doi:10.3969/j.issn.1000-436x.2013.03.019

基于子块区域分割和自嵌入技术的全盲多功能图像水印算法

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：利用子块区域分割思想和自嵌入技术提出了一种只利用攻击图像就能进行版权鉴别和内容认证的全盲多功能图像水印算法。将原始图像分割成不重叠的子块，对每个子块进行离散余弦变换，将每个子块离散余弦变换矩阵分成为区域 1 和区域 2 两部分，然后对区域 1 进行奇异值分解，通过判断奇异值均值的最高位数字奇偶性产生特征水印，然后通过调整区域 2 的 2 个离散余弦变换交流系数大小自嵌入特征水印，最后对每个子块进行逆离散余弦变换得到含水印图像。算法通过结合将区域 1 产生的特征水印自嵌入区域 2 和在区域 2 盲提取认证水印实现全盲检测和多功能。实验结果表明，算法既能实现版权保护又能实现内容认证，而且还具有区分恶意篡改和无意篡改的能力。

关键词：数字水印；多功能水印；全盲检测；区域分割；自嵌入技术
中图分类号：TN911.7 **文献标识码**：B **文章编号**：1000-436X(2013)03-0148-09

Perfectly blind image watermarking scheme with multi-purpose based on region segment for sub-block and self-embedding technology

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

Abstract: A perfectly blind image watermarking scheme with multi-purpose was proposed through region segment for sub-block and self-embedding technology, which could achieve copyright identification and content authentication only with the help of an attacked image. The original image was split into non-overlapping blocks, and each block was conducted with discreet cosine transformation. Each block's discreet cosine transformation matrix was divided into region 1 and region 2. Then, after the region 1 was conducted with singular value decomposition, a feature watermark was derived from judging the parity of the first digit from average of singular values. Moreover, the feature watermark was self-embedded into region 2 by adjusting two discreet cosine transformation alternative current coefficients. Finally, a watermarked image was obtained after conducting inverse discrete cosine transformation on each block. The proposed algorithm achieved perfectly blind detection and multi-purpose by combining self-embedding feature watermark from region 1 into region 2 and blindly extracting authentication watermark in region 2. Experimental results show that the proposed algorithm can achieve both copyright protection and content authentication, and has the ability to distinguish malicious tamper and unintentional tamper.

Key words: digital watermarking; multi-purpose watermark; perfectly blind detection; region segment; self-embedding technology

1 引言

根据数字水印的顽健性，数字水印算法一般分为顽健水印算法、脆弱水印算法和半脆弱水印算法 3 类。顽健水印算法[1~8]用于对数字载体实现版权保护；脆弱水印算法[9~11]用于对数字载体实现内容认证；半脆弱水印算法[12,13]用于对数字载体实现一定程度的版权保护和内容认证。但是，半脆弱水印算

收稿日期：2011-12-22；修回日期：2012-03-13
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)

法的局限性在于只嵌入一个半脆弱水印时很难同时具备很强的顽健性和很敏感的脆弱性。为了克服这一缺陷，多功能水印算法[14]被提出来以同时具备很强的顽健性和很敏感的脆弱性，从而能够同时实现版权保护和内容认证。

顽健图像水印算法[1~8]在嵌入端将外在的水印嵌入到原始图像，在检测端通过计算原始水印和提取的水印之间的相关度来鉴别版权。虽然提取水印时不需要用到原始图像的任何信息，盲无意义顽健水印算法[1]还是要通过借助密钥产生原始伪随机数水印序列来计算与提取水印之间的相关度以鉴别版权，盲有意义顽健水印技术[2~8]还是需要将原始水印从嵌入端传输到检测端来计算与提取水印之间的相关度以鉴别版权。本文认为一个实用的顽健图像水印算法应该做到全盲检测，即检测端不需借助原始图像和原始水印的任何信息，可以节省传输成本和防止互联网上的被动攻击。然而，目前的盲顽健水印算法[1~8]和多功能水印算法[14]仍需要借助原始水印或其相关信息，无法达到全盲检测。

脆弱数字水印领域的自嵌入技术是指嵌入端提取原始载体的某种特征量产生水印并将其自嵌入到原始载体，而不是将外在的水印嵌入到原始载体。自嵌入技术的 2 个关键要素是：1）水印是通过提取原始载体的特征产生；2）水印被自嵌入原始载体用以建立一定的映射关系。自嵌入脆弱水印算法[9~11]可以达到全盲检测，其原因在于采用了自嵌入技术。然而，目前自嵌入技术仅仅被运用到脆弱水印领域。

综合以上分析，本文的研究目的是将自嵌入技术引入顽健水印领域以设计一种全盲多功能水印算法，达到既能同时实现版权保护和内容认证，又能实现全盲检测的目的。实验结果表明，本文算法既能同时实现版权保护和内容认证，又能达到全盲检测。

2　基于子块区域分割和自嵌入技术的全盲多功能图像水印算法

2.1　算法原理

以一个 8×8 图像子块为例说明图像子块区域分割的原理。一个 8×8 图像子块共包含 64 个像素。为描述方便，将该 8×8 图像子块记为 $\boldsymbol{R}$ 。 $\boldsymbol{R}$ 由 64 个小方块组成，每个小方块由细线条围成，代表一个像素，如图 1 所示。将 $\boldsymbol{R}$ 分割成区域 1 和区域 2 两部分，区域 1 由前 3 行前 3 列的小方块组成，区域 2 由剩余部分的小方块组成，区域 1 和区域 2 的边界都用粗线条表示。显然，$\boldsymbol{R}$ = 区域1∪区域2，而且区域1∩区域2 = ∅ 。这样就完成了对 $\boldsymbol{R}$ 进行的区域分割。

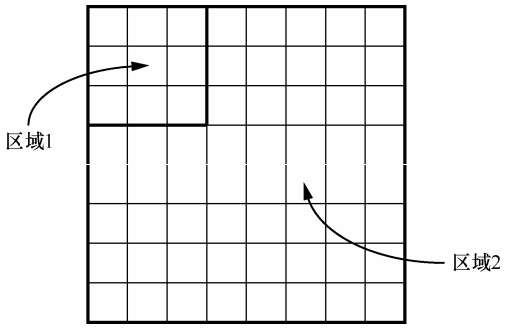


图 1　图像子块区域分割

将原始图像划分成互不重叠的子块，对每个子块进行 DCT，然后将每个子块 DCT 矩阵按照上述子块区域分割方法分割成区域 1 和区域 2。从区域 1 中提取特征产生特征水印，再自嵌入到区域 2 中。

对每个子块DCT矩阵进行区域分割使得区域1和区域 2 不重叠的目的是：保证特征水印自嵌入到区域 2 且不会对区域 1 有任何影响，进而不会对检测端从区域 1 提取特征水印造成任何影响。只要区域 1 产生特征水印的算法具有很强的抗攻击顽健性，检测端就能从区域 1 中以与原始特征水印很高的相关度提取出特征水印。另外，只要区域 2 的特征水印自嵌入算法具有很强的抗攻击顽健性和达到盲提取，检测端就能从区域 2 中以与原始特征水印很高的相关度盲提取出认证水印。这样，检测端只需要攻击图像就可以分别从区域 1 和区域 2 中提取出特征水印和认证水印，实现全盲检测；检测端通过计算提取出的特征水印和认证水印之间的相关度就可以进行版权鉴别，通过判断每比特提取的特征水印和认证水印之间的一致性就可以进行内容认证，从而实现多功能。

2.2　算法描述

2.2.1　特征水印产生

本文算法结合 DCT 的能量集中特性和 SVD 的奇异值稳定性产生特征水印。对原始图像每个子块进行 DCT，由每个子块 DCT 矩阵的直流系数和低频交流系数构成区域 1，再对区域 1 进行 SVD，通过判断奇异值均值的最高位奇偶性产生特征水印。具体步骤如下。

步骤 1 将大小为 $N\times N$ 的原始载体图像 $\boldsymbol{B}$ 划分为不重叠的 $n\times n$ 子块，每个子块记为 $\boldsymbol{B}_j$，$j=1,2,\cdots,\left(\frac{N}{n}\right)^2$。

步骤 2 对 $\boldsymbol{B}_j$ 进行 DCT，将得到的 DCT 矩阵记为 $\hat{\boldsymbol{B}}_j$。

步骤 3 对 $\hat{\boldsymbol{B}}_j$ 中由前 m 行前 m 列元素组成的区域 1 进行 SVD，得到的各奇异值为 $\lambda_j^1,\lambda_j^2,\cdots,\lambda_j^m$，其平均值记为 v_j，则

$$v_j=\frac{\lambda_j^1+\lambda_j^2+\cdots+\lambda_j^m}{m} \tag{1}$$

步骤 4 判断 v_j 的最高位数字奇偶性产生特征水印 $\boldsymbol{W}$。如果 v_j 的最高位数字为奇数，则 $w_j=1$；否则，$w_j=0$，w_j 为 $\boldsymbol{W}$ 的第 j 比特。例如：如果 $v_j=62$，最高位为 6，此时 $w_j=0$。

根据上述过程，$\boldsymbol{W}$ 的长度为 $\left(\frac{N}{n}\right)^2$ bit。由于 $\boldsymbol{W}$ 是利用原始载体图像每个子块 DCT 矩阵区域 1 自身的特征产生，所以在本文中被称为特征水印。

2.2.2 特征水印自嵌入

为了增强本文算法的安全性，特征水印自嵌入前先利用混沌映射对其进行加密处理。Logistic 混沌映射定义为

$$y_{n+1}=1-\nu y_n^2 \tag{2}$$

其中，y_0 为初值，ν 为分支参数。本文算法通过调整每个子块 DCT 矩阵区域 2 的 2 个 DCT 系数自嵌入区域 1 产生的特征水印。具体步骤如下。

步骤 1 利用 Logistic 映射对 $\boldsymbol{W}$ 进行加密，加密过程说明如下。

1) 选定初值 y_0 和分支参数 ν，利用式（2）产生混沌序列 $\boldsymbol{Y}=\{y_1,y_2,y_3,\cdots\}$。

2) 将第 $t+1$ 到第 $t+\left(\frac{N}{n}\right)^2$ 个随机数 y_{t+1}，$y_{t+2},\cdots,y_{t+\left(\frac{N}{n}\right)^2}$ 二值化为 {0,1} 序列 $\boldsymbol{L}$，即

$$l_j=\frac{\mathrm{sgn}(y_{t+j})+1}{2} \tag{3}$$

其中，sgn(·) 为符号函数，l_j 为 $\boldsymbol{L}$ 的第 j 比特，$j=1,2,\cdots,\left(\frac{N}{n}\right)^2$。混沌序列最初的随机数往往不大稳定，所以式（3）舍去混沌序列的前 t 个随机数。将 y_0 和 ν 作为本文算法的 2 个密钥。

3) 利用序列 $\boldsymbol{L}$ 对特征水印 $\boldsymbol{W}$ 进行加密，即

$$\tilde{w}_j=w_j\oplus l_j \tag{4}$$

其中，$\oplus$ 为异或运算，$\tilde{w}_j$ 为加密后的特征水印 $\tilde{\boldsymbol{W}}$ 的第 j 比特。

步骤 2 将 $\boldsymbol{B}$ 划分为不重叠的 $n\times n$ 子块，每个子块记为 $\boldsymbol{B}_j$。

步骤 3 对 $\boldsymbol{B}_j$ 进行 DCT，将得到的 DCT 矩阵记为 $\hat{\boldsymbol{B}}_j$。

步骤 4 将 $\tilde{\boldsymbol{W}}$ 自嵌入每个 DCT 矩阵 $\hat{\boldsymbol{B}}_j$ 的区域 2，即

1) 当 $\tilde{w}_j=0$ 且 $\varphi_1=\hat{B}_j(s_1,r_1)-\hat{B}_j(s_2,r_2)<\beta$ 时，令

$$\begin{cases}\hat{B}_j'(s_2,r_2)=\hat{B}_j(s_2,r_2)-(\beta-\varphi_1)/2\\ \hat{B}_j'(s_1,r_1)=\hat{B}_j(s_1,r_1)+(\beta-\varphi_1)/2\end{cases} \tag{5}$$

当 $\tilde{w}_j=0$ 且 $\varphi_1=\hat{B}_j(s_1,r_1)-\hat{B}_j(s_2,r_2)\geqslant\beta$ 时，不做改变。

2) 当 $\tilde{w}_j=1$ 且 $\varphi_2=\hat{B}_j(s_2,r_2)-\hat{B}_j(s_1,r_1)<\beta$ 时，令

$$\begin{cases}\hat{B}_j'(s_2,r_2)=\hat{B}_j(s_2,r_2)+(\beta-\varphi_2)/2\\ \hat{B}_j'(s_1,r_1)=\hat{B}_j(s_1,r_1)-(\beta-\varphi_2)/2\end{cases} \tag{6}$$

当 $\tilde{w}_j=1$ 且 $\varphi_2=\hat{B}_j(s_2,r_2)-\hat{B}_j(s_1,r_1)\geqslant\beta$ 时，不做改变。

在上述过程中，β 用于控制水印嵌入强度。$\hat{B}_j(s_i,r_i)$ 为第 j 个子块原始 DCT 矩阵 $\hat{\boldsymbol{B}}_j$ 处于 (s_i,r_i) 的系数，$\hat{B}_j'(s_i,r_i)$ 为第 j 个子块修改后的 DCT 矩阵 $\hat{\boldsymbol{B}}_j'$ 处于 (s_i,r_i) 的系数，$i=1,2$。为了保证特征水印自嵌入在每个子块 DCT 矩阵中不同的 2 个系数，要求 $s_1\neq s_2$ 和 $r_1\neq r_2$ 至少有一个成立。另外，为保证特征水印自嵌入在每个子块 DCT 矩阵的区域 2，对于同一个 i，要求 $s_i>m$ 和 $r_i>m$ 至少有一个成立。

步骤 5 将 $\hat{\boldsymbol{B}}_j'$ 进行 IDCT 后重组产生含水印图像。

2.2.3 特征水印提取

检测端从大小为 $N\times N$ 的攻击图像 $\boldsymbol{B}'$ 提取特

征水印的过程与嵌入端从原始载体图像 $\boldsymbol{B}$ 产生特征水印的过程类似。具体步骤如下。

步骤 1　将 $\boldsymbol{B}^{'}$ 划分为不重叠的 $n\times n$ 子块，每个子块记为 $\boldsymbol{B}_j^{'}$，$j=1,2,\cdots,\left(\frac{N}{n}\right)^2$。

步骤 2　对 $\boldsymbol{B}_j^{'}$ 进行 DCT，将得到的 DCT 矩阵记为 $\hat{\boldsymbol{B}}_j^{'}$。

步骤 3　对 $\hat{\boldsymbol{B}}_j^{'}$ 中由前 m 行前 m 列元素组成的区域 1 进行 SVD，得到的各奇异值为 $\lambda_j^{'1},\lambda_j^{'2},\cdots,\lambda_j^{'m}$，其平均值记为 $v_j^{'}$，则

$$v_j^{'}=\frac{\lambda_j^{'1}+\lambda_j^{'2}+\cdots+\lambda_j^{'m}}{m} \tag{7}$$

步骤 4　判断 $v_j^{'}$ 的最高位数字奇偶性产生特征水印 $\boldsymbol{W}^{'}$。如果 $v_j^{'}$ 的最高位数字为奇数，则 $w_j^{'}=1$；否则 $w_j^{'}=0$，$w_j^{'}$ 为 $\boldsymbol{W}^{'}$ 的第 j 比特。

2.2.4　认证水印提取、版权鉴别和内容认证

检测端提取认证水印、进行版权鉴别和内容认证的过程分解为以下 6 个过程。

步骤 1　将 $\boldsymbol{B}^{'}$ 划分为不重叠的 $n\times n$ 子块，每个子块记为 $\boldsymbol{B}_j^{'}$，$j=1,2,\cdots,\left(\frac{N}{n}\right)^2$。

步骤 2　对 $\boldsymbol{B}_j^{'}$ 进行 DCT，将得到的 DCT 矩阵记为 $\hat{\boldsymbol{B}}_j^{'}$。

步骤 3　从 $\hat{\boldsymbol{B}}_j^{'}$ 的区域 2 中提取认证水印 $\boldsymbol{W}^{''}$

$$w_j^{''}=\begin{cases}1, & \hat{B}_j^{'}(s_2,t_2)\geqslant\hat{B}_j^{'}(s_1,t_1)\\ 0, & \text{其他}\end{cases} \tag{8}$$

其中，$w_j^{''}$ 代表 $\boldsymbol{W}^{''}$ 的第 j 比特，$\hat{B}_j^{'}(s_i,r_i)$ 为第 j 个子块 DCT 矩阵 $\hat{\boldsymbol{B}}_j^{'}$ 处于 (s_i,r_i) 的系数，$i=1,2$。

步骤 4　利用 Logistic 映射对认证水印 $\boldsymbol{W}^{''}$ 进行解密。解密过程说明如下。

1) 利用 y_0 和 ν 2 个密钥根据式（2）产生混沌序列 $\boldsymbol{Y}=\{y_1,y_2,y_3,\cdots\}$，舍弃前 t 个随机数，将第 $t+1$ 到第 $t+\left(\frac{N}{n}\right)^2$ 个随机数 $y_{t+1},y_{t+2},\cdots,y_{t+\left(\frac{N}{n}\right)^2}$ 通过式（3）二值化为 {0,1} 序列 $\boldsymbol{L}$。

2) 对认证水印 $\boldsymbol{W}^{''}$ 进行解密，即

$$\tilde{w}_j^{''}=w_j^{''}\oplus l_j \tag{9}$$

其中，$\tilde{w}_j^{''}$ 代表解密后的认证水印 $\tilde{\boldsymbol{W}}^{''}$ 的第 j 比特。

步骤 5　计算 $\boldsymbol{W}^{'}$ 和 $\tilde{\boldsymbol{W}}^{''}$ 之间的归一化相关度（NC, normalized correlation）进行版权鉴别。NC 定义为

$$NC=\left(\sum_{j=1}^{\left(\frac{N}{n}\right)^2}(w_j^{'}\cdot\tilde{w}_j^{''})\right)/\left(\sqrt{\sum_{j=1}^{\left(\frac{N}{n}\right)^2}(w_j^{'})^2}\cdot\sqrt{\sum_{j=1}^{\left(\frac{N}{n}\right)^2}(\tilde{w}_j^{''})^2}\right) \tag{10}$$

步骤 6　通过判断每个子块提取的特征水印 $w_j^{'}$ 与解密后的认证水印 $\tilde{w}_j^{''}$ 之间是否相同来检测此子块是否遭到篡改。即如果 $w_j^{'}=\tilde{w}_j^{''}$ 成立，则认为此子块没有遭到篡改；反之，则认为此子块遭到篡改，遭篡改的子块用全黑标识。

由以上过程可知，检测端提取认证水印的过程是嵌入端自嵌入特征水印的逆过程，而且提取认证水印实现盲提取。$\boldsymbol{W}^{''}$ 解密后用于版权鉴别和内容完整性认证，所以被称为认证水印。

由 2.2.3 节和 2.2.4 节可知，检测端在无需借助原始图像和原始水印任何信息的前提下，只利用攻击图像就可以分别提取特征水印和认证水印进行版权鉴别和内容完整性认证。因此，本文算法既具有多功能，又可达到全盲检测。

3　实验结果

采用 512×512 大小的 256 级灰度 Goldhill.bmp、Peppers.bmp、Couple.bmp 3 幅图像作为原始载体图像，分别如图 2(a)、图 3(a)和图 4(a)所示。原始载体图像的分块大小为 8×8，对每个子块 DCT 矩阵的前 3 行前 3 列元素组成的区域 1 进行 SVD；Logistic 映射的初值 y_0 为 0.208 8，分支参数 ν 为 2，舍弃前 300 个混沌随机数；自嵌入特征水印的位置为每个子块 DCT 矩阵的区域 2 中处于（5,1）和（4,1）2 个系数；3 幅图像的水印嵌入强度 β 分别为 30、35 和 25。含水印 Goldhill、Peppers、Couple 图像分别如图 2(b)、图 3(b)和图 4(b)所示，与图 2(a)、图 3(a)和图 4(a)之间的 *PSNR* 分别为 36.662 5dB、36.265 8dB、36.133 9dB，所以此时本文算法具有很好的不可见性。不存在攻击时，从图 2(b)、图 3(b)和图 4(b)提取的特征水印 $\boldsymbol{W}^{'}$ 和解密的认证水印 $\tilde{\boldsymbol{W}}^{''}$ 都完全一致，它们之间的 *NC* 值都为 1，此时没有发现遭到任何篡改，定位出的遭篡改子块为 0 个。

(a) 原始 (b) 含水印

图 2 Goldhill 图像

(a) 原始 (b) 含水印

图 3 Peppers 图像

(a) 原始 (b) 含水印

图 4 Couple 图像

3.1 篡改检测实验

3.1.1 恶意篡改检测实验

按照篡改目的，篡改可分为无意篡改和恶意篡改，前者一般指添加噪声、JPEG 压缩等常规信号处理，后者一般指剪切—粘贴、叠加等操作。对 Goldhill、Peppers、Couple 3 幅图像进行恶意篡改检测实验。将原始载体图像作为背景，检测到的遭篡改子块用黑色标识。限于篇幅，这里只列出 Goldhill 图像的篡改检测实验结果图。

A1：将图 5 的 32×32 hand.bmp 二值图像每个像素乘以 50 再分别叠加到图 2(b)、图 3(b)和图 4(b)的(189:220,165:196) 区域，Goldhill 图像的篡改图像和篡改定位图像分别如图 6(a)和图 6(b)所示。此时本文算法可准确定位出篡改区域。

图 5 Hand 图像

(a) 篡改图像 (b) 定位图像

图 6 Goldhill 图像的 A1 实验结果

A2：将图 2(b)、图 3(b)和图 4(b)的(290:340, 155:205)区域每个像素数值加上 30，Goldhill 图像的篡改图像和篡改定位图像分别如图 7(a)和图 7(b)所示。此时本文算法可准确定位出篡改区域。

(a) 篡改图像 (b) 定位图像

图 7 Goldhill 图像的 A2 实验结果

A3：将图 8 的 64×64 Kids.bmp 灰度图像替换图 2(b)、图 3(b)和图 4(b)的(60:123,23:86) 区域，Goldhill 图像的篡改图像和篡改定位图像分别如图 9(a)和图 9(b)所示。可见此时本文算法可准确定位出篡改区域。

图 8 Kids 图像

(a) 篡改图像 (b) 定位图像

图 9 Goldhill 图像的 A3 实验结果

A4：剪切图 2(a)、图 3(a)和图 4(a)的(367:416, 326:385)区域替换图 2(b)、图 3(b)和图 4(b)的相同区域，Goldhill 图像的篡改图像和篡改定位图像分别如图 10(a)和图 10(b)所示。由于本文算法具有

良好的不可见性，所以从图 10（a）无法利用肉眼发现篡改。可见此时本文算法可准确定位出篡改区域。

(a) 篡改图像　　(b) 定位图像

图 10　Goldhill 图像的 A4 实验结果

A5：将图 2(b)、图 3(b)和图 4(b)的(69:124,45:90)区域替换自身的(121:176,121:166)区域，Goldhill 图像的篡改图像和篡改定位图像分别如图 11(a)和图 11(b)所示。可见此时本文算法可准确定位出篡改区域。

(a) 篡改图像　　(b) 定位图像

图 11　Goldhill 图像的 A5 实验结果

从 A1~A5 可以看出，本文算法能够较好地实现内容的完整性认证。

3.1.2　篡改类型区别

为了定量地度量遭篡改程度，定义篡改评估函数（TAF, tamper assessment function）为

$$TAF = \frac{k}{\text{图像}8\times8\text{子块总数}} \tag{11}$$

其中，k 为检测到的遭篡改 8×8 子块的总数。k 越大，TAF 也越大，表示篡改的越厉害。表 1 列出了 3 幅图像 A1~A5 恶意篡改实验的 TAF。可见，对于 3 幅实验图像，A3 篡改的最为厉害。

利用 TAF 定量度量本文表 2~表 4 各种常规信号处理的篡改程度，实验结果如各表中的“TAF”栏所示。

为区分恶意篡改和无意篡改，设定阈值 ε，如果 $TAF \leqslant \varepsilon$，则认为含水印图像遭到恶意篡改；反之，则认为遭到无意篡改。ε 的取值根据实际应用对图像内容完整性的要求而定。本文将 ε 取值为 0.012 5，表 1~表 4“篡改类型”栏列出此时各种情形下的篡改类型，本文算法能正确区分不同的篡改类型。

表 1　　**恶意篡改检测评估函数**

图像类别	实验项目	A1	A2	A3	A4	A5
Goldhill	*TAF*	0.004 2	0.006 8	0.009 5	0.004 6	0.007 8
	篡改类型	恶意	恶意	恶意	恶意	恶意
Peppers	*TAF*	0.005 1	0.007 8	0.011 0	0.005 6	0.008 3
	篡改类型	恶意	恶意	恶意	恶意	恶意
Couple	*TAF*	0.003 7	0.009 8	0.010 5	0.004 9	0.005 9
	篡改类型	恶意	恶意	恶意	恶意	恶意

3.2　抗攻击顽健性实验

表 2~表 4 分别列出了 3 幅图像抗攻击顽健性测试实验结果。尺度缩放都采用最近插值法。随机删除列是指从被删除列的右边第一列开始逐列向左移动，空余列补全黑。各表中 2 个水印所在栏的数据为它们之间的 NC。由表 2~表 4 的“$\boldsymbol{W}'$和$\tilde{\boldsymbol{W}}''$”栏可知，在各种攻击下，3 幅图像的$\boldsymbol{W}'$和$\tilde{\boldsymbol{W}}''$之间的 NC 都较高。因此，本文算法此时具有较强的顽健性。

3.3　实验结果讨论与分析

1) 特征水印嵌入位置选择

每个 8×8 子块 DCT 矩阵的前 3 行前 3 列元素组成的区域构成区域 1，特征水印产生于区域 1；剩余区域为区域 2，特征水印嵌入在区域 2 中。根据图像 JPEG 压缩原理，DCT 交流系数按照 Zigzag 扫描的顺序衰减，所以区域 2 内处于（4,1）和（5,1）的系数是该区域具有最多能量的 2 个系数，具有更大的感觉容量。因此，本文算法将特征水印嵌入在这 2 个系数中有利于提高算法的抗攻击顽健性。

2) 特征水印嵌入强度选择

特征水印嵌入强度 β 的大小与抗攻击顽健性和不可见性相关。β 越小，顽健性越弱，不可见性越好；β 越大，不可见性越差，但顽健性越强。因此，β 的选择是在顽健性和不可见性之间的一个折中。不同的原始图像可能具有较大的差异，所以无法用统一的公式来确定 β，一般根据实际应用对顽健性和不可见性的要求采用实验的方法确定。在本文中，当 3 幅图像的水印嵌入强度 β 分别为 30、35 和 25 时，从图 2(b)、图 3(b)和图 4(b)可知此时本文算法具有良好不可见性；从表 2~表 4 的“$\boldsymbol{W}'$和$\tilde{\boldsymbol{W}}''$”栏可知，此时本文算法具有较强的顽健性。

表 2　　**Goldhill 图像抗攻击顽健性实验结果**

算法	实验项目	平滑		添加噪声				JPEG 压缩		
				高斯噪声		椒盐噪声		质量因子		
		高斯滤波(窗口大小为 3×3, $\sigma=1$)	中值滤波（窗口大小为 3×3）	均值为 0, 方差为 0.000 1	均值为 0, 方差为 0.000 3	噪声密度为 0.001	噪声密度为 0.003	80	60	40
本文算法	$\boldsymbol{W}'$ 和 $\tilde{\boldsymbol{W}}''$	0.921 2	0.915 9	0.987 4	0.982 5	0.977 8	0.967 6	0.962 5	0.961 5	0.962 3
	$\boldsymbol{W}$ 和 $\boldsymbol{W}'$	0.951 5	0.963 8	0.987 4	0.982 5	0.984 1	0.979 6	0.962 5	0.961 5	0.962 3
	$\boldsymbol{W}$ 和 $\tilde{\boldsymbol{W}}''$	0.965 2	0.944 9	1.000 0	1.000 0	0.993 3	0.984 8	1.000 0	1.000 0	1.000 0
	TAF	0.085 9	0.091 6	0.013 7	0.019 0	0.024 2	0.035 4	0.041 0	0.042 0	0.041 0
	篡改类型	无意	无意	无意	无意	无意	无意	无意	无意	无意
算法[14]	*NC*	0.666 6	0.741 5	0.927 0	0.936 4	0.926 4	0.917 5	0.905 9	0.843 6	0.774 0

算法	实验项目	剪切		重采样	随机删除列		混合攻击	
		左上角 1/16	左上角 1/8	先缩小到 0.5 倍再放大到 2 倍	1 列	2 列	JPEG 70+中心剪切 1/16	JPEG 70+随机删除 1 列
本文算法	$\boldsymbol{W}'$ 和 $\tilde{\boldsymbol{W}}''$	0.972 7	0.944 6	0.947 0	0.967 2	0.929 4	0.928 9	0.940 5
	$\boldsymbol{W}$ 和 $\boldsymbol{W}'$	0.973 4	0.948 7	0.953 0	0.969 9	0.942 9	0.926 0	0.942 9
	$\boldsymbol{W}$ 和 $\tilde{\boldsymbol{W}}''$	0.975 6	0.948 9	0.992 6	0.996 4	0.979 6	0.970 8	0.997 1
	TAF	0.029 3	0.059 1	0.057 9	0.035 6	0.076 9	0.075 2	0.064 9
	篡改类型	无意	无意	无意	无意	无意	无意	无意
算法[14]	*NC*	0.902 0	0.877 6	0.803 7	0.925 1	0.920 0	0.842 9	0.881 6

表 3　　**Peppers 图像抗攻击顽健性实验结果**

算法	实验项目	平滑		添加噪声				JPEG 压缩		
				高斯噪声		椒盐噪声		质量因子		
		高斯滤波(窗口大小为 3×3, $\sigma=1$)	中值滤波（窗口大小为 3×3）	均值为 0, 方差为 0.000 1	均值为 0, 方差为 0.000 3	噪声密度为 0.001	噪声密度为 0.003	80	60	40
本文算法	$\boldsymbol{W}'$ 和 $\tilde{\boldsymbol{W}}''$	0.932 5	0.939 7	0.984 2	0.972 7	0.972 8	0.962 4	0.944 1	0.939 5	0.934 1
	$\boldsymbol{W}$ 和 $\boldsymbol{W}'$	0.951 8	0.960 3	0.984 2	0.972 7	0.977 0	0.970 1	0.944 4	0.939 7	0.934 5
	$\boldsymbol{W}$ 和 $\tilde{\boldsymbol{W}}''$	0.974 3	0.976 4	0.999 6	0.999 6	0.994 9	0.989 4	0.999 4	0.999 4	0.999 2
	TAF	0.078 4	0.070 1	0.018 3	0.031 5	0.031 5	0.043 2	0.065 4	0.070 3	0.076 7
	篡改类型	无意	无意	无意	无意	无意	无意	无意	无意	无意
算法[14]	*NC*	0.744 7	0.769 4	0.884 9	0.898 1	0.885 6	0.886 9	0.865 5	0.813 9	0.774 7

算法	实验项目	剪切		重采样	随机删除列		混合攻击	
		左上角 1/16	左上角 1/8	先缩小到 0.5 倍再放大到 2 倍	1 列	2 列	JPEG 70+中心剪切 1/16	JPEG 70+随机删除 1 列
本文算法	$\boldsymbol{W}'$ 和 $\tilde{\boldsymbol{W}}''$	0.974 0	0.944 7	0.909 6	0.950 1	0.909 9	0.909 9	0.915 5
	$\boldsymbol{W}$ 和 $\boldsymbol{W}'$	0.967 6	0.921 3	0.912 7	0.951 4	0.913 7	0.903 3	0.918 5
	$\boldsymbol{W}$ 和 $\tilde{\boldsymbol{W}}''$	0.973 5	0.944 5	0.995 6	0.998 7	0.992 0	0.974 1	0.997 0
	TAF	0.029 3	0.059 1	0.104 7	0.057 9	0.104 7	0.100 8	0.099 4
	篡改类型	无意	无意	无意	无意	无意	无意	无意
算法[14]	*NC*	0.865 5	0.852 6	0.838 7	0.880 3	0.879 6	0.817 5	0.835 9

表 4　　Couple 图像抗攻击顽健性实验结果

算法	实验项目	平滑		添加噪声				JPEG 压缩		
				高斯噪声		椒盐噪声		质量因子		
		高斯滤波(窗口大小为 3×3, $\sigma=1$)	中值滤波（窗口大小为 3×3）	均值为 0, 方差为 0.000 1	均值为 0, 方差为 0.000 3	噪声密度为 0.001	噪声密度为 0.003	80	60	40
本文算法	$\boldsymbol{W}'$ 和 $\tilde{\boldsymbol{W}}''$	0.919 6	0.917 0	0.980 5	0.974 4	0.971 0	0.948 2	0.960 9	0.956 2	0.962 9
	$\boldsymbol{W}$ 和 $\boldsymbol{W}'$	0.953 2	0.968 8	0.984 4	0.979 3	0.981 8	0.970 4	0.962 4	0.958 2	0.964 8
	$\boldsymbol{W}$ 和 $\tilde{\boldsymbol{W}}''$	0.959 4	0.939 9	0.996 1	0.995 1	0.985 7	0.967 5	0.992 4	0.991 6	0.992 0
	TAF	0.100 3	0.103 3	0.024 4	0.032 0	0.036 1	0.064 0	0.048 6	0.054 4	0.046 1
	篡改类型	无意	无意	无意	无意	无意	无意	无意	无意	无意
算法[14]	NC	0.747 8	0.794 2	0.931 4	0.945 2	0.926 4	0.931 4	0.905 2	0.864 8	0.808 8

算法	实验项目	剪切		重采样	随机删除列		混合攻击	
		左上角 1/16	左上角 1/8	先缩小到 0.5 倍再放大到 2 倍	1 列	2 列	JPEG 70+中心剪切 1/16	JPEG 70+随机删除 1 列
本文算法	$\boldsymbol{W}'$ 和 $\tilde{\boldsymbol{W}}''$	0.976 0	0.949 6	0.932 3	0.967 2	0.926 2	0.927 4	0.936 0
	$\boldsymbol{W}$ 和 $\boldsymbol{W}'$	0.969 9	0.931 0	0.949 9	0.972 0	0.944 4	0.921 6	0.941 2
	$\boldsymbol{W}$ 和 $\tilde{\boldsymbol{W}}''$	0.976 8	0.951 4	0.971 7	0.994 0	0.977 3	0.960 9	0.986 7
	TAF	0.029 3	0.059 1	0.084 0	0.041 3	0.091 8	0.087 2	0.079 8
	篡改类型	无意	无意	无意	无意	无意	无意	无意
算法[14]	NC	0.902 6	0.890 2	0.822 5	0.920 0	0.927 0	0.854 6	0.882 3

3) 顽健性分析

存在各种攻击时，由表 2~表 4 的“$\boldsymbol{W}$ 和$\boldsymbol{W}'$”栏可知，特征水印产生算法表现出很强的抗攻击顽健性；由表 2~表 4 的“$\boldsymbol{W}$ 和$\tilde{\boldsymbol{W}}''$”栏可知，特征水印自嵌入算法也表现出很强的抗攻击顽健性；由表 2~表 4 的“$\boldsymbol{W}$ 和$\tilde{\boldsymbol{W}}''$”栏可知，本文算法表现出较强的抗攻击顽健性。本文算法具有较强抗攻击顽健性的原因就在于特征水印产生和自嵌入算法本身就具有很强的抗攻击顽健性。

与文献[14]算法进行抗攻击顽健性比较。为了具有可比性，这里使 2 个算法原始载体图像与含水印图像之间的 *PSNR* 基本相同。文献[14]算法设置如下：将图 5 的 hand.bmp 作为原始水印图像；Arnold 置乱次数为 2 次；对原始载体图像的每个大小为16×16的子块进行 SVD，将水印嵌入每个子块的后 15 个奇异值；将含顽健水印的图像分为大小为 2×2 的不重叠子块，产生并自嵌入脆弱水印；Goldhill、Peppers、Couple 3 幅图像的水印嵌入强度 α 分别为 19、20 和 23，此时原始图像和含双水印图像之间的 *PSNR* 分别为 36.617 6 dB、36.149 5 dB 和 36.055 0 dB，与本文算法基本相同。对比表 2~表 4 的“$\boldsymbol{W}'$和$\tilde{\boldsymbol{W}}''$”栏和“文献[14]算法”栏可知，此时本文算法的抗攻击顽健性强于文献[14]的算法。

4) 安全性分析

本文算法在自嵌入特征水印$\boldsymbol{W}$ 前先用 Logistic 混沌映射对其进行加密，共有 y_0 和$\nu 2$ 个密钥。由于 Logistic 混沌映射具有初值敏感性，不知道 y_0 和 $\nu 2$ 个密钥的非法者是无法证实版权归属和内容真实性的。

4　结束语

本文提出了一种全盲多功能图像水印算法，算法采用子块区域分割思想和自嵌入技术，检测端只需要利用攻击图像就能进行版权鉴别和内容认证。实验结果验证了算法的有效性。所提算法具有很好的实用性和多功能性，有利于推动数字水印技术的实际应用。

参考文献：

[1] 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.

[2] 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.

[3] 李旭东.基于小波变换和相对量化的图像数字水印算法[J]. 光电子.激光, 2010, 21(9):1378-1382.
LI X D. Image digital watermarking algorithm based on wavelet transform and relative quantization[J]. Journal of Optoelectronics Laser, 2010, 21(9):1378-1382.

[4] 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.

[5] 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.

[6] 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.

[7] BEI Y L, SAN S J. Color digital watermarking based on amplitude modulation and SVR[J]. Journal of Communication and Computer, 2010, 7(2):37-42.

[8] 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.

[9] 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.

[10] 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.

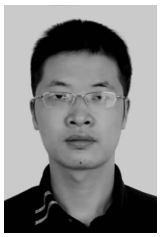
[11] 和红杰,张家树.对水印信息篡改顽健的自嵌入水印算法[J]. 软件学报, 2009, 20(2):437-450.
HE H J, ZHANG J S. Self-embedding watermarking algorithm with robustness against watermark information alterations[J]. Journal of Software, 2009, 20(2):437-450.

[12] SCHLAUWEG M, PROFROCK D, PALFNER T, *et al*. Quantization-based semi-fragile public-key watermarking for secure image authentication[A]. Proc of SPIE[C]. San Diego, California, USA, 2005. 41-51.

[13] 李春,黄继武.一种抗 JPEG 压缩的半脆弱图像水印算法[J]. 软件学报, 2006, 17(2):315-324.
LI C, HUANG J W. A semi-fragile image watermarking resisting to JPEG[J]. Journal of Software, 2006, 17(2):315-324.

[14] 叶天语,钮心忻,杨义先.多功能双水印算法[J]. 电子与信息学报, 2009, 31(3):546-551.
YE T Y, NIU X X, YANG Y X. A multi-purpose dual watermark algorithm[J]. Journal of Electronics & Information Technology, 2009, 31(3):546-551.

作者简介：



叶天语（1982-），男，浙江温州人，博士，浙江工商大学副教授，主要研究方向为信息隐藏与数字水印、量子保密通信等。