基于 Blob 算法的印刷缺陷在线检测的研究

徐 敏，唐万有，马千里，郝健强
（天津科技大学，天津 300222）

摘要：将常见的印刷缺陷分为点、线、面 3 类，介绍了印刷缺陷在线检测的基本流程，说明了 Blob 算法检测缺陷的具体步骤，并运用 Blob 算法结合缺陷面积及周长的特征，成功对印品缺陷进行了提取及归类，最终得到了缺陷的面积、位置、类型等特征参数，在程序中实现了印刷品缺陷的检测。

关键词：印刷缺陷；在线检测；Blob 算法
中图分类号：TS801.3；TS807 **文献标识码：**A **文章编号：**1001-3563(2011)09-0020-04

Research of Printing Defect On-line Detection Based on Blob Algorithm

XU Min，*TANG Wan-you*，*MA Qian-li*，*HAO Jian-qiang*
（Tianjin University of Science & Technology，Tianjin 300222，China）

Abstract：The common printing defects were divided into point，line and area. The basic flow of online defect test and detailed steps of blob algorithm were introduced. Defects in printed matter were identified and classified by blob algorithm combined with area and girth；characteristic parameters of defects such as area，position and style were extracted. Defect detection of printed matter was achieved in program.

Key words：printing defect；on-line detection；Blob algorithm

随着印刷技术的不断提高，目前胶印生产中已经实现了自动化和数字化，印刷速度也越来越高，但是印刷过程中不可避免地会出现油墨污点、纸张起皱、漏印、刮伤、套印不准、颜色失真等缺陷[1]。如何快速、准确地评定印刷质量，使有缺陷的印刷品得到快速剔除，成为了一个重要问题。

印刷品质量的全画面在线检测中[2]，要求能对印刷过程中印刷品出现的各种缺陷进行准确识别和判断，并把得到的信息及时反馈给印刷机及相应的控制系统，对印刷机进行适当调整，以实现对印刷过程的实时控制。笔者在基于数字图像处理的机器视觉检测的基础上，采用 Blob 算法成功检测和显示缺陷位置、面积、类型等特征参数，实现全画面的胶印质量在线检测，满足了快速发展的胶印技术的要求，保证印品总体质量，避免了不必要的浪费。

1 印刷品缺陷分析及检测流程

1.1 常见的印刷品缺陷

印刷质量一般是指印刷品各种外观特性的综合效果，即绝对质量[3]。文中讨论的印刷缺陷则是待检印刷品相对于标准样张的忠实性，即与标准样张有不匹配区域则认为存在缺陷。这样印刷缺陷就可分为形状缺陷和颜色缺陷两大类。形状缺陷主要关注的是缺陷的形状，例如刀丝、划痕等；而颜色缺陷则是待检印刷品与标准样张颜色的一致性，如墨色过浓、过淡等。

根据影响印刷品质量的因素，综合各种缺陷的特征，将缺陷分为 3 类，见表 1。

1.2 印刷缺陷在线检测流程

缺陷检测前，通过 CCD 采集合格的印刷品作为

收稿日期：2011-03-05
基金项目：广东省重大科技专项(2010A080402010)
作者简介：徐敏(1986－)，女，江苏连云港人，天津科技大学硕士生，主攻印刷复制技术及印刷质量控制。
通讯作者：唐万有(1955－)，男，天津人，天津科技大学教授，主要研究方向为印刷质量及质量评价系统、印刷设备及其性能。

表 1　印刷缺陷分类

Tab.1 Printing defect classification

缺陷类型	缺陷特征	实际缺陷
点	面积较小 与样张相比色差大	漏印、飞墨、污点等
线	宽度小 分布区域广	刀丝、划痕、褶皱等
面	面积大 与样张相比色差较小	脏版、糊版、墨色浓淡等

标准样张，并将其存储在计算机中[4]。当印刷品印刷完毕，经过 CCD 扫描，与样张比较，找出不同之处即缺陷，并用一定的算法对缺陷进行分析，达到缺陷检测的目的。检测的简单流程见图 1。

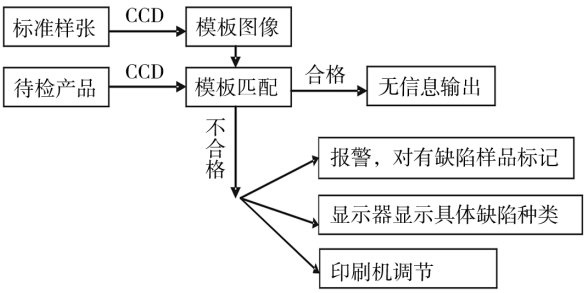

图 1　在线检测系统工作流程

Fig. 1 On-line detection system diagram

2　检测系统的关键技术

缺陷检测是在待检样张和标准样张匹配完成的基础上进行的，因此如何快速准确判断缺陷是否存在以及得到缺陷的具体特征信息，是要解决的主要问题。程序分区比较像素差异，并运用 Blob 算法对有差异区域的参数分析，从而正确区分点、线、面缺陷。

2.1　缺陷寻找

从视觉角度看，不管是形状缺陷还是颜色缺陷[5]，都是待检印品画面与标准样张的不同之处，而最终人眼都是通过颜色的不同进行判断，因此可以通过比较 2 幅图相对应位置的色差来判断缺陷的存在。为了更接近人眼的视觉判断，同时减少颜色的损失，采用 XYZ 空间作为过渡，将图像颜色由 RGB 转换到 Lab 色空间。公式如下[6]：

$$\begin{cases} X=2.768\ 9R+1.751\ 7G+1.130\ 2B \\ Y=1.000\ 0R+5.590\ 7G+0.060\ 1B \\ Z=0.056\ 5G+5.594\ 3B \end{cases}$$

$$\begin{cases} L^*=116(Y/Y_0)^{1/3}-16 \\ a^*=500[(X/X_0)^{1/3}-(Y/Y_0)^{1/3}] \\ b^*=200[(Y/Y_0)^{1/3}-(Z/Z_0)^{1/3}] \end{cases}$$

采用的 CCD 分辨率为 500 ppi，要求检测到的最小缺陷为 0.04 mm^2，而一个像素的大小为 0.002 6 mm^2，即可检测到的最小缺陷面积大概是 9 个像素的大小，因此可以将每 9 个像素看成一个整体，即一个分区，进行判断。这样整个画面就被重新划分成众多边长为 3 个像素的正方形区域。具体分区方式见图 2：以图像的左上角为坐标原点，分别作 x，y 轴，建立

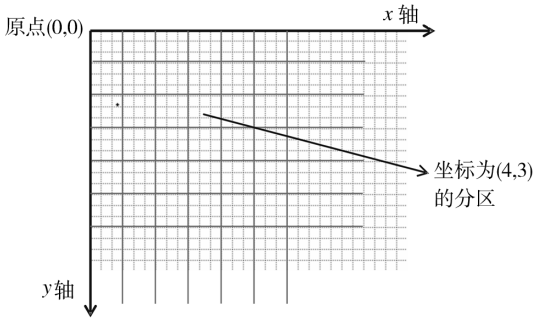

图 2　图像分区

Fig. 2 Image partition

坐标系。按照程序指定的分区尺寸，即 9 个像素，将图像分为若干区域，并分别按照坐标系定义分区坐标。

计算时用每个分区内 9 个像素的 RGB 平均值代替分区的颜色信息，然后将待测图像和标准图像相对应的分区进行对比。通过分区，大大减少了计算量，加快了程序的执行速度；而且由于连续调图像各像素间的关联性及数据的冗余性，这种方法对检测结果并没有太大的影响。

因为颜色差异应允许的范围，要根据印刷品的检测精度和其他一些具体情况而定。例如 GB/T 7705－2008《平版装潢印刷品》中规定，同批同色实地油墨颜色的色差范围：一般产品为 $\Delta E^*_{ab}\leqslant 5.00\sim 6.00$；精细产品为 $\Delta E^*_{ab}\leqslant 4.00\sim 5.00$。实际检测时，可根据检测的具体要求，参照色差程度的鉴定标准，设置合适的色差阈值。为了适应不同的印刷检测要求，在程序编制时，将色差的阈值设定为可变值，从而实现检测系统的人性化。色差程度的鉴定标准见表 2。

设 $S(L,a,b)$，$T(L,a,b)$ 分别为标准图像和待测印品分区像素的 Lab 值，结合检测精度设定色差阈值 ΔE，如果 $\Delta E<\sqrt{(SL-TL)^2+(Sa-Ta)^2+(Sb-Tb)^2}$，

表 2 色差程度鉴别标准

Tab. 2 Identification standard of color difference

NBS 单位色差值	感觉色差程度
0.0～0.5	痕迹
0.5～1.5	轻微
1.5～3.0	可觉察
3.0～6.0	可识别
6.0～12.0	大
12.0 以上	非常大

就认为该分区是色差点。其中 SL 为标准图像像素的 L 值,其它以此类推。

2.2 Blob 算法聚类缺陷

找出待检印品画面中的色差点之后,需要对缺陷区域进行分析处理。采用 Blob 算法准确找出色差点的位置并确定其大小、形状及面积等,从而对缺陷进行归类。

Blob 算法又称为斑点分析法,常用于对目标图像进行图形特征提取和分类[7]。在 VC++的编程环境下,结合数字图像处理技术,实现了 Blob 算法在印刷品缺陷在线检测中的应用[8-9]。Blob 算法实现流程见图 3。

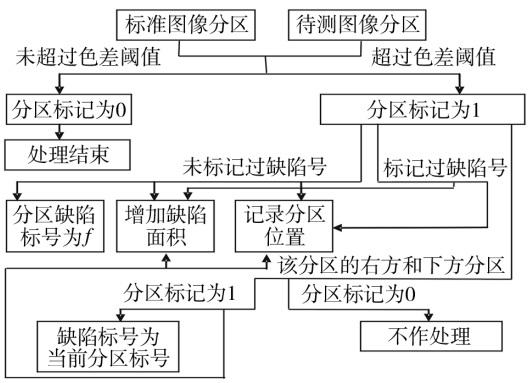

图 3 Blob 算法流程

Fig. 3 Blob algorithm diagram

具体步骤如下:

1) 使用二维数组存储各缺陷的标号和缺陷面积,即色差点的像素个数。

2) 标记有色差的分区。由左到至右、由上而下判断分区色差,如果色差值超过阈值,则将该分区标记为 1,否则为 0。

3) 缺陷标记。重新对图像扫描,如果分区标记为 1,即色差点,且未标记缺陷号,则执行以下步骤:(1)将缺陷的标号记为 f,f 变量由 1 开始递增;(2)该缺陷的面积增加相对应的像素数,即如果分区为 2,则增加 4 个像素;(3)记下分区的位置,即在以像素为单位的图像中的行号和列号。

4) 缺陷合并。如果分区标记为 1,且已经被标记过缺陷号,则以标记过的缺陷号为准,不再增加,同时执行 3)中的步骤(2)和(3),这样就将相邻区域的色差点进行了合并。

5) 缺陷膨胀。理论上可以认为是一个缺陷的不连通区域进行合并,分别检测该分区的右方和下方的 8 个分区的标记是否为 1,如果是,则这些分区的缺陷标号改为该分区缺陷标号,同时执行 3)中的步骤(2)和(3);如果不是则不进行任何处理,直到这 16 个分区检测完毕。

通过以上步骤对图像全画面检测,即可成功完成缺陷的检测。

6) 缺陷位置确定。找出同一缺陷中所有分区的最小、最大的列号、行号,以确定缺陷的外接矩形和缺陷的中心位置。

忽略算法中的复杂情况,将 Blob 算法简单理解:首先将连通的色差点找出,并赋予相应的缺陷标号;然后将相距 8 个像素以内的缺陷合并为一个缺陷,见图 4 和 5。

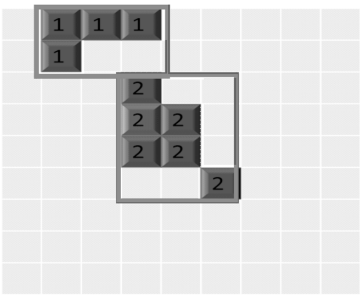

图 4 连通色差点的查找

Fig. 4 Lookup of connected color difference points

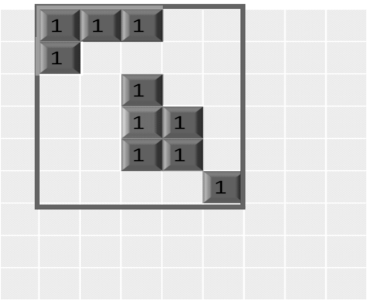

图 5 缺陷的膨胀

Fig. 5 Defect expansion

当图像经过 Blob 算法处理后，得到缺陷的标号、位置、面积及外接矩形的面积，将缺陷的视觉信息转换为形状信息，更方便下一步对缺陷类型的判断。

2.3　缺陷类型判断

通过对实际缺陷的分析、测量，结合缺陷自身特征(见表 1)以及缺陷面积与外接矩形的面积之比，可以成功区分缺陷类型。

2.3.1　点的判断

将边长大小在 3 mm 的正方形内的所有缺陷都归为点缺陷。若采用 CCD 分辨率为 500 ppi，则只要缺陷的外接矩形的长和宽均小于等于(3×500)/25.4，就可判断该缺陷为点。

2.3.2　面与线的判断

近似平行于 x，y 轴的线缺陷，如果缺陷外接矩形的较短边小于等于 2 mm，并且长边与短边之比大于 6，则认为该缺陷为线。

对于其它情况的线，因为线缺陷面积与面缺陷相比要小的多，采用缺陷面积与外接矩形面积之比区分二者。设缺陷面积为 S，相对应外接矩形大小为 W，将 $S/W \geqslant 0.3$ 的缺陷归为面，否则为线。

通过以上程序的设计，最终可以快速正确得到缺陷的各种特征信息，从而实现印刷缺陷的在线检测。分辨率为 500 ppi，ΔE 为 4 的缺陷结果见图 6－8，检测到缺陷后，将缺陷区域用蓝色矩形框进行标识，并在对应的矩形框下方标记缺陷号，以便和详细信息列表框中的内容相对应。图 7 中待检印品中的缺陷面积经检测，面积大小为 327.694 828 mm^2，与外接矩形的面积比大于 0.3，被判断为面形缺陷，符合实际人眼观察到的缺陷形状特征。

图 6　标准印品

Fig. 6 Standard prints

图 7　待检印品

Fig. 7 Prints to be detected

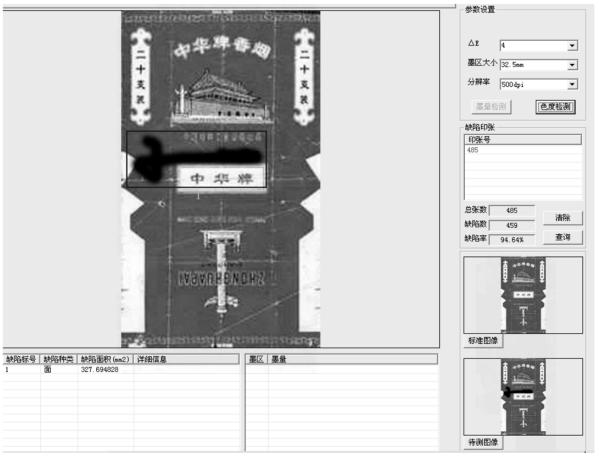

图 8　印品检测结果

Fig. 8 Defect detection results of prints

3　结语

在基于数字图像处理的机器视觉检测基础上，采用 Blob 算法结合缺陷特征，能够准确检测出印刷缺陷，并将其归类为点、线、面 3 种类型。可以通过设定分区数和色差阈值 ΔE，完成对不同精度印刷品的检测。

参考文献：

[1]　华新星. 基于线阵 CCD 的印刷品缺陷在线检测方法研究[D]. 哈尔滨：哈尔滨工业大学，2006.

[2]　陈亚军，张二虎. 基于图像处理的印刷缺陷在线检测系统研究[J]. 包装工程，2005，26(6)：65－66.

(下转第 80 页)

式中：$\Delta X(\lambda)$为测量值与预测值之间的光谱误差；n 为数据采样波段数，波长范围为 400～700 nm，采样点以 10 nm 为间隔，测量每个色块对应的 31 个采样点的光谱反射率。

根据式(9)，原有经典 Kubelka-Munk 模型和新模型相对于实际测量值的光谱均方根误差均可根据图 5 中的数据计算得出，见表 1；而相应的色差亦可计算得到，见表 2。

表 1 不同模型的光谱均方根误差

Tab. 1 The spectral RMS between different models

RMS	C	M	Y
经典 Kubelka-Munk 模型	0.383 1	0.403 9	0.348 2
新模型	0.251 2	0.177 1	0.191 5

表 2 2 种模型预测结果的色差分析

Tab. 2 The color difference between the predictions of the two different models

RMS	ΔE_C	ΔE_M	ΔE_Y
经典 Kubelka-Munk 模型	3.632 5	4.367 1	3.184 0
新模型	2.460 3	1.830 9	2.045 7

可见新模型的预测精度要高于原有模型。这主要是由于在蓝紫波段，新模型预测的反射率要略高于原有模型的预测值，这一部分的增加量主要是由于照射光中的紫外线受纸基内部荧光剂激发所致。

3 结语

以光能的辐射传递理论为基础，通过分析光在印刷品内部传播过程中的散射和吸收规律，并引入荧光激发函数，建立了新的印刷品光谱反射预测模型，修正了原有模型预测印刷品时蓝紫波段反射率偏小的问题，提高了预测精度，具有一定的创新性和实用性。但是，印刷品的呈色机理是极其复杂的，如不同颜色油墨之间的相互渗透以及底层油墨向纸基内部的非线性渗透现象，在新模型中也都没有考虑到，这些有待于进一步的研究。

参考文献：

[1] 杜艳君，张逸新. 经典色彩呈色模型[J]. 包装工程，2006，27(2)：99－101.

[2] 葛惊寰. 印刷品色彩预测模型[D]. 无锡：江南大学，2008.

[3] 梁文懂，肖时钧. 传递现象基础[M]. 北京：冶金工业出版社，2006.

[4] HERSCH R D. Spectral Prediction Model for Color Prints on Paper with Fluorescent Additives[J]. Applied Optics，2008，47(36)：6710－6722.

[5] VARGAS W E，NIKLASSON G A. Generalized Method for Evaluating Scattering Parameters Used in Radiative Transfer Models[J]. J Opt Soc Am，1997，14(9)：2243－2252.

[6] RAVEL B. A Practical Introduction to Multiple Scattering Theory[J]. Journal of Alloys and Compounds，2005，401(2)：118－126.

[7] VIGGIANO J A S. Modeling the Color of Multi-Colored Halftones[J]. Proc TAGA，1990，44(62)：19－26.

[8] MOURAD S，EMMEL P. Extending Kubelka-Munk's Theory with Lateral Light Scattering[J]. Internal Conference on Digital Printing Technologies，USA，2001：469－473.

[9] PACKARD H，DIVISION B. Spectral Modelling of Large-Format Ink-jet Printers，Research and Development Final Report[M]. RIT Press，1996.

(上接第 23 页)

[3] Intelligence Recognition of Reconnaissance Objective Based on INN Pattern Recognition[C]. Proceedings of the 5th International Symposium on Test and Measurement (Volume 1)，2003.(余不详)

[4] ZHANG Luo Z. Automatic Color Printing Inspection by Image Processing[J]. Journal of Materials Technology，2003(139)：373－378.

[5] 于佳，张二虎. 基于全画面的印刷品在线色差检测模型[J]. 包装工程，2009，30(3)：85－87.

[6] 刘武辉，胡更生，王琪. 印刷色彩学[M]. 北京：化学工业出版社，2004.

[7] china-vision. BLOB 算法简述[EB/OL]. (2003-06-12)[2011-03-01]. http://www.china-vision.net/bbs7/2003－9/29/155940867.html.

[8] RAFAEL C，RICHARD E. 数字图像处理[M]. 阮秋琦，阮宇智，译. 北京：电子工业出版社，2008.

[9] 斯特劳斯特鲁普. C＋＋程序设计语言[M]. 北京：高等教育出版社，2009.