

一种基于 Faster-RCNN 的工业产品清点算法^{* 1}

朱泽民¹, 俞芳芳¹, 董蓉², 李勃¹

(1 南京大学电子科学与工程学院, 南京 210023; 2 南通大学电子信息学院, 南通 226019)

摘要: 工业产品在出厂装箱的时候可能会出现漏装或混杂等异常情况, 为了确保出厂质量, 需要进行产品的清点。针对人工清点费时费力的问题, 提出基于深度学习 Faster-RCNN 网络构造自动清点系统。针对遮挡情况下容易漏检的问题, 创新地提出在训练阶段把产品特征碎片化, 在检测阶段把碎片化的候选框的质心分别进行水平投影和垂直投影, 并通过聚类确定产品的大致质心, 最后将碎片化的候选框融合从而实现清点。实验证明, 所提算法在提高检测效率的同时保证了非常高的可靠性和兼容性。

关键词: 产品清点; 深度学习; Faster-RCNN; 聚类; 碎片化

中图分类号: TP391 文献标志码: B 文章编号: 1671-3133(2018)07-0124-05

DOI: 10.16731/j.cnki.1671-3133.2018.07.021

Industrial product counting system based on Faster-RCNN

Zhu Zemin¹, Yu Fangfang¹, Dong Rong², Li Bo¹

(1 School of Electronic Science and Engineering, Nanjing University, Nanjing 210023, China;

2 School of Electronic Information, Nantong University, Nantong 226019, Jiangsu, China)

Abstract: There may be missing workpiece and other defects when circular workpieces are packed. To ensure the quality, need to check the number of workpieces in the box. A new auto counting system based on Faster-RCNN is proposed to save labor cost. To handle the problem that the workpiece may be undetected when be sheltered by something, a newly fragmentation training method is proposed. After that, it merges the sub-proposal by horizontal projection and vertical projection and clustering the centroid of the sub-proposals when testing. Realistic experiments prove that the method improves detecting efficiency as well as ensuring very high reliability and compatibility.

Keywords: workpiece counting; deep learning; Faster-RCNN; clustering; fragmentation

0 引言

产品在打包装箱的过程中有可能会漏装缺失等异常情况。如图 1 所示的一种零件类工业产品的装箱异常情况, 十分有必要对装箱的结果进行检测。人工检测每条产线需要一个质检员, 全天候地对流水线上经过的装箱进行检查, 容易视觉疲劳, 质量得不到有效的保证, 而且需要投入巨大的人力物力。传统的基于图像特征的机器视觉技术可以从流水线上方拍摄装箱内的图片, 然后基于图像的轮廓特征(见图 2)进行判断是否有缺陷^[1], 这在一定程度上减轻了人的工作量。然而由于现场环境复杂多变, 光照环境的变化、产品上的遮挡物、产品拍摄不完整及产品相连

等(见图 3)都会极大地影响这种方法的准确率, 造成过多的漏判和误判, 仍然需要人工进行二次检测。

随着深度学习理论不断发展, 其在实际工业生产中的应用也越来越可靠。从机器学习的角度考虑, 环形产品清点问题, 可视为一个目标检测问题。在深度学习领域, 文献[3]曾提出了基于 Caffe^[2] 框架的深度学习网络 Faster-RCNN 用于目标检测, 该网络极大地提高了目标检测算法的鲁棒性, 对光照变化具有较强的适应能力, 检测准确率得到了很大的提高, 该方法还具有参数少、更新方便的优点, 遇到新的情况重新训练, 替换训练模型即可。近年来也有相关学者对 Faster-RCNN 进一步改进^[4-9] 以应用于不同的场景。但是对于存在遮挡的情况, 由于目标不完整, 相应的

* 国家自然科学基金项目(61401239)

特征缺失,经过前向传输后得分偏低,该网络容易漏检。

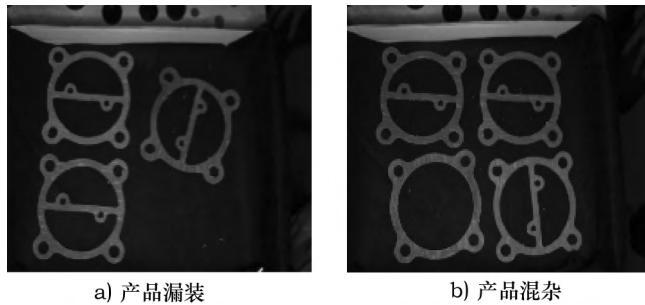


图1 装箱缺陷

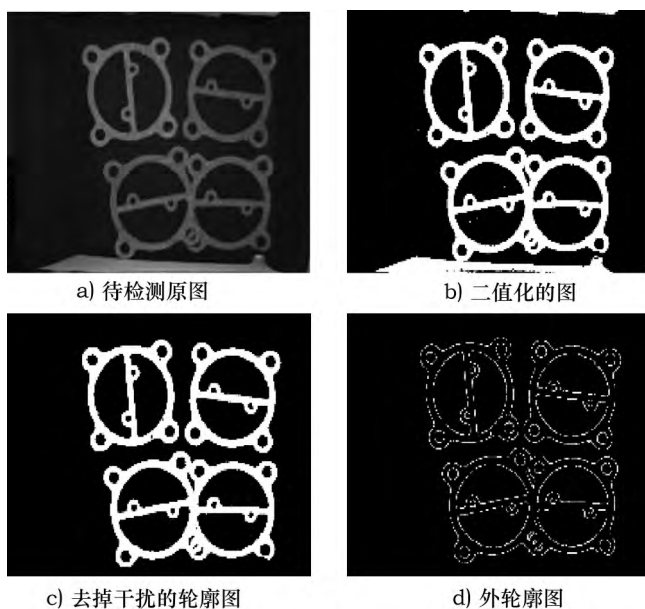


图2 基于轮廓的产品检测

本文针对 Faster-RCNN 对于遮挡目标检测的不足,提出一种新的碎片化训练方法,在检测结果中再将碎片化的结果进行融合,达到对遮挡目标较好的检测效果。

1 基于 Faster-RCNN 的清点系统

1.1 Faster-RCNN 目标检测原理

Faster-RCNN 是一种基于 CNN^[8] 的目标检测网络。作为对 Fast-RCNN^[9] 的改进, Faster-RCNN 在前者的基础上增加了区域生成网络^[2], 在训练时共享了 Fast-RCNN 部分的网络结构从而避免了大量用于提取候选框的重复计算。相比于 Fast-RCNN 的 3 fps, Faster-RCNN 目标检测能够达到 17 fps, 完全满足在线的产品清点系统对实时性的要求。Faster-RCNN 网络结构如图 4 所示。

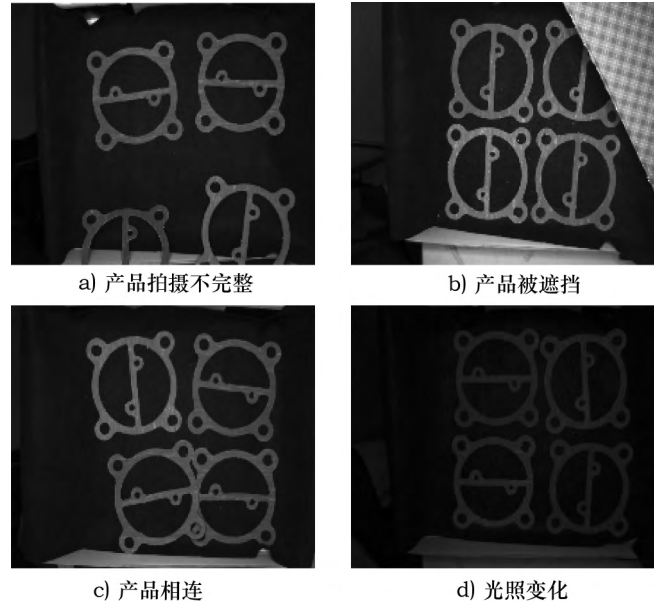


图3 各种装箱情况

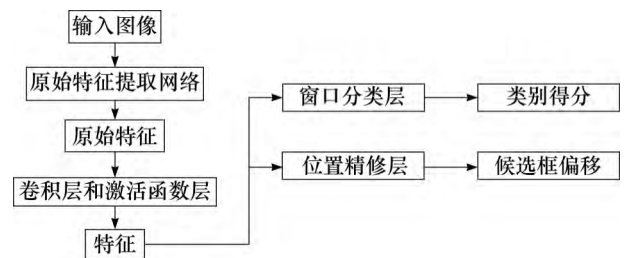


图4 Faster-RCNN 网络结构

从图 4 所示可以看出, Faster-RCNN 在网络中增加了位置精修层, 从而可以在网络中提取出待处理的候选框。由于位置精修层和窗口分类层共享了前面的原始特征提取网络、卷积层和激活函数, Faster-RCNN 减少了 Fast-RCNN 中用于提取候选框的重复计算。

1.2 Faster-RCNN 检测的问题产品

针对产品检测这种场景, 将图片的候选框分为两类, 即产品和背景, 并且选取网络层数较少的 ZF^[12] 模型进行训练。然后对迭代 4 万次的模型进行测试并对每个候选框设置 0.8 的得分阈值。对图 3 中的原图检测效果如图 5 所示, 图 5a、图 5b、图 5c、图 5d 分别为对应图 3 的检测结果。Faster-RCNN 有效地解决了产品相连等因素对检测精度的影响, 但是图 5a 中左下角的产品由于拍摄不完整造成漏检, 图 5b 中右上角的产品由于被遮挡造成漏检, 图 5d 中所有产品由于光照整体偏暗造成漏检。拍摄不完整和被遮挡可以归结为产品不完整, 在网络中的窗口分类层得分不高。如果降低得分阈值虽然能检测一部分不完整和不明显的产品, 但会将一些无关的候选框误判为产品。

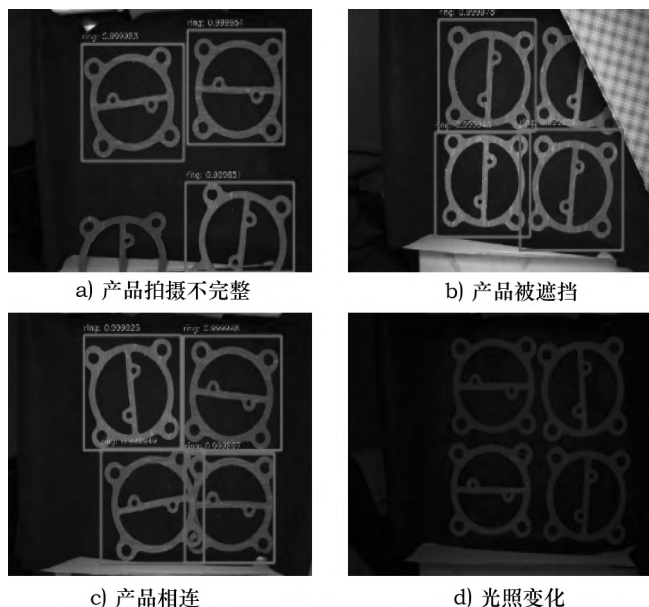


图 5 Faster-RCNN 检测效果

$$(x_1, y_1), (x_2, y_2) \Rightarrow \left[\begin{array}{cc} (x_1, y_1), \left(\frac{x_1+x_2}{2}, \frac{y_1+y_2}{2}\right) & \left(\frac{x_1+x_2}{2}, y_1\right), \left(\frac{x_1+x_2}{2}, \frac{y_1+y_2}{2}\right) \\ \left(x_1, \frac{y_1+y_2}{2}\right), \left(\frac{x_1+x_2}{2}, \frac{y_1+y_2}{2}\right) & \left(\frac{x_1+x_2}{2}, \frac{y_1+y_2}{2}\right), (x_2, y_2) \end{array} \right] \quad (1)$$

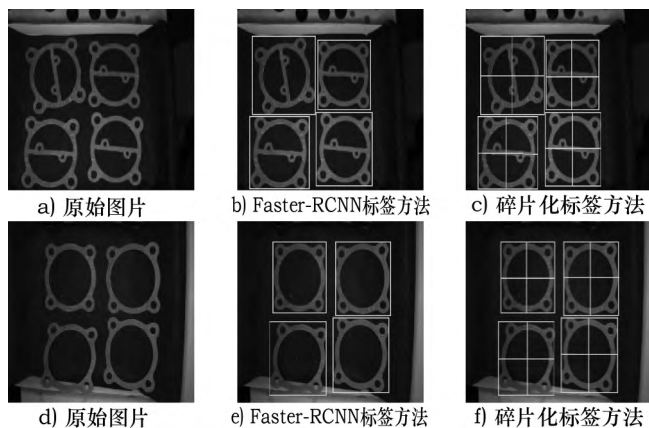


图 6 不同标签方法的对比

这样标签的好处是训练的时候每个子候选框都代表了产品的一个细节,这样虽然整个产品因为不完整或不明显而得分较低,但是在局部,却有一些细节相对完整和明显,只要能够检测出这些完整和明显的部分,就能知道相应的位置确实存在产品,从而解决不完整和不明显产品的检测问题。

2.2 检测结果融合

因训练的时候采用碎片化训练,检测结果也是碎片化的,一个产品将由几个子候选框组成,碎片化训练的检测效果如图 7 所示,图 7a ~ 图 7d 分别为对应图 3 原始图片的检测结果。

2 改进的算法

2.1 碎片化训练

针对 Faster-RCNN 检测产品的上述不足,本文提出一种全新的碎片化训练方法。在准备训练集时,将原来人工标注的候选框分成若干个代表目标局部特征的子候选框。本文将产品平均分成 N 个子候选框, N 为碎片个数。

以 $N=4$ 为例,式(1)中 (x_1, y_1) 为原候选框左上角坐标, (x_2, y_2) 为原候选框右下角坐标。

图 6 所示为 Faster-RCNN 与碎片化训练的 Faster-RCNN 训练集标签方法对比。

图 6a 与图 6d 为原始图片;图 6b 与图 6e 为对应的 Faster-RCNN 标签方法;图 6c 与图 6f 为对应的碎片化标签方法。

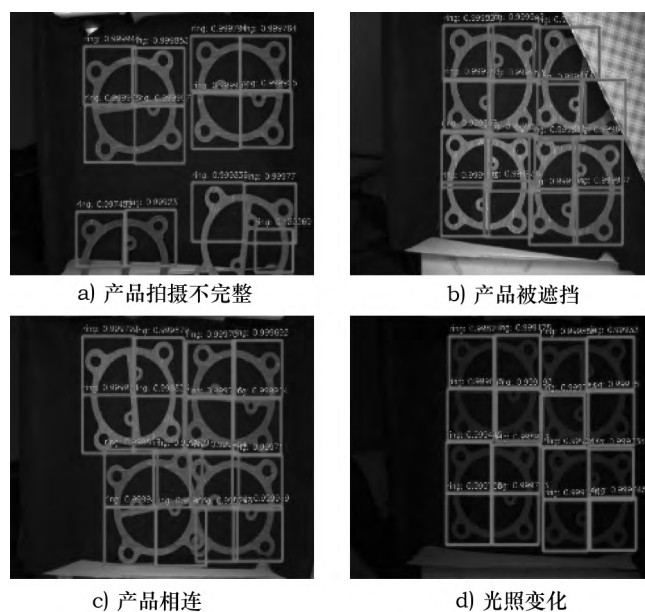


图 7 碎片化训练的检测效果

由于需要检测出产品的个数,有必要对属于同一个大候选框的子候选框进行融合,本文采用 Kmeans 算法^[13-14]对候选框质心的水平投影和垂直投影进行聚类。

当采用欧氏距离计算聚类中心时,算法终止条件的目标函数为最小化对象到其簇质心的距离的平方

和 $\min \sum_{i=1}^K \sum_{x \in C_i} \text{dist}(c_i, x)^2$, 其中, K 为待聚类点的个数; c_i 为待聚类点; C_i 为待聚类点的点集。

聚类算法原理如图 8 所示。

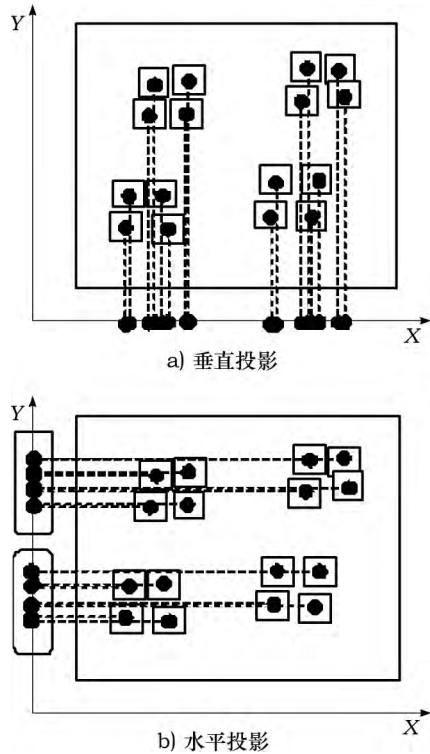


图 8 通过聚类计算产品的质心

通过一次水平投影可以计算出两个 X 坐标的聚类中心 (x_1, x_2) , 再通过一次垂直投影可以计算出两个 Y 坐标的聚类中心 (y_1, y_2) 。由此可以通过排列组合得到 4 个产品大致的中心点 (x_1, y_1) 、 (x_1, y_2) 、 (x_2, y_1) 、 (x_2, y_2) 。通过计算子候选框和所有中心点的最小欧式距离 L 可以将同一个产品上的子候选框进行融合。最小欧式距离如式 (2):

$$L = \min_{i=1,2,3,4} \sqrt{(x - x_i)^2 + (y - y_i)^2} \quad (2)$$

式中: L 为最小欧式距离; (x, y) 为子候选框的质心; (x_i, y_i) 为聚类中心。

图 3 中原图的检测效果如图 9 所示。

从图 9 所示可以看出, 碎片化训练并对检测结果进行融合能够很好地检测出不完整的产品, 对于不明显的产品也能检测出一部分从而避免漏检。

3 实验结果与分析

本文选取两种工业装箱产品进行目标检测(见图 6), 这两种产品装箱时均存在不同程度的遮挡、粘连

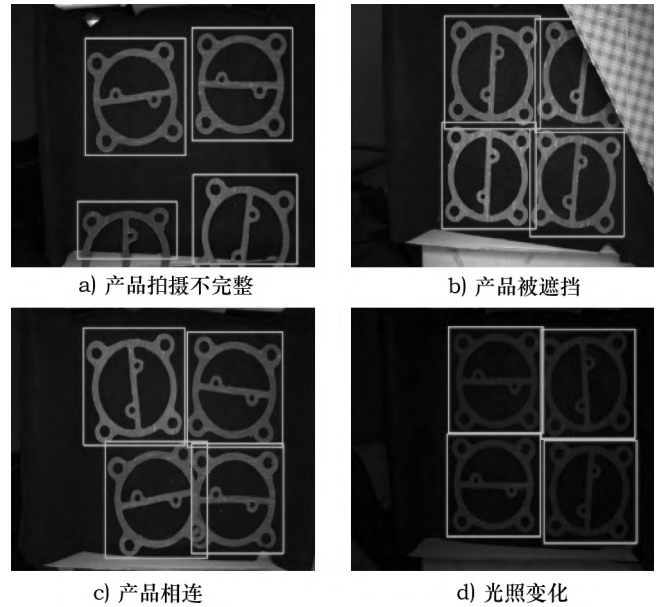


图 9 融合后的检测结果

及光照变化等复杂情况。

采用传统的基于图像轮廓的方法和 Faster-RCNN 以及改进的碎片化训练的 Faster-RCNN 进行检测时间、正确率的测试。

传统方法运行时间在 250 ~ 350 ms, Faster-RCNN 和改进的 Faster-RCNN 时间都在 50 ms 左右, 三种方法运行时间如图 10 所示。

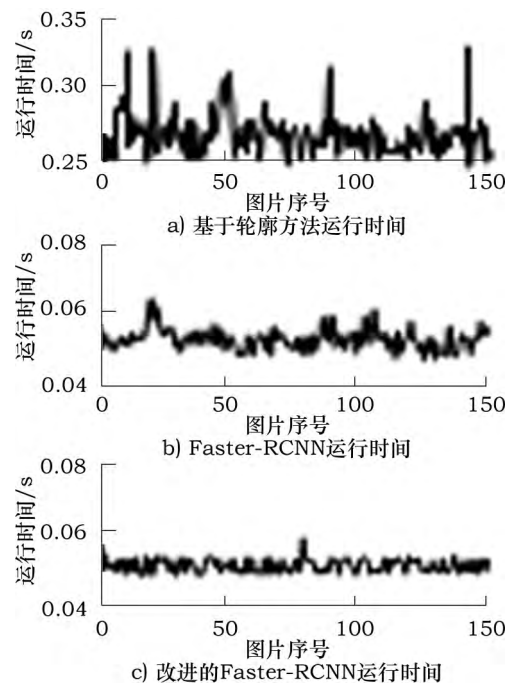


图 10 三种方法运行时间

在检测时间上 Faster-RCNN 和改进的 Faster-RCNN 由于采用基于学习的 RPN 网络提取产品的候选

框,而且采用了 GPU 加速^[15]。较传统方法有明显的提升。Faster-RCNN 的正确率相比较于传统方法有一定的提升。特别是对于产品和背景对比度不太高的场景效果比较显著。

而改进的 Faster-RCNN 采用碎片化训练的方法由于能够检测到产品的所有局部特征,有效避免了原有 Faster-RCNN 的漏检,能达到接近 100% 的正确率,三种方法检测正确率对比如表 1 所示。

表 1 三种方法检测正确率对比

算法	工业产品 1			工业产品 2		
	数量	正确数	正确率/%	数量	正确数	正确率/%
基于轮廓	5 293	5 242	99.04	5 129	5 021	97.89
Faster-RCNN	5 293	5 275	99.66	5 129	5 105	99.53
改进的 Faster-RCNN	5 293	5 293	100.00	5 129	5 128	99.98

4 结语

本文在 Faster-RCNN 的基础上提出了一种全新的目标检测算法。通过在训练阶段将训练样本碎片化和在检测阶段将子候选框融合,有效地解决了产品被遮挡而无法检测的问题。

与传统的基于轮廓方法相比,本文算法显著地减少了检测时间,提高了检测精度,并且对遮挡、产品相连等特殊情况下具有很好的兼容性,具有较高的使用价值。

参考文献:

- [1] 李华伟,李凤婷,余天洪,等.一种用于缺陷检测的瓶盖图像外轮廓确定方法:CN103606167A[P].2014.
- [2] JIA Y ,SHELHAMER E ,DONAHUE J ,et al. Caffe: Convolutional Architecture for Fast Feature Embedding. ACM International Conference on Multimedia [C]. [S. l.]: ACM ,2014: 675-678.
- [3] REN S ,HE K ,GIRSHICK R ,et al. Faster R-CNN: Towards real-time object detection with region proposal networks. Advances in neural information processing systems [C]. [S. l.]: [s. n.] 2015:91-99.
- [4] SALVADOR A ,GIRO-I-Nieto X ,MARQUES F ,et al. Faster R-CNN Features for Instance Search. Deepvision Workshop at CVPR [C]. [S. l.]: [s. n.] 2016.
- [5] JIANG H ,LEARNED-MILLER E. Face detection with the faster R-CNN. Automatic Face & Gesture Recognition (FG 2017) .2017 12th IEEE International Conference on [C]. [S. l.]: IEEE 2017: 650-657.
- [6] HOANG Ngan Le T ,ZHENG Y ,ZHU C ,et al. Multiple Scale

Faster-RCNN Approach to Driver's Cell-Phone Usage and Hands on Steering Wheel Detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops [C]. [S. l.]: [s. n.] 2016:46-53.

- [7] LI J ,LIANG X ,SHEN S M ,et al. Scale-aware Fast R-CNN for Pedestrian Detection [J]. IEEE Transactions on Multimedia ,2015(99) :1.
- [8] CAI Z ,FAN Q ,FERIS R S ,et al. A unified multi-scale deep convolutional neural network for fast object detection. European Conference on Computer Vision. Springer International Publishing [C]. [S. l.]: [s. n.] 2016: 354-370.
- [9] ZENG X ,OUYANG W ,YANG B ,et al. Gated bi-directional cnn for object detection. European Conference on Computer Vision. Springer International Publishing [C]. [S. l.]: [s. n.] 2016: 354-369.
- [10] CHUA L O. CNN: A paradigm for complexity [M]. [S. l.]: World Scientific ,1998.
- [11] GIRSHICK R. Fast R-CNN. IEEE International Conference on Computer Vision [C]. [S. l.]: IEEE ,2015: 1440-1448.
- [12] ZEILER M D ,FERGUS R. Visualizing and understanding convolutional networks. European Conference on Computer Vision [C]. [S. l.]: Springer International Publishing ,2014: 818-833.
- [13] 吴再龙,张云泉,徐建良,等.基于 OpenCL 的 Kmeans 算法的优化研究[J]. 计算机科学与探索,2014,8(10) : 1162-1176.
- [14] WAGSTAFF K ,CARDIE C ,ROGERS S. Constrained K-means Clustering with Background Knowledge. Eighteenth International Conference on Machine Learning [C]. [S. l.]: Morgan Kaufmann Publishers Inc. ,2001: 577-584.
- [15] OWENS J D ,HOUSTON M ,LUEBKE D ,et al. GPU computing [J]. Proceedings of the IEEE ,2008,96(5) : 879-899.

作者简介:朱泽民,硕士研究生,主要研究方向为机器视觉、图像处理。
李勃,通信作者,博士,副教授,主要研究方向为视频图像处理、机器学习及模式识别。

E-mail: liboee@nju.edu.cn

收稿日期:2017-03-27