图像的自嵌入及窜改的检测和恢复算法

张鸿宾,杨 成
(北京工业大学计算机学院,北京 100022)

摘 要: 提出一种图像自嵌入的水印算法.它将一个图像的主要内容嵌入该图像自身之中.该算法不仅能检测和定位对图像的窜改,而且能恢复被损坏的图像内容.算法的基本思路是把一个图像块的主要 DCT 系数经过量化、编码和加密后,嵌入另一个图像块的最低位之中.依据大量图像 DCT 量化系数的统计性质,本文仔细设计了主要 DCT 系数的编码表,使它既能满足水印负荷的要求,又能保证图像恢复时有较好的质量.分析了水印嵌入块和原图像块间偏移值应该满足的条件和设计方法.该算法为单向、不可逆的,可以抵抗可能的伪造攻击.

关键词: 多媒体认证; 数字水印; RSA 公钥算法; 自嵌入
中图分类号: TP391 **文献标识码**: A **文章编号**: 0372-2112 (2004) 02-0196-04

Tamper Detection and Self Recovery of Images Using Self-Embedding

ZHANG Hong-bin, YANG Cheng
(*Computer Institute of Beijing Polytechnic University, Beijing* 100022, *China*)

Abstract: A scheme for self-embedding an image into itself is proposed as a means for detecting and locating tamper, and self-recovering. The image is divided into small 8 ×8 blocks and each block is DCT transformed. A specified number of the lower frequency DCT coefficients are quantized and encoded using a fixed number of bits. The number of coefficients and their bit lengths are carefully chosen, according to the statistical property of these coefficients, the resulting bit-string for each block satisfies the payload of each block and has good self-recovering capability. The information about one block B (the bit string) is inserted into the LSB of another distant block, which is separated from B with an offset. We analyzed the method for choosing the offset. The proposed scheme is an one way watermarking and secure algorithm, and can well resist possible forged attack.

Key words: multimedia authentication; digital watermarking; RSA public key encryption; self-embedding

1 引言

近年来多媒体的应用取得了惊人的进展.数字媒体易于编辑、合成、复制和传播的优点在给人们带来方便的同时,也使它的知识产权保护和真实性、完整性的认证等问题成为人们关注的焦点,推动了以知识产权保护和完整性认证为目标的数字水印技术的研究.

数字水印技术利用人的视听觉特性和媒体内容的冗余性,有控制地将一些标识信息嵌入多媒体中,这些标识信息以后可以用作版权证明、完整性认证、拷贝控制和内容注释等目的.

数字水印的早期研究主要集中在用于版权保护的鲁棒水印上.随着研究的深入,人们认识到,数字水印在多媒体认证方面同样有很好的应用前景.所谓多媒体数据的认证就是要确认数据是否完整(integrity)、有无窜改,以及真实(reality)和来源可靠.多媒体数据的认证在电子商务和政务,法庭证据,新闻传媒,金融、保险、公安和医疗等电子文档的认证以及军事情报等领域都有广泛的应用.

多媒体数据的认证可以采用密码学的认证方法.传统的密码学认证有安全性较高的优点.但它不允许数据有一点变化,也很难确定窜改的位置和程度.另外,密码学的认证方法需要另外保存和传递认证码,这增加了安全上的漏洞[1].

近年来人们开始把水印技术用于多媒体完整性的认证[2~4].如果不允许原数据有一点变化,这时的认证常称为完全认证,所用的水印称为脆弱(fragile)水印.如果允许数据有一定程度的变化(如有损压缩等),只要不影响原数据的内容语义就行,这时的认证常称为内容认证,所用的水印称为半脆弱(semi-fragile)水印.完全认证和内容认证都有各自的应用场合.

目前已经有许多利用数字水印的图像认证方法发表[5~12].这些方法都在不同程度上解决了水印的不可见性、抗攻击性以及检测和定位窜改等问题,给今后认证水印的研究提供了启发.然而,总的来看,目前的认证水印在性能和安全性上仍有许多问题需要进一步的研究.作者认为,数字水印

收稿日期:2002-10-28;修回日期:2003-08-26
基金项目:国家自然科学基金(No. 60075002);北京市自然科学基金(No. 4992002);863 计划(No. 2001AA144080);北京市教委科技发展计划

和密码学方法相结合,以及认证和所有权验证相结合将是今后多媒体认证的重要研究方向.

值得注意的是,有不少的应用场合不仅要求检测和定位窜改,而且要求能够近似恢复被窜改的内容. Fridrich 曾提出了一种图像的自嵌入方法[14],将一幅图像的重要内容作为水印嵌入其自身之中,从而能够在检测窜改的同时近似地恢复被窜改的内容. 但 Fridrich 的算法为可逆算法,抵抗伪造攻击的能力不强. 另外,根据我们的实验结果发现,该算法的数据编码位长不够合理,因此恢复图像的质量不够好,需要增大嵌入的数据量来改进. 针对这些问题,本文依据 DCT 系数的统计特性,重新设计了较为合理的编码位长,并且改进了算法的安全性.

2　图像的自嵌入和自恢复算法

由于一幅图像的内容不可能完全嵌入其自身之中,因此必须抽取出其主要内容并进行压缩. 这些内容应该保证图像恢复时有足够的质量,同时又要考虑原图像的可嵌入容量,以便不影响水印嵌入后图像的质量. 借鉴 JPEG 压缩的方法,本文的图像自嵌入算法主要包括以下一些步骤:将图像划分为 8 ×8 的块并进行 DCT 变换,利用 JPEG 量化矩阵对一定数目的低频系数进行量化,对这些系数分别以一定的码长编码,使每块总的码长固定为 64 位. 将各图像块主要内容的 64 位编码(认证水印)分别嵌入相距较远的另外块的最低位(LSB)上. 这样,利用各块的认证水印即可验证各块是否被窜改. 若有窜改时,可以利用认证水印近似恢复被窜改的图像块.

下面几小节是算法的主要过程.

2.1　水印的产生和嵌入

(1) 将原图像 x 划分为 8 ×8 的块 $b_1, b_2, \cdots, b_r$, r 为块的总数.

(2) 将各块所有像素的最低位清零,作为以后水印嵌入的位置. 这样,每块的水印负荷为 64 位.

(3) 对每个最低位清零后的块 $b_i', i = 1,2,\cdots,r$,分别进行 DCT 变换. 以质量因子 0.75 时的 JPEG 量化表对 DCT 系数进行量化. 量化后的 DCT 系数矩阵记为 q.

(4) 对 q 的按之字形排序的前 14 个系数 $q_i, i = 0,1,\cdots,13$,进行 2 值编码作为水印 m_i. 每个系数的码长分配如图 1 所示. 这样每块的水印位长恰好是 64 位.

$$\begin{bmatrix} 7 & 6 & 5 & 4 & 0 & 0 & 0 & 0 \\ 6 & 6 & 4 & 3 & 0 & 0 & 0 & 0 \\ 5 & 4 & 4 & 0 & 0 & 0 & 0 & 0 \\ 4 & 3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

图 1　系数的码长分配表

(5) 对每块的水印码流 m_i 利用公钥系统的私钥 k 进行加密[13],形成加密后的水印码流 m_i'. 即

$$m_i' = E_k(m_i)$$

(6) 将块 $b_i, i = 1,2,\cdots,r$,的水印信息 m_i' 嵌入另一块 b_{i+p} 各像素的最低位. 其中 p 表示块 b_{i+p} 对块 b_i 的偏移. 如何确定 p 的合适值将在下面讨论.

水印产生和嵌入的流程如图 2 所示.

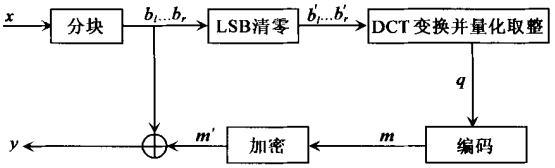

图 2　水印的产生和嵌入流程

2.2　窜改的检测与恢复

图像的认证流程如图 3 所示.

(1) 对待认证的图像 y,将其划分为 8 ×8 的块,记为 $c_1, c_2, \cdots, c_r$. 从每块像素的最低位中提取水印 m′,利用公钥系统的公钥 k 进行解密后,得到原始嵌入的水印 $m_i = D_k(m_i')$.

(2) 将每块像素的最低位清零,得到图像块 $c_1', c_2', \cdots, c_r'$. 对每个 c_i' 进行 DCT 变换、量化和二值编码,得到水印 w_i.

(3) 比较嵌入的水印 m_{i+p} 和从图像内容中提取出的水印 w_i,检测和定位窜改.

即:若 $w_i = m_{i+p}$,则块 c_i 和 c_{i+p} 均未被改动. 反之,若 $w_i \neq m_{i+p}$,则 w_i 和 m_{i+p} 哪一个被改动了,需参考 m_{i+p} 和 w_i 的情况而定.

(4) 当检测出有被窜改的块 c_i 后,可以用未被改动的块 c_{i+p} 中的水印信息 m_{i+p} 经过解码、逆量化和逆 DCT 变换后,近似地恢复被改动的块 c_i.

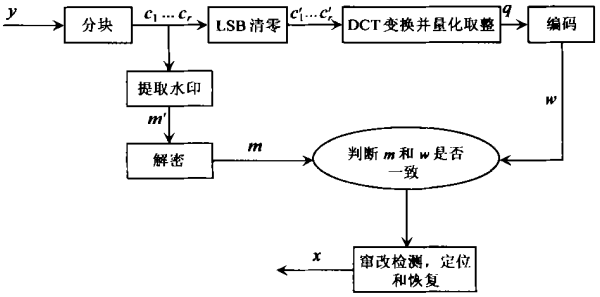

图 3　检测和恢复流程

2.3　量化系数编码位长的确定

如前所述,如果只用最低 1 位来嵌入水印的话,每个图像块的水印负荷是 64 位. 所以每块图像的主要内容也必须编码为 64位长. 文献[14]利用前11个DCT系数和图4的码长表

$$\begin{bmatrix} 7 & 7 & 7 & 5 & 0 & 0 & 0 & 0 \\ 7 & 6 & 5 & 0 & 0 & 0 & 0 & 0 \\ 6 & 5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

图 4　文献[14]的系数编码位长

使水印位长恰好为 64 位. 在实验中我们发现,用文[14]的编码表恢复出的图像质量不是很好. 对大量图像的 DCT 系数量化以后的值的分布特性进行了分析,其结果如图 5～8 所示. 其中横坐标表示量化系数的大小,纵坐标表示相应大小的量化系数出现的频率.

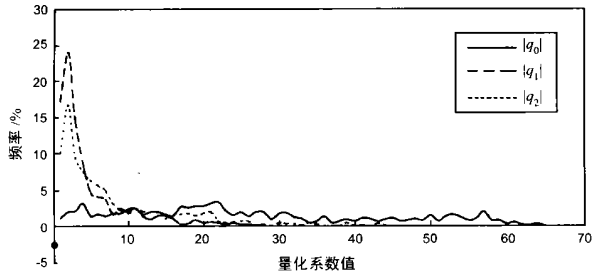

图 5 量化系数(q_0, q_1, q_2)的分布

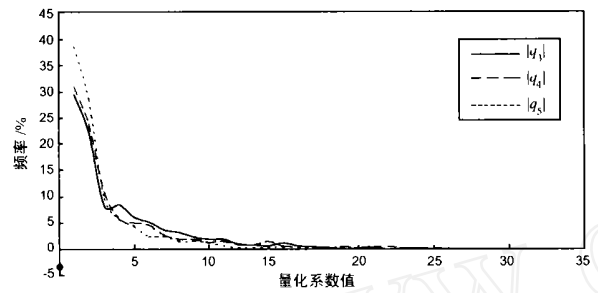

图 6 量化系数(q_3, q_4, q_5)的分布

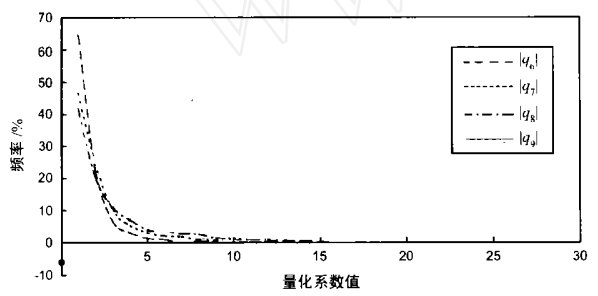

图 7 量化系数(q_6, q_7, q_8, q_9)的分布

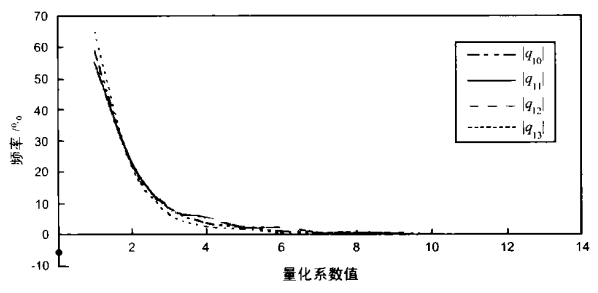

图 8 量化系数(q_{10}, q_{11}, q_{12}, q_{13})的分布

从图 5～8 中可以看出,量化后直流系数的绝对值大体上是从 0 到 64 间均匀分布. 而其他交流系数的绝对值则随值的增大而迅速衰减. 根据图 5-8 中的统计结果,可以适当分配每个系数的编码位长. 设阈值 $\alpha = 0.95$,计算量化系数值 x_i,使

$$P(q_i \leq x_i) \geq \alpha, 0 \leq i \leq 13$$

根据各 x_i 值的大小,可以给各个系数分配适当的码长,同时又使总的码长恰好为 64 位(图 1). 对于超出编码范围的系数值,以码长所能达到的最大值来代替. 实验结果表明,采用本文编码位长的恢复图像质量要优于文献[14]的.

2.4 偏移值的选取

如 2.1 节的(6)所述,块 b_i 的信息(水印) m′$_i$ 嵌入于另一块 b_{i+p} 各像素的最低位上. 其中 p 为 b_{i+p} 对 b_i 的偏移值. 在确定偏移值时有以下几个因素需要考虑:

(1) 块 b_i 和 b_{i+p} 要相距的远一些. 这样两块同时被修改的可能性就会减小.

(2) 偏移值 p 和总块数 r 之间应该满足:

$$\gcd(p, r) = 1$$

式中 gcd(·,·)为最大公约数. 如果不满足这个条件,假定 $r = np, p, n < r$,则经过 $n < r$ 次偏移后,有 $(i + np) \bmod r = i$. 即 $n(< r)$ 个块形成了一个小认证链(循环),而不是所有 r 个图像块构成一个认证链. 文献[14]中使用的偏移值 $0.3r$ 就有可能出现这个问题. 这在部分图像块出现错误的时候,增加了误判其他块的可能性. 安全上也存在漏洞. 如果所有的图像块构成一个认证链,在不是大部分的图像块都被窜改的情况下,会降低误判的可能性.

(3) 偏移值应该由用户密钥来控制,这样可以增加系统的安全性. 本文采用了如下的加密函数,它将第 i 块的主要内容嵌在了第 $f(i)$ 块的最低位. $f(i) = (k_0 + k_1 \times i) \bmod r$,且 $\gcd(k_1, r) = 1$. 式中 k_0 和 k_1 是两个常数,他们是用户掌握的密钥.

3 实验结果

我们用大量的图像进行了实验,下面是部分实验结果. 图 9(a)是原图像,图 9(b)是嵌入自恢复水印后的图像. 其峰值信噪比为 51.1,显示了很好的水印不可见性. 图 9(c)是经过窜改(在原图上添加了文字)的图像,图 9(d)是经过认证后自恢复的图像,其峰值信噪比为 50.6,表现了较好的恢复质量. 图 9(e)是对原图剪切后的图像,图 9(f)是自恢复的图像,其峰值信噪比为 33.4. 而文[14]算法恢复图像的峰值信噪比为 30.1. 图 10 是对 Baboon 图像实验的结果. 图 10(a)是剪切的图像. 图 10(b)是本文算法恢复的结果,峰值信噪比为 35.5. 而文[14]算法恢复图像的峰值信噪比为 31.7. 可以看出,本文算法的自恢复效果优于文[14]的算法.

4 小结

本文提出了一种图像自嵌入和自恢复的水印算法. 它不但能检测和定位对图像的窜改,而且能较好地恢复被窜改的内容. 为了提高算法的安全性和图像恢复的质量,本文主要完成了以下的工作.

(1) 根据对大量图像 DCT 系数量化值的统计,设计了更为合理的 DCT 低频系数的编码位长表. 与文献[14]的编码表相比较,本文的编码表更有根据也更为合理,因此图像恢复的质量也更好.

(2) 分析了偏移值选取的原则,使整幅图像形成一个认证

图 9　Lena 图像的实验结果

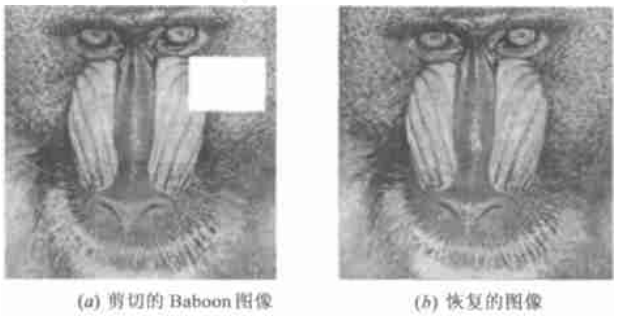

图 10　Baboon 图像的实验结果

链,提高了算法的安全性.而文献[14]的方法有可能导致许多认证链,因此安全上存在着漏洞.

(3)本文提出将非对称密码算法加密后的水印信息嵌入原始图像中,并且采用由用户密钥控制嵌入偏移量的办法,增大了图像的安全性,提高了抗伪造攻击的能力.

(4)本文提出的图像自嵌入和自恢复的方法除了可以认证和恢复图像之外,还有可能应用到图像和视频传输时的错误校正上.它的纠错能力可能比纠错码要强.

参考文献:

[1] Lin E T,Delp E J. A review of fragile image watermarks [A]. Proc. of the Multimedia and Security Workshop (ACM Multimedia '99) [C]. Orlando ,1999. 25 - 29.

[2] Walton S. Information authentication for a slippery new age [J]. Dr. Dobbs Journal ,1995 ,20(4) :18 - 26.

[3] Yeung M ,Mintzer F. Invisible watermarking for image verification [J]. Journal of Electronic Imaging ,1998 ,7(3) :578 - 591.

[4] Wolfgang R B ,Delp E J. A watermark for digital images [A]. Proc. IEEE Int. Conf. on Image Processing [C]. ACM Press ,NY,1996. 3. 219 - 222.

[5] Schynael R van ,Tirkel A ,Osborce C. A digital watermark [A]. Proc. of the IEEE ICIP [C]. Austin ,Texas ,1994. 2. 86 - 90.

[6] Wolfgang R B ,Delp E J. Fragile watermarking using the VW2D watermark [A]. Proc. SPIE ,Security and Watermarking of Multimedia Contents [C]. San Jose ,California ,1999. 204 - 213.

[7] Wong P W. A watermark for image integrity and ownership verification [A]. Proc. IS&T PIC Conf. [C]. Porland ,OR ,1998.

[8] Wong P W. A public key watermark for image verification and authentication [A]. Proc. ICIP [C]. Chicago ,IL ,Oct. 1998.

[9] Wong P W, Memon N. Secret and public key image watermarking schemes for image authentication and ownership verification [J]. IEEE Trans on Image Processing ,2001 ,10(10) :1593 - 1601.

[10] Wu M ,liu B. Watermarking for image authentication [A]. Proc. ICIP [C]. Chicago ,IL ,1998.

[11] Kundur D ,Hatzinakos D. Towards a telltale watermarking technique for tamper-proofing [A]. Proc. ICIP [C]. Chicago ,Illinois ,1998. 2. 409 - 413.

[12] Xie L ,Arce G. Joint wavelet compression and authentication watermarking [A]. Proc. IEEE ICIP [C]. Chicago ,Illinois ,1998. 2. 427 - 431.

[13] Rivest R L ,Shamir A ,Adleman L. A method for obtaining digital signature and public-key cryptosystems [J]. Commun. ACM ,1978 ,21 :120 - 126.

[14] Fridrich J ,Goljan M. Protection of digital images using self embedding [A]. Proceedings of NJIT Symposium on Content Security and Data Hiding in Digital Media ,(Newark ,NJ) [C]. New Jersty Institute of Technology May ,1999.

作者简介:

张鸿宾　男,1944 年生于河北,北京工业大学教授,博士生导师.主要研究方向为模式识别,图像处理,计算机视觉,人工神经网络,数据隐藏和数字水印等.

杨　成　男,1975 年生于河南,2002 年北京工业大学硕士毕业,目前主要从事模式识别和计算机应用的工作.