

文章编号:1000-582X(2012)08-062-05

DWT-DCT-QR 结合的音频盲水印算法

杨 榆^{a,b}, 雷 敏^{a,b}, 钮心忻^{a,b}, 杨义先^{a,b}

(北京邮电大学 a. 信息安全中心; b. 灾备技术国家工程实验室, 北京 100876)

摘 要:为保护互联网中数字音像制品的合法版权,结合离散小波变换(DWT)、离散余弦变换(DCT)、QR 分解以及音频的特性,提出了一种针对音频的盲水印算法。首先对原始音频分段,然后对每个分段进行小波变换求得近似分量,再进行离散余弦变换,对低频部分进行 QR 分解,在 QR 分解后得到的上三角矩阵中嵌入水印信息。实验结果表明,该算法具有较好的透明性和鲁棒性,能够承受高斯加噪、重采样、重量化、低通滤波、MP3 压缩、裁剪替换等常见音频信号处理攻击。

关键词:音频水印;离散小波变换;离散余弦变换;QR 分解

中图分类号:TP309.2

文献标志码:A

Audio blind watermarking scheme combined DWT-DCT-QR

YANG Yu^{a,b}, LEI Min^{a,b}, NIU Xin-xin^{a,b}, YANG Yixian^{a,b}

(a. Information Security Center; b. National Engineering Laboratory for Disaster Backup and Recovery, Beijing University of Posts and Telecommunications, Beijing 100876, P. R. China)

Abstract: In order to protect the copyright of digital audio and video in Internet, we propose a novel audio blind watermarking scheme combined discrete wavelets transform, discrete cosine transform, QR decomposition and audio characteristics. In this algorithm, the audio are split into blocks, and each block are decomposed on two-dimensional discrete wavelet transform (DWT), then the approximate sub-band coefficients are decomposed on discrete cosine transform (DCT), and the first quarter of the DCT coefficients are decomposed on QR decomposition and get a triangle matrix. At last, the watermarking information is embedded into the triangle matrix. The 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 additive white Gaussian noise, re-sampling, re-quantization, low-pass filter, MP3 compression and cutting replacement.

Key words: audio watermarking; discrete wavelets transform; discrete cosine transform; QR decomposition

随着互联网应用的广泛普及,越来越多的数字媒体信息需要在网络上发表和传播。为了保护数字媒体信息的版权,国内外许多科研人员正致力于研究数字水印技术在数字媒体信息的版权保护。由于人类听觉系统比视觉系统的敏感度要高,近年来致

力于研究图像数字水印的人员较多,而研究音频数字水印的人员较少,研究音频数字水印遇到的主要问题为数字水印的鲁棒性和不可感知性之间的平衡。文献[1-5]主要是在一些应用于图像的数字水印算法,文献[6]提出了一种 SVD 变换和抖动调制

收稿日期:2012-01-10

基金项目:国家自然科学基金资助项目(61170269、61003285、61170272、60973149、61170271)

作者简介:杨榆(1978-),女,博士,主要从事网络信息安全、信息影藏与数字水印方向研究,(Tel)13810223723;
(E-mail)leimin@bupt.edu.cn.

量化的音频水印方法,但算法中未提及嵌入水印后的音频信号的 SNR 值。文献[7-11]给出的音频水印算法,要么对某些攻击的鲁棒性不好,要么嵌入水印后音频有明显的杂音。

前期所研究的基于 DWT、DCT、SVD 的音频水印算法^[12-15],具有较好的透明性和鲁棒性,但算法效率与性能上还有待于提高。结合离散小波变换(DWT)、离散余弦变换(DCT)和 QR 分解的特性,针对音频的版权保护问题给出了一种新的音频水印方法,该算法不仅具有较好的鲁棒性,也同时具有良好的透明性,且算法效率得到了一定的提高。在仿真实验中,分别采用降噪、重采样、低通滤波、重量化、MP3 压缩、裁剪替换等常见的音频信号处理攻击手段对嵌入水印的音频信号进行处理后,重新提取水印,发现提取的水印信息清晰可见,并且计算得出的误码率、NC 值都比较理想,实验结果表明该算法对以上常见的攻击具有很强的鲁棒性,并且对嵌入水印的音频和原始音频分别找 10 人进行试听,音质与原始音频几乎一样,表明改算法具有比较好的不可感知特性。

1 DWT、DCT 变换与 QR 分解

1.1 DWT 变换

小波变换是一种新型的信号处理技术,得到越来越多的数字水印研究者的青睐。对音频这样的非平稳信号的处理与分析,选择小波变换尤为合适。DWT 变化的特性如下

1)能量集中特性。音频信号在进行一维小波变换变换前后总能量保持不变。DWT 变换后将得到近似分量和细节分量两部分,而音频的大部分能量集中在近似分量部分。

2)DWT 变换可以根据算法的特点选择小波基和小波变换级数,给 DWT 域数字水印算法的设计带来了很大的灵活性和优越性。

1.2 DCT 变换

DCT 变换是一种性能优良的正交变换方式,其特性如下:

1)DCT 域中的值所表现的是音频文件的整体分布特性而非听觉特性。

2)能量集中特性。音频信号在进行 DCT 变换前后总能量保持不变,DCT 变换后,音频信号的能量都集中在 DCT 变换的低频部分。因此,只需要用少量的低频系数就可代表信号的大部分能量。

3)稳定性。对音频信号进行一些小的干扰,音

频文件进行 DCT 变换后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。

1.3 QR 分解与 SVD 分解

QR 分解是将非奇异矩阵 $A \in R^{n \times n}$,分解成一个正规正交矩阵 Q 与上三角矩阵 R ,使得: $A=QR$,且分解唯一。由于正交矩阵是性能最好的矩阵,因此 QR 分解具有数值稳定性。

奇异值分解是对于矩阵 $A(m \cdot n)$,存在 $U(m \cdot m)$, $V(n \cdot n)$, $S(m \cdot n)$,使得: $A=U \cdot S \cdot V'$,其中 U 和 V 为正交阵, S 是一个对角矩阵。

从上述定义中,可推出:进行 QR 分解计算速度要比 SVD 分解速度快。

2 音频盲水印算法

结合 DCT、DWT、QR 变换的特性给出一种音频盲水印算法。该算法首先对音频分段后进行一维二级 DWT 变换求得近似分量,再对近似分量进行 DCT 变换后的取出前 1/4 系数转成二维矩阵,再进行 QR 分解,最后在 QR 分解后的上三角矩阵 S 中嵌入水印信息。

2.1 水印图像预处理

算法中的水印图像选取的是二值图像,大小为 $M_1 \times M_2$ 。由于音频文件是一维的,而二值图像文件为二维的。为了使二维的二值图像能作为水印信息嵌入到音频载体中,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。公式如下

$$w = \{w(i) = w(m_1, m_2), 0 \leq m_1 < M_1, 0 \leq m_2 < M_2, i = m_1 \times M_2 + m_2\}, \quad (1)$$

其中:水印 w 中的像素 $w(m_1, m_2)$ 由向量 w 中的元素 $w(i)$ 表示。

2.2 水印嵌入算法

嵌入算法流程如图 1 所示

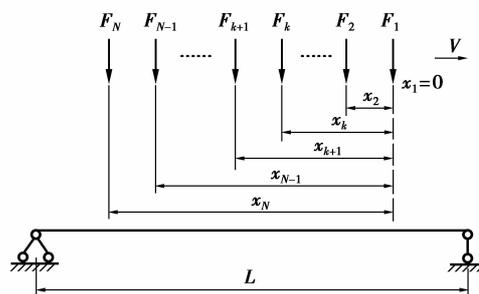


图 1 水印嵌入算法流程

算法详细描述如下:

1)对原始音频进行分段,每段长度为 1 600 样点。

2)对音频分段进行一维二级 DWT 变换,求得近似分量 $cA2$ 和细节分量 $cD1, cD2$ 。

3)对近似分量 $cA2$ 进行 DCT 变换,然后取出前 1/4 低频系数转换为转化为 10×10 的二维矩阵 jsi 。

4)对矩阵 jsi 进行 QR 变换,得到上三角矩阵 S 。

5)取出上三角矩阵 S 的第一个值 $S(1,1)$,对其进行水印嵌入,嵌入方法为

a)当 t 为偶数时,

$S_w(1,1) = S(2,2) \cdot \Delta \cdot (t+1)$, (嵌入水印是 1),

$S_w(1,1) = S(2,2) \cdot \Delta \cdot t$, (嵌入水印是 0),

b)当 t 是奇数时,

$S_w(1,1) = S(2,2) \cdot t$, (嵌入水印是 1),

$S_w(1,1) = S(2,2) \cdot (t+1)$, (嵌入的水印是 0),

令 $t = \lfloor S(1,1) / (S(2,2) \cdot \Delta) \rfloor$,仿真实验中 $\Delta = 0.5$ 。

6)对嵌入水印后的 S_w 矩阵进行 QR 逆变换,求得矩阵 S' 。

7)将矩阵 S' 转换为一维矩阵,替换步骤三中 DCT 变换后得到的前 1/4 低频系数,再进行一维 IDCT 变换,得到含水印的 $Ca1w$ 。

8)将 $Ca1w$ 替换步骤二中 DWT 变换得到的近似分量 $cA2$,再与细节分量 $cD1, cD2$ 进行一维二级 IDWT 变换,得到嵌入水印后的音频分段。

9)重复(2)~(8),直到所有分段完成水印嵌入后,将得到含水印音频文件。

2.3 水印提取算法

水印提取流程描述如下:

1)读取嵌入水印后的音频文件 Y_w ,并对其进行分段,使得每段的长度为 1 600 样点。

2)对音频分段进行一维二级 DWT 变换,求得近似分量 $cA2$ 。

3)对 DWT 变换后得到的近似分量 $cA2$ 进行 DCT 变换,然后取出前 1/4 的低频系数,转化成 10×10 的矩阵 jsi 。

4)对矩阵 jsi 进行 QR 变换,得到一个 10×10 的上三角矩阵 S_w 。

5)取出上三角矩阵 S_w 的 $S_w(1,1)$ 和 $S_w(2,2)$,如果 $S_w(1,1) / (S_w(2,2) \cdot \Delta)$ 的值接近偶数,则水印信息为 0,如果 $S_w(1,1) / (S_w(2,2) \cdot \Delta)$ 的值接近

奇数,则水印信息为 1。仿真实验中 $\Delta = 0.5$ 。

6)重复步骤(2)~(5),直至所有分段进行水印提取完毕,将得到提取的水印信息 W ,再将 W 转换为二维矩阵,即可得到水印图像。

3 仿真实验与性能分析

仿真实验中采用工具 MATLAB7.1。原始音频选取男女声对话(Speech)、古典音乐(Classic)和流行音乐(Pop)三种 wav 格式的音频文件,音频信号均为单声道、16 位编码、44.1 kHz 采样的音频。水印图像采用 32×32 的二值图像。

3.1 透明性测试

图 2、3 和 4 分别是男女声对话(Speech)、古典音乐(Classic)和流行音乐(Pop)三种 wav 格式的音频载体与未受攻击的嵌入水印后的音频的时域波形图。

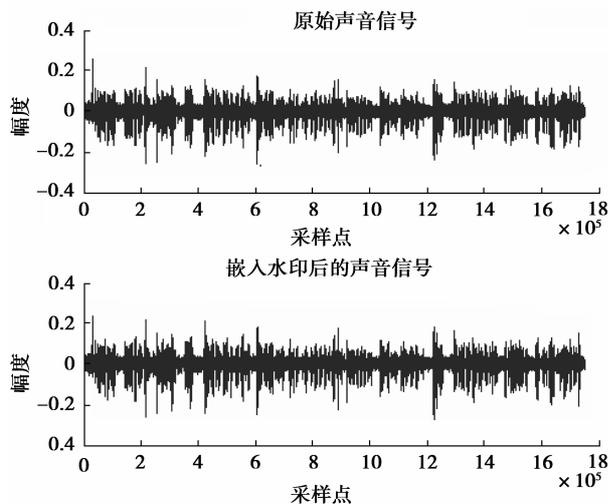


图 2 对话在嵌入水印前后的波形比较图

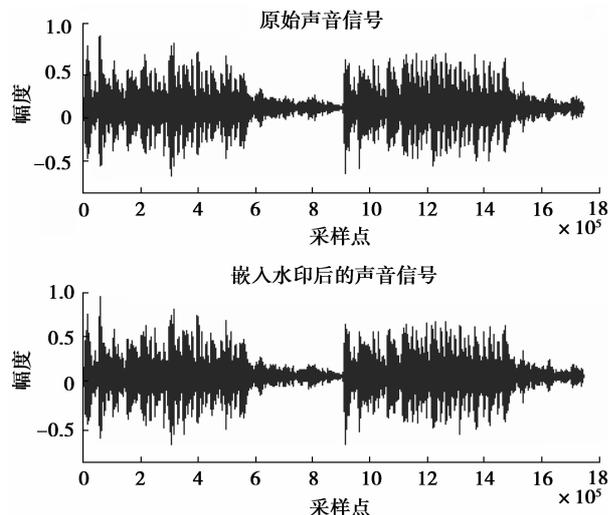


图 3 古典音乐在嵌入水印前后的波形比较图

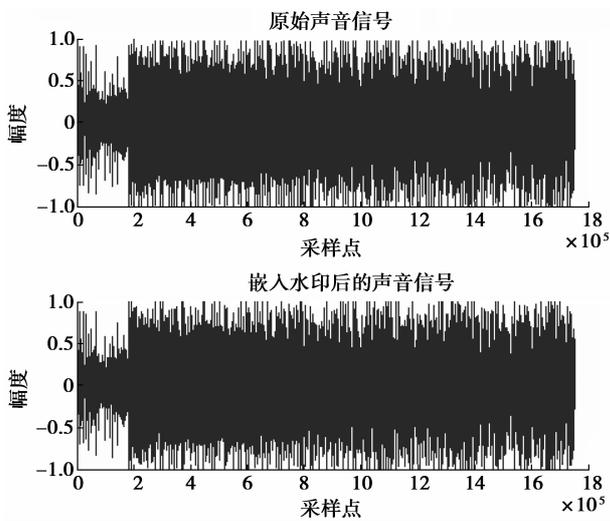


图 4 流行音乐在嵌入水印前后的波形比较图

从图中可以看出,原始载体音频与嵌入水印后的音频在波形上几乎没有什么改变。

实验中,找了 10 个同学分别试听对嵌入水印后的音频与原始音频,进行听觉上的比较与打分,最终得到的一个 MOS 值,如表 1 所示。

表 1 嵌入水印信息后 MOS 得分

音频文件名	MOS 得分
Speech	4.8
classic	4.9
Pop	4.7

3.2 健壮性测试

在仿真实验中(表 2、3、4),对嵌入水印后的音频分别进行了如下健壮性测试:无攻击、高斯噪声、重采样、低通滤波、重新量化、MP3 压缩和替换等几种常见的对音频信号的处理攻击(表 2、3、4)。

表 2 含水印 Speech 音频在受各类攻击后的比较表

攻击类型	NC	误码率	提取水印
无攻击	1	0	E
高斯噪声(20 db)	0.962 9	0.038 1	E
下采样(44100-22050-44100)	1	0	E
低通滤波(截至频率 11 025 Hz)	0.991 4	0.011 7	E
重量化	1	0	E
MP3 压缩 64 kbps	1	0	E
MP3 压缩 32 kbps	1	0	E
MP3 压缩 128 kbps	1	0	E
替换(裁剪)	1	0	E

表 3 含水印 Classic 音频在受各类攻击后的比较表

攻击类型	NC	误码率	提取水印
无攻击	1	0	E
高斯噪声(20db)	0.9791	0.0283	E
下采样(44100-22050-44100)	1	0	E
低通滤波(截至频率 1 1025 Hz)	0.987 8	0.016 6	E
重量化	1	0	E
MP3 压缩 64 kbps	1	0	E
MP3 压缩 32 kbps	1	0	E
MP3 压缩 128 kbps	1	0	E
替换(裁剪)	1	0	E

表 4 含水印 Pop 音频在受各类攻击后的比较表

攻击类型	NC	误码率	提取水印
无攻击	1	0	E
高斯噪声(20 db)	0.990 7	0.012 7	E
下采样(44100-22050-44100)	1	0	E
低通滤波(截至频率 11 025 Hz)	0.987 8	0.016 6	E
重量化	1	0	E
MP3 压缩 64 kbps	1	0	E
MP3 压缩 32 kbps	1	0	E
MP3 压缩 128 kbps	1	0	E
替换(裁剪)	1	0	E

表 2、3、4 所列数据为算法在 100% 的嵌入率下,分别对男女声对话(Speech)、古典音乐(Classic)和流行音乐(Pop)3 种不同类型音频在嵌入水印后的音频信号,进行上述攻击后提取的水印图像、求得的误码率和相关系数。从数据信息显示可知,该水印算法能够承受高斯噪声、重采样、低通滤波、MP3 压缩等音频信号处理的攻击,具有较好的稳健性。

4 结 语

结合 DWT、DCT 和 QR 的特点以及音频信号的特性,给出了一种新的音频盲水印方法。并通过 MATLAB7.1 仿真实验验证,该算法具有较好的鲁棒性和不可感知性,在水印提取上不需要原始音频和水印信息,实现了水印的盲提取。仿真实验还表明该算法高斯加噪、低通滤波、重采样、重量化、MP3

压缩、替换等常见音频信号处理攻击具有较好的稳健性。

参考文献:

- [1] Alexander S, Scott D, Ahmet M E. Secure DCT-SVD domain image watermarking: embed-ding 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, Com-munication 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] Vivekananda B K, Indranil S, Abhijit D. An audio watermarking scheme using singular value decomposition and dither-modulation quanti- zation[J]. Multimedia Tools and Applications, 2010, 52 (2-3): 369-383.
- [6] 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.
- [7] Ferraiolo D, Kuhn R. Role-based access control [C] // Proceedings of 15-th National Computer Security Conference. Washington, DC: IEEE, 1992: 554-563.
- [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. In: 7th International conference on signal processing, (ICSP'04), 2004, 1: 2393-2396.
- [10] Wu S, Huang J, Huang D, et al. Efficiently self-synchronized audio watermarking for assured audio data transmission. 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.
- [12] 雷敏. 音频数字水印与隐写分析算法研究[D]. 北京: 北京邮电大学, 2011.
- Lei Min. Research of the digital watermarking and steganalysis algorithms of audio information hiding[D]. Beijing: Beijing University of Posts and Telecommunications, 2011.
- [13] Yang Y, Lei M, Xiayun Jia et al. A new digital audio watermarking based on DCT-DWT-SVD [C] // Proceedings 2011 IEEE International Conference on Information Theory and Information Security August 7-9, 2011. Hangzhou China: IEEE, 2011, 906-909.
- [14] Lei M, Yang Y. Recent Advances in Audio-Based Steganalysis Research [C] // Proceedings of the 2nd Asia-Pacific Conference on Information Network and Digital Content Security. August 7-9, 2011. Hangzhou China: ATLANTIS Press, 2011, 208-213.
- [15] Lei M, Yang Y, Luo Shoushan, et al. Semi-fragile Audio Watermarking Algorithm in DWT Domain [J]. China Communications, 2010, 7(4): 71-75.

(编辑 侯 湘)