

基于单模板的平面多目标识别与定位方法

庞博¹, 杨夏²

1. 国防科学技术大学航天科学与工程学院, 长沙 410073; 2. 国防科学技术大学航天科学与工程学院, 长沙 410073

摘要: **目的:** 本文以机械臂实现平面工件的精确拾取为背景, 提出了一种基于归一化互相关 (Normalized Cross Correlation method, NCC) 匹配算法与非极大值抑制 (Non-maximum suppression, NMS) 相结合的平面多目标识别与定位方法。 **方法:** 该方法首先在源图像上选取矩形模板, 然后利用高斯金字塔将源图像与模板图像降采样。接下来使用NCC方法得到源图像与模板图像匹配得到的响应值图像, 通过非极大值抑制算法, 得到图像中多目标的响应点, 再通过拉普拉斯金字塔复原, 最后在源图像中标示出目标的位置。 **结果:** 本文针对该方法分别进行了仿真实验与实物实验, 准确率很高, 算法运行时间小于200ms。 **结论:** 结果证明该方法具有识别准确率高, 速度快的特点, 适合实际工程任务的需要。

关键词: 归一化互相关 (NCC); 非极大值抑制 (NMS); 单模板; 多目标; 识别; 定位

Planar multi-target recognition and localization method based on single template

Pang Bo¹, Yang Xia²

1. College of Aerospace Science and Engineering, National University of Defense Technology, Changsha, 410073

2. College of Aerospace Science and Engineering, National University of Defense Technology, Changsha, 410073

Abstract: **Objective:** For achieving the goal of picking planar object correctly by mechanical arm, this paper proposes a method of planar multi-target recognition and localization based on Normalized Cross Correlation (NCC) method and Non-maximum suppression (NMS). **Method:** Firstly, we should select a rectangular template on the source image, and then the source image and the template image are reduced by using the Gaussian Pyramid. Next, we use the NCC method to obtain the response image of the source image and the template image and get the response points of multiple targets in it through NMS. And finally, we can indicate the location of the target in the source image by the restoration through the Laplacian Pyramid. **Result:** In this paper, we carried out simulation experiments and real experiments on this method. The accuracy rate is pretty high and the running time of the algorithm is less than 200ms. **Conclusion:** The results show that this method has the characteristics of high accuracy and speed. It is really suitable for the actual engineering task.

Key words: Normalized Cross Correlation; Non-maximum suppression; single template; multi-target; recognition; localization

收稿日期: ; 改回日期:

第一作者简介: 庞博(1992-), 男, 硕士, 主要研究方向为图像测量与视觉导航。E-mail:465277043@qq.com

0 引言

近年来,随着机器视觉领域的研究与发展,越来越多的工业机器人配备了视觉系统,使机器人在自主移动和机械臂准确定位技术方面有了显著提高。与此同时,形形色色的视觉和图像算法的成功应用也显著地提升了机械臂的智能水平,而机械臂对目标物进行抓取的实时性、准确性和稳定性,也是机械臂采用视觉引导进行智能抓取的智能水平的标准体现。要提高机械臂所作的一系列动作的智能水平,使其具备智能化移动和抓取的功能,关键是使机器人对环境目标具备一定的认知能力,通过感知环境目标确定自身的行动,使其不需运动接触即可对环境目标实现抓取等动作,因此,视觉系统的研究已成为提高机器人智能化的一个较为热门的方向,目标的识别与定位方法更是这其中重点的研究问题。

目前,在该研究领域内,已经有很多国内外学者开展了相关的研究工作,并取得了较大的进展。唐宇等^[1]提出了一种基于区域划分的平面工件识别与定位方法,利用 LSD 方法进行直线段的提取,然后将杂线段分类进行剔除,在保留的轮廓线段基础上进行直线段的连接和区域划分,最后利用各个区域的形状特征向量和模板的形状特征向量进行对比,以完成工件的识别和定位。但是该方法只对具有直线轮廓的工件适用,当工件为不规则形状时无法得到正确的识别结果。王修岩等^[2]提出了一种基于单目视觉技术的解决方法,给出一种新的统计矩阵标记区域算法,利用目标本身的几何特征达到识别目的,选取质心标记法实现目标物体的定位。该方法用于识别杂乱物体中的某一个形状物体,并不适用于相同的多目标识别与定位。为了实现移动机器人的目标识别,山东大学^[3]提出了一种实时性要求高的算法,采用 SURF 算法提取视觉图像中的目标特征,利用 BBF 算法对目标特征进行匹配,并采用 RANSAC 算法消除误匹配点。尽管该方法识别效果较好,适应仿射变换,可 SURF 或 SIFT 算法只能对纹理丰富的物体进行识别。另外王红涛等^[4]给出了一种有效的基于边缘匹配的工件识别算法,该算法采用了 Canny 算子提取的边缘信息作为匹配特征,将改进的 Hausdorff 距离作为图像匹配的相似性度量,在搜索过程中采用了自适应代沟替代策略的遗传算法。优点是能有效解决具有平移、旋转和部分

遮挡等情况,但缺点也很明显,当图像分辨率较大时,算法运行时间较长,无法满足实际工程任务的实时性要求。针对相对定位技术在复杂地面背景下存在的相对定位目标选取质量不高、识别难度大的问题,陈世伟等^[5]基于最稳定极值区域(MSER)和双层匹配矫正策略提出了一种新的相对定位目标选取与识别算法。该方法识别精度高、抗干扰能力强,但也存在相对定位目标的选取数量和质量人工很难把握的问题。Koshiro Yata 等^[6]提出了一种基于特征模板并利用主成分分析(PCA)方法进行模板匹配的算法,实现了多种类别目标的识别与定位,具有一定的鲁棒性并提高了效率。A. Chayeb^[7]介绍了一种快速有效的多物体识别检测的方法,该方法采用 AdaBoost 算法对模板进行学习,对每一类的目标创建 HOG 特征,利用该特征来进行对物体的检测与识别。类似方法还有很多,在算法运行之前都需要较长的离线时间对模板或特征进行学习,无法满足工程任务需求的多变性。

现有的目标识别与定位算法主要存在以下问题:(1)基于灰度或特征匹配的算法只能识别单目标;(2)多目标的识别方法需要较长的离线时间对模板或特征进行学习,不适应工程任务多变的实际情况;(3)大多数算法可以保证识别定位的准确性,但是速度较慢,无法满足实时性的要求。

为了克服现有方法的缺点,本文以机械臂准确拾取平面工件的实际工程任务为背景,提出了一种基于归一化互相关(Normalized Cross Correlation method, NCC)匹配算法^[8]与非极大值抑制(Non-maximum suppression, NMS)方法^{[9][10]}相结合的平面多目标识别与定位方法。利用高斯金字塔降采样后的源图像与模板图像进行归一化互相关匹配,再通过结合距离阈值的非极大值抑制算法,得到目标在响应图像上的响应点,再通过拉普拉斯金字塔复原,进而确定目标在源图像上的位置。仿真实验和实物实验识别准确率很高,算法运行时间小于 200ms,验证了本文方法的准确性与实时性,满足实际工程任务需求。

1 方法原理

1.1 图像金字塔

以多个分辨率来表示图像的一种有效且概念简单的结构是图像金字塔^[11]。通常情况下,典型的图像金字塔分为两类:

(1) 高斯金字塔：向下采样，常用的图像金字塔；

(2) 拉普拉斯金字塔：从金字塔低层图像重建上层图像，配合高斯金字塔一起使用，可以对图像进行最大程度的复原，亦可以预测残差金字塔。

高斯金字塔包含低通滤波器，其截止频率从上一层到下一层以 2 为系数逐层增加，可以跨越很大的频率范围。它通过高斯平滑和亚采样获得一系列下采样图像，每一层都按照从下到上的顺序编号，层 G_{i+1} 表示尺寸小于层 G_i 的相邻层。给定一个整数变量 n 和一维取样序列 $f(n)$ ，则基为 2 的上采样序列 $f_{2\uparrow}(n)$ 如下式定义：

$$f_{2\uparrow}(n) = \begin{cases} f(n/2), n \text{ 为偶数} \\ 0, \text{其他} \end{cases} \quad (1)$$

基为 2 的下采样操作的定义为：

$$f_{2\downarrow}(n) = f(2n) \quad (2)$$

拉普拉斯金字塔是通过源图像减去先缩小后放大的图像的一系列图像构成的，可以理解为高斯金字塔的逆形式，拉普拉斯金字塔第 i 层的数学定义：

$$L_i = G_i - UP(G_{i+1}) \otimes g_{5 \times 5} \quad (3)$$

式中 G_i 表示第 i 层图像，UP 操作是将源图像中位置为 (x,y) 的像素映射到目标图像的 $(2x+1,2y+1)$ 位置，符号 \otimes 表示卷积运算， $g_{5 \times 5}$ 表示 5×5 的高斯内核。

1.2 归一化互相关匹配 (NCC)

模板匹配^[12]是一个简单过程，我们需要将模板与图像进行匹配，该模板是需要寻找的目标（子图像）。将模板的中心放在一个图像点上，然后计算模板中有多少个点与图像点相匹配。对整幅图像重复此过程，相应最大的点，即最佳匹配点，也就是目标（通过模板给出）在图像中所处的位置。模板匹配的主要优点是对噪声和遮挡的鲁棒性，固有的平均操作可以降低对噪声的敏感程度，最大化过程可以降低对遮挡的敏感程度。

归一化互相关匹配算法是一种经典的模板匹配算法，通过计算模板图像和匹配图像的互相关值，可以确定匹配的程度，互相关值最大时的搜索窗口位置决定了模板图像在待匹配图像中的位置。NCC 算法具有很高的准确性和适应性，并且对图像灰度值的线性变换具有鲁棒性，不受灰度值的线性变换

的影响^[13]。

NCC 算法响应函数的数学表达形式如下式所示：

$$R(x, y) = \frac{\sum_{x', y'} (T(x', y') \cdot I(x+x', y+y'))}{\sqrt{\sum_{x', y'} T(x', y')^2 \cdot \sum_{x', y'} I(x+x', y+y')^2}} \quad (4)$$

$I(x,y)$ 表示源图像， $T(x,y)$ 表示模板图像。

NCC 算法尽管具有能适应线性灰度畸变的优点，但是最显著的缺点是计算开销比较大，不太适应工业生产需求的需求。如果模板是正方形且尺寸为 $m \times m$ ，与尺寸为 $N \times N$ 的图像进行匹配，由于有 m^2 个像素与所有除边界外的像素点进行匹配，计算的开销为 $O(N^2 m^2)$ ，这是位置不变性模板匹配的开销^[11]。故而本文先将源图像和模板图像利用图像金字塔进行降采样，再利用 NCC 方法进行模板匹配来降低计算开销。

由机械臂与单目相机构成的手眼系统拍摄的源图像如图 1 所示：



图1 源图像

Fig.1 Source image

在源图像中利用鼠标框选出如图 2 所示的矩形模板：



图2 模板图像

Fig.2 Template image

以系数为 2 将源图像与模板图像进行降采样，然后利用降采样后的源图像和模板图像进行归一化

互相关匹配，利用式(4)计算得到的响应图像经过伪彩色处理后如图 3 所示：

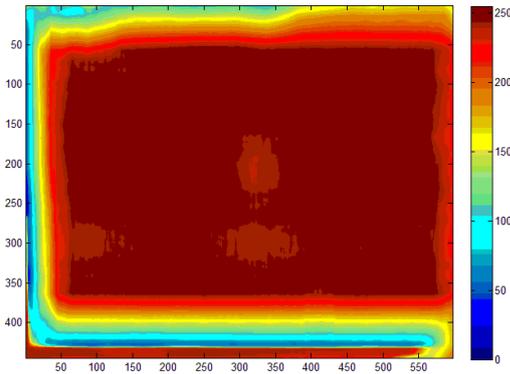


图3 响应图像

Fig.3 Response image

1.3 非极大值抑制(NMS)

模板匹配算法通过计算模板图像和匹配图像的互相关值，可以确定匹配的程度，互相关值最大时的搜索窗口位置决定了模板图像在待匹配图像中的位置，但是那仅限于识别单目标的情况。针对基于单模板的多目标识别问题，由于光照、噪声等一系列干扰的影响，需要找到多个目标就需要找到超过一定响应值的所有位置，并去除杂点（在响应图像中体现为粘连在一起的像素点）。

非极大值抑制^[14](NMS)是局部极大值的搜索问题，NMS 在机器视觉特别是物体的识别与定位中具有广泛的运用。如何设计高效的 NMS 算法对于许多应用是十分关键的，例如三维重建、模式识别、纹理分析以及视频跟踪等。

在一个窗口内，如果有多个像素点的响应值满足阈值条件，则保留响应值最大的那个像素点，排除其它像素点。这样可以避免相同目标的重复检测，提高算法效率。本文首先将符合响应阈值条件的像素点按照响应值降序排列，然后利用图像的膨胀[11]运算来达到检测极大值的目的，因为膨胀运算可以用窗口内的最大值替代当前的灰度值。另外，为了避免单一约束的不确定性，本文还设置了距离阈值对满足条件的像素点进行约束，从而达到去除误匹配的作用。两像素点间距离的表达式为：

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (5)$$

根据经验可知，当NCC响应函数值超过0.9时，匹配情况良好。设置膨胀操作窗口 18*18，距离阈值选取为 0.9*模板短边长，得到模板在源图像中的响应位置图像如图 4 所示（响应位置为黑点处）：

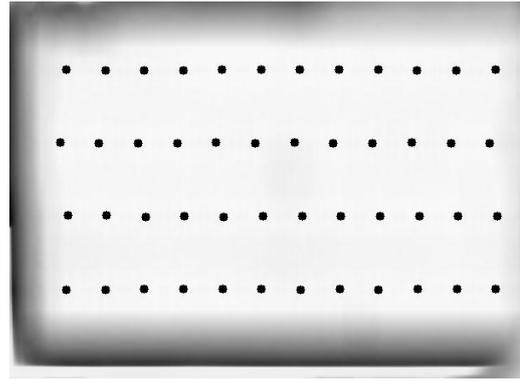


图4 响应位置图像

Fig.4 Image of responding location

最终在源图像中的识别与定位结果如图 5 所示：



图5 结果图像

Fig.5 Result image

2 实验结果与分析

2.1 仿真实验

为了验证该方法可以满足多目标识别定位的准确性与实时性的需求，本文针对不同形状的平面物体进行了仿真实验，并且计算了算法的运行时间。实验所需参数设置如下：NCC 响应阈值 0.9，膨胀窗口 18*18，距离阈值选取为 0.9*模板短边长。仿真源图像的分辨率为 1280*1024，模板图像在源图像上利用鼠标框选。实验平台为 Intel(R)Core(TM) i5-3470 CPU @3.20GHz 的 PC，未使用并行质量优化等显著优化算法。两种形状的源图像及模板图像如图 6 所示，以 2 为因数降采样次数分别为 0、1、2 时的多目标识别与定位结果如图 7、8、9 所示，算法识别率与运行时间如表 1 所示。

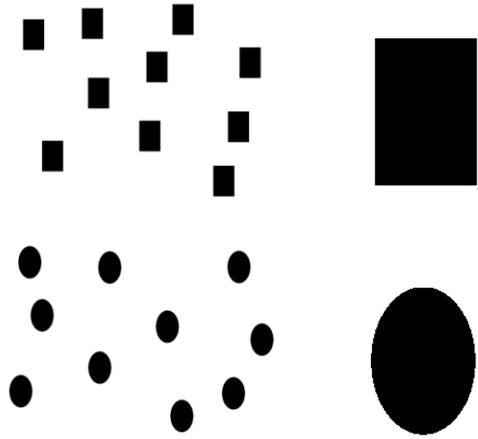


图6 源图像及模板图像

Fig.6 Source image and template image

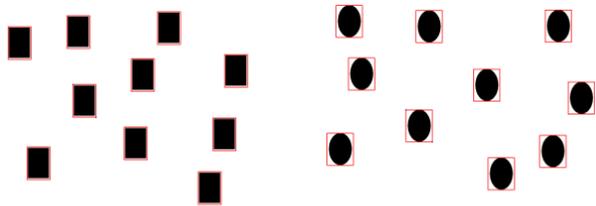


图7 结果图像1(降采样次数:0)

Fig.7 Result image 1(subsample times: 0)

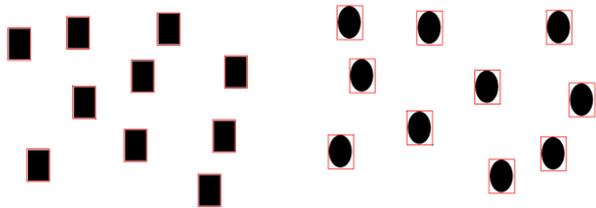


图8 结果图像2(降采样次数:1)

Fig.8 Result image 2(subsample times: 1)

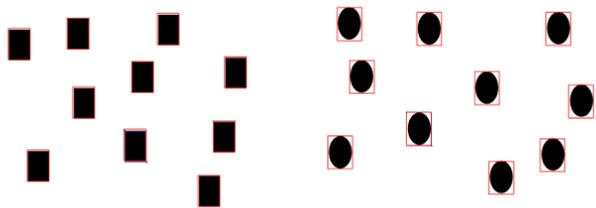


图9 结果图像3(降采样次数:2)

Fig.9 Result image 3(subsample times: 2)

表1 仿真实验结果

Table 1 Result of simulation experiments

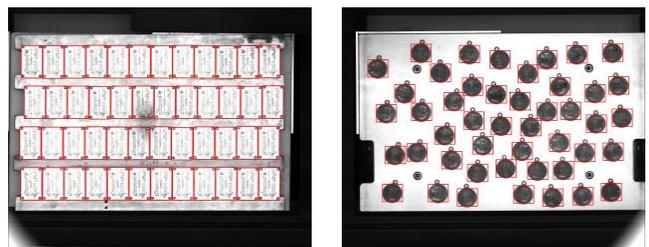
下采样次数	形状	平均运行时间(ms)	识别率
0	矩形	385.365	100%
	椭圆形	363.520	100%
1	矩形	112.333	100%
	椭圆形	117.055	100%
2	矩形	42.040	100%
	椭圆形	64.103	100%

通过仿真实验的结果可以看出, 本文所提出的基于单模板多目标的识别定位方法, 识别率很高, 没有错匹配情况, 符合准确率要求。当下采样次数为 2 时, 识别两种形状的算法运行时间均不超过 100ms, 完全符合工程上实时性的要求。

2.2 实物实验

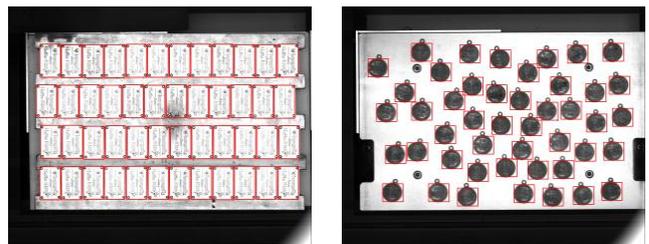
为了满足实际工程需求, 在仿真实验的基础上, 我们又对不同形状的工件进行了实物实验。本实验要求识别过程中不能出现误匹配, 算法运行时间不能超过 200ms。在工作平台上随机放置着相同工件, 工作台正上方的单目相机拍摄出的图像分辨率为 1280*1024, 模板图像是在源图像上用鼠标随机框选所得。实验所需参数及软件平台与仿真实验保持一致。

首先利用本文所提出的算法对矩形及圆形工件进行多目标的识别与定位, 不同降采样次数的实验结果如图 10, 算法识别率与运行时间如表 2 所示。



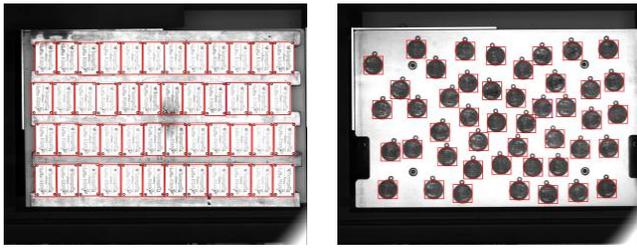
(a)降采样次数:0

(a)subsample times: 0



(b)降采样次数:1

(b)subsample times: 1



(c)降采样次数:2

(c)subsample times: 2

图10 矩形圆形工件识别定位结果

Fig.10 Result image of recognition and location of rectangle and circle

表 2 实物实验结果

Table 2 Result of real experiments

下采样次数	形状	平均运行时间 (ms)	识别率
0	矩形	333.413	100%
	圆形	346.296	100%
1	矩形	102.332	100%
	圆形	138.587	100%
2	矩形	46.021	100%
	圆形	49.134	100%

通过实物实验结果可以看出，当降采样次数大于等于 1 次时，程序平均运行时间都在 200ms 以内，符合工程任务实时性需求。识别率均为 100%，可以说明该方法识别定位准确率较高，符合实验要求。但是，从结果图上同样可以看到，当降采样次数不断增加时，工件的定位精度会逐渐降低。当源图像和模板图像分辨率很低时，该方法尽管速度很快，但可能会出现漏匹配或错匹配的情况。以矩形工件源图像为例，当降采样次数为 3 时，工件识别定位的结果如图 11 所示。



图11 结果图像（降采样次数:3）

Fig.11 Result image (subsample times: 3)

尽管程序运行时间仅仅为 26.372ms，但是从结果图像中可以看出，定位精度很差，并且出现了漏匹配的情况。所以，将该方法运用到实际工程中的时候，降采样次数不宜过多，需要平衡速度与准确度的得失。

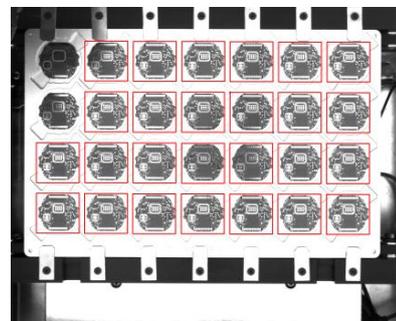
另外，在实验过程中可以发现，在模板的选取上要避免选取遮挡、磨损等情况较为严重的不完整模板，尽量选取清晰完整的模板，否则就会出现图 12 这种出现识别与定位错误的情况，严重影响工程任务的开展。



图12 磨损模板识别定位结果

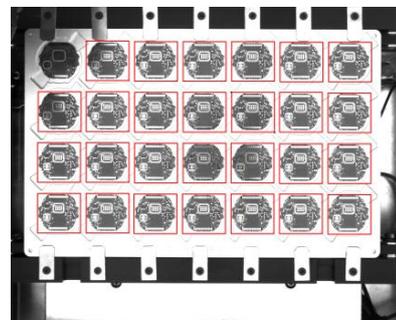
Fig.12 Result image of recognition and location by wear template

最后，为了验证算法对形状的普适性，本文选取了一种不规则形状的工件进行相同的实验，不同降采样次数的实验结果如图 13 所示。



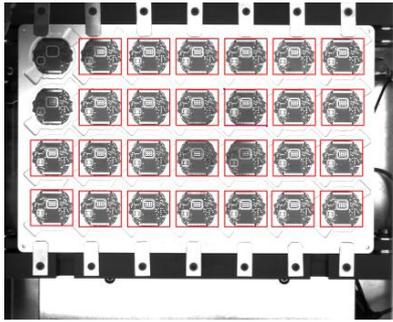
a)降采样次数:0

(a)subsample times: 0



b)降采样次数:1

(b)subsample times: 1



c)降采样次数:2

(c)subsample times: 2

图13 不规则形状工件识别定位结果

Fig.13 Result image of recognition and location of irregular shape object

从结果图中可以看出,识别率不为 100%,存在漏匹配的情况,这主要是因为源图像中左上角的一部分与其他部分存在亮度差异,模板与这一部分匹配时响应值较低,没有达到响应值的阈值要求,出现漏匹配的情况。当下采样次数大于等于 1 时,运行时间小于 200ms,符合实际工程实时性需求,识别准确率很高,无错匹配,符合准确性要求。

3 结 论

本文以机械臂实现平面工件的精确拾取为背景,提出了一种基于归一化互相关匹配算法与非极大值抑制相结合的平面多目标识别与定位方法该方法实时性强,准确率高,可以满足实际工程需要。从实验结果来看,还有以下几点可以进行完善与改进:(1)该方法对噪声及线性灰度畸变不敏感,但是对非线性灰度畸变敏感度较高,当模板图像与源图像目标所在部分的亮度相差较大时,目标识别会出现遗漏的情况,从而降低准确率;(2)本文提出的算法所涉及到的各阶段设置的阈值,均为根据实际经验设定,没有自适应调节功能。算法针对不同场合时的参数需要根据实际情况设定;(3)该方法对平面场景内多目标的定位精度为像素级,随着下采样次数的增加而降低,通过实验可知,下采样次数不能超过 3 次;(4)该方法对工件的形状及纹理具有鲁棒性,但对一定程度的遮挡和磨损敏感,需要选取的模板尽量清晰完整,才可以得到正确的识别结果;(5)该方法不具有方向性,当源图像中目标轴线与模板轴线相交超过一定的角度时,无法对目标进行识别与定位,这也是模板匹配算法的弊端所在。

参考文献(References)

- [1]Tang Y, Wu Q X, Zhu F. Research on the recognition and location method of plane workpiece [J]. Mechanical design and manufacture, 2015 (10): 172-175.[唐宇, 吴清潇, 朱枫. 平面工件的识别与定位方法研究[J]. 机械设计与制造, 2015(10):172-175.]
- [2]Wang X Y, Cheng T T. Research on target recognition technology of industrial robot based on monocular vision [J]. Mechanical design and manufacturing, 2011(4): 155-157.[王修岩, 程婷婷. 基于单目视觉的工业机器人目标识别技术研究[J]. 机械设计与制造, 2011(4):155-157.]
- [3]Liu H W. Target recognition and localization of mobile robot based on monocular vision [D]. Shandong University, 2011: 986-991.[刘洪伟. 基于单目视觉的移动机器人目标识别与定位[D]. 山东大学, 2011.]
- [4]Wang H T, Fu W P, Kang Y N. Workpiece image recognition edge matching method [J]. Instrumentation journal, 2008, 29(5): 986-991.[王红涛, 傅卫平, 康业娜. 工件图像识别的边缘匹配方法研究[J]. 仪器仪表学报, 2008, 29(5):986-991.]
- [5]Chen S W, Yang X G Zhang S X, et al. Target selection and recognition algorithm for relative localization in complex ground background [J]. Laser and infrared, 2015, 45(11): 1375-1379.[陈世伟, 杨小冈, 张胜修,等. 复杂地面背景下相对定位目标选择与识别算法[J]. 激光与红外, 2015, 45(11):1375-1379.]
- [6]Yata K, Koutaki G Uchimura K. Image matching by eigen template method for multi-class classification[C]. The Workshop on Frontiers of Computer Vision. IEEE, 2015:1-4.
- [7]Chayeb A, Ouadah N, Tobal Z, et al. HOG based multi-object detection for urban navigation[C]. IEEE International Conference on Intelligent Transportation Systems. IEEE, 2014:2962-2967.
- [8]Scandaroli G G, Meilland M, Richa R. Improving NCC-based direct visual tracking[C]. European Conference on Computer Vision. 2012:442-455.
- [9]Wang H L, Wang J, Mao S, et al. Sub pixel center of circular target based on non-maximum suppression [J]. Journal of instrumentation and instrumentation, 2012, 33(7): 1460-1468.[王静, 王海亮, 向茂生,等. 基于非极大值抑制的圆目标亚像素中心定位[J]. 仪器仪表学报, 2012, 33(7):1460-1468.]
- [10]Zhang Q, Zhang C B, Chen Z H. A simplified non maximum suppression of constraint conditions [J]. Journal

of University of Science & Technology China, 2016(1).[张强, 张陈斌, 陈宗海. 一种改进约束条件的简化非极大值抑制[J]. 中国科学技术大学学报, 2016(1).]

[11](美)冈萨雷斯(Gonzalez, R. C.), (美)伍兹(Woods,等). 数字图像处理(第3版)[M]. 电子工业出版社, 2010.

[12]Nixon M, Aguado A S. Feature Extraction & Image Processing for Computer Vision[M]. Feature extraction & image processing for computer vision, Publishing House of Electronics Industry, 2013.

[13]Dong J. Rapid and reliable matching technology of template image [D]. National University of Defense Technology, 2014[董晶. 模板图像快速可靠匹配技术研究[D]. 国防科学技术大学, 2014.]

[14] Chen J H, Ye X N. Improvement of the non-maximum suppression algorithm for pedestrian detection [J]. Journal of East China University of Science and Technology: Natural Science Edition, 2015, 41(3): 371-378.[陈金辉, 叶西宁. 行人检测中非极大值抑制算法的改进[J]. 华东理工大学学报:自然科学版, 2015, 41(3):371-378.]