

DOI: 10.13382/j.jemi.B2205887

# 自动化集装箱码头装卸目标三维测量系统设计\*

宓超<sup>1</sup> 黄世凤<sup>2</sup> 张钰洁<sup>1</sup> 张志伟<sup>3</sup> 姚磊<sup>3</sup>

(1. 上海海事大学物流工程学院 上海 201306; 2. 上海海事大学物流科学与工程研究院 上海 201306;

3. 上海海瞩智能科技有限公司 上海 201306)

**摘要:**针对传统激光雷达在自动化集装箱码头装卸过程中对集装箱三维姿态定位精度低、成本高等问题,提出了一种基于视觉的集装箱姿态三维测量系统。首先通过小规模深度学习网络快速进行集装箱锁孔粗定位,其次通过传统图像处理算法对集装箱锁孔进行二次定位得到集装箱锁孔的精确位置,最后结合装卸过程中集装箱的物理运动对集装箱姿态进行三维测量。实验结果表明,与改进前的深度学习网络相比,测量精度更高、测量速度更快;整体算法的测量精度为93.71%,约12.45 fps,集装箱姿态测量平均测量误差约为4.95%,满足自动化装卸的要求。

**关键词:** 自动化码头; 集装箱装卸; 集装箱锁孔; 三维测量

**中图分类号:** TP29; TP18; U653.921 **文献标识码:** A **国家标准学科分类代码:** 41310

## Design of 3D measurement system for loading and unloading targets in automated container terminals

Mi Chao<sup>1</sup> Huang Shifeng<sup>2</sup> Zhang Yujie<sup>1</sup> Zhang Zhiwei<sup>3</sup> Yao Lei<sup>3</sup>

(1. Logistics Engineering College, Shanghai Maritime University, Shanghai 201306, China; 2. Institute of Logistics

Science &amp; Engineering, Shanghai Maritime University, Shanghai 201306, China;

3. Shanghai Haizhu Intelligent Technology Co., Ltd., Shanghai 201306, China)

**Abstract:** Aiming at the problems of low positioning accuracy and high cost of traditional LiDAR in the process of loading and unloading of automated container terminals, a vision-based three-dimensional measurement system for container attitude is proposed. Firstly, through a small-scale deep learning network for rapid coarse positioning container corner, secondly, the traditional image processing algorithm is used to reposition the container corner pieces to obtain the precise position of the container keyhole, and the three-dimensional measurement of the container posture is carried out in combination with the physical movement of the container during the loading and unloading process. The experimental results show that compared with the deep learning network before improvement, the measurement accuracy is higher and the measurement speed is faster, the measurement accuracy of the overall algorithm is 93.71%, about 12.45 frames/s, and the average measurement error of container attitude measurement is about 4.95%, which meets the requirements of automatic loading and unloading.

**Keywords:** automated terminal; container handling; container keyhole; 3D measurement

## 0 引言

传统的集装箱码头依靠人工操作,工作效率低下、工作出错率高且无法全面保障工人人身安全,已不再适用

当前社会经济快速发展的形势,而自动化集装箱码头使用机器代替人工操作,不仅能够实现与传统集装箱码头相同的集装箱转运和暂存的功能,还大大提高了工作效率<sup>[1]</sup>,降低人工成本同时保障人员安全。轨道式龙门起重机是自动化集装箱码头堆场作业的主要设备,负责集

收稿日期: 2022-10-06 Received Date: 2022-10-06

\* 基金项目: 中央军委装备预先研究领域基金(80912020104)、上海市自然科学基金(22ZR1427700)、中国(上海)临港自贸区管理委员会核心技术研发项目(SH-LG-GK-2020-21)资助

装箱堆场与集装箱卡车(下简称集卡)之间的集装箱装卸工作。在集装箱装卸作业过程中,最常见的事故是集卡随集装箱一同被吊具吊起<sup>[2]</sup>,从而造成严重的经济、安全损失,因此在集装箱装卸过程中,了解集装箱的三维姿态是非常重要的。在现有的自动化集装箱码头中,激光雷达仍为集装箱三维姿态测量最常用的手段。在集装箱装卸过程中,激光雷达通过扫描集装箱轮廓实时分析集装箱姿态变化。激光雷达具有不易受天气和光照条件影响的优点,但在实际应用中,激光雷达标定难度较大、成本太高,不适宜大多数码头。

随着机器视觉技术的快速发展,基于视觉的测量采用非接触式工作,具有分辨率高、采集数据快、低成本、高精度以及可全天候工作等优点<sup>[3]</sup>,被广泛应用于航天<sup>[4]</sup>、航海<sup>[5]</sup>、工业<sup>[6]</sup>、装备制造<sup>[7]</sup>等领域。基于视觉的三维测量依据相机数量可分为单目视觉测量、双目视觉测量和多目视觉测量。其中双目视觉测量和多目视觉相比单目视觉测量精度更高、测量范围更广但计算量较大、标定难度大且需要更加昂贵的硬件成本以及很难保证测量实时性,而单目视觉系统只使用单个视觉传感器,具有结构简单、计算速度快等优点,因此仍被广泛应用到实际项目中。

基于单目视觉的三维测量分为传统图像处理算法和深度学习算法,传统的图像处理算法能够在特定的数据集上取得不错的效果,如柳宁等<sup>[8]</sup>通过检测目标的旋转角特征对目标定位并引导机器人跟随目标运动,但传统图像处理算法应用环境简单、实时性差且使用浅层特征导致算法鲁棒性较差;而深度学习算法使用卷积神经网络通过端到端的方式能够提取到更深层次特征从而很好地实现目标的检测和分类,如于重重等<sup>[9]</sup>提出的一种基于卷积神经网络的度量学习模型实现小样本零部件表面缺陷检测;Liu等<sup>[10]</sup>将注意力机制与时序卷积神经网络相融合设计了一种皮带损伤检测模型。

基于深度学习的姿态测量方法大致分为3个方向,1)在RGB图像中直接检测目标的三维姿态,如YOLO-6D<sup>[11]</sup>、Pose-RCNN<sup>[12]</sup>、DOPE<sup>[13]</sup>等算法。上述算法的关键都是找到RGB图像中目标的中心,然后回归到相机坐标系下目标的质心实现对目标的姿态检测,虽然能够直接地对目标进行姿态检测但算法通常过于复杂,实时性比较差,因此必须依赖于高性能计算机才能达到实时性;2)提前定义多种目标姿态作为标注信息,将姿态估计问题转化为分类问题,该方法多用于卫星姿态检测,如任笑圆等<sup>[14]</sup>提出一种结合深度学习技术的姿态估计方法,首先通过深度神经网络得到姿态初值,其次基于姿态初值建立图像和目标模型之间的特征关联,进而求解目标姿态,但实际应用中目标通常会有多种姿态变化,需要设计多种模板构建模板库进而耗费大量算力,使检测速度较

慢从而难以应用到实际项目中;3)将深度学习目标检测网络与传统图像处理方法相结合,将检测到的目标代入相机标定后的模型得到目标的三维姿态<sup>[15]</sup>。该方法将深度学习网络与传统图像处理算法相结合,在光照、天气条件较差的环境下也能很好地检测到目标且大大降低了模型的复杂度,使得在性能较低、作业机械空间有限的嵌入式计算设备中也能达到实时检测的要求。

由于传统的激光雷达测量方式难以对集装箱三维姿态进行高精度测量,且目前主流的深度学习网络算法过于复杂、计算量高,本文设计了一种基于视觉的非接触三维测量系统。为了提高装卸过程中集装箱三维姿态测量的精度和实时性,采用将深度学习网络与传统图像处理算法相结合的方式。首先通过改进的深度神经网络对集装箱锁孔快速粗定位,同时使用目标跟踪网络对装卸目标进行准确跟踪;其次通过传统图像处理算法对集装箱锁孔进行二次定位得到集装箱锁孔的精确位置;最后结合装卸过程中集装箱的物理运动对集装箱姿态进行三维测量。

## 1 基于视觉的测量系统

本文提出的用于集装箱装卸的基于视觉的测量系统由一个图像处理单元、多台工业相机和用来固定相机的支架组成,图1以安装4台相机为例展示了现场相机安装位置。相机用来采集卸箱作业时集卡侧面的图像,相机安装位置能够采集到集装箱前后锁销的位置。

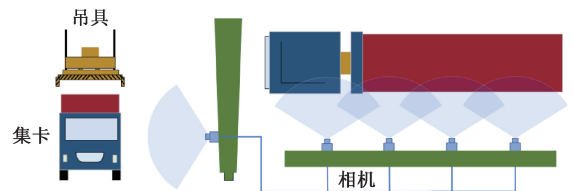


图1 基于视觉的测量系统硬件安装示意图

Fig. 1 Schematic diagram of hardware installation of vision-based measurement system

在集装箱装卸过程中,会发生集卡车被吊起、集装箱翻转等事故,为了避免此类事故的发生,本文使用相机采集集装箱锁孔在装卸过程中的三维位置变化,实现对集装箱装卸过程中的三维测量。集装箱锁孔如图2所示。

## 2 视觉测量算法

本文提出的视觉测量算法主要分为3个部分,基于检测的跟踪网络、目标二次定位网络和三维测量算法。其中基于检测的跟踪网络分为目标检测和目标跟踪两部



图 2 集装箱锁孔  
Fig. 2 Container keyhole

分,本文提出的目标检测网络使用卷积网络结构并引入注意力机制提高对集装箱锁孔特征的提取能力;目标跟踪部分使用简化后的 Deep SORT<sup>[16]</sup> 对锁孔目标进行实时跟踪;目标二次定位网络在目标检测、跟踪的基础上使用传统图像处理算法对集装箱锁孔进行二次定位,得到集装箱锁孔的精确位置;三维测量算法借助相机成像原理对集装箱锁孔三维姿态测量,进而估计集装箱姿态,实现对集装箱装卸过程的三维测量。

### 2.1 基于检测的跟踪算法

#### 1) 目标检测算法

本文检测的对象为集装箱锁孔,属于小目标,目前已经成熟的算法无论是单阶段目标检测算法还是双阶段目标检测算法,都无法直接对小目标进行精准、快速的检测。SSD<sup>[17]</sup> 算法作为单阶段目标检测算法中的典型算法,具有模型小、可移植到嵌入式设备等优点,同时检测精度、速度均有不错的效果,故本文对 SSD 算法进行改进,使其能实现对集装箱锁孔目标的快速检测。首先删减了网络层数,删除对大目标检测的高层卷积,精简网络结构,其次使用特征提取能力更强的骨干网络 Resnet-18<sup>[18]</sup> 代替 VGG-16<sup>[19]</sup> 以提高网络深度,同时引入注意力机制,在保证较快的检测速度的情况下来提高对小目标检测的精度,修改后的网络命名为 R-E-SSD。算法的网络结构如图 3 所示。

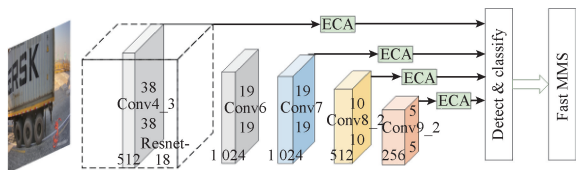


图 3 R-E-SSD 网络结构  
Fig. 3 R-E-SSD network structure

原始的 SSD 模型以 VGG-16 作为特征提取网络,但 VGG-16 网络参数量庞大,计算速度缓慢,无法达到实时分类检测集装箱锁孔的要求。因此本文将 SSD 的特征提取网络更改为 Resnet-18 深度残差网络,该网络仅有 18

层,具有更快的计算速度,且浮点计算量仅为 VGG-16 网络的 1/10,可以更好满足实时分类检测的要求,训练时可以使模型更快地收敛,从而减少训练时间。残差网络的结构如图 4 所示。

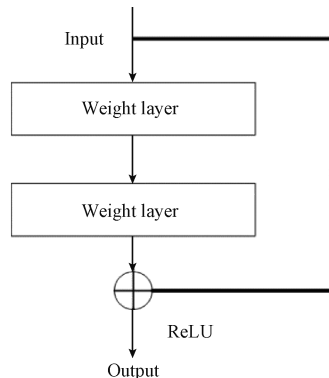


图 4 残差网络结构图  
Fig. 4 Residual network structure diagram

通道注意力机制能够很好地改善 CNN 的性能,提高目标特征提取和分类能力,因此本文引入了一种轻量级高效注意力机制 ECA-Net<sup>[20]</sup>。ECA-Net 是一种不降维的局部跨信道交互策略和自适应选择一维卷积核大小的模块,在增加极少的参数数量的同时增强网络有用信息、去除冗余特征。该网络首先对输入特征图的每个通道进行池化操作,获得全局感受野,然后直接进行局部跨通道连接,即通过考虑由池化操作得到的每个通道及其  $k$  个邻近通道来进行一维卷积操作, $k$  的取值通过通道数  $C$  自适应确定:

$$k = \psi(C) = \left\lceil \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right\rceil_{\text{odd}} \quad (1)$$

其中,  $\lceil \cdot \rceil_{\text{odd}}$  表示取离结果最近的奇数,  $\gamma$  和  $b$  表示常量,分别取值为 2 和 1。ECA-Net 网络模型的结构图如图 5 所示。

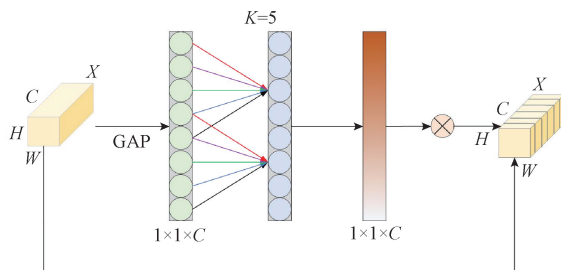


图 5 ECA-Net 网络结构图  
Fig. 5 ECA-Net network structure diagram

#### 2) 目标跟踪算法

在集装箱装卸过程中,仅使用深度学习算法对每一帧图像进行目标检测耗费大量内存且难以达到实时效

果,因此使用基于检测的跟踪网络 Deep SORT 对后续帧中集装箱锁孔的运动进行跟踪。Deep SORT 的跟踪效果很大程度上依赖检测器的检测效果,Deep SORT 检测器为上述提出的 R-E-SSD。Deep SORT 在 SORT<sup>[21]</sup> 的基础上使用更加可靠的关联方法来更好地解决目标被遮挡导致跟踪丢失的问题,不过在集装箱装卸过程中,几乎不存在集装箱锁孔被遮挡的现象,为此,本文对 Deep SORT 的关联方法进行一定的修改,删除了外观相似度匹配关联,保留运动信息关联从而保证算法的实时性,使其更能满足集装箱装卸过程中对集装箱锁孔的快速跟踪,算法结构如图 6 所示。

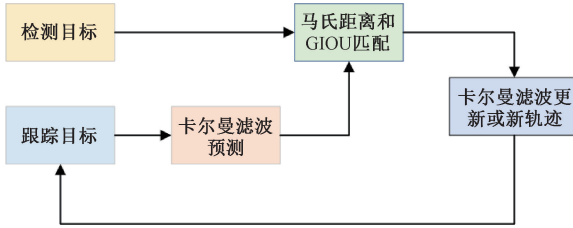


图 6 简化后的 Deep SORT  
Fig. 6 Simplified Deep SORT

本文提出的精简后 Deep SORT 工作流程分为预测、观测和更新 3 个步骤。在预测阶段,使用卡尔曼滤波对运动变量初始化,并预测下一帧的目标位置。卡尔曼滤波使用匀速和线性观测模型预测、更新目标轨迹,卡尔曼滤波状态变量构建如式(2)所示,  $u, v$  为目标检测结果的中心坐标,  $\tau, h$  分别为目标检测结果的纵横比和高度。 $\hat{u}, \hat{v}, \hat{\tau}, \hat{h}$  是由卡尔曼滤波预测的下一帧的目标位置。

$$X = [u, v, \tau, h, \hat{u}, \hat{v}, \hat{\tau}, \hat{h}] \quad (2)$$

通过 R-E-SSD 算法和卡尔曼滤波分别得到检测结果和预测结果,采用马氏距离衡量预测结果和检测结果之间的相似性,马氏距离的计算如式(3)所示,  $d_{i,j}$  为轨迹  $i$  与检测结果  $j$  之间的运动匹配值,  $S_i$  为该帧观测空间的协方差矩阵,通过卡尔曼滤波得到。

$$d_{i,j} = (d_j - y_j)^T S_i^{-1} (d_i - y_i) \quad (3)$$

在数据匹配阶段,使用匈牙利算法寻求预测结果与检测结果的匹配最优解,如式(4)所示。

$$\min Z = \sum_{i=1}^m \sum_{j=1}^n d_{i,j} x_{i,j} \quad (4)$$

匹配成功则进入卡尔曼滤波的更新阶段。式中  $m$  和  $n$  分别是跟踪到的目标数目以及检测到的目标数目。

数据匹配失败是导致集装箱装卸过程中目标跟踪失败的主要原因。数据匹配失败又分为跟踪阶段匹配失败和检测阶段匹配失败。跟踪阶段匹配失败主演原因为 R-E-SSD 网络漏检或跟踪过程中目标消失;检测阶段匹配失败产生的原因是跟踪过程中出现新目标或目标被遮

挡。由于集装箱装卸作业中几乎不存在目标被遮挡现象,故检测阶段匹配失败仅为出现新目标这一个原因。针对上述数据匹配失败问题,本文使用 GIOU 对匹配失败的跟踪目标与检测目标进行二次匹配,相比 IOU, GIOU 不仅关注重叠区域也关注非重叠区域, GIOU 如式(5)和(6)所示,并通过实验来确定最大阈值以去除相关性较低的检测框与跟踪框之间的匹配,若匹配成功则进入卡尔曼滤波的更新阶段。

$$GIOU = IOU - \frac{|A_c - U|}{|A_c|} \quad (5)$$

$$IOU = \frac{|A \cap B|}{|A \cup B|}, U = |A \cup B| \quad (6)$$

式中:  $A$  为检测结果,  $B$  为预测结果,  $A_c$  为检测结果与预测结果框的最小闭包区域面积。最接近预测结果的检测结果归类为同一目标。

### 2.2 目标二次定位算法

经过上述检测跟踪算法得到的锁孔图像如图 7 所示,接下来使用传统图像处理算法对集装箱锁孔进行二次定位,实现对集装箱锁孔的精确定位。



图 7 集装箱锁孔示意图

Fig. 7 Schematic diagram of container keyhole

从图 7 中可以看出,由于集装箱锁孔长期暴露在露天环境,集装箱锁孔周围产生了铁锈磨损,对于集装箱锁孔的精确定位造成了较大的困难。为此,首先对图像预处理,增强图像质量并剔除部分噪声,随后再精准定位集装箱锁孔,如图 8 所示。

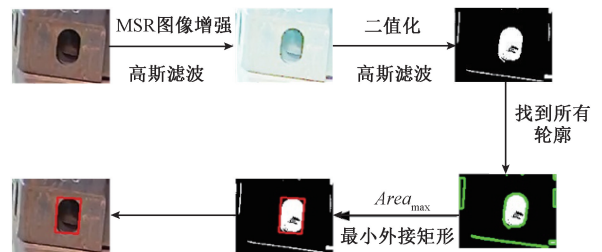


图 8 集装箱锁孔二次定位

Fig. 8 Secondary positioning of container keyhole

高斯滤波是一种线性平滑滤波,被广泛用于图像处理中的降噪,首先使用高斯滤波过滤原始图像部分噪声,再使用 MSR 算法对图像预处理增强图像质量, MSR 算法是基于 Retinex 理论<sup>[22]</sup> 发展而来用于图像增强的经典算法。MSR 算法用公式描述如式(7)所示:

$$R(x, y) = \sum_{k=1}^N \omega_k \{ \log [ I_i(x, y) ] - \log [ F_k(x, y) \cdot I_i(x, y) ] \} \quad (7)$$

式中:  $\omega_k$  为不同尺度对应的加权系数,  $\sum_{k=1}^N \omega_k$ ;  $N$  表示尺度参数的数量, 通常为 3;  $F_k(x, y)$  为不同尺度的环绕卷积函数,  $F_k(x, y)$  用公式描述如式(8)所示:

$$F_k(x, y) = \lambda e^{-\frac{(x^2+y^2)}{d_k^2}} \quad (8)$$

式中:  $d_k$  为大小不相同的尺度参数, 一般会选择 3 个不同尺度的参数, 使得尺度因子涵盖的范围较大。

集装箱锁孔图像经过图像增强后, 往往会产生一定的噪声。故使用高斯滤波对增强后的图像降噪, 随后对图像二值化处理, 得到阈值分割后的二值化图像。

从二值化图像中可以看出, 尽管高斯滤波去除了大量的噪声、保留了集装箱锁孔区域外, 仍然存在锁件边缘和锁孔内部阴影, 为了避免其他部分对于集装箱锁孔定位的干扰, 首先查找二值化图像中存在的所有闭合轮廓  $d_2$  并计算所有闭合轮廓面积  $d_2$ , 比较闭合轮廓的面积并找到最大面积  $d_2$  所代表的闭合轮廓为集装箱锁孔的外轮廓  $d_2$ , 最后使用最小外接矩形拟合集装箱锁孔轮廓实现对集装箱锁孔的精确定位。

### 2.3 三维测量算法

考虑到在集装箱装卸过程中, 一旦发生集卡吊起或集装箱翻转等事故, 需要通过三维测量获得集装箱和车架的姿态信息, 才能重新摆正集装箱, 然后再次起吊, 因此本文需要研究集卡被吊起时的集装箱姿态。

如图 9 所示, 国际标准 ISO1161 规定了集装箱锁孔的大小为 178 mm×118 mm×162 mm, 侧面的集装箱锁孔的大小约为 79 mm×52 mm。通过将集装箱锁孔实际大小与其在二次定位到的图像中成像的像素的大小建立关系从而得到集装箱锁孔的三维坐标, 实现对集装箱锁孔的三维测量。

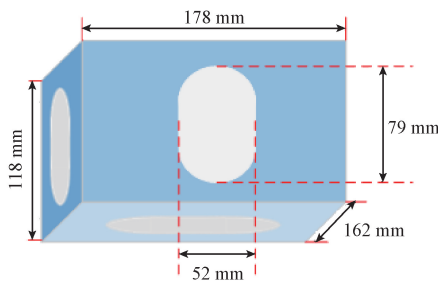


图 9 标准集装箱锁孔

Fig. 9 Standard container keyhole

#### 1) 像素距离转化为实际距离

使用相似三角形原理来得到图像像素与实际距离之间的对应关系, 原理如图 10 所示。该方法基于针孔摄像

机的特性, 其计算简单准确。

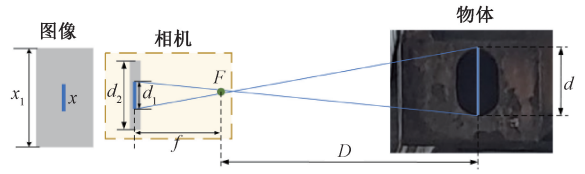


图 10 相机成像原理

Fig. 10 Camera imaging principle

如图 10 所示,  $F$  为相机的焦点,  $f$  为相机焦距,  $D$  为被测物体距相机焦点的距离,  $d$  为被测物体的实际长度,  $d_1$  为物体在相机成像面的长度,  $d_2$  是相机的内参数, 为相机像平面大小,  $x_1$  为图像总像素,  $x$  为被测物体在图像中成像占据像素坐标系下的像素个数。被测物体与相机焦点所成的三角形类似像与相机焦点形成的三角形, 根据三角形相似原理, 可以得到  $d$  与  $x$  之间的关系, 如式(9)所示。

$$d = \frac{D}{f} \cdot \frac{d_2}{x_1} \cdot x \quad (9)$$

若被测物体距相机焦点  $F$  的距离  $D$  已知, 则  $d$  与  $x$

则成正比关系, 即  $\kappa = \frac{D}{f} \cdot \frac{d_2}{x_1}$ , 其中  $\kappa$  定义为像素多少和实际距离之间的转换因子, 理想状态下, 图像水平方向转换因子与垂直方向上的转换因子相同, 即  $\kappa_y = \kappa_z$ 。

$$d = \kappa \cdot x \quad (10)$$

在计算距离之前, 需要注意的是, 大多数相机都有一些图像失真, 这是组装过程中的镜头失真和协调问题造成的。因此, 在计算位置参数之前, 使用张正友标定法<sup>[23]</sup>对图像进行校准。

#### 2) 锁孔偏移距离

为了实现在集装箱装卸过程中对集装箱目标的三维测量, 本文以竖直方向为  $Z$  轴, 集卡车行驶方向为  $Y$  轴, 以相机拍摄的方向为  $X$  轴分别测量集装箱锁孔在装卸过程中的三维位置变化。由于  $X$  方向上的偏移距离计算在实际操作如“集卡防吊起”等应用几乎并未涉及, 现实意义不大, 故本文暂时只讨论集装箱装卸作业过程中在  $Y, Z$  两个方向上的偏移距离。

如图 11 所示, 集装箱锁孔在  $Y$  方向和  $Z$  方向上的偏移距离指检测到的集装箱锁孔中心位置  $P_i$  与初始位置  $P_0$  之间在  $Y$  方向和  $Z$  方向上的移动距离。根据上文中提到的集装箱锁孔二次定位可以得到初始状态下和时刻  $i$  集装箱锁孔中心位置  $P_0(y_0, z_0)$  和  $P_i(y_i, z_i)$ 。已知集装箱距相机初始位置  $D_0$ , 根据相机成像原理得到初始状态下集装箱像素大小和实际距离之间比例关系  $\kappa_0 = \frac{D_0}{f}$ 。

$\frac{d_1}{x_1}$ ; 二次定位后集装箱锁孔的中心位置  $P_i(y_i, z_i)$ , 故集装箱锁孔在 Y 方向和 Z 方向下的偏移距离  $\Delta Y$  和  $\Delta Z$  如式 (11) 所示。

$$\begin{aligned} \Delta Y &= \kappa_0 | y_i - y_0 | \\ \Delta Z &= \kappa_0 | z_i - z_0 | \end{aligned} \quad (11)$$

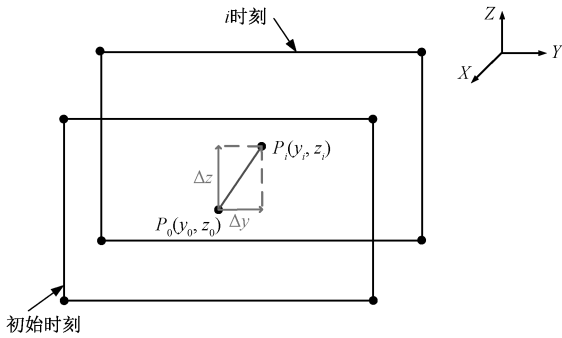


图 11 Y, Z 偏移距离  
Fig. 11 Y, Z offset distance

### 2.4 集装箱装卸目标视觉测量算法

本文提出的集装箱装卸目标视觉测量算法流程如下:

1) 输入集装箱装卸视频, 目标检测网络 R-E-SSD 对首帧图像进行集装箱锁孔检测, 得到包含集装箱锁孔目标的位置信息  $y_j$ ;

2) 后续帧中目标跟踪网络通过卡尔曼滤波得到集装箱锁孔目标的位置信息  $d_j$ ;

3) 通过式(3)计算检测目标  $y_j$  和卡尔曼预测目标  $d_j$  的马氏距离并使用式(4)数据匹配, 若匹配成功, 进入卡尔曼滤波更新阶段; 若匹配失败, 进入 GIOU 二次匹配, 若匹配成功, 进入卡尔曼滤波更新阶段; 若匹配失败, 生成新轨迹, 重复上述步骤, 直到跟踪整个视频序列;

4) 吊具停止上升, 对经过上述检测、跟踪算法得到的集装箱锁孔进行二次定位, 首先对集装箱锁孔使用式(7)进行 MSR 图像增强并去除图像多余噪声, 随后对图像二值化并检测二值化图像所有轮廓  $C_i$ , 寻找最大面积轮廓  $C_{max}$  使用最小矩形框精确定位集装箱锁孔;

5) 将二次定位算法得到的集装箱锁孔的精确位置与标准锁孔对比并通过公式(9)得到目标与像素的比例因子  $\kappa$ , 根据式(11)计算此时集装箱锁孔在竖直、水平方向上的偏移距离  $\Delta Y$  和  $\Delta Z$ , 实现对集装箱锁孔的三维测量。

## 3 实验

### 3.1 实验准备

本系统的测量精度主要取决于集装箱锁孔的检测精

度、二次定位的精度; 测量速度主要取决于集装箱锁孔的检测速度、跟踪速度以及二次定位的速度, 实验部分主要以精度和速度为指标测量这几个部分。本实验的训练环境为一个常见的工业计算机配置: GPU: Nvidia Tesla M40 (24 GB); CPU: Intel i7-6700; 操作系统: Ubuntu20.04 系统; 机器学习库: Pytorch1.3.0; 编程语言: Python3.7; 图像处理库: OpenCV4.0。

为验证本文设计的系统的可靠性, 实验使用的数据来自本团队在天津某码头使用多种尺寸集装箱进行集装箱装卸任务的实验。现场相机安装示意图如图 12 所示, 图像分辨率为  $1\ 980 \times 1\ 080$ , fps 为 24。



图 12 现场安装示意图  
Fig. 12 On-site installation diagram

本团队选取 40 尺集装箱和分别在前、中、后 3 个箱位的 20 尺集装箱做实验, 图 13 展示了不同箱位、不同相机所采集的数据。



图 13 数据集展示  
Fig. 13 Data set display

### 3.2 实验结果分析

#### 1) 目标检测实验

为对比本文设计的 R-E-SSD 对集装箱锁孔目标的检测效果, 使用原始 SSD 和 DSSD<sup>[24]</sup> 算法为对比算法并使用现场采集到的数据作为数据集, 经筛选标注后得到 2 783 张图像, 以 8:1:1 划分数据集。对比结果如表 1 所示。检测速度的评价指标为 FPS (frame per second, FPS), 定义为网络每秒处理图像的帧数; 检测精度的评价指标为 AP (average precision), 定义为单一类别的平均

精度。检测结果如表 1 所示。

表 1 目标检测性能对比

Table 1 Object detection performance comparison

	Backbone	Input	AP/%	FPS
SSD	VGG-16	300×300	71.86	36
DSSD	Resnet-101	320×320	73.46	17
Ours	Resnet-18	300×300	95.4	30.7

由表 1 可知,本文设计的 R-E-SSD 算法相比原始的 SSD 和 DSSD<sup>[24]</sup>算法在集装箱锁孔目标的检测效果均有较大提升,分别为 23.54% 和 21.94%;从检测速度上看,R-E-SSD 算法检测速度虽然相比原始的 SSD 略有降低但也符合实时检测的要求。实验结果表明,本文提出的算法能够在实时测试的基础之上有效地提高复杂背景下对集装箱锁孔目标检测准确率。

## 2) 目标跟踪实验

为了更加直观地展示本文提出的算法目标跟踪效果,本文使用了两段集装箱装卸视频作为测试并截取视频中几帧图像展示目标跟踪效果,如图 14 所示,图中左下角展示了视频序列帧,分别为两段视频的首帧、120 帧和第 5、135 帧,从图 14 中可以看出,本文算法并未出现跟踪目标丢失的情况。



图 14 目标跟踪结果

Fig. 14 Target tracking results

为了验证本文所提出的目标跟踪算法的性能,将本文算法与原始跟踪算法的结果进行对比,选取多目标跟踪中使用的跟踪准确度 (MOTA) 及每秒传输帧数 (FPS) 为指标来判断算法的准确度和实时性。对比结果如表 2 所示。由表 2 可知,优化后的目标检测算法 (R-E-SSD) 与精简后的 Deep SORT 相结合能够保证较高的跟踪精度的同时保持较快的跟踪速度,说明本文提出的算法明显优于原始算法,更加适用于集装箱装卸过程中集装箱锁孔的检测和跟踪。

## 3) 二次定位实验

由于深度学习网络使用大量数据、通过端到端的方式提取数据更为深层的特征,实现对数据更为准确地表

表 2 目标跟踪算法性能对比结果

Table 2 Comparison results of target tracking algorithm performance

	MOTA/%	FPS
SSD+Deep SORT	79.73	14.2
Ours	96.52	21.7

达,使其在复杂环境下也能达到较高的精度。在上述深度学习网络的基础上,在集装箱装卸过程中即使在光照条件不足的情况下,也能对集装箱锁孔目标实现较高精度的粗定位。对集装箱锁孔的二次定位是在深度学习网络跟踪的结果上采用传统图像处理算法对目标进行二次定位,其二次定位的精度主要取决于深度学习网络的精度。因此,本实验主要测试传统图像处理算法的定位误差以及检测速度。

图像处理算法的实现基于 Python 和 OpenCV。实验部分,本文使用了一组 278 张经 R-E-SSD 检测后的集装箱锁孔的,这些图像是在白天和晚上不同时刻对不同箱型的集装箱采集所得,文中仅使用几张在典型光照条件下采集的图像作为实验代表,经传统图像处理算法检测后的结果如表 3 所示。表中前两行使用的图像为白天不同光照下采集所得,后两行使用的图像分别为晚上起重机开启大灯作为辅助照明和傍晚起重机未开启大灯时采集的图像。从表中可以看出,传统图像处理算法对于集装箱锁孔的二次定位具有良好的检测效果,且使用 MSR 图像增强算法能够有效去除噪声并且保持图像质量,突出锁孔目标,有利于实现对集装箱锁孔的精确定位。

结合上述的检测、跟踪算法,对集装箱锁孔进行二次定位实验以判断整体检测算法的性能。整体算法的检测时间  $T$  是 R-E-SSD 对集装箱锁孔的粗定位、装卸过程中的跟踪以及传统图像算法对集装箱锁孔二次定位时间之和,检验整体算法的实时性,即  $T = T_{detect} + T_{track} + T_{secondary}$ ; 检测准确率  $Acc$  是 R-E-SSD 算法的检测率、跟踪算法的准确率以及传统图像处理算法的乘积,检验整体算法的精确度,即  $Acc = Acc_{detect} \times Acc_{track} \times Acc_{secondary}$ 。经计算,整体算法的检测时间  $T$  约为 80.1 ms,约 12.45 fps,检测准确率  $Acc$  约为 93.71%。

## 4) 三维测量实验

为验证本文提出的算法在实际应用场景下的有效性,使用激光雷达作为对照组,在相同场景下使用本文算法与激光雷达分别测量,检验本文算法测量的精确度。本文使用的是 16 线激光雷达,其水平扫描角度为  $360^\circ$ ,垂直扫描角度为  $\pm 15^\circ$ 。对于本文提出的视觉测量方法是以集装箱锁孔为检测目标,通过测量集装箱锁孔在装卸过程中的姿态变化实现对集装箱的姿态测量;对于 16 线激光雷达,由于其无法分辨集装箱的细节,本文标记几个

表 3 二次定位结果  
Table 3 Secondary positioning results

图像	MSR 增强	二值化	闭合轮廓 $C_i$	$\max Area_i$	$C_{\max}$

点作为测量的基准点,通过测量这几个点的变化实现对集装箱姿态测量。图 15 为视觉测量示意图,图 16 为激光雷达测量示意图,图中标记了集卡车头、集卡平板、集装箱以及测量基准点的位置。

的距离与实际吊具移动的距离的差值,分别为  $Y$ -error,  $Z$ -error。

初始状态下,集装箱距相机的实际距离  $D_0$  为 2.7 m,相机分辨率为 1 920×1 080,使用 40 尺集装箱进行装卸作业模拟,40 尺集装箱尺寸为 12.192 m×2.438 m×2.591 m。根据上文提到的像素距离与实际距离的比例因子得到初始状态下  $\kappa_y = 3.59, \kappa_z = 3.25$ 。当吊具在  $Z$  和  $Y$  方向上移动时,像素距离与实际距离的比例因子  $\kappa_y$  和  $\kappa_z$  保持不变。每完成一次两个方向上的移动所采集的数据为一组实验数据,共 10 组。激光雷达与视觉测量偏移距离实验结果如表 4 所示。



图 15 视觉姿态测量

Fig. 15 Visual attitude measurement

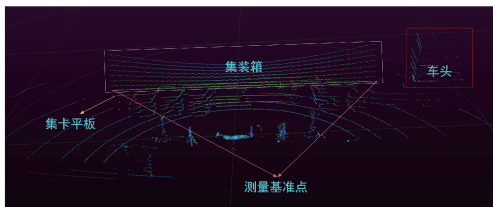


图 16 激光雷达姿态测量

Fig. 16 LiDAR attitude measurement

控制吊具每次仅在单一方向上移动来检验本文提出的集装箱偏移距离测量的准确性,其中吊具在不同方向上移动固定距离。 $Z$  方向上移动距离为 30 cm,  $Y$  方向上移动距离为 30 cm,集装箱偏移距离评价指标为计算得到

表 4 激光雷达与视觉测量对比实验表

Table 4 LiDAR and visual measurement comparison experiment table

实验次数	视觉测量		激光雷达测量	
	$Y$ -error/cm	$Z$ -error/cm	$Y$ -error/cm	$Z$ -error/cm
实验 1	1.48	1.27	19.63	17.05
实验 2	1.90	1.39	17.16	15.24
实验 3	1.83	1.26	16.01	13.65
实验 4	1.88	1.08	17.65	16.77
实验 5	1.93	1.41	16.24	13.72
实验 6	1.95	1.47	17.21	15.77
实验 7	1.77	1.01	17.54	14.16
实验 8	1.34	1.11	18.63	17.12
实验 9	1.18	1.20	15.84	12.67
实验 10	1.93	1.41	18.97	17.30

从表 4 中可以得到,在集装箱装卸过程中,激光雷达



测量方法测得集装箱锁孔在  $Y/Z$  方向的偏移距离的平均值分别约为 17.49 cm 和 15.3 cm, 而本文设计的视觉测量方法测得集装箱锁孔在  $Y/Z$  方向的偏移距离的平均值分别约为 1.72 cm 和 1.26 cm, 激光雷达测量误差约为视觉测量误差的 8~10 倍, 视觉测量精度明显高于激光雷达。集装箱仅在竖直提升和水平方向移动时, 集装箱像素大小和实际距离之间比例因子  $\kappa$  几乎不变, 从而保证了测量误差较小, 能够满足集装箱码头自动化装卸的要求。

## 4 结 论

为了解决自动化集装箱码头装卸过程中对集装箱实时、准确地姿态测量的问题, 与现有的激光雷达和直接在图像中进行三维测量的深度学习网络等视觉测量算法不同, 本文针对自动化码头复杂环境下、全天候工作的特点提出了一种深度学习与传统图像处理算法相结合的目标三维测量算法, 设计了一种基于视觉的集装箱装卸三维测量系统。首先, 使用基于检测的目标跟踪网络对装卸目标进行快速粗定位, 在目标检测网络中引入注意力机制并且在目标跟踪网络中剔除外观特征匹配阶段来提升模型的精度和速度; 其次结合传统图像处理算法对装卸目标实现二次精准定位; 最后分析装卸过程中集装箱的运动情况建立模型测量集装箱的三维姿态。由实验结果可知, 本文提出的算法具有快速、准确等优点, 能够对集装箱姿态实现较高的检测准确率, 满足集装箱码头自动化装卸的要求。在后续的研究工作中, 尝试使用多个相机对装卸过程中的集装箱进行多视角测量并尝试引入新变量对其测量, 如集装箱的偏移角度等。

## 参考文献

- [ 1 ] MI C, HUANG Y G, FU C H, et al. Vision-based measurement: Actualities and developing trends in automated container terminals[J]. IEEE Instrumentation & Measurement Magazine, 2021, 24(4) : 65-76.
- [ 2 ] HUANG Q F, HUANG Y G, ZHANG Z W, et al. Truck-lifting prevention system based on vision tracking for container-lifting operation[J]. Journal of Advanced Transportation, 2021, 2021 : 1-9.
- [ 3 ] 张宗华, 刘巍, 刘国栋, 等. 三维视觉测量技术及应用进展[J]. 中国图象图形学报, 2021, 26(6) : 1483-1502.  
ZHANG Z H, LIU W, LIU G D, et al. 3D visual measurement technology and application progress [J]. Chinese Journal of Image and Graphics, 2021, 26(6) : 1483-1502.
- [ 4 ] 杨岳航, 陈武雄, 朱明, 等. 基于机器视觉的无人机自  
主着陆技术[J]. 国外电子测量技术, 2020, 39(4) : 57-61.  
YANG Y H, CHEN W X, ZHU M, et al. Autonomous landing technology of UAV based on machine vision[J]. Foreign Electronic Measurement Technology, 2020, 39(4) : 57-61.
- [ 5 ] 匡翠萍, 刘旭, 夏子龙, 等. 基于计算机视觉的平水桥海滩风暴潮响应分析[J]. 同济大学学报(自然科学版), 2022, 50(7) : 1009-1016.  
KUANG C P, LIU X, XIA Z L, et al. Analysis of storm surge response of Pingshuiqiao beach based on computer vision [J]. Journal of Tongji University (Natural Science Edition), 2022, 50(7) : 1009-1016.
- [ 6 ] 田如安, 李筠, 杨海马, 等. 基于机器视觉的汽车减震杆检测系统[J]. 电子测量技术, 2019, 42(22) : 103-106.  
TIAN R AN, LI J, YANG H M, et al. Automotive shock absorbing rod detection system based on machine vision[J]. Electronic Measurement Technology, 2019, 42(22) : 103-106.
- [ 7 ] 黄海鸿, 刘文杰, 钱正春, 等. 零件表面损伤对再制造修复件承载能力影响的超声相控阵检测研究[J]. 电子测量与仪器学报, 2020, 34(7) : 32-41.  
HUANG H H, LIU W J, QIAN ZH CH, et al. Ultrasonic phased array detection of the effect of part surface damage on the bearing capacity of remanufactured repaired parts [J]. Journal of Electronic Measurement and Instrumentation, 2020, 34(7) : 32-41.
- [ 8 ] 柳宁, 张治国, 李德平, 等. 基于视觉跟踪实时引导的传送带跟随方法[J/OL]. 计算机集成制造系统, [2022-12-03] : 1-21.  
LIU N, ZHANG ZH G, LI D P, et al. Conveyor belt following method based on visual tracking real-time guidance [J/OL]. Computer Integrated Manufacturing System; [2022-12-03] : 1-21.
- [ 9 ] 于重重, 萨良兵, 马先钦, 等. 基于度量学习的小样本零器件表面缺陷检测[J]. 仪器仪表学报, 2020, 41(7) : 214-223.  
YU ZH ZH, SA L B, MA X Q, et al. Surface defect detection of small samples of zero devices based on metric learning [J]. Chinese Journal of Scientific Instrument, 2020, 41(7) : 214-223.
- [ 10 ] LIU M, ZHU Q G, YIN Y F, et al. Damage detection method of mining conveyor belt based on deep learning[J]. IEEE Sensors Journal, 2022, 22(11) : 10870-10879.
- [ 11 ] TEKIN B, SINHA S N, FUA P. Real-time seamless

- single shot 6D object pose prediction [ C ]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, 292-301.
- [ 12 ] BRAUN M, QING R, WANG Y K, et al. Pose-RCNN: Joint object detection and pose estimation using 3D object proposals [ C ]. IEEE 19th International Conference on Intelligent Transportation Systems ( ITSC ), 2016, 1546-1551.
- [ 13 ] WEINZAEPFEL P, BREGIER R, COMBALUZIER H, et al. DOPE: Distillation of part experts for whole-body 3D pose estimation in the wild [ C ]. Computer Vision-ECCV 2020, 2020, 12371: 380-397.
- [ 14 ] 任笑圆, 蒋李兵, 钟卫军, 等. 基于视觉的非合作空间目标三维姿态估计方法 [ J ]. 电子与信息学报, 2021, 43(12): 3476-3485.  
REN X Y, JIANG L B, ZHONG W J, et al. Vision-based 3D pose estimation method for non-cooperative space targets [ J ]. Journal of Electronics and Information, 2021, 43(12): 3476-3485.
- [ 15 ] 崔海华, 姜涛, 杜坤鹏, 等. 基于深度学习位姿估计的多视结构光三维成像方法 [ J ]. 光学学报, 2021, 41(17): 72-81.  
CUI H H, JIANG T, DU K P, et al. Multi-view structured light 3D imaging method based on deep learning pose estimation [ J ]. Journal of Optics, 2021, 41(17): 72-81.
- [ 16 ] WOJKE N, BEWLEY A, PAULUS D. Simple online and realtime tracking with a deep association metric [ C ]. 2017 IEEE International Conference on Image Processing ( ICIP ), 2017: 3645-3649.
- [ 17 ] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multiBox detector [ C ]. European Conference on Computer Vision, 2016: 21-37.
- [ 18 ] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [ C ]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [ 19 ] SIMONYANIM K, ANDREW Z. Very deep convolutional networks for large-scale image recognition [ C ]. International Conference on Learning Representations ( ICLR ), 2015: 1-14.
- [ 20 ] WANG Q L, WU B G, ZHU P F, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks [ C ]. IEEE/CVF Conference on Computer Vision and Pattern Recognition ( CVPR ), 2020, 11531-11539.
- [ 21 ] BEWLEY A, GE Z Y, OTT L, et al. Simple online and real-time tracking [ C ]. IEEE International Conference on Image Processing ( ICIP ), 2016, 3464-3468.
- [ 22 ] ZOTIN A. Fast algorithm of image enhancement based on multi-scale retinex [ J ]. Procedia Computer Science, 2018, 131: 6-14.
- [ 23 ] ZHANG Z Y. A flexible new technique for camera calibration [ J ]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(11): 1330-1334.
- [ 24 ] CHENG Y F, WEI L, ANANTH R, et al. DSSD: Deconvolutional single shot detector [ J ]. arXiv preprint arXiv:1701.06659, 2017.

### 作者简介



宓超 (通信作者), 2007 年于上海海事大学获得学士学位, 2009 年于法国南特大学获得硕士学位, 2010 年于上海海事大学获得硕士学位, 2019 年于上海海事大学获得博士学位, 现为上海海事大学物流工程学院副教授, 主要研究方向为物流工程、自动化控制、模式识别等。

E-mail: chaomi@shmtu.edu.cn

**Mi Chao** (Corresponding author) received a B. Sc. degree from Shanghai Maritime University in 2007, a M. Sc. degree from Nante University in France in 2009, a M. Sc. degree from Shanghai Maritime University in 2010, a Ph. D. degree from Shanghai Maritime University in 2019. Now he is an associate professor of Logistics Engineering College of Shanghai Maritime University. His main research interests include logistics engineering, automation control, pattern recognition, etc.



黄世凤, 2020 年于北方民族大学获得学士学位, 现为上海海事大学物流科学与工程研究院硕士研究生, 主要研究方向为机器学习。

E-mail: 202030510010@stu.shmtu.edu.cn

**Huang Shifeng** received her B. Sc. degree from North Nationalities University in 2020. Now she is a M. Sc. candidate of Logistics Science and Engineering Research Institute of Shanghai Maritime University. Her main research interests include machine learning.



张钰洁, 现为上海海事大学物流工程学院硕士研究生, 主要研究方向为机器学习。

E-mail: 202130210133@stu.shmtu.edu.cn

**Zhang Yujie** is now a M. Sc. candidate of Control Science and Engineering, College of Logistics Engineering, Shanghai Maritime University, China. Her research interest includes machine learning.

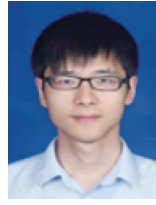


张志伟, 2016 年于上海海事大学物流工程学院机电工程获得硕士学位。目前为上海海瞩智能科技有限公司的工程师, 主要研究方向为机器视觉。

E-mail: zhangzhiwei@smuvision.com

**Zhang Zhiwei** received the M. Sc.

degree in Mechatronic Engineering, Logistic Engineering School, Shanghai Maritime University, Shanghai, China, in 2016. He is now an engineer with the Shanghai SMUVision Smart Technology Ltd, Shanghai, China. His research interest includes machine vision.



姚磊, 于中国南京航空航天大学航天工程学院取得飞机设计与工程学士学位, 目前在上海 SMUVision 智能科技担任高级工程师, 主要研究方向为机器视觉和图像测量。

E-mail: 962174173@qq.com

**Yao Lei** is now working in Shanghai

SMUVision Smart Tech Ltd. as a senior engineer. He received the B. Sc. degree in Aircraft Design and Engineering with the College of Aerospace Engineering, Nanjing University of Aeronautics and Astronautics, China. His main research interests include machine vision and image measurement.