

doi:10.16055/j.issn.1672-058X.2021.0005.011

# 面向机器人抓取的双目视觉单步多目标检测方法应用研究\*

李世裴, 李春琳, 韩家哺, 朱新龙

(上海工程技术大学 机械与汽车工程学院, 上海 201600)

**摘要:**针对工业机器人的抓取技术普遍采用离线示教的方式,目标物的改变就会导致抓取不准的问题,研究了一种面向机器人抓取的双目视觉单步多目标检测方法,强化了机器人的抓取能力;首先搭建机器人抓取视觉检测系统,系统采用双目测距模型定位深度,单步多目标检测器定位像素坐标;通过相机标定和 eye-to-Hand 手眼标定将图像中的像素坐标以及双目测距模型的深度坐标转换为机器人基坐标的位姿,建立机器人运动学模型完成机器人基坐标到机器人末端坐标的转换,通过 ROS 实现上位机与机器人通讯,并将定位结果发送给 ROS;经过多次实验,表明方法的目标检测误差在 3.5 mm 以内,深度误差在 1.2 mm 之内,具有很好的实用推广价值。

**关键词:**单步多目标检测;双目视觉;机器人抓取;相机标定

**中图分类号:**TP249

**文献标志码:**A

**文章编号:**1672-058X(2021)05-0068-07

## 0 引言

机器人抓取技术是指机器人从复杂的实际环境中通过相机识别出目标物体<sup>[1]</sup>,完成对目标物的定位,进而进行抓取操作。目标检测作为最核心问题之一,是一种基于目标几何和统计特征的图像分割技术<sup>[2]</sup>,尤其在复杂场景中进行处理时,目标检测的准确性和快速性显得尤其重要。传统视觉<sup>[3]</sup>利用图像的几何特征、颜色特征、局部特征等来进行图像的特征提取,这种方法无法检测多种目标物,也无法模拟真实的复杂工业环境,将深度学习与图像相结合<sup>[4]</sup>,可以取得很好的效果,深度学习在图像上具有很高的特征提取能力<sup>[5]</sup>,能快速有效地对深度图像进行特征提取,将特征进行一系列的处理,最后输出完成对目标物的分类和定位<sup>[6]</sup>。Navneet Dalal 等人使用描述方向梯度直方图的 HOG<sup>[7]</sup>将图像边缘信息和形状信息由梯度信息描述,并使用线性 SVM<sup>[8]</sup>作为分类器完成对目标物的提取,这种方法

可以完成对单个物体的识别,这种基于机器学习的目标检测方法为目标检测提供了一种新的思路,但是这种方法仅仅局限于一种物体,且在复杂的工业场景下,难以实现对目标物的精准识别。夏晶等<sup>[9]</sup>提出一套基于 Mask R-CNN 的目标检测方法提取三维场景中的可抓取区域,利用深度学习完成对目标物的精准识别具有重大意义,它标志着机器人技术与深度学习的结合,这种利用深度学习对目标进行定位的机器人定位技术在深度检测上具有较大的误差,这些机器人抓取定位技术存在着定位精度不足,或者无法满足对深度精确度的要求。或者只能完成对一种物体的定位,当目标物种类变多或者工业场景发生改变,这些方法的定位便难以完成。牛津大学计算机视觉组提出构建 VGGNet-16 等深度学习模型<sup>[10]</sup>应用于 SSD (Single Shot MultiBox Detector)深度学习模型中,能够达到目标检测的精度和速度。将这个模型基于机器人抓取实际场景,与深度模型结合,完成对深度和像素位置的双重精准定位,通过深度模型对多种不同物体进行精准识

收稿日期:2020-09-11;修回日期:2020-11-28.

\* 基金项目:上海市科学技术委员会科研计划项目(18030501200);上海工程技术大学新工科建设项目(X202018001).

作者简介:李世裴(1995—),安徽宣城人,男,硕士研究生,从事机器人视觉研究.

别,完成对目标物深度的精准测量,二者在视觉系统中完成协作,相互通讯,通过完成对目标物的精准定位。这种方法的精确度高,误差在毫米级别,并且可以适应多种不同环境,具有很强的鲁棒性。

依托搭建的协作机器人抓取视觉检测系统,提出通过双目相机采集目标物的图像,经过数据扩增,创建数据集,从而研究目标检测和机器人抓取技术。通过将采集的图像输送至单步多目标检测模型<sup>[11]</sup>中进行训练,完成对目标物的平面定位实验。通过双目测距原理完成对目标物的深度定位实验。由相机标定<sup>[12]</sup>和手眼标定<sup>[13]</sup>获取目标物的机器人基坐标位置信息,在机器人操作平台 ROS<sup>[14]</sup>中建立机器人与上位机通讯之后,将获取到的机器人基坐标信息发送给 ROS。

### 1 机器人抓取视觉检测系统的搭建

如图 1 所示,双目相机固定安装在工作空间正上方,工作空间中放置各种类的目标物,三手指灵巧手安装在协作机器人末端。双目相机经过相机标定后,采集图像输送至单步多目标检测器,完成训练识别出目标物类别和像素坐标,同时双目相机根据双目测距原理建立模型,输出得到目标物的深度。像素坐标和深度坐标信息通过手眼标定转换到机器人基坐标。在 ROS 中建立机器人与上位机的通讯,将机器人基坐标信息发送给 ROS。

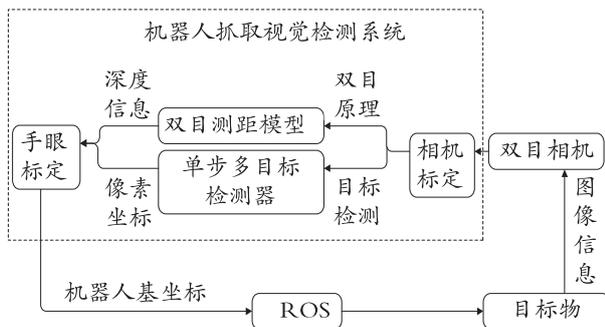
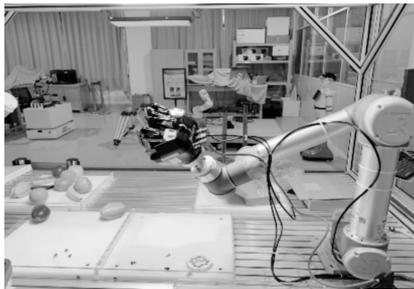


图 1 机器人抓取视觉检测系统实验平台

Fig. 1 Robot grasping vision detection system experimental platform

### 2 视觉标定原理及方法

机器人视觉系统由双目相机、相机标定系统和手眼标定系统构成,主要完成像素坐标和目标点的深度计算,并转换为机器人基坐标,经过视觉系统转换的坐标点会提供给 ROS。

#### 2.1 双目相机标定原理

(1) 双目相机成像模型。单目相机标定的目的是获取相机的内参和外参,内参主要是内部结构参数 $(\frac{1}{dx}, \frac{1}{dy}, Cx, Cy, f)$ ,外参主要是相机的旋转矩阵  $R$  和平移向量  $t$ 。dx 和 dy 分别代表着相机单个感光单元芯片的长度和宽度。相机感光芯片中心点在  $x$  和  $y$  上的偏移量即  $Cx$  和  $Cy$ ,  $f$  代表相机的焦距。双目相机标定首先要获取左右相机的内外参数,接着通过立体标定对左右图像进行立体校准,最后确定两个相机的相对位置关系<sup>[15]</sup>,即中心距。如图 2 所示为理想双目相机成像模型,假设两个相机光轴平行,且相机参数一致,则深度值的推导过程下:

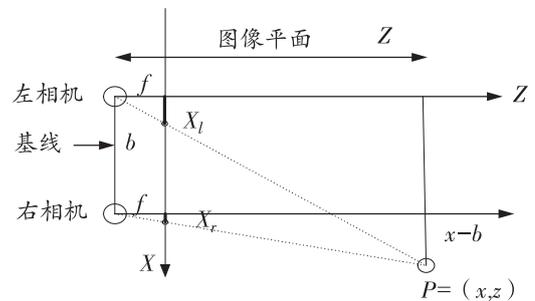


图 2 理想相机成像模型

Fig. 2 Ideal camera imaging model

$P$  为空间中一点,  $f$  为相机焦距,  $b$  为两个相机之间的间距,  $z$  为  $P$  点到双目相机的深度。根据三角形相似定律:

$$\frac{z}{f} = \frac{y}{y_l} = \frac{y}{y_r} = \frac{x}{x_l} = \frac{x-b}{x_r}$$

由上解得:

$$x = \frac{x_l * b}{x_l - x_r}, z = \frac{f * b}{x_l - x_r}, y = \frac{y * b}{x_l - x_r}$$

由视差  $d = x_l - x_r$  可得

$$z = \frac{b * f}{d}, x = \frac{x_l * z}{d}, y = \frac{z * y}{f} \tag{1}$$

由式(1)推导可知,左右相机之间的间距  $b$  和焦距  $f$  是相机本身的参数,数值也是固定的,相机中

的唯一变量是  $d$ ,故双目相机测距的核心问题是计算左右相机的视差,即左相机的每个像素点  $(x_l, y_l)$  和右相机中对应点  $(x_r, y_r)$  的对应关系。

(2) 极线约束和图像矫正技术。相机坐标系与目标物构成的三维坐标系中,目标物  $P$  点与左右相机的中心点  $(c_1, c_2)$  形成三维空间中的一个平面  $Pc_1c_2$ ,这个平面称为极平面。极平面与左右图像相交的两条直线称之为极线,对于在左图像中的  $P$  点,寻找到在右图像中的对应点,便可求解出  $P$  点的深度,极线约束就是当同一个空间点在两幅图像上分别成像时,已知左图投影点  $p_1$ ,那么对应右图投影点  $p_2$  一定在相对于  $p_1$  的极线上,这样可以极大的缩小匹配范围。如图 3 所示,可以直观地看到,沿着极线搜索,最终可以找到与左图像对应的点  $p_2$ 。

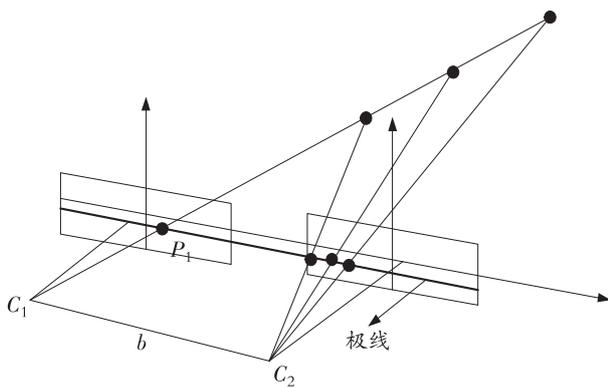


图 3 极线约束

Fig. 3 Polar constraints

但是在有些场景下两个相机独立固定,和  $c_2$  不会完全水平,即使是固定在同一个基板上也会因为装配的原因导致  $c_1$  和  $c_2$  不完全水平,图像校正技术就是通过分别对两张图片求单应矩阵得到的,这种技术把两个不同方向上的图像平面重新投影到一个平面且光轴相互平行,将非理想情况下的模型变成理想模型。

2.2 相机标定

相机标定是完成双目测距的必要条件,使用张正友标定法,打印一张棋盘格标定板,双目相机从至少 3 个方向采集标定板的照片,就可以求出左右相机的内外参数和双目相机的径向畸变  $K_1, K_2, K_3$  和切向畸变  $P_1, P_2$ 。基于 opencv 将采集的图像裁剪为左右两张图像,调用 matlab 中的 stereo Camera Calibrator 函数,输入棋盘格单位格大小以及裁剪完

的图像,便可求解出相机内外参数及畸变参数。如表 1 所示,为双目相机标定的过程与结果。 $Rot$  代表相机 2 相对于相机 1 的旋转矩阵,  $Trans$  代表相机 2 相对于相机 1 的偏移矩阵。

表 1 双目相机标定结果

Table 1 Binocular camera calibration results

	1.000 0	-4.915 1	0.005 9
$Rot$	-9.273 0	1.000 0	0.002 4
	-0.005 9	-0.002 4	1.000 0
$Trans$	-64.446 3	-0.108 7	0.454 5

2.3 手眼标定

手眼标定是为了建立相机坐标系与机器人基坐标系之间的关系,将视觉系统识别到的结果转移到机器人坐标系下。根据相机固定的位置不同,可以将手眼标定分为眼在手 (eye-in-hand) 和眼在外 (eye-to-hand) 两种形式,选用眼在外的形式进行手眼标定,相机固定在机器人的上方,在这种关系下,机器人末端和标定板的位姿关系始终保持不变。手眼标定的主要方法是控制机器人末端带动标定板在不同位置移动,如图 4 所示,  ${}^wH$  为求解目标,表示相机坐标系相对于机器人基坐标系的变换矩阵。 ${}^eH_i, {}^wH_j$  分别代表基坐标系相对于相机机器人末端 TCP 的第  $i, j$  次变换矩阵,  ${}^cH_i, {}^gH_j$  分别代表棋盘格相对于摄像机坐标系的第  $i, j$  次变换矩阵。 ${}^eH$  代表棋盘格相对于机器人末端 TCP 的变换矩阵。

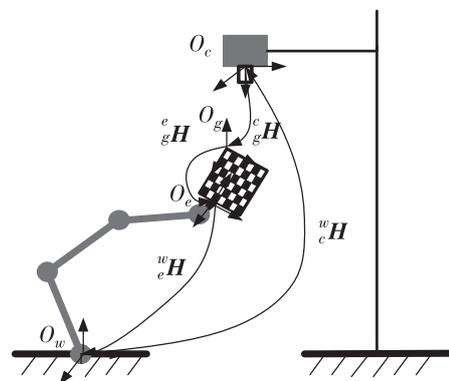


图 4 手眼标定模型

Fig. 4 Hand-eye calibration model

通过求解不同位置时相机外参和机器人工具坐标系相对于机器人基坐标系的关系来求出相机坐标系相对于机器人基坐标系的关系。基于选取的手眼

标定类型为眼在手,手眼变换矩阵为:

$${}^w\mathbf{H} * {}^c\mathbf{H} = {}^w\mathbf{H}_i * {}^c\mathbf{H} \quad (2)$$

$${}^w\mathbf{H} * {}^c\mathbf{H} = {}^w\mathbf{H}_j * {}^c\mathbf{H} \quad (3)$$

由于 ${}^w\mathbf{H}_i$ 和 ${}^w\mathbf{H}_j$ 都是可逆的,于是可将式(2)和式(3)分别左乘 $({}^w\mathbf{H}_i)^{-1}$  $({}^w\mathbf{H}_j)^{-1}$ ,消除掉常量 ${}^c\mathbf{H}$ 之后得:

$$({}^w\mathbf{H}_j)^{-1} * {}^w\mathbf{H} * {}^c\mathbf{H}_j = ({}^w\mathbf{H}_i)^{-1} * {}^w\mathbf{H} * {}^c\mathbf{H}_i \quad (4)$$

再由式(4)两边同时左乘 ${}^w\mathbf{H}_j$ 同时右乘 $({}^w\mathbf{H}_j)^{-1}$ ,得到:

$${}^w\mathbf{H}_j * ({}^w\mathbf{H})^{-1} * {}^w\mathbf{H} * {}^c\mathbf{H}_j = {}^w\mathbf{H} * {}^c\mathbf{H}_i$$

令 $A = {}^w\mathbf{H}_j * ({}^w\mathbf{H}_j)^{-1} * {}^w\mathbf{H} * {}^c\mathbf{H}_j$ ,  $B = {}^c\mathbf{H}_j * ({}^c\mathbf{H}_i)^{-1}$ ,  $X = {}^w\mathbf{H}$ 可得:

$$AX = XB \quad (5)$$

手眼标定的主要问题就是求解如式(5),opencv提供5种方法求解 $AX = XB$ ,选择Tsai两步法速度最快,通过C++定义一个手眼标定函数`calibrateHandeye()`,输入机器人末端相对于机器人基坐标系的旋转矩阵与平移向量,标定板相对于双目相机的齐次矩阵,计算出手眼矩阵。在深度坐标上,基于双目相机测距原理,建立一个双目测距模型,通过输入像素坐标便可得对应的深度。

### 3 单步多目标检测原理与方法

#### 3.1 单步多目标检测原理

SSD(Single Shot MultiBox Detector)是一种单步多目标检测器,单步多目标检测器采用不同的尺度在不同的位置进行密集采样,用CNN得到的特征提取直接进行分类和回归。

在被SSD处理之前,图像数据会被转化为 $300 * 300$ 的尺寸,SSD首先会经过VGG-16卷积神经网络的前5个卷积组,每个卷积组都由卷积层和池化层构成,SSD使用Astrous算法将第6层和第7层全连接层转换为卷积层,接着再通过3个尺寸不同的卷积组和一个平均池化层检测不同尺寸的物体。再每个卷积神经网络输出对应的特征图后,卷积滤波器会产生不同的预测,得到大量的回归框,SSD使用NMS(非极大值抑制)的方法来去除大量的重复框,当阈值设定为0.01后,候选框便被消除掉,节约了大量计算,他将所有的计算封装在单个网

络中,使得速度明显加快。

#### 3.2 单步多目标检测方法

在对机器人视觉系统搭建之后,使用相机采集了9种目标物体在不同的位姿情况下的图像。由于采集的图像数量达不到理想的数量,通常会采用一些方法去扩充这些图像,进而扩大样本集,防止过拟合的出现。接着需要对扩充后的样本集进行标注,此时样本集中的有效信息将会被进一步提取,经过训练最终生成SSD目标检测模型,完成目标检测的任务。

##### 3.2.1 样本集的制作

深度学习中要在大型数据集上进行训练,优化相应的参数最终得到一个合适的模型。如图5所示,在实际实验中选取了9种工业中常见机械部件(轴承、螺钉等)。

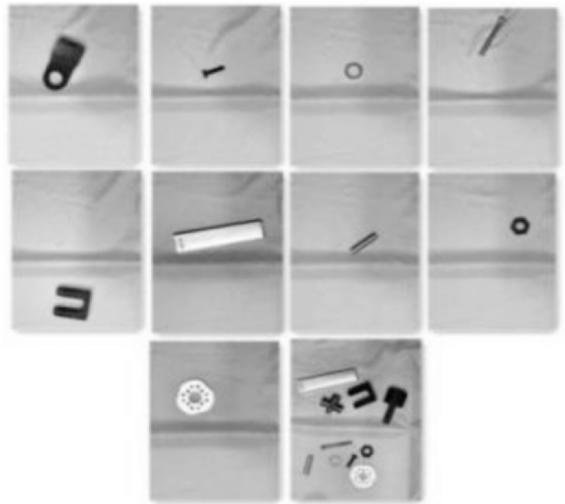


图5 实验中所用的9种机械部件

Fig. 5 Nine kinds of mechanical parts used in the experiment

##### 3.2.2 样本集的标注

在训练SSD目标检测模型中需要对目标物进行标注,让计算机知道目标物的类别和目标物的像素坐标位置。使用LabelImg标注工具来对950张图像进行标注,标注的最终结果保存在“.xml”文件下,“.xml”下包含样本集的所有信息。

完成标注工作之后,要对数据集进行划分,将数据集划分为训练集、测试集和验证集。使用python编写一个划分的脚本,脚本按照设置的比例自动拆分训练集、测试集和验证集最终生成以.txt结尾的文件。实验选取基于Tensorflow深度学习框架的

Object Detection API 训练目标检测模型,其中的 create\_pascal\_tf\_record.py 程序将数据集最终转换为 TFRecord 格式。

### 3.2.3 SSD 模型的训练

SSD 模型训练首先要进行实验模型的预训练,选择 SSD-300 VGG-based 作为预训练模型,编写模型训练配置文件,修改相应的训练则值为 NCHW,在 NVIDIA 1080TI 的 GPU 环境下运行 conda 进行模型的训练。设置初始学习率 0.000 6,训练次数为 45 000 次。使用 TensorFlow 中的 TensorBoard 可以完成对训练过程可视化,查看模型训练情况,模型的损失值随着训练次数的增加逐渐减小。最后利用 Object Detection API 中的 python 脚本将模型结构以及参数相关的文件转化为 PB 可执行文件。

### 3.3 目标检测

在 ROS 平台中完成 SSD 目标检测模型的搭建,则需要先创建 detect\_ros 节点,这个节点集成了 SSD 目标检测的 PB 文件、Object Detection API 以及数据集的 TFRecord 文件,对相机中的图像数据进行检测,将置信度、回归框、坐标中心等信息发布对话,完成订阅便能得到置信度、回归框、坐标中心。

## 4 实验及结论

### 4.1 目标检测实验

在目标检测实验中,取出 9 个目标物中的 4 种不同物体进行实验,用训练好的 SSD 模型进行检测,SSD 模型会输出每个物体的回归框、编号、置信度。每个编号的回归框也用不同颜色表示,如图 6 所示。

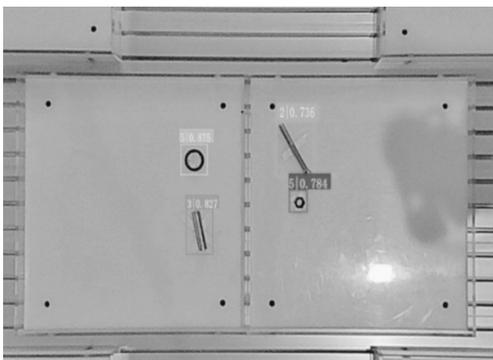


图 6 目标检测结果

Fig. 6 Target detection results

### 4.2 物体定位实验

SSD 在输出矩形框的同时,会输出矩形框的四点像素坐标,如图 7 所示,根据坐标位置可以确定目标物的中心点,物体的中心点经过双目测距模型可得目标物的深度,实验使用 HBV-1780 双目相机,相机参数如下:焦距 2.1 mm,FOV 为  $100^\circ$ ,相机间距 60 mm,分辨率  $2\ 560 \times 720$ 。实验选择了 4 个位置,其像素坐标分别为 (540, 270)、(550, 300)、(690, 260)、(700, 330),将 9 种不同的物体分别放置在这 4 个位置,通过目标检测系统得到的像素坐标平均值 ( $X_m, Y_m$ ) 与真实像素坐标 ( $X_r, Y_r$ ) 的偏移量来计算  $x$  方向和  $y$  方向上的绝对误差,结果如表 2 所示。通过双目测距模型得到的深度  $Z_m$  和真实的深度平均值  $Z_r$  来计算深度方向上的绝对误差,结果如表 3 所示。

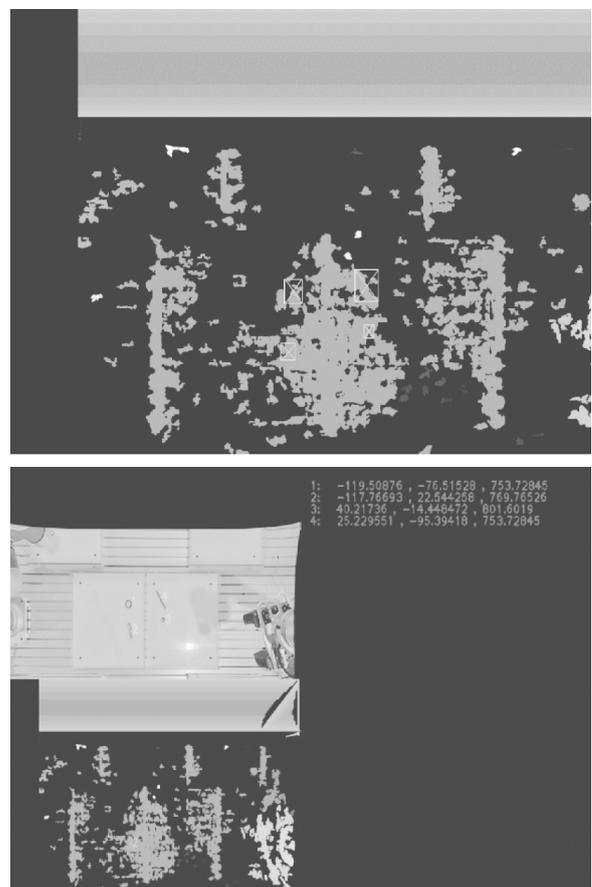


图 7 双目测距结果

Fig. 7 Binocular ranging results

表2 像素坐标平均值与方差

Table 2 Mean and variance of pixel coordinates

方向	坐标/mm	平均值/mm	方差/mm
x	540	538.4	2.7
	550	549.1	3.3
	690	690.7	6.9
	700	699.4	1.4
y	270	267.6	3.8
	300	298.5	2.0
	260	261.8	5.1
	330	330.2	2.8

表3 深度测量平均值

Table 3 Average depth measurement

编号	$Z_r$ /mm	$Z_m$ /mm
1	705.3	707.2
2	710.0	708.8
3	769.7	770.9
4	782.4	779.6
5	767.5	768.1
6	733.7	734.5
7	720.2	721.0
8	742.8	743.1
9	754.5	753.3

试验最终结果表明:整个抓取检测系统能够快速准确地测量出多个物体在平台的像素坐标和深度值,且具有同时检测多个目标物的能力,像素坐标位置误差在 3.5 mm 以内,深度误差在 1.2 mm 之内,整体定位误差较小,可以满足抓取操作的要求。

## 5 结论

(1) 针对机器人抓取技术难以适应真实工业环境的问题,面向机器人抓取给出了一种双目视觉单步多目标检测方法。

(2) 采用单步多目标检测器替代传统视觉检测器检测多种物体,获得类别、回归框、置信度,对获得的回归框进一步处理获取目标物的中心点。

(3) 通过目标检测实验表明,测试集表现良好,方法的目标检测误差在 3.5 mm 以内,深度误差在 1.2 mm 之内,具有很好的应用推广价值。下一步单

步多目标检测算法可进一步优化,进一步对像素坐标的偏差进行修正。

## 参考文献(References):

- [1] 孟娜. 双目视觉机器人的目标识别与抓取技术研究[D]. 沈阳:沈阳工业大学,2019  
MENG N. Research on Target Recognition and Grasping Technology of Binocular Vision Robot [D]. Shenyang: Shenyang University of Technology,2019(in Chinese)
- [2] 宋金鸿. 基于 DSP 的对空目标检测的研究与实现[D]. 西安:西安工业大学,2013  
SONG J H. Research and Implementation of Space Target Detection Based on DSP[D]. Xi'an: Xi'an University of Technology,2013(in Chinese)
- [3] 林刚,王波,彭辉,等. 基于改进 Faster-RCNN 的输电线巡检图像多目标检测及定位[J]. 电力自动化设备,2019,39(5):213—218  
LIN G, WANG B, PENG H, et al. Multi-target Detection and Location of Transmission Line Patrol Image Based on Improved CALY-RCNN[J]. Electric Automation Equipment, 2019,39(5):213—218(in Chinese)
- [4] 尹宏鹏,陈波,柴毅,等. 基于视觉的目标检测与跟踪综述[J]. 自动化学报,2016,42(10):180—187  
YIN H P, CHEN B, CHAI Y, et al. Overview of Target Detection and Tracking Based on Vision[J]. Journal of Automation, 2016, 42(10):180—187(in Chinese)
- [5] KRIZHEVSHY A, SUTSKEVER I, HINTON G E. Image Classification with Deep Convolution Neural Networks [A]. In: Advances in Neural Information Processing Systems,2012
- [6] DALAL N, TRIGGS B. Histograms of Oriented Gradients for Human Detection [C]// IEEE Computer Society Conference on Computer Vision & Pattern Recognition, 2005
- [7] SCH C, LAPTEV I, CAPUTO B. Recognizing Human Actions: A Local SVM Approach [C]// International Conference on Pattern Recognition, 2004
- [8] 夏晶. 基于 RGB-D 和深度学习的机器人抓取检测[D]. 南京:东南大学,2019  
XIA J. Robot Grasping Detection Based on RGB-D and Deep Learning [D]. Nanjing: Southeast University, 2019 (in Chinese)
- [9] ZHANG Z. A Flexible New Technique for Camera Calibration[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2000,22(11):1330—1334
- [10] 孙元辉,徐智勇,张建林,等. 基于深度学习的单步目

- 标检测器特征增强算法[J]. 半导体光电, 2019, 40(1):108—111
- SUN Y H, XU Z Y, ZHANG J L, et al. Feature Enhancement Algorithm of Single Step Target Detector Based on Deep Learning [J]. Semiconductor Optoelectronics, 2019, 40(1):108—111 (in Chinese)
- [11] 支健辉,董新民,孔星炜,等. 相机标定的外界影响因素分析[J]. 应用光学, 2014, 35(2): 286—291
- ZHI J H, DONG X M, KONG X W, et al. Analysis of External Influencing Factors of Camera Calibration [J]. Applied Optics, 2014, 35(2): 286—291 (in Chinese)
- [12] 唐新星,倪涛,何丽鹏,等. 基于立体视觉的遥操作工程机器人自主作业系统[J]. 农业机械学报, 2012, 43(10): 224—228
- TANG X X, NI T, HE L P, et al. An Autonomous Operating System of Teleoperation Engineering Robot Based on Stereo Vision [J]. Journal of Agricultural Machinery, 2012, 43(10): 224—228 (in Chinese)
- [13] 李铁军,王晓撰,杨冬. 基于 ROS 的机器人运动控制研究[J]. 组合机床与自动化加工技术, 2019(10): 73—77
- LI T J, WANG X Z, YANG D. Research on Robot Motion Control Based on ROS [J]. Modular Machine Tools and Automatic Processing Technology, 2019(10): 73—77 (in Chinese)
- [14] 柯科勇. 基于双目视觉的散乱堆放工件拾取系统[D]. 广州:广东工业大学, 2016
- KE K Y. The System for Picking up Scattered and Stacked Workpiece Based on Binocular Vision [D]. Guangzhou: Guangdong University of Technology, 2016 (in Chinese)

## Application of Binocular Vision Single Step Multi-target Detection Method for Robot Grasping

LI Shi-pei, LI Chun-lin, HAN Jia-bu, ZHU Xin-long

(School of Mechanical and Automobile Engineering, Shanghai University of Engineering Science, Shanghai 201600, China)

**Abstract:** Grasping technology of industrial robots is generally adopted in the way of off-line teaching, the change of the object will lead to the problem of inaccurate grasping. This paper studies a binocular vision monocular multi-object detection method for robot grasping, which enhances the robot's grasping ability. Firstly, the robot grasping vision detection system is built. This system adopts binocular ranging model to locate depth and single-step multi-target detector to locate pixel coordinates. Through the camera calibration and eye-to-hand hand-eye calibration image pixel coordinates and the depth of the binocular distance measurement model of the coordinate transformation for the robot base coordinate position, robot kinematics model is set up to complete the transformation from robot base coordinate to robot end coordinate, by ROS, superordination machine and the robot localization is realized, and the result is sent to ROS. After many experiments, it is shown that the target detection error of this method is within 3.5 mm and the depth error is within 1.2 mm, which has a good practical popularization value.

**Key words:** single-step multi-target detection; binocular vision; robot grasping; camera calibration

责任编辑:田 静

引用本文/Cite this paper:

李世裴,李春琳,韩家哺,等. 面向机器人抓取的双目视觉单步多目标检测方法应用研究[J]. 重庆工商大学学报(自然科学版), 2021, 38(5): 68—74

LI S P, LI C L, HAN J B, et al. Application of Binocular Vision Single Step Multi-target Detection Method to Robot Grasping [J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2021, 38(5): 68—74