

# 基于视觉定位的机器人智能拆垛系统<sup>①</sup>

刘宝临<sup>1</sup>, 邹汶材<sup>2</sup>

<sup>1</sup>(西南交通大学 信息科学与技术学院, 成都 611756)

<sup>2</sup>(西南交通大学 计算机与人工智能学院, 成都 611756)

通信作者: 刘宝临, E-mail: [blLiu6@163.com](mailto:blLiu6@163.com)



**摘要:** 针对传统示教再现机器人仅能进行位置确定、轨迹固定的拆垛任务, 局限于固定场景的问题, 设计了一个基于视觉定位的机器人智能拆垛系统. 该系统利用目标像素中心坐标转换求得对应世界坐标. 针对眼在手外的安装相机方式, 导致目标经图像处理算法求得的旋转角度可能由于相机自身的偏转而产生误差的问题, 提出利用相机外参系数补偿目标旋转角度. 最后设计拆垛策略, 通信引导机器人以由近及远的抓取顺序执行拆垛任务, 并无需人工干预自动完成整垛拆卸. 经过实验数据表明, 该系统可在未知工作场景中对未知位置目标进行抓取, 位置误差可达 1.1 mm, 角度误差可达 1.2°, 堆垛一层定位时间为 1.2 s 左右, 满足工业场景中对拆垛机器人的精度与效率需求.

**关键词:** 视觉定位; 坐标转换; 位姿补偿; 拆垛策略; 自动拆卸

引用格式: 刘宝临, 邹汶材. 基于视觉定位的机器人智能拆垛系统. 计算机系统应用, 2023, 32(7): 138-144. <http://www.c-s-a.org.cn/1003-3254/9171.html>

## Intelligent Robot Depalletizing System Based on Visual Positioning

LIU Bao-Lin<sup>1</sup>, ZOU Wen-Cai<sup>2</sup>

<sup>1</sup>(School of Information Science and Technology, Southwest Jiaotong University, Chengdu 611756, China)

<sup>2</sup>(School of Computing and Artificial Intelligence, Southwest Jiaotong University, Chengdu 611756, China)

**Abstract:** An intelligent robot depalletizing system based on visual positioning is designed to solve the problem that the traditional teaching and playback robot can only perform depalletizing tasks with given positions and fixed trajectories and thus is limited to fixed scenes. The system uses the coordinate transformation of the target pixel center to obtain the corresponding world coordinates. For the problem that the eye-in-hand camera may lead to the inaccurate rotation angle of the target obtained by the image processing algorithm due to the deflection of the camera, it is proposed to use the extrinsic parameter coefficient of the camera to compensate for the rotation angle of the target. Moreover, a depalletization strategy is designed, and the communication guides the robot to automatically perform the depalletization task by grabbing from nearest to farthest without manual intervention. The experimental data shows that the system can grab the target with an unknown position in an unknown work scene, with a position error of 1.1 mm and an angle error of 1.2°, and the time to position the stacking layer is about 1.2 s. The system meets precision and efficiency requirements for depalletizing robots in the industrial scenes.

**Key words:** visual positioning; coordinate transformation; pose compensation; depalletization strategy; automatic depalletizing

拆垛任务广泛存在于物流、制造、生产行业中<sup>[1,2]</sup>, 传统示教再现机器人按照编程指令以固定的轨迹, 按照预先设定的方式对确定的位置进行抓取. 这种拆垛

方式虽然稳定快速, 但局限于某个固定场景, 缺乏对工作环境改变、目标实际位姿变动的感知能力, 同时机器人末端定位精度和工作效率受环境影响. 因此现阶段

① 收稿时间: 2022-12-14; 修改时间: 2023-01-20; 采用时间: 2023-02-23; csa 在线出版时间: 2023-04-28

CNKI 网络首发时间: 2023-04-28

段机器人主要应用于某固定场景中,进行相对标准且重复性较高的任务<sup>[3]</sup>。近年来,在示教再现工业机器人的基础上,为了使机器人智能化,研究者们将传感器集成在工业机器人上<sup>[4]</sup>,用于检测机器人工作环境以避免碰撞<sup>[5,6]</sup>,提高机器人末端传感器的位置估计精度<sup>[7,8]</sup>,这些改进使工业机器人对工作环境具备一定感知能力,形成了具备一定智能化能力的工业机器人。但智能化程度还无法满足当前制造业、物流业复杂多样的需求。

基于视觉定位的智能化工业机器人可以通过视觉设备感知周围环境,利用图像处理算法从视觉捕获的图像中提取所需信息<sup>[9]</sup>,引导机器人对未知位置的目标进行抓取,从而能够在未知环境进行工作。并且依靠低成本、易使用的优点,基于视觉定位的智能化工业机器人极大地降低了生产成本、提高了工作效率。所以目前基于视觉的工业机器人引导技术是当下机器人研究领域的热门方向。杜学丹<sup>[10]</sup>和夏晶等人<sup>[11]</sup>将机器视觉与深度学习相结合用于指导机器人的拆垛工作,将 Faster R-CNN (faster region-based convolutional neural network)、R-FCN (region-based fully convolutional network) 等优秀的目标检测与识别网络迁移到姿态估计问题中,实现对物体的位姿估计与抓取。Shin 等人<sup>[12]</sup>采用实例分割算法 Mask R-CNN (region-based convolutional neural network) 识别物体,眼在手外安装相机,并以相机拍摄到的场景图像信息为基础结合重心和方向算法获得物体姿态。但未考虑相机自身的旋转角度可能导致待抓取目标位姿不准确的情况。朱贺勇<sup>[13]</sup>和王德明等人<sup>[14]</sup>利用机器视觉从场景点云提取待抓取目标几何信息,使用基于局部特征描述子的 PPF (point pair feature)、4PCS (4-points congruent sets) 等确定目标位姿,供机器人抓取。Shin 等人<sup>[13]</sup>通过计算三维模型点对特征全局建模,存储成哈希表,然后计算场景点云点对特征,索引哈希表查找对应点,通过投票和 ICP (iterated closest points) 选取最优位姿。此方法精度较高,但算法成本较高且耗时较长,应用于拆垛场景中效率不高。Hinterstoisser 等人<sup>[15]</sup>和 Fan 等人<sup>[16]</sup>将机器视觉与模板匹配结合,使用彩色图像的梯度信息和物体表面的法向量、形状作为模板匹配的依据,获取物体位姿信息。Fan 等人<sup>[16]</sup>应用 SBM (shape-based matching) 方法识别物体,三维重建获取目标三维位姿,并提出了一种拆垛策略使机器人先按行后按列的方式自动拾取堆垛每一层物体。但是此策略并未考虑机器人与堆垛相对位置关系。如果固定优先抓取第一行目标,容易发生与其他目标碰撞。

针对以上将视觉应用在机器人拆垛任务中出现的相机旋转容易带来位姿误差、效率不够高、拆垛过程发生碰撞等问题,设计了一个智能拆垛系统。结合相机外参系数解决相机本身偏转对待抓取目标位姿带来误差的情况,生成更为准确的旋转角度。与二维像素中心坐标结合深度信息经坐标转换得到的世界三维坐标相结合,共同组成所有待抓取目标的四维位姿。利用深度信息进行非顶层目标过滤,一次性实现堆垛顶层所有目标定位。通过生成机器人抓取顺序为相较机器人底座由近及远的点位信息串,解决先按行后按列进行抓取时导致夹具容易与其他目标发生碰撞的问题, TCP 通信引导机器人安全地自动执行整垛拆卸。最后在自采垛型数据集上和实际搭建的硬件抓取执行系统上对所提出的系统进行实验,验证机器人智能拆垛系统可在工作场景中以较高的效率进行准确、安全地拆垛任务。

## 1 系统框架

机器人智能拆垛系统通过网线与工业相机连接,以太网与机器人连接。待抓取目标位于托盘上,工业相机位于工作区域上方。整个系统的框架如图 1 所示。首先利用张正友棋盘格标定法<sup>[17]</sup>对相机进行标定,标定参数一方面用于确定像素坐标系到托盘坐标系的转换关系。另一方面用于补偿相机本身偏转对目标位姿带来的误差。在实际应用时,使用工业相机对工作区域进行拍照,获取当前堆垛的深度图像与彩色图像。采用深度学习算法 Mask R-CNN<sup>[18]</sup>对彩色图像进行分割识别,并通过获取分割实例掩膜最小外接矩形的图像处理得到所有目标像素中心和外接矩形框的旋转角度。利用像素坐标系到托盘坐标系的转换关系与深度图像信息将目标二维像素中心坐标转换得到目标在托盘上的三维坐标,并同时利用位姿补偿公式补偿最小外接矩形框的旋转角度。共同组成目标的四维位姿。顶层目标过滤后,一次实现堆垛顶层所有待抓取目标定位。最后通过设计拆垛策略生成相较机器人底座由近及远的点位信息串发送给拆垛机器人,引导机器人以由近及远的抓取顺序完成准确、安全的拆垛工作。

## 2 视觉定位

待抓取目标的位姿由其托盘坐标和旋转角度构成,用于描述目标在托盘上的状态,默认待抓取目标位于平面上,用 4 自由度位姿 $(X, Y, Z, R)$ 表示, $(X, Y, Z)$ 描述目标在托盘上的三维坐标, $R$ 描述目标在平面的旋转状态即目标绕托盘坐标系 $z$ 轴的旋转角度。为实现目标的视觉定位,需

首先通过标定出当前工作场景下的相机的外参数, 确定像素坐标系与托盘坐标系之间的转换关系, 生成像素坐标

系到托盘坐标系的转换公式. 并利用外参数补偿相机自身偏转对目标旋转角度带来的误差, 生成位姿补偿公式.

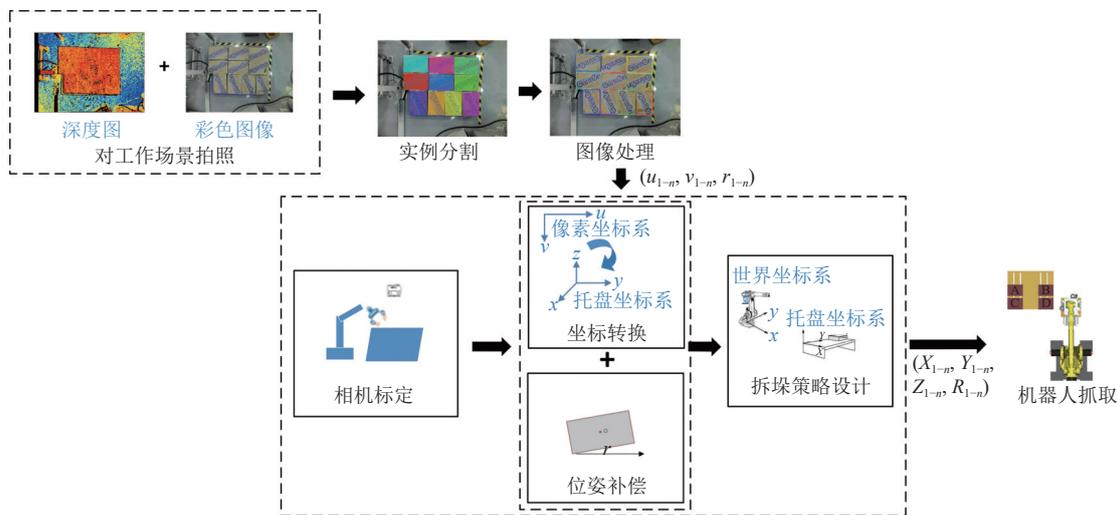


图1 机器人智能拆垛系统框架

## 2.1 相机标定

为获取像素坐标系与托盘坐标系的转换参数并消除相机自身偏转对目标旋转角度的影响, 本文使用张正友棋盘格标定法<sup>[17]</sup>对工业相机 FMB11-E2 进行标定. 首先通过编译相机 SDK 导出相机的内参数与畸变系数, 使用畸变系数校正由于镜头畸变引起的图像的变形失真<sup>[19]</sup>. 然后将棋盘格置于托盘左上角, 借助棋盘格对相机外参矩阵进行标定. 对于工业相机位于工作区域上方, 棋盘格位于托盘左上角的工作场景, 标定出托盘坐标系到相机坐标系的旋转平移矩阵即相机外参矩阵  $[R_{TrayToCamera} | T_{TrayToCamera}]$ . 为后续生成像素坐标系到托盘坐标系的转换公式与位姿补偿公式提供参数.

### 2.1.1 坐标转换

根据针孔成像模型<sup>[20]</sup>, 结合相机标定中获得的相机外参数和编译相机 SDK 导出的相机内参数, 可以确定像素坐标系与托盘坐标系之间的转换关系. 某一待抓取目标中心在托盘坐标系坐标下的位置  $(X, Y, Z)$  与其像素投影点坐标  $(u, v)$  的转换关系表示为式 (1):

$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = Z_c \begin{pmatrix} R_{TrayToCamera} & T_{TrayToCamera} \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \frac{f}{dx} & 0 & u_0 & 0 \\ 0 & \frac{f}{dy} & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}^{-1} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \quad (1)$$

其中,  $f$  为像距,  $dx$ 、 $dy$  分别表示  $x$ 、 $y$  方向上的一个像素在相机感光板上的物理长度,  $u_0$ 、 $v_0$  分别表示相机感光板中心在像素坐标系下的坐标, 这 5 个参数为工业相机的内部参数, 共同组成内参矩阵.  $Z_c$  是工业相机采集到的当前目标表面中心距离相机的深度信息, 从深度图像中获得.

### 2.1.2 位姿补偿

确定待抓取目标角度, 若仅以相机拍摄到的彩色图像为基础求取旋转角度, 当相机发生旋转会导致获取到的目标旋转角度叠加上相机自身的旋转角度, 使最后获取到的位姿误差较大. 针对此问题, 本文提出通过相机外参数对目标的旋转角度进行补偿的方法.

假设在工作场景中相机水平未旋转时, 将托盘坐标系与相机坐标系的方向重合, 需以内旋的方式、 $z-y-x$  的旋转顺序将托盘坐标系分别旋转  $\alpha$ 、 $\beta$ 、 $\gamma$ . 由于目标处于平面上, 位姿补偿可以通过求得当前工作场景下托盘坐标系相较于相机坐标系的绕  $z$  轴的旋转角度, 以此补偿待抓取目标因相机旋转导致的角度误差. 具体做法是将当前工作场景下标定出的旋转矩阵  $R_{TrayToCamera}$  经罗德里格斯 (Rodrigues) 变换<sup>[21]</sup>, 转为旋转向量  $\theta$ 、 $\varphi$ 、 $\psi$ , 分别指当前工作场景下将托盘坐标系与相机坐标系方向重合, 需分别绕  $x$ 、 $y$ 、 $z$  轴的旋转角度. 再以式 (2) 补偿待抓取目标的旋转角度, 得到目标在水平面上的实际旋转角度. 其中  $r$  是图像中待抓取目标的旋转角度, 定义  $R$  为位姿补偿后物体在托盘上的实际旋转角度.

$$R = r - \gamma + \psi \quad (2)$$

## 2.2 目标定位实现

在实际工作场景中,利用训练好的实例分割框架与目标最小外接矩形的图像处理方法对工业相机获取到的当前堆垛的彩色图像进行处理.并结合像素坐标系到托盘坐标系的转换式(1)和位姿补偿式(2),获取到所有待抓取目标四维位姿,顶层目标过滤后实现堆垛顶层所有目标的视觉定位.

### 2.2.1 实例分割框架

本文使用的实例分割框架,采用深度学习算法 Mask R-CNN<sup>[18]</sup>作为主网络,网络结构如图2所示.

Mask R-CNN<sup>[18]</sup>改进了 Faster R-CNN 采用的 FCN

特征提取网络,采用 FPN 结构进行特征融合,有效提升了目标检测的精度,能对箱体有较好的检测和定位.同时 Mask R-CNN<sup>[18]</sup>中的 Mask 分支能够对目标进行掩膜分割,为每个待抓取实例生成一个高质量掩膜 Mask,从而利用掩膜计算目标更精确的最小外接矩形,最终获得拣选目标中心以及水平旋转角度信息.算法实现过程中,以相机对当前工作场景下不同箱型码放形成的堆垛进行采集而生成的彩色图像作为数据集,对数据集进行标注训练后,可以精确识别出箱体、罐装等目标.使用训练好的 Mask R-CNN<sup>[18]</sup>实例分割模型对相机采集得到的彩色图像进行处理,可以实现二维图像中待抓取目标实例掩膜的精确像素级提取,如图3所示.

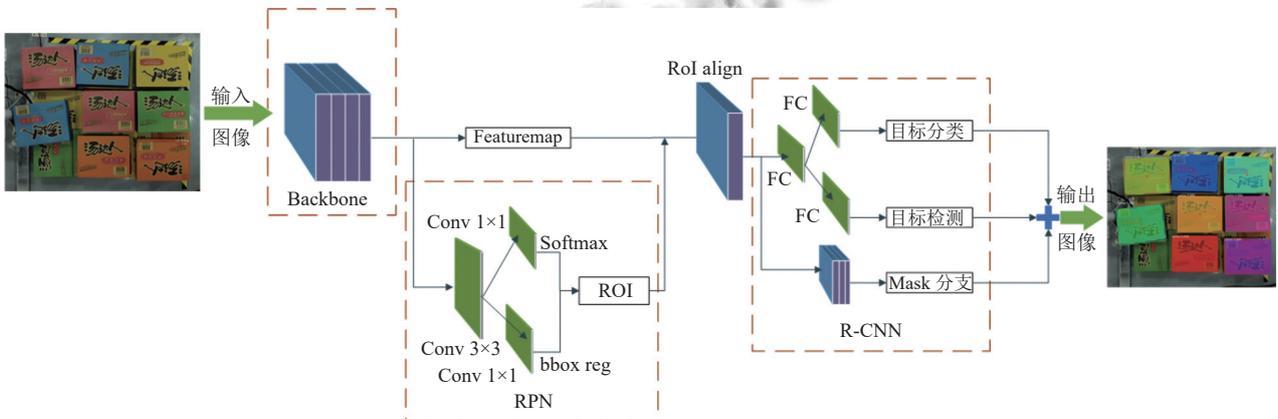


图2 Mask R-CNN 网络结构图

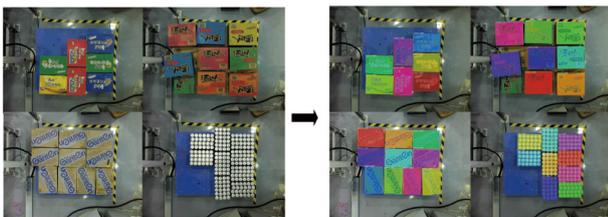


图3 目标实例分割

### 2.2.2 定位方法

在实际工作场景中使用抓手或者吸盘进行拆垛任务,需要定位目标中心位置以及目标旋转位姿以保证抓取运送过程中的准确性与稳定性,对于工业化纸箱箱体和热缩膜易拉罐等方形目标,目标中心可以通过最小外接矩形中心来获得.利用抓取目标的实例掩膜,使用凸包算法和旋转卡壳算法,循环计算分割出的每个目标实例掩膜的轮廓点集的最小外接矩形,从而获得外接矩形像素中心和旋转角度用作拣选.定义目标像素中心为 $O$ ,定义旋转角度为 $r$ .处理方法如下.

(1) 获取最小凸边形.在一个实数向量空间中,对于给定集合 $M$ ,所有包含 $M$ 的凸集的交集 $P$ 被称为 $M$ 的凸包,即刚好箱体某实例轮廓点集的一个凸多边形.对于实例轮廓,先使用凸包算法中的斯卡兰奇三硬币算法获得实例轮廓的最小凸多边形 $E$ .

(2) 在获得 $E$ 后,需要计算其最小外接矩形.通过不断枚举实例轮廓的边做外接矩形,比较外接矩形的面积,取其中面积最小的外接矩形作为实例轮廓的最小外接矩形.并且在枚举过程中记录外接矩形的左上角顶点 $(x_1, y_1)$ 、长度 $l$ 、宽度 $w$ 、旋转角度 $r$ .

定义像素中心坐标 $O$ 为 $(x_1 + l/2, y_1 + w/2)$ 旋转角度为 $r$ .其中 $r$ 以一个箱体的最小外接矩形为例,当水平轴逆时针旋转碰到最小外接矩形的第1条边为外接矩形长边,则以角度 $r$ 作为目标的旋转角度,如图4(a)所示.当碰到最小外接矩形的第1条边为外接矩形短边,且此时旋转角度为 $r_0$ ,则待抓取目标的旋转角度 $r$ 为 $r_0$ 的余角并取相反数,如图4(b)所示.

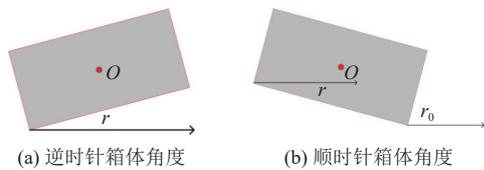


图4 基于最小外接矩形的旋转角度获取方法

获得目标的像素中心与旋转角度后,利用坐标转换式(1),结合深度图像信息对目标像素坐标进行转换得到目标托盘坐标,同时利用位姿补偿式(2)结合外接矩形框的旋转角度,补偿得到目标在托盘上的实际旋转角度.共同组成图像中所有目标的四维位姿 $(X, Y, Z, R)$ .去除在托盘坐标系中 $Z$ 小于高度阈值的点位信息进行顶层过滤,实现堆垛顶层所有待抓取目标的定位.其中阈值根据箱型的高度以及码放的层数来决定.

### 3 拆垛控制

拆垛控制由拆垛策略设计、机器人通信组成.策略设计根据不同的应用场景生成符合该应用场景下的拆垛顺序,以由近及远的拆垛顺序进行拆垛任务,有效地避免夹具在进行抓取的过程中碰撞到其他目标.机器人通信控制机器人末端进行抓取任务,并在完成顶层所有目标抓取后自动拍照分割识别进行下一层抓取.另外,为了处理异常情况,系统中设计了异常捕捉与反馈,如:目标位姿识别错误、通信异常等,并提供暂停和继续按钮,供现场工作人员使用.

#### 3.1 拆垛策略设计

根据机器人进行拆垛任务时需要先抓取离自己最近目标的需求,设计拆垛策略.需将目标在托盘坐标系中的位姿转换为机器人基坐标系下的位姿,通过校准托盘坐标系原点在机器人基坐标系的位置设定用户坐标系自动生成从托盘坐标系到世界坐标系的坐标转换矩阵 $[R_{\text{TrayToRobot}} | T_{\text{TrayToRobot}}]$ .定义机器人基坐标系为世界坐标系原点.

利用托盘坐标系到世界坐标系的坐标转换矩阵 $[R_{\text{TrayToRobot}} | T_{\text{TrayToRobot}}]$ 将堆垛顶层所有待抓取目标的位姿转换为机器人坐标系下的位姿.取四维位姿中的目标 $X, Y$ 值计算目标与世界坐标系原点即机器人基座的距离.再根据此距离进行从小到大排序,以此顺序发送堆垛顶层所有待抓取目标的四维位姿,满足工业上机器人需要优先抓取离自己最近目标的需求.

#### 3.2 机器人通信

为实现机器人的自动化运行,减少人工干预, TCP

通信实现机器人智能拆垛系统和机器人抓取过程相互通信.并通过设计监听线程,实现在顶层所有目标抓取完成后,自动拍照分割识别进行下一层抓取,直到完成整垛拆卸.

拆垛系统通过 TCP Socket 套接字与机器人的 Kernal 程序进行 TCP 通信连接.机器人端通过编写生成 Kernal 程序接收从设定端口中发来的待抓取目标中心位姿信息串.发送位姿串定义为 $data$ ,例如:

$$data = xyzr$$

表示一个待抓取目标的位姿信息,为实现自动抓取完一个堆垛中所有目标,添加一个额外线程用于监听该层是否发送完毕,当发送完毕后自动启动拍照分割流程进行下一层拆垛.以此循环,直到拍摄图片中未分割出待抓取实例,停止发送.

### 4 实验结果与分析

网络训练中,训练集和测试集为4种箱型(如图5所示)组成的混合垛型图像.其中未增强图像180张,分辨率为 $1280 \times 960$ .为了增加模型训练集的鲁棒性和泛化性,采用改变图像亮度、模糊图像、翻转平移图像、增加噪声等图像增强技术来模拟现场中不同环境下采集的质量不一的图像,增强后训练集为340张,测试集60张.利用Labelme标注软件对图像进行实例轮廓标注,生成JSON文件并转化为COCO数据集格式的训练集和测试集.训练显卡为GTX 1080Ti.基于TensorFlow-GPU 1.14版本深度学习框架,训练时学习率设置为0.001,训练步数为80个epoch.

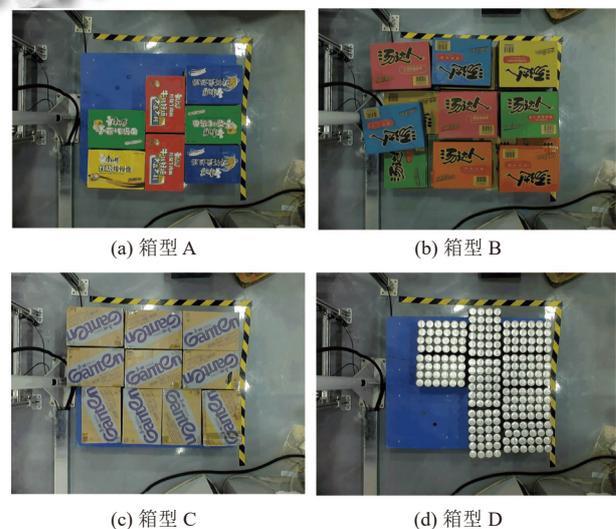


图5 4种箱型

实际操作中,整个实验平台由智能拆垛软件控制系统和硬件抓取执行系统组成.智能拆垛软件控制系统的开发环境为 Windows 10 x64,在 Visual Studio 2019 中编译,由 OpenCV 3.4.7 提供图像处理程序的相关类库和 Eigen-3.4.0 提供矩阵向量等线性代数操作运算的相关类库支持.硬件抓取执行系统由 FMB11-E2 工业相机、示教器、6 轴工业机器人、托盘、待抓取目标组成.针对包括箱体、罐装的 4 种箱型(A、B、C、D)组合而成的标准码垛和混箱码垛进行定位实验.共测试 8 种垛型视觉定位结果,验证机器人智能拆垛系统能对未知位置进行识别定位,可在工作场景中准确、安全地进行拆垛任务.

#### 4.1 位姿补偿实验

首先将相机分别顺时针和逆时针旋转  $20^\circ$  后,对 A 型标准码垛进行位姿补偿实验,将 A 型标准码垛的 7 个箱体位姿补偿前的旋转角度、位姿补偿后的旋转角度、多次人工测量结果的平均值进行对比,如表 1 所示.可以看到位姿补偿前待抓取目标旋转角度  $r$  与真实测量的旋转角度  $R$  相差的度数绝对值接近相机自身的旋转角度.顺时针旋转相机后,图像中箱体的角度平均误差达到  $20.7^\circ$ ,逆时针旋转相机,角度平均误差达到  $20.2^\circ$ .

表 1 A 型标码首层位姿补偿实验结果( $^\circ$ )

补偿类型	箱体1	箱体2	箱体3	箱体4	箱体5	箱体6	箱体7
位姿补偿前 $r$ (顺时针)	-69.5	81.5	64.5	52.8	-11.8	-24.5	-38.1
位姿补偿后 $R$ (顺时针)	-89.5	61.5	44.5	32.8	-31.8	-44.5	-58.1
实际测量 $R$	-90	60	45	30	-30	-45	-60
位姿补偿前 $r$ (逆时针)	71.0	38.9	24.3	12	-52	-64.5	-81
位姿补偿后 $R$ (逆时针)	91	58.9	44.3	32	-32	-44.5	-61
实际测量 $R$	90	60	45	30	-30	-45	-60

由此误差可知,当相机自身存在偏转时,以相机拍摄到的图像信息为基础求取目标的旋转角度时,会导致最终求得的目标自身的旋转角度叠加上相机的偏转角度,误差较大.分别对相机顺时针旋转后和逆时针旋转后求得所有目标的旋转角度进行位姿补偿.顺时针旋转相机并位姿补偿后,箱体旋转角度误差降至  $1.4^\circ$ .逆时针旋转相机并位姿补偿后,箱体旋转角度误差降至  $1.2^\circ$ .证明了位姿补偿的有效性.

#### 4.2 视觉定位实验结果与分析

再针对 4 种箱型组合而成的 8 种标准码垛和混箱码垛进行定位实验.表 2 为 8 种垛型在 X 轴、Y 轴、Z 轴方向的平均定位误差与旋转角度的平均计算误差.其中每个垛形顶层的定位误差为视觉定位计算所得的三维坐标与多次人工测量结果的平均值之间的误差.

实验数据表明,视觉定位精度稳定在  $1.1\sim 5.5\text{ mm}$  之间.位姿补偿后对平面上的物体的旋转角度定位的误差在  $1.2^\circ\sim 2.7^\circ$  之间.

表 2 不同垛型平均定位误差

待拆垛型	视觉定位平均误差(mm)			角度( $^\circ$ )
	X	Y	Z	R
A型标码	1.6	2.1	2.9	1.2
B型标码	2.9	2.4	2.4	1.2
C型标码	2.7	3.4	5.4	2.5
D型标码	2.4	1.1	2.7	1.2
A+B混码	3.7	3.1	2.5	1.7
A+C混码	2.8	1.9	5.5	2.7
B+C混码	2.3	3.5	4.2	1.3
A+B+C+D混码	3.7	4.1	1.6	2.5

误差产生的原因主要有两点:首先是实例分割模型的训练数据集的大小会影响在实际工作场景中对彩色图像的分割结果,训练集越大,分割效果越好,像素中心与世界坐标就越准确,误差越小.另外人为测量时由于精度等问题也会导致误差.但是微小的位置位移并不影响吸盘的成功抓取,也从一定的角度上说明了搭建的智能拆垛软件系统的稳定性.

#### 4.3 抓取实验

完成策略设计与通信设置后,工业机器人以拆垛策略中生成的顺序接受抓取点位信息,控制带有吸盘的机器臂到达指定位置进行抓取任务,图 6 为抓取中间过程,对堆垛顶层的 4 个箱体进行依次抓取.从抓取过程中可以发现,机器人优先抓取离机器人基坐标系最近的目标,并以由近及远的顺序地实现堆垛顶层所有目标的抓取.



图 6 抓取过程

定义机器人智能拆垛系统的响应时间为 $T_1$ ,包括相机采集时间、堆垛顶层目标位姿计算时间.对于堆垛每层目标的抓取,只需执行一次智能拆垛系统.再定义硬件抓取执行系统的工作周期 $T_2$ ,包括一个目标位姿的通讯时间和机器臂从开始抓取到再放置目标到指定地点的时间,经过实验过程计时,发现 $T_1$ 为1.2 s左右, $T_2$ 为15 s左右.执行系统的工作周期大部分是机器臂运动时间,响应时间相对于堆垛顶层所有目标抓取的工作时间占比很小,可以忽略不计.机器人智能拆垛系统的响应时间不会对整个抓取系统的效率造成制约.

## 5 总结

针对传统示教再现机器人仅能进行位置确定、轨迹固定的拆垛任务,局限于某个固定场景的问题,设计了基于视觉定位的机器人智能拆垛系统,通过双目工业相机引导机器人对未知位置的目标执行抓取任务,在未知环境中进行工作.系统经过对4种箱型码放而成的8种垛型进行定位抓取实验,结果证明,基于深度学习实例分割算法的视觉定位方法能够有效实现堆垛顶层的未知位置的目标识别定位,误差稳定在1.1–5.5 mm之间;改善相机本身偏转对待抓取目标旋转角度带来误差的位姿补偿方法对待抓取目标旋转角度的识别误差稳定在 $1.2^\circ$ – $2.7^\circ$ 之间;堆垛顶层所有目标地位姿计算时间为1.2 s左右.拆垛策略实现机器人按照由近及远的拆垛顺序到达待拆目标的准确位置执行抓取.经验证,系统准确性、适用性可满足实际工业拆垛任务需求.

## 参考文献

- 徐翔斌,马中强.基于移动机器人的拣货系统研究进展.自动化学报,2022,48(1):1–20.[doi:10.16383/j.aas.c190728]
- 陈阳,郑甲红,王婧.双机器人协同控制研究综述.计算机系统应用,2022,31(2):13–21.[doi:10.15888/j.cnki.csa.008319]
- 王诗宇.智能化工业机器人视觉系统关键技术研究[博士学位论文].沈阳:中国科学院大学(中国科学院沈阳计算技术研究所),2021.
- 蒋新松.未来机器人技术发展方向探讨.机器人,1996,(5):30–36.
- Hoffmann A, Poepfel A, Schierl A, et al. Environment-aware proximity detection with capacitive sensors for human-robot-interaction. Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems. Daejeon: IEEE, 2016. 145–150.
- Antonelli D, Bruno G. Human-robot collaboration using industrial robots. Proceedings of the 2nd International Conference on Electrical, Automation and Mechanical Engineering. Amsterdam: Atlantis Press, 2017. 99–102.
- Posada JRD, Schneider U, Pidan S, et al. High accurate robotic drilling with external sensor and compliance model-based compensation. Proceedings of the 2016 IEEE International Conference on Robotics and Automation. Stockholm: IEEE, 2016. 3901–3907.
- 孟少华,胡瑞钦,张立建,等.一种基于机器人的航天器大型部件自主装配方法.机器人,2018,40(1):81–88,101.
- 尹仕斌,任永杰,刘涛,等.机器视觉技术在现代汽车制造中的应用综述.光学学报,2018,38(8):0815001.
- 杜学丹,蔡莹皓,鲁涛,等.一种基于深度学习的机械臂抓取方法.机器人,2017,39(6):820–828,837.
- 夏晶,钱堃,马旭东,等.基于级联卷积神经网络的机器人平面抓取位姿快速检测.机器人,2018,40(6):794–802.[doi:10.13973/j.cnki.robot.170702]
- Shin H, Hwang H, Yoon H, et al. Integration of deep learning-based object recognition and robot manipulator for grasping objects. Proceedings of the 16th International Conference on Ubiquitous Robots (UR). Jeju: IEEE, 2019. 174–178.
- 朱贺勇.物体三维重构与位姿估计[硕士学位论文].杭州:浙江大学,2022.
- 王德明,颜熠,周光亮,等.基于实例分割网络与迭代优化方法的3D视觉分拣系统.机器人,2019,41(5):637–648.
- Hinterstoisser S, Lepetit V, Ilic S, et al. Model based training, detection and pose estimation of texture-less 3D objects in heavily cluttered scenes. Proceedings of the 11th Asian Conference on Computer Vision. Daejeon: Springer, 2012. 548–562.
- Fan XJ, Liu XG, Wang XL, et al. An automatic robot unstacking system based on binocular stereo vision. Proceedings of the 2014 IEEE International Conference on Security, Pattern Analysis, and Cybernetics. Wuhan: IEEE, 2014. 86–90.
- Zhang ZY. Flexible camera calibration by viewing a plane from unknown orientations. Proceedings of the 7th IEEE International Conference on Computer Vision. Kerkyra: IEEE, 1999. 666–673.
- He KM, Gkioxari G, Dollár P, et al. Mask R-CNN. Proceedings of the 2017 IEEE International Conference on Computer Vision. Venice: IEEE, 2017. 2980–2988.
- 周前飞,刘晶红.航空变焦镜头非线性畸变快速校正方法.光学学报,2015,35(4):0411001.
- 康佳.基于复杂工业场景的拆垛机器人视觉定位技术的研究[硕士学位论文].北京:北京邮电大学,2021.
- 茅凌波,史金龙,周志强,等.基于单视图关键点投票的机器人抓取方法.计算机集成制造系统,1–16.https://kns.cnki.net/kcms2/article/abstract?v=3uoqlhG8C45S0n9fL2suRadTyEVl2pW9UrhTDCdPD65zPhUaeS--asbRNezWrTHW8H82BDL2JS03FgvdttbFyhF\_DzNcSGHi&uniplatform=NZKPT.(2021-08-24)[2022-07-16].

(校对责编:孙君艳)