

港口集装箱自主识别定位技术

丁小虎, 陈 宁

(集美大学机械与能源工程学院, 福建 厦门 361021)

[摘要] 针对港口集装箱自动装卸问题, 结合双目视觉和深度学习技术, 设计了一种基于 Faster R-CNN (faster regions with convolutional neural network) 模型的集装箱三维识别定位方法。首先利用双目摄像头采集集装箱图像, 用于训练 Faster R-CNN 模型, 然后使用该模型检测图像中的集装箱目标, 对识别出的目标添加矩形框, 并提取其中心点图像坐标, 接着通过对双目摄像头进行标定和匹配, 获取集装箱矩形框中图像坐标点的深度, 实现对集装箱的三维定位, 最后将集装箱的三维坐标转换到轮胎吊吊具坐标系下, 获得所有集装箱目标中心和吊具中心距离。实验结果表明, 系统运行速度可以达到 30 fps, 平均定位误差在 5 mm 以内, 系统可以有效解决集装箱三维实时识别和定位问题, 提升港口集装箱自动化装卸能力。

[关键词] 集装箱; 双目视觉; 深度学习; 三维定位

[中图分类号] TP 273

Autonomous Recognition and Location Technology of Port Container

DING Xiaohu, CHEN Ning

(School of Mechanical and Energy Engineering, Jimei University, Xiamen 361021, China)

Abstract: In view of the problems of automatic container operation, and combining binocular vision with deep learning, a container 3D recognition and location method based on Faster R-CNN model is presented. The paper firstly uses the image information of the containers acquired by binocular camera to train Faster R-CNN model, then detects the targets according to the model for getting the rectangular box of the container and its center coordinate of frame. In order to get the depth of the containers to realize the three-dimensional positioning, the binocular depth camera has been calibrated and matched. Finally, the coordinates are converted into the coordinate system of the tire hanging spreader to calculate the distance between all the containers center and spreader center. The results show that the system can reach 30fps, and the average position of error is within 5mm. The system can effectively solve the three-dimensional real-time recognition and location problem of container, and improve the automatic operation level for handling the containers at port.

Keywords: container; binocular vision; deep learning; three-dimensional positioning

0 引言

提升集装箱的自主搬运能力是港口设备智能化升级的主要研究方向, 目前有越来越多的先进技术在港口设备智能化改造中得到应用, 如视觉检测技术、防摇摆控制技术和自动导引技术等。文献

[收稿日期] 2018-03-17

[基金项目] 福建省科技厅资助项目 (JK214024); 国家海洋局 (省海洋渔厅) 资助项目 (2014FJPT03); 福建省自然科学基金资助项目 (2016J01755)

[作者简介] 丁小虎 (1994—), 男, 硕士生, 主要研究方向为机器视觉、嵌入式智能控制。通信作者: 陈宁 (1972—), 男, 副教授, 博士, 主要研究方向为机器视觉、工业机器人。

[1] 提出了一种基于机器视觉的集装箱装卸桥系统闭环防摇切换控制方法; 文献 [2] 研究了基于图像处理的起重机吊装防摇晃系统; 文献 [3] 研究了一种港口集装箱卡车的视觉自动导引停靠技术。但是传统的机器视觉检测和定位技术的鲁棒性较差, 其特征提取或者特征融合等目标提取算法难以适应复杂的港口实际环境。随着深度学习技术的不断发展, 该技术在物体目标检测识别等领域已逐渐代替传统视觉处理算法, 成为处理图像问题的主流算法^[4-7], 机器视觉结合深度学习技术也将成为提升轮胎吊智能化装卸水平的突破点。目前, 在物体目标检测方面, Faster R-CNN 模型利用和整个目标检测网络共享卷积特征的区域建议网络 RPN (region proposal network)^[8], 生成包含目标的候选区域, 检测精度相对较高。目标定位方面, 为了获取目标的三维位置信息, 考虑到双目视觉系统能够适用于室外光照变化环境, 并且具有高精度和非接触测量的特点, 选择双目视觉系统测量与目标的距离。本文提出使用 MYNT EYE D1000-IR 双目摄像头作为集装箱图像采集器, 研究轮胎吊下集装箱目标的识别和定位技术, 首先利用 Faster R-CNN 模型实时判别及检测集装箱, 然后通过张正友标定法^[9]对双目摄像头进行标定, 采用半全局立体匹配算法求解视差, 并计算集装箱目标的三维信息, 最后经过坐标系转换计算获得吊具中心和所有集装箱目标中心的距离, 从而实现轮胎吊系统对集装箱目标装卸定位。

1 轮胎吊模拟装卸系统

轮胎吊模拟装卸系统台架结构图如图 1 所示, 该台架可以实现手动控制小车水平移动、吊具垂直升降、吊锁开闭和集卡等基本功能。

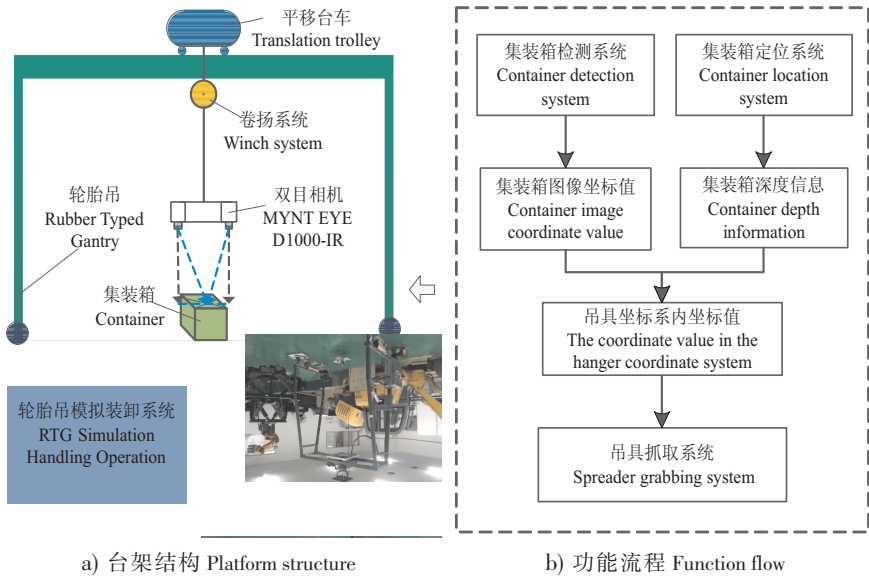


图 1 系统结构组成
Fig.1 System composition

吊具和双目视觉系统的安装位置如图 2 所示, 吊具传感器坐标原点为 $O_d - X_d Y_d Z_d$ 。双目视觉传感器型号为 MYNT EYE D1000-IR, 主要由一对基线为 120 mm 的平行的 RGB 镜头、一个惯性测量单元 IMU 和两颗 IR 红外主动光探测器组成。该双目摄像头能够以 60 帧/s 的速率同步传送最高 2560 × 720 分辨率的 RGB 彩色图像, 惯性测量单元可以测量和调节运动中吊具的姿态, 红外主动光探测器可以通过发射的红外结构光来增加视野纹理, 适用于室外以及完全黑暗环境。

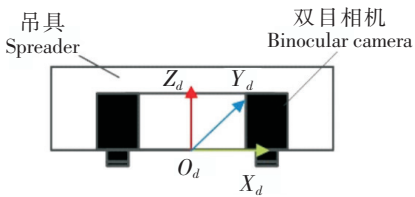


图 2 安装位置图
Fig.2 Installation position

2 集装箱检测系统

集装箱检测系统通过 D1000-IR 双目相机采集视野内的彩色 RGB 图像，存储含有集装箱目标在不同环境和姿态下的图像，对其中的集装箱目标进行标记，并用于训练 Faster R-CNN 模型，训练后的 Faster R-CNN 模型可以检测集装箱目标以及返回其位置的图像坐标点。

2.1 集装箱图像采集和扩增

选取 8 种不同船运公司的 20 尺（5.89 m×2.32 m×2.31 m）集装箱货柜的缩比模型作为实验图像采集目标，比例为 1:20，如图 3 所示。集装箱搬运的工作区域主要都是灰黑色的水泥地面，吊具的垂直运动最大高度设置为 2.5 m，使用 D1000-IR 进行图像采集，按照不同运动高度，并考虑昼夜、雨天和雾天等天气，同时采集的图像序列涵盖边缘部分残缺和多种位姿状态等情况，共采集集装箱图像 1 000 张。

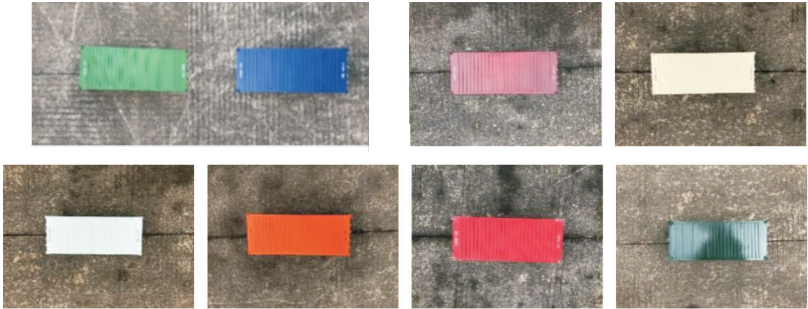


图 3 8 种集装箱缩比模型
Fig.3 Eight container scaling models

为了进一步增加集装箱数据集的多样性，对已采集的集装箱数据样本图像进行扩增，分别对原图像作 $\pm 45^\circ$ 旋转变换，同时对获得的旋转图像进行椒盐噪声和锐化处理，最后对所有 $\pm 45^\circ$ 旋转变换后的图像进行居中截取，如图 4 所示。

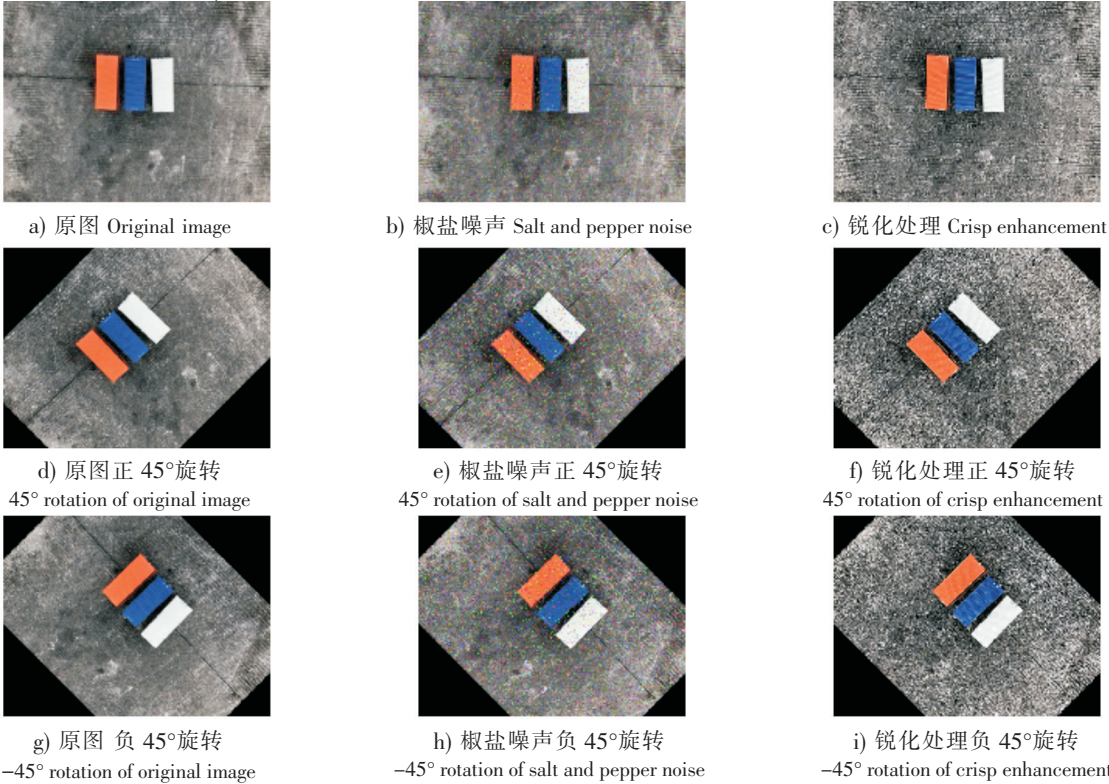


图 4 集装箱图像扩增效果图

Fig.4 Enhancement effect of container images

扩增后的集装箱数据集总共包含了 9 000 张含有集装箱目标的图像, 随机选取 6 750 张图像 (含有 20 250 个集装箱目标) 为训练集, 其余剩下 2 250 张图像 (含有 5 625 个集装箱目标) 作为测试集。

2.2 Faster R-CNN 模型参数初始化及训练

Faster R-CNN 是一种检测精度较高的深度学习模型, 模型结构如图 5 所示。

Faster R-CNN 模型通过卷积层提取目标特征映像图谱 (feature map), 输入给区域建议网络 RPN (region proposal network), 产生候选区域, 再连接到感兴趣区域池化层 (region pooling layer), 然后将池化后的特征输入给全连接层 (fully connected layer), 最后通过分类器进行分类及回归, 获得目标的类别和位置^[8]。

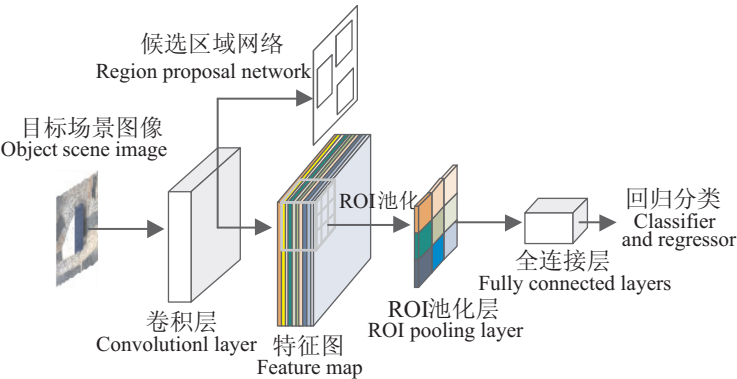


图 5 Faster R-CNN 模型结构
Fig.5 Faster R-CNN model structure

采用集装箱数据集微调 Faster R-CNN 预训练模型的方式训练整个模型, 其中预训练模型在 Oxford-IIIT Pet 数据集下训练获得, 同时将模型训练参数中的学习率设为 0.0002, 批量大小设定为 24, 迭代次数设定为 65 000。在训练过程中, 对于 RPN 网络中每个滑动窗口, 分别取 4 种面积尺度 {492, 642, 1 282, 2 562} 和 3 种长宽比 {1:1, 1:2, 2:1}, 共组合成 12 个不同尺度和长宽比的锚框, 在深度学习框架 Tensorflow 下, 实现 Faster R-CNN 检测模型算法。

服务器平台配置: Intel (R) Core (TM) i7-7700K CPU@4.2GHz 处理器; 16GB 内存; 1T 硬盘; 8GB Geforce GTX 1080 GPU。

3 集装箱定位系统

集装箱检测系统可生成集装箱目标的识别矩形包络框, 取其矩形框的中心点作为装卸搬运定位点, 利用集装箱定位系统可以获得该点的深度值, 得到定位目标点的三维坐标后, 通过坐标变换将其世界坐标位置转换到吊具坐标系下, 以实现集装箱目标的抓取。

3.1 摄像头投影关系

为了获得集装箱上一点的三维坐标与其对应像素坐标的关系, 需要知道视觉系统模型的投影关系和相关参数, 根据针孔相机模型可建立成像几何关系, 如图 6 所示。图 6 表示视觉系统的世界坐标系 {W}、摄像机坐标系 {C}、图像坐标系 {1} 和像素坐标系 {2} 之间投影关系。

假设集装箱目标上某点世界坐标值为 (X_w, Y_w, Z_w) , 在像素坐标系中的坐标是 (u, v) , 可得到成像模型齐次坐标

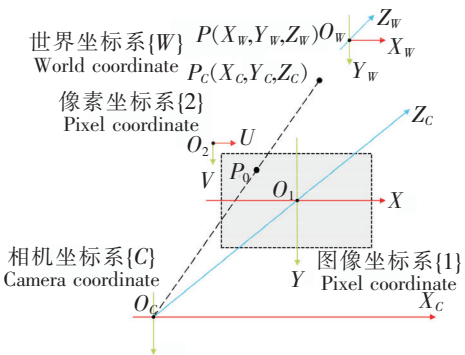


图 6 摄像机成像投影关系
Fig.6 Imaging projection relation in camera

投影变换关系:

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{M} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = \mathbf{M}_1 \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \mathbf{M}_1 \mathbf{M}_2 \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x & 0 \\ 0 & f_y & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}. \quad (1)$$

式中: f_x, f_y 为尺度因子; c_x, c_y 为图像中心与原点偏差像素数; \mathbf{R} 为旋转矩阵; \mathbf{T} 为平移矩阵; \mathbf{M} 为摄像头的投影矩阵。其中: \mathbf{M}_1 为摄像头的内部参数矩阵; \mathbf{M}_2 为摄像头的外部参数矩阵。

3.2 双目相机标定及匹配

3.2.1 双目相机标定

为了获取相机模型的对应参数, 采用 Stereo Camera Calibrator Toolbox 工具对 D1000-IR 双目相机进行标定。在标定板上选取世界坐标系 ($Z_w = 0$), 标定板采用的是 8×8 黑白相间的正方形棋盘格, 每个正方形边长为 28 mm。确保足够的光照后, 调整标定板的姿态, 并采集 20 组左右标定板图片, 其标定重构结果如图 7 所示。

通过对 D1000-IR 双目摄像头进行标定, 可以获取左右摄像头的内部参数 \mathbf{M}_1^L 和 \mathbf{M}_1^R 及相机的畸变系数 K_1^L 和 K_1^R , 如下所示:

$$\mathbf{M}_1^L = \begin{bmatrix} 731.63243 & 0 & 676.75722 \\ 0 & 735.86286 & 392.60994 \\ 0 & 0 & 1 \end{bmatrix}; \quad \mathbf{M}_1^R = \begin{bmatrix} 723.63620 & 0 & 621.96609 \\ 0 & 724.63474 & 382.67816 \\ 0 & 0 & 1 \end{bmatrix};$$

$$\mathbf{K}_1^L = [0.32557 \quad 0.11672 \quad 0 \quad 0 \quad 0]; \quad \mathbf{K}_1^R = [-0.32557 \quad 0.11672 \quad 0 \quad 0 \quad 0]。$$

在确定双目摄像头的内部参数后, 即可进一步求出外部参数矩阵 \mathbf{M}_2 的旋转矩阵 \mathbf{R} 和平移矩阵 \mathbf{T} , 如下所示:

$$\mathbf{R} = \begin{bmatrix} 0.99956 & -0.00446 & 0.02913 \\ 0.00411 & 0.99991 & 0.01219 \\ -0.02918 & -0.01206 & 0.99950 \end{bmatrix};$$

$$\mathbf{T} = [-120.84019 \quad 0.76369 \quad -3.11056]。$$

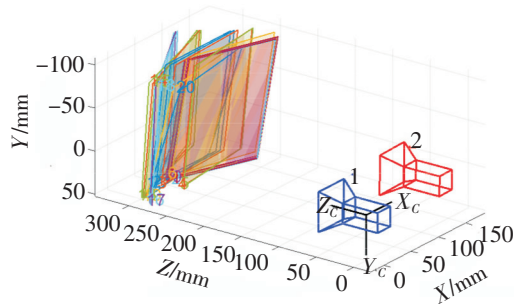


图 7 双目标定外参重构

Fig.7 Reconstruction of external parameters of binocular calibration

由标定结果可知, 两摄像头之间存在径向畸变, 双目视觉系统能够满足光轴平行的理想状态。

3.2.2 双目立体匹配

半全局立体匹配 SGM (semi-global matching) 算法是一种稠密立体匹配方法^[10], 本文采用 SGM 算法对双目图像进行立体匹配, SGM 算法具有较好的鲁棒性, 使用双目相机的 GPU 处理器加速可以取得比较高的计算效率, 主要算法步骤如下:

1) 利用 Census 变换^[11]计算匹配代价值, 即将集装箱图像局部窗口区域的中心像素灰度和邻域内像素灰度映射为一个比特串, 然后将映射得到的比特串来代替中心像素的灰度值。其变换公式如下:

$$C_s(u, v) = \bigotimes_{i=-l}^l \bigotimes_{j=-r}^r \xi(I(u, v), I(u+i, v+j)); \quad (2)$$

$$\xi(I_p, I_q) = \begin{cases} 0 & (I_p \leq I_q), \\ 1 & (I_p > I_q). \end{cases} \quad (3)$$

其中: $I(u, v)$ — 集装箱图像中心 (x, y) 的灰度值; $I(u+i, v+j)$ — 图像坐标 $(u+i, v+j)$ 的灰度值; $\xi(I_p, I_q)$ — 比特位, 像素点 p, q 的灰度值 I_p 和 I_q 比较运算; $C_s(u, v)$ — 集装箱中心像素的点 (u, v) 的

Census 的变换码; l — $1/2$ 局部图像窗口宽度; r — $1/2$ 局部窗口高度; \otimes —一位串联符。

集装箱场景图像经过 Census 变换后, 根据汉明距离相似性测度函数^[12]计算匹配代价, 其公式如下:

$$C(u,v,d) = \text{Hamming}(C_{sl}(u,v), C_{sr}(u-d,v))。$$

(4)

式中: $C_{sl}(u,v)$, $C_{sr}(u-d,v)$ 分别为集装箱左图像和右图像的 Census 的变换码; $C(u,v,d)$ 为匹配代价值; d 为左、右图像横坐标像素点的差值。

2) 由于代价计算只考虑局部相关性, 无法直接获得视差, 通过匹配代价聚合, 可以使计算得到的匹配代价值更准确地反应像素相关性。代价聚合采用大小固定矩阵窗口策略, 计算窗口内的聚合代价,

$$C_{\text{sum}}(u,v,d) = \sum_{i=-m}^m \sum_{j=-n}^n C(i,j,d)。$$

(5)

式中: m —聚合窗口的宽度; n —聚合窗口的高度;

利用公式 (5) 求出聚合代价之后, 计算集装箱图像的每个像素点坐标对应的视差

$$D(u,v) = \underset{d_{\min} \leq d \leq d_{\max}}{\operatorname{argmin}} C_{\text{sum}}(u,v,d)。$$

(6)

3) 对视差图进一步优化处理, 采用唯一性检测方法剔除错误视差, 并通过中值滤波的方法来抑制图像噪点。

3.3 目标深度坐标计算

轮胎吊模拟装卸系统通过卷扬系统的钢丝绳连接吊具和平移小车, 吊具和平移小车的位置相对平行, MYNT EYE D1000-IR 左右摄像头的距离相对固定且与光轴平行, 原理如图 8 所示。图 8 中双目视觉系统以左摄像头为原点建立左摄像机坐标系 $O_1-u_l v_l$, 并令世界坐标系与其重合, 左摄像头的图像像素坐标系 $O_1-u_l v_l$, 右摄像头的图像像素坐标系 $O_2-u_r v_r$; f 为双目摄像头的焦距; b 为基线距离; H_0 为平移小车和吊具之间的高度; 集装箱上的某点 $P(X_Q, Y_Q, Z_Q)$ 在左右摄像头图像像素坐标系的对应点分别是 $P_l(u_l, v_l)$ 和 $P_r(u_r, v_r)$ 。根据相似三角形的几何关系易得:

$$X_Q = [u_l / (u_r - u_l)] \cdot b,$$
$$Y_Q = [v_l / (u_r - u_l)] \cdot b,$$
$$Z_Q = [f / (u_r - u_l)] \cdot b。$$

3.4 世界坐标系转换为吊具坐标系

集装箱检测和定位系统在确定集装箱目标和所在位置后, 将信息发送给吊具抓取系统。由于双目摄像头安装在吊具抓取系统上, 所以在对集装箱的装卸过程中摄像头会随着平移小车水平移动和卷扬升降而运动。又因为抓取时只考虑集装箱目标和吊具的相对位置, 所以在吊具上的双目系统完成对集装箱的三维测距后, 需要将结果由双目摄像机的世界坐标系转换到吊具坐标系。

实验过程中, 双目摄像头安装在吊具的正中间, 所以吊具坐标系的原点选在双目摄像头的中点处。对双目摄像头原始标定时, 标定板和吊具平面平行, 所以旋转矩阵近似为 0, 左右摄像头初始标定的外参的平移矩阵分别是 $T_1(t_{x_l}, t_{y_l}, t_{z_l})$, $T_2(t_{x_r}, t_{y_r}, t_{z_r})$, 所以左右摄像头相对于原始世界坐标系的坐标分别为 $(-t_{x_l}, -t_{y_l}, -t_{z_l})$ 和 $(-t_{x_r}, -t_{y_r}, -t_{z_r})$, 可得吊具坐标系原点相对于原始世界坐标系的坐标为 $(-(t_{x_l} + t_{x_r})/2, -(t_{y_l} + t_{y_r})/2, -(t_{z_l} + t_{z_r})/2)$ 。

集装箱在世界坐标系的某点 $P(X_w, Y_w, Z_w)$, 则可得该点在吊具坐标系下 $P'(X_d, Y_d, Z_d)$ 坐标的转换关系:

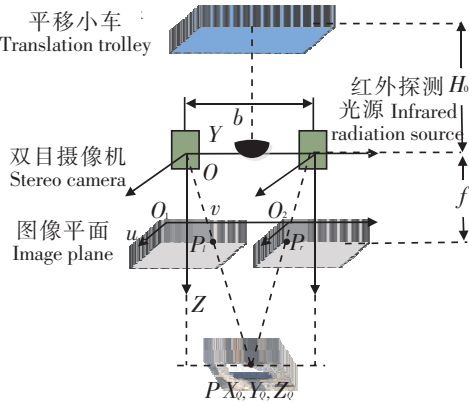


图 8 双目视觉获取深度坐标原理

Fig.8 Acquisition principle of depth coordinate of binocular vision

$$\begin{cases} X_d = X_w + (t_{xl} + t_{xr})/2, \\ Y_d = Y_w + (t_{yl} + t_{yr})/2, \\ Z_d = Z_w + (t_{zl} + t_{zr})/2. \end{cases} \quad (7)$$

4 结果与分析

4.1 集装箱目标识别实验

为检测 Faster R – CNN 算法对集装箱识别的准确性，本文进行了集装箱识别实验。图 9 是 Faster R – CNN 的网络训练损失函数值变化曲线图，训练迭代次数到 5 万次后，损失函数值基本保持稳定，提取训练的模型文件，对集装箱的图像帧进行检测识别。

将不同类别的集装箱箱体随机放在轮胎吊工作区域内，使用 Faster R – CNN 模型进行检测识别，Faster R – CNN 模型使用矩形框将所有集装箱目标包络起来，并在其矩形包络框上方显示检测到的集装箱的名称和准确度，同时根据包络框的位置可以确定集装箱中心点在整个图像中的位置，如图 10 所示。

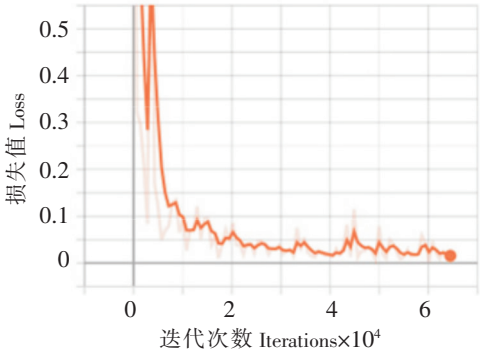


图 9 网络损失函数下降曲线图

Fig.9 Recession curve of loss function of network

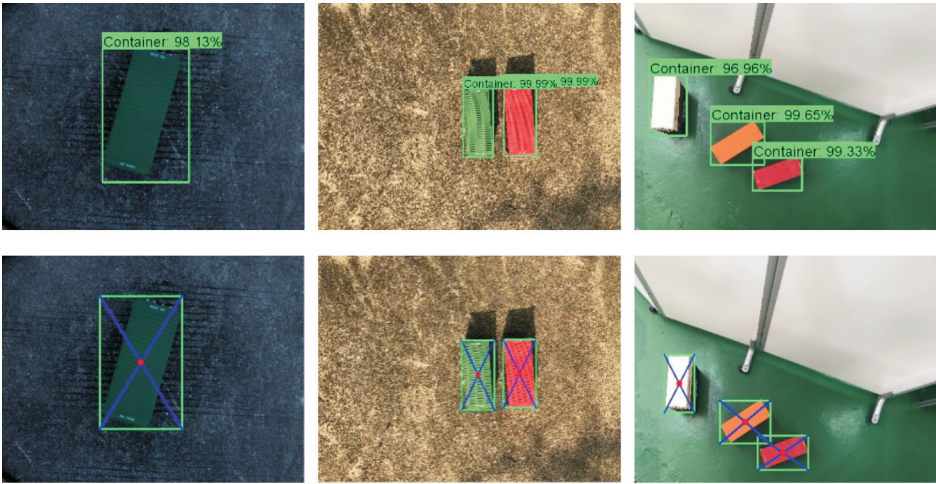


图 10 集装箱目标识别结果

Fig.10 Recognition results of container target

本文将正确识别集装箱目标个数和错误识别集装箱目标个数作为识别准确率的标准，集装箱错误识别包括漏识别和误识别两种情况。统计 2 250 张集装箱测试集图像，共有 5 625 个集装箱目标，集装箱检测识别统计结果如表 1 所示。5 625 个集装箱中有 5 468 个集装箱被正确识别，识别正确率为 97.21%，错误识别的集装箱个数为 157 个，识别错误率为 2.79%，其中漏识别 29 个，误识别 128 个。

表 1 集装箱测试集检测结果统计

Tab.1 Statistics of measure and recognition results of containers

类别 Classes	集装箱目标数 The number of containers	百分比 Percentage/%
正确识别 Correct	5 468	97.21
漏识别 Skipping	29	0.50
误识别 Missing	128	2.29

4.2 双目深度定位实验

本文采用 Faster R – CNN 和基于 SGM 的双目立体匹配算法组成的轮胎吊双目 Faster R – CNN 系

统, 进行集装箱的定位实验, 控制吊具在距离集装箱目标 400 ~ 2 000 mm 范围内进行测量实验, 其中双目深度定位测距率如图 11 所示。

根据矩形框位置确定集装箱目标中心点二维位置坐标点后, 同时由双目定位系统可获得其目标矩形框中心点深度坐标信息, 中心点的坐标代表集装箱目标的图像坐标, 在吊具坐标系下, 识别定位图像坐标原点转换到其图像正中央, 如图 12 所示, 图 12(a)为彩色图像, 图 12(b)为深度图像, 图中集装箱的图像坐标深度分别是 $Z_1 = 1246.37\text{ mm}$, $Z_2 = 1139.48\text{ mm}$, $Z_3 = 1246.90\text{ mm}$, 同时根据公式 (7) 计算得到吊具坐标系原点和集装箱目标识别矩形框的直线距离分别为 1250.79 mm, 1158.73 mm 和 1278.95 mm。

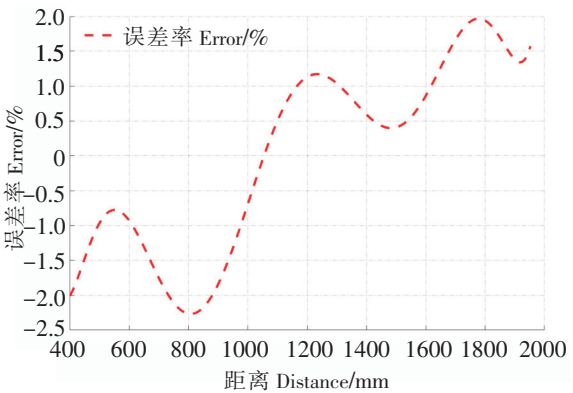
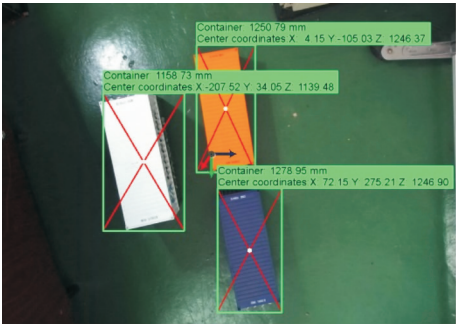
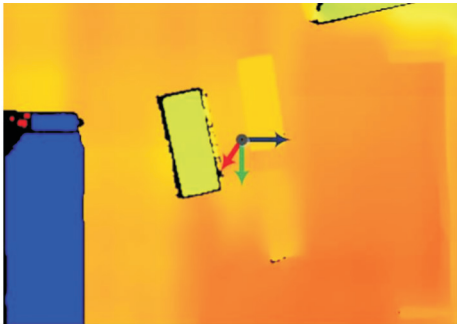


图 11 双目摄像头深度误差
Fig.11 Depth error of binocular camera



a) 集装箱三维坐标 Three-dimensional coordinates of container



b)双目深度图 Binocular depth map

图 12 定位图像
Fig.12 Localization image

4.3 系统性能分析

表 2 和表 3 分别是系统误差和实时性能分析, 集装箱检测系统识别单帧图像耗时0.027 s, 集装箱定位系统进行一组左右图像匹配耗时 0.0065 s, 整个识别和定位系统对双目左右图像进行一次图像匹配和各一次集装箱目标检测, 总共耗时 0.035 s (~ 30 帧/s)。分析表明, 该系统精度和实时性可以满足实际工作的要求。

表 2 集装箱坐标值平均误差	
Tab.2 Average error of container coordinate value	
定位坐标	误差 Error
Location coordinate	/ %
水平坐标 X	1.23
Horizontal coordinates	
垂直坐标 Y	0.94
Vertical coordinates	
深度坐标 Z	1.18
Depth coordinates	

表 3 系统处理速度	
Tab.3 Processing speed of system	
系统	耗时 Time
System	/s
整个系统	0.035
Entire system	
检测识别模块	0.027
Detection and recognition module	
双目定位模块	0.0065
Binocular positioning module	

5 结论

本文基于轮胎吊装卸模拟系统台架设计了一套集装箱三维自动识别和定位系统, 系统利用双目摄像机实时检测识别作业场景内的集装箱目标, 并准确计算集装箱的三维空间定位信息。在该系统中,

本文引入了一种基于 Faster R-CNN 的集装箱目标识别模型,并设计和制作满足数据多样性要求的集装箱图像数据集来训练该目标识别模型,通过调整训练参数实现了模型网络收敛,同时使用 GPU 处理器加速集装箱目标检测模型算法和 SGM 双目匹配算法,最终系统展现了对集装箱目标识别定位的鲁棒、实时和稳定检测的优良性能。本文将双目视觉和深度学习技术同时应用在港口装卸中,实现了对集装箱的自动三维识别与定位。

[参 考 文 献]

- [1] 付主木,高爱云,费树岷,等. 集装箱桥吊多模型参考切换双闭环防摇控制 [J]. 机械工程学报, 2011, 47(23): 161-166.
- [2] 李普林,柏伟. 基于图像识别技术的自动化岸边集装箱起重机设计 [J]. 起重运输机械, 2012(11): 33-35.
- [3] BYUN S, JUNG S, KIM M. Container chassis alignment and measurement based on vision for loading and unloading containers automatically [C] //International Conference on Hybrid Information Technology. USA: IEEE, 2006: 582-587. DOI:10.1109/ICHIT.2006.253665.
- [4] 唐聪,凌永顺,杨华,等. 基于深度学习物体检测的视觉跟踪方法 [J]. 红外与激光工程, 2018, 283(5): 148-158.
- [5] 王全东,常天庆,张雷,等. 基于深度学习算法的坦克装甲目标自动检测与跟踪系统 [J]. 系统工程与电子技术, 2018, 468(9): 252-265.
- [6] 宋焕生,张向清,郑宝峰,等. 基于深度学习方法的复杂场景下车辆目标检测 [J]. 计算机应用研究, 2018, 35(4): 1270-1273.
- [7] SHIN H C, ROTH H R, GAO M C, et al. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning [J]. IEEE Transactions on Medical Imaging, 2016, 35(5): 1285-1298. DOI:10.1109/TMI.2016.2528162.
- [8] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 37(6): 1137-1149.
- [9] ZHANGZ. A flexible new technique for camera calibration [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(11): 1330-1334.
- [10] HIRSCHM L LERH. Accurate and efficient stereo processing by semi-global matching and mutual information [J]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005, 2(2): 807-814.
- [11] 杨荣坚,王芳,秦浩. 基于双目图像的行人检测与定位系统研究 [J]. 计算机应用研究, 2018, 35(5): 1591-1600.
- [12] 代伟,赵杰,杨春雨,等. 基于双目视觉深度感知的带式输送机煤量检测方法 [J]. 煤炭学报, 2017, 42(增刊): 547-555.

(责任编辑 陈 敏 英文审校 郑青榕)