]。通过视觉引导的抓取方法是目前

普遍的做法:通过图像处理技术获取目标物体的位置,建立像素坐标与机械臂坐标的关系,实现目标抓取。目前比较成熟的自动化抓取系统主要基于二维的图像特征,如瑞典 ABB 公司开发的 Flex Picker 机器人^[3]基于二维图像处理算法实现对物体平面位姿检测。但是堆叠管纱属于典型的 Bin-Picking 问题。近年来针对 Bin-Picking 出现了多种算法:王照华等^[4]将深度数组

收稿日期:2021-01-07;修回日期:2021-05-30

基金项目:西安市现代智能纺织装备重点实验室项目(2019220614SYS021CG043)。

第一作者简介:李乐乐(1996),男,陕西榆林人,硕士研究生,主要研究方向为计算机视觉、机器人抓取。E-mail:ll19961105@outlook.com

可视化,转为对二维图像中的物体进行识别,结合深度信息完成物体定位;Oh 等^[5]提出了基于双目相机的立体视觉算法(geometric pattern matching, GPM),通过对对象的几何特征来配准图像,极大地减少变形和光照的变化对系统稳定性的影响;Kong 等^[6]通过 SIFT 算法确定关键点,对比采集点云与预设场景点云的关键点信息,建立对应位置关系;Gong 等^[7]将采集到的点云进行超像素聚类,利用零件的凹凸性进行分割,再与模板数据进行比对。以上方法均可以实现三维空间下的物体识别与位姿检测,但仅适用于简单的识别目标和特定场景,普适性较差。

针对管纱上纱的抓取场景,课题组设计了一种基于 Yolov4-Tiny 目标检测模型的管纱抓取系统,系统通过 Yolov4-Tiny 网络进行纱管识别,结合裁剪点云和 RANSAC 算法进行点云配准计算出管纱位姿,最后驱动机械臂进行抓取、调整姿态并放置。

1 识别抓取系统构成

管纱抓取系统的实验平台主要由 6 自由度机械臂、深度相机和工业控制计算机组成,如图 1 所示。本系统选用的是 6 轴机械臂 S6H4D,该机械臂采用电动 2 指式末端执行器完成抓取操作,能够快速、平稳地完成各种位姿下物体的抓取任务。课题组选用了 Inter Realsense D435i 深度相机,平行安装于纱框正上方,作为数据采集系统。工控机中已预装 Pycharm2019、相机驱动软件和相关数据库,完成上位机控制指令的发送,实现相机与机械臂控制系统的通信。

管纱抓取系统组成及算法流程如图 1 所示。首先采集 RGB-D 图像信息,使用 RGB 图像对管纱进行训练识别,然后利用采集到的 RGB 图像结合深度信息,创建三维点云并进行下采样,依据识别结果对点云进行裁剪,并采用点快速特征直方图(fast point feature histograms, FPFH)对三维点云特征进行描述,而后采用 RANSAC 算法进行点云配准,得到位姿变换矩阵,最后通过手眼标定将抓取点位姿信息发送给机械臂的控制柜,引导机械臂完成管纱抓取。

2 基于 Yolov4-Tiny 模型的点云位姿检测算法

2.1 Yolov4-Tiny 模型

Yolov4 是 Yolo^[8]的第 4 个版本,属于单次级检测算法,该类算法将检测问题转换为回归问题,将物体的

定位和分类放在一个步骤中完成,因此检测速度上有较大提升^[9]。Yolov4-Tiny 是 Yolov4 的简化版,通过删减一些结构,减少参与计算的参数量级,在保证满足特定对象的精度的同时,进一步提升速度。Yolov4-Tiny 网络结构如图 2 所示,主要包括 CSPdarknet53-tiny 组成的主干特征提取网络、FPN 特征金字塔结构组成的加强特征提取网络以及 2 个尺度的 Yolo Head 预测框结构。Yolov4-Tiny 将所有激活函数修改为收敛速度更快的 LeakyReLU 函数:

$$y = \max(0, x) + l_{\text{leak}} * \min(0, x) \quad (1)$$

式中: y 为输出变量; x 为输入变量; l_{leak} 是一个很小的常数。

2.2 点云预处理与配准

得到 Yolov4-Tiny 网络的预测框后,依据预测框的信息对点云数组进行裁剪,同时保留其位置信息,预测框内除了所需的管纱数据外,还存在纱框及其他干扰物体,需进行预处理,包括体素下采样以及平面分割,以得到较为纯净的管纱点云信息。将裁剪后的点云与模板点云进行配准,得到变换矩阵。

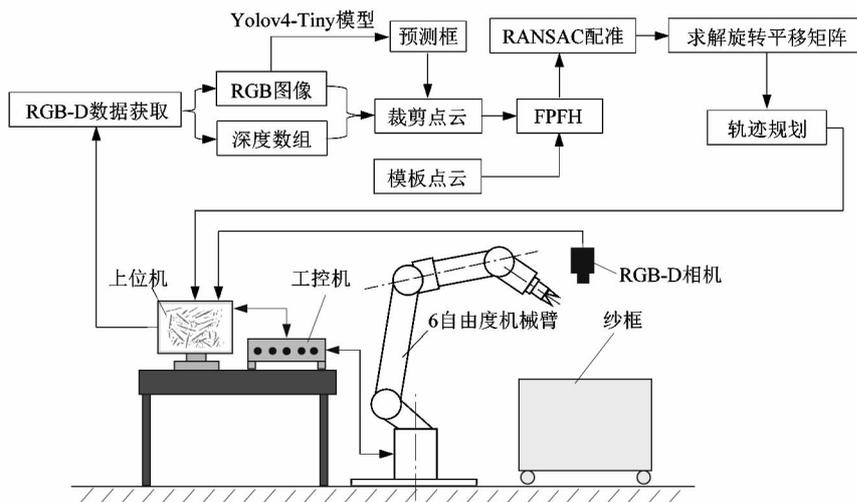


图 1 系统组成及算法流程图

Figure 1 System composition and algorithm flow chart

2.2.1 体素下采样

点云信息处理通常因数据量过大导致耗时较长,需按照一定的规则对获取到的点云进行“滤波”,其中体素下采样就是一种针对点云常用的滤波方法,以减少计算量。

主要包括 2 个步骤:

- 1) 将点规整地划分到三维体素栅格中;
- 2) 在每个体素栅格中,通过平均内部的所有点来精确生成一个点。

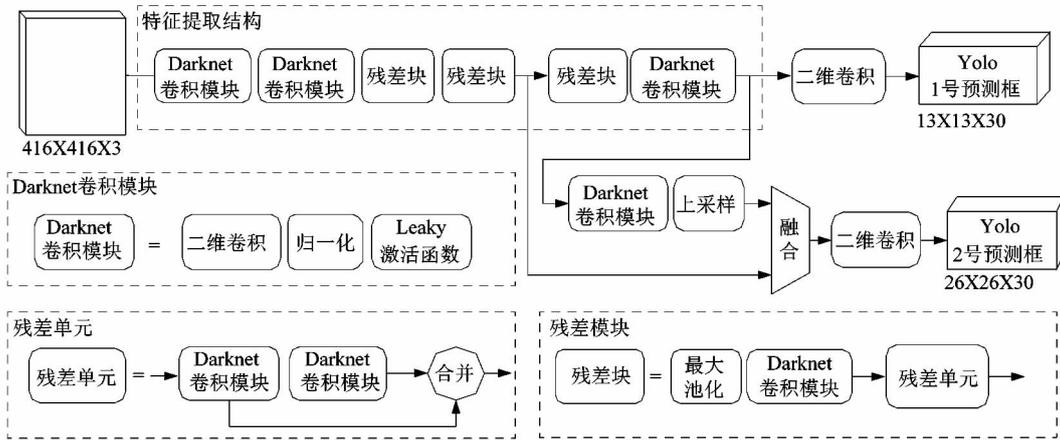


图 2 Yolov4-Tiny 网络结构图

Figure 2 Yolov4-Tiny network structure diagram

2.2.2 点云平面分割

为保证配准精度,需去除预测框内多余的点云数据,如纱框侧壁平面数据以及环境噪点。依靠平面拟合方法的平面方程可表示为: $ax + by + cz + d = 0$ 。

记平面上不共线的 3 个点 $P_i = [x_i, y_i, z_i]^T, i = 1, 2, 3$ 。根据平面方程可以得到:

$$A\theta = 0. \tag{2}$$

式中:

$$A = \begin{bmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{bmatrix}; \tag{3}$$

$$\theta = [a \quad b \quad c \quad d]. \tag{4}$$

求解 θ 即可。得到拟合出的平面方程后,设定阈值 m ,将 z 轴范围在 $(P_i + m, P_i - m)$ 内的值记为平面内点,其余点记为平面外点,对平面内点进行去除。

2.2.3 点云配准

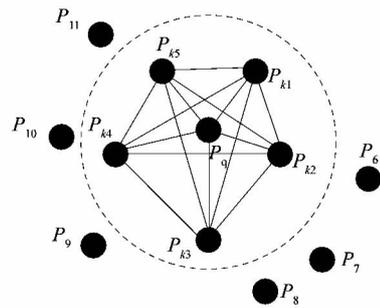
点特征直方图(point feature histogram, PFH)是通过将目标点及其临近点的几何特征关系参数化实现的,其采用高维直方图进行描述。点特征相对法线和曲面这类标量,具备更大的差异性。同时点特征能够保持平移旋转的不变性,具有较强的鲁棒性。

如图 3(a)所示,以点 P_q 为中心,以 R 为半径的邻域范围,表示一个查询点 P_q 的 PFH 计算影响区域, P_q 在圆球中间位置,与点 P_q 的距离小于半径 R 的所有点全部互相连接在一个网络中。最终的 PFH 描述是通过计算邻域内所有 2 点间关系而得到的直方图^{[10]36}。

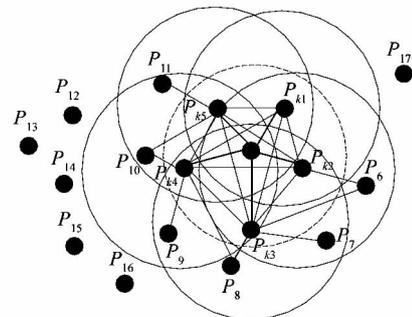
图 3(b)所示为快速点特征直方图 FPFH,记为 F ,从点特征直方图变化而来,可有效降低其运算时间。

对于每个计算的目标点,取其所有的邻域点组成点集,统计目标点与所有邻域点的三元数组 $(\alpha, \varphi, \theta)$ 形成直方图,记为 S 。对于点集中的每个点,都有其对应的 S 。之后使用所有点的 SPFH 的加权和来表示快速点特征直方图^{[9]38}。

$$F(P_q) = S(P_q) + \frac{1}{k} \sum_{i=1}^k \frac{1}{\omega_k} S(P_k). \tag{5}$$



(a) 快速点特征直方图 FPFH



(b) 快速点特征直方图 FPFH

图 3 PFH 和 FPFH 原理图

Figure 3 Schematic of PFH and FPFH

随机采样一致性 (random sample consensus, RANSAC) 配准是一种经典的点云配准方法,该算法在

全局范围内对待匹配对象和场景对象进行随机采样,并利用每组采样匹配的结果对所对应的转换质量进行排序,遴选出所有匹配结果中最佳的转换关系^[11]。设 $\{P_i^s\}$ 和 $\{P_i^m\}$ 分别为从场景中分割出来的点云对象及其对应的模型点云对象提取出来的 FPFH 特征集合,则:

1) 从 $\{P_i^s\}$ 中随机选出 s 个样本点,选取的样本点两两间距不得小于某一由用户设置的阈值 d_{\min} ,即选取的采样点的配准距离大于等于 d_{\min} ;

2) 对于每个样本点,在 $\{P_i^m\}$ 查找所有与样本点特征相近的点,并且将它们放在同一列表中,并从此列表中随机选取一个点与此样本点进行配对;

3) 计算样本点对应的刚性变换矩阵,并对其质量进行评价,选取胡贝尔(Huber)损失函数对误差进行度量。

胡贝尔(Huber)损失函数公式如下:

$$L_{\delta}(y, f(x)) = \begin{cases} \frac{1}{2}(y - f(x))^2, & \text{for } |y - f(x)| \leq \delta; \\ \delta|y - f(x)| - \frac{1}{2}\delta^2, & \text{除此以外。} \end{cases} \quad (6)$$

式中: δ 为设定的参数; y 表示目标值; $f(x)$ 表示预测值。

重复以上3步骤直到迭代达到一定次数。RANSAC算法的迭代次数无法精确确定,但可以根据其运行时间预测出大概范围,文中的迭代次数数量级为 $10^3 \sim 10^4$ 。最后从所有迭代过程产生记录的质量转换中选取评价指数最高的变换关系作为最佳变换关系进行点云配准^[12]。



图5 管纱识别结果

Figure 5 Results of cop identification

在实际的情况下,单个机械臂每次仅能完成1根管纱抓取任务,故每次仅需识别出少量待抓取管纱即可,实验保留2个管纱作为候选。故设定总预测框数量大于等于2即视为识别成功,本实验进行了100次,

3 实验结果

实验在Pycharm2019开发平台下,结合Open3D设计了抓取系统的控制界面,采用Inter RealSense D435i深度相机及其加持装置组成图像采集系统。实验环境为搭载Inter Core i5 CPU及NVIDIA GTX 960M图像处理器和Python3.6的工控机。

3.1 管纱检测实验

Yolov4-Tiny网络模型训练时,设置动量为0.9,每次送入网络中训练的数据数量为16,初始学习率 l_r 为0.001,学习率采用余弦退火的更新方式,衰减系数为0.1,总迭代次数为300。课题组所用的数据集取自实验室抓取场景,拍摄了一共400张照片。使用TensorFlow自带的可视化工具tensor board可查看模型训练情况,如图4所示,损失值随着迭代次数增加逐渐减小,并最终降至8.73。训练结束后,将训练好的模型在配置Inter Core i5 CPU和NVIDIA GTX 960M图像处理器的工控机上进行测试,部分实验结果如图5所示。

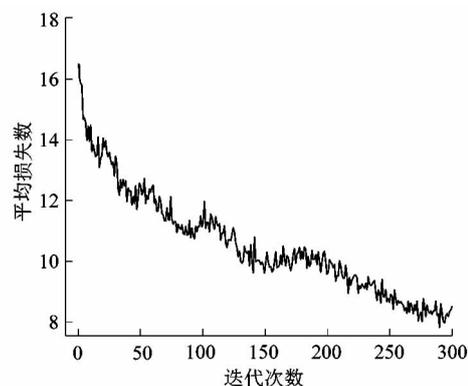


图4 损失值变化图

Figure 4 Loss diagram

其中检测结果预测框数量为3,2,1,0的测试图片张数分别为10,78,6,6,成功率达到88%。

3.2 点云预处理与配准实验

首先以5 cm为半径进行体素划分,并计算单位体

素中所有点组成的重心点,以此进行体素下采样,然后根据 Yolov4-Tiny 神经网络的输出结果获得预测框的

位置,根据置信度大小选取较大者对点云进行裁剪,效果如图 6 所示。

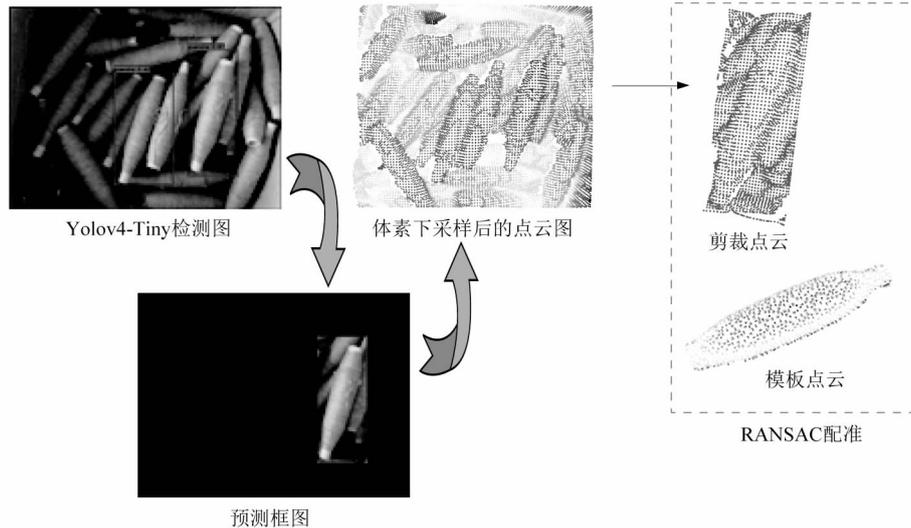


图 6 预处理结果

Figure 6 Preprocessing results

点云配准通过 FPFH 特征提取后用 RANSAC 算法进行多次迭代和验证,获得从模板点云到场景点云的变换矩阵。预测框点云裁剪、配准所耗时长及误差如表 1 所示。

实验验证表明:提出的基于 Yolov4-Tiny 模型的位姿检测算法在保证一定程度上准确率的同时,耗时更短。

表 1 预测框点云裁剪和配准耗时及误差

Table 1 Prediction frame point cloud clipping, registration time and error

项目	耗时/s	误差/mm
预测框点云裁剪	0.011	-
点云配准(FPFH + RANSAC)	0.080	0.5

3.3 管纱抓取实验

实验中的抓取系统采用眼在手外的模式,在光照充足的条件下进行了管纱的抓取实验,依照随机摆放的原则,其框内位置和姿态不同,进行了 6 批次抓取测试,每批次进行 10 次抓取,总共抓取 60 次。实验中使用 Yolov4-Tiny 模型对管纱进行识别,系统计算出机械臂抓取的位姿信息,并通过 SOCKET 通信发给机械臂控制柜,引导机械臂进行抓取,抓取过程如图 7 所示。

在相同硬件条件下,实验对比了基于双目视觉的位姿检测方法、基于点云聚类分割的位姿检测方法以及本系统中基于 Yolov4-Tiny 模型的点云位姿检测算法,实验结果如表 2 所示。

抓取实验结果如表 3 所示。Yolov4-Tiny 模型能稳定地识别到管纱,抓取总成功率达到 65%。

表 2 位姿检测方法对比

Table 2 Comparison of pose estimation methods

方法	耗时/s	准确率/%
文献[4]方法	3.3	63.8
文献[6]方法	5.2	73.9
本系统算法	2.3	71.2

表 3 管纱抓取实验结果

Table 3 Experimental results of cop capture

抓取批次	总抓取次数	成功次数	抓取率/%
1	10	6	60
2	10	7	70
3	10	5	50
4	10	8	80
5	10	5	50
6	10	8	80

4 结论

由表 2 可知,相较于传统的基于双目和点云聚类分割的位姿检测方法,本系统算法在检测耗时上平均低于 2 种传统方法 1.95 s,准确率介于二者之间。实

课题组提出一个基于点云数据,结合 Yolov4-Tiny 模型进行管纱识别和 RANSAC 算法进行点云配准的管纱抓取系统,系统通过 Yolov4-Tiny 网络模型能够较



图7 管纱抓取过程

Figure 7 Capture process of cop

为准确地识别和获取管纱位置信息,通过预测框的使用,避开了传统点云配准中最为耗时的点云聚类分割的步骤,从而大大提高了抓取过程的效率。由于管纱体积较大,对抓取位置精度要求不高,仅使用 RANSAC 算法进行配准即可获得较好的效果,省去了处理数据量更大的 ICP 算法,在点云配准上节约了大量时间,在保证一定精度的前提下,提高了抓取效率,为上纱系统提供了一种可行的方案。

参考文献:

- [1] 张帆,张团善,冀永乐,等. 基于机器视觉的纺纱管颜色分拣算法研究[J]. 西安工程大学学报,2018,32(5):560-566.
- [2] 宋薇,仇楠楠,沈林勇,等. 面向工业零件的机器人单目立体匹配和抓取[J]. 机器人,2018,40(6):951-958.
- [3] CONNOLLY C. A new integrated robot vision system from FANUC Robotics[J]. Industrial Robot,2007,34(2):103-106.
- [4] 王照华. 基于机器视觉的堆叠矩形纸盒识别方法研究[J]. 机械与电子,2019,37(6):75-80.
- [5] OH J K,LEE S,LEE C H. Stereo vision-based automation for a Bin-picking solution[J]. International Journal of Control, Automation and Systems,2012,10(2):362-373.
- [6] KONG M X,NIU C Y. An inertial navigation pre-integration algorithm based on multi-sensor fusion for indoor mobile robot positioning[C]//Proceeding of the 2019 2nd International Conference on Service Robotics Technologies. New York:ACM Press,2019:43-47.
- [7] GONG X J,CHEN M,YANG X J. Point cloud segmentation of 3D scattered parts sampled by RealSense[C]//2017 IEEE International Conference on Information and Automation. New York:IEEE,2017:1-6.
- [8] REDMON J,DIVVALA S,GIRSHICK R,et al. You only look once: unified, real-time object detection[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA:IEEE,2016:779-788.
- [9] 丘海斌,陈丹,王孝顺. 基于机器视觉的水表抓取系统[J]. 计算机系统应用,2020,29(3):81.
- [10] 龚学健. 基于 RealSense 的散乱零件三维目标识别[D]. 哈尔滨:哈尔滨工业大学,2018:36-38.
- [11] FISCHLER M A, BOLLES R C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography[J]. Graphics and Image Processing,1981,24(6):381-395.
- [12] 宋景豪. 三维点云数据配准方法研究[D]. 哈尔滨:哈尔滨工业大学,2014:35.