

一种可抵抗 MP3 压缩的音频水印算法

王向阳^{1,2} 杨红颖¹ 赵红¹

摘要 提出了一种稳健的自适应混合域数字音频水印算法, 该算法具有以下特点: 1) 通过采纳更为稳健的同步信号及其全新嵌入策略, 提高了音频水印的抗攻击能力; 2) 充分利用离散小波变换 (Discrete wavelet transform, DWT) 的多分辨率特性和离散余弦变换 (Discrete cosine transform, DCT) 的能量压缩特性, 改善了数字水印的隐藏效果; 3) 结合听觉掩蔽特性自适应确定量化步长, 提高了数字水印的不可感知性; 4) 以保证不可感知性和稳健性的良好平衡为前提, 实现了数字水印的盲检测. 仿真实验表明, 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、低通滤波、随机剪切等攻击均具有较好的稳健性. 特别地, 该算法对于最为普通的 MP3 压缩攻击具有极强的抵抗能力 (尤其是高压比下).

关键词 数字水印, 数字音频, MP3 压缩, 量化步长

中图分类号 TP391

A Robust Adaptive Digital Audio Watermarking Scheme Against MP3 Compression

WANG Xiang-Yang^{1,2} YANG Hong-Ying¹ ZHAO Hong¹

Abstract In this paper, a robust adaptive digital audio watermarking scheme in the hybrid domain is presented. The features of the proposed algorithm are as follows: 1) More robust synchronization codes and a new embedded strategy are adopted to resist the attack effectively. 2) The multi-resolution characteristics of discrete wavelet transform (DWT) and the energy-compression characteristics of discrete cosine transform (DCT) are combined to improve the transparency of digital watermark. 3) The adaptive quantization steps are identified according to human auditory masking. 4) The algorithm can extract the watermark without the help of the original digital audio signal. Experimental results show that the proposed watermarking algorithm is imperceptible and robust against various signal processings such as re-sampling, re-quantifying, noising adding, low pass filtering and cropping. Especially, our audio watermarking scheme is remarkably effective in resisting MP3 compression.

Key words Digital watermarking, digital audio, MP3 compression, quantization step

1 引言

现有数字音频水印算法基本上可以被归纳为两

类^[1], 即空间域方法和变换域方法. 其中, 变换域音频水印方法因其具有较好的稳健性而得到了广泛使用. 为了发挥 DCT 域的能量压缩特性 (聚能作用), 文献 [2] 提出将水印信号嵌入到 DCT 域的低频系数上, 具有较好的稳健性, 但由于采用固定量化的方法嵌入水印, 所以其不可感知性有待进一步改善. 另外, 小波域音频水印技术因其具有良好的多分辨率表示、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^[3~7], 但现有小波域音频水印算法普遍具有抵抗有损压缩能力不强的弱点 (特别是高压比下). 文献 [8] 提出了一种稳健数字音频水印算法, 该算法能够结合 MP3 压缩对原始音频的影响嵌入水印信息, 具有一定创新且提供了较强的抗有损压缩能力, 但其存在如下不足: 1) 提取水印信息时需要原始音频信号的参与, 不利于实际应用; 2) 算法共涉及四个重要阈值 (p, ϵ, q, d), 且均需要结合实验进行取值, 故通用性极差; 3) 所采纳的固定嵌入深度, 影响了数字水印的不可感知性和稳健性; 4) 抗随机剪切等去同步攻击能力较弱.

收稿日期 2005-9-20 收修改稿日期 2006-6-23

Received September 20, 2005; in revised form June 23, 2006

辽宁省自然科学基金 (20032100)、视觉与听觉信息处理国家重点实验室 (北京大学) 开放基金 (0503)、大连市科技基金 (2006J23JH020)、江苏省计算机信息处理技术重点实验室 (苏州大学) 开放课题基金 (KJS0602), 和“图像处理与图像通信”江苏省重点实验室 (南京邮电大学) 开放基金 (ZK205014) 项目资助

Support by the Natural Science Foundation of Liaoning Province of P.R.China (20032100), the Open Foundation of State Key Laboratory of Vision and Auditory Information Processing (Beijing University, 0503), the Natural Science Foundation of Dalian City of P.R.China (2006J23JH020), the Open Foundation of Jiangsu Province Key Laboratory for Computer Information Processing Technology (Soochow University, KJS0602), the Open Foundation of Key Laboratory of Image Processing and Image Communication (Nanjing University of Posts and Communications, ZK205014)

1. 辽宁师范大学计算机与信息技术学院 大连 116029 2. 北京大学视觉与听觉信息处理国家重点实验室 北京 100871

1. School of Computer and Information Technique, Liaoning Normal University, Dalian 116029 2. State Key Laboratory of Vision and Auditory Information Processing, Beijing University, Beijing 100871

DOI: 10.1360/aas-007-0248

本文提出了一种新的自适应混合域数字音频盲水印算法, 该算法选取更为稳健的 16 位巴克码作为同步标记, 并通过修改多个采样值的统计均值嵌入同步码, 同时结合听觉掩蔽特性采用自适应量化策略将水印信息在混合域内嵌入到数字音频中.

2 数字水印的嵌入

本文所采用的数字水印嵌入策略为: 在空间域内嵌入同步信号, 在混合域 (DCT 和 DWT) 内嵌入水印信息. 假设原始数字音频信号为 $A = \{a(i), 0 \leq i < Length\}$. 其中, $Length$ 为音频数据的个数, $a(i) \in \{0, 1, 2, \dots, (2^p - 1)\}$ 是第 i 个音频数据的幅度值, p 为表示每个数据所使用的比特数. 假设二值水印图像为 $W = \{w(i, j), 0 \leq i < M, 0 \leq j < N\}$. 其中, $w(i, j) \in \{0, 1\}$ 代表二值水印图像的第 i 行、第 j 列像素值. 假设同步码为 $F = \{f(i), 0 \leq i < Lsyn\}$, 其中, $f(i) \in \{0, 1\}$ (为了便于使用, 这里已对原巴克码进行简单变换, 即 $+1 \rightarrow 0, -1 \rightarrow 1$), $Lsyn$ 为同步码的长度. 则自适应混合域数字音频盲水印嵌入过程 (关键步骤) 可描述如下.

2.1 预处理

为了消除二值水印图像的像素空间相关性, 提高整个数字水印系统的安全性能, 确保数字音频某一部分受到破坏后仍能全部或部分地恢复水印, 宜首先对二值水印图像进行置乱加密. 为此, 本文首先引入 Arnold 变换, 将二值水印图像 W 置乱加密为安全水印矩阵 W_1 , 其中 $W_1 = \{w_1(i, j), 0 \leq i < M, 0 \leq j < N\}$.

另外, 为进一步提高抵抗剪切、平移等去同步攻击能力, 以使得数字音频部分数据丢失后仍能保证水印检测的有效可靠, 本文算法首先对原始数字音频信号进行分段处理, 然后再将每个音频数据段分割成长度为 L_1 和 L_2 的两部分并分别嵌入同步码与水印信息. 其中, 第 i 个音频数据段可以表示为

$$A(i) = \{a(iL + k), 0 \leq k < L\} (0 \leq i < \lfloor Length/L \rfloor)$$

这里, $L = L_1 + L_2$, $L_1 = Lsyn \times n$, $L_2 = M \times N \times 2^{H+1}$, H 为小波变换级数, n 为常数 (本文选取为 5).

设某个音频数据段为 A^0 (其长度为 L_1 和 L_2 的两部分为 A_1^0 和 A_2^0), 则同步码与数字水印信息的嵌入过程如下.

2.2 同步码的嵌入

为了保证不可感知性与稳健性的良好平衡, 本文采用修改多个采样值 (n 个采样值) 的方法, 将同步信息嵌入到采样值的统计均值中, 具体步骤为:

1) 将 A_1^0 按同步码长度 $Lsyn$ 分成 $Lsyn$ 段, 每一段 $PA_1^0(m)$ 含有 n 个采样点, 即

$$PA_1^0(m) = \{pa_1^0(m)(i) = a_1^0(i + m \times n), 0 \leq i < n, 0 \leq m < Lsyn\}$$

2) 计算 $PA_1^0(m)$ 的平均值, 即

$$\overline{PA_1^0(m)} = \frac{1}{n} \sum_{i=0}^{n-1} pa_1^0(m)(i), (0 \leq m < Lsyn)$$

3) 采用量化方法嵌入同步码, 即对每一段 $PA_1^0(m)$, 修改其均值 $\overline{PA_1^0(m)}$, 以嵌入一位同步码. 修改策略为

$$pa_1^{0'}(m)(i) = pa_1^0(m)(i) + (\overline{PA_1^{0'}(m)} - \overline{PA_1^0(m)})$$

其中, $PA_1^0(m) = \{pa_1^0(m)(i), 0 \leq i < n\}$ 为修改前采样值, $PA_1^{0'} = \{pa_1^{0'}(m)(i), 0 \leq i < n\}$ 为修改后的采样值, 且有

$$\overline{PA_1^{0'}(m)} = \begin{cases} IQ(\overline{PA_1^0(m)}) \times S_1 + s_{1/2}, & \text{如果 } Q(\overline{PA_1^0(m)}) = f(m) \\ IQ(\overline{PA_1^0(m)}) \times S_1 - s_{1/2}, & \text{如果 } Q(\overline{PA_1^0(m)}) \neq f(m) \end{cases}$$

$$IQ(\overline{PA_1^0(m)}) = \left\lfloor \frac{\overline{PA_1^0(m)}}{S_1} \right\rfloor$$

$$Q(\overline{PA_1^0(m)}) = \text{mod}(IQ(\overline{PA_1^0(m)}), 2)$$

其中, mod 为取模运算, S_1 为量化步长.

2.3 水印信号的嵌入

1) 音频数据段 (后面部分) 的 DWT. 对每个音频数据段 A_2^0 进行 H 级小波变换, 并得到小波系数 $A_2^{0H}, D_2^{0H}, D_2^{0H-1}, \dots, D_2^{01}$. 其中, A_2^{0H} 为音频数据段 A_2^0 小波变换的第 H 级近似分量, $D_2^{01}, \dots, D_2^{0H-1}, D_2^{0H}$ 分别为音频数据段 A_2^0 小波变换的第 $1 \sim H$ 级的细节分量.

2) 近似分量的 DCT. 考虑到小波变换域的近似分量能够有效抵御各种干扰, 这里选择 H 级小波变换的近似分量 A_2^{0H} 进行 DCT 及嵌入水印.

$$A_2^{0HC} = DCT(A_2^{0H}) = \{a_2^0(t)^{HC}, 0 \leq t < L_2/2^H\}$$

3) 水印信号的嵌入. 为保证音频水印的稳健性及不可感知性, 本文根据人类听觉系统的频域掩蔽特性, 采用量化策略将水印信息嵌入到 DCT 域的低中频部分, 即量化调制前 $M \times N$ 个 DCT 系数, 得到 $A_2^{0'HC}$, 其中

$$A_2^{0'HC} = \begin{cases} a_2^{0'}(t)^{HC}, & 0 \leq t < M \times N \\ a_2^0(t)^{HC}, & M \times N \leq t < L_2/2^H \end{cases}$$

量化规则如下:

$$a_2^{o'}(t)^{HC} = \begin{cases} IQ(a_2^0(t)^{HC}) \times S_2 + S_2/2, & \text{如果 } Q(a_2^0(t)^{HC}) = w_1(i, j) \\ IQ(a_2^0(t)^{HC}) \times S_2 - S_2/2, & \text{如果 } Q(a_2^0(t)^{HC}) \neq w_1(i, j) \end{cases}$$

$$t = (i - 1) \times N + j$$

其中, $0 \leq i < M$, $0 \leq j < N$, S_2 为量化步长.

$$IQ(a_2^0(t)^{HC}) = \lfloor (a_2^0(t)^{HC})/S_2 \rfloor$$

$$Q(a_2^0(t)^{HC}) = \text{mod}(IQ(a_2^0(t)^{HC}), 2)$$

这里, mod 是取模运算.

4) 逆 DCT. 对嵌入水印后的音频数据段 $A_2^{o'HC}$ 做逆 DCT.

$$A_2^{o'H} = IDCT(A_2^{o'HC})$$

5) 逆 DWT. 以 $A_2^{o'H}$ 代替 $A_2^{o'H}$ 并进行 H 级逆 DWT, 得到含有水印的音频数据段信号 $A_2^{o'}$.

3 量化步长的自适应选取

对基于量化的音频水印嵌入方案来说, 量化步长的选取至关重要. 因为量化步长与水印嵌入强度密切相关, 量化步长取值越大, 数字水印不可感知性能越好 (但同时也更容易给音频引入失真). 选取确定量化步长应充分考虑数字音频自身特点和视觉掩蔽特性. 然而, 现有音频水印嵌入方案^[4] 普遍采用了均匀量化策略, 即对整个数字音频采用一个相同的量化步长. 这样, 如果量化步长取值较小, 则会影响数字水印的抗攻击性; 而如果量化步长取值较大, 则会给数字音频质量带来较大影响. 另外, 不同的载体音频, 只有采纳不同的量化步长值才能各自达到比较好的隐藏效果.

为了有效克服均匀量化策略所存在的不足, 同时兼顾算法的实际应用 (若给每个嵌入点都确定一个量化步长, 计算量太大), 本文将在时域内结合整个数字音频内容确定量化步长 S_1 , 在频域内结合音频数据段内容确定量化步长 S_2 , 即

首先, 根据时域掩蔽特性确定出量化步长 S_1 (用于同步码嵌入), 有

$$S_1 = 2^{\lfloor \log_2 s_a \rfloor} \quad (\lfloor \cdot \rfloor \text{ 为向下取整函数}), \quad s_a = \overline{|A|}$$

其中, $s_a = \overline{|A|}$ 表示原始数字音频 A 的绝对值均值 (即峰值的均值).

然后, 结合频域掩蔽特性自适应确定量化步长 S_2 (用于水印信息嵌入), 具体步骤为

1) 音频数据段 A_2^0 (A_2^0 的含义同上) 的 DWT.

对音频数据段 A_2^0 进行 H 级小波变换, 并得到小波系数 $A_2^{0H}, D_2^{0H}, D_2^{0H-1}, \dots, D_2^{01}$. 其中, A_2^{0H} 为音频数据段 A_2^0 小波变换的第 H 级近似分量, 而 $D_2^{0H}, D_2^{0H-1}, \dots, D_2^{01}$ 分别为音频数据段 A_2^0 小波变换的第 $1 \sim H$ 级的细节分量.

2) 近似分量的 DCT. 这里选择 H 级小波变换的近似分量 A_2^{0H} 进行 DCT.

$$A_2^{0HC} = DCT(A_2^{0H}) = \{a_2^0(t)^{HC}, 0 \leq t < L_2/2^H\}$$

3) 量化步长 S_2 的自适应选取. 由于水印信息已嵌入到 DCT 域的前 $M \times N$ 个系数内, 故为了减少数字水印干扰, 以确保水印提取过程能够有效计算出量化步长, 本文结合音频数据段内容并利用 DCT 域的后 $M \times N$ 个系数自适应确定量化步长 S_2 . 即

$$S_2 = 2^{\lfloor \log_2 s_b \rfloor} \quad (\lfloor \cdot \rfloor \text{ 为向下取整函数}), \quad s_b = \frac{\sum_{t=L_2/2^H-1}^{L_2/2^H-1} |a_2^0(t)^{HC}|}{M \times N}$$

4 数字水印的提取

本文讨论的自适应混合域数字音频水印算法属于盲水印算法, 即检测数字水印信号时不需要原始音频载体.

4.1 同步码的检测

这里所涉及到的同步码检测, 是指检测窗口 (大小为 L_1) 所覆盖的音频数据段内是否含有所嵌入的同步码信息. 其工作步骤为:

1) 按照第 3 节工作步骤, 自适应确定出量化步长 S_1^* .

2) 按照第 2.2 节工作步骤, 分别计算音频数据段 (即窗口所覆盖部分) 前 $n \times m$ 个采样值 $PA^*(m)$ 的平均值 $\overline{PA^*(m)}$

$$PA^*(m) = \{pa^*(m)(i) = a^*(i + m \times n), 0 \leq i < n, 0 \leq m < L_{syn}\}$$

3) 提取同步码信息 F' . 提取策略为

$$F' = \{f'(m) = \lfloor \overline{PA^*(m)} / S_1^* \rfloor \text{ mod } 2, 0 \leq m < L_{syn}\}$$

4) 为了有效避免假同步现象发生, 本文采用了通信领域的帧同步码逐位比较方法确定同步码, 即如果所提取出的同步码 F' 和原同步码 F 完全相同, 则认为已检测到所嵌入的同步码信息.

4.2 水印信号的提取

这里所涉及到的水印信号提取, 是指从同步码后的音频数据段 (令该音频数据段为 A_2^{0*} , 其长度为 L_2) 内提取水印信息. 其工作步骤为:

1) 按照第 3 节工作步骤, 结合音频数据段 A_2^{0*} 内容自适应确定出量化步长 S_2^* .

2) 对音频数据段 A_2^{0*} 进行 H 级 DWT, 得到小波系数

$$A_2^{0*H}, D_2^{0*H}, D_2^{0*H-1}, \dots, D_2^{0*1}$$

3) 对近似分量 A_2^{0*H} 作 DCT 变换

$$A_2^{0*HC} = DCT(A_2^{0*H}) = \{a_2^{0*}(t)^{HC}, 0 \leq t < L_2/2^H\}$$

4) 提取水印信息, 提取公式为

$$W' = w'(i, j) = \left[a_2^{0*}(t)^{HC} / S_2^* \right] \bmod 2$$

$$\hat{w}^*(i, j) = \begin{cases} 1 & \text{如果 } w_r^*(i, j) \text{ 中有一半以上 (不含一半) 为 } 1 \\ 0 & \text{否则} \end{cases} \quad (r = 1, 2, \dots, R)$$

这里, $w_r^*(i, j) \in W_r^*$, $\hat{w}^*(i, j) \in \hat{W}^*$.

5 实验结果与结论

为了验证本文数字音频水印算法的高效性, 本文进行了检测性能测试、抗攻击能力测试的实验, 并与文献 [8] 算法进行了对比. 实验中, 所选用的原始载体是采样频率为 44.1 kHz, 分辨率为 16 比特, 长度为 9.75 秒的单声道数字音频信号. 小波变换采用了常见的 Daubechies-1 小波基. 小波变换级数选取为 $H = 2$, 并选用了码长为 16 位的巴克码 1111100110101110 作为同步信号. 由于采用的水印在感觉上是可视的, 所以提取的水印信息很容易辨别.

表 1 给出了本文算法和文献 [8] 算法的抗攻击能力对照结果 (包括所提取出的水印图像、数字水印的归一化相关系数 NC 和失真率 BER、数字音频的

$$(0 \leq i < M, 0 \leq j < N, t = (i - 1) \times N + j)$$

5) 对所提取出的水印信息 W' 进行逆置乱加密, 便可得到二值水印图像

$$W^* = \{w^*(i, j), 0 \leq i < M, 0 \leq j < N\}$$

需要说明的是, 由于本文对同步码及水印信息采用了循环嵌入策略, 故数字水印检测过程将提取出多个二值水印图像 $W_r^*(r = 1, 2, \dots, R)$. 为此, 需进一步按照“多数原则”计算出最终的数字水印 \hat{W}^* , 即

峰值信噪比 PSNR).

本文提出了一种稳健的自适应混合域数字音频水印算法, 该算法能够通过选取更为稳健的 16 位巴克码作为同步标记、结合数字音频内容自适应确定量化步长、将同步码嵌入到多个采样值的统计均值等措施, 将数字水印信号在混合域内嵌入到原始数字音频内. 同时, 仿真实验还表明, 本文所提出的音频水印算法不仅具有较好的不可感知性, 而且对叠加噪声、重新采样、重新量化、随机剪切等音频处理与攻击均具有较好的稳健性. 特别地, 该算法对于最为普通的 MP3 压缩攻击具有极强的抵抗能力 (尤其是高压缩比下). 此外, 本文算法还具有计算简单、容易实现、抽取水印时无需原始载体音频等特点, 这大大增强了其用于 MP3 音乐作品著作权保护的实用性.

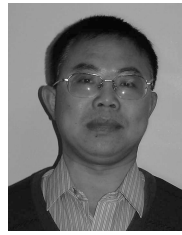
表 1 数字水印对部分常规攻击的抵抗能力

Table 1 The watermark detection results for various attacks

	未攻击	重量化	重采样	重采样	重采样	高斯噪声	剪切前面	剪切中间	Mp3	Mp3	Mp3	Mp3	Mp3	Mp3	
															22.05kHz
文献	NC	1.0000	0.9996	0.3733	0.3773	0.3702	0.9161	0.3851	0.3739	1.0000	1.0000	0.9948	0.3619	0.3638	0.3667
[8]	BER	0	0.0002	0.4983	0.4939	0.4954	0.0435	0.4833	0.4817	0	0	0.0029	0.5010	0.4932	0.5215
算法	PSNR	37.5415	37.3409	37.5850	22.8633	20.8454	27.9397	13.9636	20.7441	37.3108	37.3107	36.8539	36.3864	35.3564	35.8408
本文	NC	1.0000	1.0000	1.0000	0.9232	0.8978	0.9612	1.0000	1.0000	1.0000	1.0000	1.0000	0.9528	0.9748	0.9304
算法	BER	0	0	0	0.0645	0.1414	0.0443	0	0	0	0	0	0.0273	0.0144	0.0527
	PSNR	37.8174	37.4785	37.6815	22.8996	20.8620	28.0767	13.9789	20.7483	37.7747	37.9777	36.9516	36.5663	35.7669	35.8712

References

- 1 Kim H J. Audio watermarking techniques. In: Proceedings of Pacific Rim Workshop on Digital Steganography, Kyushu Institute of Technology, Kitakyushu, Japan, July 3~4, 2003
- 2 Lie W N, Chang L C. Robust and high-quality time-domain audio watermarking subject to psycho acoustic masking. In: Proceedings of the IEEE International Symposium on Circuits and Systems. Arizona, USA, 2002. **2**: 45~48
- 3 Yeo Kwon, Kim H J. Modified patchwork algorithm: A novel audio watermarking scheme. *IEEE Transactions on Speech and Audio Processing*, 2003, **11**(6): 381~386
- 4 Li Wei, Yuan Yi-Qun, Li Xiao-Qiang, Xue Xiang-Yang, Lu Pei-Zhong. Overview of digital audio watermarking. *Journal of Communications*, 2005, **26**(2): 100~111
(李伟, 袁一群, 李晓强, 薛向阳, 陆佩忠. 数字音频水印技术综述. 通信学报, 2005, **26**(2): 100~111)
- 5 Xu C, Wu J, Sun Q, Xin K. Applications of digital watermarking technology in audio signals. *Journal of the Audio Engineering Society*, 1999, **47**(10): 805~812
- 6 Swanson D., Zhu B., Tewfik H. Robust audio watermarking using perceptual masking. *Signal Processing*, 1998, **66**(6): 337~355
- 7 Zhao Chun-Hui, Li Fu-Chang. A blind audio watermark algorithm based on DWT and DCT. *Journal of Electronics & Information Technology*, 2003, **25**(11): 1570~1574
(赵春晖, 李福昌. 一种 DWT 与 DCT 结合的盲音频水印算法. 电子与信息学报, 2003, **25**(11): 1570~1574)
- 8 Megías D, Herrera-joancomartí J, Minguillón J. A robust audio watermarking scheme based on MPEG 1 layer III compression. *Communications and Multimedia Security - CMS 2003*, Springer-Verlag, LNCS **963**: 226~238. 2003



王向阳 辽宁师范大学计算机与信息技术学院教授, 主要研究领域包括网络信息安全技术、多媒体信息处理技术. 本文通信作者. E-mail: wxy37@263.net
(**WANG Xiang-Yang** Professor. His research interest covers information security and multimedia processing. Corresponding author of this paper.)



杨红颖 辽宁师范大学计算机与信息技术学院副教授, 主要研究领域包括多媒体技术、计算机辅助教学.
(**YANG Hong-Ying** Associate professor. Her research interest covers multimedia technique and CAI.)



赵红 辽宁师范大学计算机与信息技术学院硕士研究生, 主要研究领域包括信息隐藏与数字水印.
(**ZHAO Hong** Master student at Liaoning Normal University. Her research interest covers information security and digital watermarking.)