

一种多级数据嵌入的信息隐藏方法研究*

蔡龙飞¹, 赵慧民²

(1. 广东工程职业技术学院, 广东 广州 510520;
2. 广东技术师范学院, 广东 广州 510665)

摘要: 利用 Bandelet 变换实现数字图像特征信息自适应提取, 对水印信息实现自适应多级调制嵌入, 强化了对特征信息的保护和恢复, 以及对检测信号进行进一步的信息筛选。实验结果表明, 提出的数字水印特征多级嵌入调制及其检测的研究方法, 提高了水印图像的恢复性和抗攻击能力。

关键词: 数据调制; 信息隐藏; Bandelet; 数字水印

中图分类号: TP309 文献标志码: A 文章编号: 0529-6579 (2013) 02-0023-05

An Information Hiding Method Research of Multilevel Data Modulation Embedding

CAI Longfei¹, ZHAO Huimin²

(1. College of engineering occupation technology, Guangzhou 510520, China;
(2. Guangdong Polytechnic Normal University, Guangzhou 510665, China)

Abstract: Utilizing Bandelet transform to achieve adaptively features extraction of the digital image, the method can improve performances of protecting and recovering image features by multilevel data modulation & embedding of watermarking information, and choose information needed using detection technology at receiver. Experimentation results shows the proposed method can enhance recover performances of the digital watermark image and resistance to attacks.

Key words: data modulation; information hiding; Bandelet; digital watermarking

信息隐藏技术是实现保密通信和数字版权保护的一种有效方法, 已成为多媒体信息安全研究领域持续发展的一个研究焦点。目前, 基于小波变换的信息隐藏方法是实现信息隐藏的焦点技术^[1-5]。

小波较 Fourier 分析更擅长的是表示局部奇异特征, 但是严格采样的小波分析在表示多维信号时存在两大不足: 变换不具备平移不变性; 对具有方向性的几何奇异特征表示不稀疏。近年来, 为了克服以上缺陷, 同时保持小波多尺度分析和局部化时频分析能力, Daubechies^[6], Vetterli^[7], Mallat^[8], Donoho^[9]等国际权威小波研究专家提出了一系列全新的多尺度分析方法, 由于这些方法和理论均具有小波的重要特征, 我们称之为新型小波理论。

2000年, E Le Pennec 和 Stéphane, Mallat 提出了 Bandelet 变换^[8]。Bandelet 变换是一种基于边缘的图像表示方法, 能自适应地跟踪图像的几何正则方向。Bandelet 基并不是预先确定的, 而是优化最终的应用结果来自适应的选择具体的基的组成。Pennec 和 Mallat 给出了 Bandelet 变换的最优基快速寻找算法^[8]。初步实验结果表明, 与普通的小波变换相比, Bandelet 变换在去噪和压缩方面体现出了一定的优势和潜力^[10-11], 其压缩方法可以达到比 JPEG2000 等主流方法更好的压缩效果。根据 Bandelet 变换的性质, 文献 [12] 提出了一种基于 Bandelet 变换采用零水印思想的全频域无损水印算法, 该算法检测了载体图像特征, 但没有应用隐藏

* 收稿日期: 2012-10-09

基金项目: 广东省自然科学基金资助项目 (S2012010008639); 广东省科技计划资助项目 (2012B010100035)

作者简介: 蔡龙飞 (1976年生), 男, 副教授; E-mail: cailongfei2@126.com

技术。文献 [13] 提出了基于第 2 代 Bandelet 变换的认证水印算法, 利用了图像几何矢量线方向流特征, 但该算法没有充分利用水印图像的边缘正则性和水印信息不同分布特性, 因此, 信息隐藏提取与认证精度有限。

目前, Bandelet 变换通过自适应处理设计各向异性的基函数, 建立了各向异性基函数字典, 因此, 可以实现对图像中的边缘进行稀疏表示。由于 Bandelet 变换具有边缘信息稀疏表示的能力, 本文从一个新的角度来研究数字水印, 即通过研究数字水印的边缘 (特征信息) 的稀疏表示, 设计出一种对特征信息和一般信息采用不同隐藏策略的多级数据调制嵌入的数字水印算法, 并建立了数字水印特征级检测和识别的数学模型, 从而达到增加水印信息量, 提高数字水印抗攻击能力和检测识别的准确性的目的。

1 Bandelet 理论及其应用的研究

Bandelet 是一组正交基, 所以表示图像时信息

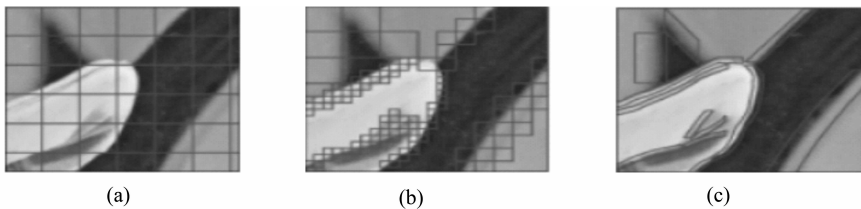


图 1 JPEG, JPEG 2000, Bandelet 三种变换编码思想图示

Fig. 1 Transform coding base principle of JPEG and JPEG 2000 as well as Bandelet

2 基于 Bandelet 的特征级数字图像水印算法

目前国内外鲜有应用 Bandelet 变换实现水印特征多级隐藏和检测的研究的报道。从对 Bandelet 变换的性质的介绍, 我们知道其克服了经典小波变换的弊端 (不具备平移不变性; 对具有方向性的几何奇异特征表示不稀疏)。利用 Bandelet 变换对水印图像 (尤其是数字指纹图像) 几何边缘的稀疏表示能力, 可以预见其应用于数字水印特征信息的提取和检测具有良好的性能。

2.1 基于 Bandelet 变换的图像几何特征条件分割

按照图 1 的分割模型对水印图像几何特征进行条件分割。这时, 信源全局和局部数据之间建立联系的基础是对信源信号联合分布 $p(x, \sigma_x^2)$ 的分析, 其获得的概率形式为:

$$p_{x, \Sigma_x^2}(x, \sigma_x^2) = p_{\Sigma_x^2}(\sigma_x^2) p_{x| \Sigma_x^2}(x | \sigma_x^2) \quad (1)$$

无冗余。它的自适应策略使得可以用较少的 Bandelet 系数表示图像在边缘处的奇异性质^[10], 如图 1 中 (a)、(b)、(c) 分别为 JPEG、JPEG2000、Bandelet 编码)。Pennec 和 Mallat 首先定义了一种能表征图像局部正则方向的几何矢量线 (Geometric flow of vectors); 再对图像的支撑区间 S 进行二进剖分 $S = \cup_i \Omega_i$, 当剖分足够细时, 每一个剖分区 Ω_i 中最多只包含图像的一条轮廓线 (边缘)。对于包含轮廓线的局部区域, 几何正则的方向就是轮廓的切线方向。根据局部几何正则方向, 在全局最优的约束下, 计算区域 Ω_i 上矢量场 $\tau(x_1, x_2)$ 的矢量线。再沿矢量线将定义在 Ω_i 上的区间小波以生成 Bandelet 基, 从而能够充分利用图像本身的局部几何正则性。所有剖分区 Ω_i 上的 Bandelet 的集合构成了一组 $L_2(S)$ 上的标准正交基。构造 Bandelet 变换的中心思想是定义图像中的几何特征为矢量场, 而不是看成普通的边缘的集合。矢量场表示了图像空间结构的灰度值变化的局部正则方向。

这里, $x \in X$ 为水印图像信号, σ_x^2 代表信号方差, $p_{\Sigma_x^2}(\sigma_x^2)$ 代表了水印信号的临界方差分布, 而条件分布 $p_{x| \Sigma_x^2}(x | \sigma_x^2)$ 表示了 Bandelet 变换获得局部系数数据的一种统计行为。对应临界分布条件, 全局数据统计特性为:

$$p_x(x) = \int_0^\infty p_{x, \Sigma_x^2}(x, \sigma_x^2) d\sigma_x^2 = \int_0^\infty p_{x| \Sigma_x^2}(x | \sigma_x^2) p_{\Sigma_x^2}(\sigma_x^2) d\sigma_x^2 \quad (2)$$

这样图像的统计形式就是一种 Bandelet 变换模型。这种关系为把图像数据分割后局部和全局系数之间的几何特征统计关系提供了一个基本 Bandelet 变换数据链。

2.2 基于 Bandelet 特征数据多级结构水印形式

根据 Bandelet 变换系数的不同分布及其不同重要性, 本文提出一种基于 Bandelet 变换结构的多级数据调制嵌入的信息隐藏实现方法。这种多级结构

是把相同空间位置和几何方向上的不同频率系数划分为同一组并构成一个“父-子”树结构的多级形式，它把数字水印图像特征点的不同数据按照 Bandedet 变换“父-子节点”关系的不同分解成不同能量的数字水印向量基，在多级结构的不同节点自适应生成不同强度的水印信息。

本文研究时设 Bandedet 变换系数树结构的深度为 $L=4$ （即为四级），每一级节点都有相同的子节点数目 $D_l, l = 1, 2, \dots, L$ ，对于树中的每个节点（根节点除外） $i = [i_1, \dots, i_l]$ ， $i_l = 1, 2, \dots, D_l$ ，唯一的水印数据向量基 $\{a_{YZ}^{i_1, i_2, \dots, i_l}\}$ （YZ 代表多小波树中的某个子带系数）满足高斯分布 $N(0, \sigma_w^2)$ 且相互独立。按照 SS 技术，Bandedet 变换水印数据向量基 $\{a_{YZ}^{i_1, i_2, \dots, i_l}\}$ 生成的形式如下：设 Bandedet 变换得到的水印图像几何特征数据的二进制序列为 G ，经过扩频系数 cr 形成扩展序列 E ，再与随机种子函数 $Sk = \text{randseed}(K, M)$ 进行置乱调制形成水印数据向量基 $a^{i_1, \dots, i_L} = \{m_1, m_2, \dots, m_{N-1}\}$ （这里，随机种子和密钥可结合视频内容的 Hash 变换获得）。

$$G = \{x_1, x_2, \dots, x_n\}, x_j \in \{0, 1\}, j \in \mathbf{N} \quad (3)$$

$$E = \{e_1, e_2, \dots, e_n\}, e_i = a_j,$$

$$j \cdot cr \leq i \leq (j + 1) \cdot cr, cr > 1, i, j, cr \in \mathbf{N} \quad (4)$$

$$S_k = \{k_1, k_2, \dots, k_n\}, k_i \in \{0, 1\}, i \in \mathbf{N} \quad (5)$$

$$a^{i_1, \dots, i_L} = \{m_1, m_2, \dots, m_{N-1}\}, m_i = e_i \cdot k_i, i \in \mathbf{N} \quad (6)$$

结合 Bandedet 变换分割和 Bandedet 变换树结构的多级数据调制嵌入，设信号 X 按照 Bandedet 变换几何特征分割进行统计，并实现了图像分割后，在 Bandedet 变换树中有 N 个可嵌入的子带系数，水印数据向量基 $\{a\}$ 具有同样的长度和能量，则按照多级“父-子节点”的关系，在树的第 i 个节点上（即对应第 i 个子信道， $1 \leq i \leq L$ ）嵌入的水印数据形式为：

$$W^{(i)} = \sqrt{\sigma_{x_1}^2} a^{i_1} + \sqrt{\sigma_{x_2}^2} a^{i_1, i_2} + \dots + \sqrt{\sigma_{x_L}^2} a^{i_1, i_2, \dots, i_L} \quad (7)$$

其中， $\sigma_{x_1}^2, \dots, \sigma_{x_L}^2 \in \sigma_x^2$ 分别为对应 L 个高斯子节点的信号几何特征数据的方差。设水印数据总的信号功率为 σ_w^2 ，失真 $d^N(Y', Y) = \frac{1}{N} \sum_{i=1}^N d(Y'_i, Y_i) \leq \sigma_w^2$ ，失真对应的加性高斯噪声的方差为 σ_z^2 ，数字水印数据的总长度为 N 。这样，对应于树的第 i 个节点上嵌入的水印数据，第 i 个并行信道上（ $1 \leq i \leq L$ ）分配的数据内容为：

$$Y^{(i)} = \alpha X^{(i)} + W^{(i)} \quad (8)$$

$$\alpha = \frac{\sigma_w^2}{\sigma_w^2 + \sigma_z^2}$$

表示了水印系统失真引入的补偿因子（或叫补偿系数）。

2.3 数字水印数据的多级调制嵌入与检测处理

把水印数据长度为 N 的隐藏信息 $a^{i_1, \dots, i_L} = \{m_1, m_2, \dots, m_{N-1}\}$ ， $m \in \{1, 2, \dots, 2^{NR}\}$ 映射编码成随机数字水印序列 $w^N = W^N(m, X^N, K^N)$ 。调制嵌入失真对应于数据隐藏（数字水印）信号的方差，其约束条件为 $E[\sum_{i=1}^N W_i^2] \leq N\sigma_w^2$ ，决定了由于水印信号嵌入引起的误码率。

这里，水印图像的几何特征数据呈现的状态 X 对编码而言是一种非因果关系，它由信源信息分布、攻击数据、和失真补偿组成。因此，按照 2.2 水印多级数据结构，本文采取的技术实现方案如图 2，图 3 所示。对应的实现方案，数字水印嵌入过程如下。

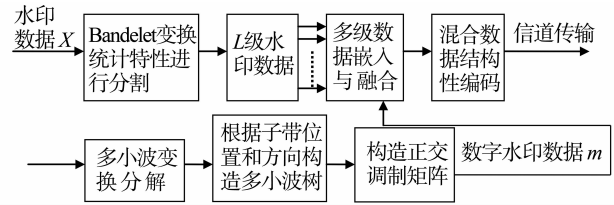


图 2 多级数据嵌入的信息隐藏实现方案（发送端）

Fig. 2 A information hiding scheme based multi-level data embedding at transmission

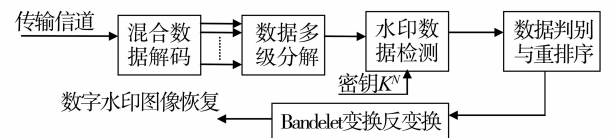


图 3 多级水印数据恢复的信息隐藏实现方案（接收端）

Fig. 3 A information hiding scheme based multi-level data embedding at receiving

1) 将待嵌入的水印图像进行 Bandedet 变换，得到一维不同子带的指纹数据和不同重要性的特征数据。

2) 对载体图像 S 用多小波变换进行 4 级分解，对小波系数进行重排，并把相同空间位置和方向上的不同频率子带系数划分为同一组作为多小波树的“父-子节点”分支树。

3) 在树的 $1 \leq l \leq L$ 级，对不同分支的子带系数设置高斯分布 $N(0, \sigma_w^2)$ 水印向量基 $\{a_{YZ}^{i_1, i_2, \dots, i_l}\}$ ；

水印数据由公式 (3) - (6) 产生。

4) 按照 $W^{(i)} = \sqrt{\sigma_{x_1}^2} a^{i_1} + \sqrt{\sigma_{x_2}^2} a^{i_1, i_2} + \dots + \sqrt{\sigma_{x_L}^2} a^{i_1, i_2, \dots, i_L}$ 的关系, 对树的各级 $1 \leq l \leq L$ 调制水印数据, 则在树的 $1 \leq l \leq L$ 级嵌入数据后的, 混合数据 (即包含载体数据和水印数据) 内容为 $Y_l^{(i_1, \dots, i_l)} = \alpha X_l + W_{ll}^{(i_1, \dots, i_l)}$ 。

5) 对嵌入水印数据后的多小波系数做反变换, 得到含数字水印的混合图像。

对应图 3 的实现方案, 数字水印检测与提取过程如下:

1) 对含水印的混合图像用 Bandelet 变换进行 4 级分解, 重排系数, 得到 Bandelet 变换树及其分组。

2) 对每个多级树结构的子带系数计算奇偶, 提取嵌入的水印数据 m 。

3) 相似度测量: 用归一化相关系数定义如下:

$$\eta = \frac{\sum_{l=1}^L W_l^{(i_1, \dots, i_l)} \cdot \overline{W}_l^{(i_1, \dots, i_l)}}{\sqrt{\sum_{l=1}^L W_l^{(i_1, \dots, i_l)^2} \cdot \sum_{l=1}^L \overline{W}_l^{(i_1, \dots, i_l)^2}}} \quad (9)$$

如果 $\eta > T$ (设置的门限), 则指纹图像数据已经检测出来。显然, η 越接近 1 越好。

3 实现结果及其性能分析

实验用 DGHM 多小波变换对 512×512 像素的 256 级灰度 Lena 载体图像进行 4 级多小波分解, 用 Bandelet 变换对待隐藏的 128×128 像素的指纹图像 (PSNR = 38.37 dB) 进行 4 级

Bandelet 变换分解^[13]。对称指纹树结构具有 $L = 4$ 级, 其中 $[\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_4^2] = [\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{2}]$ 。实验时, 指纹水印化数据基 $\{a\}$ 服从高斯分布 $N(0, \sigma_w^2)$, $\sigma_w^2 = 1$ 。图 4 (a), (b), (c), (d) 分别显示了 $512 * 512$ Lena 的原始图像, $128 * 128$ 被嵌入的指纹图像, Bandelet 变换后基于树结构多级嵌入后 Lena 图像, 多小波变换后嵌入后 Lena 图像。在接收端, 图 4 (e) 为本文提出的加入高斯噪声后接收端恢复的指纹图像, 图 4 (f) 为高斯噪声后接收端恢复的指纹图像。由图 4 可见, 由于 Bandelet 变换很好地利用了指纹形状的几何流特征, 因此, 本文多级嵌入指纹后混合图像具有很好的不可见性, 且提取的指纹图像 PSNR 高于常用 Wavelet 方法的 PSNR 约 7 dB, 说明本文信息隐藏失真较小, 同时具有较高的保真度和恢复能力。

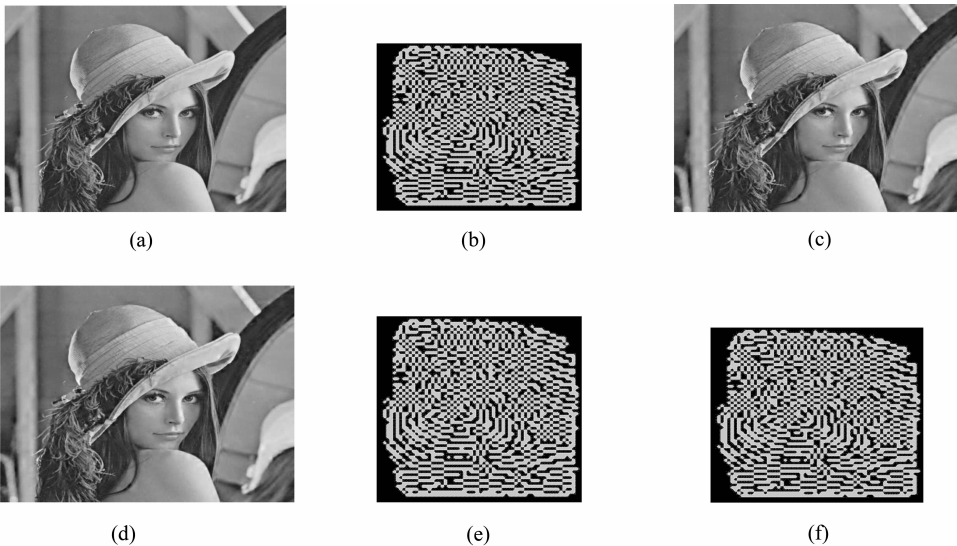


图 4 Lena 图像, 指纹图像, 混合图像和检测恢复的指纹图像

Fig. 4 Lena image, fingerprint image, stereo-image and received fingerprint image

(a) Lena 载体图像, PSNR = 38.37 dB; (b) 指纹水印图像, PSNR = 36.31 dB; (c) 多级嵌入后 Lena 图像, PSNR = 35.28 dB; (d) 多小波变换嵌入图像, PSNR = 34.69 dB; (e) Bandelet 变换检测恢复的指纹; (f) 多小波变换检测恢复的指纹图像, PSNR = 28.21 dB.

4 结 论

以 Bandelet 变换理论和算法研究为主线, 本文从基于 Bandelet 变换的数字水印图像分割和几何信息特征描述表示出发, 建立了一种有效的树结构多级水印数据分配模型。和基于小波变换的单级信息隐藏算法比较, 实验结果说明了本文方法的有效性。未来, 借鉴 Bandelet 变换的多级数据处理, 我们将从指纹水印信息的特征提取、隐藏、检测等算法构造一种密钥密码实现技术, 为数字水印的电子信息技术作为实际应用背景展开深入地研究。

参考文献:

[1] 赵耀. 基于小波变换的抵抗几何攻击的鲁棒视频水印[J]. 中国科学 E 辑, 2006, 36(2): 137 - 152.

[2] EDWARD J, WONG P W, DITTMANN J, et al. Security, forensics, steganography, and watermarking of multimedia contents X (proceedings volume) [DB/OL]. SPIE Digital Library, April 2008.

[3] VOLOSHYNOVSKIY S, PEREIRA S, PUN T. Attack modeling: Towards a second generation watermarking benchmark[J]. Signal Processing, Special Issue on Information Theoretic Issues in Digital Watermarking, June 2010, 90(6): 1177 - 1214.

[4] 赵慧民, 张惠乐, 胡学骏. 一种基于树结构的高斯分布扩频指纹嵌入技术的方法研究[J]. 中山大学学报: 自然科学版, 2006, 45(1): 42 - 45.

[5] 赵慧民, 康显桂, 黄继武. Laplacian 信源分割及其并行

Gaussian 信道传输的信息隐藏技术研究[J]. 中山大学学报: 自然科学版, 2008, 47(4): 29 - 34.

[6] DAUBECHIES I, HAN B, RON A, et al. Framelets: MRA based construction of wavelet frames[J]. Applied Comput Harmon Anal, 2004, 14(1): 1 - 46.

[7] DO M N, VETTERLI M. Contourlets, beyond wavelets [C]. G V Welland ed, Academic Press, 2003.

[8] PENNEC E L, MALLAT S. Image compression with geometrical wavelets[C]. In Proc. of ICIP2000. Vancouver, Canada, September, 2000: 661 - 664.

[9] DONOHO D, HUO X. Beamlets and multiscale image analysis, multiscale and multiresolution methods [J]. Springer Lecture Notes in Computational Science and Engineering, 2002(20): 149 - 196.

[10] PENNEC E L, MALLAT S. Bandelet representations for image compression[C]. Proceedings. 2001 International Conference on Volume 1, 2001: 12.

[11] PENNEC E L, MALLAT S. Sparse geometric image representations with bandelets [J]. IEEE Trans. Image Processing, 2005, 14(4): 423 - 438.

[12] 杨岳湘, 罗永, 叶昭辉. 基于 Bandelet 与自适应矩阵范数的全频域无损水印方法[J]. 计算机研究与发展, 2007, 44(12): 1996 - 2003.

[13] 刘绪崇, 罗永, 王建新, 等. 基于第二代 Bandelet 变换的图像认证水印算法[J]. 通信学报, 2010, 31(12): 123 - 130.

[14] 赵慧民, 郭一缜, 丁晓艳. 一种数字图像和指纹水印位平面合成的优化方法[J]. 中山大学学报: 自然科学版, 2010, 49(6): 64 - 67.

(上接第 22 页)

[6] 陈夺峰, 黄岁樑. 改进遗传算法确定水资源优化配置模型经济效益系数[J]. 水资源保护, 2008, 24(2): 6 - 9.

[7] 刘思峰, 党耀国, 方志耕等. 灰色系统理论及其应用[M]. 北京: 科学出版社, 2005.

[8] 王艳菊, 王珏, 吴泽宁. 基于灰色关联分析的支持向量机需水预测研究[J]. 节水灌溉, 2010, 10(4): 49 - 52.

[9] TIBSHIRANI, R. Regression shrinkage and selection via the lasso [J]. Journal of the Royal Statistical Society.

1996, 58(1): 267 - 288.

[10] VOLKER R. The generalized LASSO [J]. IEEE Transactions on Neural Networks, 2004, 15(1): 16 - 28.

[11] 李霞, 刘超. 基于收缩机制的若干回归模型比较研究[J]. 统计与决策, 2008, 5: 30 - 32.

[12] KOH K, KIM S J, BOYD S. An interior-point method for large-scale L1-regularized logistic regression [J]. Journal of Machine Learning, 2007(8): 1519 - 1555.