

图文信息技术

一种基于差影法及 SVM 的在线纸病检测分类方法

曲蕴慧^{1,2}, 汤伟¹, 冯波¹

(1. 陕西科技大学 电气与信息工程学院, 西安 710021; 2. 西安医学院计算机教研室, 西安 710021)

摘要: **目的** 解决目前纸病分类算法存在的实时性差、难以适应生产线在线检测要求等问题。**方法** 提出一种基于差影法和支撑向量机的在线纸病检测分类方法。首先使用差影法来判断纸张是否含有纸病; 对含有纸病的纸张进行打标机打标, 同时存储图像, 提取纸病区域外接矩形的特征向量; 最后使用支撑向量机对纸病进行分类。**结果** 将该方法与已有的 BP 神经网络以及朴素贝叶斯方法进行对比可知, 分类正确率高于目前已有的分类方法, 对于 4 种纸病的分类正确率均在 90% 以上, 而且实时性好, 更加适合于在线检测。**结论** 该方法可以有效地对纸病进行分类, 满足生产线实时检测分类的要求。

关键词: 差影法; 纸病分类; 特征向量; 支撑向量机

中图分类号: TS75 **文献标识码:** A **文章编号:** 1001-3563(2018)23-0176-05

DOI: 10.19554/j.cnki.1001-3563.2018.23.030

On-line Detection and Classification Method Based on
Background Subtraction and SVMQU Yun-hui^{1,2}, TANG Wei¹, FENG Bo¹(1. College of Electric and Information Engineering, Shaanxi University of Science & Technology, Xi'an 710021, China;
2. Computer Teaching and Research Section, Xi'an Medical University, Xi'an 710021, China)

ABSTRACT: The work aims to solve the problems of current paper defect classification algorithm, including poor real-time ability and difficulty in adapting to the requirements of on-line detection of the production line. An on-line paper defect classification method based on background subtraction and support vector machine (SVM) was proposed. Firstly, background subtraction method was used to determine whether the paper contained defects. Then, the paper with defects was marked by the marking machine and the images were stored. The eigenvectors of enclosing rectangle in the paper defect area were extracted. Finally, the paper defects were classified by SVM. Based on the comparison of the proposed method and the existing BP neural network as well as the naive Bayesian method, the classification accuracy was higher than that of the existing classification method. The four kinds of paper defects with classification accuracy of over 90% and good real-time ability were more suitable for on-line detection. The proposed method can effectively classify paper defects and meet the requirements of real-time detection and classification of the production line.

KEY WORDS: background subtraction; paper defect classification; eigenvector; SVM

现代造纸工业正向着设备大型化、自动化、高效率、低能耗的方向发展。提高纸张质量是我国造纸工业技术自主研发的重要方向之一^[1]。传统的人工纸病瑕疵识别需工人通过肉眼对纸张进行瑕疵的检测, 然后再人工进行瑕疵分类, 容易受到工人水平和眼睛疲劳等因素的干扰, 效率及准确率都很低, 远远赶不

上当前高速纸机的需求, 因此纸张质量检测已经逐步由人工检测发展成自动化检测。基于机器视觉和人工智能技术的瑕疵检测分类技术节约了人力资源成本, 提高了生产效率, 已经成为了自动化造纸行业的研究热点之一^[2-3]。

目前常用于纸病分类的模式分类方法主要有 BP

收稿日期: 2018-07-09

基金项目: 陕西省教育厅自然专项 (17JK0645); 陕西省科技统筹创新工程计划 (2012KTCQ01-19); 陕西省重点科技创新团队计划 (2014KCT-15)

作者简介: 曲蕴慧 (1983—), 女, 西安医学院副教授, 主要研究方向为图形图像处理与机器学习。

神经网络方法、朴素贝叶斯方法以及主成分分析 (PCA) 方法等^[4-5]。BP 神经网络分类方法需要大量的样本, 而对于纸张瑕疵检测, 由于没有现成的, 大规模的数据集用于训练, 故分类效率低, 准确率不高; 朴素贝叶斯的方法假设属性之间相互独立, 这个假设在实际应用中往往是不成立的, 在属性个数比较多或者属性之间相关性较大时, 分类效果不好; 主成分分析 (PCA) 算法主要用于降低特征向量的维数, 从而提取主要的特征值, 而分类时大多使用的是最小距离分类器, 分类效率较低。

研究针对当前纸病分类算法存在的识别率低, 算法通用性差, 耗时较长, 无法适用于生产线实时检测等问题, 提出一种基于差影法及支持向量机的纸病检测分类方法。首先对生产线上采集到的图像进行预处理, 使用差影法判断其是否含有纸病; 然后对含有纸病的图像提取纸病区域最小外接矩形; 最后对纸病区域的最小外接矩形部分进行特征提取, 分类器分类以得到最终分类结果。实验结果表明研究所提出的算法对于样本量较小的纸病分类正确率可达到 95% 以上, 并且内存占用率小, 计算量小, 耗时短, 基本可达到生产线上实时检测分类的要求。

1 纸病分类的原理及流程

文中所提出的基于差影法及支持向量机的在线纸病检测分类方法的流程见图 1。

研究中, 为了保证纸病的在线检测, 算法首先对

工业相机采集到的纸张图像进行预处理, 使用差影法快速检测纸病, 在纸病检测过程中, 只需检测出纸张是否含有纸病。对含有纸病的纸张, 在生产线上进行打标处理, 同时提取纸病区域的最小外接矩形, 对纸病区域最小外接矩形, 应用支持向量机的方法对纸病进行分类识别, 给出最后的分类结果, 提高了纸病的在线检测效率, 节约内存, 使得最终设计出的系统可以很好的应用于纸张在线生产检测过程。

2 差影法判定纸病

为了满足生产线上对纸病区域实时进行打标处理的要求, 研究选用差影法进行快速纸病判定^[6-7]。即使用采集到的图像序列和一幅无纸病的正常纸张图像做差, 差影前后效果见图 2。然后对做差之后的图像的灰度值进行判定, 若大于给定阈值, 则判定纸张含有纸病, 进行生产线上打标处理, 同时保存此帧图像, 提取纸病区域的外接矩形, 送后续纸病分类程序进行纸病分类处理; 否则判定纸张无纸病。

图 2a 为一帧正常纸张图像。图 2b 为一帧含有孔洞的纸张图像, 将其与图 2a 的图像进行差影, 得到图 2c 图像; 图 2d 为一帧含有脏点的纸张图像, 将其与图 2a 的图像进行差影, 得到图 2e 图像。对于脏点图像, 由于脏点部分的灰度值小于正常纸张的灰度值, 故做差影时, 直接做差将会得到一个负值, 而对于灰度图像, 灰度值的范围为 $[0, 255]$, 所以在实现时, 取差的绝对值作为输出, 解决差后灰度值为负值的情况。

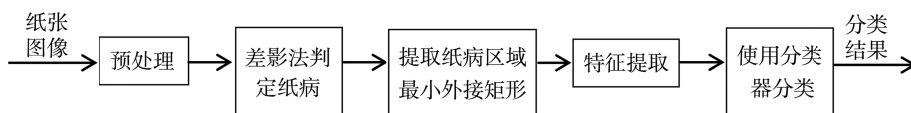


图 1 纸病分类流程

Fig.1 Flow chart of paper defect classification

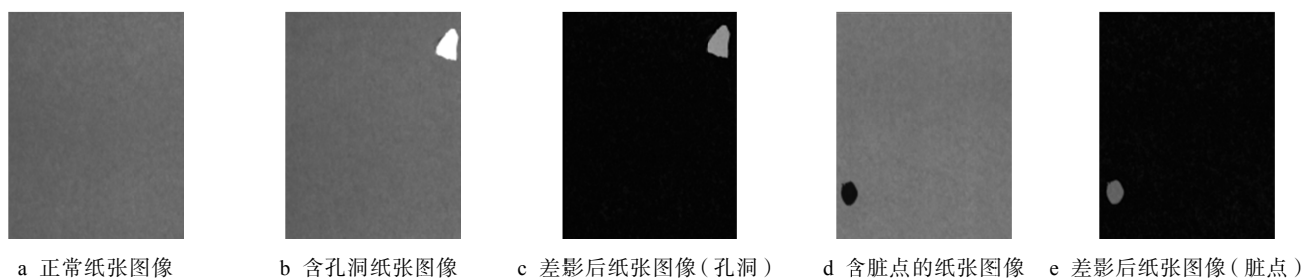


图 2 差影效果

Fig.2 Background subtraction effect

3 特征提取

特征选择的标准^[8-9]如下所述。

1) 特征应该容易提取。对于纸病的在线检测, 特征的提取不能花费太大的时间复杂度, 否则不利于识别的实时性。

2) 选取的特征应对噪声和不相关的转换不敏感。

3) 尽量选取最具区分能力的特征。

基于以上规则, 研究选取了纸病区域的以下特征组成纸病分类器的输入特征向量。

3.1 灰度均值

图像的灰度均值即为图像各像素点的灰度平均

值^[10]。纸病区域的最小外接矩形的灰度均值 M_{can} 见式(1)。

$$M_{\text{can}} = \frac{1}{M \cdot N} \sum f(i, j) \quad (i, j \in S) \quad (1)$$

式中: $f(i, j)$ 为纸病图像 (i, j) 点的灰度值; S 为纸病区域最小外接矩形的区域; M, N 为纸病区域最小外接矩形的长和宽; $M \cdot N$ 则为纸病区域最小外接矩形部分的总像素点的个数。

3.2 灰度标准差

图像的灰度标准差反应图像各像素点灰度值与图像平均灰度值的离散程度。纸病区域的最小外接矩形的灰度标准差 σ^2 见式(2)。

$$\sigma^2 = \frac{1}{M \cdot N} \sum (f(i, j) - M_{\text{can}})^2 \quad (i, j \in S) \quad (2)$$

式中: $f(i, j)$ 以及 S 的含义与式(1)相同; M_{can} 为式(1)计算出来的灰度均值。

3.3 圆形度

圆形度反应了所求区域和圆形的相似度。圆形度的值越大, 表明纸病区域与圆形的相似程度越高。纸病区域的圆形度见式(3)。

$$R = \frac{4\pi A}{L^2} \quad (3)$$

式中: A 为纸病区域最小外接矩形的面积; L 为纸病区域最小外接矩形的周长。对上述 $M \cdot N$ 的纸病外接矩形区域, 即有: $A = M \cdot N$; $L = 2(M + N)$, 带入(3)式可得式(4)。

$$R = \frac{\pi(M \cdot N)}{(M + N)^2} \quad (4)$$

3.4 熵

图像熵描述了图像信源的平均信息量。图像的熵可以表示为图像灰度级集合的比特平均数, 单位(比特/像素), 见式(5)。

$$H = - \sum_{i=1}^M \sum_{j=1}^N f(i, j) \log[f(i, j)] \quad (i, j \in S) \quad (5)$$

4 分类器介绍

对于纸病分类, 目前还没有公认的纸病公共数据库来进行训练以及性能测试。所以目前的研究在图像数据库方面, 所用实验数据只能通过工业相机在实际生产线上采集足够数量的图像, 然后通过人工在所有图像中找出瑕疵, 并对瑕疵进行分类建立数据库来进行研究比对。故在分类器的选择方面, 选择了对训练样本数目要求相对较低的支持向量机(Support Vector Machine, SVM)^[11]。

SVM 是一种训练机器学习的算法, 可以用于解决分类和回归问题。支持向量机方法是建立在统计学

理论的 VC 维理论和结构风险最小原理基础上的。SVM 是以结构化风险最小化为原则, 即兼顾训练误差与测试误差的最小化^[12]。

SVM 本身是一个二值分类器, 当处理多类问题时, 就需要构造合适的多类分类器。目前, 构造 SVM 多类分类器的方法主要分为直接法和间接法 2 类^[13]: 直接法是指直接对目标函数进行修改, 将多个分类面的参数进行求解, 然后合并转化为一个最优化问题。即通过求解最优化问题实现多类分类。这种方法计算复杂度比较高, 实现起来比较困难; 间接法是通过多个二分类器的联合使用来实现多分类器的构造。

研究按照间接法的思想, 构造多个二分类器来完成多分类的任务。研究中需要对 4 类样本进行分类, 见图 3, 4 类样本一共需要构造 6 个二分类的 SVM 来完成分类。对于 4 类纸病: 脏点、孔洞、褶皱和裂痕, 在分类时, 先问分类器“脏点 V 裂痕”(即回答是脏点还是裂痕), 如果是裂痕, 则往左走, 回答“孔洞 V 裂痕”(是孔洞还是裂痕), 以此类推。

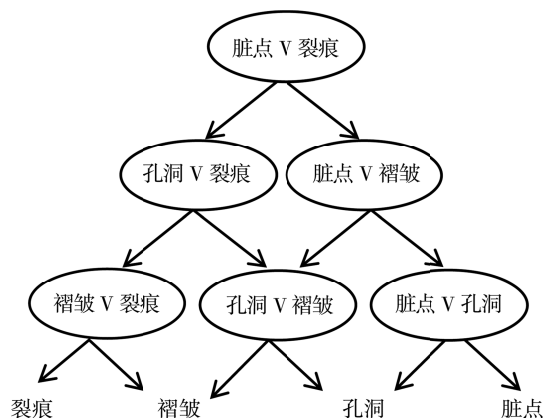


图3 SVM 实现多分类
Fig.3 SVM multi-classification

这种方法在对 4 种纸病进行分类时, 虽然构造了 6 个二分类的分类器, 但实际上进行分类时, 只需调用其中 3 个分类器, 故该方法耗时相对较短, 而且不会出现重叠或不可分类等现象。

在实现时, 对于根节点(即第 1 个分类器)的选取, 为了提高分类效率, 期望根节点少犯错, 即尽可能降低在根节点处分错的几率, 所以在根节点进行第 1 次二分类的 2 个类别, 在编程时选取了特征相差最大的 2 个: 脏点和裂痕。这 2 个节点无论从灰度均值还是圆形度等特征值上, 相差都比较大, 分类器进行分类时, 分错的概率比较低。

5 实验结果及分析

实验过程中, 采用 Visual Studio 2010+Open CV 编程^[14-15]实现研究所提出的算法, 对脏点、孔洞、褶皱和裂痕等 4 类纸病的分类, 并进行了实验室仿真

测试。纸病图像数据库来源于实验室纸病检测设备实际采集到的二维纸病图像序列，每帧图像大小为 4096×1048。实验室纸病检测装置见图 4。实验过程中选取包含脏点、孔洞、褶皱和裂痕的纸病图像共 280 个样本进行实验，其中 200 个样本用于分类器的训练，80 个用于测试。

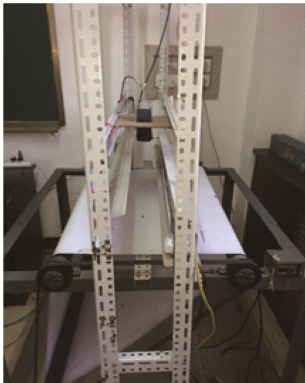


图 4 实验室纸病检测装置
Fig.4 Paper defect detection device for the lab

给出了随机选取的脏点、孔洞、褶皱、裂痕的特征向量见表 1。

表 1 各类纸病特征向量
Tab.1 Eigenvectors of all kinds of paper defects

纸病类型	灰度均值	灰度标准差	圆形成度	熵
脏点	84.6	0.78	0.68	8.17
孔洞	235.5	0.94	0.78	7.71
褶皱	100.5	0.69	0.35	5.35
裂痕	180.6	0.82	0.30	7.45

采用文中所提出的算法对上述纸病样本进行分类，同时编程实现了朴素贝叶斯分类器以及 BP 神经网络分类器，与文中所提出的方法进行了对比，结果见表 2。

由表 2 中数据可得，文中所提出算法对于各类纸病，分类正确率均在 90%以上，整体的正确率可达到 95%，满足工业生产的要求。对 3 种分类算法进行对比，文中算法的正确率明显优于其他 2 种分类方法。

表 2 3 类分类算法效果对比
Tab.2 Comparison of effects of three kinds of classification algorithms

分类方法	脏点（20）		孔洞（20）		褶皱（20）		裂痕（20）		合计（80）	
	正确数目	正确率/%	正确数目	正确率/%	正确数目	正确率/%	正确数目	正确率/%	正确数目	正确率/%
朴素贝叶斯	18	90	18	90	15	75	16	80	67	83.75
BP神经网络	20	100	19	95	17	85	18	90	73	91.25
文中算法	20	100	20	100	18	90	19	95	77	96.25

实验过程中，对 3 种方法分类错误的纸病进行了进一步的分析，其中朴素贝叶斯方法对 80 例数据进行了分类，争取数目为 67 例，其余 13 例数据中，错误 10 例，无法识别 3 例，对于有些特征不是很明显的纸病，如褶皱等，其错误率较高或出现无法识别现象；BP 神经网络分类方法在实验数据量较少的情况下，分类正确率赶不上文中的方法，并且训练速度较慢。

目前纸机速度可达 30 m/s，纸幅宽度约 9 m，传统的在线检测方法需要记录纸张的整体信息，数据庞大复杂。实际造纸过程中，纸病出现概率相对较低，并且纸病区域一般较小，故文中所提方法只用记录纸病区域少量信息来进行进一步处理识别，大大降低了存储容量，提高了处理速度，无论在时间复杂度和空间复杂度上均优于传统方法，满足工业生产在线检测的需求。

6 结语

文中针对当前纸病分类算法耗时较长，无法适用于生产线实时检测的问题，提出一种纸病在线检测分类算法。首先通过差影法判断纸张是否含有纸病，对含有纸病的纸张进行生产线打标机打标，同时存储纸

病图像，提取纸病外接矩形区域的特征向量，使用 SVM 对纸病进行分类。实验结果表明，研究所提出的方法实时性好，内存使用率低，分类效果满足工业生产线在线检测分类的要求。

参考文献：

[1] 李茜, 文豪, 曲蕴慧, 等. 边缘跟踪算法在多纸病图像分割中的应用[J]. 中国造纸, 2017, 36(8): 41—45.
LI Qian, WEN Hao, QU Yun-hui, et al. Application of Edge Tracking Algorithm to Multiple Paper Defects Segmentation[J]. China Pulp & Paper, 2017, 36(8): 41—45.

[2] 曲蕴慧, 汤伟, 文豪, 等. 基于 DCT 同态滤波的纸病图像去噪算法研究[J]. 中国造纸, 2018, 37(5): 45—49.
QU Yun-hui, TANG Wei, WEN Hao, et al. Paper Defects De-noising Algorithm Based on Homomorphic Filtering and Discrete Cosine Transform[J]. China Pulp & Paper, 2018, 37(5): 45—49.

[3] 张学兰. 基于机器视觉的纸病图像分析[D]. 广州: 华南理工大学, 2013.
ZHANG Xue-lan. Analysis on Image of Paper Defects

- Based on Machine Vision[D]. Guangzhou: South China University of Technology, 2013.
- [4] 周强, 陈颖, 沈天宇, 等. 基于机器视觉的纸病检测系统发展综述[J]. 中国造纸, 2016, 35(5): 72.
ZHOU Qiang, CHEN Ying, SHEN Tian-yu, et al. Review of Paper Defect Detection System Based on Machine Vision Technology[J]. China Pulp & Paper, 2016, 35(5): 72.
- [5] 院金彪, 周强, 郑海英, 等. 基于朴素贝叶斯分类器的纸病离线辨识方法研究[J]. 中国造纸学报, 2014, 29(1): 58—62.
YUAN Jin-biao, ZHOU Qiang, ZHENG Hai-ying, et al. Paper Defects Offline Static Identification Based on Naive Bayes Classifier[J]. Transactions of China Pulp and Paper, 2014, 29(1): 58—62.
- [6] CHEN Ben-zhi, FANG Zhi-hong, XIA Yong, et al. Accurate Defect Detection Via Sparsity Reconstruction for Weld Radiographs[J]. NDT and E International, 2018, 94: 62—69.
- [7] LI X, JIANG H, YIN G. Detection of Surface Crack Defects on Ferrite Magnetic Tile[J]. NDTE Int 2014, 62: 6—13.
- [8] BAHRAMPOUR S, NASRABADI N M, RAY A, JENKINS W K. Multimodal Task-driven Dictionary Learning for Image Classification[J]. IEEE Trans Image Process, 2016, 25(1): 24—38.
- [9] 张铮, 徐超, 任淑霞, 等. 数字图像处理与机器视觉——Visual C++与 Matlab 实现[M]. 北京: 人民邮电出版社, 2016.
ZHANG Zheng, XU Chao, REN Shu-xia, et al. Digital Image Processing and Machine Vision: Implemented in Visual C++ and Matlab[M]. Beijing: Posts & Telecom Press, 2016.
- [10] 尼克松, 阿瓜多. 计算机视觉特征提取与图像处理[M]. 北京: 电子工业出版社, 2014.
NIXON M S, AGUADO A S. Feature Extraction & Image Processing for Computer Vision[M]. Beijing: Publishing House of Electronics Industry, 2014.
- [11] 郝淑丽. 基于支持向量机的金属表面细小瑕疵检测算法[D]. 上海: 上海交通大学, 2008.
HAO Shu-li. Flaw Detection in Metallic Surface Based on Support Vector Machine[D]. Shanghai: Shanghai Jiao Tong University, 2008.
- [12] 周强, 齐璐, 张慧. 基于 SVD 和 SVM 的复杂背景噪声图像的纸病辨识[J]. 中国造纸学报, 2016, 31(2): 49—54.
ZHOU Qiang, QI Lu, ZHANG Hui. Recognition of Paper Defect Image with Complex Background Noise Based on SVD and SVM[J]. Transactions of China Pulp and Paper, 2016, 31(2): 49—54.
- [13] 弗拉赫. 机器学习(Machine Learning)[M]. 北京: 人民邮电出版社, 2016.
PETER F. Machine Learning[M]. Beijing: Posts & Telecom Press, 2016.
- [14] 张平. OpenCV 算法精解[M]. 北京: 电子工业出版社, 2017.
ZHANG Ping. OpenCV Algorithm Precision Solution[M]. Beijing: Publishing House of Electronics Industry, 2017.
- [15] 望熙荣, 望熙贵. OpenCV 和 Visual Studio 图像识别应用开发[M]. 北京: 人民邮电出版社, 2017.
WANG Xi-rong, WANG Xi-gui. OpenCV & Visual Studio Application and Development of Image Recognition[M]. Beijing: Posts & Telecom Press, 2017.