doi:10.3969/j.issn.1000-436X.2012.10.002

自嵌入完全盲检测顽健数字水印算法

叶天语
（浙江工商大学 信息与电子工程学院，浙江 杭州 310018）

摘 要：针对现有顽健水印算法无法实现完全盲检测，将自嵌入思想引入到顽健水印领域，提出一种完全盲检测顽健水印算法。首先将原始图像分割成互不重叠的子块，对每个子块进行离散余弦变换，通过比较每个子块的直流系数与所有子块直流系数的均值之间的大小关系产生特征水印，利用 Logistic 混沌序列对特征水印进行加密，然后调整每个子块的 2 个离散余弦变换中低频系数的大小自嵌入加密的特征水印，最后进行逆离散余弦变换得到含水印图像。算法结合自嵌入加密的特征水印和盲提取认证水印实现完全盲检测。实验结果表明：算法在抵抗平滑、添加噪声、JPEG 压缩、重采样、剪切和几何攻击如中间随机删除行、向下偏移行、向右偏移列上表现出很强的顽健性。

关键词：数字水印；自嵌入；完全盲检测；顽健性；特征水印；认证水印
中图分类号：TN911.7 **文献标识码**：A **文章编号**：1000-436X(2012)10-0007-09

Self-embedding robust digital watermarking algorithm with perfectly blind detection

YE Tian-yu
(College of Information & Electronic Engineering, Zhejiang Gongshang University, Hangzhou 310018, China)

Abstract: Those existing robust watermarking algorithms couldn't achieve perfectly blind detection. A robust watermarking algorithm with perfectly blind detection was proposed by introducing the self-embedding idea into robust watermarking literature. At first, the original image was split into non-overlapping blocks, and each block was transformed by DCT. The feature watermark was derived from judging the numerical relationship between each block's DC coefficient and average of DC coefficients form all blocks. After encrypted by Logistic chaos sequence, the feature watermark was self-embedded into each block by adjusting two middle or low frequency DCT coefficients. Finally, the watermarked image was obtained after IDCT. The proposed algorithm achieved perfectly blind detection by combining self-embedding the encrypted feature watermark and blindly extracting authentication watermark. Experimental results show that the proposed algorithm has strong robustness to resist various attacks such as smoothing, adding noise, JPEG compression, resampling, cropping and geometric attacks like random row removal, downward shifting and right shifting.
Key words:digital watermarking; self-embedding; perfectly blind detection; robustness; feature watermark; authentication watermark

1 引言

根据数字水印的内容进行分类，水印可分为有意义水印和无意义水印。有意义水印一般指水印的内容具有具体的含义，比如二值图像、灰度图像、商标、作者个人信息等；无意义水印一般对应一个

收稿日期：2011-03-28；**修回日期**：2011-08-24
基金项目：浙江省教育厅科研基金资助项目（Y201017916）
Foundation Item: The Scientific Research Fund of Zhejiang Provincial Education Department (Y201017916)

没有具体含义的伪随机序列。除此之外，温泉等[1]提出的“零水印”实质上代表原始载体的某个特征，可看作是一种特殊形式的有意义水印。

根据检测端是否需要借助原始载体，数字水印技术可分为盲数字水印技术和非盲数字水印技术。非盲数字水印技术在检测端检测水印时往往需要借助与原始载体相关的信息，盲数字水印技术在检测端不需要借助任何与原始载体相关的信息。盲数字水印技术比非盲数字水印技术更具有实用性。

顽健数字水印算法在检测端通常通过计算原始水印与提取的水印之间的相关度来判断版权。根据检测端是否需要借助与原始载体、原始水印相关的信息，目前的顽健数字水印算法的检测行为可分为 4 类。第 1 类：检测端既需要借助原始载体的相关信息，又需要借助原始水印的相关信息。第 2 类：检测端需要借助原始载体的相关信息，但不需要原始水印的相关信息。第 3 类：检测端不需要借助原始载体的相关信息，但需要借助原始水印的相关信息。第 4 类：检测端既不需要借助原始载体的相关信息，也不需要借助原始水印的相关信息。传统意义上的非盲顽健水印算法具有第 1 类或第 2 类检测行为，实用性比较差。具有第 3 类检测行为的盲无意义顽健水印算法[2~5]在检测端不借助任何原始载体的相关信息从攻击载体提取出水印，但要借助密钥产生原始伪随机水印序列，然后计算两者的相关度判断版权。具有第 3 类检测行为的盲有意义顽健水印算法[6~16]在检测端不借助任何原始载体的相关信息从攻击载体提取出水印，然后计算从嵌入端传输过来的原始水印与提取的水印之间的相关度判断版权。零水印算法[1,17~20]在检测端需要把存储在第三方公证中心的原始零水印取出来，然后计算与提取的零水印之间的相关度判断版权，也具有第 3 类检测行为。具有第 4 类检测行为的顽健水印算法国内外几乎没有出现，本文称这类算法为完全盲检测顽健水印算法。

目前的盲顽健水印算法具有第 3 类检测行为，仍然无法满足对实用性要求很高的场合，其实用性需要进一步改善。这是因为：具有第 3 类检测行为的盲顽健水印算法虽然在检测端不需要借助任何原始载体的相关信息，但还是需要借助原始水印或原始水印的部分信息来衡量原始水印与提取的水印之间的相关度以判断版权。那么，嵌入端传输原始水印或其部分信息到检测端(或第三方公证中心)进行存储就需要一定的传输成本和存储成本，而且传输的过程很难完全防止互联网上普遍存在的被动攻击。例如，攻击者成功“偷听”所传递的原始水印或原始水印的部分信息，通过分析得到所嵌入的水印并进一步伪造水印传递给检测端，会使得真正的原始水印无法用来鉴别版权，达到干扰版权鉴别的目的，从而使得水印算法无法抵抗解释攻击[21]。

自嵌入脆弱水印技术[22~26]的显著特点是嵌入端提取原始载体的特征产生水印并自嵌入到原始载体以达到内容完整性认证。目前，“自嵌入”思想基本上仅仅被应用到脆弱水印技术领域。

综合以上分析，本文试图将自嵌入脆弱水印算法的“自嵌入”思想引入到顽健数字水印领域，设计具有第 4 类检测行为的完全盲检测顽健水印算法，进一步改善现有盲水印算法的实用性。

2 自嵌入完全盲检测顽健数字水印算法描述

2.1 特征水印产生算法

图像经离散余弦变换（DCT, discrete cosine transformation）将得到直流（DC, direct current）系数、低频交流系数、中频交流系数和高频交流系数。相比于交流系数，DC 系数往往很大，集中了图像的主要能量。因此，图像分块后各子块的 DC 系数与所有子块 DC 系数均值之间的大小关系对外在干扰会表现出较好的稳定性。本文正是利用这种稳定性产生顽健特征水印[17]。

根据以下过程从大小为 $M\times M$ 的原始图像产生特征水印[17]。

1) 原始图像分成大小为 $m\times m$ 的不重叠子块。

2) 每个子块进行 DCT，$\boldsymbol{B}_i(0,0)$ 代表第 i 个子块的 DC 系数，$\boldsymbol{B}_v(0,0)$ 代表所有子块 DC 系数的均值，$i=1,2,\cdots,\left(\frac{M}{m}\right)^2$。

3) 特征水印 $\boldsymbol{W}$ 通过对比每个子块 DC 系数 $\boldsymbol{B}_i(0,0)$ 与所有子块 DC 系数均值 $\boldsymbol{B}_v(0,0)$ 的大小关系来产生：

$$w_i=\begin{cases}1, & \boldsymbol{B}_i(0,0)>\boldsymbol{B}_v(0,0)\\ 0, & \text{其他}\end{cases} \tag{1}$$

其中，w_i 代表 $\boldsymbol{W}$ 的第 i 比特。容易知道，特征水印 $\boldsymbol{W}$ 的长度为 $\left(\frac{M}{N}\right)^2$ bit。

由以上过程可以知道，$\boldsymbol{W}$ 是通过提取原始图像的 DC 系数特征产生的，这个过程其实没有对原始图像做任何的修改。因此，本文把 $\boldsymbol{W}$ 称为特征水印。

2.2　特征水印加密

不可见性、抗攻击顽健性和安全性是数字水印技术的基本要求。根据混沌动力学理论，混沌映射往往具有初值敏感性，初值发生微小变化都会导致产生的随机序列发生较大的改变。为了增强本文算法的安全性，本文利用 Logistic 混沌映射的初值敏感性对产生的特征水印进行加密。不知道密钥的攻击者无法准确解密出特征水印。

在非线性动力系统中，混沌现象是确定性的、类似随机的过程，不收敛又没有周期性，对初始值非常敏感。Logistic 混沌映射是一个广泛应用的非线性动力系统，定义为

$$x_{n+1}=1-\gamma x_n^2 \tag{2}$$

x_{n+1} 的取值范围为 $(-1,1)$，$n=0,1,2,3,\cdots$。如果 γ 在 $[0,2]$ 内取值，则该映射处于混沌状态。利用 Logistic 映射加密特征水印的过程说明如下。

1) 在 $(-1,1)$ 选择一个实数作为初值 x_0，在 $[0,2]$ 选择一个实数作为 γ，然后通过式(2)进行迭代产生混沌随机数序列 $\{x_1,x_2,x_3,\cdots\}$。将初值 x_0 和参数 γ 作为 Logistic 映射的前 2 个密钥。

2) 舍去混沌序列的前 κ 个随机数，因为混沌序列最初的随机数往往不稳定。将第 $\kappa+1$ 个到第 $\kappa+\left(\frac{M}{N}\right)^2$ 个随机数 $x_{\kappa+1},x_{\kappa+2},\cdots,x_{\kappa+\left(\frac{M}{N}\right)^2}$ 通过式(3)二值化为 GF(2)域的 $\{0,1\}$ 序列 $\boldsymbol{L}$，即

$$l_i=\frac{\operatorname{sgn}(x_{\kappa+i})+1}{2} \tag{3}$$

其中，sgn(·) 为符号函数，l_i 为 $\boldsymbol{L}$ 的第 i 比特，$i=1,2,\cdots,\left(\frac{M}{m}\right)^2$。Logistic 映射均值为 0，所以式(3)选择 0 作为阈值进行二值化。将参数 κ 作为 Logistic 映射的第 3 个密钥。

3) 利用二值化得到的 $\{0,1\}$ 序列 $\boldsymbol{L}$ 对特征水印 $\boldsymbol{W}$ 进行加密。加密方法为

$$w_i^e=w_i\oplus l_i \tag{4}$$

其中，$\oplus$ 为异或运算，w_i^e 为加密的特征水印 $\boldsymbol{W}^e$ 的第 i 比特，$i=1,2,\cdots,\left(\frac{M}{m}\right)^2$。

2.3　加密的特征水印自嵌入算法

DC 系数集中了图像各原始子块的主要能量，将加密的特征水印 $\boldsymbol{W}^e$ 自嵌入每个子块的 DC 系数容易使含水印图像产生方块效应。每个子块的 DCT 高频交流系数往往很小，很容易被 JPEG 压缩等常见信号处理过滤掉，将加密的特征水印 $\boldsymbol{W}^e$ 自嵌入每个子块的 DCT 高频交流系数会使算法无法具备较强的抗攻击顽健性。相比于 DCT 高频交流系数，每个子块 DCT 中低频系数具有更多的能量，有更大的感觉容量。将加密的特征水印 $\boldsymbol{W}^e$ 自嵌入每个子块的 DCT 中低频交流系数，一方面不会容易使水印图像产生方块效应，从而使算法具备较好的不可见性，另一方面会使算法具备较强的抗攻击顽健性。

调整原始图像各子块的 2 个 DCT 中低频交流系数自嵌入加密的特征水印 $\boldsymbol{W}^e$。

1) 原始图像分成大小为 $m\times m$ 的不重叠子块。

2) 每个子块进行 DCT。

3) 将加密的特征水印自适应自嵌入原始图像的每个子块 DCT 中低频交流系数：

$$\begin{cases}\boldsymbol{B}_i(r_1,s_1)\leftrightarrow\boldsymbol{B}_i(r_2,s_2),\\ \text{如果 } w_i^e=0\text{且 }\boldsymbol{B}_i(r_1,s_1)<\boldsymbol{B}_i(r_2,s_2)\\ \boldsymbol{B}_i(r_j,s_j)=\boldsymbol{B}_i(r_j,s_j),\\ \text{如果 } w_i^e=0\text{且}\boldsymbol{B}_i(r_1,s_1)\geqslant\boldsymbol{B}_i(r_2,s_2)\end{cases}$$

$$\begin{cases}\boldsymbol{B}_i(r_1,s_1)\leftrightarrow\boldsymbol{B}_i(r_2,s_2),\\ \text{如果 } w_i^e=1\text{且 }\boldsymbol{B}_i(r_1,s_1)\geqslant\boldsymbol{B}_i(r_2,s_2)\\ \boldsymbol{B}_i(r_j,s_j)=\boldsymbol{B}_i(r_j,s_j),\\ \text{如果 } w_i^e=1\text{且}\boldsymbol{B}_i(r_1,s_1)<\boldsymbol{B}_i(r_2,s_2)\end{cases}$$

$$\begin{cases}\begin{cases}\boldsymbol{B}_i(r_1,s_1)=\boldsymbol{B}_i(r_1,s_1)+\eta_i/2\\ \boldsymbol{B}_i(r_2,s_2)=\boldsymbol{B}_i(r_2,s_2)-\eta_i/2\end{cases}\\ \text{如果 } 0<\boldsymbol{B}_i(r_1,s_1)-\boldsymbol{B}_i(r_2,s_2)<\eta_i\\ \boldsymbol{B}_i(r_j,s_j)=\boldsymbol{B}_i(r_j,s_j),\\ \text{如果 } \boldsymbol{B}_i(r_1,s_1)-\boldsymbol{B}_i(r_2,s_2)\geqslant\eta_i\end{cases}$$

$$\begin{cases}\begin{cases}\boldsymbol{B}_i(r_1,s_1)=\boldsymbol{B}_i(r_1,s_1)-\eta_i/2\\ \boldsymbol{B}_i(r_2,s_2)=\boldsymbol{B}_i(r_2,s_2)+\eta_i/2\end{cases}\\ \text{如果 } 0<\boldsymbol{B}_i(r_2,s_2)-\boldsymbol{B}_i(r_1,s_1)<\eta_i\\ \boldsymbol{B}_i(r_j,s_j)=\boldsymbol{B}_i(r_j,s_j),\\ \text{如果 } \boldsymbol{B}_i(r_2,s_2)-\boldsymbol{B}_i(r_1,s_1)\geqslant\eta_i\end{cases} \tag{5}$$

其中，符号“$\leftrightarrow$”表示交换左右 2 个数的大小；$\eta_i=\mu\times\boldsymbol{B}_i(0,0)$；$\mu$ 为 w_i^e 自适应嵌入的尺度因子，根据实际应用场合对不可见性和顽健性的要求进

行折中选择；$\boldsymbol{B}_i(r_j,s_j)$ 代表第 i 个子块 DCT 矩阵在 (r_j,s_j) 处的系数，$j=1,2$，要求满足 $r_1=r_2$ 和 $s_1=s_2$ 不能同时成立，而且 $r_1=0$ 和 $s_1=0$ 不能同时成立，$r_2=0$ 和 $s_2=0$ 不能同时成立。

4) 各子块进行逆 DCT，重组后得到含水印图像。

2.4　特征水印提取算法

检测端特征水印提取过程与嵌入端特征水印产生过程类似，通过对比图像各子块的 DC 系数与所有子块 DC 系数均值之间的大小关系来提取特征水印。从攻击图像提取特征水印的过程说明如下[17]。

1) 攻击图像分成大小为 $m\times m$ 的不重叠子块。

2) 每个子块进行 DCT，$\boldsymbol{B}_i^a(0,0)$ 代表第 i 个子块的 DC 系数，$\boldsymbol{B}_v^a(0,0)$ 代表所有子块 DC 系数的均值，$i=1,2,\cdots,\left(\frac{M}{m}\right)^2$。

3) 特征水印 $\boldsymbol{W}^a$ 通过对比每个子块 DC 系数 $\boldsymbol{B}_i^a(0,0)$ 与所有子块 DC 系数均值 $\boldsymbol{B}_v^a(0,0)$ 的大小关系提取出来：

$$w_i^a=\begin{cases}1, & \boldsymbol{B}_i^a(0,0)>\boldsymbol{B}_v^a(0,0)\\ 0, & \text{其他}\end{cases} \tag{6}$$

其中，w_i^a 代表 $\boldsymbol{W}^a$ 的第 i 比特。

2.5　认证水印盲提取算法和解密

检测端认证水印的提取过程为嵌入端加密的特征水印自嵌入过程的逆过程，通过对比每个子块 2 个 DCT 中低频交流系数的大小关系提取认证水印。认证水印的提取和解密过程说明如下。

1) 攻击图像分成大小为 $m\times m$ 的不重叠子块。

2) 每个子块进行 DCT。

3) 从每个子块 DCT 中低频交流系数提取出认证水印 $\boldsymbol{W}^{a'}$：

$$w_i^{a'}=\begin{cases}0, & \boldsymbol{B}_i(r_1,s_1)>\boldsymbol{B}_i(r_2,s_2)\\ 1, & \text{其他}\end{cases} \tag{7}$$

其中，$w_i^{a'}$ 代表 $\boldsymbol{W}^{a'}$ 的第 i 比特。

4) 利用 x_0 和 γ 2 个密钥产生 Logistic 映射混沌序列 $\{x_1,x_2,x_3,\cdots\}$，舍弃前 κ 个随机数，将第 $\kappa+1$ 个到第 $\kappa+\left(\frac{M}{N}\right)^2$ 个随机数 $x_{\kappa+1},x_{\kappa+2},\cdots,x_{\kappa+\left(\frac{M}{N}\right)^2}$ 通过式(3)二值化为 GF(2)域的{0,1}序列 $\boldsymbol{L}$。

5) 通过式(4)相对应的解密方法对提取出的认证水印 $\boldsymbol{W}^{a'}$ 进行解密，即

$$w_i^{a''}=w_i^{a'}\oplus l_i \tag{8}$$

其中，$w_i^{a''}$ 代表解密后认证水印 $\boldsymbol{W}^{a''}$ 的第 i 比特。

6) 计算攻击后提取出的特征水印 $\boldsymbol{W}^a$ 和解密后的认证水印 $\boldsymbol{W}^{a''}$ 之间的归一化相关度（NC, normalized correlation）评价算法的顽健性以判断版权。NC 定义为

$$\theta=\left(\sum_{i=1}^{\left(\frac{M}{m}\right)^2}(w_i^a\cdot w_i^{a''})\right)\Big/\left(\sqrt{\sum_{i=1}^{\left(\frac{M}{m}\right)^2}(w_i^a)^2}\cdot\sqrt{\sum_{i=1}^{\left(\frac{M}{m}\right)^2}(w_i^{a''})^2}\right) \tag{9}$$

显然，检测端提取认证水印时达到盲提取。本文把 $\boldsymbol{W}^{a'}$ 称为认证水印的原因是检测端对 $\boldsymbol{W}^{a'}$ 进行解密后用于进行版权认证。由第 2.4 节和第 2.5 节可知，检测端只需要利用攻击后的含水印图像就可以分别提取出特征水印和认证水印来计算 NC，不需要借助任何原始图像和原始水印的相关信息。因此，本文算法可以实现完全盲检测。

3　实验结果

3.1　实验参数说明

选择 256 级大小为 512×512 的 Lena、Barbara 和 Elain 3 幅灰度图像作为实验图像，分别如图 1、图 2 和图 3 所示。子块的大小都为 16×16，所以特征水印 $\boldsymbol{W}$ 的长度为 1 024bit。Logistic 混沌映射初值 x_0 取值为 0.28，参数 γ 取值为 1.5，Logistic 混沌序列加密特征水印 $\boldsymbol{W}$ 时舍去前 $\kappa=200$ 个随机数。加密的特征水印 $\boldsymbol{W}^e$ 自适应自嵌入原始 Lena、Barbara 和 Elain 图像每个子块处于(3,5)和(4,4)位置的 DCT 系数。w_i^e 自适应嵌入的尺度因子 μ 取值为 0.028。得到的含水印 Lena、Barbara 和 Elain 图像分别如图 4、图 5 和图 6 所示，与原始 Lena、Barbara 和 Elain 图像之间的 PSNR 为 35.959 5dB、36.121 1dB 和 35.763 7dB。因此，此时算法对 3 幅测试图像都具有良好的不可见性。

图 1　原始 Lena 图像

图 2　原始 Barbara 图像

图 3　原始 Elain 图像

图 4　含水印 Lena 图像

图 5　含水印 Barbara 图像

图 6　含水印 Elain 图像

3.2　抗攻击顽健性实验

用 NC 来衡量算法抵抗各种攻击的顽健性。中间随机删除行指从被删除行的下边第 1 行开始逐行向上移动，空余行补全黑。向下偏移行指整个图像下移几行，上面几行补全黑，最后几行移出丢失。向右偏移列指整个图像右移，前面几列补全黑，最后几列移出丢失。重采样采用 nearest 插值法。各表中“/”下方为攻击后的 3 幅含水印图像与原始图像之间的 PSNR，“/”上方为 2 个水印序列之间的 NC。

1) 提取出的特征水印 $\boldsymbol{W}^{a}$ 和解密后的认证水印 $\boldsymbol{W}^{a''}$ 之间的 NC。

攻击后的 3 幅含水印图像各自提取出的特征水印 $\boldsymbol{W}^{a}$ 和解密后的认证水印 $\boldsymbol{W}^{a''}$ 之间的 NC 如表 1 所示。从表 1 可以看出，算法对各种攻击都表现出很强的顽健性。

2) 提取出的特征水印 $\boldsymbol{W}^{a}$ 和原始特征水印 $\boldsymbol{W}$ 之间的 NC。

攻击后的3幅含水印图像各自提取出的特征水印 $\boldsymbol{W}^{a}$ 和原始特征水印 $\boldsymbol{W}$ 之间的NC 如表2 所示。因此，本文的特征水印产生算法对各种攻击都具有很强的顽健性。另外，对比表 1 和表 2 的数据可以发现，特征水印产生算法的抗攻击顽健性稍好于本文算法。

3) 解密后的认证水印 $\boldsymbol{W}^{a''}$ 和原始特征水印 $\boldsymbol{W}$ 之间的 NC。

攻击后的 3 幅含水印图像各自解密后的认证水印 $\boldsymbol{W}^{a''}$ 和原始特征水印 $\boldsymbol{W}$ 之间的 NC 如表 3 所示。因此，本文的加密特征水印自嵌入算法对各种攻击都具有很强的顽健性。另外，对比表 1 和表 3 的数据可以发现，加密的特征水印自嵌入算法的抗攻击顽健性稍好于本文算法。

4　分析与讨论

将文献[19]算法和文献[20]的算法二与本文的完全盲检测算法进行抗攻击顽健性对比。文献[19]算法的子块大小为 4×4。文献[20]的算法二以 haar 小波为基进行 1 级离散小波变换（DWT, discreet wavelet transformation），子块大小为 4×4。各种攻击时，文献[19]算法和文献[20]算法二产生的原始零水印和提取的零水印之间的 NC 如表 4 所示。对比表 1 和表 4 的数据可知，除了剪切攻击和中间随机删除 2 行攻击外，本文的完全盲检测算法的抗攻击顽健性基本上强于文献[19]算法和文献[20]算法二。另外，本文的算法能够实现完全盲检测，在检测端只需要攻击图像就可以判断版权。然而，文献[19]算法和文献[20]算法二无法实现完全盲检测，无法做到只利用攻击图像就判断版权，因为在检测端需

表 1　攻击后的 3 幅含水印图像各自提取出的特征水印 W^a 和解密后的认证水印 $W^{a''}$ 之间的 NC

图像	平滑				加噪声		JPEG 压缩		
	高斯低通滤波(窗口大小为 $4\times4,\sigma=2$)	高斯低通滤波(窗口大小为 $3\times3,\sigma=1$)	中值滤波（窗口大小为 3×3）	中值滤波（窗口大小为 5×5）	高斯噪声(均值为 0,方差为 0.001)	椒盐噪声(噪声密度为 0.005)	质量因子		
							75	50	25
Lena	0.976 9/	0.993 5/	0.994 4/	0.970 6/	0.994 4/	0.990 7/	0.998 1/	0.996 2/	0.991 6/
	27.836 9	31.537 4	33.130 9	30.305 8	28.687 4	27.497 3	33.656 6	32.670 7	31.474 5
Barbara	0.977 3/	0.997 0/	0.993 0/	0.958 7/	0.998 0/	0.987 2/	1.000 0/	0.995 1/	0.983 4/
	23.134 9	25.491 8	24.419 6	22.546 7	28.797 8	27.603 5	32.889 0	30.781 9	28.054 8
Elain	0.983 3/	0.992 1/	0.997 1/	0.981 4/	0.994 1/	0.997 1/	0.996 1/	0.998 0/	0.993 1/
	28.154 2	30.424 6	31.384 3	30.742 6	28.552 1	27.295 5	31.637 3	30.937 2	30.193 7

图像	重采样		剪切		几何攻击			
	先缩小到 0.8 倍再放大到 1.25 倍	先缩小到 0.5 倍再放大到 2 倍	左上角 1/32	右上角 1/32	中间随机删除行		向下偏移行	向右偏移列
					行数为 1	行数为 2	行数为 1	列数为 1
Lena	0.965 0/	0.980 5/	0.955 3/	0.939 1/	0.982 3/	0.920 0/	0.959 2/	0.957 5/
	25.931 3	28.171 6	21.465 7	18.011 5	28.187 7	24.723 3	26.997 0	25.702 9
Barbara	0.937 3/	0.952 2/	0.924 8/	0.932 2/	0.982 2/	0.908 6/	0.978 5/	0.949 4/
	21.300 1	21.965 9	17.507 2	18.999 7	25.449 8	21.284 1	24.086 7	19.297 1
Elain	0.976 5/	0.973 7/	0.948 4/	0.926 9/	0.979 5/	0.930 6/	0.983 3/	0.973 6/
	25.938 6	28.541 4	23.294 0	19.457 5	26.696 6	24.545 1	25.170 0	25.811 5

表 2　攻击后的 3 幅含水印图像各自提取出的特征水印 W^a 和原始特征水印 W 之间的 NC

图像	平滑				加噪声		JPEG 压缩		
	高斯低通滤波(窗口大小为 $4\times4,\sigma=2$)	高斯低通滤波(窗口大小为 $3\times3,\sigma=1$)	中值滤波（窗口大小为 3×3）	中值滤波（窗口大小为 5×5）	高斯噪声(均值为 0,方差为 0.001)	椒盐噪声(噪声密度为 0.005)	质量因子		
							75	50	25
Lena	0.987 9/	0.993 5/	0.994 4/	0.991 6/	0.996 2/	0.996 2/	0.998 1/	0.996 2/	0.994 4/
	27.836 9	31.537 4	33.130 9	30.305 8	28.660 5	27.438 3	33.656 6	32.670 7	31.474 5
Barbara	0.989 1/	0.997 0/	0.997 0/	0.988 1/	0.997 0/	0.995 0/	1.000 0/	0.997 0/	0.994 1/
	23.134 9	25.491 8	24.419 6	22.546 7	28.823 8	27.569 4	32.889 0	30.781 9	28.054 8
Elain	0.986 2/	0.994 1/	0.999 0/	0.996 1/	0.997 1/	0.995 1/	0.998 0/	1.000 0/	0.995 1/
	28.154 2	30.424 6	31.384 3	30.742 6	28.551 1	27.507 0	31.637 3	30.937 2	30.193 7

图像	重采样		剪切		几何攻击			
	先缩小到 0.8 倍再放大到 1.25 倍	先缩小到 0.5 倍再放大到 2 倍	左上角 1/32	右上角 1/32	中间随机删除行		向下偏移行	向右偏移列
					行数为 1	行数为 2	行数为 1	列数为 1
Lena	0.983 3/	0.989 7/	0.970 5/	0.938 1/	0.993 4/	0.981 3/	0.979 2/	0.988 8/
	25.931 3	28.171 6	21.465 7	18.011 5	28.187 7	24.723 3	26.997 0	25.702 9
Barbara	0.984 0/	0.993 0/	0.923 6/	0.942 2/	0.988 1/	0.979 1/	0.993 1/	0.980 0/
	21.300 1	21.965 9	17.507 2	18.999 7	25.449 8	21.284 1	24.086 7	19.297 1
Elain	0.985 2/	0.987 2/	0.972 1/	0.928 0/	0.990 2/	0.977 3/	0.986 3/	0.988 2/

25.938 6	28.541 4	23.294 0	19.457 5	26.696 6	24.545 1	25.170 0	25.811 5

表 3　攻击后的 3 幅含水印图像各自解密后的认证水印 $W^{a''}$ 和原始特征水印 W 之间的 NC

图像	平滑				加噪声		JPEG 压缩		
	高斯低通滤波(窗口大小为 $4\times4, \sigma=2$)	高斯低通滤波(窗口大小为 $3\times3, \sigma=1$)	中值滤波（窗口大小为 3×3）	中值滤波（窗口大小为 5×5）	高斯噪声(均值为 0,方差为 0.001)	椒盐噪声(噪声密度为 0.005)	质量因子		
							75	50	25
Lena	0.988 9/ 27.836 9	1.000 0/ 31.537 4	1.000 0/ 33.130 9	0.978 8/ 30.305 8	1.000 0/ 28.660 5	0.997 2/ 27.438 3	1.000 0/ 33.656 6	1.000 0/ 32.670 7	0.997 2/ 31.474 5
Barbara	0.988 2/ 23.134 9	1.000 0/ 25.491 8	0.996 0/ 24.419 6	0.970 4/ 22.546 7	0.997 0/ 28.823 8	0.997 0/ 27.569 4	1.000 0/ 32.889 0	0.998 0/ 30.781 9	0.989 2/ 28.054 8
Elain	0.997 1/ 28.154 2	0.998 0/ 30.424 6	0.998 0/ 31.384 3	0.985 3/ 30.742 6	0.998 0/ 28.551 1	0.997 1/ 27.507 0	0.998 0/ 31.637 3	0.998 0/ 30.937 2	0.998 0/ 30.193 7

图像	重采样		剪切		几何攻击			
	先缩小到 0.8 倍再放大到 1.25 倍	先缩小到 0.5 倍再放大到 2 倍	左上角 1/32	右上角 1/32	中间随机删除行		向下偏移行	向右偏移列
					行数为 1	行数为 2	行数为 1	列数为 1
Lena	0.981 5/ 25.931 3	0.990 7/ 28.171 6	0.955 3/ 21.465 7	0.939 1/ 18.011 5	0.988 8/ 28.187 7	0.936 1/ 24.723 3	0.979 7/ 26.997 0	0.966 5/ 25.702 9
Barbara	0.949 2/ 21.300 1	0.959 0/ 21.965 9	0.985 1/ 17.507 2	0.983 0/ 18.999 7	0.994 1/ 25.449 8	0.927 0/ 21.284 1	0.985 4/ 24.086 7	0.967 5/ 19.297 1
Elain	0.991 2/ 25.938 6	0.986 4/ 28.541 4	0.978 2/ 23.294 0	0.989 2/ 19.457 5	0.989 3/ 26.696 6	0.952 7/ 24.545 1	0.997 1/ 25.170 0	0.985 3/ 25.811 5

表 4　文献[19]算法和文献[20]算法二产生的原始零水印和提取的零水印之间的 **NC**

文献	图像	平滑				加噪声		JPEG 压缩		
		高斯低通滤波(窗口大小为 $4\times4, \sigma=2$)	高斯低通滤波(窗口大小为 $3\times3, \sigma=1$)	中值滤波（窗口大小为 3×3）	中值滤波（窗口大小为 5×5）	高斯噪声(均值为 0,方差为 0.001)	椒盐噪声(噪声密度为 0.005)	质量因子		
								75	50	25
文献[19]	Lena	0.922 8/ 28.278 5	0.937 4/ 32.999 1	0.954 0/ 35.111 0	0.939 3/ 30.908 9	0.922 8/ 29.974 3	0.990 1/ 28.187 3	0.966 2/ 37.502 1	0.955 5/ 35.477 8	0.943 1/ 33.378 9
	Barbara	0.913 1/ 23.273 5	0.925 2/ 25.803 0	0.943 3/ 24.618 1	0.932 6/ 22.620 3	0.912 3/ 29.986 4	0.987 0/ 28.428 1	0.967 4/ 35.693 3	0.957 8/ 32.286 4	0.939 6/ 28.788 2
	Elain	0.917 2/ 28.639 5	0.921 2/ 31.537 8	0.937 0/ 32.489 1	0.928 0/ 31.302 0	0.959 1/ 29.990 9	0.992 3/ 28.676 0	0.956 0/ 33.758 5	0.941 2/ 32.668 9	0.929 6/ 31.584 5
文献[20]的算法二	Lena	0.954 9/ 28.278 5	0.982 5/ 32.999 1	0.985 6/ 35.111 0	0.968 9/ 30.908 9	0.982 7/ 29.965 1	0.984 8/ 28.362 5	0.992 2/ 37.502 1	0.984 8/ 35.477 8	0.976 8/ 33.378 9
	Barbara	0.927 6/ 23.273 5	0.966 5/ 25.803 0	0.978 6/ 24.618 1	0.957 1/ 22.620 3	0.970 5/ 29.995 2	0.982 0/ 28.394 3	0.994 0/ 35.693 3	0.983 1/ 32.286 4	0.965 6/ 28.788 2
	Elain	0.969 9/ 28.639 5	0.988 4/ 31.537 8	0.991 5/ 32.489 1	0.988 4/ 31.302 0	0.986 4/ 29.981 6	0.992 0/ 28.458 7	0.996 3/ 33.758 5	0.992 4/ 32.668 9	0.983 4/ 31.584 5

文献	图像	重采样		剪切		几何攻击			
		先缩小到 0.8 倍再放大到 1.25 倍	先缩小到 0.5 倍再放大到 2 倍	左上角 1/32	右上角 1/32	中间随机删除行		向下偏移行	向右偏移列
						行数为 1	行数为 2	行数为 1	列数为 1
文献[19]	Lena	0.945 1/ 25.713 5	0.939 7/ 27.967 9	0.981 3/ 21.619 2	0.984 3/ 18.080 3	0.965 6/ 28.922 0	0.952 2/ 24.998 4	0.960 1/ 27.573 1	0.940 2/ 26.081 9
	Barbara	0.942 9/ 21.148 2	0.932 8/ 21.809 2	0.985 0/ 17.566 1	0.983 0/ 19.082 2	0.976 3/ 25.843 2	0.964 4/ 21.427 1	0.958 5/ 24.347 0	0.944 9/ 19.382 1
	Elain	0.960 7/ 25.633 2	0.940 2/ 28.265 3	0.980 8/ 23.539 4	0.983 5/ 19.558 2	0.982 2/ 27.246 2	0.971 7/ 24.846 0	0.970 9/ 25.552 5	0.968 4/ 26.245 1
文献[20]的算法二	Lena	0.946 3/ 25.713 5	0.954 1/ 27.967 9	0.980 4/ 21.619 2	0.985 8/ 18.080 3	0.974 8/ 28.922 0	0.952 6/ 24.998 4	0.960 1/ 27.573 1	0.932 5/ 26.081 9
	Barbara	0.912 7/ 21.148 2	0.938 1/ 21.809 2	0.986 3/ 17.566 1	0.985 0/ 19.082 2	0.968 8/ 25.843 2	0.947 8/ 21.427 1	0.943 2/ 24.347 0	0.911 4/ 19.382 1
	Elain	0.964 4/ 25.633 2	0.970 2/ 28.265 3	0.981 1/ 23.539 4	0.985 7/ 19.558 2	0.980 7/ 27.246 2	0.968 7/ 24.846 0	0.970 4/ 25.552 5	0.967 7/ 26.245 1

要把存储在第三方公证中心的原始零水印取出来，然后计算与提取的零水印之间的相关度判断版权。

本文算法对各种攻击具有强顽健性，其原因在于同时达到以下 2 点：1)特征水印产生算法本身就具有很强的抗攻击顽健性；2)加密的特征水印自嵌入算法本身就具有很强的抗攻击顽健性。另外，本文算法的抗攻击顽健性稍差于特征水印产生算法和加密的特征水印自嵌入算法的原因在于加密的特征水印自嵌入会对产生特征水印的 DC 系数有一定的影响。

从第 2 节可以看出，本文算法能够实现完全盲检测，其原因在于同时达到以下 2 点：1)从原始图像通过特征提取产生特征水印，经加密后自嵌入到原始图像，而不是将额外的水印形式嵌入到原始图像；2)加密的特征水印自嵌入算法在检测端本身就可以达到盲提取认证水印。然而，国内外现有顽健水印算法[1~20]仍然无法实现完全盲检测。因此，本文算法改善了现有顽健水印算法的实用性。

5　结束语

目前的盲顽健水印算法在检测端还是需要借助原始水印或原始水印的部分信息，从而无法真正实现完全盲检测，实用性受到限制。针对这一问题，本文将脆弱水印算法的“自嵌入”思想引入到顽健水印领域，提出一种自嵌入完全盲检测顽健水印算法。本文的算法同时具备以下 4 个特性。

1) “自嵌入”特性：通过特征提取方法从原始图像产生特征水印，再将加密的特征水印自嵌入到原始图像产生含水印图像。

2) “完全盲提取”特性：检测时既不需要借助任何与原始载体相关的信息，又不需要借助任何与原始水印相关的信息。在检测端，首先从攻击图像利用与嵌入端相同的特征提取方法提取特征水印；然后从攻击图像利用嵌入端加密的特征水印自嵌入算法的逆过程盲提取出认证水印；最后计算提取的特征水印与认证水印之间的归一化相关度来判断版权。因此，检测端只需要利用攻击后的含水印图像。

3) 顽健性：特征水印产生算法和加密的特征水印自嵌入算法对各种攻击都具有很强的顽健性，进一步地，本文的算法具有很强的抵抗各种攻击的顽健性。

4) 安全性：特征水印自嵌入原始图像前先要进行 Logistic 混沌序列加密，不知道加密密钥的攻击者是无法准确解密出特征水印的。

本文提出的自嵌入完全盲检测顽健水印技术改善了现有盲顽健水印技术的实用性，有利于推动顽健数字水印技术的实用化进程。这体现在以下 2 个方面。

1) 节省传输成本和存储成本。检测端只需要攻击后的含水印图像就可以进行版权认证，嵌入端无需传输原始载体和原始水印的任何信息到检测端(或第三方公证中心)进行存储，从而节省传输成本和存储成本。

2) 阻断解释攻击。嵌入端无需传输原始水印的任何信息到检测端(或第三方公证中心)，可以有效阻断解释攻击。

本文算法将原始图像每一子块产生的特征水印嵌入到该块的 DCT 系数中，使得算法易受矢量量化（VQ, vector quantization）攻击。下一步的研究将集中在使自嵌入完全盲检测水印算法能够抵抗 VQ 攻击。

参考文献：

[1] 温泉,孙锬锋,王树勋.零水印的概念与应用[J].电子学报, 2003, 31(2):214-216.
WEN Q, SUN T F, WANG S X. Concept and application of zero-watermark[J]. Acta Electronica Sinica, 2003,31(2):214-216.

[2] WANG X Y, HOU L M, WU J. A feature-based robust digital image watermarking against geometric attacks[J]. Image and Vision Computing, 2008, 26(7):980-989.

[3] WANG X Y, YANG Y P, YANG H Y. Invariant image watermarking using multi-scale Harris detector and wavelet moments[J]. Computers and Electrical Engineering, 2010, 36(1):31-44.

[4] QI X J, QI J. A robust content-based digital image watermarking scheme[J]. Signal Processing,2007, 87(6):1264-1280.

[5] WANG X Y, REEVES S D. Robust correlation of encrypted attack traffic through stepping stones by flow watermarking[J]. IEEE Transactions on Dependable and Secure Computing, 2011,8(3):434-449.

[6] 牛少彰,钮心忻,杨义先等.半色调图像中数据隐藏算法[J].电子学报, 2004 ,32 (7):1180-1183.
NIU S Z, NIU X X, YANG Y X, *et al*. Data hiding algorithm for halftone images[J]. Acta Electronica Sinica , 2004 ,32 (7):1180-1183.

[7] WANG X Y, CUI C Y. A novel image watermarking scheme against desynchronization attacks by SVR revision[J]. Journal of Visual Communication and Image Representation, 2008, 19(5):334-342.

[8] 李旭东,张振跃.图像双层划分和奇异值分解的数字水印算法[J].浙江大学学报(工学版), 2006,40(12):2088-2092.

LI X D, ZHANG Z Y. Two-layer partition and singular value decomposition based image watermarking[J]. Journal of Zhejiang University (Engineering Science) ,2006,40(12):2088-2092.

[9] 袁大洋,肖俊,王颖.数字图像水印算法抗几何攻击顽健性研究[J].电子与信息学报,2008,30(5):1251-1256.
YUAN D Y, XIAO J, WANG Y. Study on the robustness of digital image watermarking algorithms to geometric attacks[J]. Journal of Electronics & Information Technology, 2008,30(5):1251-1256.

[10] 李旭东.抗几何攻击的空间域图像数字水印算法[J].自动化学报,2008,34(7):832-837.
LI X D. Geometric attack resistant image watermarking in spatial domain[J]. Acta Automatica Sinica, 2008,34(7):832-837.

[11] 许文丽,李磊,王育民. 抗噪声、几何失真和 JPEG 压缩攻击的顽健数字水印方案[J].电子与信息学报,2008,30(4):933-936.
XU W L, LI L, WANG Y M. Robust digital watermarking scheme resistant to gaussian noise, geometric distortion and JPEG compression attacks[J]. Journal of Electronics & Information Technology, 2008, 30(4):933-936.

[12] 李雷达,郭宝龙,表金峰.基于奇偶量化的空域抗几何攻击图像水印算法[J].电子与信息学报,2009,31(1):134-138.
LI L D, GUO B L, BIAO J F. Spatial domain image watermarking scheme robust to geometric attacks based on odd-even quantization[J]. Journal of Electronics & Information Technology,2009,31(1):134-138.

[13] LI L D, QIAN J S, PAN J S. Characteristic region based watermark embedding with RST invariance and high capacity[J]. International Journal of Electronics and Communications,2011,65(5):435-442.

[14] CHANG C C, LIN P Y, YEH J S. Preserving robustness and removability for digital watermarks using subsampling and difference correlation[J]. Information Sciences, 2009, 179(13):2283-2293.

[15] THORAT C G, JADHAV B D. A blind digital watermark technique for color image based on integer wavelet transform and SIFT[A]. Proceedings of the International Conference and Exhibition on Biometrics Technology[C]. Coimbatore, TamilNadu, India, Procedia Computer Science, 2010. 236-241.

[16] WU X Y, GUAN Z H. A novel digital watermark algorithm based on chaotic maps[J]. Physics Letters A, 2007, 365(5-6):403-406.

[17] 叶天语.离散余弦变换域抗二次打印-扫描顽健零水印算法[J].光子学报,2011,40(1):142-148.
YE T Y. A robust zero-watermarking algorithm against dual print-and-scan process based on discrete cosine transformation[J]. Acta Photonica Sinica, 2011,40(1):142-148.

[18] 叶天语,马兆丰,钮心忻,杨义先.强顽健零水印技术[J].北京邮电大学学报,2010,33(3):126-129.
YE T Y, MA Z F, NIU X X,*et al*. A zero-watermark technology with strong robustness[J]. Journal of Beijing University of Posts and Telecommunications, 2010, 33(3):126-129.

[19] YE T Y. A watermarking algorithm for certificate forgery prevention[A]. Proceedings of the International Conference on Consumer Electronics Communications and Networks[C]. Xianning, China, 2011. 3888-3890.

[20] YE T Y. Robust zero-watermark algorithms in hybrid transform domains based on the parity of norm's highest digit[A]. Proceedings of the First International Conference on Electrical and Electronics Engineering[C]. Wuhan, China, 2011.73-81.

[21] 李庆诚,窦毅.数字水印的解释攻击与关联性特征[J].计算机应用, 2005, (5):115-117.
LI Q C, DOU Y. Interpretation attack and relating characteristic of digital watermarking[J]. Application Research of Computers, 2005, (5):115-117.

[22] 张鸿宾,杨成.图像的自嵌入及窜改的检测和恢复算法[J].电子学报,2004,32(2):196-199.
ZHANG H B, YANG C. Tamper detection and self-recovery of images using self-embedding[J]. Acta Electronica Sinica, 2004 ,32(2):196-199.

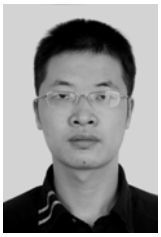
[23] 和红杰,张家树.基于混沌置乱的分块自嵌入水印算法[J].通信学报,2006,27(7):80-86,93.
HE H J, ZHANG J S. Chaos-based scramble self-embedding watermarking algorithm[J]. Journal on Communications, 2006, 27(7): 80-86, 93.

[24] 张宪海,杨永田.基于脆弱水印的图像认证算法研究[J].电子学报, 2007,35(1):34-39.
ZHANG X H, YANG Y T. Image authentication scheme research based on fragile watermarking[J]. Acta Electronica Sinica, 2007, 35(1):34-39.

[25] 王国栋, 刘粉林, 刘媛等. 一种能区分水印或内容篡改的脆弱水印算法[J]. 电子学报,2008, 36(7):1349-1354.
WANG G D, LIU F L, LIU Y, *et al*. An image authentication scheme with discrimination of tampers on watermark or image[J]. Acta Electronica Sinica, 2008, 36(7):1349-1354.

[26] 和红杰,张家树.对水印信息篡改顽健的自嵌入水印算法[J].软件学报,2009,20(2):437-450.
HE H J, ZHANG J S. Self-embedding watermarking algorithm with robustness against watermark information alterations[J]. Journal of Software ,2009,20(2):437-450.

作者简介：



叶天语（1982-），男，浙江温州人，博士，浙江工商大学讲师，主要研究方向为信息隐藏与数字水印。