Computer Science and Application 计算机科学与应用, 2023, 13(8), 1507-1514
Published Online August 2023 in Hans. https://www.hanspub.org/journal/csa
https://doi.org/10.12677/csa.2023.138149

Hans汉斯

基于DCT的安全抗剪切攻击数字图像水印算法

李芸伟

北京印刷学院信息工程学院，北京

收稿日期：2023年7月3日；录用日期：2023年8月1日；发布日期：2023年8月7日

摘　要

为解决嵌入水印后不抗剪切攻击的问题，围绕Arnold置乱，结合分块的离散余弦变换(Discrete Cosine Transform, DCT)水印算法，本文提出了基于DCT域的安全抗剪切攻击的盲数字水印算法。首先将水印图像进行置乱加密，并对它进行上下对换、左右对换得到变换水印，进行二次水印嵌入，提取时对两个水印进行重组拼接来得到完整水印，从而提高鲁棒性。实验结果表明，在二次嵌水印后，只要剪切面积不超过图像大小的1/2，该算法对剪切攻击就有100%的抵抗力，并且透明性也能得到很好的保证。

关键词

数字图像水印，DCT，PSNR，BER，Arnold

Digital Image Watermarking Algorithm Based on DCT Security against Shear Attack

Yunwei Li

School of Information Engineering, Beijing Institute of Graphic Communication, Beijing

Received: Jul. 3rd, 2023; accepted: Aug. 1st, 2023; published: Aug. 7th, 2023

Abstract

To address the issue of not being resistant to cropping attacks after embedding watermarks, this paper proposes a secure blind digital watermarking algorithm based on Arnold and the block wise Discrete Cosine Transform (DCT) watermarking algorithm, which is secure against cropping attacks in the DCT domain. Firstly, the watermark image is scrambled and encrypted, and it is swapped up and down and left and right to obtain the transformed watermark, and the secondary watermark is embedded. When extracting, the two watermarks are reassembled and spliced to obtain the complete watermark, thus improving the robustness. The experimental results show that the algorithm has 100% resistance to the shear attack, and its transparency can be well

文章引用: 李芸伟. 基于 DCT 的安全抗剪切攻击数字图像水印算法[J]. 计算机科学与应用, 2023, 13(8): 1507-1514.
DOI: 10.12677/csa.2023.138149

guaranteed, as long as the shear area does not exceed 1/2 of the image size after the secondary watermarking.

Keywords

Digital Image Watermarking, DCT, PSNR, BER, Arnold

1. 引言

随着科技的快速发展，我们可以快速处理和接收电子信息，并进行各种数字媒体的编辑、修改、存储和传播。然而，这也引起了各种版权侵犯问题。因此，越来越多的人开始重视个人隐私和版权保护，并将保护版权作为数字水印工作者的核心目标。

数字水印的概念由 Tirkel 等人[1]于 1993 年提出，对该项技术的研究一直延续至今。Tirkel 等人[1]提出的最低有效位(Least Significant Bit, LSB)算法尽管实现方式简单，但在面对恶意攻击时的稳健性很差，容易造成水印提取不出来。Cox [2]等人提出了一种扩频通信的水印技术，把水印信息直接插入到数据的频谱分量中，提高了水印对恶意攻击的抵抗能力。Zhao [3]等提出对载体图像进行 8 × 8 大小分块，对子块进行 DCT 变换，每块嵌入一位水印，利用合适的步长对各子块的直流系数根据当前水印信息位进行量化，之后进行 DCT 反变换得到含水印图像。刘志军[4]等提出对图像进行 8 × 8 分块后进行 DCT 变换，将水印嵌入到人视觉上的最重要的地方，即图像的低频和中频区域。

基于中频系数替换的 DCT 域水印算法是一种常见的数字水印算法。该算法通过选择合适的 DCT 系数作为水印信息的嵌入位置，将水印信息嵌入到 DCT 系数中。其中，中频系数因为在图像中所占比例较大，因此更加适合嵌入水印信息。该算法的优点是具有较好的鲁棒性和透明性。同时，由于该算法的实现比较简单，因此在实际应用中得到了广泛的应用。然而，该算法也存在一些缺点，例如水印容量较小，且对于某些攻击(如低通滤波)容易受到破坏。因此我们需要对水印信息进行相应的处理，如加密、置乱等。

针对现有一些误码率高且不抗剪切的鲁棒水印算法，本文提出了一种基于加密且抗剪切的数字水印算法。首先对水印图像进行 Arnold 置乱，并通过 DCT 变换进行第一次水印嵌入，但初步的实验发现，这种单次嵌入不抗剪切攻击，所以我们进行了二次嵌入。第二遍水印的嵌入规则与之前一样，但是对水印图像进行了上下左右的变换。水印提取阶段，通过对两个水印图像进行重构即可得到完整的原始水印信息。

2. 数字水印特征及评价标准

数字水印算法的性能评价标准包括透明性、鲁棒性、不可伪造性和水印容量四个方面，我们在设计和评估水印方案时必须考虑以上四方面的重要因素，其中水印容量和鲁棒性是相互矛盾的一对评价指标。一般来说，提高水印容量，往往会降低水印鲁棒性；而提高水印的鲁棒性，则往往会减少水印容量。因此，在设计和选择数字水印算法时，必须在两者之间进行权衡和取舍，以达到最优的性能表现[5]。

2.1. 透明性

透明性指含水印图像和载体图像相比，不能有肉眼可见的视觉质量的下降，且不影响载体的正常使

用。一般用峰值信噪比(Peak Signal to Noise Ratio, PSNR)进行判定。其计算公式如(1)所示：

$$\mathrm{PSNR}=10\lg\left[\frac{M\times N\times 255^2}{\sum_{x-1}^{M}\sum_{y-1}^{N}\left[I'(x,y)-I(x,y)\right]^2}\right] \tag{1}$$

其中，$I(x,y)$为载体图像在(x,y)处的像素值，$I'(x,y)$为含水印图像同一位置像素值，M和N分别为图像的长和宽。

2.2. 鲁棒性

鲁棒性指水印在遭受各种有意或无意攻击后而不易被破坏，水印仍能保持部分完整性并被准确提取出来，可能的攻击形式包括剪切、信道噪声、滤波、JPEG 压缩、几何攻击等。水印鲁棒性的评价指标是误码率(Bit Error Rate, BER)，计算公式如(2)所示：

$$\mathrm{BER}=\frac{\sum_{i-1}^{M}\sum_{j-1}^{N}abs\left(W'(i,j)-W(i,j)\right)}{M\times N} \tag{2}$$

其中$W(x,y)$为原始二值水印，$W'(x,y)$为经过检测得到的二值水印信息，M 和 N 分别为水印图像的长和宽，$abs(\ )$为求绝对值函数。

2.3. 不可伪造性

不可伪造性是指数字水印应该是难以被伪造和删除的，以避免恶意攻击和盗版。具有不可伪造性的水印可以帮助保护数字媒体内容的版权和安全，从而保护数字媒体内容的创作者和所有者的利益。它能够唯一地标识原始图像的相关信息，任何第三方都不能伪造他人的水印信息。

2.4. 水印容量

水印容量是指在 PSNR 和 BER 允许范围内，水印信息所占用的数据比例，一般用水印嵌入率表示。对于图像水印而言，用嵌入的水印信息比特数与载体信息比特数的比值来表示嵌入率。视频水印则用每一帧图像中或者每几帧图像中可以嵌入的水印信息比特数来衡量[6]。

3. 算法概述

本文实验中将原始图像和水印图像都分为 8 × 8 的图像块，分别对每个图像块进行 DCT 变换，每个 DCT 变换后的图像块包含 64 个 DCT 系数，找出要嵌入水印的点的坐标和该坐标周围上下左右四个点的系数值，对这四个 DCT 系数求均值。因为实验发现，只嵌入一次水印时对剪切攻击无抵抗性，因此我们选择二次嵌入水印。

3.1. 水印预处理

置乱算法是一种相对简单的加密算法，但是由于是初始条件的改变，使得变换后的图像与原始图像找不到任何关联，因此当应用到水印系统时，可以起到安全性的作用[7]。置乱的算法有许多种类型，如利用 Arnold 变换、Baker 变换等算法[8]，本文对图像信息的预处理主要采用 Arnold 变换进行预处理。

Arnold 置乱又称猫脸变换(Cat Mapping)，是一种常见的数字图像加密技术。它通过多次应用一个特定的置换矩阵来改变图像像素的位置。该过程可以使图像的像素在空间上分散，从而增加了图像的随机性和复杂性。Arnold 置乱逻辑简单，可有效的降低系统的复杂度。其定义式如(3)所示：

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} (\mathrm{mod}\, N) \tag{3}$$

其中 $x, y \in \{0,1,\cdots,N-1\}$，$(x,y)$ 是原图像像素中的坐标，(x',y') 为经过变换后在新图像中的坐标位置，N 为图像矩阵的阶数。Arnold 置乱具有周期性，对于不同阶数的图像，置乱的周期不同，本实验中水印图像的置乱周期为 48。

3.2. 基于 DCT 域的盲水印嵌入

水印嵌入流程如图 1 所示。设原始图像 $I = \{I(x,y); 0 \le x \le M, 0 \le y \le N\}$，将 I 分为互不重叠的 8 × 8 的图像块，每个图像块用 Block 表示。对于 $M \times N$ 的图像 I 来说，一共分为 $(M \times N)/(8 \times 8)$ 个图像块，同时，水印图像 W 也要被分割。本实验将 I 分为 64 × 64 个 8 × 8 的图像块，W 也是 8 × 8 的图像块。对每个 Block 分别进行 DCT 变换，每一个 DCT 变换后的 Block 包含 64 个 DCT 系数，水印嵌入完成后得到的含水印图像记为 I'。

在嵌入水印阶段，首先要选择 Block 中的某个位置作为要嵌入水印的坐标点，以 Block(5,5)为例，找出该坐标周围上下左右四个点的系数值，对这四个 DCT 系数求均值，将均值记为 Emb_mean。值得注意的是，为保证水印信息的安全性，在嵌入水印之前，我们可以对水印进行加密操作，将含密水印嵌入载体图像中，具体的加密方式可以按需选择。如此，即使水印信息能被提取出来，也需要进一步的解密操作才能看到真实的水印内容。嵌入规则如式(4)所示，其中 Q 为嵌入强度[9]。将嵌入水印后的 DCT 系数做二维 IDCT 变换，所有 $(M \times N)/(8 \times 8)$ 个子块都处理完成后，即生成含水印的图像。

$$I'(x,y) = \begin{cases} \text{Emb\_mean} + Q, W(x,y) = 1\text{时} \\ \text{Emb\_mean} - Q, W(x,y) = 0\text{时} \end{cases} \tag{4}$$

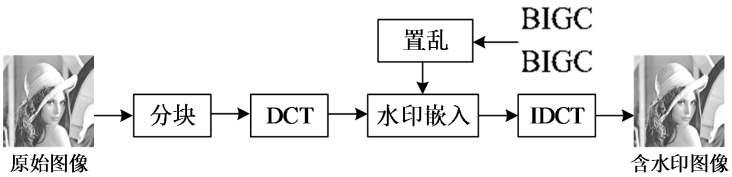


Figure 1. Flow chart with embedded blind watermark
图 1. 盲水印嵌入流程图

3.3. 基于 DCT 域的盲水印提取

水印提取流程如图 2 所示。预处理阶段，我们要对含水印图像进行分块处理，依旧分为互不重叠的 8 × 8 的块(记为 Block′)，对每个 Block'进行二维 DCT 变换。

提取水印阶段，首先找出各个块坐标位置为(5,4)、(5,6)、(4,5)、(6,5)四个点的 DCT 系数值，并对其做均值，记均值为 Ext_mean，按以下规则提取水印：当 Block′(5,5) > Ext_mean 时，提取水印信息为 0；当 Block′(5,5) < Ext_mean 时，提取水印信息为 1。对提取的水印信息逆置乱即可得到恢复后的水印图像。

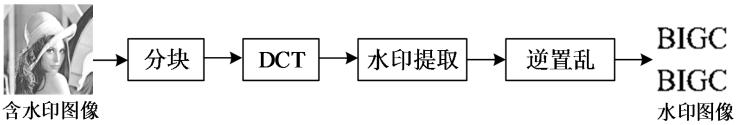


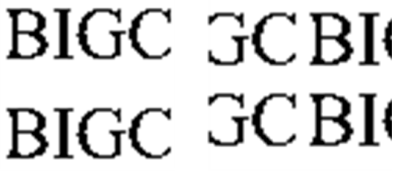
Figure 2. Flow chart with extracting blind watermark
图 2. 盲水印提取流程图

3.4. 水印二次嵌入

本实验首先尝试了对载体图像进行一次水印嵌入，但局限性表现在不抗剪切攻击，即在对图像进行剪切后，提取出的水印信息也会相应的缺失一块。这主要是和基于分块的 DCT 变换思想有关。经过分块以后，水印信息分别嵌入各个子块，最终使得水印在整个载体图像中的分布是均匀的。图 3 展示了不抗剪切的水印嵌入算法，如我们对含水印图像进行如图 3(a)所示的剪切攻击，水印信息就会不可避免地缺失对应位置的水印内容如图 3(b)。我们需要考虑如何改进才能使含水印图像抗剪切攻击？

(a)剪切含水印图像　　(b)水印提取图

Figure 3. Non-clipping sample diagram
图 3. 不抗剪切示例图



(a)水印变换前　　(b)水印变换后

Figure 4. Watermark transformation before and after comparison
图 4. 水印变换前后对比图

如图 4 所示，我们需要对水印图像(记为水印 1)进行上下左右对换，即可得到水印 2 [10]。将得到的水印 1 和水印 2 按图 5 的流程图依次进行水印嵌入，即可得到抗剪切攻击的含水印图像。由于嵌入水印 1 时，选择 Block(5,5)作为坐标点，因此我们选择 Block(3,3)进行水印 2 的嵌入。在水印提取阶段，我们要对提取出的两个水印图像进行重构，即可组成完整的水印信息，重构规则如公式(5)所示，其中 $x_1(i,j)$ 和 $x_2(i,j)$ 分别表示提取的水印 1 和水印 2，$y(i,j)$ 表示重组水印。图 6 展示了在嵌入水印的图像遭受剪切攻击后水印信息的提取样例。

$$\begin{cases} y(i,j)=x_1(i,j), & x_1(i,j)=x_2(i,j)\text{时} \\ y(i,j)=x_1(i,j)+x_2(i,j), & x_1(i,j)\neq x_2(i,j)\text{时} \end{cases} \tag{5}$$

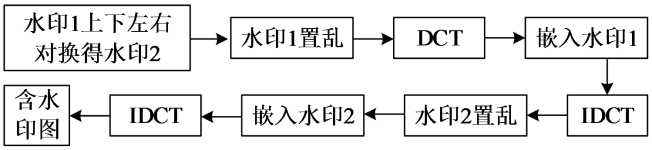


Figure 5. Flowchart with two watermarks embedded
图 5. 两次水印嵌入流程图

Figure 6. Reconstruction of watermark after clipping attack
图 6. 剪切攻击后水印的重构

4. 算法的统计与分析

在通过实验验证算法的过程中，采用 512 × 512 的灰度图像 Lena 作为载体图像，并分为 8 × 8 的子块，共 4096 块；采用图 6 中所示的 64 × 64 的二值图像 BIGC 作为水印图像，共 64 块，实验平台为 MatlabR2017a。

4.1. 透明性测试

透明性测试可以反应水印信息的不可见程度。图 7 所示共有 4 组图片，我们设置嵌入强度为 30 时，分别采用 DCT 算法进行水印嵌入，其中，(a) (b) (c) (d)为原始图像，(e) (f) (g) (h)为含水印图像。可以观察到，人眼几乎看不出差别。另外，我们还统计了该算法在不同嵌入强度下的 PSNR 值，选择的 4 个嵌入水印强度分别是 10、20、50、100，从图 8 可看出，在不加任何攻击，嵌入水印相同的情况下，随着嵌入强度的增加，PSNR 逐步降低。

当水印嵌入强度为 20 时，对嵌入水印后的图像与原始图像进行峰值信噪比 PSNR 的计算，发现四幅图像的 PSNR 值为 40.6275，40.3765，39.8056，38.2305。而当两幅图像的 PSNR 值大于 25 时，人眼就不能很明显的识别出两幅图像的差别。因此可以说明，本水印透明性比较好。

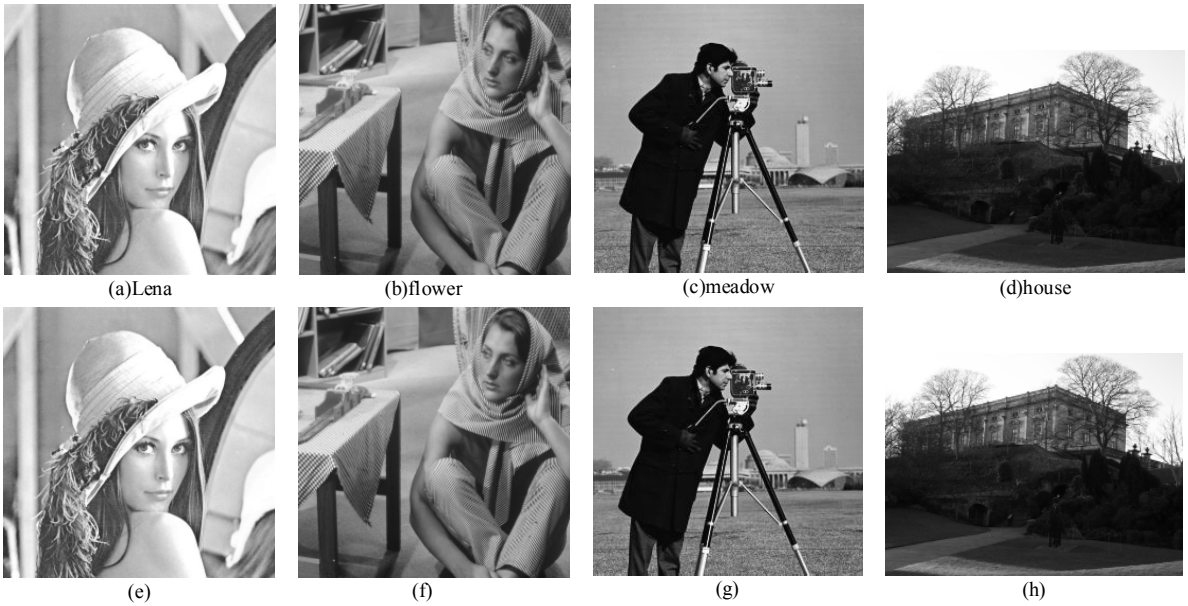


Figure 7. Image comparison before and after embedding watermark
图 7. 嵌入水印图像前后对比

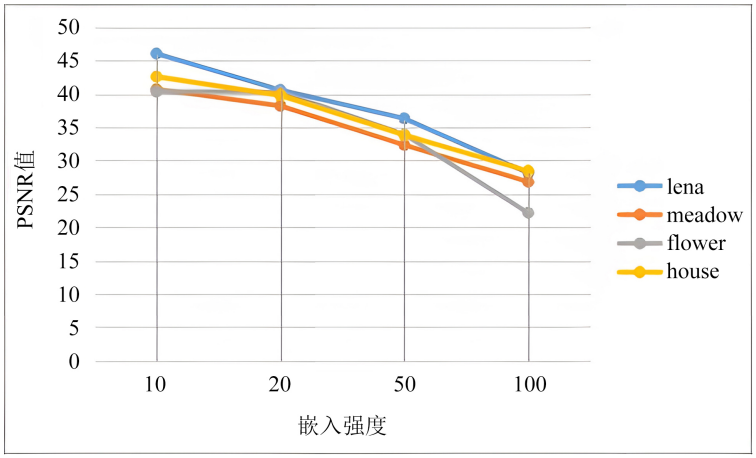


Figure 8. Statistical line plots of PSNR values at different embedding intensities
图 8. 不同嵌入强度下 PSNR 值的统计折线图

4.2. 鲁棒性测试

为了说明算法的鲁棒性，我们在嵌入水印强度为 10、20、50、100、200 的情况下，针对嵌入水印后的 Lena 图像分别进行剪切攻击、高斯噪声攻击、均值滤波攻击，然后再提取攻击后的水印图像，并计算与原始水印信息的误码率 BER，结果见表 1。可以看出在均值滤波攻击下，误码率逐渐降低，一直低于 0.5；在高斯噪声攻击下，误码率先缓慢增加，自嵌入强度为 100 后开始骤降，误码率低于 0.5；在剪切攻击下，误码率缓慢降低最后保持不变。由此可见，该算法在经过二次嵌水印后，鲁棒性较强。

在嵌入水印强度为 20 时，如图 9~11 所示，我们分别展示了剪切、高斯、均值滤波攻击方式下提取的各个水印，其中子图(a)表示攻击图，子图(b)表示提取的含密水印 1，子图(c)表示提取的含密水印 2，子图(d)表示解密的重构水印。从图中可以看出，该水印的鲁棒性较好，在各种攻击下均能较大程度的还原水印信息。

Table 1. BER values for various attacks
表 1. 各种攻击下的误码率数值

嵌入强度	剪切攻击 BER	高斯攻击 BER	均值滤波攻击 BER
10	0.1135	0.4944	0.4807
20	0.1121	0.4973	0.4539
50	0.1106	0.5049	0.2612
100	0.1104	0.5044	0.0723
200	0.1104	0.2422	0.0231

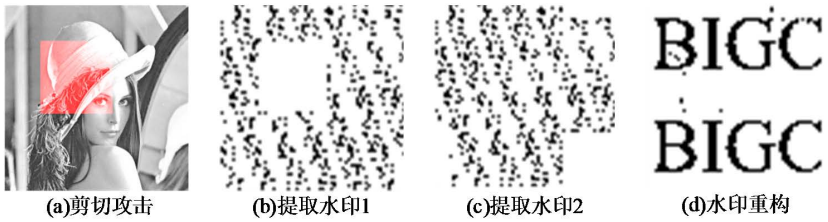
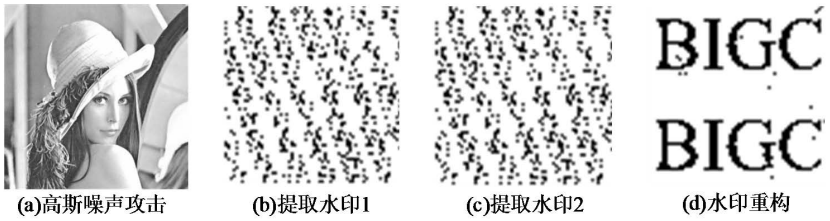
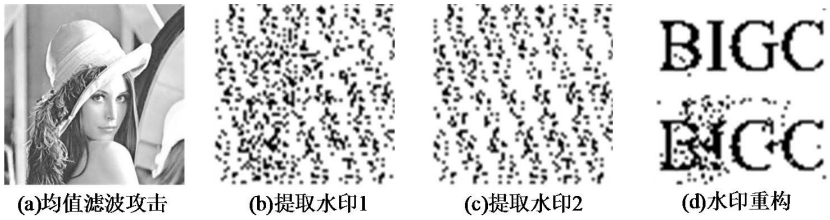


Figure 9. Extract watermark under shear attack
图 9. 剪切攻击提取水印



(a)高斯噪声攻击 (b)提取水印1 (c)提取水印2 (d)水印重构

Figure 10. Extract watermark under Gaussian noise attack
图 10. 高斯噪声攻击提取水印



(a)均值滤波攻击 (b)提取水印1 (c)提取水印2 (d)水印重构

Figure 11. Extract watermark under mean filtering attack
图 11. 均值滤波攻击提取水印

5. 结论

本文在基于 DCT 水印嵌入算法的基础上，采用二次水印嵌入和 Arnold 水印置乱算法，实现了水印的抗剪切攻击特性，实现的核心关键就在于二次水印的变换过程以及盲水印的嵌入过程。实验结果表明，只要剪切面积不超过图像大小的 1/2，该算法对剪切攻击就有一定的抵抗力。而对于常见的高斯噪声攻击、均值滤波攻击等鲁棒性还是很强的，并且透明性也能得到很好的保证。

参考文献

[1] Tirkel, A.Z. and Rankin, G.A. (1993) Electronic Watermark. *Digital Image Computing*, *Technology and Applications*, **3**, 666-673.

[2] Cox, I.J., Kilian, J., Leighton, T., *et al*. (1996) A Secure, Robust Watermark for Multimedia. In: Anderson, R., Ed., *Information Hiding*. *IH* 1996. *Lecture Notes in Computer Science*, Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-61996-8_41

[3] Zhao, Y.L., Zheng, X.S., Li, N., Liu, G.Q. and Wang, Q.X. (2006) A Digital Image Watermark Algorithm Based on DC Coefficients Quantization. *The 6th World Congress on Intelligent Control and Automation*, **2**, 9734-9738.

[4] 刘志军. 基于 DCT 域多通道彩色图像盲检水印算法[J]. 山东大学学报(工学版), 2011, 41(3): 31-35.

[5] 何美娟. DCT 域图像数字水印嵌入技术的研究[D]: [硕士学位论文]. 北京: 北京邮电大学, 2009.

[6] 田敏. 基于数字水印的多媒体信息版权保护技术[D]: [硕士学位论文]. 济南: 山东大学, 2014.

[7] 苏娜. 基于 DCT 域的数字图像水印算法的研究与应用[D]: [硕士学位论文]. 成都: 电子科技大学, 2016.

[8] 杜江, 裴珂, 谢维信. 一种新的图像数字水印产生方法[J]. 西安电子科技大学学报, 2000, 27(4): 491-495.

[9] 李天举. DCT 域图像盲水印算法研究及 FPGA 实现[D]: [硕士学位论文]. 西安: 西安科技大学, 2013.

[10] 白伟. DCT 水印算法在抗剪切攻击方面的研究[J]. 太原师范学院学报(自然科学版), 2014, 13(3): 63-66.