

# 印刷纸介质文档中的信息隐藏技术研究

国伟<sup>1</sup>, 刘宇鑫<sup>2</sup>, 杨斌<sup>2</sup>, 齐文法<sup>2</sup>

(1. 中国检验检疫科学研究院, 北京 100123; 2. 北京大学 计算机科学技术研究所, 北京 100871)

**摘要** 随着信息技术和数码图像复制设备的飞速发展, 纸介质文档中的信息安全面临很大的挑战, 如何解决纸介质文档中的信息安全问题已成为众多学者关注的研究课题之一。本研究提出一种基于网点调制底纹的纸质印刷文档中信息隐藏的新方法, 该底纹由大量随机排列的印刷网点构成, 包括编码点、定位点以及视觉干扰点。通过编码点特定位置的偏移来隐藏信息, 采用定位点来探测编码块和相邻编码点偏移前的初始位置, 同时通过干扰点的随机扰动保证隐藏信息的安全性和视觉隐蔽性。实验结果表明, 利用此方法设计出的网点调制底纹隐藏的信息量大, 并且对打印扫描攻击具有较强的鲁棒性。该技术可广泛用于防伪印刷、防复印、文档内容防篡改以及文档图像检索等应用领域。

**关键词** 防伪印刷; 信息隐藏; 网点调制

中图分类号 TP391; TS853+.6

文献标识码 A

文章编号 1674-5752(2013)02-30-05

## Research on Information Hiding Technology in Paper-Based Document

GUO Wei<sup>1</sup>, LIU Yu-xin<sup>2</sup>, YANG Bin<sup>2</sup>, QI Wen-fa<sup>2</sup>

(1. Chinese Academy of Inspection and Quarantine, Beijing 100123, China; 2. Institute of Computer Science and Technology, Peking University, Beijing 100871, China)

**Abstract** With the rapid development of the information technology and digital image copy devices, the paper media still plays an important role in many areas and correspondingly the information security of paper media documents are facing great challenges. So how to solve the information security problems of paper media has become one of the hottest research topic. In this study, an information hiding method used in printed documents image was proposed based on the modulating shade composed of a large number of randomly arranged screen dots. These dots included the coding dots, anchoring dots and perturbing dots. Experimental results showed that with the new method, screen dots modulating shade has larger information capacity and also has more strong robustness against the attacks of printing and scanning processes. Moreover, this technology also can be widely used in many other areas, such as anti-counterfeit printing, anti-copy, document content tamper-proofing and document image retrieval, etc.

**Key words** Anti-counterfeit printing; Information hiding; Screen dots modulating

## 0 引言

当今社会, 随着信息技术的发展, 诸如电子书、存储器等新型数据记录介质虽然不断涌现, 但纸介质仍

凭借其原始性、便捷性, 在信息交换领域具有不可替代的作用。而随着高速计算机、高分辨率打印机和扫描仪等的出现, 载有机密信息的文档、证件、票据、有价证券等可以很容易地被伪造复制, 这必然会给个人、企业乃至国家造成巨大损失。因此, 如何解决纸

介质文档的信息安全问题以防止相关文件被伪造盗用已成为备受学者关注的研究课题之一。

数字水印技术可以有效地应用于防伪印刷领域<sup>[1-4]</sup>，但是用于水印嵌入的印刷图像中大部分为网目调图像，其水印信息容量十分有限。为了实现在印刷防伪的同时提高水印信息容量，顾泽仓教授提出利用几何学或物理学形态不同的网点来实现信息嵌入的网屏编码技术<sup>[5-6]</sup>。该技术将网点组合起来称为网屏，以网屏作为载体可实现大信息量的嵌入，从而实现纸介质文档的印刷防伪<sup>[7]</sup>。网屏编码技术理论上可以在一张 A4 幅面的纸张上隐藏 38000 个汉字，但在实际应用中，考虑到打印和扫描误差，可嵌入的水印信息容量仅为理论值的 1/4<sup>[8]</sup>，因此利用网屏编码技术可实际嵌入的信息量不足 10000 个汉字。

为了进一步提高底纹载体中可嵌入的水印信息容量，并且增强其在经过打印扫描攻击后的鲁棒性，同时保证生成的底纹具有美观的视觉效果，本研究提出了一种基于网点调制底纹的印刷文档中信息隐藏新方法。其中底纹由大量网点构成，并通过编码点的相对位置偏移实现水印信息的嵌入。同时，通过信息提取实验，论证了该方法提升水印信息容量的可行性。

## 1 基于网点调制底纹的信息隐藏技术

### 1.1 信息隐藏原理

构成底纹的网点可以分为定位点、编码点及视觉干扰点三种类型。其中定位点在信息嵌入的过程中位置始终保持不变。因此，通过探测定位点的位置便可确定编码图像块（即单元底纹图像块）的起始位置及相邻编码点在偏移前的初始位置。编码点根据所要嵌入信息的不同向 8 个方向进行位置偏移。视觉干扰点不代表任何水印信息，其位置可以随机扰动，一方面改善视觉效果，另一方面也打乱了点阵排布的规律，从而保证了信息隐藏的安全性和视觉隐蔽性。

如图 1 所示，中间的黑点为编码点偏移前的初始位置，一个编码点可向周围 8 个邻域方向进行位置偏移，每个偏移方向可代表三位水印信息，例如向左偏移代表水印信息“101”。一个字节可使用三个编码点来表示，其中最后一位可作为奇偶校验位，用于位串

信息的校验，判断位串信息的对错。字节流的信息嵌入方式如图 2 所示。

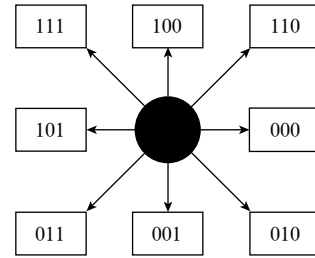


图 1 编码点的偏移方向  
Fig. 1 Deflection direction of the coding dot

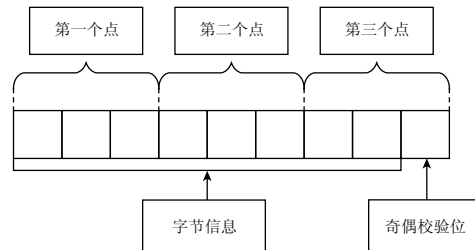


图 2 信息字节流的嵌入方式  
Fig. 2 Embedding method of the information byte data

### 1.2 信息嵌入方法

在信息嵌入前编码图像块中的网点为规整矩阵式排列，如图 3 所示。假设  $G[i, j]$  为点阵中位于第  $i$  行第  $j$  列的网点，如果它满足式(1)，则被视为一个定位点，否则被视为编码点。

$$(i\%4 = 1)\text{and}(j\%4 = 1)\text{or}(i\%4 = 3)\text{and}(j\%4 = 3) \quad (1)$$

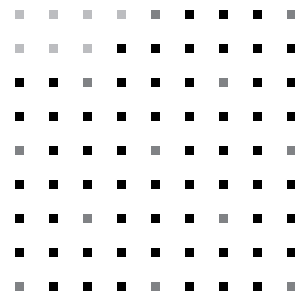


图 3 原始网点为矩阵式排列  
Fig. 3 Original dots in matrix arrangement

在图 3 中，深灰色网点为定位点，其余的黑色网点为编码点，网点排列时水平或垂直方向上每两个定位点之间存在 3 个编码点。令每个完整编码图像块网点行数及列数均为 4 的倍数，则定位点和编码点所占比例分别为 1/8 和 7/8。最后，在编码图像块起始位置处排列固定形状的网点图案作为信息嵌入图像块的起始标志，如图 3 中左上角浅灰色标识的网点图案。

在信息嵌入过程中,根据待嵌入水印位串信息,按照从上到下、从左到右的顺序依次对编码点进行位置偏移,偏移方式遵循图1所示规则。另外在间距过大的网点位置中间增加额外的网点作为视觉干扰点,从而改善编码后图像块的整体随机视觉效果,如图4中的灰色标识网点即为额外增加的视觉干扰点。考虑到信息成功识别率,视觉干扰点的位置不能与编码点的偏移位置发生冲突,最终嵌入信息后的网点排列效果示意图如图4所示。

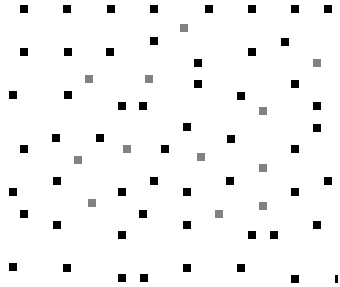


图4 信息嵌入后的点阵排列效果  
Fig.4 Result image with information embedded

信息嵌入前需要先确定整体编码图像块的大小,而一个完整的单元编码图像块而言,其大小随着待嵌入水印信息量的变化而变化。假设  $N$  为嵌入的有效字节数,  $k$  为编码图像块起始位置处的定位点所占用的可编码字节数,  $M$  为一个最小的单元编码图像块应包含的网点的行数或列数。记  $n = FIXED\_REDUNDANCY + N$ , 其中  $FIXED\_REDUNDANCY$  为冗余信息量,比如  $CRC$  校验码或特定标识码等。令一个单元编码图像块内网点的行数与列数比值为  $R$ , 则编码图像块的列数  $f(x)$  的计算如公式(2)。

$$f\left(\frac{\sqrt{24(n+k)/7}}{R}\right) = \begin{cases} \frac{\sqrt{24(n+k)/7}}{R}, \frac{\sqrt{24(n+k)/7}}{R} \% 4 \\ = 0 \text{ and } \frac{\sqrt{24(n+k)/7}}{R} \geq M \\ 4 \times (\text{int})\left(\frac{\sqrt{24(n+k)/7}}{4R} + 0.75\right), \\ \frac{\sqrt{24(n+k)/7}}{R} \% 4! \\ = 0 \text{ and } \frac{\sqrt{24(n+k)/7}}{R} \geq M \\ M, \frac{\sqrt{24(n+k)/7}}{R} \leq M \end{cases} \quad (2)$$

若原始点阵中每两个相邻编码点网点中心距离为  $P$

个像素,则编码图像块的宽度  $X(N)$  和高度  $Y(N)$  分别可由公式(3)计算得到。

$$\begin{cases} X(N) = P \times f\left\{\frac{\sqrt{24(n+k)/7}}{R}\right\} \\ Y(N) = 4 \times (3 + (\text{int})\frac{R \times X(N)}{4}) \end{cases} \quad (3)$$

## 2 信息提取方法

信息提取的原理是通过搜索每个单元编码图像块内所有的定位点,进而确定编码点偏移前的初始位置  $O$ ;接着以每个编码点的初始位置  $O$  为中心,分别在其周围8个方向进行搜索进而确定编码点偏移的方向,并根据图1所示的偏移规则还原出原始嵌入的水印信息位串。

为了在已知网点的基础上搜寻到其余未知网点,需要确定信息嵌入前相邻网点的中心间距  $D$ 。由于图像经过打印扫描后,会发生旋转、缩放等变化,因此在信息提取时,需要先对扫描图像进行图像纠偏或缩放等预处理工作。具体的信息提取处理过程如下。

步骤1:由于定位点在信息嵌入过程中位置始终保持不变,根据点阵的排列规律,存在着大量由上、中、下、左和右5个定位点构成的最小对称“+”形状。搜索构成该形状的定位点,并将其位置坐标分别记为  $U(X_U, Y_U)$ ,  $C(X_C, Y_C)$ ,  $B(X_B, Y_B)$ ,  $L(X_L, Y_L)$  和  $R(X_R, Y_R)$ 。

步骤2:在每个特定的“+”形状图案中,  $U$  和  $B$  两点间相距均为  $D$  的8倍,则相对于原始点阵中网点中心间距为  $D$  的两个网点而言,水平方向上的偏移误差向量  $(X_{HR}, Y_{HR})$  可分别表示为:

$$\begin{cases} X_{HR} = \frac{X_R - X_L - 8D}{8} \\ Y_{HR} = \frac{Y_R - Y_L}{8} \end{cases} \quad (4)$$

垂直方向上的偏移误差向量  $(X_{VR}, Y_{VR})$  分别为:

$$\begin{cases} X_{VR} = \frac{X_B - X_U}{8} \\ Y_{VR} = \frac{Y_B - Y_U - 8D}{8} \end{cases} \quad (5)$$

步骤3:假设一个搜索到的定位点坐标为  $(X_P, Y_P)$ , 相邻两个定位点间的距离为  $D$  的4倍,则该定位点右侧和下侧另外两个相邻定位点的搜索中心位置分别为:

$$O_1(X_p + 4(D + X_{HR}), Y_p + 4Y_{HR}) \quad (6)$$

$$O_2(X_p + 4X_{VR}, Y_p + 4(D + Y_{VR})) \quad (7)$$

步骤4：根据公式(6)和(7)可以找到所有的定位点，因此可以确定其附近的编码点偏移前的位置，进而根据图1所示的网点位置偏移规则确定隐藏的水印位串信息。

### 3 试验结果与讨论

在本实验中，采用 HP LaserJet 1200 打印机进行纸质文档的打印输出，通过 HP ScanJet 4890 扫描仪扫描纸质文档从而获得文档的灰度图像数据，打印和扫描的分辨率同为 600dpi。单元编码图像块的分辨率也为 600dpi，其中每两个相邻网点中心间距为 16 个像素，相应的每个编码点的尺寸大小为 2 像素×2 像素，编码点在水平和垂直方向上的偏移距离为 4 个像素。

在一张 A4 页面文档中，分别嵌入 300B、500B、3KB、10KB、20KB、26KB 和 32KB 的有效信息，经过打印扫描后的信息提取结果见表 1 所示。

表 1 打印扫描后的测试结果  
Tab.1 Test results after printing and scanning

隐藏的信息量	识别正确率 (%)
300B	100
500B	100
3KB	99.9685
10KB	99.9713
20KB	99.9891
26KB	99.9523
32KB	99.9432

由表 1 可以看出，在没有纠错码和重复嵌入的情况下，信息提取识别正确率达到 99.9% 以上。整幅 A4 (4832 像素×6888 像素) 页面可嵌入的最大信息量计算如下：

编码块的列数： $4832/16 \approx 300$  (取能被 4 整除的最小整数)

编码块的行数： $6888/16 = 428$

最大可隐藏的信息量为： $(300 \times 428 - 7) \times 7/8/3/1024 \approx 36.5\text{KB}$

36.5KB 相当于大约 18000 个汉字，考虑打印扫描过程的误差容错处理，以及需要额外增加校验码和纠错码

等冗余信息以提高信息的正确识别率，在实际应用中可以隐藏大约 32KB 左右的有效信息，相当于大约 16000 个汉字的信息量。

### 4 结论

目前，关于纸介质文档的安全问题已经是一个全球关注的问题。每年因机密文件的泄密和伪造带来的损失是巨大的。基于纸介质文档的信息隐藏技术，将需要隐藏的信息嵌入底纹图像中，凭借其隐藏信息的安全性、生成底纹的难以复制性等特点，可被广泛应用于文件追踪、防复印以及证件防伪等领域<sup>[9]</sup>。

本研究提出了一种基于网点调制底纹的印刷文档图像中的信息隐藏方法。通过编码点特定位置的偏移来隐藏信息，采用定位点来探测编码块和相邻编码点偏移前的初始位置，同时通过干扰点的随机扰动保证隐藏信息的安全性和视觉隐蔽性。通过该方法可以实现大信息量的水印嵌入，并且对打印扫描等操作具有较强的鲁棒性。这种信息隐藏方法有很大的应用前景，比如在办公安全打印领域中，文件正常打印的同时在正文内容下面附加打印背景底纹，其中隐藏有重要的文件溯源追踪信息(文件打印的机器名、用户名、打印机名称、打印时间等)或者特殊的密文信息。截获文件原件或者复印件后，通过普通扫描仪和专门的识别程序提取事先嵌入的信息，可以很快地追溯到文件非法传播的源头，以便追究相关人员和组织的责任，从而防止泄密事件的发生。

但是就其技术本身而言，还需要从以下几个方面进行深入研究：1) 如何进一步提高信息容量，即用最少的网点位置偏移代表更多的信息；2) 在基于方形块的信息嵌入方案中缺少了视觉干扰点，在一定程度上降低了随机底纹的视觉效果，如何在保证大信息量嵌入的前提下，进一步改进视觉效果；3) 如何将网点底纹图像自然地融合到网目调背景图案中，达到更加理想的视觉隐藏效果。

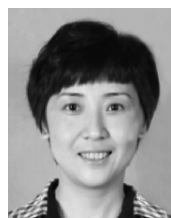
### 参考文献

[1] 郑海红, 王义峰, 曾平, 等. 基于视觉特性的半色调图像水印算法[J]. 吉林大学学报: 工业版, 2006, 36(5): 782 - 786.

- ZHENG Hai-hong, WANG Yi-feng, ZENG Ping, et al. Watermarking Algorithm for Halftone Images Based on Human Visual System [J]. Journal of Jilin University: Engineering and Technology Edition, 2006, 36(5): 782 - 786.
- [2] MIZUMOTO Tadashi, MATSUI Kineo. Robustness Investigation of DCT Digital Watermarking for Printing and Scanning [J]. Electronics and Communications in Japan, 2003, 86(4): 11 - 19.
- [3] MING S F, OSCAR C A. Data Hiding Watermarking for Halftone Images [J]. IEEE Transaction on Image Processing, 2002, 11(4): 477 - 484.
- [4] FU M S, AU O C. A Robust Public Watermark for Halftone Images [C]// Proceedings of IEEE International Symposium on Circuits and Systems, Washington, USA: IEEE Computer Society, 2002, 3: 639 - 642
- [5] 顾泽苍. 在印刷媒体上埋入信息的方法: 中国, 200610014468 [P]. 2008 - 01 - 02.  
GU Ze-cang. A Method that Information is Hidden in the Print Media: China, 200610014468 [P]. 2008 - 01 - 02.
- [6] 顾泽苍. 网屏编码证件防伪处理方法: 中国, 200610013990 [P]. 2007 - 12 - 05.  
GU Ze-cang. An Anti - Fake Method of Certificate Based on Screen Coding: China, 200610013990 [P]. 2007 - 12 - 05.
- [7] 赵立龙, 顾泽苍, 方志良, 等. 一种基于视觉特性及形态网屏编码的纸介质信息防伪方法[J]. 光电子·激光, 2008, 19(11): 1526 - 1529.  
ZHAO Li-long, GU Ze-cang, FANG Zhi-liang, et al. An Anti - Fake Method Based on Visual Characteristics and Morphology Screen Coding [J]. Journal of Optoelectronics · Laser, 2008, 19(11): 1526 - 1529.
- [8] 贾凤美. 网屏编码信息安全技术的研究[D]. 天津: 南开大学, 2007.  
JIA Feng-mei. Researches on the Information Security Technology Based on Screen Coding [D]. Tianjin: Nankai University, 2007.
- [9] 国伟, 杨斌, 亓文法, 等. 图像浮雕算法及其在防伪印刷中的应用[J]. 中国印刷与包装研究, 2011, 3(1): 24 - 28.  
GUO Wei, YANG Bin, QI Wen-fa, et al. Image Relief Algorithm and Its Application in Security Printing [J]. Chi-

na Printing and Packaging Study, 2011, 3(1): 24 - 28.

### 主要作者



国伟 (1978 年 -), 博士, 副研究员; 主要研究方向为食品包装防伪技术、包装 CAD 和防伪印刷。

Ph. D. GUO Wei, born in 1978. She is an associate professor and her research interests mainly include food packaging anti - counterfeit technology, packaging CAD and security printing.



刘宇鑫 (1985 年 -), 硕士, 工程师; 主要研究方向为数字图像处理、信息安全和安全印刷。

LIU Yu-xin, born in 1985. He got the master degree and now is an engineer. His research interests mainly include digital image processing, information security and security printing.



杨斌 (1969 年 -), 硕士, 研究员; 主要研究方向为数字图像处理、信息安全和安全印刷。

Professor YANG Bin, born in 1969. He got the master degree and his research interests mainly include digital image processing, information security and security printing.



亓文法 (1978 年 -), 硕士, 高级工程师; 主要研究方向为数字图像处理、信息安全和安全印刷。

QI Wen-fa, born in 1978. He got the master degree and now is a senior engineer. His research interests mainly include digital image processing, information security and security printing.

E-mail: qiwenfa@pku.edu.cn