

半色调图像中数据隐藏算法

牛少彰, 钮心忻, 杨义先, 胡文庆

(北京邮电大学信息安全中心, 北京 100876)

摘 要: 在许多打印和出版应用中, 需要在半色调图像中嵌入数据用于作品的版权保护和认证. 本文提出了一种半色调图像中数据隐藏的算法, 该算法利用 DCT 的中频系数比较法进行数字水印信息的嵌入. 在对数字水印信息进行嵌入前, 首先进行编码, 转化成比特流, 然后通过改变 DCT 中频系数的序关系嵌入数字水印, 经打印/扫描后提取出水印, 实验结果显示: 该算法能够抵抗打印/扫描攻击, 并且隐藏效果好, 恢复水印时不需要原始图像. 该算法也可用于普通证件的防伪.

关键词: 打印/扫描; 数字水印; 半色调图像

中图分类号: TP3091.9 **文献标识码:** A **文章编号:** 0372-2112 (2004) 07-1180-04

Data Hiding Algorithm for Halftone Images

NIU Shao2zhang, NIU Xin2xin, YANG Yi2xian, HU Wen2qing

(Information Security Center, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: In many printer and publishing applications, it is desirable to embed data in halftone images for copyright and authentication purposes. In this paper, a novel data hiding algorithm for halftone images is proposed. The digital watermarking is embedded by DCT mid2frequency coefficients comparison method. Before embedding, the digital watermarking is transformed into binary data, and then they are embedded into DCT mid2frequency subbands by changing the order of these coefficients. Experiments show that this algorithm is robust for printing and scanning, and the original picture is not needed in the recover progress of the digital watermarking. This algorithm can also be used for forgery detection in certificates.

Key words: printing/scanning; digital watermarking; halftone images.

1 引言

数字化技术和 Internet 的发展正在改变着文化传播的载体和方式, 各种数字化产品和模拟与数字化的转换产品目不暇接, 仅靠传统的密码技术来完成数据的加密、认证和保护工作是远远不够的. 数字水印作为数据安全领域的新生事物, 由于具有很高的技术含量和很强的生命力, 在版权保护、网络安全、数据防伪上具有不可替代的作用.

按照水印的嵌入方法可以将水印分为两类: 空间域水印和变换域水印. 空间域方法通过改变载体信息的空间域特性来隐藏水印; 最低有效位 (LSB) 法是最具代表性的一种^[1, 2], 其原理就是通过修改表示数字图像颜色 (或者颜色分量) 的较低位平面, 即通过调整数字图像中对感知不重要的像素低比特位来表达水印信息, 达到嵌入水印信息的目的. 变换域方法通过改变数据 (主要指图像、音频、视频等) 变换域的一些系数来隐藏水印. 它在图像变换域改变数据, 再进行反变换得到嵌入水印后的图像, 基于变换域的数字水印算法主要有:

离散余弦变换 (DCT)、离散傅立叶变换 (DFT) 和离散小波变换 (DWT) 等^[3~5].

在印刷品中进行信息隐藏是目前信息隐藏领域的一个新的研究应用方向. 通过印刷水印系统, 我们可以使用人眼看不见的水印形式将一些标识信息嵌入在印刷图像中, 以解决印刷品的版权保护、来源认证和防止伪造等问题.

数字图像大都由激光打印机输出, 激光打印机属于点阵打印机, 使用半色调技术用黑白点阵打印灰度图像, 目前激光打印机中常用的半色调技术是抖动法, 一般来说, 抖动法输出的图像只是大体上和原图一致, 局部细节上往往失真. 其次, 由半色调复合点的形状、激光束的扩散、纸张的吸水特性和光滑度等因素造成的半色调复合点变化 (包括复合点增益和曝光), 也常会导致输出图像变得模糊不清, 在使用扫描仪进行扫描的过程造成图像畸变. 因而印刷水印成为当前数字水印研究的一个难点, 这方面的文献相对较少. 文献^[6]给出的算法是在空间域上实现的, 由于空间域上像素值的改变较大, 使得提取过程比较复杂. 文献^[7]给出了在变换域上通过线性运算

实现的水印算法, 并指出通过多次的重复实验可以用来判断图像是经过一轮印刷扫描(正版) 还是经过两轮印刷扫描(盗版). 文献^[8]详细研究了打印扫描对图像水印的影响, 通过直方图给出了打印扫描前后图像 RGB 值的变化情况, 由于差异较大, 难以通过 RGB 值进行水印的检测, 但打印扫描对频域的干扰较小, 并指出在 DCT 域中嵌入水印鲁棒的条件为在 DCT 的中频系数上.

因此寻找合适的水印算法, 使得数字图像由激光打印机输出, 局部细节上失真, 使用扫描仪进行扫描过程中造成图像变化的情况下, 仍能检测到所嵌入的水印是在印刷图像中进行数据隐藏的关键. 我们通过将水印信息/ 放大0, 利用 DCT 的中频系数比较法提出了一种在印刷图像中嵌入数字水印的算法, 并在打印/ 扫描及在 Photoshop 软件中对扫描图像进行调整后, 能提取出隐藏的数据.

2 数字水印嵌入算法

离散余弦变换是一种典型的数字图像变换. 数字图像可看作是一个二元函数在离散网格点处的采样值, 可以表示为非负矩阵形式. 此外, 图像的内容通常具有某种程度的自相关性, 即图像局部的内容往往变化不大. 对数字图像进行离散余弦变换后, 可以充分利用图像的自相关性来减少数据量, 以达到压缩数字图像的目的. 目前流行的数字图像和视频压缩标准如 JPEG 和 MPEG 都以离散余弦变换为核心.

二维离散余弦变换定义如下:

$$F(u, v) = a(u) a(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos\left[\frac{(2x+1)uP}{2N}\right] \# \cos\left[\frac{(2y+1)vP}{2N}\right]$$

逆变换定义为:

$$f(x, y) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} a(u) a(v) F(u, v) \cos\left[\frac{(2X+1)uP}{2N}\right] \# \cos\left[\frac{(2Y+1)vP}{2N}\right]$$

其中 $a(0) = \sqrt{\frac{1}{N}}$ 且 $a(m) = \sqrt{\frac{2}{N}}$, $1 \leq m \leq N$,

$f(x, y)$ 为图像的像素值, $F(u, v)$ 为图像做 DCT 变换后的系数.

考虑到数字图像经打印和扫描后, 会有较大的几何失真, 虽然经过一定的处理后, 将这种失真尽我们所能降到了最低点, 但是失真依然存在, 为了使水印信息经过半色调打印及扫描仪扫描后仍能提取出来, 我们对原始图像按 8@8 像素分块, 并将水印信息/ 放大0, 即在每 8@8 像素分块中只包含一个比特的水印信息. 我们用 DCT 中频系数比较法来嵌入水印信息, 算法如下:

将要嵌入的水印信息记为 $b(j)$ ($j = 1, 2, \dots, N$), 其中 $b(j)$ 为 0 或 1. 将原始图像按 8@8 像素分块, 对每一块作 DCT 变换, 得到如表 1 所示的 DCT 系数矩阵, 其中每个格代表一个系数. 有字母的格子代表嵌入水印信息的位置.

我们将中频系数进行了分组, 如表 1 中所示, 每三个系数一组, 共分了七组, $b(j)$ 代表该块要嵌入的第 j 比特位水印信息. 水印数据的第 j 比特位信息由 $a(i, 1)$, $a(i, 2)$, $a(i, 3)$, i

$= 1, 2, \dots, 7$) 的相互关系表示, 嵌入是通过调整这三个数据的位置来实现的. 当水印的第 j 比特位为 1 时, 即 $b(j) = 1$ 时, 将 $a(i, 2)$ 位置的系数与 $a(i, 1)$, $a(i, 2)$, $a(i, 3)$ 中绝对值最大的系数调换位置, 若 $a(i, 2)$ 为绝对值最大的系数, 则位置不变, 否则, 将 $a(i, 1)$, $a(i, 2)$, $a(i, 3)$ 中绝对值最大的系数放在 $a(i, 2)$ 位置上, 而 $a(i, 1)$ 的位置放绝对值次大的系数, $a(i, 3)$ 的位置放绝对值最小的系数; 当水印的第 j 比特位为 0 时, 即 $b(j) = 0$ 时, 将 $a(i, 2)$ 位置的系数与 $a(i, 1)$, $a(i, 2)$, $a(i, 3)$ 中绝对值最小的系数调换位置, 而 $a(i, 1)$ 的位置放绝对值最大的系数, $a(i, 3)$ 的位置放绝对值次大的系数.

表 1 DCT 变换的系数矩阵

						$a(1, 1)$	$a(1, 2)$
					$a(2, 1)$	$a(2, 2)$	$a(1, 3)$
				$a(3, 1)$	$a(3, 2)$	$a(2, 3)$	
			$a(4, 1)$	$a(4, 2)$	$a(3, 3)$		
		$a(5, 1)$	$a(5, 2)$	$a(4, 3)$			
	$a(6, 1)$	$a(6, 2)$	$a(5, 3)$				
$a(7, 1)$	$a(7, 2)$	$a(6, 3)$					
	$a(7, 3)$						

因为嵌入水印时修改的 DCT 系数均为中频系数, 系数值相对都比较接近, 为增加水印的强度, 需要选择适当的参数 d , 通过参数 d 来调节嵌入水印的强度. 但是若参数 d 的选取过大, 将对图像质量造成影响. 检测时是根据系数的大小关系来提取水印的, 如果系数之间的差别很小, 经打印/ 扫描后, 相应的系数关系就有可能改变, 因此检测时发生误判的概率就很大, 相对来说就等于鲁棒性很差. 为此, 需要预先选定一个门限, 同时也是为了控制水印的强度, 即当

$$|\max(a(i, 1), a(i, 2), a(i, 3))| - |\min(a(i, 1), a(i, 2), a(i, 3))| < \forall th$$

时, 就认为水印强度不足, 需要在嵌入水印过程中使用 d 作为调节, 否则就直接嵌入.

3 数字水印的提取

水印的提取算法也就是水印嵌入算法的逆向算法, 但是由于数字图像在打印/ 扫描后, 失真较大, / 放大0 了的水印信息在按块提取时, 提取出的数据往往是不一致的. 下面我们考虑用模糊模式识别原理来解决数据在提取时的不一致问题.

对经过打印/ 扫描后的图像按 8@8 进行分块, 对于图像的每一 8@8 块, 做 DCT 变换, 所得系数仍用表 1 表示. 根据图像的每第 j 个 8@8 分块决定隐藏数据的第 j 个比特位 $b(j)$ 的值.

具体的提取算法如下:

$$\text{if } |a(i, 2)| > (|a(i, 1)| + |a(i, 2)|) / 2$$
$$b(i, j) = 1;$$
$$\text{else } b(i, j) = 0;$$

由于存在误判问题, 由此提取的七个比特 $b(1, j)$, $b(2, j)$, \dots , $b(7, j)$ 可能是不一致的, 通过提取出的 $b(1, j)$, $b(2, j)$, \dots , $b(7, j)$, 定义 $b(j)$ 对 1 的隶属度为

$$b(1, j), b(2, j), \dots, b(7, j) \text{ 中 } 1 \text{ 的个数} / 7$$
$$\text{对 } 0 \text{ 的隶属度为}$$

参考文献:

- [1] Van Schyndel R G, Tirkel A Z, Osborne C F. A digital watermark[A]. Proc IEEE Int. l Conference on Image Processing[C]. Los Alamitos, California, USA: IEEE CS Press, 1994. 2. 86- 90.
- [2] Wolfgang R B, Delp E J. A watermark for digital image[A]. Proc IEEE Int. l Conference on Image Processing[C]. Los Alamitos, California, USA: IEEE CS Press, 1996. 3. 219- 222.
- [3] Cox I J. Secure spread spectrum watermarking for multimedia[J]. IEEE Trans on Image Processing. 1997, 6(12): 1673- 1687.
- [4] Cox I J, Linnartz J P M G. Some general methods for tampering with watermarks[J]. IEEE Journal on Selected Areas in Communications, 1998, 16(4): 587- 593.
- [5] 钮心忻, 杨义先. 基于小波变换的数字水印隐藏与检测算法[J]. 计算机学报, 2000, 23(1): 21- 27.
- [6] 张静, 张春田. 能抗打印))) 扫描过程的数字图像水印算法[A]. 全国第三届信息隐藏学术研讨会论文集[C]. 西安: 机械工业出版社, 2001. 151- 157.
- [7] 宋玉杰, 刘瑞祯, 谭铁牛, 王蕴红. 数字水印技术在印刷品防伪中的应用[J]. 中国图像图形学报. 2001, 6(5): 450- 454.
- [8] Tadashi Mizumoto, Kineo Matsui. Robustness investigation of DCT digital watermark for printing and scanning[J]. Electronics and Communications in Japan (Part III: Fundamental Electronic Science), 2003, 86(4): 11- 19.
- [9] 牛少彰, 钮心忻, 杨义先, 陈宇丰. 基于密钥分存原理的数字水印分存技术[A]. 全国第四届信息隐藏学术研讨会论文集[C]. 大连: 机械工业出版社, 2002. 145- 151.

作者简介:



牛少彰 男, 1963 年生于北京, 教授, 1985 年和 1988 年在北京师范大学分别获得理学学士和硕士学位, 主要研究领域为图像处理、信息隐藏、数字水印、网络信息安全和应用数学。



钮心忻 女, 1963 年生于北京, 副教授, 1985 年和 1988 年在北京邮电大学分别获得工学学士和硕士学位, 1997 年在香港中文大学获博士学位, 主要研究领域为信息安全、信号信息处理、信息隐藏和数字水印。