# 一种新的自适应量化数字音频水印算法*

王向阳[1], 杨红颖[1], 赵 红[2]
(1. 辽宁师范大学计算机与信息技术学院, 辽宁大连 116029;
2. 中国科学院软件研究所 信息安全国家重点实验室, 北京 100039)

摘要: 文章中提出了一种新的自适应量化数字音频水印算法, 该算法首先将视觉可辨的二值水印图像降维成一维水印序列, 并对水印序列进行随机置乱与 BCH 纠错编码, 再将原始数字音频信号划分成音频数据段, 最后选择音频段进行快速傅立叶变换(FFT), 并依据人类听觉系统(HAS)模型自适应确定量化步长量化 FFT 系数嵌入水印信息。该算法提取水印信息时不需要原始数字音频信号。仿真结果表明: 该自适应量化数字音频水印算法不仅具有较好的透明性, 而且对诸如叠加噪声、有损压缩、低通滤波、重新采样等攻击均具有较好的鲁棒性。

关键词: 音频水印; 人类听觉系统(HAS); FFT; 量化
中图分类号: TP391 文献标识码: A 文章编号: 1000-3630(2004)02-0117-04

## Digital audio watermarking algorithm based on adaptive quantization

WANG Xiang-yang[1], YANG Hong-ying[1], ZHAO Hong[2]
(1. *School of Computer and Information Technology, Liaoning Normal University, Liaoning Dalian* 116029, *China;*
2. *State Key Laboratory of Information Security, Institute of Software of Chinese Academy of Sciences, Beijing* 100039, *China*)

**Abstract:** In this paper, a new digital audio watermarking algorithm based on adaptive quantization is presented. A visually recognizable binary watermark image is first converted into a one-dimensional sequence. The sequence is permuted pseudo-randomly and modulated using BCH coding. Then the original digital audio signal is converted into segments and some audio segments are selected. Finally the modulated watermarking signal is embedded into the FFT coefficients by means of adaptive quantizaiton. The proposed algorithm can extract the watermark without the original audio signal. Experimental results show that the proposed watermarking scheme is inaudible and robust against various signal-processing operations such as noise addition, lossy compression, low pass filtering, re-sampling, and re-quantization.

**Key words:** digital audio; human auditory system (HAS); FFT; quantization

## 1 引 言

作为数字媒体作品知识产权保护的一种有效手段, 数字水印(Digital Watermarking)技术已得到广泛关注, 并成为国际学术界的一个研究热点[1]。依据应用范围, 通常可以把数字水印划分为静态图像水印、视频图像水印和音频水印。近几年, 对静态图像水印和视频图像水印(尤其是静态图像水印)的研究很多, 而对于数字音频水印技术的研究却较少报道[2-6], 这是由于: (1)与图像和视频相比, 音频信号在每个时间间隔内采样的点数要少得多, 意味着音频信号中可嵌入的信息量要比可视媒体少得多. (2)人类听觉系统(HAS)要比人类视觉系统(HVS)灵敏得多, 听觉上的不可知觉性实现起来要比视觉上困难得多。

本文以二值灰度图像为数字水印, 提出了一种新的自适应量化数字音频水印算法, 该算法首先将视觉可辨的二值水印图像降维成一维水印序列, 并对水印序列进行随机置乱与 BCH 纠错编码, 再将原始数字音频信号划分成音频数据段, 最后选择音频数据段进行 FFT 处理, 并依据人类听觉系统模型, 自适应确定量化步长量化 FFT 系数嵌入水印信息。该算法提取

收稿日期: 2003-05-10; 修回日期: 2003-10-08
基金项目: 辽宁省自然科学基金(20032100)和信息安全国家重点实验室开放基金(03-02)资助
作者简介: 王向阳(1965-), 男, 辽宁铁岭人, 教授, 硕士生导师, 主要研究领域为信息隐藏与数字水印、算法设计与分析。

水印信息时不需要原始数字音频信号。仿真结果表明: 该自适应量化数字音频水印算法不仅具有较好的透明性, 而且对诸如叠加噪声、有损压缩、低通滤波、重新采样等攻击均具有较好的鲁棒性。

## 2 二值水印图像的预处理

本文算法所采纳的数字水印信号不再是传统的伪随机序列, 而是一幅有意义的二值灰度图像。假设二值水印图像为 $W=\{w(i,j), 0\leqslant i<P, 0\leqslant j<Q\}$, 其中, $w(i,j)\in\{0,1\}$ 代表二值水印图像的第 $i$ 行、第 $j$ 列象素的灰度值。

考虑到要将二维的水印图像嵌入到一维的数字音频信号中, 故需对其进行降维处理, 即将二值水印图像转变为一维的二进制水印序列

$$S=\{s(k)=w(i,j),\ 0\leq i<P, 0\leq j<Q,\ k=iQ+j\}$$

为了消除二值水印图像的象素空间相关性, 提高数字水印算法鲁棒性, 确保含水印数字音频信号某一部分受到破坏后仍能全部或部分地恢复水印, 宜对上面得到的一维二进制序列 $S$ 进行置乱变换。本文采用文献[7]中的方法, 由密钥 $K_1$ 生成由 1 到 $P\times Q$ 的随机排列, 然后根据随机排列重新调整 $S$ 中向量元素的位置, 以得到由 $S$ 转换而来的新一维二进制序列 $V=\{v(k), 0\leqslant k<P\times Q\}$。

一般说来, 含有水印信息的数字音频信号在传输过程中很容易受到各种干扰和攻击(如低通滤波、叠加噪声、重新采样、重新量化、有损压缩等), 从而导致水印检测时产生错误。为了降低检测水印的差错率, 从而进一步提高数字水印的抗攻击能力, 本文算法采用了误差纠错编码技术[8]。当然这也会增加数字水印信息冗余量, 考虑到计算量和实现方便, 这里采用 BCH(Bose-Chaudhuri-Hocquenghem) 二进制分组纠错编码, 对水印信息进行预编码。在本文, 经过权衡编码效率与纠错能力两者关系, 选用了 BCH(7, 4) 编码, 码长为 7, 信息位为 4, 可以在 7 位码元中纠错 1 位. 经过 BCH 纠错编码, 可以得到待隐藏的水印序列 $X=\{x(k), 0\leqslant k<7\times P\times Q/4\}$。

## 3 数字音频水印嵌入算法

假设原始数字音频信号为 $A=\{a(i), 0\leqslant i<N\}$。其中, $N$ 为采样数据个数, $a(i)\in\{0, 1, 2, \cdots, (2^p-1)\}$ 是第 $i$ 个音频采样数据的幅度值, $p$ 表示每个采样数据所使用的比特数。

本文将采用对数字音频信号进行分段的方法嵌入水印信息, 即首先将整个音频信号划分为若干个音频数据段, 进而选取部分音频数据段用于嵌入水印(每个音频段仅嵌入 1 个比特的水印信息)。

### 3.1 音频信号的分段处理

假设每个音频段含有 $L$ 个采样数据(为了保证数字水印的透明性与鲁棒性, 音频段长度 $L$ 的取值以大于等于 8 为宜), 则可将整个音频信号划分成 $N/L$ 个音频段, 即

$$A=\{A(k),\ 0\leqslant k<N/L\},$$
$$A(k)=\{a(kL+i),\ 0\leqslant i<L\}$$

为了保证数字水印信息能够均匀地扩散到整个数字音频信号中, 以进一步提高数字水印的透明性和鲁棒性, 本算法首先计算每个音频段的能量值, 然后通过密钥控制, 在能量值较大(通过设置阈值) 的音频数据段中随机选取 $7\times P\times Q/4$ 个音频数据段 $A_e=\{A_e(k), 0\leqslant k<7\times P\times Q/4\}$ 用于水印嵌入[7]。

### 3.2 快速傅氏变换(FFT) 系数的选取

对所选取音频数据段 $A_e=\{A_e(k), 0\leqslant k<7\times P\times Q/4\}$ 进行快速傅氏变换(FFT):

$$F_e(k)=FFT(A_e(k)),\ 0\leqslant k<7\times P\times Q/4$$

式中, $F_e(k)=\{f_e(k, m), 0\leqslant m<L\}$, $f_e(k, m)$ 为第 $k$ 音频数据段 $A_e(k)$ 的快速傅立叶变换 $f_e(k)$ 的第 $m$ 个系数。

由人类听觉系统(HAS) 的掩蔽特性知: 频域中大幅度频率分量两侧的成分不易被感知。因此, 宜从第 $k$ 音频数据段 $A_e(k)$ 的 FFT 变换域内选取较大系数用于水印序列元素 $x(k)$ 的嵌入。本文选择了第三大 FFT 系数(通过 FFT 系数排序) 用于水印嵌入, 令该 FFT 系数为 $f_e(k, u)$。

### 3.3 数字水印的嵌入

水印嵌入是通过对所选择的 FFT 系数 $f_e(k, u)$ 进行特殊量化完成的。设 $f_e(k, u)$ 表示待量化的 FFT 系数, $f'_e(k, u)$ 表示量化处理后的 FFT 系数, $x(k)$ 为待嵌入水印比特, $\Delta$ 表示量化间隔. 与现有算法[3, 4, 6] 普遍采用均匀量化($\Delta$ 为常数) 不同, 本文将量化间隔选取为基于人类听觉系统而计算出的可感知噪声阈值 $JND$[9], 即 $\Delta=JND(k, u)$, 以便于进一步提高数字水印的透明性与鲁棒性。计算可感知噪声阈值 $JND$ 时, 采用了类似于 MPEG 音频中的模型 I (参见文献[9])。其基本步骤为: (1)通过

FFT 计算信号声压值; (2)根据信号短时功率分布情况确定音频信号的特性, 识别出类音调信号和类窄带噪声信号; (3)计算每个掩蔽者的掩蔽阈值; (4)计算出可感知噪声阈值 $JND$。

首先定义量化规则。本文将量化函数定义为用量化间隔 $JND(k, u)$ 对待量化 FFT 系数 $f_e(k, u)$ 进行二次取模求余运算, 即

$$Q(f_e(k,u)) = [f_e(k,u) \bmod (JND(k,u))] \bmod \quad (2)$$

其量化规则为: 如果 $f_e(k, u) \bmod (JND(k, u))$ 为奇数, 则 $Q(f_e(k, u))$ 取 1, 即 $f_e(k, u)$ 归于 1 类; 如果 $f_e(k, u) \bmod (JND(k, u))$ 为偶数, 则 $Q(f_e(k, u))$ 取 0, 即 $f_e(k, u)$ 归于 0 类。然后量化系数嵌入水印信息, 即对所选 FFT 系数 $f_e(k, u)$ 进行适当修改并赋予二值信息的过程。本文所采用的量化嵌入方案为:

如果 $Q(f_e(k, u))$ 不等于 $x(k)$, 则

如果 $x(k) = 0$, 则

$$f'_e(k,u) = \begin{cases} q_1 \cdot JND(k,u) - 1, \\ \quad 若 f_e(k,u) > JND(k,u) \\ q_1 \cdot JND(k,u) + 1, \\ \quad 若 f_e(k,u) < -JND(k,u) \\ f_e(k,u), \\ \quad 若 |f_e(k,u)| \leqslant |JND(k,u)| \end{cases}$$

否则

$$f'_e(k,u) = \begin{cases} q_2 \cdot JND(k,u), \\ \quad 若 |f_e(k,u)| > |JND(k,u)| \\ f_e(k,u), \\ \quad 若 |f_e(k,u)| \leqslant |JND(k,u)| \end{cases}$$

否则

不对 $f_e(k, u)$ 做任何处理;

其中, $q_1 = \left[\frac{f_e(k,u)}{JND(k,u)}\right]$, $q_2 = \left[\frac{f_e(k,u)}{JND(k,u)}\right]$

### 3.4 逆 FFT 及音频段的重新组合

为了得到含水印的数字音频信号, 需首先对含有水印信息的 $F'_e$ 做逆 FFT, 即

$$A'_e = IFFT(F'_e) = \{IFFT(F'_e(k)), \quad 0 \leqslant k \leqslant 7 \times P \times Q/4\}$$

将 $A'_e$ 代替 $A_e$, 再结合未被选用(嵌入水印)的音频数据段, 即可最终得到含有水印的数字音频信号 $A'$。

值得注意的是, 在嵌入数字水印过程中生成了抽取水印信息的密钥 $K_2$, 即用于嵌入数字水印信息的 FFT 系数位置。

## 4 数字音频水印检测算法

本文给出的水印检测算法无需原始音频信号。设待检测数字音频信号为 $A'$, 提取水印信息时, 首先采用水印嵌入算法中与 3.1 节相同的操作, 对待检测音频信号 $A'$ 进行分段处理, 并利用密钥控制选取出相应的音频段 $A'_e$

$$A'_e = \{A'_e(k), \ 0 \leqslant k < 7 \times P \times Q/4\}$$

再对所选音频数据段进行 FFT, 并利用密钥 $K_2$ 选择 FFT 系数 $f'_e(k, u)(0 \leqslant k < 7 \times P \times Q/4)$ 抽取数字水印信息 $X' = \{x'(k), 0 \leqslant k < 7 \times P \times Q/4\}$, 其水印信息的抽取规则为 $x'(k) = Q(f'_e(k, u))$。

最后对所抽取出的数字水印信息 $X'$ 按照水印图像预处理的相反操作(即 BCH 解码、逆随机置乱和升维), 便可以计算出嵌入的二值水印图像 $W' = \{w'(i, j), 0 \leqslant i < P, 0 \leqslant j < Q\}$。

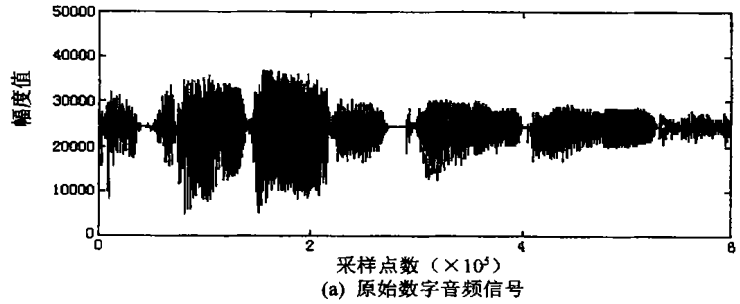
此外, 为了消除观测者的经验、身体条件、实验条件和设备等主、客观因素的影响, 需采用归一化相关系数(Normalized Cross-Correlation, NC)[10] 对抽取的水印和原始水印的相似性进行定量评价, 其定义为:

$$NC(W', W) = \frac{\sum_{i=0}^{P-1}\sum_{j=0}^{Q-1} w(i,j) \cdot w'(i,j)}{\sqrt{\sum_{i=0}^{P-1}\sum_{j=0}^{Q-1} w^2(i,j)} \sqrt{\sum_{i=0}^{P-1}\sum_{j=0}^{Q-1} w'^2(i,j)}}$$

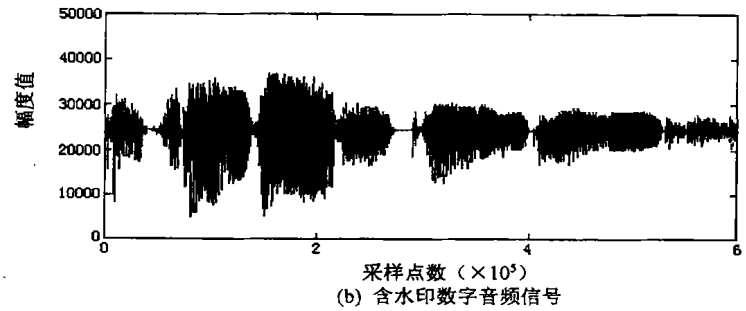
其中, $W'$ 是待判决的二值水印图像, $W$ 是嵌入的原始二值水印图像。如果该归一化相关系数 $NC$ 超过某一阈值 $T$, 我们就判定数字音频信号中存在此二值水印图像(阈值 $T$ 是用户根据数字水印的具体应用背景而确定的)。

## 5 实验结果

为了验证本文数字音频水印算法的高效性, 以下给出了检测性能测试、抗攻击能力测试的实验结果。实验中, 原始数字音频信号(流行音乐)的采样频率为 44.1kHz, 分辨率为 16 bit, 取 600000 个采样点, 其波形如图 1(a) 所示。二值水印图像(64 × 64)如图 1(c) 所示。音频数据段长度选取为 $L = 16$。水印相似性判别阈值 $T = 0.5$。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。另外, 本文还采用归一化相关系数($NC$)定量分析了提取水印与原始水印的相似度, 采用信噪比



(c) 原始二值水印图像



(d) 提取出的二值水印图像

图 1 数字水印嵌入与检测结果

(*SNR*)评价了原始数字音频信号与含水印数字音频信号之间的差别。

**5.1** 检测性能测试

图 1 的(a)、(b)、(c)、(d)分别为原始数字音频信号、嵌入水印后的数字音频信号(*SNR* = 44.2 dB)、原始二值水印图像、提取出的二值水印图像(*NC* = 0.9983)(未受到任何攻击)。

表 **1** 数字水印对常见攻击的抵抗能力

| 攻击方式 | 归一化相关系数 | |
|---|---|---|
| | 无水印音频信号 | 含水印音频信号 |
| 低通滤波 | 0.2302 | 0.9127 |
| 叠加噪声 | 0.1673 | 0.8436 |
| 重新采样 | 0.2185 | 0.9744 |
| 有损压缩 | 0.1621 | 0.8909 |

注: 水印相似性判别阈值 $T = 0.5$

**5.2** 抗攻击能力测试

通常, 对含水印数字音频信号的攻击方式主要包括低通滤波、叠加噪声、重新采样、有损压缩等。为了检测算法的鲁棒性, 仿真实验对未嵌入水印的原始音频信号和含水印音频信号分别进行了以上处理, 表 1 给出了本文算法对常见信号处理与攻击的抵抗能力(归一化相关系数)。其中

⑴低通滤波。采用长度为 6 阶, 截止频率为 22.05kHz 的巴特沃兹低通滤波器。

⑵叠加噪声。对数字音频信号在时域中加入高斯白噪声。

⑶重新采样。将音频信号采样频率提升为 48kHz, 再还原为原采样频率 44.1kHz。

⑷有损压缩。先对音频信号进行压缩(压缩比为 12: 1), 再进行解压缩。

## 6 结束语

本文以快速傅立叶变换(FFT)、人类听觉系统(HAS)为基础, 提出了一种自适应量化数字音频水印算法。该算法具有以下特点: ⑴数字水印选用了有意义的二值灰度图像; ⑵嵌入水印信息时, 采用了基于人类听觉系统(HAS)的自适应量化间隔; ⑶数字水印的提取不需要原始音频信号。同时仿真结果也表明: 该数字音频水印算法不仅具有较好的透明性, 而且对诸如叠加噪声、有损压缩、低通滤波、重新采样等攻击均具有较好的鲁棒性。

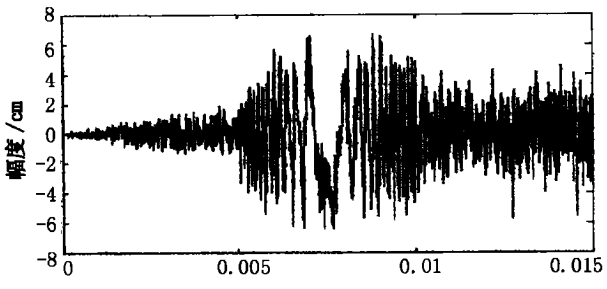
## 参 考 文 献

[1] Cox I J, Matt L Miller. The first 50 years of electronic watermarking [J]. Journal of Applied Signal Processing, 2002, **2002** (2): 126-132.

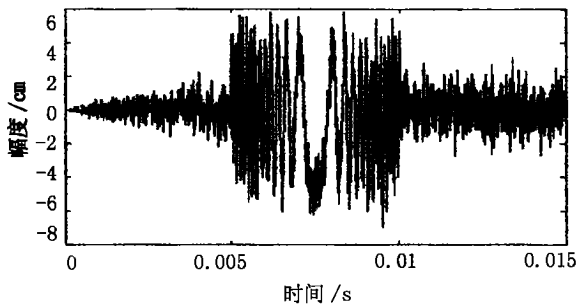
[2] Bassia P, Pitas I, Nikolaidis N. Robust audio watermarking in the time domain [J]. IEEE Trans on Multimedia, 2001, **3**(2): 232-241.

[3] 孙锐, 孙洪, 姚天任. 基于量化的小波域数字音频水印算法[J]. 华中科技大学学报(自然科学版), 2002, **30**(5): 12-15.

(下转第 **127** 页)




(g)　噪声功率改变后的GACI算法输出波形




(h)　噪声功率改变后的改进的 GACI 算法输出波形

图 1　MATLAB 仿真结果( $SNR = -5$dB)

## 3　仿真结果及其分析

本文中主要利用前述的四种算法, 采用线谱增强器的形式, 提取高斯白噪声背景下的线性调频脉冲信号。不同输入信噪比下, 高斯白噪声、频宽 2kHz 脉冲宽度 50ms 的线性调频信号以及输出信号的波形如图 1 所示( 图 1 为 MATLAB 仿真结果)。

图 1 中的( c)、( d)、( e) 分别是在相同的信噪比( $SNR = -5$dB) 下 LMS 算法、ACI 算法、GACI 算法的仿真结果。显然 GACI 算法的性能优于 ACI 算法和 LMS 算法。图 1 中的( f)、( g)、( h) 分别是在噪声功率增加, 而 $SNR$ 不变时 ACI 算法、GACI 算法和改进的 GACI 算法的仿真结果。可以看出在 $SNR$ 不变而噪声功率变化时, 改进的 GACI 算法的滤波器参数仍然最佳, 因而性能远远高于其他算法。

## 4　结束语

本文研究了 LMS 算法、ACI 算法、GACI 算法的原理, 在比较了各算法性能的基础上对 GACI 算法加以改进, 以降低其参数对背景噪声变化的敏感性; 用软件仿真器进行仿真, 并对结果加以分析。从而发现改进后的 GACI 算法的参数对背景噪声功率变化的敏感性明显降低, 并证明了其可靠性以及工程上的可实现性, 预计将有广泛的实际应用前景。

## 参　考　文　献

[1] Tugay M A, Tanik Y. Properties of the momentum LMS Algorithm[J]. Signal Processing. 1989, **18**: 117-127.

[2] Bershad N J, CHANG Y H. Time correlation statistics of the LMS adaptive algorithm weights[J]. IEEE Trans Acoustics, Speech, Signal Processing, 1985, **33**(1): 309-312.

[3] 林山, 李志舜, 马远良. 一种推广的自适应相干累积算法——基本原理与应用[J]. 西北工业大学学报, 1992, **10**(2): 147-151.

[4] 马远良. 检测单频或调频脉冲信号的自适应相干累积[A]. 第一届全国信号处理学术会议论文集[C]. 1984. 971-974.

[5] 沈福民. 自适应信号处理[M]. 西安: 西安电子科技大学出版社, 2001.

[6] 苏涛, 吴顺君, 李真芳, 宋万杰. 高性能 DSP 与高速实时信号处理(第二版)[M]. 西安: 西安电子科技大学出版社, 2002.

(上接第 **120** 页)

[4] Michael Arnold, Sebastian Kanda. MP3 robust audio watermarking[A]. IEEE International Conference on Multimedia and Expo (ICME)[C]. 2000, (1, 2), New York, NY, USA, 2000. 1013-1016.

[5] 王秋生, 孙圣和. 一种在数字音频信号中嵌入水印的新算法[J]. 声学学报, 2001, **26**(5): 465-467.

[6] Emilia Gomez, Pedro Cano, et al. Mixed watermarking-fingerprinting approach for integrity verification of audio recordings[A]. International Telecommunications Symposium-ITS2002[C]. Natal, Brazil, 2002.

[7] Hsu Chiou-Ting, Wu Ja-ling. Hidden digital watermarks in images [J]. IEEE Trans on Image Processing, 1999, **8**(1): 58-68.

[8] 归绍升. 纠错编码技术与应用[M]. 上海: 上海交通大学出版社, 1988.

[9] Information technology-coding of moving pictures and associated audio for digital storage media at up to about 1.5 m bits/s-part 3: audio. British standard, BSI, London, 1993.

[10] Martin Kutter, Fabien A P Petitcolas. A fair benchmark for image watermarking system[A]. In: Proceedings of Security and Watermarking of Multimedia Contents [C]. San Jose, CA, USA, 1999. 226-239.