

数字水印技术在印刷品防伪中的应用

宋玉杰 刘瑞祯 谭铁牛 王蕴红

(模式识别国家重点实验室,中国科学院自动化所,北京 100080)

摘要 防伪技术一直是人们关注的一个热点,针对印刷品盗版特定场合防伪存在成本较高的问题,在分析印刷扫描给电子图象带来误差的基础上提出了一种用于印刷品防伪的数字水印技术.该方法中水印的添加采用了线性运算,而水印的提取则采用了阈值判断法;在提高算法的鲁棒性方面,水印的构造采用了简单但非常有效的重复构造法.另外,为了保证实验结果的稳定性,算法还采用了 CIElab 色彩空间.实验结果表明,在对参数优化后,该算法能较好地满足防伪的一般要求,同时与其他防伪技术相比,不仅大大减少了防伪成本,而且,该算法还可用于普通证件的防伪.

关键词 数字水印 防伪 版权保护

中图法分类号: TP309.9 文献标识码: A 文章编号: 1006-8961(2001)05-0450-05

Digital Watermarking for Forgery Detection in Printed Materials

SONG Yu-jie, LIU Rui-zhen, TAN Tie-niu, WANG Yun-hong

(National Lab of Pattern Recognition Institute of Automation, Chinese Academy of Sciences, Beijing 100080)

Abstract Forgery prevention and detection are of great social importance. This paper attempts to apply digital watermarking into forgery prevention. The digital watermarks are embedded into the most important DCT coefficients except the DC coefficient of the original image through linearly additive operation. The corresponding watermarks are extracted through threshold judgement and the threshold is determined through the difference of two cycle print&scan and one cycle print&scan of the images. To improve the robustness of the algorithm, the simple and effective repetition is applied to watermark configuration. And the CIElab color space is chosen to guarantee the repetition of the results and minimize the error brought by the transference of images between different equipments. Experimental results show that the proposed algorithm can satisfy the requirements of forgery prevention for ordinary images, that is to say the digital watermark is robust to one cycle print&scan and fragile to two cycle print&scan. Compared with other forgery prevention and detection techniques, the algorithm described in this paper reduced the computation cost considerably as no other special material or equipment is required and ordinary user can verify the authentication of printing materials. This algorithm can also be used in printing materials copyright protection, such as the owner verification in certificates, passports, etc.

Keywords Digital watermarking, Forgery prevention, Copyright protection

0 引言

随着当今网络技术的飞速发展,对多媒体的版权保护与内容鉴别提出了较高的要求,而电子水印技术就是应这种要求而产生并发展起来的,但不容

忽视的是,大量图象是用印刷方法产生的,如:商标、产品包装、书刊、证书、证件、邮票、磁卡、出版物、影音制品封面、货币、有价票券、入场券、单据、广告、挂历、名片、护照等,当然还有大量用印刷方法产生的文档(书籍、材料等).近些年来,由于随着商品经济的发展,假冒行为也随之而来,其中印刷品的假

冒也十分严重,虽然印刷技术保证了印刷品的同一性,但假冒者可按常规的作业方式轻易地复制这些印刷包装品,假冒行为屡屡得手,因此,这些用印刷方法形成的图象、文档的版权保护就成了急待解决的问题。尽管人们曾在彩色印刷上研发了不少防伪方法,如激光防伪、专用水印纸防伪、电话电码防伪、电脑网络防伪等等,在 1999 年北京国际博览会上又推出了几种新的技术^[1](特殊加网防伪技术、分期解密防伪技术、隐形标识防伪技术、隐形图象防伪技术),所有这些防伪方法虽在一定范围内起到了一定程度的防伪作用,然而产品的成本也随之大幅度地增加,如国内茅台酒厂每年要花大量资金从国外购进防伪标签,因此,人们渴望一种成本低廉的防伪方法。

数字水印技术是一个新兴的领域,在印刷领域也得到了应用,如目前最流行的软件 PhotoShop,其中就有由 Digimark 公司开发的添加水印功能,其嵌入的水印虽在一定程度上能够抵抗打印扫描,但加入的水印内容非常有限,仅仅限于 Digimarc ID、Distributor ID 和 Copyright Year,而且对于几何变换非常敏感,其目前主要用于版权保护。AlpVision 公司开发的 PhotoCheck 软件^[2]则用于护照的真伪鉴别。在护照中,由于水印是加在护照持有人的照片中,如果照片被换掉(即伪造),水印自然就丢失了,因此是一种基于内容的水印版权保护算法。Lin 曾利用傅立叶变换所具有的旋转、剪切、平移等特点,较早开展了脆弱性水印的研究,并研究了抵抗印刷扫描的水印算法^[3,4]。李忠源等采用 Photoshop 本身所带的功能进行了印刷品信息隐藏方面的研究^[5],但尚缺乏理论方面的根据。

由于印刷品防伪对于数字水印技术有一定要求,因此本文先做一个假定,即盗版者是通过扫描已发行的正版产品,然后再进行印刷、销售,来实现盗版的。图版 II 中的图 1 和图 2 给出了水印的添加与提取示意图,从图中可以看出,在水印提取过程中,对盗版产品而言,由于盗版过程的存在,使得水印提取前比正版产品多了一轮印刷扫描,而一般防伪技术的要求是对正品与盗版品要有不同识别方式,因此,如果将水印技术用于防伪的话,则要求从正品中提取出的水印与从盗版品中提取的水印有明显区别。由于数字水印一个非常重要的测试指标就是它的鲁棒性,因此,水印能够实现防伪的第 1 步就是要求水印对一轮印刷扫描是鲁棒的,第 2 步则要求对

两轮印刷扫描是脆弱的,而实现印刷品的版权保护只需满足第 1 步要求即可。从国内外公开发表的文献来看,虽有一些学者已经开始试图把数字水印技术用于印刷品的版权保护,但将水印技术用于防伪尚未见公开报道,而且所有这些算法均仅于抵抗一轮印刷扫描^[2~5],故仅仅实现了防伪的第 1 要求。本文提出了用数字水印技术来解决印刷品防伪的方案,并取得了较理想的结果。

1 实现方法

1.1 印刷扫描过程失真分析

水印的鲁棒性和脆弱性是与其印刷扫描过程中的失真紧密联系在一起的,而且在印刷扫描过程中,所有图象的失真均可归结为如下 4 类:一是人眼可见像素点的失真,这种失真主要包括亮度对比度的变化、gamma 修正、图象的半色调、颜色变化、点扩散、点检测光学扩散噪声及边缘高阶噪声等带来的失真,二是人眼可见的几何失真,主要有由旋转、尺度变化及剪切等带来的失真,三是由实验的客观条件带来的随机误差;四是由于在对原稿进行印刷和对印刷品进行扫描过程中包含了数字模拟之间的转换,且同一图象在不同设备上具有不同色彩空间描述方式,而造成的失真。这是因为各个色彩空间色域的不同,所有图象在不同空间转换时,不可避免地要丢失信息。

对于前 3 类失真,只能依靠算法和控制客观实验条件来减少,而对于第 4 类失真,则可通过色彩空间的选择来减少。在目前诸多的色彩空间中,如计算机图形学中的色彩空间 RGB、CMY、HIS (HLS、HSB) 等,以及彩色电视色彩空间 YUV、YIQ、YCr-Cb 等,还有基于 CIE 的色彩空间 CIE_{xyY}、CIE_{Lab}、CIE_{Luv} 等,其中只有基于 CIE 的色彩空间是独立于设备的,由于其中 CIE_{Lab} 色彩模式与 YUV 色彩空间类似,且均是用亮度和色差来描述色彩分量,因此它是目前唯一既能直观感觉,又能用于参数描述的色彩空间,它不仅能客观准确地描述、传递、管理与复制色彩,而且能独立于设备。由于 CIE_{Lab} 可以保证在进行色彩模式转换时,CMYK 范围内的色彩没有损失,因此,本文采用色彩空间最广的 CIE_{Lab} 色彩模式来作为转换的基础,而且水印的添加与提取都是在 Lab 空间进行的。

1.2 水印添加与提取

添加水印时,先将原始图象转化为 CIELab 描述方式,即仅将其中的亮度分量加入水印;然后对水印进行构造,这里以加入水印“Tan tieniu”为例(水印长度记为 l)来介绍如何构造水印:首先将每个字符所对应的 ASCII 码各位取出,并依次排为一 $\{0, 1\}$ 序列,为了提高算法鲁棒性,需将该序列编码为 $\{-1, 1\}$ 序列,并重复该序列 m 次,以得到总长度为 $8lm$ 的序列 $\{w_i\}$;然后对原图取 DCT 变换,再找出 DCT 系数中最大的前 $8lm$ 个 d_i ,接着再将水印通过线性运算 $\tilde{d}_i = d_i(1 + \alpha w_i)$ 加入 DCT 系数中^[6],其中,水印强度由 α 控制;最后进行 IDCT 即可得到添加水印的图象,再印刷该图象即是所需印刷品。

提取水印时,首先扫描任一份印刷品,将扫描后的图象处理成与原图大小相同的图;然后将扫描图象转化为 CIELab 描述方式,并提出其中的亮度分量;接着进行相反的减法运算 $w'_i = (\tilde{d}'_i / d'_i - 1) / \alpha$,即得到一序列 $\{w'_i\}$,设定一域值 M ,如果 w'_i 大于 M (M 的选择将在 3.3 节论述),则认为提出的为 1,否则就认为提出的为 -1. 以字母 T 的第 8 位 T_8 为例,如果提出的 1 的个数大于 -1 的个数,则认为 $T'_8 = 1$,否则, $T'_8 = 0$. 用同样的方法还可以提出字母 T' 的第 7、第 6、 $\dots$ 、第 1 位,将 T' 的各位信息结合起来(即字符 T' 的 ASCII),就得到对应于加入字符 T 所提出的字符 T';最后将提出的每个字符结合起来,即是所提出的水印,并且可以同时计算出每一位的正确率,以 $T_8 = 1$ 为例, T_8 的位错误率 = (所有 $T'_8 = 0$ 位的个数) / (重复次数 m).

1.3 参数优化

由于通道容量是一个与水印鲁棒性及水印的不可察觉性紧密联系在一起的重要因素,其鲁棒性要求容量尽可能地多,而不可察觉性则要求加入的容量有一个上限,因此三者之间是相互矛盾的. MIT 实验室的 Smith 和 Comiskey 从信息论的角度提出了一个简单的基于静止图象的数据隐藏系统信道模型^[7],即原始图象认为一个近似连续的二维带限通道,它具有较大的平均噪声能量,而隐藏信息则被认为是通过这一通道的传递信号. 从这一模型出发,并以高斯噪声通道为基本信道模型,就可以计算出在给定带宽的情况下,为达到信道容量 C 所需的 S/N (信号与噪声能量之比)的下限. 在低信噪比 S/N 的情况下,信道容量与系统的信噪比成正比. 该模型有局限性,因为它仅假设信息是在空间域叠加.

Ramkumar 等在以上模型的基础上进行了改进^[8],并设计了在加噪条件下的数据隐藏模型,尽管该模型关于信息容量的估计仍有许多不足之处,但已为容量的估计提供了一个非常好的思路. 在图象质量评价指标 PSNR 选 38dB,水印强度系数 α 选 0.2 的情况下,对于一幅 256×256 的灰度图象,可以最多加入长度为 1 000bit 左右的水印序列.

对于一次印刷扫描的图象,如果水印序列每个字符的各位都可以准确提出(或各位的正确率都在 50% 以上),那么就认为水印对一次印刷扫描是鲁棒的;如果对于二次印刷扫描的图象,有相当一部分水印序列的字符无法提出,那么就认为水印对于二次印刷扫描是脆弱的. 由此可见,在水印的提取过程中,错误率阈值的选择是非常重要的. 由于现在盗版技术非常高超,印刷机扫描仪精度也越来越高,经过一次印刷扫描的图象与经过二次印刷扫描的图象从视觉上来看,差别非常小. 如果以上方案可行的话,则必须找出一次与二次印刷扫描的差异.

从多次实验结果(100 次)可看出,如果印刷条件基本相同的话,那么从正品扫描后的图象中提出水印序列的均值 m 基本稳定在某一值附近,且对于每次实验, m 值上下浮动 $-0.1 \sim 0.1$,将 100 次实验中计算的各个 m 的均值记为 m_1 ;而对盗版产品扫描,从印刷后的图象中提出序列的均值 m 也基本稳定在某一值附近,而对于每次实验,其 m 值上下浮动 $-1 \sim 1$,将 100 次实验中计算的各个 m 的均值记为 m_2 . 通过多次实验结果发现, m_2 与 m_1 的差值要明显大于 m_1 和 m_2 本身的波动. 因此在提取时其阈值选 m_1 .

2 实验结果

2.1 Stirmark 3.1 鲁棒性测试

首先用目前比较流行的测试软件 Stirmark 3.1 对算法的鲁棒性进行测试,测试结果表明,本文算法不仅对 JPEG 压缩、尺度变化、 x - y 轴显示比例变化、对称及非对称地移去图象行和列等操作的鲁棒性非常好,而且可以抵抗一定的滤波方法(如 3×3 的中值滤波)、抵抗很小程度的中心裁剪(1%~2%)、抵抗一定程度的旋转(带裁剪,带尺度或不带尺度变换)($-0.75^\circ \sim 0.75^\circ$)及 x - y 方向的修剪(1%~2%)所造成的影响;但对通用线性几何变换则非常脆弱. 表 1 列出了算法对 Stirmark 3.1

攻击正确检测个数. 该算法的鲁棒性能够满足一轮印刷扫描的要求.

表 1 算法对 Stirmark 3.1 攻击正确检测的个数

测试类型	测试数目	正确检测数目
对称及非对称地移去图象的行和列	5	5
滤波(中值、高斯、FMLR、锐化)	6	5
JPEG 压缩	12	12
中心裁剪	9	2
通用线性几何变换	3	0
改变 x - y 轴的显示比例	8	8
带裁剪、不带尺度变换的旋转	16	6
带裁剪和尺度变换的旋转	16	6
尺度变换	6	6
x - y 方向的修剪	6	3
Stirmark 随机弯曲	1	0
合计	88	53

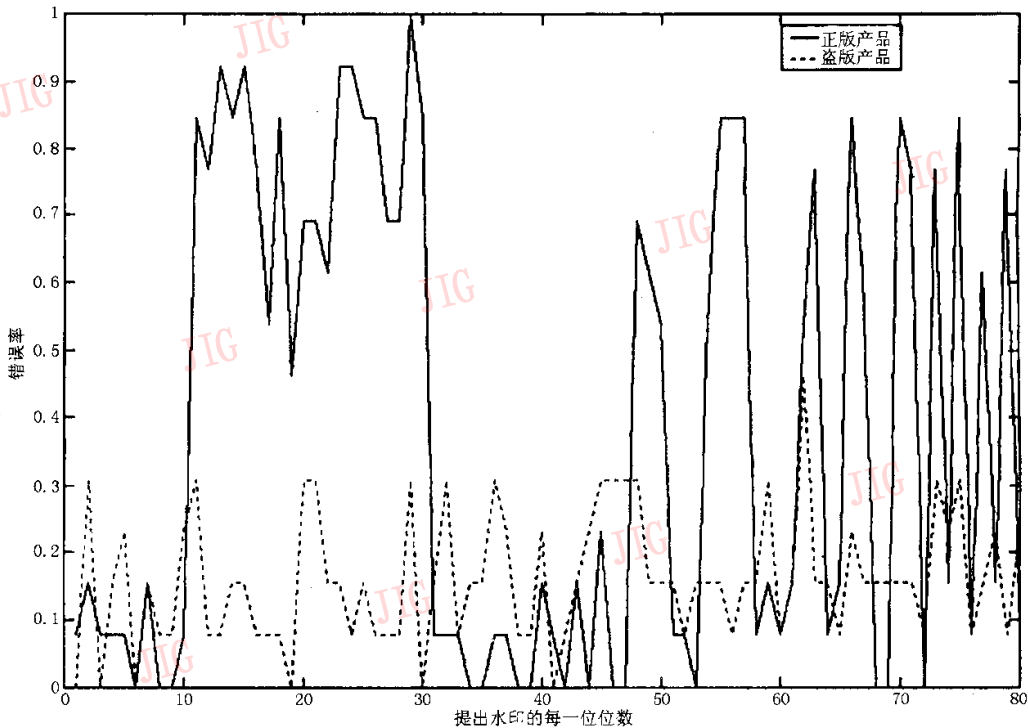


图 1 盗版产品与正版产品提取水印位错误率的比较

3 结论及问题

打假防伪是人们广泛关注的社会问题,而本文则试图利用数字水印技术来解决印刷品防伪难题. 由实验结果可见,本文提出的数字水印防伪方案能够满足一般图象防伪的基本要求,而且还可以用于仅仅对一轮印刷扫描有鲁棒性要求的场合,如身份

2.2 水印提取

图 1 给出了在 100 次重复实验中一组添加水印“Tan tieniu”时的典型实验结果. 图中横坐标是水印每个字母所对应 ASCII 码的位数,纵坐标是各位所对应的错误率. 该图给出了从正版产品与盗版产品提取出的水印位错误率对照. 从图 1 中可清楚地看出,从正版产品与盗版产品中分别提取出的水印位错误率有明显差异,由此可见,本文方法能满足防伪的要求. 对于提出的每一位水印信息,如果错误率在 35% 以下,就认为是正确提取;否则认为是错误提取. 由图 1 可见,从正版产品中提取水印的错误率都在 35% 以下,因此水印完全能够提取出来,而盗版产品则仅仅可以从其中提出一部分水印(50%左右). 为此可以定义一个门限,即如果能够提出水印的 85%,则认为是正品;否则就被怀疑为盗版产品.

证,护照等的防伪. 与其他防伪技术相比,该算法最突出的优点表现在它不需要特殊材料或特殊工艺等,因而真伪鉴别的成本大大降低,一般用户可以很容易地实现产品的真伪鉴别. CIElab 空间的选择保证了数据在不同的设备之间转换时的完整性. 为检验算法的鲁棒性,重复实验是在同样印刷扫描设备、同样印刷扫描条件下,采用同一幅图象、同样水印、同样参数设定而进行的,若设备或条件发生变化时,

参数的优化会有所不同. 当然水印技术在印刷品防伪领域的应用研究仍有大量工作要做, 如果把色差的因素考虑在内, 就可以加入更多的水印, 这样图象的视觉评价指标也将有所变化等等.

参 考 文 献

- 1 陈建猛, <http://baozhen.363.net/jj.htm>, Beijing, 1999.
- 2 AlpVision, <http://www.alpvision.com>, Switzerland, 2000.
- 3 Lin C Y, Chang S F, Distortion modeling and invariant extraction for digital image print-and-scan process. In: Intl. Symp. on Multimedia Information Processing (ISMIP 99), Taipei, Taiwan, Dec. 1999.
- 4 Lin C Y. Public watermarking surviving general scaling and cropping: An application for print-and-scan process, In: Multimedia and Security Workshop at ACM Multimedia 99, Orlando, FL, Oct. 1999.
- 5 李忠源, 付震中. 印刷图象中的信息隐含. 第二届信息隐藏学术研讨会论文集, 北京, 2000; 34~36.
- 6 Cox I J, Kilian J, Leighton T *et al.* Secure spread spectrum watermarking for images, audio and video. In: International conference on image processing (ICIP'96), (III) 1996; 243~246.
- 7 Smith J R, Comiskey B O. Modulation and information hiding in images. In: Workshop on Information Hiding, University of Cambridge, UK, May 1996; 463~470.
- 8 Ramkumar M. Data hiding in multimedia-theory and applications (Ph. D. Dissertation), Jan. 2000.

宋玉杰 1975 年生, 1999 年进入中科院自动化所模式识别室攻读博士学位. 研究兴趣包括多媒体真伪鉴别、数字水印、图象编码等领域.

刘瑞祯 1969 年生, 1998 年进入中科院自动化所模式识别室读博士. 研究兴趣包括信息隐藏和数字水印、图象处理与编码、计算机视觉等领域.

谭铁牛 1964 年生, 研究员、博士生导师. 主要研究领域为图象处理、计算机视觉和模式识别等. 发表论文 100 余篇, 获准和申请专利 5 项.

王蕴红 1968 年生, 副研究员, 1995 年和 1998 年先后获南京理工大学硕士与博士学位. 研究领域为模式识别、信号处理、神经网络及计算机视觉. 目前已发表文章 20 多篇, 申请专利 3 项.

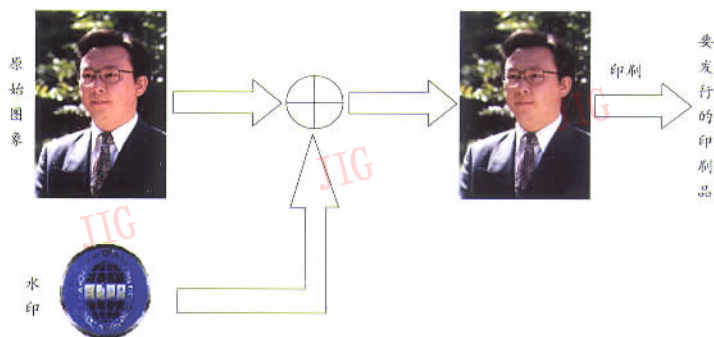


图1 水印添加过程

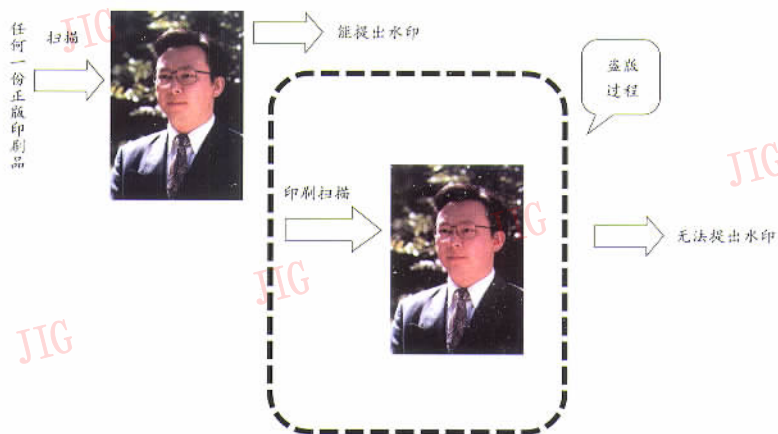


图2 水印提取过程