

图像内容可恢复的半脆弱水印技术研究

刘东彦¹⁾ 刘文波¹⁾ 张弓²⁾

¹⁾ (南京航空航天大学自动化学院, 南京 210016) ²⁾ (南京航空航天大学信息科学与技术学院, 南京 210016)

摘要 提出了一种图像内容可恢复的半脆弱水印技术, 采用自适应水印对图像进行认证、定位篡改, 再利用零水印技术从第三方处提取篡改处的恢复水印, 重构被篡改作品。认证方案采用基于非抽样 Contourlet 域奇异值分解的自适应水印认证方法, 该方法具有更高的篡改定位精度, 能有效区分有意攻击与无意攻击。恢复水印是通过将原始图像分块、置乱后, 分别进行 4×4 编码, 再对码流加密后交给第三方获得的。实验结果表明, 本算法能够在容许一定压缩比的情况下检测发生的篡改, 并有效地恢复被篡改的区域。

关键词 半脆弱水印 篡改定位 图像恢复 非抽样 Contourlet 变换 混沌

中图分类号: TP309.7 文献标志码: A 文章编号: 1006-8961(2009)01-0020-06

Research of Recoverable Semi-fragile Watermarking for Image Content

LIU Dong-van¹⁾, LIU Wen-bo¹⁾, ZHANG Gong²⁾

¹⁾ (College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016)

²⁾ (College of Communication Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016)

Abstract In this paper, a new semi-fragile watermarking scheme based on image content for tampering localization and recovery is proposed. The authentication watermark is embedded in the SVs (singular values) with an adaptive quantized method within the domain of nonsubsampled contourlet transform (NSCT). This method can improve the tampering localization capability without causing noticeable artifacts and can distinguish involuntary modification from vicious attacks. The recovery watermarking is achieved by dividing the original image, scrambling, encoding and encrypting at last handing on the third party in which zero-watermarking technique is used. The results of the experiment show that the proposed scheme can detect and localize the tamper as well as can recover the tamper image when it undergoes JPEG compression.

Keywords semi-fragile watermarking, tamper localization, image recovery, nonsubsampled contourlet transform, chaos

0 引言

数字水印 (digital watermarking) 技术是将一些标识信息 (即数字水印) 直接嵌入数字载体当中, 但不影响原载体的使用价值, 也不容易被人的知觉系统 (如视觉或听觉系统) 觉察或注意到。通过这些隐藏在载体中的信息, 可以达到确认内容创建者、购买者、传送隐秘信息或判断载体是否被篡改等目的。

数字水印按功能可分为认证水印和版权保护水印, 前者根据对篡改的识别能力又可分为完全脆弱水印和半脆弱水印。完全脆弱性水印, 能够检测出任何对图像像素值的改变或对图像完整性的破坏。在许多实际应用中, 往往需要水印能够抵抗一定程度的有益的数字信号处理操作, 有效区分有意攻击 (如剪切、擦除) 与无意攻击 (如 JPEG 压缩)。故半脆弱水印技术被广泛研究和关注。其中文献 [1] 提出了通过修改小波高频系数进行水印的嵌入。然而

基金项目: 航空科学基金项目 (04D52032); 江苏省自然科学基金项目 (K2001047)

收稿日期: 2008-09-10 改回日期: 2008-10-22

第一作者简介: 刘东彦 (1983—), 女。南京航空航天大学测试计量技术及仪器专业, 硕士研究生。主要研究领域为数字水印、信号处理。E-mail: dy030230305@163.com

水印直接嵌入到小波高频系数中, 虽然可以保护原始图像的可视性, 但却极易受到压缩等攻击。因此文献 [2] 采用对图像小波分解低频系数进行奇异值分解的方法来嵌入水印, 但由于尺度变换的影响, 该算法使图像的定位精度降低。本文提出采用非抽样 Contourlet 变换代替小波变换, 将子带图像分块后进行奇异值分解 (SVD), 通过自适应量化修改奇异值来实现水印嵌入。本文的算法中由于整个分解过程没有任何抽样环节, 各尺度下子带均与原始图像具有相同的尺寸, 不但提高了篡改定位的精度, 同时能有效区分有意攻击与无意攻击。

当由认证过程知道作品是否被改变及何处被改变后, 人们自然会想到如何重构被篡改的作品, 比较典型的重构算法有 Fridrich 等人提出的自嵌入算法思想^[3], 将低清晰度的压缩图像嵌入到自身的 LSB (least significant bit) 最不重要像素位, 但其鲁棒性较差。文献 [4] 则基于盲重构思想, 首先判断作品经历了何种篡改, 再对这些篡改进行反变换恢复出原始作品, 该算法只适合处理可进行反变换的失真。文献 [5] 提出了一种基于 4×4 分块的图像编码算法, 通过映射函数将码流嵌入到变换域系数中, 实现了专门针对感兴趣区域的篡改恢复。文献 [6] 利用第三方及“零水印”技术得到了 DWT (discrete wavelet transform) 域抗几何攻击的鲁棒水印算法, 文献 [7] 中也利用此方法提出了一种实现多个隐形签名的数字水印算法。本文的恢复水印并不是直接嵌入图像中, 而是采用“零水印”技术的原理, 将其交给第三方, 这样不会因为恢复水印的嵌入影响认证精度及图像质量。同时图像经历压缩以后, 恢复水印并没有被破坏, 故能够对经历过压缩的图像上的恶意篡改进行恢复。

1 非抽样 Contourlet 变换及图像置乱

1.1 非抽样 Contourlet 变换

小波变换凭借对信号良好的时频局部分析, 在信号处理中得到了广泛的应用。然而, 小波变换的局限性也受到了广泛的关注, 其小波基只具有有限个方向。Contourlet 变换 (contourlet transform, CT) 作为一种多尺度几何分析工具极大地解决了小波存在的问题^[8], 在很多领域 (如图像的消噪和融合, 数字水印^[9]) 得到了很好的应用。但 CT 中的抽样操作会产生频谱混叠, 使变换不具有平移不变性。非

抽样 Contourlet 变换是由 Do 等人在 2005 年提出的, 由非采样塔式结构 (nonsampled pyramid, NSP) 实现多尺度分析, 再通过非抽样方向滤波器组 (nonsampled directional filter bank, NSDFB) 实现多方向性分析^[10]。图 1(a) 为非抽样 Contourlet 变换 (nonsampled contourlet transform, NSCT) 的流程图, 首先由非抽样塔式结构将图像分解为一个高频子带, 一个低频子带, 接着用非抽样方向滤波器组将高频子带分解为几个方向子带, 类似的分解步骤可以在非抽样塔式结构输出的低频子带上反复进行从而实现信号的多尺度、多方向分析。图 1(b) 为三尺度, 方向数分别为 8, 4, 4 对应的 NSCT 分解频谱图 w_1, w_2 表示不同的频率方向。由于整个分解过程中没有任何抽样环节, 因此, 各尺度下各方向子带均与原输入图像具有相同的尺寸。众所周知冗余变换对添加噪声稳健, 而 NSCT 较小波变换具有更多的冗余, 水印又可以看作是在强背景下添加弱噪声, 故 NSCT 较小波变换能够增加水印的嵌入量而又不会造成图像视觉质量的明显下降, 并且水印嵌入量的增加提高了篡改定位的精度。

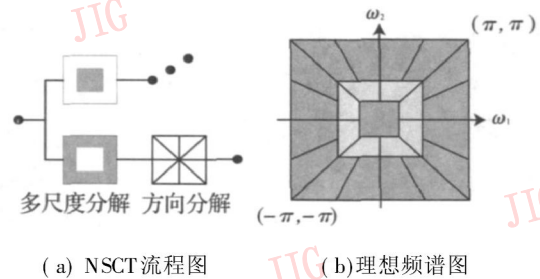


图 1 非抽样 Contourlet 变换

Fig 1 Nonsampled contourlet transform

1.2 Tent 映射及其 2 维置换网络

在恢复水印的生成过程中, 利用 1 维 Tent 映射^[11]产生 2 维置换网络, 对图像块进行置乱。Tent 映射如式 (1) 所示:

$$x_{k+1} = \begin{cases} x(k)/a & x(k) \leq a \\ (1-x(k))/1-a & x(k) > 1-a \end{cases} \quad (1)$$

$$k = 1, 2, 3, \dots$$

当参数 $a = 0.5$ 时, 式 (1) 为典型的混沌系统。基于混沌的初值敏感性及其遍历性, 利用混沌动力系统 (式 (1)) 产生的伪随机序列作为图像置乱序列, 可以有效地增强嵌入水印的安全性^[12]。

2 维置换网络的形成过程如下: 任选一初始值, 按式 (1) 迭代产生混沌序列 x_k , 随机抽取 M 个序

列, 经处理后使其值为在 $[1, M]$ 之间的整数, 组成新的序列 x_m ; 同样再从序列 x_k 中随机抽取 N 个值, 处理后组成新的整数序列 x_n 。由 x_m, x_n 形成一个 2 维置换网络, 经 2 维网络置换后, 坐标 (i, j) 的像素被置到 $(x_m(i), x_n(j))$ 的位置。表 1 所示即横坐标 X 分别为 25, 26, 27, 28, 纵坐标 Y 分别为 38, 39, 40 时, 经置乱后所对应的新坐标。如原始坐标为 (26, 38) 的图像块经置乱后的新坐标为 (55, 1)。

表 1 置乱后部分坐标的对应关系

Tab 1 The relationship of partial coordinates after scrambled

	25	26	27	28
38	(39, 1)	(55, 1)	(35, 1)	(24, 1)
39	(39, 38)	(55, 38)	(35, 38)	(24, 38)
40	(39, 53)	(55, 53)	(35, 53)	(24, 53)

2 算法原理

2.1 认证水印的嵌入

认证水印的嵌入算法流程图如图 2 所示, 其中非抽样 Contourlet 变换 (NSCT) 的 NSP 采用 “maxflat” 塔式滤波器组, NSDFB 采用 “dmxflat7” 方向滤波器组具有最佳的视觉效果, 两者的基本组成单元都是满足完全重构条件 $H_0(z)G_0(Z) + H_1(z)G_1(z) = 1$ 的双通道非抽样滤波器组^[10]。“maxflat” 由 2 维分解、重构低通滤波器, 2 维分解、重构高通滤波器组成; “dmxflat7” 由分解扇形滤波器、重构扇形滤波器组成。

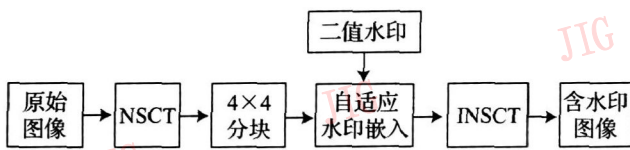


图 2 认证水印嵌入流程图

Fig 2 The embedding flow chart of authentication watermarking

水印嵌入采用自适应量化修改奇异值的方法, 分为 3 步。

1) 自适应量化: 量化步长的计算公式如下^[4]:

$$Q_q = Q_{min} + (Q_{max} - Q_{min}) \times \frac{(C_q - C_{min})}{(C_{max} - C_{min})} \quad (2)$$

$$C_q = m_q + 3\sigma_q, \quad q = 1, 2, \dots, \frac{m}{4} \times \frac{n}{4}$$

Q_{min}, Q_{max} 分别为量化步长的最小值和最大值,

可以由实验者设定; m_q, σ_q 分别表示第 q 个 4×4 图像块的平均值和标准偏差; C_{min}, C_{max} 分别为 $\{C_q\}$ 的最小值和最大值。式 (2) 中的 Q_q 基于 NSCT 域中图像块的统计意义, 能够自适应于每一块不同的图像;

2) 水印嵌入: w 为二值水印图像中的一位, 如果 $w = 1, D$ 是奇数, $D = D + 1$; D 是偶数时, D 不变。如果 $w = 0, D$ 是偶数, $D = D + 1$; D 是奇数时, D 不变。 $D = \lfloor N_v / Q_q \rfloor, N_v = \|\mathbf{v}\|$; \mathbf{v} 是块 E_q 的奇异值所组成的向量, $\|\cdot\|$ 表示取范数, E_q 表示第 q 个 4×4 的图像块;

3) 修改奇异值: 利用

$$\mathbf{v}^* = \mathbf{v} \times (Q_q \times D + Q_q / 2) / N_v \quad (3)$$

对奇异值进行修改, 并进行奇异值反变换得到含水印的块。奇异值代表的是图像固有的几何性质, 具有很好的稳定性, 对它进行小的修改不会影响图像的视觉质量, 从而能够将水印嵌入到奇异值中来传递信息。

2.2 恢复水印的生成

恢复水印生成流程图如图 3 所示, 首先对原始图像分块, 利用 2 维置换网络对分块后的图像进行置乱, 分别对每块进行 4×4 编码, 同时取图像稳定信息与 Logistic 混沌产生的码流异或, 再与先前得到的编码码流异或加密, 交给第三方形成恢复水印。

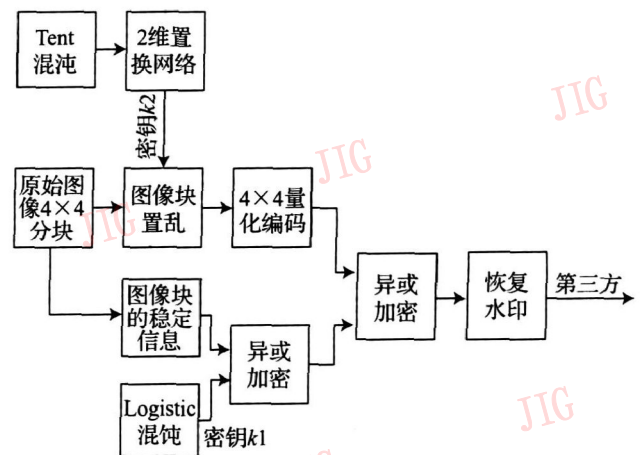


图 3 恢复水印生成流程图

Fig 3 The forming process of recovery watermarking

NSCT 域认证水印的定位精度为 4×4 故量化编码采用文献 [7] 中提出的 4×4 编码方法, 量化系数矩阵 O 及编码位数矩阵 L 如下:

$$O = O_0 \times 1.5 = \begin{bmatrix} 16 & 11 & 10 & 16 \\ 12 & 12 & 14 & 19 \\ 14 & 13 & 16 & 24 \\ 14 & 17 & 22 & 29 \end{bmatrix} \times 1.5$$

$$L = \begin{pmatrix} 7 & 6 & 5 & 0 \\ 6 & 5 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

式中, O_0 为基本量化表。每一图像块的稳定信息取最高有效位 (MSB) 及次 MSB 共 32 比特二进制码流。水印的生成可表示为

$$W = (A \oplus B(k1)) \oplus C(k2) \quad (4)$$

式中, \oplus 是指进行异或操作, W 为恢复水印, A 为图像的稳定信息, $B(k1)$ 表示以密钥 $k1$ 产生的混沌码流, $C(k2)$ 表示以密钥 $k2$ 置乱后的图像块编码码流。置乱的目的, 其一, 能够增加算法的安全性; 其二, 本文的恢复水印生成与图像块的稳定信息有关, 考虑到故意篡改一般集中在图像的某个区域, 若该区域图像的稳定信息也被篡改后, 利用提取密钥找到恢复水印所在位置, 解码重构依然能够正确恢复被篡改作品。

2.3 篡改的检测与恢复

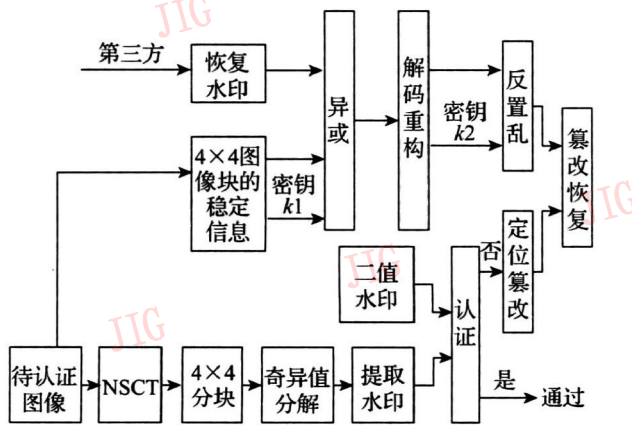


图 4 篡改认证及恢复流程图

Fig 4 The tamper authentication and recovery flow chart

按照水印嵌入的过程对含水印图像进行 NSCT 分解, 分块计算得到 D 值, 如果 D 为奇数则 $w = 1$, 否则 $w = 0$ 。为了估计篡改的程度, 采用了下面的篡改评估函数:

$$BER = \frac{1}{M_1 \times M_2} \sum_k \sum_l w_{k,l} \oplus \hat{w}_{k,l} \quad (5)$$

$w_{k,l}$ 是二值水印, $\hat{w}_{k,l}$ 是提取的水印, $M_1 \times M_2$ 是水印的大小, BER 的值在 $[0, 1]$ 之间。

水印使用者可以根据应用场合的不同设定不同的域值 τ 当 $BER \geq \tau (0 \leq \tau \leq 1)$ 时, 认为图像被篡改过, 在安全性要求较高的场合可以把 τ 值设小。

若判定图像被篡改, 则通过提取的水印定位被篡改的位置, 再利用提取密钥向第三方申请获得恢复水印, 以密钥 $k2$ 找到被篡改块的恢复信息所在的位置, 以密钥 $k1$ 恢复编码码流, 解码重构恢复被篡改的图像。如下:

$$C(k2) = W \oplus (A \oplus B(k2)) \quad (6)$$

3 实验结果分析

3.1 抗 JPEG 压缩实验

本算法能够有效地抵抗 JPEG 压缩, 但对恶意篡改敏感。本文的实验采用了多幅 256×256 的图像, 在 NSCT 算法中嵌入 64×64 的水印量, DWT 算法中嵌入 32×32 的水印量。表 2 中列出了标准图像 “Lena”, “peppers”, “woman” 在不同的量化步长及相应的 JPEG 压缩比下, 两种算法分别对应的峰值信噪比 (PSNR) 与误码率 (BER) 的值, 其中 PSNR 用于描述图像视觉质量。

表 2 NSCT 域算法与 DWT 域算法抗 JPEG 压缩性能比较

Tab 2 The contrast between NSCT scheme and DWT scheme against JPEG

图像	PSNR		不同 JPEG 压缩比下的误码率 BER 的值 %								参数	
	NSCT	DWT	90%		70%		50%		30%		Q_{min}	Q_{max}
			NSCT	DWT	NSCT	DWT	NSCT	DWT	NSCT	DWT		
Lena	39.02	40.77	0	0	0.09	0.68	1.51	3.32	10.94	19.04	6	31
Peppers	38.06	39.74	0	0	0.13	0.74	1.93	4.01	12.39	19.65	9	45
Woman	41.42	42.03	0	0	0.08	0.59	1.32	3.24	10.20	18.79	12	40

从表 2 可以看出, 虽然 NSCT 算法嵌入的水印量是 DWT 算法中的 4 倍, 但是含水印图像的视觉质量 (PSNR 值对比) 并没有明显下降, 且其抗 JPEG 压

缩的能力较 DWT 算法有大幅度提高 (BER 值对比)。NSCT 域算法嵌入水印后 Lena 图像的 $PSNR = 39.02$ 当 JPEG 压缩比为 50% 时, 提取认证水印的

BER 值为 1.51%，说明此算法形成的水印图像不但拥有很高的视觉质量，而且具有很好的抗 JPEG 压缩的鲁棒性。而此时 DWT 域算法中 PSVR = 40.77，JPEG 压缩比为 50% 时，提取认证水印的 BER 值为 3.32%，在视觉质量提高不大的情况下抗压缩性能大幅下降。图 5 分别为 NSCT 域中 Lena 含水图像在压缩比为 90%，70%，50%，30% 下提取到的水印图像，其中 BER 的值分别为 0.009%，1.51%，10.94%。



图 5 不同 JPEG 压缩下提取的水印图像
Fig 5 The extracted watermark after different JPEG compression ratios

3.2 检测篡改及定位精度分析

文献 [2] 中将 256×256 的图像单尺度小波变换后对低频子图进行分块，分块的大小为 4×4 嵌入的水印量为 $32 \times 32 = 1024$ bits，然而由于小波域的尺度变换，使得实际的定位精度为 8×8 。本算法

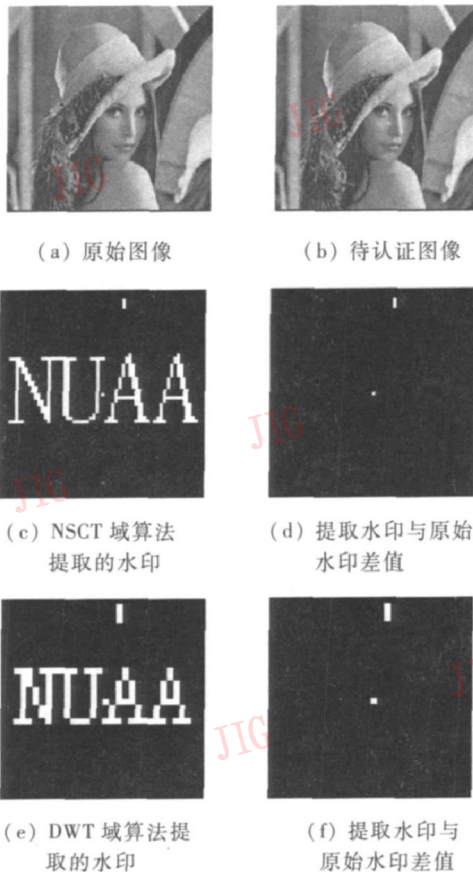


图 6 篡改定位精度分析

Fig 6 The analysis of the ability for location after modified

中，单尺度 NSCT 变换后输出的子带图像与原始图像具有同样的大小，实际的定位精度为 4×4 水印嵌入量为 $64 \times 64 = 4096$ bits。下面以 Lena 图像为例进行说明。

实验中采用 photoshop 的铅笔工具对含水图像的 Lena 图像 (图 6(a)) 进行修改形成待认证图像 (图 6(b))。然后分别采用本文的方法和文献 [4] 中的方法检测篡改并对篡改进行定位，结果分别如图 6(c) 和图 6(e) 所示。为了对比定位的精度，将提取的水印图 6(c) 和图 6(e) 与原始水印相减再取绝对值得图 6(d) 和图 6(f)，可以明显地看出本文的算法能够有效地检测微小篡改并将篡改锁定在更精确的范围内，具有较高的定位精度。

3.3 篡改恢复

本算法利用了第三方的优势，恢复水印不会由于数据的压缩而去除或者被破坏，在经历压缩的图像上进行剪切篡改，本算法依然能够检测出来篡改并能够对篡改区域进行恢复。图 7(a) 为原始图像，图 7(b) 当 JPEG 压缩比为 80% 的含水图像，此时的 PSVR = 33.71，BER 近似等于零，图 7(c) 是在图 7(b) 的基础上剪切后的图像，Lena 图像的面部信息已无完全被去除，图 7(d) 是由本算法恢复出的近似图像，其 PSVR = 26.36 能够基本满足对图像内容的识别要求。



图 7 篡改恢复能力分析

Fig 7 The analysis of the ability for rehab after modified

4 结论

本文提出的可恢复的半脆弱数字水印技术，能

够在图像经历一定压缩比的情况下检测并定位发生的篡改, 同时有效地恢复被篡改的区域。认证水印利用非抽样 Contourlet 变换冗余及整个分解过程无抽样的特性, 增加了水印嵌入量, 提高了篡改定位的精度。嵌入阶段采用自适应量化修改奇异值的方法提高了算法抗 JPEG 压缩的能力。通过将恢复水印交给第三方而不是直接嵌入图像中, 减小了对认证过程的影响及保证了水印图像的质量。在加密过程中充分利用了混沌系统的特性, 增加了算法的安全性。

参考文献 (References)

- [1] Xie L, Arce G R. Joint wavelet compression and authentication watermarking [C] // Proceedings of the IEEE International Conference on Image Processing Chicago, IL, USA: IEEE Press, 1998: 427-431.
- [2] Bao P, Ma Xiao-hu. Image adaptive watermarking using wavelet domain singular value decomposition [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2005, 15(1): 96-102.
- [3] Fridrich J, Goljan M. Image with self-correcting capabilities [C] // Proceedings of the IEEE International Conference on Image Processing Kobe, Japan: IEEE Press, 1999: 792-796.
- [4] Kundur D, Hatzinakos D. Blind image deconvolution [J]. IEEE Signal Processing Magazine, 1996, 13(3): 43-64.
- [5] Ren Juan, Wang Yun-hong, Tan Tien-ju. A self-recovery algorithm based on region of interest [J]. Journal of Automation Engineering, 2004, 30(6): 833-843. [任娟, 王蕴红, 谭铁牛. 基于感兴趣区域的图像认证与自恢复算法 [J]. 自动化学报, 2004, 30(6): 833-843.]
- [6] Li Jing-bing, Huang Xiyue. A DWT-based watermarking algorithm robust to geometric attacks [J]. Computer Simulation, 2007, 24(3): 303-306. [李京兵, 黄席樾. 一种基于 DWT 抗几何攻击数字水印鲁棒算法 [J]. 计算机仿真, 2007, 24(3): 303-306.]
- [7] Li Jing-bing, Huang Xiyue, Zhou Yaxun. An algorithm of multiple invisible signatures [J]. Application Research of Computer, 2005, 22(8): 252-254. [李京兵, 黄席樾, 周亚训. 一种实现多个隐形签名的数字水印算法研究 [J]. 计算机应用研究, 2005, 22(8): 252-254.]
- [8] Do M N, Vetterli M. The contourlet transform: An efficient directional multiresolution image representation [J]. IEEE Transactions on Image Processing, 2005, 14(6): 760-769.
- [9] Jayalakshmi M, Merchant S N, Desai Uday B. Significant pixel watermarking in contourlet domain [C] // Proceedings of the IET International Conference on Visual Information Engineering Vienna, Austria: IEEE Press, 2006: 416-421.
- [10] Do M N, Zhou Jian-ping. The nonsubsampled contourlet transform: theory, design and applications [J]. IEEE Transactions on Image Processing, 2005, 14(1): 469-472.
- [11] Zhao Dawei, Chen Guang-rong, Liu Wen-bo. A chaos-based robust wavelet-domain watermarking algorithm [J]. Chaos, Solitons & Fractals, 2004, 22(1): 47-55.
- [12] Peng Zhen-ni, Liu Wen-bo. Color image authentication based on spatiotemporal chaos and SVD [J]. Chaos, Solitons & Fractals, 2008, 36(4): 946-952.