

基于小波域的鲁棒音频数字水印算法

林克正, 钟媛, 范波

(哈尔滨理工大学 计算机科学与技术学院, 黑龙江 哈尔滨 150080)

摘要: 为有效保护音频数字产品, 增强其抗同步攻击能力, 在小波域的基础上, 提出了一种新的盲检测数字音频水印算法. 先将完整音频进行三层小波分解, 取其低频系数, 再对每个音频帧进行三层小波分解, 也取其低频系数, 最后分别求出两组低频系数的平均值, 并进行比较判断水印嵌入位置. 嵌入水印后的音频信号与原音频信号在听觉上是相似的. 实验结果表明, 该方法既可增强水印的不可见性, 又能抵御多种攻击, 具有很好的鲁棒性和实用性.

关键词: 音频水印; 离散小波变换; 同步; 鲁棒性

中图分类号: TP391.41 **文献标志码:** A **文章编号:** 1007-2683(2010)04-0081-05

Robust Audio Watermarking Scheme Using Wavelet Transform

LIN Ke-zheng, ZHONG Yuan, FAN Bo

(School of Computer Science and Technology, Harbin University of Science and Technology, Harbin 150080, China)

Abstract: In order to protect digital audio products more effectively and improve the ability of anti-synchronization attack, a new blind audio watermarking algorithm based on wavelet transform is proposed. First, the algorithm do three-level wavelet decomposition to the whole audio signal and get the low frequency coefficients, then do three-level wavelet decomposition to each audio frame and get the low frequency coefficients, finally find the watermark embedding positions according to comparing the two average values from the two groups low frequency coefficients. The watermarked audio signal sounds similar to the original audio signal. Experiments show that this algorithm can not only enhance invisibility of the watermark more effectively, but also resist several attacks. This algorithm is robust and practical.

Key words: audio watermark; discrete wavelet transform; synchronization; robust

0 引言

音频水印的算法主要分两种: 时域的算法和变换域的算法. 时域水印的嵌入是通过修改原始音频数据的时域采样值强度实现的^[1]. 由于水印要均衡于鲁棒性和透明性之间, 因而可选择的属性范围较小, 鲁棒性较差. 变换域的算法是通过修改原始音频信号的变换域系数进行水印嵌入的. 常见的变换域

算法除了离散傅立叶变换(Discrete Fourier Transform, DFT)以外, 还有离散余弦变换(Discrete Cosine Transform, DCT)^[2]和小波变换^[3]等. 变换域的算法要对原始音频数据进行变换, 计算量较大, 但往往会获得较好的鲁棒性, 因此变换域算法是目前研究的热点. 其中, 小波域音频水印技术具有良好的多分辨表示, 时频局部分析, 且易于兼容 MPEG4 压缩标准等特点, 引起了人们的高度重视^[4].

本文提出了一种基于小波的鲁棒音频数字水印

收稿日期: 2008-12-24

基金项目: 黑龙江省教育厅科学技术研究项目(10051054)

作者简介: 林克正(1962—), 男, 教授, 博士, E-mail: link@hrbust.edu.cn.

算法,利用小波变换后的低频系数对与大多数攻击比较稳定的特性,改变大于全局参考均值的三个音频帧之间的局部小波低频系数的均值比率关系,嵌入水印比特。

1 基于小波的多尺度分析

1.1 多尺度分析

多尺度分析^[1] (Multi-scale Analysis) 也叫多分辨率分析 (Multi-resolution Analysis) 或多分辨率逼近 (Multi-resolution Approximation)。这一概念的提出是小波分析理论的一大突破,它揭示了函数结构的自相似性,以及小波分析与分形几何 (fractal geometry) 的内在联系。1988年, Mallat 在构造正交小波时提出了多分辨率分析 (或逼近) 概念。它从空间的概念上形象地说明了小波的多分辨率特性,不仅能将以前的各种正交小波基的构造法巧妙地统一起来,而且还给出了构造其他小波基的方法^[1]。

关于多分辨率分析,以一个三层的分解进行说明,其小波分解树如图1所示。分解具有关系: $S = A_1 + D_1 = A_2 + D_2 + D_1 = A_3 + D_3 + D_2 + D_1$ 。如果要进一步分解,则可以把低频部分 A_3 再分解成低频部分 A_4 和低频部分 D_4 ,以下再分解则依此类推^[7-8]。从图1可以看出,多分辨率分析只是对低频部分进行逐步分解,而高频部分则不予考虑。

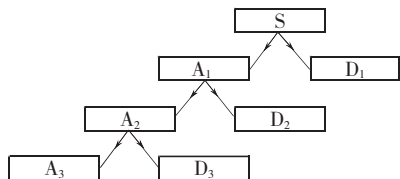


图1 信号的三层多分辨率分解

1.2 小波函数的构造

设尺度向量是满足条件

$$\sum_k h_0(k) = \sqrt{2} \text{ 且 } \sum_k h_0(k) h_0(k+2l) = \delta(l) \quad (1)$$

的一个序列,则存在一个函数,即

$$\Phi(t) = \sum_k h_0(k) \Phi(2t-k) \quad (2)$$

称为尺度函数。

它可以通过自身半尺度 $\Phi(2t-k)$ 复制,即以 $\Phi(2t-k)$ 的加权和来构造,权重为 $h_0(k)$ 。

由 $h_0(k)$ 能够产生函数 $\Phi(t)$,也能产生 $h_1(k)$,并由这两者产生基本小波 $\Psi(t)$ 。事实上,如果有 $h_0(k)$ 和 $\Phi(t)$ 中的任何一个,都能用它来产生另一

个,但通常由 $h_0(k)$ 开始更容易。

一旦有了 $\Phi(t)$ 和 $h_0(k)$,就可以定义如下的一个被称为小波向量的离散高通脉冲响应,即

$$h_1(k) = (-1)^k h_0(-k+1) \quad (3)$$

然后,由此得到基本小波

$$\Psi(t) = \sum_k h_1(k) \Phi(2t-k) \quad (4)$$

再由此得出正交归一小波集

$$\Psi_{j,k}(t) = 2^{j/2} \Psi(2^j t - k) \quad (5)$$

利用正交多分辨率分析以及尺度方程和小波方程的系数,可以使用矩阵形式构造离散数字信号和数字图像小波变换的递推算法,即矩阵金字塔算法或矩阵 Mallat 算法,为小波变换^[1]的应用编程提供一种代数形式的程序结构。

2 水印的嵌入与提取

在嵌入或提取水印之前,先对整个音频进行三层小波分解,计算第三层小波分解的低频系数的绝对平均值 V_g 。然后,将音频段分成固定长度的音频帧,对每个音频帧进行三层小波分解,并计算其第三层小波分解后低频系数绝对平均值 V_f ,选取阈值系数 k ,从中筛选出水印嵌入点 (即 $V_f > k \times V_g$ 的音频帧) 进行嵌入或提取。由于水印嵌入后 V_g 被提升,在水印嵌入之前将 V_g 作为修正值以及头标志嵌入至第一个以及最后一个符合阈值要求的音频帧,从而在提取水印时,该标志既可以作为边界界定,也能作为全局小波分解低频系数的绝对平均值的修正。

2.1 水印空间选取

对原始音频进行三层小波分解,根据其第三层小波分解低频系数计算低频系数绝对平均值 V_g ;然后以固定帧长 N ,将音频分段为若干音频帧,类似地对每个音频帧进行三层小波分解,并计算第三层小波分解低频系数绝对平均值 V_f ;选取阈值系数 k ,用队列 Q 记录下 V_f 和 V_g 满足阈值关系 (即 $V_f > k \times V_g$) 的音频帧的开始位置。重复以上步骤直至队列记录下需要嵌入水印比特及水印头尾标志,即 $\text{Len}(Q) \geq 3 \times \text{Len}(\text{水印}) + 2$ 。

2.2 水印的嵌入

如果位于队列 Q 的第一个或最后一个位置,则从队列中取出该位置,根据其从音频中得到对应音频段,对其进行三层小波分解,调整其第三层小波分解低频系数绝对平均值为 V_g ,然后,根据修正后的第三层低频系数以及其他各层高频及低频系数进行逆小波变换,

更新对应音频中的音频帧完成水印头尾标志的嵌入。

否则从队列 Q 中依次取出三个位置,对应在音频中取出三帧 F_1, F_2, F_3 , 分别对 F_1, F_2, F_3 进行三级小波变换,利用对应的第三层小波分解低频系数 cA_3 , 计算其第三层小波分解低频系数绝对平均值 $V_{F_1}, V_{F_2}, V_{F_3}$ 。

当嵌入水印 $W(i) = 1$ 时,有

$$\left. \begin{aligned} cA_3(F_1) &= cA_3(F_1) \left(1 + \frac{2 \times V_{F_2} - \alpha \times (V_{F_1} + V_{F_3})}{V_{F_1} + V_{F_3}} \right), \text{ if } \frac{V_{F_1} + V_{F_3}}{2V_{F_2}} < 1 \\ cA_3(F_3) &= cA_3(F_3) \left(1 + \frac{2 \times V_{F_2} - \alpha \times (V_{F_1} + V_{F_3})}{V_{F_1} + V_{F_3}} \right), \text{ if } \frac{V_{F_1} + V_{F_3}}{2V_{F_2}} < 1 \\ \text{do nothing,} & \text{ if } \frac{V_{F_1} + V_{F_3}}{2V_{F_2}} \geq 1 \end{aligned} \right\} \quad (6)$$

当嵌入水印 $W(i) = 0$ 时,

$$\left. \begin{aligned} \text{do nothing,} & \text{ if } \frac{V_{F_1} + V_{F_3}}{2V_{F_2}} < 1 \\ cA_3(F_2) &= cA_3(F_2) \left(1 + \frac{(V_{F_1} + V_{F_3}) - 2 \times \alpha \times V_{F_2}}{2V_{F_2}} \right), \text{ if } \frac{V_{F_1} + V_{F_3}}{2V_{F_2}} \geq 1 \end{aligned} \right\} \quad (7)$$

为了加大识别区分度,式(6)、(7)中 $0 < \alpha < 1$ 。根据调整后各帧的第三层小波分解低频系数及各层小波分解高频和低频系数进行逆小波变换,更新原始音频中的对应三帧内各点采样值完成水印比特嵌入。如图2所示为水印嵌入的详细流程。

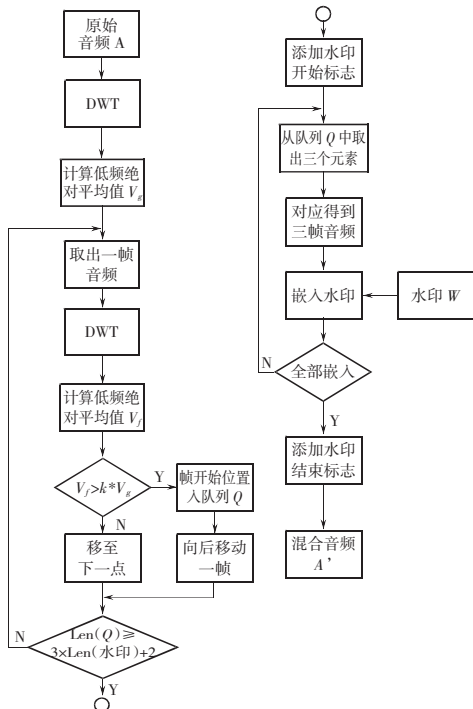


图2 水印嵌入流程

为了提高数字水印的隐藏效果,本文选取小波分解第三层低频系数作为水印的嵌入域,这是因为水印编码可看作是在强背景(原始音频)下叠加1个弱信号(水印)。只要叠加的信号低于对比度门限,听觉系统就无法感觉到信号的存在。而根据 Weber 定律,对比度门限和背景信号的幅值成正比例。由于低频系数的幅值一般远大于高频系数,故低频系数具有比较大的感觉容量。也就是说,在低频系数内嵌入水印信息不会引起原始音频听觉质量的明显改变,从而保证了水印信息的透明性。

2.3 水印信息的提取

水印的提取过程是嵌入的逆过程,与嵌入相对应的水印提取第三层小波分解低频系数绝对平均值的幅值关系即式(8)来提取水印,故有

$$\left. \begin{aligned} w(i) &= 0, \text{ if } \frac{V_{F_1} + V_{F_3}}{2V_{F_2}} < 1 \\ w(i) &= 1, \text{ if } \frac{V_{F_1} + V_{F_3}}{2V_{F_2}} \geq 1 \end{aligned} \right\} \quad (8)$$

3 实验结果及分析

实验采用长度为 7s,分辨率为 32bits,采样频率为 44.1kHz 的一段音频信号。将该音频信号进行三层小波分解,得到如图3所示的低频系数和低频系数。

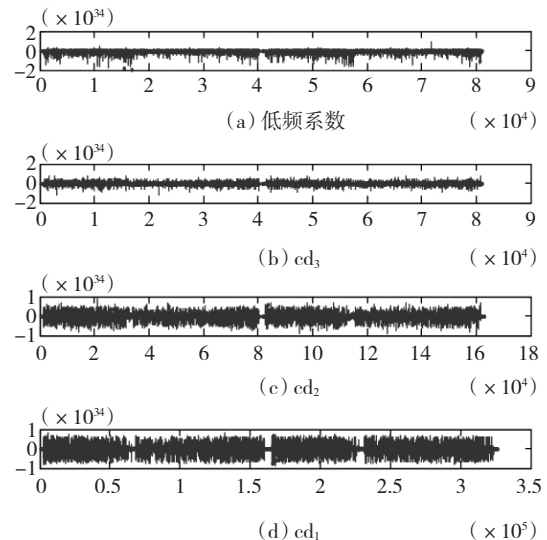


图3 小波系数

从图3中可以看出,低频系数用来刻画信号的整体架构, cd_3, cd_2, cd_1 则表示高频系数,用来刻画信号的细节部分。

3.1 不可见性

选取一段长为 1 024 的随机数序列作为待嵌入水印, $\alpha = 0.4$. 图 4 所示为原始音频、水印及嵌入后的混合音频信号. 发现嵌入水印信息后的音频波形与原音频波形改变很小, 几乎不被察觉. 说明该算法具有很好的不可见性.

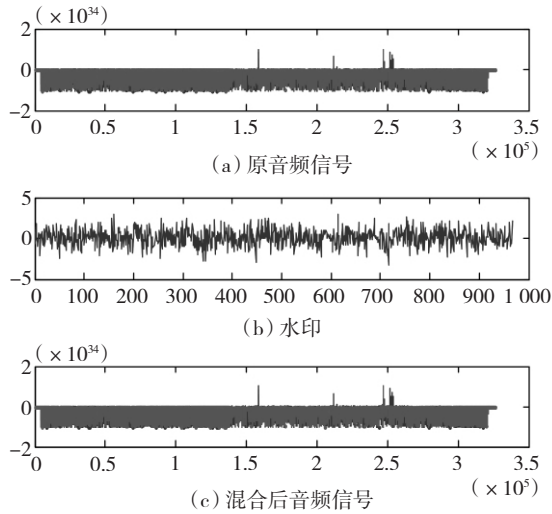


图 4 水印嵌入效果

3.2 抗攻击性

这里选用归一化相关系数 (Normalized Cross-Correlation, NC) 对提出的水印和原始水印的相似性进行定量评价, 其定义为

$$NC(W, W') = \frac{\sum_{i=1}^{M_1} \sum_{j=1}^{M_2} W(i, j) W'(i, j)}{\sqrt{\sum_{i=1}^{M_1} \sum_{j=1}^{M_2} W^2(i, j)} \sqrt{\sum_{i=1}^{M_1} \sum_{j=1}^{M_2} W'^2(i, j)}} \quad (3)$$

在实验中, 小波变换采用 Daubechies-1 小波基进行三级离散小波变换及反变换. 表 1 给出了本文提出水印算法在受到各种攻击时提取水印的相关结果与文 [10] 提取结果的比较.

在鲁棒测试中, 对嵌入水印的混合音频进行了如下攻击:

- 1) 重新量化 (Re-quantization): 将音频信号从 16b 量化为 8b, 然后再量化 16b.
- 2) 重新采样 (Re-sampling): 将音频信号采样频率降为 22.05kHz 和 11.025kHz, 然后提升至 44.1kHz.
- 3) 叠加噪声 (Additive noise): 加入能量为信号能量 10% 的白噪声.
- 4) Mp3 压缩 (Mp3 compression): 用不同比特率 (112kbit/s, 64kbit/s, 48kbit/s, 32kbit/s) 对音频进行

压缩、解压缩处理.

表 1 受攻击后水印提取结果性能比较

攻击类型	NC	
	文 [10] 算法	本文算法
原始水印	1.000 0	1.000 0
重新量化	1.000 0	1.000 0
重采样 22.05kHz	0.408 7	0.993 8
重采样 11.025kHz	0.361 8	0.783 5
叠加噪声	0.961 9	1.000 0
随机剪切 10%	1.000 0	1.000 0
Mp3-112k	1.000 0	1.000 0
Mp3-64k	0.392 5	0.995 6
Mp3-48k	0.383 8	0.949 2
Mp3-32k	0.372 2	0.809 8

实验表明, 该算法在保证水印不可感知性的同时提高了水印抵抗攻击的能力.

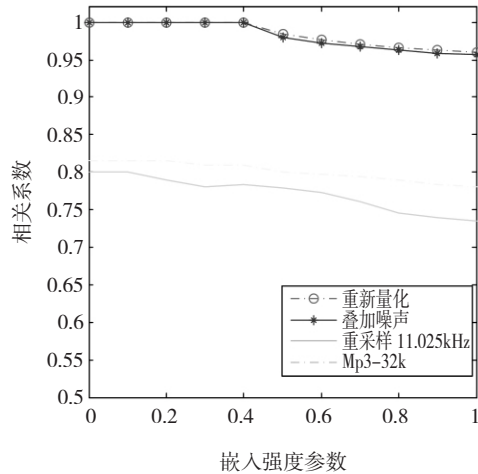


图 5 α 为不同值的相关系数

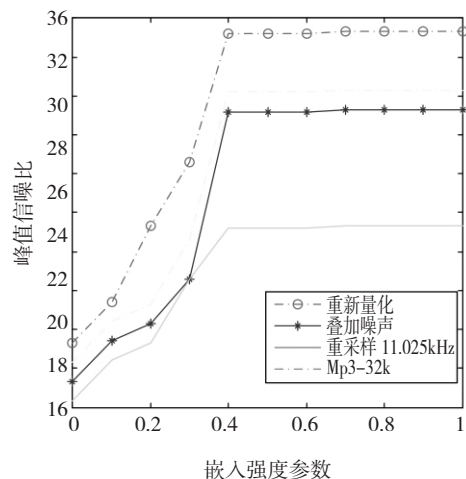


图 6 α 为不同值的峰值信噪比

由图5,图6可以看出随着嵌入强度的逐渐加大,相关系数呈减弱趋势,但关联程度仍然较大.峰值信噪比随着强度的增强逐渐升高,说明该算法有很好的鲁棒性.

4 结 语

目前,音频水印的算法有很多,但是抗同步攻击的算法较少.尤其是随机裁剪对于音频水印算法的鲁棒性,更是一个挑战,而对于不降低音频听觉效果的随机裁剪,不会对小波低频系数绝对平均值与全局小波低频系数绝对平均值接近的音频帧进行裁剪,因而本算法中选取帧内小波变换低频系数均值接近全局参考均值的多个音频帧作为水印嵌入点的同步策略,具有较强的鲁棒性.同时,无需增加额外的同步标志信息,嵌入点位置不固定,需计算之后得出,从而增加了水印信息的透明性.实验表明,本文所提出的算法不仅具有较好的不可感知性,而且对诸如重新采样、重新量化、叠加噪声、Mp3压缩等攻击均具有较好的鲁棒性.同时,本文提出算法无需原始载体音频,属于盲水印算法,增强了其实用性.

参 考 文 献:

- [1] 林克正,师 晶,唐远新. 基于视频帧的彩色水印嵌入算法研究 [J]. 哈尔滨理工大学学报, 2009, 14(6): 22-29.
- [2] 王永琦,杨 洋. 基于听觉特性的 DCT 域数字音频水印算法 [J]. 计算机工程与应用, 2008, 44(3): 88-90.
- [3] 赵 红,崔永瑞,郑素娥. 小波域音频水印综述 [J]. 计算机安全, 2007, 13(9): 1-3.
- [4] 师 晶,林克正,秦丽萍. 基于 DWT 和 SVD 的改进分块图像水印算法 [J]. 哈尔滨理工大学学报, 2009, 14(5): 35-38.
- [5] 施成湘,杨 丹,尚 晋,等. 扩展的多尺度模糊边缘检测 [J]. 计算机工程与应用, 2006, 23(7): 65-68.
- [6] SERDEAN C V. DWT-based High-capacity Blind Video Watermarking Invariant to Geometrical Attack [J]. IEEE Proc - Vis Image Signal Process, 2005, 150(1): 51-58.
- [7] WANG Xiangyang, NIU Panpan. A Robust, Digital-Audio Watermarking Method [J]. IEEE Computer Society, 2009, 16(3): 60-69.
- [8] 白红哲,王 亮,孙晓岩. 盲音频数字水印算法 [J]. 哈尔滨理工大学学报, 2005, 10(3): 117-120.
- [9] 雷 赞,刘 建,严 波,等. 窄带的自同步音频水印算法 [J]. 计算机学报, 2008, 31(7): 1283-1290.
- [10] HUANG Ji-Wu, WANG Yong, Shi Y. Q. A Blind Audio Watermarking Algorithm with Self Synchronization [J]. Proceeding of IEEE International Symposium on Circuits and System, 2002(2): 627-630.

(编辑: 付长缨)

(上接第80页)

- [1] CICERONE S, DI Felice P. Cardinal Relations Between Regions with a Broad Boundary [C] // 8th ACM Symposium on Advances in Geographic Information Systems (GIS'00). New York: ACM Press, 2000: 15-20.
- [2] DU Shihong, WANG Qiao, YANG Yipeng, et al. Fuzzy Description of Spatial Direction Relations [J]. Journal of Computer-Aided Design & Computer Graphics, 2005, 8: 1744-1751.
- [3] GAU W L, BUEHRER D J. Vague Sets [J]. IEEE Transactions on System Man and Cybernetics, 1993, 23: 610-614.
- [4] PAPADIAS D, THEODORIDI Y, STELL T. The Retrieval of Di-

rection Relations Using R-tree [C] // Database and Expert Systems Applications. Athen, Greece, 1994: 78-85.

- [5] 刘永山,郝忠孝. 基于 MBR 的主方向关系一致性检验 [J]. 软件学报, 2006, 17(5): 976-982.
- [6] 杨 宁,毛宇光. 基于 Vague 集的模糊数据库研究 [D]. 南京航空航天大学, 2005.

(编辑: 付长缨)