

一种 DWT 与 DCT 结合的盲音频水印算法¹

赵春晖 李福昌

(哈尔滨工程大学信息与通信工程学院 哈尔滨 150001)

摘 要 该文提出了一种新的盲音频水印算法。它充分利用离散小波变换的多分辨率特性和离散余弦变换的能量压缩能力,通过量化处理把降维后的二值图像水印嵌入到原始的音频信号中,在水印的提取过程中不需要原始音频信号的参与。仿真实验证明了算法的鲁棒性和不可感知性。

关键词 数字水印, 离散小波变换, 离散余弦变换, 量化

中图分类号 TN911.73, TP391.4

1 引言

多媒体技术和网络技术的发展,显著地提高了信息传播的效率和准确度。但随之而来的副作用是,有恶意的个人和团体有可能在没有得到作品所有者许可的情况下,随意复制、篡改、销售有版权的内容。因此,如何实施信息交换中的版权保护和信息安全措施,已成为数字通信中亟待解决的问题之一。

数字水印技术被认为是解决版权问题的一种有效方式。它充分利用人类视觉和听觉系统的冗余(即视觉和听觉分辨率的局限性),嵌入与版权所有者相关的信息,以证实数字产品的版权归属^[1]。从应用范围上,可把水印分为静态图像水印、视频水印和音频水印。目前,关于静态图像水印和视频水印的研究很多,而对于音频水印的研究却鲜有报道。但随着 mp3 等新一代压缩标准的广泛应用,对音频数据产品的保护就显得越来越重要。

音频水印算法的实现有两种方法:时域的方法和变换域的方法。早期的水印算法通过修改原音频信号中最不重要的部分,以达到嵌入水印的目的。Won-Gyum Kim 提出了在数字音频信号中直接嵌入数字印章的方法^[2],该方法在水印提取过程中不需要原始音频信号的参与。但这有个缺点:水印信号是与最不重要分量捆绑在一起,就无法抵抗一些常见信号处理操作。在文献^[3]中,作者利用 FFT 计算其频谱,求其频率遮掩和时间遮掩函数,再变换回时域嵌入水印,但算法复杂,计算量大。王秋生等提出了基于分段离散余弦变换的数字音频水印算法^[4],但在水印的提取过程中需要原始音频信号的参与。文献^[5]提出了基于小波变换的音频水印算法。同样,在水印的提取过程中需要原始音频信号的参与,不便于实际工程应用。文献^[6]提出了基于离散余弦变换的音频水印算法,但由于嵌入的水印信息是高斯白噪声,水印信息不具有直观的特点。

本文利用小波变换的多分辨率特性和余弦变换的解相关能力和能量压缩能力(聚能作用),以直观、丰富的二值图像作为水印信号,给出一种新的音频数字水印算法。由于在水印的嵌入过程中采用了量化方案,在水印的提取过程中就不需要原始音频信号的参与。仿真实验证明了算法的稳健性和不可感知性。

2 水印信号的嵌入

设 W 是水印,它是一幅大小为 $M_1 \times M_2$ 的二值图像,可表示为

$$W = \{w(i, j), 0 \leq i \leq M_1 - 1, 0 \leq j \leq M_2 - 1\} \quad (1)$$

其中 $w(i, j) \in \{0, 1\}$ 为二值图像的像素值。

¹ 2002-06-24 收到, 2002-12-19 改回

哈尔滨市后备学科带头人基金(0171218025), 全国优秀博士学位论文作者专项基金(200037)和教育部优秀青年教师资助计划项目, 及高等学校优秀青年教师教学科研奖励计划资助

A 表示原始的数字音频信号, 其数据个数为 K , 它可表示为

$$A = \{a(k), 0 \leq k \leq K-1\} \quad (2)$$

其中 $a(k) \in \{0, 1, 2, \dots, (2^p - 1)\}$ 是第 k 个数据的幅度值, p 为表示每个数据所用的比特数。

由于水印是作为噪声嵌入到原始数字音频信号中的, 因此水印的嵌入必须以不影响数字音频信号的使用价值为前提。一般地, 嵌入一个水印信息至少需要 N 个音频数据。要嵌入水印的全部 $M_1 \times M_2$ 个像素, 音频数据段的长度必须满足

$$K \geq M_1 \times M_2 \times N \quad (3)$$

水印信号的具体嵌入步骤如下:

(1) 降维操作 若想将二值图像嵌入到一维的数字音频信号中, 必须先对其进行降维操作, 以使二维的二值图像转化成一维的数字序列。其操作方法如下:

$$V = \{v(l) = w(i, j), 0 \leq i \leq M_1 - 1, 0 \leq j \leq M_2 - 1, l = i \times M_1 + j\} \quad (4)$$

(2) 数据划分 为讨论方便, 将 (4) 式中的原始数字音频信号分成两部分:

$$A = A_e + A_r \quad (5)$$

其中 A_e 是与嵌入水印相关的部分, A_r 是与嵌入水印无关的部分。根据 (3) 式, 与嵌入水印相关部分的长度为 $M_1 \times M_2 \times N$ 。

(3) 离散小波变换 对与水印嵌入有关的音频数据部分 A_e , 选择合适的小波基, 作二级离散小波分解。

$$D_e = \text{DWT}(A_e) = C_e(2) \oplus D_e(2) \oplus D_e(1) \quad (6)$$

其中 $C_e(2)$, $D_e(2)$ 是二级小波分解的近似分量和细节分量, $D_e(1)$ 是一级小波分解的细节分量。

(4) 离散余弦变换 由于小波分解的近似分量能有效地抵御各种干扰, 这里选择二级小波分解的近似分量 $C_e(2)$ 进行离散余弦变换, 然后嵌入水印。设近似分量长度为 L_1 , 则

$$C_{e,\text{dct}}(2) = \text{DCT}(C_e(2)) = \{c_{e,\text{dct}}(2)(n), 0 \leq n \leq L_1 - 1\} \quad (7)$$

(5) 水印信号嵌入 通常把水印信号嵌入在离散余弦变换的低中频系数上, 这样做的好处是低中频系数通常具有较大的数值, 水印的嵌入对其影响不大。

水印的嵌入是通过对选定的变换域系数进行量化来完成的。设这里需要进行量化的系数为 f , 量化间隔为 ρ , 水印比特为 $v(l)$, 把经过量化处理后得到的系数记为 f' 。

通过量化处理嵌入水印信号的思想为: 按图 1 把系数分为两类 A 和 B :

$$A = \{2p\rho < f \leq (2p+1)\rho\} \quad (8a)$$

$$B = \{(2p+1)\rho < f \leq (2p+2)\rho\}, p = \dots, -2, -1, 0, 1, 2, \dots, \quad (8b)$$

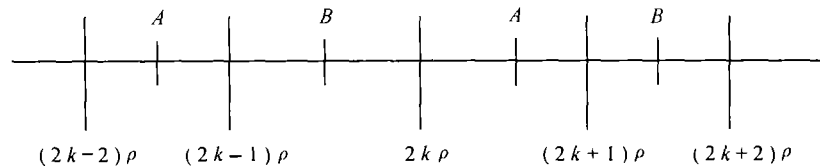


图 1 通过量化处理嵌入水印信号示意图

当 $v(l)$ 为 1 时, f 量化为与之最接近的 A 类的中点; 而当 $v(l)$ 为 0 时, f 量化为与之最接近的 B 类的中点.

设水印信号在数据段 $C_{e,\text{dct}}(2)$ 中开始嵌入的位置为 t_2 , 则嵌入过程为

当 $n = t_2 + l (l = 0, \dots, M_1 \times M_2 - 1)$ 时, 根据 $v(l)$ 的值为 1 或 0, 按图 1 的方法嵌入水印信号; 当 n 为其他值时 $c_{e,\text{dct}}(2)(n)$ 保持不变. 设嵌入水印信号后 $c_{e,\text{dct}}(2)(n)$ 变为 $c'_{e,\text{dct}}(2)(n)$, 则嵌入水印后的数据段 $C_{e,\text{dct}}(2)$ 变成

$$C'_{e,\text{dct}}(2) = \{c'_{e,\text{dct}}(2)(n), 0 \leq n \leq L_1 - 1\} \quad (9)$$

(6) 离散余弦反变换 对嵌入水印后的数据段 $C'_{e,\text{dct}}(2)$ 做离散余弦反变换:

$$C''_e(2) = \text{IDCT}(C'_{e,\text{dct}}(2)) = \{c''_e(2)(n), 0 \leq n \leq L_1 - 1\} \quad (10)$$

(7) 离散小波反变换 以 $C''_e(2)$ 代替 $C_e(2)$ 代入 D_e , 得到嵌入水印后的小波变换域表示式为 D'_e :

$$D'_e = C''_e(2) \oplus D_e(2) \oplus D_e(1) \quad (11)$$

再做离散小波反变换, 就得到时域中含有水印信号的音频信号 A'_e :

$$A'_e = \text{IDWT}(D'_e) \quad (12)$$

将 A'_e 代替 A_e , 最终得到含水印的数字音频信号:

$$A_w = A'_e + A_r \quad (13)$$

3 水印信号的提取

设 A_s 是待检测的数字音频信号, 水印的提取过程可描述如下:

(1) 对待检测的数字音频信号 A_s 作分段处理, 即 $A_s = A_{se} + A_{sr}$, 然后对含水印部分 A_{se} 作二级离散小波分解, 即

$$D_{se} = \text{DWT}(A_{se}) = C_{se}(2) \oplus D_{se}(2) \oplus D_{se}(1) \quad (14)$$

(2) 离散余弦变换: 选择二级离散小波分解的近似分量, 对其进行离散余弦变换. 根据嵌入算法, 近似分量长度为 L_1 , 则

$$C_{se,\text{dct}}(2) = \text{DCT}(C_{se}(2)) = \{c_{se,\text{dct}}(2)(n), 0 \leq n \leq L_1 - 1\} \quad (15)$$

(3) 根据水印嵌入位置和量化方案, 提取水印序列 $\{v_1(l)\}$, 即

$$v_1(l) = \begin{cases} 1, & c_{se,\text{dct}}(2)(t_2 + l) \in A \\ 0, & c_{se,\text{dct}}(2)(t_2 + l) \in B \end{cases} \quad (16)$$

(4) 对 $\{v_1(l)\}$ 升维处理, 将一维的序列变为二维的二值图像

$$W_s = \{w_s(i, j) = v_1(l), 0 \leq i \leq M_1 - 1, 0 \leq j \leq M_2 - 1, l = i \times M_1 + j\} \quad (17)$$

为消除主观因素的影响, 体现版权保护的公正性, 采用归一化相关系数对原始水印和提取水印的相似性作定量评价. 水印含有与否的判断标准为若 $NC(W, W_s) > T$, 可以判定被测音频

信号中含有水印; 否则不含有水印。一般, T 在 $0 \sim 1$ 之间取值。 T 值小, 虚警概率提高而漏警概率降低; T 值大, 则反之。折中考虑, 本文取 $T = 0.5$ 。归一化相关系数被定义为

$$NC(W, W_s) = \frac{\sum_{i=0}^{M_1-1} \sum_{j=0}^{M_2-1} w(i, j) w_s(i, j)}{\sqrt{\sum_{i=0}^{M_1-1} \sum_{j=0}^{M_2-1} w(i, j)^2} \times \sqrt{\sum_{i=0}^{M_1-1} \sum_{j=0}^{M_2-1} w_s(i, j)^2}} \quad (18)$$

4 仿真实验

实验中选取水印信息为 64×64 二值图像, 如图 3(a) 所示。首先将原始音频信号分成两部分; 其次将用于嵌入水印的部分做二级离散小波分解, 小波基取 'db4', 对二级小波分解的近似分量进行离散余弦变换; 最后将水印序列通过量化处理后嵌入到原始音频信号中, 嵌入过程中取量化间隔 $\rho = 0.035$ 。图 2(a) 是原始的音频信号, 为单声道、22.05kHz 采样率、8bit 量化, 长度为 4096×8 , 这样根据 (3) 式可确定 $N = 8$; 图 2(b) 是含水印的音频信号, 嵌入水印后的信噪比 SNR 为 39.1859, 和原始音频信号听起来几乎没有差别。

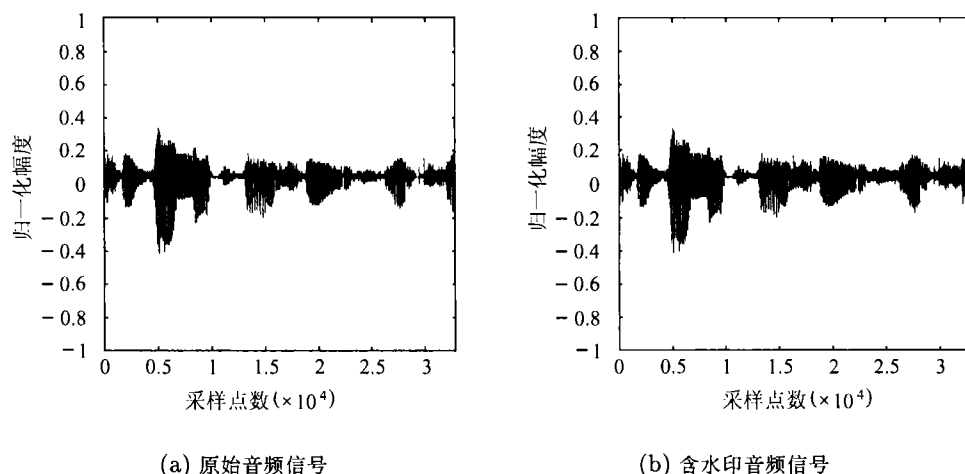


图 2

为检测算法的稳健性, 对含水印音频信号分别进行下面的处理: (1) 加入高斯白噪声 (均值为 0, 均方差为 0.01); (2) 低通滤波。取长度为 9, 截止频率为 4kHz 的契比雪夫低通滤波器; (3) 有色噪声干扰; (4) 重新采样。分别对信号进行一次抽取和一次插值处理, 抽取和插值的系数为 2。

在这几种情况下提取的水印分别如图 3(b), 3(c), 3(d), 3(e) 所示, 其归一化相关系数分别为 0.8163, 0.9076, 0.8278, 0.5623。从未经任何处理的含水印音频信号中提取的水印如图 3(f) 所示, 其相关系数为 1.0000。

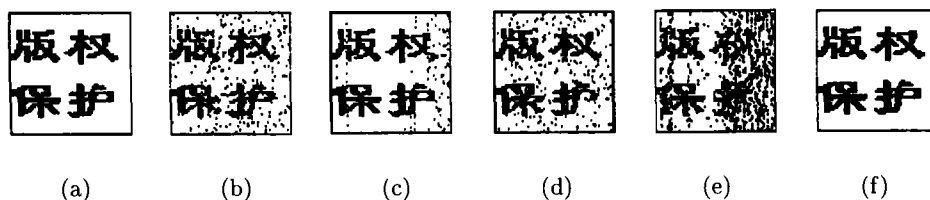


图 3 水印图像

在相同的条件(原始音频信号数据长度为 4096×8 , 嵌入水印后信噪比 SNR 为 39.1859)下, 我们进行 3 组仿真对比实验: (1) 单独采用离散余弦变换; (2) 单独采用离散小波变换; (3) 采用离散余弦变换和小波变换相结合; 并求出这 3 种算法在前面所述的 4 种水印攻击下的归一化相关系数值, 比较结果如表 1 所示。

从表 1 可知, 在原始音频信号数据长度和嵌入水印信号后原媒体信号被影响程度(用信噪比表征)相同的情况下, 本文提出的算法比单独利用余弦变换或小波变换提取水印的归一化相关系数大, 具有更好的抵抗水印攻击的能力。

表 1 3 种水印算法在相同条件下提取的归一化相关系数值的比较

算法	白噪声干扰	低通滤波	色噪声干扰	重采样
基于余弦变换的算法	0.7628	0.8321	0.7951	0.4966
基于小波变换的算法	0.5125	0.9058	0.5348	0.5615
基于两种变换结合的算法	0.8163	0.9076	0.8278	0.5623

5 结 论

本文研究了把离散余弦变换与离散小波变换相结合的水印算法。充分利用小波变换的多分辨率特性及离散余弦变换的能量压缩能力和解相关能力, 可提高水印嵌入的鲁棒性和不可感知性。由于在水印嵌入的过程中采用了量化方案, 提取水印信号时不需要原始音频信号的参与, 便于实际工程应用。仿真实验证明本文提出的算法能有效地抵抗常见的白噪声污染、色噪声污染、低通滤波、重采样等对水印的攻击。

参 考 文 献

- [1] J. Zhao, E. Koch, In business today and tomorrow, Communications of the ACM, 1998, 41(7), 67-72.
- [2] Won-Gyum Kim, Jong Chan Lee, Won Don Lee, An audio watermarking scheme with hidden signatures, International Conference on Signal Processing, Aug. 21-24, 2000, Beijing, China, 2000, 250-253.
- [3] D. Swanson, B. Zhu, H. Tewfik, L. Boney, Robust audio watermarking using perceptual masking, Signal Processing, 1998, 66(6), 337-355.
- [4] 王秋生, 孙圣和, 一种在数字音频信号中嵌入水印的新算法, 声学学报, 2001, 26(5), 464-467.
- [5] 钮心忻, 杨义先, 基于小波变换的数字水印隐藏与检测算法, 计算机学报, 2000, 23(1), 21-27.
- [6] 陈琦, 王炳锡, 一种基于 DCT 变换的语音数字水印算法研究, 信号处理, 2001, 17(3), 238-241.

A BLIND AUDIO WATERMARK ALGORITHM BASED ON DWT AND DCT

Zhao Chunhui Li Fuchang

(College of Info. & Communication Eng., Harbin Engineering Univ., Harbin 150001, China)

Abstract This paper presents a novel blind audio digital watermark algorithm. It makes full use of the multi-resolution of DWT and the energy compression of DCT and embeds binary image watermark in the audio by quantification process. The watermark can be extracted without the original audio signal. Experimental results demonstrate the algorithm is robust and imperceptible.

Key words Digital watermark, DWT, DCT, Quantification

赵春晖: 男, 1965 年生, 教授, 博士生导师, 主要研究方向为信号与图像处理, 非线性滤波等。
李福昌: 男, 1977 年生, 博士生, 主要研究方向为图像处理, 数字水印技术。