# 基于 K-means 聚类十字线的套印偏差检测方法研究

## 吕明珠,吴学毅,成刚虎,岳喜娜

(西安理工大学,西安 710048)

摘要:目的 研究并改进基于十字线套准标识符的印刷品套印偏差检测算法,以提升偏差量检测精度。 方法 使用 K-means 聚类算法对分色后得到的 C, M, Y, K 图像分别进行聚类处理, 然后分割图像并提 取出各色前景像素, 进而使用 C, M, Y 图像中的前景十字线分别减去 K 图像中的十字线, 以消除四色 重叠区域的干扰, 最后提取十字线中心并计算偏差量。结果 使用 15 幅印刷品图像样张进行实验, 改进 的基于 K-means 聚类十字线的套印偏差检测算法的检测结果, 比大津法更接近人工测量值, 误差在±0.04 mm 以内。结论 基于 K-means 聚类的十字线分割法对前景十字线的提取, 较大津法抗叠印干扰性更强。 关键词:K-means 聚类; 印刷品; 十字线; 套印偏差

中图分类号:TS807 文献标识码:A 文章编号:1001-3563(2020)01-0143-06 DOI:10.19554/j.cnki.1001-3563.2020.01.022

#### **Overprint Deviation Detection Method Based on K-means Clustering Cross Line**

LYU Ming-zhu, WU Xue-yi, CHENG Gang-hu, YUE Xi-na (Xi'an University of Technology, Xi'an 710048, China)

**ABSTRACT:** The paper aims to research and improve the printing overprint deviation detection algorithm based on cross line alignment identifier to improve the accuracy of the deviation detection. Firstly, the K-means clustering algorithm was used to cluster the C, M, Y and K images obtained after color separation. Secondly, the fourimage arerespective segmented and the each foreground pixels were extracted. Then the foreground cross lines in C, M and Y images were respectively subtracted from the cross lines in K images to eliminate the interference of four-color overlapping regions. Finally, the center of the cross line was extracted and the deviation detection algorithm based K-means clustering cross line was closer to the manual measurement results than Otsu method, and the error was within ( $\pm 0.04$  mm). The cross-line segmentation method based on K-means clustering is more resistant to overlay interference than Otsu method in foreground cross-line extraction.

**KEY WORDS:** K-means clustering; print; cross line; overprint deviation

近年来随着生活水平的提高,人们对印刷品表观 质量的要求不断提高,基于机器视觉的印刷品表面缺 陷检测方式已是大势所趋<sup>[1]</sup>。

针对基于机器视觉获取的印刷品叠印式标识符 检测过程中存在的颜色重叠区域分割困难问题,有学 者使用大津法(最大类间方差法,简称 Otsu)分别对 分色后的 C,M,Y,K 灰度图像作阈值分割<sup>[2]</sup>,提 取出目标色十字线,然后提取出十字线骨架,并使用 霍夫直线检测出骨架线十字中心,进而计算 C,M, Y 十字线中心分别相对于 K 十字中心的偏移量,即为

收稿日期: 2019-06-12

作者简介:吕明珠(1991-),女,西安理工大学硕士生,主攻基于机器视觉的表面缺陷检测。

通信作者:吴学毅(1966-),男,西安理工大学硕导、副教授,主要研究方向为三维图形可视化与虚拟现实。

套印偏差量。文中研究基于 K-means 聚类十字线标识 符的套印偏差检测算法,解决分色后图像前景十字线 提取困难的问题,以提升检测准确度。

## 基于十字线套准标识符的套印偏差 检测

传统的套印偏差检测以人工目视叠印标识符的 方式实现,角线、十字线标识符是使用最为广泛的叠 印式套准标识符,见图1,当偏移宽度大于一个线宽 时,即超过了0.1mm便判定该印品套印不合格。



图 1 常用叠印式套准标识符 Fig.1 Common overprint overlay quasi-identifier

使用机器视觉方式检测十字线 C, M, Y 相对于 K 色偏差量的过程中,最大的难点在于从获取到的十 字线原始图像中准确分解出 C, M, Y, K 四色十字 线。由 RGB 和 CMYK 补色关系理论可知<sup>[3]</sup>, RGB 并 不能一一对应 CMYK。若要从 RGB 转至 CMYK,必 须定义一个"黑色替代值",从 C, M, Y 中提取若干 成分来组成黑色通道,即 K 值。其中, RGB 取值范 围是 0~255, CMYK 取值范围是 0~100。基于上述 理论,给定 RGB 颜色空间的值(*R*, *G*, *B*),则转换 到 CMYK 空间的(*C*, *M*, *Y*, *K*)值计算式为:

 $\begin{bmatrix} c \\ m \\ y \end{bmatrix} = \begin{bmatrix} G_{\max} \\ G_{\max} \\ G_{\max} \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$  $K=\min(c, m, y)$ C=c - KM=m - KY=y - K

式中: *G*<sub>max</sub> 指颜色最大值,一般是 255。根据上述分色公式对中值滤波<sup>[4]</sup>去噪后的十字线图像(见图 2a)作 RGB 至 CMYK 的转换,得到的 C, M, Y, K

图像见图 2b—e。由图 2 可以看出,分色后的图像中 不仅包括目标色十字线,还包括其他色的干扰十字 线。如图 2d 所示,黄色 Y 图像中,不仅包括 RGB 图像中的黄色所在位置十字线,还包括了青、品红、 黑色所形成的弱灰度值干扰十字线。其中,青色十字 线和黑色十字线灰度值较接近前景黄色,品红色相对 较弱,背景黑色与青、品红、黑色十字线灰度差又相 对较大,因此,仅靠二分类法单纯地将干扰十字线和 背景划为一类,前景划为另一类的做法,易使提取结 果产生较大误差<sup>[5]</sup>。考虑将图像二分类改进为四分 类,这4个种类分别是:前景十字线、背景、较接近 前景十字线的干扰色、灰度更弱一些的干扰色。立足 于上述理论,文中研究的基于 K-means 聚类法检测套 印偏差的算法流程见图 3。

### 1.1 K-means 聚类分色后图像

K-means 算法主要以距离相似性为聚类依据,通 过计算类中样本均值,将相似样本划分到同一类别 中。聚类算法属于无监督学习<sup>[6]</sup>,算法步骤如下:第 1步,随机初始化 K 个聚类中心;第2步,计算每个 样本与每个聚类中心之间的距离相似度,将样本划分 到最相似的类别中;第3步,计算划分到每个类别中 的所有样本特征的均值,并将该均值作为每个类新的 聚类中心;第4步,重复计算第2、第3步过程,直 至聚类中心不再变化;第5步,得到最终的每个聚类 中心和每个样本的类标签<sup>[7]</sup>。

为了从分色后的 C, M, Y, K 图像中准确提取 出各自的前景十字线,使用 K-means 算法聚类 C, M, Y, K 图像,初始化参数如下所述。

1)初始化 K=4 类,初始聚类中心(灰度值)为: 第1 类中心 A<sub>1</sub>=7;第2 类中心 A<sub>2</sub>=60;第3 类中心 A<sub>3</sub>=100;第4 类中心 A<sub>4</sub>=170。

2) 聚类中心间距值 *b*<sub>1</sub>=3; *b*<sub>2</sub>=3; *b*<sub>3</sub>=3; *b*<sub>4</sub>=3。

3)初始化限制迭代次数: *c*<sub>1</sub>>0.01 或 *c*<sub>2</sub>>0.01 或 *c*<sub>3</sub>>0.01 或 *c*<sub>4</sub>>0.01;迭代上限 20次。

C, M, Y, K 等 4 幅图像聚类中心迭代过程见表 1。表 1 中记录聚类中心的数值变化过程,由于图像灰 度值一般用整数表示,故表 1 中计算结果小数部分全 部向上取整<sup>[8]</sup>。其中, C 图像共迭代 7 次; M, Y,



图 2 RGB 转 CMYK 后的 C, M, Y, K 图像 Fig.2 C, M, Y, K images after RGB to CMYK





K 图像分别迭代 5 次。初始中心数值的选取主要依据 4 个种类各自大致的分布特征,实验验证 4 个中心数 值间隔约 30~70,迭代结果一致。综上,初始中心 的选取对结果的敏感度不高,验证了 K-means 在十字 线聚类分割过程的适用性。

使用不同灰度值标记分类结果,可以方便后续阈 值分割。对 K-means 聚类结果第1类(最暗背景)赋 灰度值0;第2类(较弱十字线干扰)赋灰度值60; 第3类(较接近前景十字线的干扰色十字线)赋灰度 值150;第4类(前景十字线)赋灰度值240,标记 分类结果的聚类图像见图4a—d。使用固定阈值230 做图像分割<sup>[9]</sup>,提取出每幅图像前景十字线,即提取 出第4类。例如从4a图像提取青色十字线见图4e, 4b图像提取出品红色十字线见图4f,从4c图像提取 黄色十字线见图4g,4d图像提取出黑色十字线见图 4h。至此完成了C,M,Y,K等4幅灰度图像的目 标十字线提取,从结果图来看,提取效果比较准确。

#### 1.2 CMY 前景十字线减 K 十字线

图像差运算指 2 幅长宽一致的图像之间对应像 素做减法运算,通过做差可以确定出 2 幅图像之间的 差异处。该项技术可以快速提取出待检图像和标准图 像之间的缺损信息<sup>[10]</sup>。图像减运算公式:

#### $g(x, y) = f_1(x, y) - f_2(x, y)$

叠印式套准标识符大都为宽度 0.1 mm 的黑色细 线,四色套印越准确,重叠区(黑线)就越明显。将 标识符原图从 RGB 分解为 C,M,Y,K 等 4 色时, 重叠区域的黑色十字线将出现在每个色图中,即分色 后图像上原黑色十字线所在位置像素的灰度值大于 0。对于 C,M 或 Y 有明显套印偏差的图像,经过

图像	类中心	1	2	3	4	5	6	7	8
C图像	A1	7	4	2	1	1	1	1	1
	A2	60	45	42	38	35	35	32	32
	A3	100	115	109	102	98	97	96	96
	A4	170	156	146	141	141	140	139	139
M 图像	A1	7	3	2	1	1			
	A2	60	40	31	29	29			
	A3	100	105	90	66	66			
	A4	170	162	122	118	118			
Y 图像	A1	7	4	4	4	4			
	A2	60	42	31	33	33			
	A3	100	65	72	70	70			
	A4	170	145	116	113	113			
K图像	A1	7	6	9	9	9			
	A2	60	51	52	53	53			
	A3	100	101	96	93	93			
	A4	170	177	173	172	172			

表 1 K-means 聚类中心迭代过程 Tab.1 Iteration process of K-means clustering center



图 4 基于 K-means 十字线聚类和分割后效果图 Fig.4 Design sketch based on K-means cross line clustering and segmentation

K-means 聚类、阈值分割后的二值图像中,黑色十字 线与前景十字线像素重叠较少,但是像素分割较完 整,见图 4f。同理,套印合格的颜色偏移量微小,黑 色十字线与前景像素重叠像素多,见图 4g。发生弱 小偏移的颜色前景十字线易紧贴黑色十字线,此时, 黑色十字线部分易出现零散的不太连续的部分像素。 综上,阈值分割后的 C,M,Y,K 等 4 幅图像仍然 受到 K 十字线的干扰,使得各色十字线中心计算误 差大,解决上述问题:从阈值分割后的 C,M,Y 图 像中减去 K 十字线。

图 4e—f 图像依据上述减运算公式做差后图像见 图 5a—c。至此,基于 K-means 聚类的前景十字线提 取算法得到的效果图见图 5d,与原图对比发现,各 色十字线骨架均被准确提取了出来。

### 1.3 十字线中心计算

基于 K-means 聚类的前景十字线提取算法得到 的 C, M, Y, K 十字线骨架, 会出现十字线不完整、 间断、缺失部分等缺点<sup>[11]</sup>,但是并不影响十字线骨架 中心的位置计算,通过一定的数学逻辑即可求出,以 品红色(M)前景十字线为例,计算过程如下所述。 1) 扫描图 5b 二值图像, 找到前景像素数最多的 行和列, 分别标记为 MFcols, MFrows。

2)扫描图 5b 二值图像,找到前景像素数第二多的行和列,分别标记为 MScols, MSrows。

3) 如果 MSrows>(MFrows/2),则列中心 MCol= (MFcols+MScols)/2 行中心 MRow 同理。如果 MSrows

(MFrows /2),列中心 MCol=MFcols,行中心 MRow=MFrows。

4) 对图 5a, c, d 重复 1) —3) 步骤, 分别得到 青色行中心 CCol, 青色列中心 Crow, 黄色行中心 YCol, 黄色列中心 YRow, 黑色行中心 KCol, 黑色 列中心 KRow。

5) 依次使用计算得到的 C, M, Y 行中心和列 中心减去 K 的行中心、列中心,得到行方向上和列 方向上相较于 K 的像素距离值,转化为实际长度 (mm), 即为 C, M, Y 相对于 K 的套准偏移量<sup>[12]</sup>。

### 2 实验结果

使用 15 幅四色胶印机印刷的样张做实验对象, 通过视觉扫描设备获取印刷品数字图像,对每幅图像



图 5 C, M, Y 分别与 K 做差后图像 Fig.5 C, M, Y image after making difference with K

做预处理后<sup>[13]</sup>,分别使用大津法<sup>[14]</sup>和文中的 K-means 四聚类法提取出前景十字线,然后计算出各色套印偏 差量。再使用 8 倍放大镜和单位 0.01 mm 的标准线规菲 林尺对每个十字线进行人工偏移量测量,结果见图 6。

由于人工测量结果可信度高,上述2种算法检测 结果越接近人工精确测量值,说明其正确性越高。由 图6中的6幅图可以看出,虚线较实线更接近0值, 即文中的基于 K-means 聚类的十字线套印偏差检测 结果较基于大津法的十字线检测结果准确性更高,较 人工精确测量值误差在±0.04 mm 以内,符合套印偏 差检测精度要求<sup>[15]</sup>。

对上述 15 幅图像使用 2 种算法所耗时间见表 2, 由表 2 可知, K-means 算法处理一幅十字线图像所耗 时间大致保持在 111 ms 左右,使用大津法所需时间



图 6 算法检测偏差值与人工精确测值的对比

Fig.6 Comparisons between algorithm detection deviation and precise manual measurement result

表 2 文中方法和大津法检测时间的对比

		Iai	5.2 CO	npariso	on detv	ween K	-mean	s met	nou al	na Ots	u test	time				ms
方法	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	均值
文中方法	116	110	120	115	108	109	106	99	108	126	116	115	111	106	95	111
大津法	120	137	133	152	112	122	150	99	109	136	138	127	119	120	99	132

大致在 132 ms 左右, 文中算法的时间性能优于大津 法。同时, 根据现有四色胶印机印刷每小时 15 000 张, 即每秒 4.16 张的速度, 文中算法速度可以满足 在线检测要求。

## 3 结语

对机器视觉系统获取到的印刷品图像做 RGB 至 CMYK 色彩空间转换的过程中,CMYK 四色并不能 实现绝对的独立提取,相互之间不可避免将发生一定 干扰,这使得叠印式十字线套印标记前景十字线提取 困难,影响检测结果的正确性。文中使用 K-means 聚类算法对分色后得到的 C,M,Y,K 图像分别做 聚类处理,然后分割图像并提取出各色前景像素,进 而使用 C,M,Y 图像中的前景十字线分别减去 K 图 像中的十字线,以消除四色重叠区域的干扰,最后提 取十字线中心并计算偏差量。实验使用 15 幅印刷品 图像,分别将大津法提取前景和文中算法提取的前景 检测结果与人工精确测量值进行对比,结果表明文中 提出的检测方法更加接近人工精确测量结果,即准确 性更高。

基于数字图像处理的表面缺陷检测精度较大地 受限于缺陷检测算法,另一方面,随着图像采集设备 的飞速发展,机器视觉系统获取的印刷品图像质量正 在逐渐提升,包括分辨率和抗噪性能等,未来表面缺 陷检测系统的性能和精确度仍然有较大的提升空间。

#### 参考文献:

- 刘伟.基于机器视觉的印刷品缺陷在线检测系统
   [D].兰州:西北师范大学,2018.
   LIU Wei. On-line Defect Detection System for Printed Matter Based on Machine Vision[D]. Lanzhou: Northwest Normal University, 2018.
- [2] 杨顺波,龙永红,姚佳成.基于图像的目标区域分割 算法研究[J].电子产品世界,2019,26(2):64—68.
   YANG Shun-bo, LONG Yong-hong, YAO Jia-cheng. Research on Image-based Target Region Segmentation Algorithm[J]. Electronic Products World, 2019, 26(2): 64—68.
- [3] 沈静,陆静,李宏.基于 RGB 相关性的彩色图像可逆信息隐藏方法[J].计算机工程与应用,2018,54(20):69—73.
  SHEN Jing, LU Jing, LI Hong. Reversible Information Hiding Method for Color Image Based on RGB Correlation [J]. Computer Engineering and Application, 2018, 54 (20): 69—73.
- [4] FANG Z, YI X, TANG L. An Adaptive Boosting Algorithm for Image Denoising[J]. Mathematical Problems in Engineering, 2019, 2019(3): 1–14.
- [5] 赵丹,丁金华,孙秋花.基于计算机图像处理的印刷

品缺陷检测[J]. 包装工程, 2008, 29(12): 120—121. ZHAO Dan, DING Jin-hua, SUN Qiu-hua. Printing Defect Detection Based on Computer Image Processing [J]. Packaging Engineering, 2008, 29(12): 120—121.

- [6] 黄吉. 一种 K-means 聚类改进算法研究及应用[D]. 武汉: 湖北工业大学, 2018.
   HUANG Ji. Research and Application of an Improved K-means Clustering Algorithm[D]. Wuhan: Hubei University of Technology, 2018.
- [7] 张长勇, 吴智博. 基于 K-means 与关键点的组合行 李码放算法[J]. 包装工程, 2019, 40(9): 90—95.
   ZHANG Chang-yong, WU Zhi-bo. Combined Baggage Code Placement Algorithm Based on K-means and Key Points[J]. Packaging Engineering, 2019, 40(9): 90—95.
- [8] 朱霞,伍岳,陈俊斌,等.基于主成分聚类的军品防 护包装等级分类[J].包装工程,2017,38(23):21—25. ZHU Xia, WU Yue, CHEN Jun-bin, et al. Classification of Military Protective Packaging Grade Based on Principal Component Clustering[J]. Packaging Engineering, 2017, 38(23): 21—25.
- [9] 柴江松,王琪,刘洪豪.印刷网点微观图像阈值分割 算法研究[J].包装工程,2015,36(13):115—121. CHAI Jiang-song, WANG Qi, LIU Hong-hao. Research on Threshold Segmentation Algorithm for Micro-Image of Printing Dots[J]. Packaging Engineering, 2015, 36(13): 115—121.
- [10] ALARD C, LUPTON R. A Method for Optimal Image Subtraction[J]. The Astrophysical Journal, 1998, 503(1): 325-331.
- [11] 李俊锋,唐正宁,李森鑫.印刷十字线套准检测算法的研究[J]. 计算机工程与应用, 2012, 48(28): 174—177.
  LI Jun-feng, TANG Zheng-ning, LI Miao-xin. Research on the Alignment Detection Algorithms for Printed Cross Lines[J]. Computer Engineering and Applications, 2012, 48(28): 174—177.
- [12] 田敏, 刘全香. 分区域分等级的印刷品缺陷检测方法[J]. 包装工程, 2015, 36(21): 122—127.
  TIAN Min, LIU Quan-xiang. Subregional Graded Printing Defect Detection Method[J]. Packaging Engineering, 2015, 36(21): 122—127.
- [13] 赵翔宇,周亚同.工业干扰环境下基于模板匹配的印刷品缺陷检测[J]. 包装工程, 2017, 38(11): 187—192.
   ZHAO Xiang-yu, ZHOU Ya-tong. Printing Defect Detection Based on Template Matching in Industrial Interference Environment[J]. Packaging Engineering, 2017, 38(11): 187—192.
- [14] SHEN Xuan-jing, LIU Xiang, CHEN Hai-peng. Fast Computation of Threshold Based on Multi-threshold Otsu Criterion[J]. Journal of Electronics & Information Technology, 2017, 39(1): 144—149.
- [15] 李雪梅,唐万有.印刷品缺陷检测的方法研究[J]. 包装工程,2007,28(12):286—288.
  LI Xue-mei, TANG Wan-you. Research on Defect Detection Methods of Printed Matter[J]. Packaging Engineering, 2007, 28(12):286—288.