一种基于部分 MP3 编码原理的音频水印*

刘 伟，王朔中，张新鹏
(上海大学通信与信息工程学院，上海 200072)

摘 要：提出一种可抵抗 MP3 编码的音频水印方案。此方案对原始的 PCM 音频进行部分 MP3 编码，通过调整相邻的一对量化 MDCT 系数的大小关系嵌入水印信息，然后进行局部 MP3 解码得到含水印的 PCM 信号。水印提取时不需原始信号，只要依次比较相应频带中相邻的一对 MDCT 系数的大小即可。由于水印算法考虑到 MP3 编码的性质，所以对 MP3 有优良的稳健性。实验表明该水印方案在隐蔽性和嵌入量方面也有很好的性能。

关键词：音频水印；MP3 编码；MDCT
中图分类号：TP391；TP309 **文献标识码**：A **文章编号**：0529-6579 (2004) S2-0026-04

数字水印主要用于多媒体产品的版权保护，早期大部分研究集中在数字图像，随着数字音乐制品的大量制作和发行，特别是 MPEG－1 Layer Ⅲ即 MP3 的流行，数字音频产品版权保护的重要性日益突出，随之出现了各种音频水印嵌入技术[1,2]。本文的研究正是基于上述应用的需求而展开的。

压缩编码是音频水印必须面对的一种常规处理，近年来针对某些编码标准如 MP3 和 AAC 等提出了许多音频水印的算法。这些算法总结起来可分为以下几类：一类是在编码过程中进行信息隐藏[3]，输入 PCM 码流，输出含水印的压缩域码流；另一类算法输入输出都是压缩编码后的比特流，但在嵌入水印时需要对输入比特流进行部分或者完全解码，数据嵌入后再进行编码得到含水印的码流[4]；还有一些算法是直接将信息嵌入在压缩域码流中[5]。本文算法也是针对 MP3 编码标准而设计的，但与上述几类算法不同的是，本算法输入输出的皆为 PCM 信号。

本文所提出的方案是对输入的 PCM 信号，严格按照 MP3 编码原理[6]进行处理，在达到编码过程中的某一阶段，对某些参数进行修改以嵌入水印信息，然后按照解码的处理方式将带水印的信号还原成 PCM 码。因此，本算法的针对性强，对 MP3 压缩具有很好的稳健性。

1 MP3 编码原理的水印嵌入方案

本算法首先是将输入的 PCM 信号严格按照 MP3 编码过程进行处理，即经过心理声学模型、分析滤波器组、MDCT 变换以及量化与编码后，通过修改位于中频带小值区中的 MDCT[7] 系数完成水印的嵌入。嵌入时，用相邻的一对 MDCT 系数的大小关系来表征水印信息。图 1 给出了基于部分 MP3 编码原理的水印嵌入方案的框图。

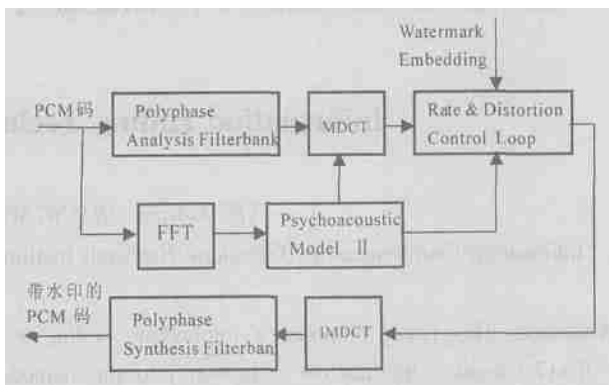

图 1 基于部分 MP3 编码原理的水印嵌入方案的框图

Fig.1 The frame of the proposed watermarking procedure

在 MP3 编码中，分析滤波器组和 MDCT 变换主要是对输入信号进行时频转换，心理声学模型为 MDCT 和量化编码提供必要的参数。对 MDCT 系数的量化编码通过内外两层迭代循环来实现：外层迭代循环即误差控制循环，使量化误差被控制在容许的范围之内（由心理声学模型提供）；内层迭代循环即比特率控制循环，决定着量化步长的大小，使输出后的比特率满足一定的限制。量化后的 MDCT 系数分成 3 个区间，编码器依据不同的区间使用不同的霍夫曼表进行编码。编码器将高频分量中一串

* 收稿日期：2004－09－11
基金项目：国家自然科学基金资助项目（60372090）；上海市教委青年科研基金资助项目
作者简介：刘伟（1978 年生），男，硕士研究生；E-mail：whatease@citiz.net

连续的零视为一个区间，称为“零区（rzero region）”，在该区是不用编码的。第二个区间称为“小值区（count1 region）”，是由一连串的绝对值小于等于 1 的值组成的，该区间的量化值以四个为一组进行霍夫曼码编码。剩下的部分为“大值区(big values region)”，此区间会出现较大的数值，霍夫曼编码对这一部分是成对编码的。

图 2 为某一帧信号在迭代循环结束时得到的量化后的 MDCT 系数，其中编码后的比特率为 64 kbps。左边部分为“大值区”，中间部分为“小值区”，右边空白部分为“零区”。

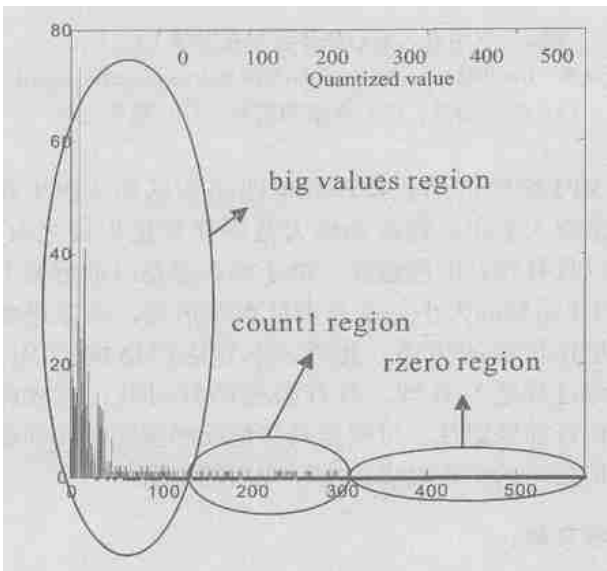

图 2　量化后的系数及分区

Fig.2　The quantized coefficients in three regions

因本文选择在小值区中进行水印的嵌入，并用相邻的一对 MDCT 系数之间的大小关系来表征水印信息，所以每个系数的最大修改量为 1。而且小值区一般位于整个频率域的中频段，因此可满足水印系统对稳健性与隐蔽性的综合要求。MP3 把整个频率域分为 21 个比例因子频带（scalefactor band），具体选择哪个比例因子频带进行水印的嵌入，可根据输出的比特率来定。例如，当编码速率为 64 kbps（单声道）时，第 18 个比例因子频带位于小值区的概率是比较大的，所以本文选择该比例因子频带进行修改。以 44.1 kHz 的采样频率为例，第 18 个比例因子频带包括第 196 到第 237 条谱线，其频率范围在 7.5～9.2 kHz 之间。

嵌入时，对一帧信号严格按照 MP3 编码过程进行处理，迭代循环结束后，首先判断第 18 个比例因子频带中的 MDCT 系数是否位于小值区，如果是，则该帧作为水印的嵌入帧，否则处理下一帧。本方案以相邻两个 MDCT 系数为一组，通过修改两者之间的大小关系来嵌入水印信息，具体的嵌入规则为：如果待嵌入 1 bit，则使得前一个 MDCT 系数的幅值为 1，其正负与量化前的 MDCT 系数保持一致，同时将后一个 MDCT 系数的幅值修改为 0；如果待嵌入 0 bit，则使得前一个 MDCT 系数幅值为 0，将后一个 MDCT 系数的幅值修改为 1。在接收端，不需要进行量化编码，只要判断该比例因子频带中每一对 MDCT 系数幅值的大小关系即可提取出水印信息。因为第 18 个比例因子频带有 42 条谱线，所以每一帧可嵌入 21 bit 水印。

2　实验结果及性能分析

本实验所使用的载体信号为 44.10 kHz 采样、量化精度为 16 bit 的单声道高保真音乐。为了能够抵抗压缩比为 11∶1（单声道 64 kbps）的压缩，在水印嵌入过程中进行量化编码时采用同样的比特率，帧长取 576，对应于 13.06 ms。

图 3 显示了某一帧信号中第 18 个比例因子频带在嵌入 21 bit 水印信息前后的变化。水印信号为：101011001011110011010。该帧信号在水印嵌入后的信噪比为 29.90 dB。上图为迭代循环结束时量化后的 MDCT 系数的幅值，下图为水印嵌入后的 MDCT 系数的幅值。从理论上可证明水印嵌入前后每条谱线被修改的概率为 50%。假设有前后两条谱线 A 和 B，每条谱线的幅值为的概率为 x，那么幅值为 0 的概率为 $1-x$。假设此时嵌入的比特为“1”的概率为 α，根据嵌入规则，那么两条谱线被修改的概率为：

$$p_1 = \alpha(1-x) + \alpha x = \alpha \tag{1}$$

同理，此时嵌入的比特为“0”的概率为 $1-\alpha$，两条谱线被修改的概率为：

$$P_0 = (1-\alpha)x + (1-\alpha)(1-x) = 1-\alpha \tag{2}$$

所以，每条谱线被修改的概率为：

$$\bar{p} = \frac{p_1 + p_0}{2} = \frac{\alpha + (1+\alpha)}{2} = 0.5 \tag{3}$$

由图 3 可见，水印嵌入前后共有 20 条谱线的幅值发生了变化，其修改率为 20/42＝47.6%，接近于理论上的修改概率 50%。

图 4 是图 3 中带水印的信号在 MP3 压缩后，再将其经过分析滤波器组和 MDCT 变换后得到的第 18 个比例因子频带的 MDCT 系数，由前后两个系数的大小关系可依次提取出所嵌入的水印信息为：101011001011110011010，水印的正确提取率为 100%，可见该算法对 MP3 具有很好的稳健性。对其他音频信号的实验也得到同样的结果。

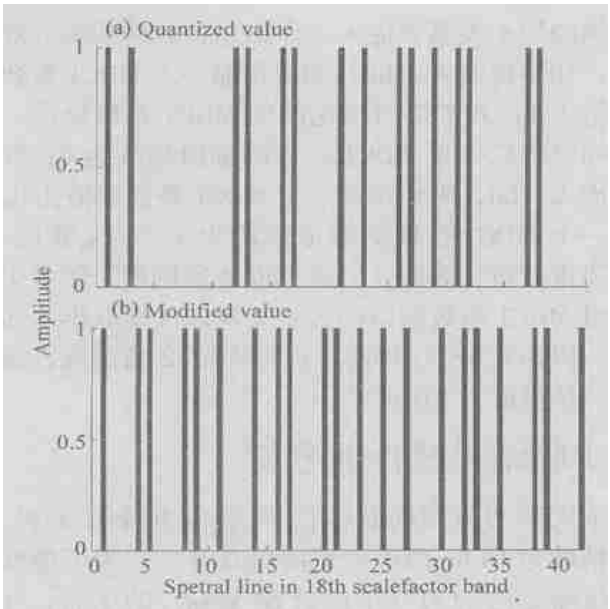

图 3 水印嵌入前后第 18 个比例因子频带中频谱的变化

Fig.3 The difference of the 18th scalefactor band between the host and watermarked signal

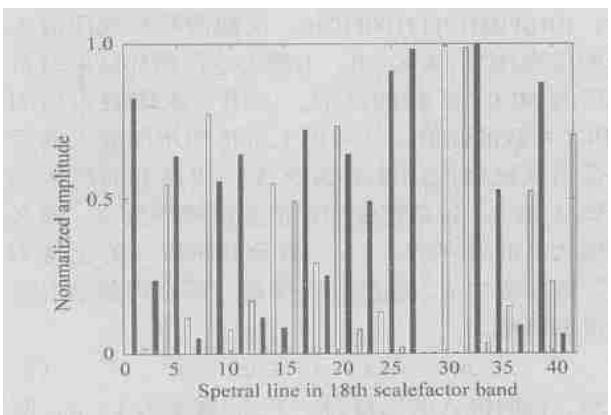

图 4 MP3 压缩后第 18 个比例因子频带的 MDCT 系数

Fig.4 The MDCT coefficients after MP3 encoding

图 5 显示一段 23 s 的古典音乐在水印嵌入前后的波形变化。该段音乐在水印嵌入后的信噪比为 36.56 dB。上图为原始信号，中图为水印嵌入后的信号，下图为两者之差。2 个嵌入帧之间至少相距 10 帧（可根据实际要求的嵌入量调整），嵌入量为 2 625 bit。

3 结 论

本文提出了一种基于 MP3 编码的音频水印算法。本算法对 PCM 音频信号进行部分 MP3 编码，

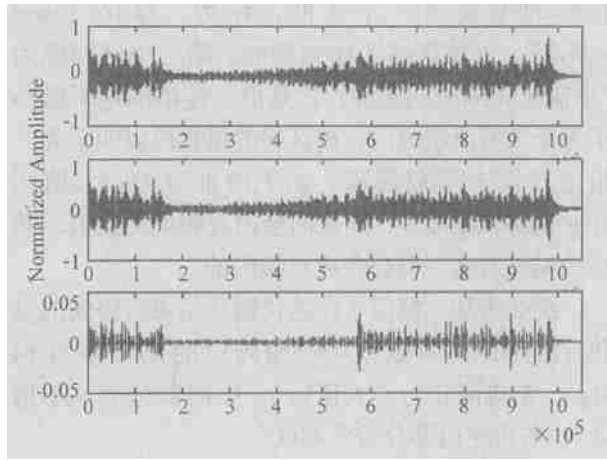

图 5 水印嵌入前后载体信号在时域上的变化

Fig.5 The difference between the host and watermarked signal

（a）原始信号；（b）含水印信号；（c）两者之差

在 MP3 编码的迭代循环后修改小值区的 MDCT 系数来嵌入水印。修改的最大范围在量化步长之内，所以具有较好的隐蔽性。由于本算法是以前后两个 MDCT 系数的大小关系来表征水印信息，所以避免了量化步长的传递。此外，本算法严格按照 MP3 编码过程进行处理，具有很强的针对性，对 MP3 有很好的稳健性。可根据具体情况调整隐蔽性和嵌入量之间的关系以满足实际应用的要求。

参考文献：

[1] BASSIA P，PITAS I，NIKOLAIDIS N. Robust audio watermarking in the time domain[J]. IEEE Trans Multimedia，2001，3：232－241.

[2] WANG S，ZHANG X，ZHANG K. Data hiding in digital audio by frequency domain dithering[J]. Lecture Notes in Computer Science，2003，2776：383－394.

[3] PETITCOLAS F. MP3 Stego[C]. Computer Laboratory，Cambridge，1998.

[4] NEUBAUER C，HERRE J. Audio watermarking of MPEG－2 AAC bitstreams[C]. In 108th AES Conv Rance，2000：5101.

[5] KOUKOPOULOS D K，STAMATIOU Y C. A compressed-domain watermarking algorithm for mpeg audio layer 3[C]. ACM 2001.

[6] ISO/IEC 11172－3，Information Technology －Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbits/s，Part 3：Audio，1993.

[7] NOLL P，PAN D. Modified discrete cosine transform － its implications for audio coding and error concealment[J]. J Audio Engineering Society(AES)，2003，51(1/2).

（下转第 33 页）

image encryption and decryption [J]. Proc IEEE Int Conference Circuits and Systems, 2000,4:49—52.
[7] WU X X, LU J H. Dynamical behavior of logistic equation on a discrete interval[J]. J Shan Hai Teacher University, 1995 (4):31—33.
[8] SCHARINGER J. Fast encryption of image data using chaotic kolmogorov flows[J]. J Electron Imaging, 1998,7(2):318—325.
[9] WANG D S, CAO L. Chaos, fractal and their applications [M]. Hefei: The Press of The University of Science & Technology of China, 1995.
[10] AIGRIER M, Kombinatorik[M]. Springer Verlag, Berlin, Heidelberg, New York, 1975.
[11] JEGER M, HRUNG E. In Die Kombinatorik II. Klett, Stuttgart, 1973.

A New Image Encryption Algorithm Based on Chaotic Map

XIONG Chang-zhen[1,2], *ZOU Jian-cheng*[2], *QI Dong-xu*[1,3]
(1. Department of Electronics, Sun Yat-sen University, Guangzhou 510275, China;
2. CAD Center, North China University of Technology, Beijing 100041, China;
3. College of Information Sciences, Macao University of Sciences and Technology, China)

Abstract: Many digital image encryption algorithms based on chaotic map were presented; most of them use the quantization of coefficients. In finite precision, chaotic map will generate a circulation or an unmovable point, which does not accord with perfect density distributing function. This paper proposes a new encryption algorithm based on chaotic map without transcendental knowledge of orbit distributing. So that people can choose any chaos model, which extends the key space of algorithm. The algorithm reduces iterative number and makes full use of chaotic property of map. This paper also introduces the segmentation mechanism of general baker transformation, which enhances the key space and security of algorithm. The experiment shows that it has the property of key sensitivity, high key space and pixel distributing uniformity.

Key words: chaotic map; arrangement; image encryption

(上接第 28 页)

Audio Watermarking Based on Partial MP3 Encoding

LIU Wei, *WANG Shuo-zhong*, *ZHANG Xin-peng*
(School of Communication and Information Engineering, Shanghai University, Shanghai 200072, China)

Abstract: An audio watermarking scheme based on partial MP3 encoding is proposed. The original PCM signal is first undergone a partial MP3 encoding to produce quantized MDCT coefficients. An appropriate band in the count1 region, in which magnitude of any quantized MDCT coefficient is either 1 or 0, is chosen for data embedding. Each pair of the quantized MDCT coefficients is used to hide 1 bit. After modification of the coefficients, partial decoder is performed to return to the PCM domain. As the algorithm has taken account the characteristics of MP3 encoding, it is robust against MP3. The method also has good inaudibility and large embedding capacity.

Key words: audio watermarking; MP3 encoding; MDCT