基于SVD和神经网络的鲁棒水印算法 *)

吕英华[1,2] 王 巍[1] 孔 俊[1] 侯 刚[1,3] 韩佳伶[1] 周颜军[1]
(东北师范大学计算机学院计算机系 长春 130024)[1]
(吉林大学计算机学院 长春 130012)[2] (东北师范大学人文学院 长春 130117)[3]

摘 要 本文提出了一种基于奇异值分解和小波包分解相结合的全新水印算法。综合利用奇异值和方差的特征来对宿主图像进行预处理之后,提取出两个具有不同掩蔽效应的子图,分别对子图的奇异值和小波包系数使用抖动调制方法嵌入不同强度的水印。本文算法不仅实现了水印的盲提取,而且实现了水印鲁棒性与透明性的最佳折中。本文的另一创新之处是提出了一种基于神经网络的新的水印复原方法。实验结果表明本文方法较传统方法而言有更高的灵活性和鲁棒性,尤其具有很强的抗JPEG压缩能力。

关键词 小波包,奇异值,抖动调制,主成分分析

A Robust Watermarking Algorithm Based on SVD and Neural Network

LU Ying-Hua[1,2] WANG Wei[1] KONG Jun[1] HOU Gang[1,3] HAN Jia-Ling[1] ZHOU Yan-Jun[1]
(Computer School, Northeast Normal University, Changchun 130024)[1] (Computer School, Jilin University, Changchun 130012)[2]
(College of Humanities and Science, Northeast Normal University, Changchun 130117)[3]

Abstract In this paper, we present a novel digital image watermarking algorithm based on singular value decomposition and wavelet packet decomposition. A new method which successfully combines the characteristics of singular values and variance is developed to preprocess the host image to obtain two sub-images with different masking effect. The application of dither modulation embedding scheme makes our method achieve the goal of watermark blind-extraction. Another significant innovation of our algorithm is that we propose a different means of watermark recovery based on neural network, which contributes to accurate data tracing. Simulation results show that our method is more flexible than traditional methods and fulfills a good compromise between the robustness and transparency.

Keywords Wavelet packets, Singular value, Dither modulation, Principal component analysis

1 引言

作为一种解决数字产品版权问题的有效手段,数字水印技术近年来受到人们的广泛关注,并得到充分发展。数字水印技术就是将数字、序列号、文字、图像标志等版权信息嵌入到多媒体数据中的一种技术,用来跟踪多媒体数据的发布与使用。数字水印的特殊应用要求嵌入的水印信息不仅具有透明性,而且具有强鲁棒性,能够抵抗各种类型的攻击操作。

根据水印嵌入的域不同,水印嵌入技术可分为两类:空域水印技术[1]及频率域水印技术[2]。虽然空域水印技术算法简单且计算量少,但是由于其对攻击的敏感性,越来越多的人开始关注频率域水印算法。离散傅立叶变换(DFT)、离散余弦变换(DCT)、离散小波变换(DWT)等都经常应用于频率域水印算法中来保证嵌入水印信息的鲁棒性和透明性。奇异值分解(SVD)是一种强大有效的数值分析技术而且广泛应用于各种技术领域中。

如果A表示尺寸为$m\times n$的灰度图像,那么每个实数矩阵A可以被SVD分解为如式(1)所示的三个矩阵的形式,其中U和V是大小分别为$m\times m$和$n\times n$的正交矩阵,S是和A大小一样的对角矩阵,而且S矩阵中的奇异值满足$\lambda_{1,1}\geqslant\lambda_{2,2}\geqslant\cdots\geqslant\lambda_{r,v}>\lambda_{r+1,r+1}=\cdots=\lambda_{t,t}=0$,其中$r$是矩阵$A$的秩($r\leqslant t, t=\min(m,n)$)。

随着SVD技术在数字水印领域中的应用,最近出现了一些基于SVD的数字水印算法[3~5]。Zhang Zhi-Ming等人在文[4]中提出一个新的基于SVD的水印算法,利用Turbo代码来提高系统鲁棒性。在文[5]中,利用DWT将宿主图像分解成四个子带以后,作者利用修改奇异值的方法来嵌入水印。本文中,我们充分利用奇异值的特点,首先对宿主图像预处理,然后结合小波包分解在不影响视觉效果的条件下实现水印嵌入操作。算法的框图如图1所示。

$$A=USV^T=\begin{bmatrix}u_{1,1} & \cdots & u_{1,m}\\ u_{2,2} & \cdots & u_{2,m}\\ & \vdots & \\ & \vdots & \\ u_{n,1} & \cdots & u_{m,m}\end{bmatrix}\begin{bmatrix}\lambda_{1,1} & 0 & \cdots & \cdots & 0\\ 0 & \lambda_{2,2} & \cdots & \cdots & 0\\ & & \lambda_{r,r} & \cdots & \\ & & \cdots & \lambda_{t,t} & \\ 0 & 0 & \cdots & \cdots & 0\end{bmatrix}\begin{bmatrix}v_{1,1} & \cdots & v_{1,n}\\ v_{2,1} & \cdots & v_{2,n}\\ & \vdots & \\ & \vdots & \\ v_{n,1} & \cdots & v_{n,n}\end{bmatrix}=\sum_{i=1}^{r}\lambda_i u_i v_i^T \quad (1)$$

*)基金项目:本文得到国家教育部重点项目(编号:02090)的资助。**吕英华** 教授,主要研究方向:图像处理,面向并行对象计算机系统的模拟。**王 巍** 硕士研究生,主要研究方向:数字信息安全。**孔 俊** 副教授,主要研究方向:图像处理,模式识别,生物认证。**侯 刚** 硕士研究生,主要研究方向:智能规划,数字水印。**韩佳伶** 硕士研究生,主要研究方向:数字水印。**周颜军** 副教授,主要研究方向:数据挖掘,数字库与信息管理系统。

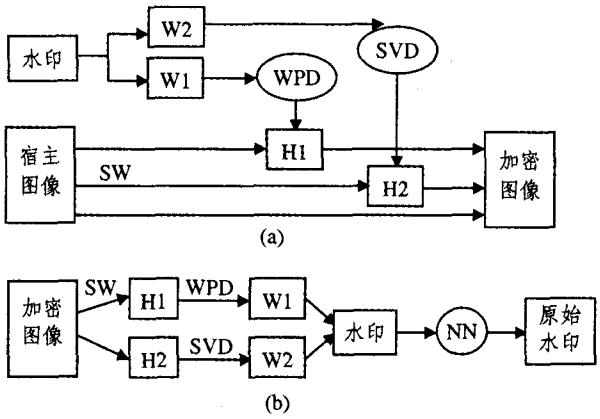


图 1　水印算法框图　(a)水印嵌入过程　(b)水印复原过程

本文第 2 部分详细介绍了宿主图像预处理算法。第 3 部分介绍了水印的嵌入算法和提取算法。水印的不可见性度量及对攻击鲁棒性的实验结果列在了第 4 部分。第 5 部分提出了基于神经网络的水印复原方法。最后为全文的总结。

2　宿主图像预处理

每幅宿主图像都有不同的灰度信息和特征，根据人类视觉系统的特点，人眼对不同亮度和纹理区域的噪声具有不同的敏感度。因此本文根据图像特征将宿主图像分割成不同区域来保证水印嵌入强度的可控性和不可见性。

为了更好地表征图像的区域特征，本文提出了一个结合奇异值(Singular Value)和方差(Variance)特征处理宿主图像的新方法(SVV)。方差衡量的是一个图像区域的灰度对比度和平滑程度，而奇异值描述的是区域的亮度信息[5]。一个区域，如果它的方差值越大，则纹理越复杂；如果奇异值越大，则区域越亮。因此我们充分利用奇异值和方差的特点来分割宿主图像，算法如下：

(1)将宿主图像在空域中进行分块，分块大小为 8×8。

(2)计算所有图像块的方差，并且升序排序。

(3)对每一图像分块进行 SVD 变换，利用式(2)计算 E，即矩阵 S 中最大两个奇异值的比值。

$$[USV]=\mathrm{svd}(A),E=S(1,1)/S(2,2) \tag{2}$$

其中 A 代表图像子块矩阵，svd 代表奇异值分解操作，$S(1,1)$和 $S(2,2)$代表的是对角矩阵 S 中两个最大的奇异值。

(4)画出示意图来表征宿主图像的 E 和方差关系。如图 2 所示，是图像 Airplane (见图 4(a))和 Baboon (见图 5(a))的关系示意图。其中横坐标代表的是按照方差值升序排列的所有图像子块的标识，纵坐标代表的是相应于每一图像块的 E 值。观察示意图，我们可以发现奇异值的比值随着图像块方差的增大而降低，但是其中有噪声点。

(5)观察方差值大的区域，标记出噪声点。这些点所对应的图像子块平均亮度较高，包含边缘但是边缘灰度差很小。这些子块对水印嵌入是敏感的。

(6)通过分析示意图，去除噪声点，选择具有最大方差值的 512 个图像子块构成纹理高活动区域子图 $H1$。

(7)依照同(6)一样的规则，选择接下来的 2048 个图像子块构成纹理中等活动区域子图 $H2$。

经过对宿主图像预处理以后，我们获得两个子图 $H1$ 和 $H2$，每个子图中的图像子块都具有一致的视觉掩蔽效应而且易于控制水印嵌入强度。

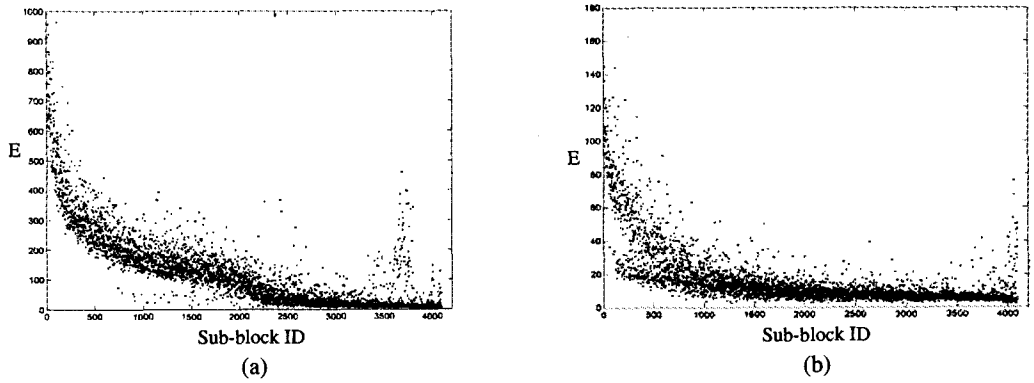


图 2　(a)Airplane 的示意图　(b)Baboon 的示意图

3　水印嵌入算法和提取算法

3.1　水印嵌入算法

本文的水印嵌入算法包括两部分：基于 SVD 的嵌入算法和基于小波包(WP)分解的嵌入算法。首先将原始水印分割成相等的两部分，然后经过上述对宿主图像分块处理以后，我们分别利用这两种嵌入算法将两部分水印嵌入到两幅子图 $H1$ 和 $H2$ 中。

3.1.1　基于 SVD 的嵌入算法

利用 SVD 将水印嵌入到纹理中等活动区域子图 $H2$ 中，不仅因为子图 $H2$ 可以抵抗高通、低通滤波等操作的影响，而且因为奇异值分解得到的最大奇异值对于图像处理过程中的攻击操作是稳定的。算法描述如下：

(1)对纹理中等活动区域子图中的每一 8×8 分块进行 SVD。

(2)选择 S 矩阵中的最大值记为 $S(1,1)$。

(3)根据式(3)应用量化方法将水印信号嵌入到图像子块中。

$$\begin{aligned} &L=fix(S(1,1)/Q)\\ &\text{if}\quad w=1\quad S(1,1)=L\times Q\\ &\text{else}\quad w=0\quad S(1,1)=L\times Q+Q/2 \end{aligned} \tag{3}$$

其中 L 代表 $S(1,1)/Q$ 结果的整数部分，w 代表水印信息，Q 代表量化阶距。

(4)对每一分块进行奇异值分解逆变换得到嵌入水印后的子图。

(5)将嵌入水印后的纹理中等活动区域子图嵌回到宿主图像中。

3.1.2　水印提取算法

对于纹理高活动区域子图 $H1$，我们采用小波包(WP)分解方法嵌入水印。小波变换的多尺度特征[6]以及适应于

JPEG-2000 静态图像压缩标准[7]的特性，使得小波域水印嵌入算法对于压缩等攻击操作具有强鲁棒性。而且，小波包变换的使用大大扩展了水印嵌入空间，使得本文算法将大容量水印信息嵌入到了很小的子图 $H1$ 中。嵌入算法如下：

(1)对纹理高活动区域子图 $H1$ 进行二维小波包分解得到系数阵 $H\_wavelet^{wp}$ 和小波包树 T [8]。

(2)选择 $H\_wavelet^{wp}$ 中的最优系数来嵌入水印信息。

(3)对于小波包系数 WPC_i 采用量化方法嵌入水印，其中 Δ 为量化阶距，l 为整数，WPC_i^w 为嵌入水印后的小波包系数。

若待嵌入的水印 $w_i=1$，则量化公式为

$$WPC_i^w=\begin{cases} l\Delta & l\Delta<WPC_i\leqslant(l+0.5)\Delta \\ (l+1)\Delta' & (l+0.5)\Delta<WPC_i<(l+1)\Delta \end{cases} \quad (4)$$

若待嵌入的水印 $w_i=0$，则量化公式为

$$WPC_i^w=(l+0.5)\Delta, l\Delta<WPC_i<(l+1)\Delta \quad (5)$$

(4) 将嵌入水印的小波包系数写入小波包中。

(5) 根据小波包树 T 和小波包系数重构嵌入水印后的子图。

(6) 将嵌入水印后的纹理高活动区域子图嵌回到宿主图像中。

需要特别提出的是，因为小波包分解的不同子带系数具有不同的掩蔽效应，所以我们选择不同的量化阶距来嵌入水印。量化阶距的选择会在实验结果部分作详细介绍。

3.2 水印提取算法

本文中水印提取算法即为嵌入算法的逆过程。通过传递密钥获得所有加密信息，即可实现水印的盲提取。水印提取算法如下：

(1)根据宿主图像预处理算法，处理宿主图像，获得两个子图 $H1^*$ 和 $H2^*$。

(2)对于纹理中等活动区域子图 $H2^*$，在每一分块上进行 SVD 并得到 $S^*(1,1)$。

(3)根据式(6)计算提取的水印值 w^*。

$$Z=\mathrm{mod}(S^*(1,1),Q)$$

$$\mathrm{if}(Z-Q/2)<Q/4 \quad W^*=0 \ \mathrm{else} \quad W^*=1 \quad (6)$$

其中 Z 代表 $S^*(1,1)$ 和 Q 的模。

(4) 对于纹理高活动区域子图 $H1^*$，进行小波包分解以获得小波包树 T^* 和嵌入水印信息的相应节点的小波包系数。

(5)根据式(6)计算水印值，将式中 $S^*(1,1)$ 替换为小波包系数。

(6)将提取出的两部分水印重构成一个水印整体。

4 实验结果

水印系统的性能可以被如下 4 个特征来衡量：不可见性、容量、鲁棒性、安全性。我们通过实验来说明本文算法的性能。图 3 显示的是尺寸为 64×64 的原始水印和从相应宿主图像中提取出来的水印图像。图 4、图 5 是采用本文算法对尺寸为 512×512、256 级灰度的图像 Airplane 和 Baboon 分别嵌入水印后的结果。以 Airplane 图像为例，实验中首先根据纹理信息从宿主图像中提取出两幅活动区域子图，然后将图 3(a)所示的水印分割成相等的两部分并嵌入到相应子图中。嵌入过程中，纹理中等活动区域子图的量化阶距置为 50。对于纹理高活动区域子图，首先使用“Haar”小波基函数做 2 次小波包分解，然后选择小波包树的节点 5 和 9 嵌入水印，量化阶距分别为 30，20。

本文中我们采用峰值信噪比 $PSNR$ 来评价算法的透明性，定义如下：

$$PSNR=10\times\log_{10}\left(\frac{M\times N\times\max(I^2(i,j))}{\sum_{i=1}^{M}\sum_{j=1}^{M}[I(i,j)-I^*(i,j)]^2}\right) \quad (7)$$

其中 $I(i,j)$ 代表宿主图像像素值，$I^*(i,j)$ 代表嵌入水印后的宿主图像像素值，M 和 N 分别代表宿主图像的高和宽。

采用归一化互相关系数 NC 来评价算法对攻击的鲁棒性，定义如下：

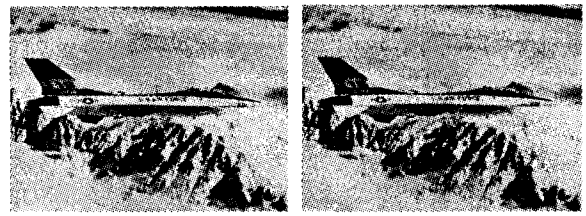
$$NC=\frac{\sum_{i=1}^{H}\sum_{j=1}^{W}WM(i,j)WM^*(i,j)}{\sum_{i=1}^{H}\sum_{j=1}^{W}[WM(i,j)]^2} \quad (8)$$

其中 $WM(i,j)$ 代表水印图像像素值，$WM^*(i,j)$ 代表提取出的水印图像像素值，H 和 W 分别代表水印图像的高和宽。

在不受任何攻击的情况下，Airplane 图像嵌入水印后的 $PSNR$ 值为 42.5283，提取出来的水印 NC 值为 1。Baboon 图像嵌入水印后的 $PSNR$ 值为 42.8793，NC 值为 1。

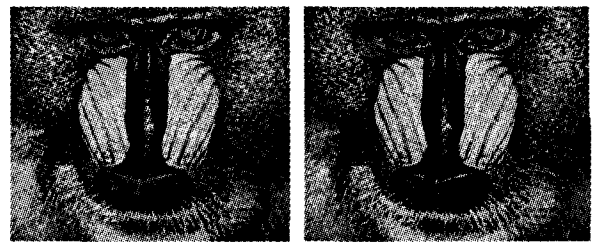
(a) (b) (c) (d)

图 3 (a)原始水印；(b)从图 4(b)中提取出来的水印；(c)(d)从图 4(b)分别经过 80%和 50%JPEG 压缩后提取出来的水印



(a) (b)

图 4 (a)原始宿主图像 (b)嵌入水印后的图像



(a) (b)

图 5 (a)原始宿主图像 (b)嵌入水印后的图像

图 3(c)和(d)显示从受到 80%和 50%JPEG 压缩攻击之后的 Airplane 图像中提取出来的水印图像。如图所示，经过攻击的水印图像质量很好，具有很少的失真。表 1 和表 2 列出了图像 Airplane 和 Baboon 经过各种攻击后的实验结果，其中“SVD”和“WP”代表基于 SVD 和 WP 方法嵌入的水印，“OW”是原始水印的缩写。“QF”代表 JPEG 压缩质量因子，“CR”代表剪切率，“D”代表椒盐噪声密度，“M”和“V”代表加入高斯白噪声的均值和方差。从表 1 和表 2 列出的实验结果可以看出，我们的算法对攻击具有很强的鲁棒性。而且，当 JPEG 压缩质量因子取为 30%时，水印的 NC 值仍然很高。因此本文算法可以完全解决水印隐藏在图像传输过程中的问题。

为了比较算法的有效性，我们仿真了 Sun 的算法[9]，在同样的攻击操作下其算法的实验结果如表 3 所示。和表 1、表 2

比较，本文算法的性能要远远优于 Sun 的方法，尤其在抵抗椒盐噪声和 JPEG 压缩方面。

表 1 图像 Airplane 经过不同攻击后的水印提取结果，用 NC 衡量

攻击类型		SVD	WP	OW
JPEG 压缩	QF=80	1	1	1
JPEG 压缩	QF=50	0.9966	0.9607	0.9744
JPEG 压缩	QF=30	0.8591	0.8489	0.8528
剪切	CR=25%	0.7718	0.7288	0.7452
椒盐噪声	D=0.02	0.6174	0.8427	0.7567
高斯噪声	M=0 V= 0.01	0.9986	0.9966	0.9976

表 2 图像 Baboon 经过不同攻击后的水印提取结果，用 NC 衡量

攻击类型		SVD	WP	OW
JPEG 压缩	QF=80	1	1	1
JPEG 压缩	QF=50	0.9974	0.9966	0.9970
JPEG 压缩	QF=30	0.9191	0.9228	0.9212
剪切	CR=25%	0.7660	0.7351	0.7467
椒盐噪声	D=0.02	0.6957	0.7851	0.7404
高斯噪声	M=0 V=0.01	0.9994	0.9977	0.9985

表 3 Sun 的算法仿真经过不同攻击后的水印提取结果，用 NC 衡量

攻击类型	JPEG	JPEG	JPEG	剪切	噪声	噪声
图像	QF=80	QF=50	QF=30	CR=25%	椒盐	高斯
Airplane	0.9988	0.9885	0.7644	0.8169	0.5324	0.9962
Baboon	0.9999	0.9424	0.8079	0.8028	0.5967	0.9952

5 水印复原

从上文实验结果可以看出，本文算法对多种攻击操作都具有强鲁棒性，但是对剪切操作攻击的抵抗能力有些不理想。这是存在于大容量水印算法中的普遍问题。因此本文提出一个基于神经网络(NN)[10]的水印复原方法。根据提取出来的受攻击水印图像的残留信息，运用神经网络可以辨别出水印原型。经过训练的神经网络可以联想出原始水印的精确内容，对跟踪多媒体数据发行具有深远意义。

充分利用神经网络的灵活性和适应性，本文采用后向传播神经网络(BP)对残留水印的特征进行识别。水印特征由主成分分析方法(PCA)[11]提取。

5.1 神经网络训练

这一部分，要完成神经网络的训练以使神经网络具有记忆并识别不同水印特征的能力。首先，将原始水印进行不同类型和不同参数的攻击，得到一个大的受攻击过的水印图像集合，将其作为神经网络的训练集。然后，利用主成分分析方法提取这些图像的特征值。我们将每幅水印图像看作矩阵 X，对其进行 PCA 变换并选择前 12 个特征值作为训练样本的输入模式向量。神经网络输出层的每个结点代表不同的水印类型。12 个特征值记为 $B_k=(a_1,a_2,\cdots,a_p)$ 作为输入向量和神经网络输出 $Y_k=(y_1,y_2,\cdots,y_q)$ 共同构成了一组训练模式对，其中 p 为输入层结点数，q 为输出层结点数，k 为训练模式对数。BP 神经网络的结构如图 6 所示，网络包括输入层 12 个节点，隐含层 20 个结点，输出层 q 个结点，q 值可以根据实际水印类型数而改变。

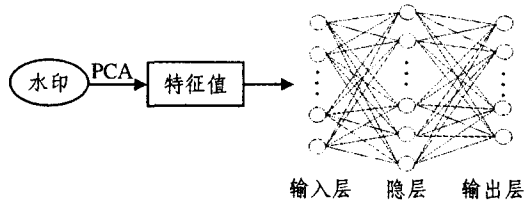


图 6 本文算法所使用的神经网络结构

5.2 基于神经网络的水印恢复

经过训练的神经网络在识别水印方面具备了很强的自适应能力。以一幅经过攻击的宿主图像中提取出来的水印作为测试图像，选择 PCA 变换后最大的 12 个特征值作为输入模式送入神经网络，被攻击水印的原型会被神经网络准确地识别出来。在实验中，即使受攻击水印的 NC 值已经低为 0.2394，神经网络仍然可以联想记忆出它的原始水印。因此本文中提出的水印复原算法是水印技术的一个重要补充。

总结 本文提出一个基于鲁棒水印算法来保护图像版权的技术。我们的方法是结合 SVD 和小波包分解实现水印嵌入的新算法。首先利用奇异值和方差的特征来预处理宿主图像，然后将分割的水印分别嵌入到宿主图像的不重叠位置，完全摆脱了水印重复嵌入的相互影响。而且，由于神经网络的灵活性，它的应用成功地实现了本文的水印复原算法。实验结果有力地表明我们的方法显著地增加了鲁棒水印信息的嵌入量及水印的抗攻击能力。

参 考 文 献

1 Cheung W N. Digital Image Watermarking in the Spatial and Transform Domains. In: Proc. of TENCON'2000, 2000

2 Chen Wen-Yuan, Chen Chin-Hsing. A Robust Watermarking Scheme using Phase Shift Keying with the Combination of Amplitude Boost and Low Amplitude Block Selection. Pattern Recognition, 2005, 38: 587～598

3 Shieh Jieh-Ming, Lou Der-Chyuan, Chang Ming-Chang. A semi-blind digital watermarking scheme based on singular value decomposition. Computer Standards & Interfaces, 2005

4 Zhang Zhi-Ming, Wang Lei. A novel SVD watermarking method with turbo code enhanced robustness. In: Proc. of the Intl. Computer Congress 2004 on Wavelet Analysis and its Applications, and Active Media Technology, v 1, 2004

5 Ganic E, Eskicioglu A M. Robust DWT-SVD domain image watermarking: embedding data in all frequencies. In: Proc. of the 2004 Multimedia and Security Workshop on Multimedia and Security, 2004. 166～174

6 Kundur D, Hatzinakos D. Digital Watermarking using Multiresolution Wavelet Decomposition. In: Proc. of the 1998 IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing, 1998, 5: 2969～2972

7 Suhail M A, Obaidat M S, Ipson S S, et al. A Comparative Study of Digital Watermarking in JPEG and JPEG 2000 Environments. Information Sciences, 2003, 151: 93～105

8 Gonzalez R C, Woods R E, Eddins S L. Digital Image Processing Using Matlab: Publishing House of Electronics Industry, Beijing, 2004

9 Sun R, Sun H, Yao T. A SVD and quantization based semi-fragile watermarking technique for image authentication. In: Proc. Internat. Conf. Signal Process, 2002, 2: 1952～1955

10 Yu Pao-Ta, Tsai Hung-Hsu, Lin Jyh-Shyan. Digital watermarking based on neural networks for color images. Signal Processing, 2001, 81: 663～671

11 Fortuna J, Capson D. Improved support vector classification using PCA and ICA feature space modification. Pattern Recognition, 2004, 37: 1117～1129