

基于插值检测的扫描图像源辨识方法*

房穗韬, 方艳梅

(中山大学信息科学与技术学院, 广东 广州 510006)

摘要: 针对扫描图像真实性的问题, 提出了应用 EM 算法与扫描仪传感器噪声的自然统计特性相结合的识别方法。该识别方法一共生成 63 维特征, 用 SVM 进行训练, 通过大量实验图像, 文中对取证方法效果进行了分析研究。结果显示, 扫描图像虽然在生成过程中不存在 CFA 插值过程, 但是在被显示前, 仍经过了插值运算, 傅立叶频谱图仍具有周期性。该算法对 SG 图像(扫描图像)与 CG 图像、SG 图像与 DG 图像(数码相机生成图像)之间的识别, 达到了较理想的效果。

关键词: 图像取证; 扫描图像; 传感器噪声; 插值检测; EM 算法

中图分类号: TP391.4 **文献标志码:** A **文章编号:** 0529-6579(2013)05-0019-07

Scanned Image Source Identification Based on Interpolation Detection

FANG Suitao, FANG Yanmei

(School of Information Science and Technology, Sun Yat-sen University Guangzhou, Guangdong, 510006, P. R. China)

Abstract: Since scanned images are not generated by CFA process, a part of researchers have studied scanned image forensic around sensor pattern noise. A method is proposed to contain the statistical features of the noise image and features derived from EM algorithm for the authenticity of the scanned image. This method generates a total of 63 dimensional feature classified by SVM. By analyzing the evidence effectiveness and testing a large number of images, the work results show that although scanned images is not generated with CFA process, scanned images also has experienced interpolation on its Fourier spectrum map. Moreover, for the purpose of distinguish between scanned images and other images (CG, DG), this method achieves the desired results.

Key words: image forensic; scanned image; sensor noise; interpolation detection; EM algorithm

随着数字照片时代的到来, 用于商业交往的传统支票运作方式, 逐渐被通过扫描传递的数字影像支付方式替代, 大大提高了资金转帐和利用效率。然而, 扫描的数字图像因其易操作性, 往往会被恶意利用。因此, 开展针对图像真实性的扫描图像取证研究, 对于确保公共信任秩序、打击犯罪、维护司法公正和新闻诚信具有十分重要的社会意义。

国内对扫描图像认证的研究极少, 国外研究集中在基于成像过程的一致性上。Gloe^[1]把对传感器噪声源的识别算法用于扫描图像的辨识, 并用

TIFF 格式图像在三种扫描仪模型中进行了实验分析, 但因扫描仪的玻璃板上可能存在划痕或灰尘^[2], 使其结果受到影响; Delp 等^[3-4]一方面应用扫描仪的模式噪声尝试区分三种扫描仪, 对 TIFF 格式的图像库区分度达到 96%, 一方面通过传感器模式噪声的自然统计特性尝试对扫描图像、数码相机图像及 CG 这三类的 TIFF 格式图像进行区分, 三类图像两两区分度均在 85%; 此外, Gou^[5]等提出了一种基于图像去噪、邻域分析、小波分析等多角度的自然统计特征提取方法, 对扫描

* 收稿日期: 2013-03-13

基金项目: 国家 973 计划资助项目(2011CB302204); 国家科技支撑计划资助项目(2012BAK16B00)

作者简介: 房穗韬(1987年生), 男; 研究方向: 数字图像取证; 通讯作者: 方艳梅; E-mail: fangym@mail.sysu.edu.cn

图像源进行认证。实验中尝试了对 7 种扫描仪进行区分, 每种扫描仪的图像为 26 幅, 每幅图像均提取 60 维特征, 平均准确率约为 90%。由于实验中数据库样本偏少, 实验结果仍需进行大量的测试。

本文通过分析扫描图像 (scanned image, SG)、数码相机图像 (digital image, DG)、计算机生成图像 (computer image, CG) 三类图像的成像过程中的固有区别, 针对目前流行的 JPEG 格式图像提出了应用 EM 算法与图像噪声的自然统计特性相结合的分类源辨识方法, 达到了区分扫描仪图像与其他两类图像的目的, 其中区分 SG 与 CG 及 DG 混合类图像, 达到了接近 95% 的识别率。

1 扫描仪成像过程分析

常见的扫描仪一般包括主板、扫描头、光学系统、机械机构等四个部分^[6]。扫描仪的原理如图 1 所示, 扫描仪工作时, 光源会把光线照射在准备输出的图像上, 为了能均匀地照在 SG 图像, 扫描仪光源为长度与工作平台宽度一样宽的长条形, 扫描时做垂直于扫描方向的动作, 每扫描一行就能得到原图像横向一行的图像信息。照射到原图上的光线经反射后, 穿过一个较细的缝隙, 形成一条横向光带, 再经过一组反射镜的折射, 最后由一个光学透镜聚焦再照到感光器件 CCD 上, CCD 实现数字模拟转换功能, 把聚焦产生的光信号转化为对应的模拟信号后, 生成数字信号, 暂存在缓存器中。扫描仪在机械传统机构与扫描介质进行相对运动后, 会得到全部的图像信息, 并按照扫描的前后顺序把数据传递给连接的计算机中。最终图稿被显示在屏幕前, 还需在扫描仪中经过自带软件的重新处理, 其中包括插值、颜色变换、白平衡、曝光率控制、降噪等预处理^[7]。

由上述成像过程分析可知, SG 与其他生成设备图像存在以下区别。

1) 扫描仪是逐行扫描的, 采用一维的 CCD 来获取二维图像, 只存在水平方向的像素点, 扫描仪的最大分辨率由步进电机的最小位移和水平像素点的数量共同决定。

2) 扫描仪分一次扫描与三次扫描两种扫描技术, 分别由一个 CCD 扫描 RGB 三通道三次或使用三个线性的 CCD 一次性扫描 RGB 三个通道, 两种技术均不存在颜色插值过程。

3) 扫描介质一般不可能完全覆盖扫描仪的玻璃板, 扫描过程中通常只有一部分单元参与扫描过程, 而且由于是逐行扫描, 图像在存储前要经过扫

描仪自带软件进行内部处理, 其中包括插值、颜色变换、白平衡、曝光率控制、降噪等等。

4) 常年使用的扫描仪上经常会存在一些物理上的小缺陷, 如划痕与灰尘, 这些物理特性即使在压缩率较高的 SG 上, 也依然会被检测出来。

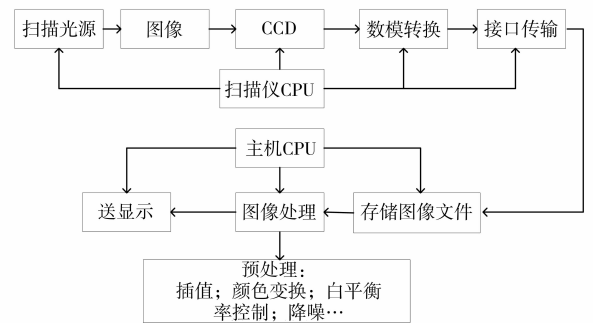


图 1 扫描仪原理

Fig. 1 The principle of scanner

由于扫描仪在成像过程中均是采用逐行扫描, 导致了行与行之间没有运用插值运算, SG 与 DG 图像在傅立叶频谱图上的差异如图 2 所示, SG 的横向周期性与列向周期性不一样, 导致了在频谱图上有明显的亮线, 而 DG 图像则是横纵周期一致, 中间有一明显的亮点, 高频低频成分均匀分布。根据此差别, 本文提取 EM 算法结果中的残留误差的加权平均系数与噪声图像的自然统计特性作为统计特征, 区分 SG 与其他类型图像, 具体的流程如下。

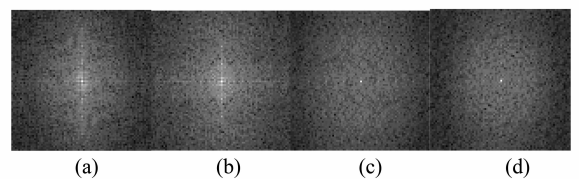


图 2 傅里叶频谱图 (a) 和 (b) SG 图像; (c) 和 (d) DG 图像

Fig. 2 Fourier spectrum (a) and (b) for SG image, (c) and (d) for DG image

2 EM 算法

现实世界中经常存在着一些“不完全数据”, 即数据集中的数据是不完整的, 通过引入额外的参数, 可以对这些数据集合进行优化, EM 算法^[8], 即期望最大化算法。它是一种当观测数据不完整时, 求解最大似然估计的迭代算法。

而在数字图像中, 图像的样本点可以分为两类: 第一类是与相邻样本相关联的样本点集合, 用

C_1 表示；第二类是与它们的相邻点不相关的样本点集合，用 C_2 表示，图像中的每个样本点要么属于 C_1 ，要么属于 C_2 ，EM 算法就是求解图像像素点属于 C_1 的迭代算法^[9]。一幅图像应用 EM 算法迭代求解图像中的像素相关性的流程图见图 3。

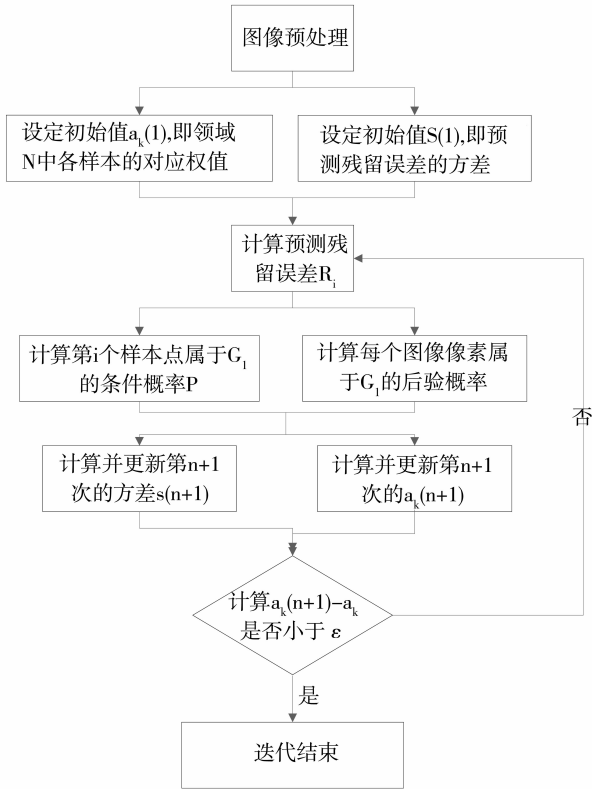


图 3 EM 算法流程图

Fig. 3 The flow chart of the EM algorithm

在上述流程中，本文选定邻域 N 为 3，即假设样本点周围 3×3 的范围为样本点的插值区间，每幅图像均提取邻域的加权系数作为该图像的第一类特征，即应用 EM 算法后，每幅图像的分类特征维数为 48 维。

后验概率图是一幅图像经过 EM 算法后的结果之一，它反映了一幅图像中的像素点属于 C_1 的概率，如图 4 所示。由于 C_1 类像素点是由邻域像素点插值达到的，因而根据不同类型的图像，呈现一定周期性。SG 经过 EM 算法后得到的插值概率图像也是一种周期性信号，所以本文可通过傅立叶变换观察其中的周期特性。

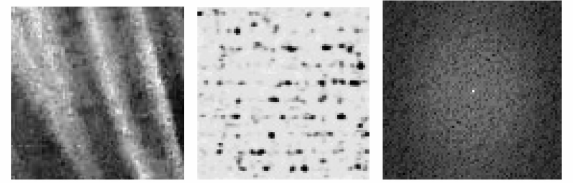
图 4 为图像经过 EM 算法后的运算结果，图 4 (c) 列表明，DG 图像的行与列都参与插值运算，故在傅立叶频谱图上，DG 图像是行与列周期性一致，呈均匀分布；而 SG 则因成像过程是逐行扫描，经扫描仪内部软件处理时，其只有列方向参与插值

运算，其傅立叶频谱图的行列方向周期性不一致，在频谱图上呈现出一条明显的亮线；而 CG 图像则因其生成过程，而在傅立叶频谱图上毫无规律性。

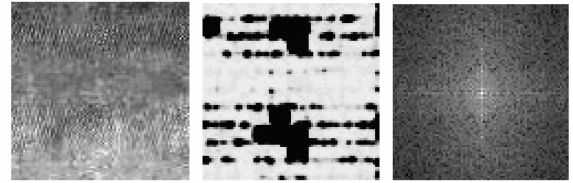
CG:



DG:



SG:



(a) 原图 (b) 插值概率图像 (c) 傅立叶频谱图

图 4 三类 JPEG 图像经 EM 算法的结果

Fig. 4 EM results of images

实验过程中为了更好地观察傅立叶频谱图，对它进行了相应变换，把频谱移频到原点。变换后的图像频率分布会以原点为中心呈对称分布，中心区域表示图像的低频成分，外部为图像的高频成分。

3 扫描仪传感器噪声

除了应用 EM 算法对 SG 进行源取证，本文尝试引入传感器噪声的自然统计特性作为宏观统计特征对上述算法进行补充。传感器模式噪声的识别算法已经成功应用于数码相机的分类取证^[10]，在这些研究中发现，模式噪声中存在着两种不同的成分：一种是随机噪声，它会随着不同图像的变化而改变，不利于识别，另外一种固定噪声，它不会随着图像的变化而改变，是代表传感器内部的固有的噪声，这种噪声在扫描仪中也存在^[11]。利用这种固有的特征，我们把两种成份分离，得到想要的固有部分，然后提取其自然统计特征，用 SVM 进行区分^[12]。

图像扫描是逐行进行的，故产生的每一行数字图像都是由相同传感器像素产生的，所以对于 SG，

行平均值可以近似地被称为行传感器噪声参数, 这种均值处理在减低随机噪声的影响时, 也同时增强了固有噪声。同理, 将列平均值称为列传感器噪声参数^[13]。实现过程如下, 通过去噪算法获取噪声图像, 设该图像的分辨率大小为 $M \times N$, 于是

$$I_n = I - I_{de} \quad (1)$$

然后, F_n 和 F_n^c 分别表示行平均噪声参数和列平均噪声参数, 并且

$$F_n(1, j) = \frac{1}{M} \sum_{i=1}^m I_n(i, j); \quad 1 \leq j \leq N \quad (2)$$

$$F_n^c(1, j) = \frac{1}{M} \sum_{i=1}^m I_n(i, j); \quad 1 \leq j \leq N \quad (3)$$

这里, X 与 Y 的协方差函数定义为, $C(X, Y)$

$$C(X, Y) = \frac{(X - \bar{X}) - (Y - \bar{Y})}{\|X - \bar{X}\| \|Y - \bar{Y}\|} \quad (4)$$

用 $\rho_{row}(i)$ 表示行平均噪声参数与每一行的传感器噪声参数的协方差函数^[14], 其定义为

$$\rho_{row}(i) = C(F_n, F_n(i, \cdot)); \quad 1 \leq i \leq M \quad (5)$$

同理, 用 $\rho_{col}(i)$ 表示列平均噪声参数与每一列的传感器噪声参数的协方差函数, 其定义为

$$\rho_{col}(j) = C(F_n, I_n(\cdot, j)); \quad 1 \leq j \leq M \quad (6)$$

实验过程中, 发现用扫描仪较低光学分辨率进行扫描的时候, 会产生较大的随机噪声, 从而导致照明上的波动, 导致 ρ_{row} 的值会比较小, 和 ρ_{col} 的值接近; 相反, 用扫描仪较高光学分辨率进行扫描时, 产生的随机噪声比较小, ρ_{row} 值会比较大, 远远大于 ρ_{col} 值。并且, F_n 和 F_n^c 的统计特性在扫描仪光学分辨率较低时要比光学分辨率较高时较明显。针对上述情况, 我们对和的均值、标准差、峰度和偏度的这些特性进行了自然统计, 作为 1~8 的特征值, 我们再对和的标准差、峰度和偏度这 3 个特征值进行统计, 作为 9~14 的特征值。第 15 个特征 F_{15} 用公式 (7) 表示,

$$F_{15} = \left(1 - \frac{\frac{1}{N} \sum_{j=1}^N \rho_{col}(j)}{\frac{1}{N} \sum_{i=1}^N \rho_{col}(i)} \right) * 100 \quad (7)$$

对于光学分辨率较高的扫描图片, 在大部分图片上都会有很好的效果, 它反应行列相关系数的比值。

通过预处理, 我们把图像转换成灰度图, 再经过双密度双树小波变换的去噪算法^[15], 得到其噪声图像。把噪声图像经过如上处理后, 统计其自然特性, 获取所需的 15 个特征值, 作为本文算法的第二类特征。

4 实验结果分析

实验中采用支持向量机 SVM^[16-17] 作为分类

器。对每个图像我们只取了部分大小 (如 1024×1024 、 512×512) 作特征提取对象。对图像库的图像提取了 EM 算法特征和噪声特征, 然后分别选取了部分图像特征对 SVM 进行了训练, 其余图像特征用于测试^[18]。

实验中的数据库分为三部分, 分别是 DG 图像库、CG 图像库和 SG 图像库, 其中 SG 图像库分别来自 18 台不同的扫描仪, 当中包括了 150 dpi 到 600 dpi 的 747 幅图像, 而 CG 图像均是从网络下载的 747 幅图像, DG 图像库包含了 6 个品牌相机的 1494 幅图像。扫描仪、相机的列表分别如表 1、表 2 所示。

表 1 扫描仪图像库¹⁾

Table 1 Scanner image database

编号	扫描仪类型	有效像素	数量	大小
1	CanoScan 500F	2 480 × 3 507	80	1 024 × 1 024
2	AvisonFB2100_ A	2 441 × 2 797	80	1 024 × 1 024
3	BenQ 5800	3 259 × 2 487	100	1 024 × 1 024
4	Canon DR - 7580	2 176 × 1 680	100	1 024 × 1 024
5	HP Scanjet G4050	2 256 × 1 632	60	1 024 × 1 024
6	Epson Perfection V330	2 176 × 1 704	60	1 024 × 1 024
7	CanoScan 9000F	3 510 × 2 482	100	1 024 × 1 024
8	各种机型 (小数量)	2 534 × 3 238	167	1 024 × 1 024

1) 格式均为 JPEG

表 2 数码相机图像库¹⁾

Table 2 Digital image database

编号	数码相机类型	有效像素	数量	大小
1	NIKON D300	4 288 × 2 848	150	1 024 × 1 024
2	OLYMPUS - u20D, S400D	2 272 × 1 704	150	1 024 × 1 024
3	PANASONIC - DMC - FZ30	3 264 × 2 448	127	1 024 × 1 024
4	SONY DSC S800	2 048 × 1 536	110	1 024 × 1 024
5	CANON EOS 600D	1 600 × 1 200	110	1 024 × 1 024
6	OLYMPUS VR - 310	2 364 × 1 564	100	1 024 × 1 024

1) 格式均为 JPEG

4.1 检测率分析

首先本文应用 EM 算法提取插值概率图像与傅立叶频谱图, 区分三类图像。每幅图像应用 EM 算法后得到的残留误差的加权平均系数为分类特征, 共 48 维。三类图像特征分成如下三组实验对象: SG 图像与 CG 图像、SG 图像与 DG 图像、SG 图像与 CG&DG 混合图像 (MG)。按上述实验组别, 用

SVM 对特征进行训练和预测，其中数据库中 70% 的图像样本训练，30% 的图像用于预测，结果由图 5 (a) 的 ROC 性能曲线表示，其中曲线图中横坐标表示虚警率，纵坐标为检测率。

表 3 显示，SG 图像与 CG 图像的源分类达到了 91.79%，具有较好的效果，因为 CG 类图像在成像过程中不存在插值运算，而 SG 类图像在列方向上存在着插值运算；同理，扫描仪与数码相机的成像过程不同，插值算法及插值邻域也会不同，相机的颜色滤波阵列是周期性排列的，因此，经过插值后图像傅里叶频谱图同样呈现周期性，DG 图像行与列的邻域均参与了插值运算，而 SG 图像只在列方向上存在着插值操作，故 SG 图像与 DG 图像的源取证也达到了很不错的效果，有 93.96%。最后，SG 图像与其余两类图像组成的混合图像也有较好的分类效果，区分率为 91.30%。

表 3 SG 图像源取证 ROC 曲线性能指标

Table 3 Scanned image source identification ROC curve parameter

算法	类别	检测率	虚警率	漏检率	AUC
		/%	/%	/%	
EM-48D	SG vs. CG	91.79	11.11	5.31	0.958 5
	SG vs. DG	93.96	9.18	2.90	0.980 8
	SG vs. MG	91.30	11.58	5.80	0.975 0
SN-15D	SG vs. CG	87.28	14.29	11.16	0.926 8
	SG vs. DG	90.40	7.60	11.61	0.954 7
	SG vs. MG	89.51	12.05	8.93	0.933 3
SN&EM-63D	SG vs. CG	93.96	6.21	3.86	0.971 8
	SG vs. DG	96.14	2.90	4.83	0.991 9
	SG vs. MG	94.93	6.76	3.38	0.985 5
文献 4-45D	SG vs. CG	77.03	31.09	14.86	0.826 4
	SG vs. DG	80.63	29.73	9.09	0.867 7
	SG vs. MG	77.48	27.03	18.02	0.819 0
文献 5-60D	SG vs. CG	93.59	8.42	5.41	0.970 2
	SG vs. DG	95.52	3.25	5.71	0.987 7
	SG vs. MG	94.72	7.71	5.85	0.975 8

由于实验中数据库采用的是 JPEG 格式图像^[16]，JPEG 图像在其压缩量化过程中，围绕其减少图像存储空间的目的，必不可少地会损失部分高频的细节特征，削弱图像像素间的相关性，故用作实验对象难度较大。本文尝试引入传感器噪声的自然统计特性作为宏观的统计特征对上述算法进行补充。一幅图像通过提取噪声图像后，计算上述提及的噪声自然统计特性共 15 维统计特征，通过 SVM 训练及预测，观察算法的识别率，结果如表 3 所示。

表 3 表明，若只应用噪声图像的统计特征区分

SG 图像与其他两类图像，效果并不显著。SG 图像与 CG 图像的源辨识率达到 87.28%，SG 图像与 DG 图像的源辨识率为 90.4%，SG 图像与其余两类混合类图像的源辨识率为 89.51%。因为 CG 图像是通过电脑生成的，不存在固有的传感器噪声；同理，扫描仪与数码相机的成像过程不同，固有的传感器噪声自然统计特性也会不同，所以传感器噪声的自然统计特性能从统计学的角度对三类图像稍作区分。

最后，本文对两种算法特征进行合并整理，用 SVM 训练图像库的特征值，仍用 70% 进行训练，剩下 30% 进行测试，结果由图 5 (a) 的 ROC 性能曲线展示。

综合实验结果得出，当把两类算法特征协同应用时，SG 与 CG 类图像的源取证达到了理想的效果，有 93.24%；SG 与 DG 图像的源取证更达到了明显的效果，检测率为 96.14%，SG 与混合类图像的区分也有 94.93%，比在单独应用一种算法的效果有所提高，ROC 各性能参数如表 3 所示。文献 [4] 中，Khanna 的方法应用于三色通道中提取得到的 45 维特征对三类 TIFF 格式辨识率为 85%，JPEG 格式图像为 79.8%，如图 5 (b) 所示。本文算法针对 JPEG 格式图像平均准确率为 94.93%，对比文献 [4]，更具有实用性及广泛性。

当实验图像数据库是 JPEG 图像格式时，图片在压缩量化过程中特征有损失，文献 [5] 中提出了它们在这方面的缺陷，本文算法在一定程度上弥补了 JPEG 图像压缩过程中造成的特征损失，提高了利用盲检测辨识 SG 与其他两类图像的准确率。本文算法与文献 [5] 的对比结果如图 5 (a) 与图 5 (c) 所示，从准确率的角度观察，本文算法比文献 [5] 略优；从算法的计算复杂度观察，本文算法为 $O(n)$ ，文献 [5] 所提算法为 $O(n^2)$ ，本文算法运行时间远远小于文献 [5] 所提算法。

在实验过程中，本文实验对象采用 JPEG 格式而不采用 TIFF 格式，原因在于，从实用性及普适性角度来看，TIFF 格式所占存储空间比 JPEG 格式要大很多，一般来说，一幅 300 万像素照相产生的 DG 图片，JPEG 格式 1 M 大小，但是 TIFF 格式 20 M 大小；正因 TIFF 格式图像所需存储空间大，其包含的信息也比 JPEG 格式图像多，从信息量的角度上看，区分 JPEG 格式的源分类图像，要比 TIFF 格式困难得多。

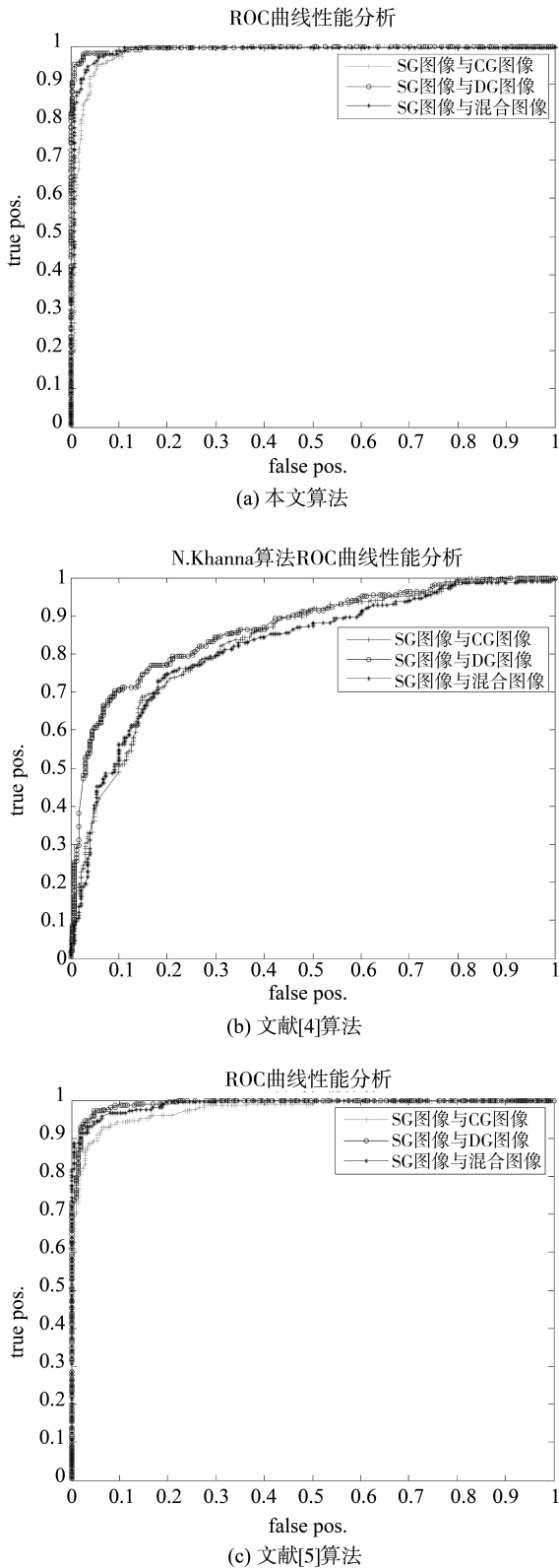


图5 SG图像源取证ROC曲线 (a) 本文综合算法的ROC曲线; (b) 文献[4]所提算法; (c) 文献[5]所提算法

Fig. 5 Scanned image source identification ROC curve (a) Proposed Synthesis algorithm; (b) Khanna's algorithm [4]; (c) Gou's algorithm [5]

由于EM算法中相关因子个数 N 的选取对结果有较大影响, 本文通过比较邻域 N 分别为3, 4, 5时, EM算法的准确率, 最后选取了 $N=3$ 完成了上述实验, 邻域 N 为4和5时, 本文算法结果如表4所示。从表中可以看出, 邻域 N 越大, 准确率反而会降低, 选取合适的邻域 N 对实验非常重要。

表4 改变邻域 N 时EM算法的准确率

N	类别	检测率/%	TP/%	TN/%	AUC
3	SG vs. CG	91.79	93.72	89.86	0.963 4
	SG vs. DG	93.96	92.75	95.17	0.980 9
	SG vs. MG	91.30	88.41	94.21	0.961 7
4	SG vs. CG	89.85	94.20	85.51	0.960 5
	SG vs. DG	91.54	89.85	93.24	0.967 2
	SG vs. MG	89.61	90.82	88.41	0.956 8
5	SG vs. CG	88.33	83.81	92.86	0.944 1
	SG vs. DG	90.24	89.52	90.95	0.959 5
	SG vs. MG	89.29	91.43	87.14	0.951 8

4.2 稳定性分析

为了验证算法的稳定性, 我们采取“ M -重交叉验证”的方法。在操作过程中, 我们把CG类跟SG类的图像库特征各分成10份, 分别为 $M_1, M_2, M_3 \dots M_{10}$, 每份74张, 每次用其中的9份组成各自的图像库, 然后放入SVM中进行训练及测试, 最后实验结果如图6所示。从图中, 我们可以看出, 各性能指标上下波动不大, 表示算法相对稳定。

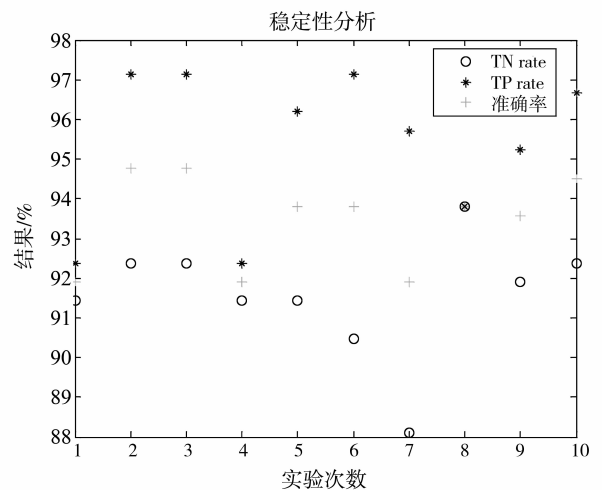


图6 M 重交叉验证稳定性分析

Fig. 6 Stability of M -fold cross validation

此外, 我们还针对图像的大小对本算法进行稳定性分析。针对 1024×1024 、 512×512 及 $256 \times$

256 这三种不同尺寸, 我们对 SG 图像与 CG 图像进行了稳定性分析, 结果如表 5 所示。

表 5 对不同尺寸的 SG 与 CG 图像应用综合算法

Table 5 ROC curve parameter by SG and CG

图像大小	检测率/%	TP/%	TN/%	AUC
1 024 × 1 024	93.96	91.79	96.14	0.971 8
512 × 512	94.20	90.82	97.58	0.979 3
256 × 256	92.51	92.75	92.27	0.973 5

从表 5 可以看出, 随着图像的大小不断降低, 本算法的正确率并没有大幅下降, 说明算法较稳定, 针对小尺寸图像, 也能达到较好的效果。

5 结 论

本文提出了一种基于图像插值检测, 并与扫描仪传感器噪声的行列自然统计特性相结合的一种新算法来对扫描图像进行源取证。本文证明了插值检测不仅能应用于数码相机的源分类检测, 也同样能应用于扫描图像取证。检测图像的插值相关性还有很多方法, 下一步可以从这方面进一步研究, 提高算法的准确率。

参考文献:

[1] GLOE T, FRANZ E, WINKLER A, Forensics for flatbed scanners[C] // Proc of SPIE Security, Steganography, Watermarking of Multimedia Contents IX, 2007:650511.

[2] DIRIK A E., SENCAR H T, MEMON N, Flatbed scanner identification based on dust and scratches over scanner platen[C] // IEEE International Conference on Signal Processing, 2009: 1385 - 1388.

[3] KHANNA NITIN, MIKKILINENI A K, DELP Edward J. Scanner identification using sensor pattern noise [C] // Proc of SPIE Security, Steganography, and Watermarking of Multimedia Contents IX, 2007:65051K.

[4] KHANNA NITIN, GEORGE T C, DELP Edward J. Forensic techniques for classifying scanner, computer generated and digital camera images[C] // IEEE International Conference on Signal Processing, 2008: 1653 - 1656.

[5] GOU H, SWAMINATHAN A, Robust scanner identification based on noise features [C] // Proc of SPIE Security,

ty, Steganography, Watermarking of Multimedia Contents IX, 2007:65050S.

[6] 张桂兰. 扫描仪的原理及其使用方法[J]. 印刷技术, 1994, 11: 14.

[7] HOLST G C, LOMHEIM T S, CMOS/CCD Sensors and Camera Systems [C] // The International Society for Optical Engine, 2007.

[8] 黄颖, 李伟, EM 算法与 K-Means 算法的比较[J]. 计算机与现代化, 2007, 9: 12.

[9] POPESCU A C, FARID Hany. Exposing digital forgeries by detecting traces of re-sampling [J]. IEEE Tran on Signal Processing, 2005, 53 (2): 758 - 767.

[10] LUKAS J, FRIDRICH J, Digital camera identification from sensor pattern noise [J]. IEEE Trans on Information Forensics and Security, 2006, 1(2): 205 - 214.

[11] GOU H, SWAMINATHAN A, WU Min. Intrinsic sensor noise features for forensic analysis on scanners and scanned images [J]. IEEE Trans on Information Forensics and Security, 2009, 4(3): 476 - 491.

[12] CRISTIANINI N, SHAWE-Taylor J, An introduction to support vector machines (and other kernel-based learning methods) [M]. Cambridge University Press, 2000.

[13] CHOI Chang-Hee, LEE Min-Jeong. Scanner identification using spectral noise in the frequency domain[C] // IEEE International Conference on Image Processing, 2010: 2121 - 2124.

[14] KHANNA Nitin, MIKKILINENI A K, DELP Edward J, Scanner identification using feature-based processing and analysis[J]. IEEE Trans on Information Forensics and Security, 2009, 4(1): 123 - 139.

[15] 张春梅, 张太镒. 基于双密度双树复数小波变换的图像去噪[J]. 西安交通大学学报, 2006, 40(2): 4 - 8.

[16] CHANG C C, LIN C J. LIBSVM: A library for support vector machines[EB/OL]. (2001 - 04 - 17) [2013 - 03 - 13]. Available: <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.

[17] KEETTHI Sathiya. Training and testing low-degree polynomial data mappings via Linear SVM [J]. Journal of Machine Learning Research, 2010(11): 1471 - 1490.

[18] 尹京, 方艳梅. 数码翻拍图像取证算法[J]. 中山大学学报: 自然科学版, 2011, 50(6): 48 - 52.