

文章编号: 1007-5321(2011) 增-0051-04

## 基于 DWT-DCT-SVD 的音频盲水印算法

雷 敏<sup>1,2</sup>, 杨 榆<sup>2</sup>

(1. 北京邮电大学 软件学院, 北京 100876; 2. 北京邮电大学 信息安全中心, 北京 100876)

**摘要:** 提出了一种结合离散小波变换(DWT)、离散余弦变换(DCT)和奇异值分解(SVD)的音频盲水印方法,将原始音频数据分段后进行二级小波变换,再将近似分量进行余弦变换,把余弦变化后的前 1/4 系数进行奇异值分解后嵌入水印。水印的嵌入容量为 27.56 bit/s。实验结果表明,该算法具有较好的透明性,并且对于 MP3 压缩、重量化、重采样、低通滤波、裁剪替换、高斯加噪等常见音频信号处理攻击具有很强的鲁棒性。

**关键词:** 小波变换; 奇异值分解; 盲水印

中图分类号: TN309

文献标志码: A

## An Audio Blind Watermarking Scheme Based on DWT-DCT-SVD

LEI Min<sup>1,2</sup>, YANG Yu<sup>2</sup>

(1. School of Software Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China;

2. Information Security Center, Beijing University of Posts and Telecommunications, Beijing 100876, China)

**Abstract:** A new audio blind watermarking scheme based on discrete wavelet transform-discrete cosine transform-singular value decomposition (DWT-DCT-SVD) is proposed. In this algorithm, the audio is split as blocks, and each block is decomposed on two-dimensional DWT, then the approximate sub-band coefficients are decomposed on DCT, and the first quarter of the DCT coefficients are decomposed on SVD transform and get a diagonal matrix. At last, the watermarking information is embedded into the diagonal matrix. The watermark embedding capacity is 27.56 bit/s. Experiments show that the algorithm can get better balance between transparency and robustness of watermark, and it has strong robustness against the common audio signal processing such as MP3 compression, re-quantization, re-sampling, low-pass filter, cutting replacement and additive white Gaussian noise.

**Key words:** discrete wavelets transform; singular value decomposition; blind watermarking

近年来,对图像水印算法的研究成果较多<sup>[1-4]</sup>,而音频数字水印方面的成果相对较少。人类的听觉系统比视觉系统敏感得多,因此,实现音频数字水印鲁棒性和透明性之间的平衡就更加困难,音频数字水印具有更大的挑战性。

目前的音频水印算法主要分为时域算法和变换域算法。在时域上,大部分水印算法不能抵抗一些常见的攻击,如文献[5]中提出的回声隐藏算法等。

在变换域中,一般都是将水印信号或音频载体信号进行离散小波变换、离散余弦变化或奇异值分解,然后在变换域中嵌入水印信号,但这种方法一般都不能实现水印盲提取,不利于实际应用<sup>[6-7]</sup>。目前也有一些半盲水印<sup>[8]</sup>和盲水印。

2010 年,Vivekananda B K 等<sup>[6]</sup>提出一种利用 SVD 变换和抖动调制量化的音频水印方案,该算法中未提及嵌入水印后的音频信号的信噪比(SNR)。

收稿日期: 2011-06-05

基金项目: 国家重点基础研究发展计划项目(2007CB311203); 国家自然科学基金项目(60821001, 90604022); 中央高校基本科研业务费专项资金项目(BUPT2009RC0215)

作者简介: 雷 敏(1979—),男,博士生, E-mail: leimin@bupt.edu.cn.

一些文献中给出了音频水印算法,但它们对某些攻击的鲁棒性不好,如对低通滤波和重采样攻击的误码率超过10%<sup>[9]</sup>;对MP3压缩攻击的误码率达到24%<sup>[10]</sup>.

针对上述问题,本文给出了一种结合DWT、DCT和SVD的音频盲水印方法,该算法具有较好的透明性,且效率较高.对该算法进行仿真实验后的结果表明,该算法对于MP3压缩、重量化、重采样、低通滤波、降噪、裁剪替换等常见音频信号处理攻击具有很强的鲁棒性.

## 1 背景知识

### 1.1 DCT特性分析

DCT作为一种性能优良的正交变换方式,具有3个特性.

1) DCT域中的值所表现的是音频文件的整体分布特性,非听觉特性.

2) 能量保持与集中特性.音频信号在DCT前后总能量保持不变,同时DCT变换需对能量重新分配,音频能量集中特性是DCT的一个显著特点,它只用少量低频系数就可代表信号的大部分能量.

3) 稳定性.对音频信号施加小的干扰,对应到DCT的变换系数将不会产生大的变化;反之,DCT域中的变换系数发生小的变动,对应到时域信号会将变动分散到整个信号中,也不会产生大的变化.

### 1.2 DWT特性分析

DWT变换具有2个特性.

1) 能量保持与集中特性.音频信号在一维DWT变换前后总能量保持不变,同时DWT变换对能量重新分配,DWT变换以后信号分为近似分量和细节分量.音频的大部分能量集中在近似分量.

2) DWT变换具有多分辨率分析的特点,而小波基和小波变换级数的选择更是给小波变换域数字水印算法的设计带来了很大的灵活性和优越性.进行DWT变换时可以根据算法的特点选择小波基和小波变换级数,基于DWT变换算法的复杂度和性能等根据自己的要求选择小波变换的级数和小波基.

### 1.3 SVD特性分析

SVD是一种将矩阵对角化的数值方法,是线性代数中最有用和最有效的工具之一,它在统计分析、信号处理、图像处理和系统理论中被广泛应用,具有如下特性:

1) 调整 $S$ 矩阵的值不会影响到信号的质量;

2)  $S$ 矩阵具有较好的稳定性,各种常规的信号处理不会对其值产生较大影响.

根据DWT变换、DCT变换、SVD变换和音频的特点,提出一种结合DWT、DCT和SVD三种变换的盲音频水印方法.

## 2 音频盲水印算法

结合DCT-DWT-SVD变换的音频盲水印算法首先对音频进行分段,对每个分段进行一维 $N$ 级DWT变换;再将DWT变换后的近似分量进行DCT变换;取出DCT变换后的前 $1/M$ 系数,再将这些系数进行SVD变换;在变换后的 $S$ 矩阵中嵌入水印信息.因为 $S$ 矩阵具有较强的稳定性,将 $S$ 矩阵的 $S(1,1)$ 和 $S(2,2)$ 取出来,利用 $\lfloor S(1,1)/(S(2,2)\Delta) \rfloor$ 的奇偶性嵌入水印信息.这里,取 $N=2$ , $M=4$ , $\Delta=0.5$ .

### 2.1 水印嵌入算法

水印嵌入算法的流程描述如下.

1) 水印图像预处理.选择二值图像作为水印图像,其大小为 $M_1M_2$ .因为载体是一维的音频文件,为了能将二维的二值图像作为水印信息嵌入载体中,需要对二值图像进行降维处理,把二维图像转化为一维向量,通过 $w = \{w(i) = w(m_1, m_2) | 0 \leq m_1 \leq M_1, 0 \leq m_2 \leq M_2, i = m_1M_2 + m_2\}$ 降维操作,水印 $w$ 中的像素 $w(m_1, m_2)$ 用元素 $w(i)$ 表示.

2) 将原始音频分段,每段的长度为1600样点,对每个分段进行一维二级DWT变换,取出DWT变换后的近似分量 $c_r$ 的长度为400.

3) 对近似分量 $c$ 进行DCT变换,取出变换后的系数 $y$ 前 $1/4$ 组成长度为100的向量 $Y$ ,并将这个向量转化为 $10 \times 10$ 的矩阵 $j$ .

4) 对每个 $10 \times 10$ 的矩阵 $j$ 进行SVD变换,得到一个 $10 \times 10$ 的对角阵 $S$ .

5) 将每个对角矩阵 $S$ 的第1个值 $S(1,1)$ 取出来,对其进行水印嵌入,嵌入方法为:

当 $\lfloor S(1,1)/(\Delta S(2,2)) \rfloor$ 是偶数,若嵌入的水印是1,则 $T(1,1) = S(2,2) \Delta (\lfloor S(1,1)/(S(2,2)\Delta) \rfloor + 1)$ ,若嵌入的水印是0,则 $T(1,1) = S(2,2) \Delta \lfloor S(1,1)/(S(2,2)\Delta) \rfloor$ ;当 $\lfloor S(1,1)/(S(2,2)\Delta) \rfloor$ 是奇数,若嵌入的水印是1,则 $T(1,1) = S(2,2) \lfloor S(1,1)/(S(2,2)\Delta) \rfloor$ ;若嵌入的水印是0,则 $T(1,1) = S(2,2) (\lfloor S(1,1)/(S(2,2)\Delta) \rfloor + 1)$ .

6) 对嵌入水印后的 $T$ 矩阵进行SVD反变换,

得到矩阵  $S'$ .

7) 将  $S$  变成一维向量, 替换步骤 3) 中得到的矢量  $y$  的前  $1/4$  分量得到新的  $y'$ , 对  $y'$  进行一维 IDCT 变换, 得到含水印的  $C$ .

8) 用含水印的  $C$  代替步骤 2) 得到近似分量, 与步骤 2) 得到的二级细节分量和一级细节分量进行一维二级逆小波变换, 得到嵌入水印后的音频分段.

9) 所有分段进行水印嵌入后, 得到含水印音频文件.

## 2.2 水印提取算法

水印提取流程描述如下.

1) 读取嵌入水印后的含水印音频文件, 得到含水印音频数据  $Y_1$ .

2) 将含水印音频分段, 每段的长度为 1 600 样点, 对每个分段进行一维二级 DWT 变换, 得到近似分量  $c_1$ .

3) 将  $c_1$  进行 DCT 变换, 取出 DCT 变换后的系数前  $1/4$  组成长度为 100 的向量  $Y_1$ , 并将这个向量转化为  $10 \times 10$  的矩阵  $j_1$ .

4) 对每个  $10 \times 10$  的矩阵  $j_1$  进行 SVD 变换, 得到一个  $10 \times 10$  的对角阵  $S$ .

5) 将每个对角矩阵  $S$  的第 1 个值  $S(1, 1)$  和第 2 个值  $S(2, 2)$  取出来进行对比, 得到 1 位水印信息. 判别公式为若  $S(1, 1) / (S(2, 2) \Delta)$  接近偶数, 则水印信息为 0; 若  $w(1, 1) / (S(2, 2) \Delta)$  接近奇数, 则水印信息为 1.

6) 所有分段进行以上步骤后, 得到提取后水印信息  $W$ , 将  $W$  变为二维图像矩阵, 得到水印图像.

## 3 实验与性能分析

为验证本文算法的透明性和鲁棒性, 选取了 3 类有代表性的音频: 男女声对话 (speech)、古典音乐 (classic) 和 流行音乐 (pop). 音频格式均为 44.1 kHz 采样、16 bit 编码、单声道.

图 1 ~ 图 3 分别为实验样本 speech、classic 和 pop 的载体音频和未受攻击嵌入水印后音频的时域波形.

试验中对加载水印音频进行以下顽健性测试:

1) 无攻击; 2) 添加 20 dB 的高斯噪声; 3) 以 0.5 倍的采样率做非整数倍下采样; 4) 以 11.025 kHz 为截止频率的低通滤波; 5) 8 bit 重新量化; 6) 在 64 kbit/s 的比特率下进行 MP3 压缩; 7) 在 32 kbit/s 的比特率

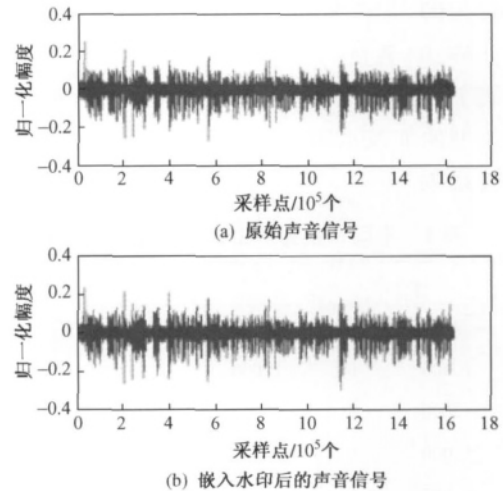


图 1 对话在嵌入水印前后的波形比较

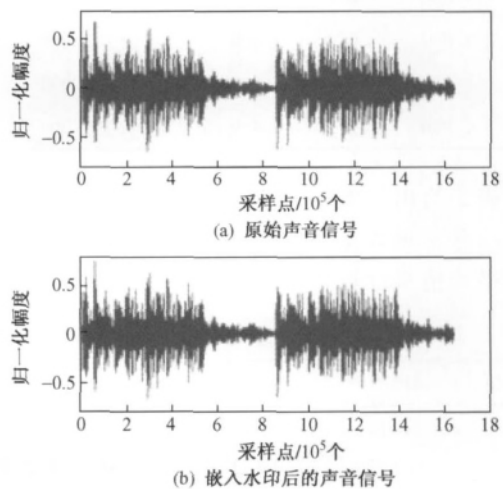


图 2 古典音乐在嵌入水印前后的波形比较

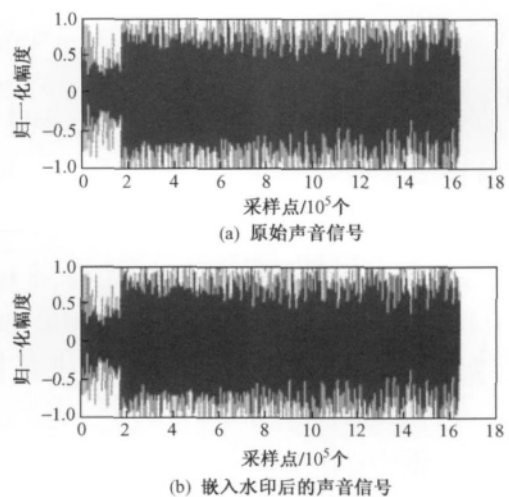


图 3 流行音乐在嵌入水印前后的波形比较

下进行 MP3 压缩; 8) 在 128 kbit/s 的比特率下进行 MP3 压缩; 9) 替换.

表 1 给出了本文算法在 100% 的嵌入率下 3 种不同类型的加载水印音频对上述攻击方式的误码率和相关系数.

表 1 不同音频类型在音频攻击后的结果

攻击	speech		classic		pop	
	NC	误码率	NC	误码率	NC	误码率
1	1.000	0	1.000	0	1.000	0
2	1.000	0	0.997	0.003	1.000	0
3	1.000	0	1.000	0	1.000	0
4	0.987	0.014	0.963	0.038	0.973	0.027
5	1.000	0	1.000	0	1.000	0
6	1.000	0	1.000	0	1.000	0
7	1.000	0	1.000	0	1.000	0
8	1.000	0	1.000	0	1.000	0
9	1.000	0	0.999	0.001	1.000	0

表 2 给出了本文算法和现有算法在加载水印音频对一些常见攻击方式的误码率. 这里按照 3 种音频的平均值来计算.

表 2 不同算法在音频攻击后的误码率 %

算法	高斯噪声	重采样	低通滤波	MP3 压缩
本文	0.10	0	2.63	0
	(20 dB)	(22.05 kHz)	(11.025 kHz)	(32 kbit/s)
文献[9]	5.13	13.64	18.06	5.71
	(36 dB)	(22.05 kHz)	(11.025 kHz)	(128 kbit/s)
文献[10]	4.98	0	未提及	24.18
	(16.12 dB)	(22.05 kHz)		(32 kbit/s)
文献[11]	2.34	2.03	2.19	4.38
	(10 dB)	(22.05 kHz)	(8 kHz)	(64 kbit/s)
文献[6]	0	1.00	0	2.00
	(20 dB)	(22.05 kHz)	(11.025 kHz)	(32 kbit/s)

本文的算法中,每个音频分段含1 600个样点,每个音频分段中嵌入 1 位水印信息,音频信号的采样率为 44.1 kHz,因此,该算法的水印容量为 44 100/1 600 = 27.6 bit/s.

4 结束语

提出一种结合 DWT、DCT 和 SVD 的音频盲水印方法,该算法具有较好的透明性,且效率较高.实

验结果表明,该算法对于 MP3 压缩、重量化、重采样、低通滤波、高斯加噪、裁剪替换等常见音频信号的处理攻击具有很强的鲁棒性.

参考文献:

[1] Alexander S, Scott D, Ahmet M E. Secure DCT-SVD domain image watermarking: embedding data in all frequencies [C]//ACM Multi-media and Security Workshop. 2004.

[2] Santhi V, Rekha S, Tharini S. A hybrid block based watermarking algorithm using DTW-DCT-SVD techniques for color images [C]//The International Conference on Computing, Communication and Networking. 2008: 1-7.

[3] Satyanarayana M P, Rajesh K. A robust digital image watermarking scheme using hybrid DWT-DCT-SVD technique [J]. International Journal of Computer Science and Network Security, 2010, 10(10): 185-192.

[4] El-Taweel G S, Onsi H M, Samy M, et al. Secure and Non-blind watermarking scheme for color images based on DWT [J]. GVIP Journal, 2005, 5(4): 1-5.

[5] Ferraiolo D, Kuhn R. Role-based access control [C]//Proceedings of 15<sup>th</sup> National Computer Security Conference. Washington, DC: IEEE, 1992: 554-563.

[6] Vivekananda B K, Indranil S, Abhijit D. An audio watermarking scheme using singular value decomposition and dither-modulation quantization [J]. Multimedia Tools and Applications, 2010, 52(2-3): 369-383.

[7] Ali A H. Digital audio watermarking based on the discrete wavelets transform and singular value decomposition [J]. European Journal of Scientific Research, 2010, 39(1): 6-21.

[8] Hamza O, Sankur B, Nasir M. An SVD-based audio watermarking technique [C]//ACM Multi media Conference. 2005: 51-56.

[9] Wang R, Xu D, Chen J, et al. Digital audio watermarking algorithm based on linear predictive coding in wavelet domain [C]//7<sup>th</sup> International Conference on Signal Processing. 2004: 2393-2396.

[10] Wu S, Huang J, Huang D, et al. Efficiently self-synchronized audio watermarking for assured audio data transmission [J]. IEEE Trans Broadcast, 2005, 51(1): 69-76.

[11] Zezula R, Misurec J. Audio digital watermarking algorithm based on SVD in MCLT domain [C]//Third International Conference on Systems. 2008: 140-143.