

DCT 和 DWT 域音频水印幅值变化规律分析

张金全^{1,2}, 王宏霞¹

(1. 西南交通大学信息科学与技术学院, 四川成都 610031; 2. 成都信息工程学院网络工程学院, 四川成都 610225)

摘要: 嵌入水印对音频信号波形的影响受到的关注不多. 分析得出改变采样序列离散余弦变换 (Discrete Cosine Transform, DCT) 域某个交流系数时, 逆 DCT 后幅值变化服从余弦曲线, 改变离散小波变换 (Discrete Wavelet Transform, DWT) 域某个近似系数时, 重构后采样序列幅值变化呈现为所选小波的尺度函数. 并对在 DWT-DCT 混合域嵌入水印做了讨论. 通过推导和实验, 得到采样点的改变量与 DCT 域低频系数的改变量、DWT 域近似系数的改变量及混合域低频系数的改变量成线性关系. 所得结论有助于发现嵌入算法带给采样点序列的不连续变化而使含水音频产生可听噪声, 也有助于提高水印的嵌入效率.

关键词: 音频水印; 离散余弦变换; 离散小波变换; 混合域; 可听噪声

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112 (2013) 06-1193-05

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2013.06.023

Analysis on Law of Distortion of Audio Signal for Embedding Watermark in DCT and DWT

ZHANG Jin-quan^{1,2}, WANG Hong-xia¹

(1. School of Information Science and Technology, Southwest Jiaotong University, Chengdu, Sichuan 610031, China;

2. School of Network Engineering, Chengdu University of Information Technology, Chengdu, Sichuan 610225, China)

Abstract: The distortion of audio signal associated with watermarking has received little attention. We have obtained some distortion laws in this paper. When a certain alternative coefficient is modified in DCT (Discrete Cosine Transform) domain, the law of audio amplitude distortion meets the cosine curve after inverse DCT. For the DWT (Discrete Wavelet Transform) domain audio watermarking, the distortion of audio signal shows the shape of scaling function after inverse DWT if a certain approximation coefficient in DWT domain is justified. Furthermore, analysis on embedding watermark in DWT-DCT domain is performed. By derivations and experiments, we draw the conclusions that the relationships between the variation of sample value and that of the low frequency coefficients in DCT, that of the approximation coefficient in DWT, and that of the low frequency coefficient in hybrid domain, are linear respectively. Our conclusions are helpful to find the discontinuities due to embedding watermark, which may cause audible noise, and improve the efficiency of embedding watermark.

Key words: audio watermarking; DCT; DWT; hybrid domain; audible noise

1 引言

近年来,随着互联网的飞速发展,数字音频的传播、存储变得方便快捷.由此带来的版权保护和内容可靠性等问题受到了人们的广泛关注,数字水印是解决上述问题的技术之一.

对音频嵌入水印时,研究者提出了许多在频域嵌入水印的算法.其中 DCT (Discrete Cosine Transform)、离散傅里叶变换、DWT (Discrete Wavelet Transform) 等被广泛使用.由于 DCT 是静态数字图像压缩标准 JPEG 的核心算

法,故 DCT 域水印方案在图像水印中得到了深入的研究,但在音频水印中用的比较少.文献[1]讨论了选择 DCT 域系数嵌入水印和引入的噪声不可听性之间的关系,文献[2]对文献[1]的结论进行了分析,认为应该将水印嵌入在幅值较大的音频部分.文献[3]则利用 DCT 域中的直流系数数据正负性不易改变的特性来嵌入水印信息,实现鲁棒性与不可感知性之间的平衡.

随着压缩标准 MPEG-4 的推出, DWT 域水印技术日益得到重视.文献[4]修改 DWT 域的近似系数嵌入同步信号与水印,在改善同步信号稳健性的同时,利用 DWT

的时频局部特性,较好地解决了同步信号稳健性与其搜索计算量之间的矛盾.文献[5]在时域重复嵌入同步信号,在 DWT 域低频系数的统计均值上嵌入水印信息,能抵御常规信号处理操作和一定程度上的同步攻击.

结合 DWT 和 DCT 的特性,文献[6]提出了在 DWT 和 DCT 构成的混合域嵌入水印,并用实验证明算法有较强的鲁棒性和较好不可听性,文献[7]采用在 DWT 和 DCT 的混合域嵌入水印.其他一些水印算法中也用到了 DCT^[8]或 DWT^[9].

在音频中嵌入水印,最终都会反应到音频采样点幅值的变化上.文献[10]注意到了由于嵌入时幅值的修改,节边界音频波形的不连续性而产生可听噪声.文献[11]注意到了同步码的嵌入对音频波形的影响,并指出文献[6]的含水印音频会产生轻微的“喀”声.文献[12]比较了含水印音频与原始音频幅值差值.由于 DCT 和 DWT 在音频水印技术中的广泛使用,本文关注了这两类变换在嵌入水印对音频波形的影响.文中默认使用抖动调制规则嵌入水印.

2 DCT 域音频水印幅值变化规律分析和仿真

由于音频是一维时间序列信号,故一般使用一维 DCT 处理.设音频信号为 $x = (x_0, x_1, \dots, x_{N-1})$, 一维 DCT 和 IDCT 的表达式如下:

$$y_k = c_k \sum_{n=0}^{N-1} x_n \cos[(2n+1)k\pi/2N], k \in [0, N-1] \quad (1)$$

$$x_n = \sum_{k=0}^{N-1} c_k y_k \cos[(2n+1)k\pi/2N], n \in [0, N-1]$$

其中 c_k 是常数.

设在嵌入信息时对第 i 系数的修改量为 e ,则执行 IDCT 后^[2],

$$x'_n = x_n + c_i e \cos[(2n+1)i\pi/2N], n \in [0, N-1] \quad (2)$$

通常水印算法是修改 DCT 的低频系数嵌入水印信息.如果修改的是直流系数,即式(2)中 $i=0$,所有样本序列的幅值的变化值是相同的,变化值为:

$$x'_n = x_n + c_0 e = x_n + e/\sqrt{N}, n \in [0, N-1] \quad (3)$$

如果修改的是第 i 个交流系数,可以得到:

$$x'_n = x_n + e\sqrt{2} \cos((2n+1)i\pi/2N)/\sqrt{N}, n \in [0, N-1] \quad (4)$$

音频采样点的幅值变化服从 $e\sqrt{2} \cos((2n+1)i\pi/2N)/\sqrt{N}$, $n \in [0, N-1]$,近似于 $e\sqrt{2} \cos\alpha/\sqrt{N}$, $\alpha \in (0, i\pi)$ 上以 $i\pi/2N$ 为周期的均匀采样,当 N 较大时, $\cos(\pi/2N) \approx 1$,即对于交流系数的修改,最大值约为 $e\sqrt{2}/\sqrt{N}$.

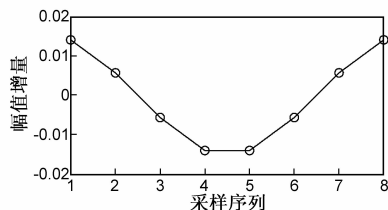
由上面的分析,如果通过修改 DCT 域直流系数嵌

入水印,与原来音频相比,分帧边界处可能出现 $2e/\sqrt{N}$ 的变化,因为相邻帧可能都是最大修改值且修改方向相反.而在交流系数嵌入水印时,这个值接近 $2e\sqrt{2}/\sqrt{N}$.

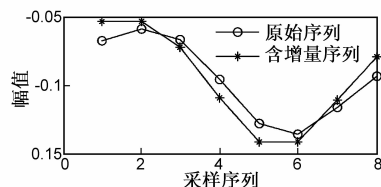
由于 DCT 是正交变换,变换前后信号能量不变,故对任一个 DCT 域系数的修改增量为 e ,其在频域和时域引入的噪声的能量都为 e^2 .故修改值 e 的大小决定了嵌入算法的信噪比,这也与文献[2]推导的结论一致.故对于不同系数的修改,如果修改值相等,噪声的能量相等,但反应到时域上,音频采样点幅值的变化却是不同的.由文献[10]可知,嵌入水印时,对于幅值的修改,分帧边界处可能产生音频波形不连续的情况,从而引起可听噪声,故嵌入水印需要注意分帧边界处音频幅值的平滑性.

由式(2)可知,音频某个采样点幅值的改变量与 DCT 域系数的变化量是线性关系,这可以用来提高算法的嵌入效率.

下面是时域采样序列因修改 DCT 域系数引起的幅值变化的仿真.为便于观察,DCT 的序列长度为 8,对 DCT 域系数的修改增量为 0.03.图 1(a)是修改第二个交流系数,经 IDCT 变换后采样序列的增量,图 1(b)是采样序列和修改后采样序列的对比.实验表明,前面的分析与仿真的结果是吻合的.



(a) 修改第2个交流系数时采样序列幅值变化



(b) 修改第2个交流系数,所得序列与原来序列的波形
图1 修改DCT交流系数对序列的影响

以文献[1]的仿真实例为例,执行 DCT 的长度为 8,水印嵌入在第 1 个系数(直流系数)或者第 4 个系数.设某帧音频信号为 $x = (x_0, x_1, \dots, x_7)$,如果嵌入在直流系数上,故只需计算直流系数.由式(1),直流系数为 $y_1 = e/\sqrt{8} \times \sum_{i=0}^7 x_i$,若修改量为 e ,由式(3),单个采样点幅值改变量 $e/\sqrt{8}$,即与低频系数修改量 e 成为线性关系,则含水印的音频信号为 $x' = (x_0 + e/\sqrt{8}, x_1 + e/\sqrt{8}, \dots, x_7 + e/\sqrt{8})$,不需要执行 DCT 和 IDCT.嵌入在

第 4 个系数上的情形也可以类似得到。

3 DWT 域音频水印幅值变化规律分析和仿真

先以 Haar 小波为例进行说明。

设 $(x_0, x_1, \dots, x_{2^n-1})$ 是长度为 2^n ($n > 1$) 的音频采样序列, 进行 1 级 Haar 小波分解, 可得近似系数^[13]:

$$A_{1,k} = (x_{2k} + x_{2k+1})/\sqrt{2}, k \in [0, 2^{n-1} - 1],$$

细节系数:

$$D_{1,k} = (x_{2k} - x_{2k+1})/\sqrt{2}, k \in [0, 2^{n-1} - 1]$$

如果近似系数 $A_{1,k}$ 都有一个增量 e , 即 $A'_{1,k} = A_{1,k} + e$, 则进行小波重构后,

$$x_{2k} = (A'_{1,k} + D_{1,k})/\sqrt{2} = (A_{1,k} + e + D_{1,k})/\sqrt{2}, k \in [0, 2^{n-1} - 1]$$

$$x_{2k+1} = (A'_{1,k} - D_{1,k})/\sqrt{2} = (A_{1,k} + e - D_{1,k})/\sqrt{2}, \in [0, 2^{n-1} - 1] \quad (5)$$

也即所有的采样序列的幅值有同一个增量 $e/\sqrt{2}$ 。

进一步的推导表明, 如果进行 j 级小波分解, 所有近似系数有相同的改变值 e , 则采样序列所有的采样点的值也都有同一个增量 $e/(\sqrt{2})^j$ 。如果修改某一个近似系数, 进行 j 级小波分解, 由 Haar 小波的分解和重构表达式, 则与这个近似系数对应的 2^j 个采样点的值有相同增量 $e/(\sqrt{2})^j$ 。

由前面的推导结果可知, 对于 Haar 小波, 无论小波分解的级数为多少, DWT 域近似系数的改变量与音频的采样点幅值的变化量是线性关系。进一步的实验表明, 对于 dbN ($N \geq 2$) 小波, 这个线性关系仍然是成立的。

下面是修改 DWT 域近似系数后, 时域采样序列幅值变化的仿真。实验中的序列长度为 32, 进行 2 级小波分解后, 对第 3 个近似系数修改增量为 0.03, 重构后采样序列的幅值增量见图 2。

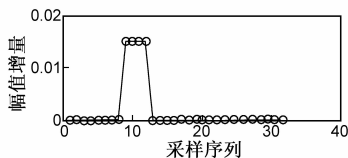


图2 修改Haar小波一个近似系数时, 序列幅值的变化

对于 dbN ($N \geq 2$) 小波, 实验表明, 如果进行 j 级小波分解, 当所有近似系数有相同的增量 e , 序列所有采样点的值也都有同一个增量 $e/(\sqrt{2})^j$ 。由小波的性质可知, 这与选择的小波无关, 相当于同向改变该段音频的幅值。进一步的实验选用了 Daubechies、Coiflets、Symlets 三个系列的多个小波函数, 执行多级分解。对近似系数有相同的增量的修改时, 各个采样点增量有一点差异,

但这个差异小, 接近采样点增量的万分之一甚至更小, 故可以忽略。由此可见, 可以认为对小波分解的所有近似系数一个相同增量, 实际是对所有采样点有相同的增量。

如果修改某一个近似系数, 由于 dbN ($N \geq 2$) 小波在 N 取不同值时, 尺度函数和小波函数不一样, 支撑区间各不相同。但不论进行多少级小波分解, 反映到采样序列上的增量图形呈现为所选小波的尺度函数的形状, 这与 Haar 小波是一致的。如果修改的是近似系数中的前几个或者后几个, 反映到采样序列上的增量图形呈现为所选小波的尺度函数的部分形状。

实验中, 采样序列长度为 64, 选用 $db5$ 进行 2 级分解, 修改第 10 个近似系数, 增量为 0.03, 图 3 是序列幅值增量图。可以看到, 重构后幅值的变化呈现为所选小波的尺度函数的形状。

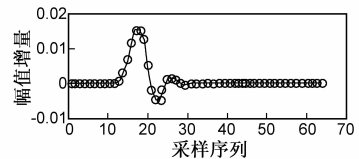


图3 修改db5小波的一个近似系数时, 序列幅值的变化

以文献[4]为例, 算法是将水印嵌入在音频段的近似系数上。在实验中, 作者比较了使用 Daubechies、Coiflets、Symlets 小波的结果, 指出小波基的选择对算法的性能影响很小, 故采用了最简单的 Haar 小波, 执行的是 8 级小波分解。

继续按照式(5)的推导, 执行 8 级 Haar 小波分解就是由 $2^8 = 256$ 个采样点获得一个近似系数。设音频采样点为 $(x_0, x_1, \dots, x_{255})$, 该近似系数可如下计算:

$$A_{8,1} = \sum_{k=0}^{255} x_k/(\sqrt{2})^8 = \sum_{k=0}^{255} x_k/16, k = 0, 1, \dots, 255$$

嵌入时 $A_{8,1}$ 的改变量为 e , 由前面推导, 嵌入后的音频采样点的改变量与近似系数的改变量 e 成线性关系, 为 $e/(\sqrt{2})^j = e/(\sqrt{2})^8 = e/16$, 则含水印的音频采样点的幅值即为 $(x_0 + e/16, x_1 + e/16, \dots, x_{255} + e/16)$, 嵌入过程不需要执行小波分解和小波重构。

若是选用 dbN ($N \geq 2$) 小波, 则需要通过如上面实验中的方法, 获得近似系数的单位改变量与对应的音频采样点的变化值的线性关系, 乘以嵌入强度后加上音频采样点的值, 就可得到含水印音频。这样, 在嵌入水印后不需执行小波重构就可以获得含水印音频。

4 DWT 和 DCT 混合域音频水印幅值变化规律分析和仿真

下面的讨论是假定先对采样序列进行 DWT, 再对小波分解后的近似系数进行 DCT, 也即对音频的低频部

分进行 DCT.

设 $\mathbf{x} = (x_0, x_1, \dots, x_{2^n-1})$ 是一个长度为 2^n ($n > 1$) 的采样序列, 先分析选用 Haar 小波的情况.

设对 \mathbf{x} 进行 j 级小波分解后的近似系数为 $A_k, k \in [0, 2^{n-j} - 1]$, 然后对近似系数 A_k 进行 DCT 变换, 设直流系数为 dA , 交流系数为 $aA_k, k \in [1, 2^{n-j} - 1]$. 下面讨论修改 DCT 低频系数对于音频幅值的影响.

如果修改的是直流系数 dA , 设增量为 e , 由(3)式, 经 IDCT 变换后, 所有的近似系数增加一个相同的幅值 $e/2^{(n-j)/2}$, 再经逆 DWT, 从而所有的采样点的增量为 $e/2^{n/2}$. 这是一个与小波分解级数无关的表达式, 也就是说, 如果修改的是直流系数 dA , 等同于对该段音频所有样本序列一个相同的增量 $e/2^{n/2}$.

如果修改的是第一个交流系数 aA_1 , 设增量为 e , 由式(4), 经 IDCT 变化后, 小波分解的近似系数的增量变化为 $e\sqrt{2}\cos((2m+1)\pi/2^{n-j+1})/2^{(n-j)/2}, m \in [0, 2^{n-j} - 1]$, 每一个近似系数会影响 2^j 个音频采样点的值, 使得这 2^j 个采样点有相同的增量.

修改 DCT 其他交流系数的结果可以按此分析方法得到.

由前面的推导可知, 如果选用 Haar 小波, 无论小波分解的级数为多少, 混合域低频系数的改变值与音频采样点幅值的变化量是线性关系. 进一步的实验表明, 如果选用 dbN ($N \geq 2$) 小波, 这个结论仍然成立.

实验中, 取 $n = 5$, 小波分解级数 $j = 2$, 对于 DCT 系数的修改增量取 $e = 0.01$. 图 4 是修改第二个交流系数时, 序列幅值增量图. 如分析的一样, 由于小波分解级数为 2, 故每个近似系数影响 $2^2 = 4$ 个采样点, 总的趋势服从 $\cos\alpha, \alpha \in (0, 2\pi)$.

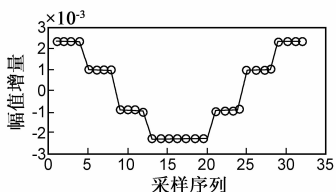


图4 选用Haar小波, 修改混合域第2个交流系数时序列幅值的变化

下面是对选用 dbN ($N \geq 2$) 小波的情况的分析.

如果修改的是直流系数 dA , 设增量为 e , 由于是对所有近似系数做的 DCT, 则经过 IDCT 后, 所有近似系数有相同增量, 但由于小波的支撑区间增大, 近似系数的增量小于 Haar 小波的 $e/2^{(n-j)/2}$, 故再经逆 DWT 变换后, 采样序列幅值增量相同, 但增量小于选用 Haar 小波的 $e/2^{n/2}$, 且 N 越大, 增量越小.

如果修改的是 DCT 的某个交流系数, 反应到音频序列幅值上的增量与小波分解级数相关. 随着分解的

级数增大, 采样序列幅值增量的趋势不再服从 $\cos\alpha$.

下面是采用“db2”小波的仿真实验, 小波分解级数为 2, 增量为 0.03, 修改第二个交流系数时采样点幅值增量分布如图 5 所示, 趋势基本服从 $\cos\alpha, \alpha \in (0, 2\pi)$. 曲线比使用“db1”小波时更平滑, 但幅值的最大增量减小了, 且没有使用“db1”时具有的对称性质.

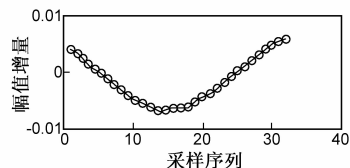


图5 选用db2小波, 修改混合域第2个交流系数时序列幅值的变化

结合前面分析结果应用于文献[1,4]中算法的仿真, 所得结论可用于文献[6,7]的仿真实现.

5 结论

由于在音频中嵌入水印时对音频的修改会引入噪声, 本文分析了在 DCT 域、DWT 域和它们的混合域嵌入水印时, 对于系数的修改, 反映到时域音频序列采样值的变化规律. 并在分析后, 结合现有文献算法, 把所得结论应用于算法的仿真实现. 所得结论可以提高算法的效率, 降低算法的复杂度, 以及发现算法中可能因嵌入水印产生音频波形不平滑而引入可听噪声的缺陷.

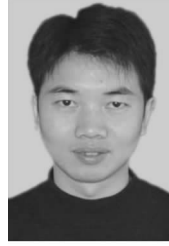
参考文献

- [1] 马翼平, 韩纪庆. DCT 域音频水印: 嵌入对策和算法[J]. 电子学报, 2006, 34(7): 1260 - 1264.
Ma Y P, Han J Q. Audio watermarking in DCT: embedding strategy and algorithm[J]. Acta Electronica Sinica, 2006, 34(7): 1260 - 1264. (in Chinese)
- [2] Zeng G R, Qiu Z D. Audio watermarking in DCT embedding strategy and algorithm[A]. Proceedings of the 9th International Conference on Signal Processing (ICSP) [C]. Beijing, China: IEEE Press, 2008. 2193 - 2196.
- [3] 温泉, 王树勋, 年桂君. DCT 域音频水印: 水印算法和不可感知性测度[J]. 电子学报, 2007, 35(9): 1702 - 1705.
Wen Q, Wang S X, Nian G J. Audio watermarking in DCT domain: algorithm and measurement of imperceptibility[J]. Acta Electronica Sinica, 2007, 35(9): 1702 - 1705. (in Chinese)
- [4] Wu S Q, Huang J W, Huang D R, et al. Efficiently self-synchronized audio watermarking for assured audio data transmission[J]. IEEE Transactions on Broadcasting, 2005, 51(1): 69 - 76.
- [5] Wang X Y, Niu P P, Yang H Y. A robust digital audio watermarking based on statistics characteristics[J]. Pattern Recognition, 2009, 42(11): 3057 - 3064.
- [6] Wang X Y, Zhao H. A novel synchronization invariant audio

- watermarking scheme based on DWT and DCT [J]. IEEE Transactions on Signal Processing, 2006, 54(12): 4835 – 4840.
- [7] Wang H X, Fan M Q. Centroid-based semi-fragile audio watermarking in hybrid domain [J]. Science China Information Sciences, 2010, 53(3): 619 – 633.
- [8] Lei B Y, Soon I Y, Li Z. Blind and robust audio watermarking scheme based on SVD-DCT [J]. Signal Processing, 2011, 91(8): 1973 – 1984.
- [9] V Bhat K, Sengupta I, Das A. An adaptive audio watermarking based on the singular value decomposition in the wavelet domain [J]. Digital Signal Processing, 2010, 20(6): 1547 – 1558.
- [10] Lie W N, Chang L C. Robust and high-quality time-domain audio watermarking based on low-frequency amplitude modification [J]. IEEE Transactions on Multimedia, 2006, 8(1): 46 – 59.
- [11] Megías D, Serra-Ruiz J, Fallahpour M. Efficient self-synchronised blind audio watermarking system based on time domain and FFT amplitude modification [J]. Signal Processing, 2010, 90(12): 3078 – 3092.
- [12] Wang X Y, Niu P P, Lu M Y. A robust digital audio watermarking scheme using wavelet moment invariance [J]. Journal of Systems and Software, 2011, 84(8): 1408 – 1421.

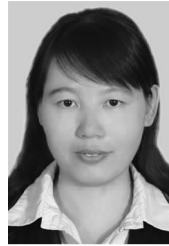
- [13] 孙延奎. 小波分析及其应用 [M]. 北京: 机械工业出版社, 2005, 10 – 16.

作者简介



张金全 男, 1974 年出生于四川邻水, 现为西南交通大学信息科学与技术学院博士研究生, 主要研究方向为音频数字水印.

E-mail: zhjq@cuit.edu.cn



王宏霞 女, 1973 年出生于河北省赵县, 现为西南交通大学信息科学与技术学院教授、博士生导师, 主要研究方向为多媒体信息安全技术.

E-mail: hxwang@home.swjtu.edu.cn