用于篡改检测及认证的脆弱音频水印算法

全笑梅　张鸿宾
(北京工业大学计算机学院　北京　100022)

摘　要： 该文提出一种新的用于确保音频作品可信性的脆弱水印算法。心理声学模型控制音频信号的小波包分解和小波包域内水印的嵌入量，增强了水印嵌入的透明度和灵活性。与现有算法不同，该算法在单一系数上嵌入多位水印，可以更加准确地反映篡改的程度。通过引入篡改检测函数，可同时进行完整性认证和在时/频域定位篡改并给出篡改程度的度量。仿真实验结果显示该算法能有效地检测和定位篡改，可用于衡量法庭证据及新闻广播等的可信度。

关键词： 多媒体认证，数字音频水印，心理声学模型，小波包分解
中图分类号： TP391　　**文献标识码：** A　　**文章编号：** 1009-5896(2005)08-1187-06

Fragile Audio Watermarking Algorithm for Telltale Tamper Proofing and Authentication

Quan Xiao-mei　Zhang Hong-bin
(*College of Computer Science, Beijing University of Technology, Beijing* 100022, *China*)

Abstract A novel watermarking scheme to ensure the credibility of audio is proposed in this paper. Psychoacoustic model controls the wavelet packet decomposition as well as the watermark insertion in the wavelet packet domain, which enhances both the transparency and the flexibility of the algorithm. Unlike previously proposed algorithms, the presented method embeds several bits of watermark instead of just one in a single coefficient. By introducing the detecting function, the new scheme can not only authenticate the integrity of the audio signal but also locate the tampered region and decide to what extent the audio being tampered. Experimental results show that the proposed scheme has the desired property and can be used for applications including authentication for forensic identification and news broadcasting.

Key words Multimedia authentication, Digital audio watermarking, Psychoacoustic model, Wavelet packets

1　引言

数字化技术和互联网的迅猛发展使得数字媒体的复制，修改和传播简单易行。但同时数字媒体的真实性、完整性也受到了严重威胁。数字水印技术作为一种可能的保护手段，近年来已成为一个活跃的研究领域，出现了许多相关的算法。例如：文献[1,2]提出的水印算法就是用于认证和进行篡改检测的。文献[1]提出鸡尾酒水印算法，通过正负调制在音频信号的 FFT 域一次嵌入水印，水印信号为原音频信号的量化信息。该算法可以使用不同的检测过程用于版权保护或进行篡改检测。文献[2]提出的算法针对静止图像，通过量化小波系数嵌入二值水印，量化步长的选择取决于小波变换的尺度大小。

为了满足不可感知性的要求，一个有效的水印算法必须应用相应的感知模型[3,4]。对于音频信号，嵌入水印的过程就是改变原始音频信号的过程，改变到什么程度就能被人觉察出来由人的听觉系统模型 HAS(Human Auditory System)决定。该模型的主要目的就是通过分析输入音频信号，确定量化噪声在输入音频信号频谱的哪些位置可以被掩蔽(Masking)掉，同时定量地确定掩蔽的程度。掩蔽是人耳的一种感知局限性，指一个音频信号的听觉阈值由于另一个音频信号的存在而升高，即一个音频信号的存在会使人耳听不到在其频谱附近或时域附近出现的其它信号。大量的心理声学实验表明人耳的掩蔽特性是具有频率选择性的，在不同的频谱范围其敏感程度也不同。具有相同频率选择性的频谱构成一个临界频带(Critical band)。由于临界频带的存在，在某一频率的噪声掩蔽阈值完全由它所处的临界频带内的信号决定。因此音频信号的子带分解越接近临界频带，听觉系统模型就能越好地模拟人耳的生理机能[5]。

2004-02-23 收到，2005-01-28 改回
国家自然科学基金(60075002)，国家 863 计划 (2003AA144080)和北京市自然科学基金资助课题

但是现有水印算法选用的听觉系统模型一般都应用快速傅里叶变换(FFT) 划分子带(例如文献[1]提出的算法)，传统的 FFT 虽然能在频域准确定位，却毫无时域定位性，只是人为把某些频谱划分为一个子带，所以不能充分利用人类听觉系统特性。近年来提出的小波域音频水印，利用小波多分辨率分析的优点使水印性能有所提高。小波变换遵循小尺度大频窗、大尺度小频窗的时频分布特性可以分析任意尺度的信号，但是由于小波变换的频率分辨率随频率升高反而降低，因此若只对某些特定频率点或时间点的信息感兴趣时，小波变换的这种时频窗口分布显然不是最优选择。具体表现为一般小波变换中的频率分析与已有的心理声学模型的临界频带划分并不一致，因此限制了一般小波域的水印性能。

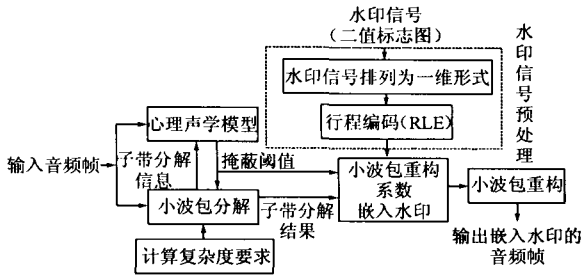
本文提出一种脆弱水印算法，对适用于 MPEG1 层 1 的心理声学模型 II 进行改进，使之适用于小波包域，根据子带掩蔽阈值和水印嵌入和提取端要求的计算复杂度自适应地选择最好小波包基对音频信号进行接近临界频带的分解与重构。然后采用量化小波包系数的方法自适应地嵌入水印信号。算法采用的水印为二值标志图像。与已有算法[1,2]相比，本文算法有如下优点：(1)引入心理声学模型控制子带分解和水印嵌入过程，确保了水印嵌入的不可感知性，增强了算法的透明度。(2)通过对 MPEG1 层 1 心理声学模型 II 的改进，进行自适应小波包分解，使子带划分接近临界频带，因此得到的子带掩蔽阈值大于文献[1]中通过 FFT 计算得到的阈值。当用该掩蔽阈值作为步长量化变换系数时，可以更好地控制水印的嵌入量。(3)采用自适应小波包分解具有更好的时频域定位性，因此本文的脆弱水印不仅可以检测给定音频文件是否被篡改而且可以在时域和频域更为精确地定位篡改。(4)文献[2]中提出的算法对篡改检测的敏感性是与尺度有关系的，即小/大尺度的小波系数对小/大篡改敏感，这与人类的感知特性不符。本文应用心理声学模型的研究成果，根据掩蔽阈值量化小波包系数避免了上述问题的发生。

2 系统描述

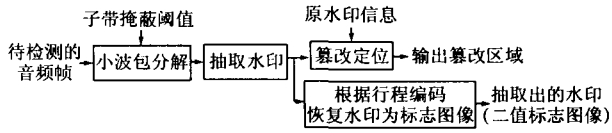
2.1 概述

在本文的音频水印系统(图 1)中一帧信号由 512 个采样点组成，同时作为心理声学模型和小波包分解部分的输入。心理声学模型计算子带掩蔽阈值，控制小波包分解。分解结束后，函数空间二剖分形成一棵二叉树，该二叉树叶子结点对应的函数构成了符合当前应用的小波包分解的最好基。小波包分解结束后，保存每个子带的掩蔽阈值，用于水印嵌入和检测过程。水印在量化小波包系数时嵌入。

检测时首先进行小波包分解，然后抽取水印，根据所得水印判断被测音频是否与原始文件一致。在不一致的情况下



(a) 水印嵌入过程结构图



(b) 水印抽取和定位过程结构图

图 1 数字音频水印系统结构图

可通过篡改检测函数定位篡改并估计篡改程度。

2.2 心理声学模型及自适应小波包分解

本文选用的心理声学模型是 MPEG1 层 1 心理声学模型 II 的改进版本。MPEG1 标准的第 3 部分(11172-3)是基于感知的子带压缩和编码方案[6]。子带划分模拟临界频带，充分利用人耳的掩蔽效应进行压缩编码。临界频带表征了人类最主要的听觉特性，它是在研究纯音对窄带噪声掩蔽量的规律时被发现的，在加宽噪声带宽时，最初是掩蔽量增大，但带宽超过某一定值后，掩蔽量就不再增加，这一带宽就称为临界频带。因此临界频带对应的子带掩蔽阈值是最大的。所以子带分解与临界频带的接近程度就决定了应用心理声学模型的有效性。

ISO/IEC 11172-3 把输入的音频信号分成等宽的频率子带，用来模拟临界频带。虽然 MPEG1 子带划分方案较为简便，但是如同文献[5,7]指出的那样，这种方案不能准确反映人类听觉系统的频率依赖特性，因为它的子带划分与临界频带有较大差距。为了解决这一不足，本文把小波包分解与 MPEG1 层 1 心理声学模型 II 相结合，使心理声学模型用于小波包域，综合考虑掩蔽阈值与要求的算法复杂度动态进行子带划分。

首先，根据前两个阈值计算块(即前两帧信号)的幅度，相位值计算本帧信号的预测幅度 $\hat{r}_w$ 和相位 $\hat{f}_w$：

$$\left.\begin{aligned}\hat{r}_w &= 2.0r_w(t-1)-r_w(t-2)\\ \hat{f}_w &= 2.0f_w(t-1)-f_w(t-2)\end{aligned}\right\} \tag{1}$$

其中 t 标记当前计算块，$t-1$，$t-2$ 分别标记前一和前两个阈值计算块。由此得到非预测量度 C_w：

$$C_w=\frac{\left(\left(r_w\cos f_w-\hat{r}_w\cos\hat{f}_w\right)^2+\left(r_w\sin f_w-\hat{r}_w\sin\hat{f}_w\right)^2\right)^{0.5}}{r_w+\mathrm{abs}\left(\hat{r}_w\right)} \tag{2}$$

预测值及非预测度量用来判别每帧音频信号的有调和无调

成分。模型 II 通过引入阈值计算分割区域(Threshold calculation partitions)，为阈值计算提供了一个分辨率，近似等于 FFT 谱线和 1/3 临界频带两者之间更宽的一个。这样就把频域值转化为与临界频带宽度相关的量，确定了相应频谱的感知定位。对人类听觉系统的研究表明某一音频信号的掩蔽能力取决于其频谱的感知定位和响度(Loudness)，响度是用信号的能量度量的。所以，模型 II 在每个分割区域内计算信号的能量 e_b 和加权非预测量 C_b：

$$e_b = \sum_{\omega=\omega\text{low}b}^{\omega\text{high}b} r_w^2, \qquad C_b = \sum_{\omega=\omega\text{low}b}^{\omega\text{high}b} r_w^2 C_w \tag{3}$$

其中 ωlow$b$, ωhigh$b$ 分别为分割区中的最低和最高频率线。上述两个量用来计算输入信号所在临界频带内的掩蔽阈值。同时用扩展函数 sprdngf(spreading function)来描述该信号对周围临界频带的掩蔽能力。模型 II 用分割区域能量和非预测量与扩展函数作卷积：

$$\begin{aligned} ecb_b &= \sum_{bb=1}^{b\max} e_{bb} * \text{sprdngf}(\text{bval}_{bb}, \text{bval}_b) \\ ct_b &= \sum_{bb=1}^{b\max} c_{bb} * \text{sprdngf}(\text{bval}_{bb}, \text{bval}_b) \end{aligned} \tag{4}$$

其中 bb 为分割区域标记，bmax 为分割区总数，bval_{bb} 为以 bb 为标记的分割区的平均巴克值(Bark 临界频带率的单位，一个临界频带的宽度为 1 巴克)，bval_b 为信号所在分割区的平均巴克值。因为 ct_b 是信号能量加权值，所以对 ct_b 归一化后得到 cb_b，进而得到调标志 tb_b：

$$tb_b = -0.299 - 0.43\ln(cb_b), \qquad 0<tb_b<1 \tag{5}$$

由调标志 tb_b 计算每个阈值计算分割区域所要求的信噪比 SNR_b：

$$SNR_b = \max\left(\min val_b, tb_b \times TMN_b + (1-tb_b)\times NMT_b\right) \tag{6}$$

上式中 minval_b 是控制立体声无掩蔽效应的 SNR 低限，TMN_b 是在阈值计算分割区域表中可查到的调掩蔽噪声值(dB)，NMT_b 为噪声掩蔽调值，对所有阈值分割区域都为 5.5dB。根据 SNR_b 可计算每个阈值计算分割区域的实际能量，该能量在 FFT 谱线上展开后得到该阈值分割区域的实际掩蔽阈值 nb_w。每个阈值分割区域的最后掩蔽阈值还要考虑由经验决定的绝对掩蔽阈值(Absolute masking threshold)，用 $absthr_w$ 表示，即安静时候的阈值。最后求得的每个阈值分割区域的掩蔽阈值 thr_w 为：

$$thr_w = \max\ (nb_w,\ absthr_w) \tag{7}$$

因为在高频段安静状态阈值远远大于实际掩蔽阈值，所以把绝对掩蔽阈值考虑在内，提高了心理声学模型的效果。

本文把模型 II 计算所得掩蔽阈值作为小波包分解确定最好基的代价函数。

与小波分解不同，小波包分解不仅对尺度空间 V_i 进行二剖分，而且也可以对小波空间 W_i 进行类似的二剖分。小波包分解的灵活性使得它更适于分析时变的音频信号。本文的小波包分解是由心理声学模型控制的，所以最终的子带结构更加接近临界频带，符合人耳的听觉特性。每次分解从左至右，按照当前子带掩蔽阈值从小到大的顺序进行，避免对一个频段信号的反复分解，同时考虑到时间分辨率的要求，分解自顶向下进行。若分解后新子带的掩蔽阈值提高，则进行分解，否则转向下一个子带。分解终止后，形成树形滤波器组，该滤波器组输出与当前最好基对应的系数。

在模型 II 中，阈值分割区域和 32 个子带是通过查编码分割区域表对应起来的。由于 32 个子带是等带宽的，因此在低频区域，子带带宽比该区域临界频带带宽要宽，即一个子带包含多个临界频带，而在高频区域一个临界频带带宽跨越数个子带。因此模型 II 在低频区域用该子带包含的掩蔽阈值的最小值作为子带掩蔽阈值(见图 2(c))，在高频区域则选取该子带包含的掩蔽阈值的平均值作为子带掩蔽阈值(见图 2(d))。而本文进行动态小波包分解划分子带，与模型 II 不同，每次分解对应的频谱范围与阈值分割区域动态对应。每个子带的掩蔽阈值是包含在该子带内的所有阈值计算分割区域对应掩蔽阈值的最小值，如图 2, 图 3 所示。

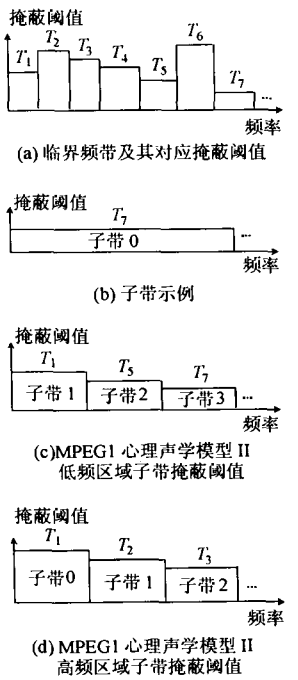


图 2　MPEG1 音频部分子带分解及阈值选取示意图

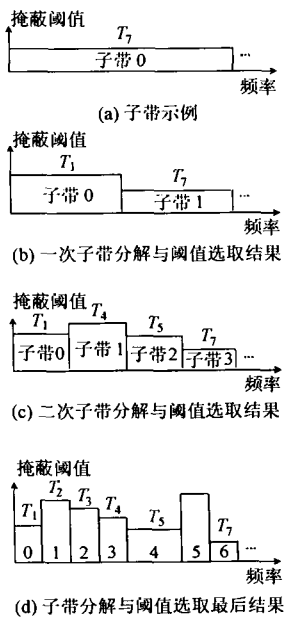


图 3　本文算法子带分解及阈值选取示意图

2.3 脆弱水印嵌入算法

2.3.1 水印信号预处理　　算法首先把二值标志图像(水印信号)变为按行或按列排列的一维 0，1 序列；其次根据二值图像的特点，对其一维序列进行行程编码(Run length encoding)，水印信号经过这样的无损压缩后，仍传递相同的

信息，但需要嵌入的数据量却成倍减少。

例如：若由二值标志图像转化的 0，1 序列为…000000111111111…，则行程编码后变为…6091…，保留相同像素的个数和第一段像素的值，变为…69…(0)。在本算法中设该像素个数为嵌入的水印单位，这样便于在提取端把压缩后的水印信号恢复为二值图像。

2.3.2 水印信号嵌入过程　设 f 为要嵌入水印的小波包系数，△为 f 对应子带的掩蔽阈值，以△为步长量化 f，得到整数量化值 q 或 Q：

$$q=\left\lfloor \frac{f}{\Delta} \right\rfloor, \quad Q=\left\lceil \frac{f}{\Delta} \right\rceil \tag{8}$$

其中运算 $n=\lfloor x \rfloor$ 表示对 x 下取整；运算 $n=\lceil x \rceil$ 表示对 x 上取整。显然，当 $Q\neq q$ 时，必有 $Q=q+1$，所以在 Q 和 q 中必有一个奇数一个偶数。设嵌入水印后的小波包系数为 f_w，当前要嵌入的水印单位为 w，按照如下步骤进行：

(1) $m=w \bmod \Delta$，$r=w-m\times\Delta$，mod 为取模运算符。

(2) (a) 若 $m=0$，则

$$f_w=\begin{cases} q\times\Delta+\dfrac{\Delta+2r}{4}, & f\geq 0 \\ Q\times\Delta-\dfrac{\Delta+2r}{4}, & f<0 \end{cases} \tag{9}$$

(b) 若 $m=1$ 且 $r=0$，

$$f_w=\begin{cases} q\times\Delta, & \text{若}q\text{为奇数} \\ (q+1)\times\Delta, & \text{若}q\text{为偶数} \end{cases} \tag{10}$$

(c) 若 $m\neq 1$ 且 $m\neq 0$，$w=w-\Delta$，则

$$f_w=\begin{cases} q\times\Delta, & \text{若}q\text{为偶数} \\ (q+1)\times\Delta, & \text{若}q\text{为奇数} \end{cases} \tag{11}$$

(3) 转向下一个小波包系数。

0 由上述嵌入过程可知，算法对小波包系数的改变量没有超过当前子带的掩蔽阈值，即小波包系数改变量 $f_\Delta=|f_w-f|\leq\Delta$，因此保证了水印嵌入是不可感知的。

对嵌入水印的小波包系数进行小波包重构，恢复嵌入水印的音频信号。

2.3.3 水印信号的抽取和篡改定位　本文算法在抽取水印时无需原始音频信号。

提取水印时首先根据保存的掩蔽阈值进行小波包分解，以保证子带划分结果与嵌入时完全一致。

假设待检测的小波包系数为 f'，抽取的水印单位为 w_i'，△为 f' 对应子带的掩蔽阈值。按照如下步骤抽取水印：

(1) 当前抽取水印单位清零：$w_i'=0$

(2) (a) 当 $f'\geq 0$ 时，$q'=\left\lfloor \dfrac{f'}{\Delta} \right\rfloor$，$u=f'-q'\times\Delta$

(b) 当 $f'<0$ 时，$q'=\left\lceil \dfrac{f'}{\Delta} \right\rceil$，$u=|f'-q'\times\Delta|$

(3) 若 $(1/4)\Delta\leq u<(3/4)\Delta$，则 $w_i'=w_i'+2(u-\Delta/4)$

(4) 若 $(3/4)\Delta\leq u<\Delta$，则 $q'=q'+1$

(5) (a) 若 $0\leq u<(1/4)\Delta$ 或 $(3/4)\Delta\leq u<\Delta$，且 q' 为奇数，则 $w_i'=w_i'+\Delta$，当前水印单位抽取完毕，$i=i+1$，转向(1)。

(b) 若 $0\leq u<(1/4)\Delta$ 或 $(3/4)\Delta\leq u<\Delta$，且 q' 为偶数，则 $w_i'=w_i'+\Delta$，转向(6)。

(6) 当前水印单位 w_i' 在下一个小波包系数继续抽取，转向(2)。

把全部抽取出的水印单位按照下标 i 升序排列，再根据保存的第一个水印单位的像素值确定每个水印单位对应的像素值，恢复水印信号的行程编码形式。进而恢复 64×64 的二值图像表示。该二值图像即为抽取出的水印。

通过比较抽取出的水印图像和原始水印图像，可确定给定音频信号是否与原始信号完全一致。

本文通过引入篡改检测函数，定位篡改区域、确定篡改程度。假设 u_i 和 u_i' 分别为在同一个小波包系数上嵌入的水印和检测出的水印，n 为嵌入水印的小波包系数的总数，p 为水印篡改程度。引入下面的篡改检测函数在频域定位篡改：$\text{tamper}(i)=\dfrac{|u_i-u_i'|}{u_i}$，则篡改程度为

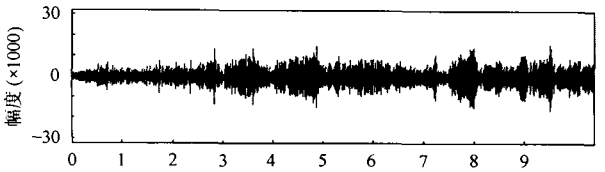
$$p=\sum_{i=1}^{n}\text{tamper}(i) \tag{12}$$

用值 $u_i'-u_i$ 代替当前小波包系数，进行小波包重构，恢复后的时域信号可在时域确定篡改位置，同时根据信号的幅值可判断篡改程度。

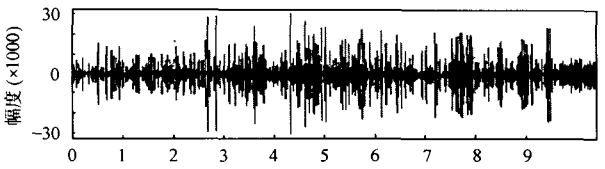
3 仿真实验结果

在 Windows XP 平台上用 Visual C++ 6.0 实现了本文算法，并对世界名曲、歌曲、器乐曲、流行音乐等多种音频信号进行实验，所选音频均为采样率44.1kHz，16位量化的PCM信号。

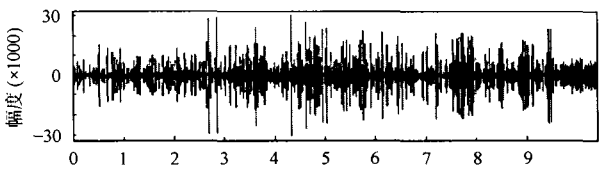
小波函数的选择综合考虑了小波的正交性、正则性、消失矩阶数和支集的大小。正交小波的正则性对应其频域局部性，支集大小对应小波的时域局部性。消失矩对应变换的快速衰减速度。一般来说，小波的时、频局部性越好，变换的效果越佳。而且变换时希望能够快速衰减。但对于正交小波来说，如果有 p 阶消失矩，那么其支集长度至少是 $2p-1$，因此消失矩的增加会直接导致支集长度的增加，而小波的支集长度对应数字滤波器的拍数，直接决定计算量的大小，同时支集的增长也会使变换的频率分辨率增加，时间分辨率减小。对应给定的消失矩阶数，Daubechies 小波具有最小的支集。所以综合考虑以上因素，实验中选取有 8 阶消失矩、支集长度为 15 的 Daubechies 小波。图 4 为部分实验结果。



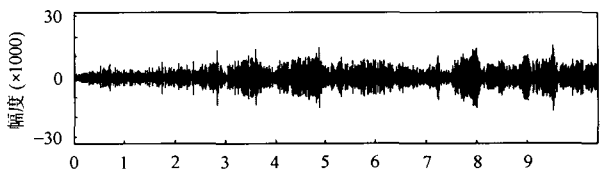
(a) 原始音频信号



(b) 嵌入水印前对 409600 个采样点进行自适应小波包分解结果



(c) 嵌入水印后的 409600 个采样点小波包系数



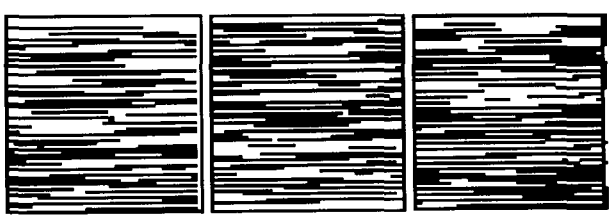
(d) 嵌入水印后恢复的音频波形图


(e) 嵌入的水印信号（二值标志图像）


(f) 本文算法提取的水印信号

图 4　水印嵌入和提取结果

对嵌入水印的音频信号进行攻击实验，分别加入白噪声(图 5 (a) SNR=28.04dB) 和随机噪声(图 5 (b) SNR=32.20dB)，进行低通滤波(低通滤波器长度为 6，截止频率为 2kHz)(图 5(c))，mp3 压缩(压缩比 12：1) (图 5(d))，重新采样(采样频率为 22.05kHz)(图 5(e))，攻击后抽取的水印如图 5 所示。


(a)　(b)　(c)

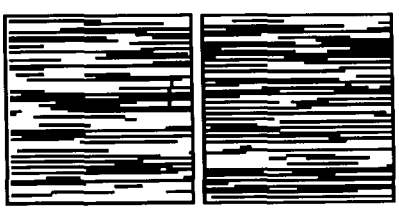
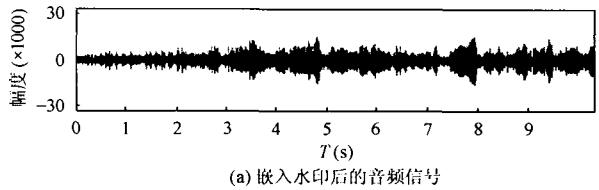
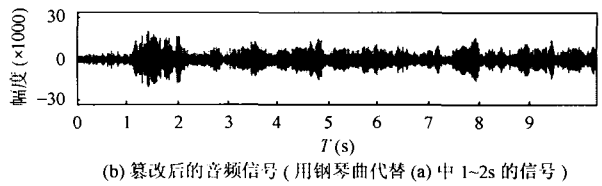

(d)　(e)

图 5　水印攻击实验结果

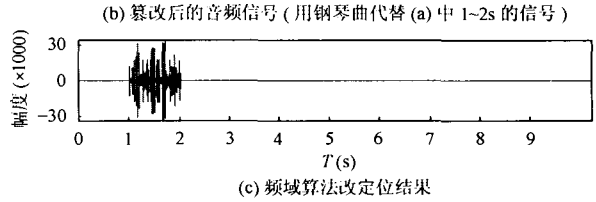
对嵌入水印的音频信号进行剪切及本算法篡改检测定位结果如图 6 所示。利用篡改检测函数对嵌入水印的音频信号进行时域和频域定位，整个音频信号的篡改程度 p=10.41%。从图 6 (c)，图 6 (d)可看出本文算法无论在时域还是在频域对篡改都有良好的定位性。



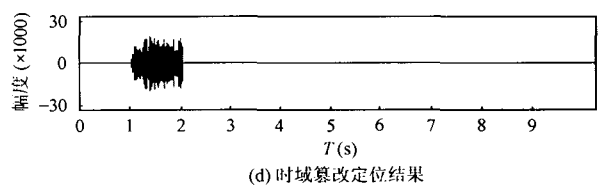
(a) 嵌入水印后的音频信号



(b) 篡改后的音频信号 (用钢琴曲代替 (a) 中 1~2s 的信号)



(c) 频域算法改定位结果



(d) 时域篡改定位结果

图 6　定位篡改实验结果

4　讨论

在本文的脆弱水印算法中，单一小波包系数根据所在子带的掩蔽阈值自适应地对应若干位水印，因此系数的改变可能导致其上嵌入的多位水印发生改变，所以本文的脆弱水印算法可以敏感地捕捉到音频信号的变化；本文所选水印是视觉易辨识的二值图像，便于在提取端判断给定音频文件是否与原始文件完全一致。但是在有些情况下，虽然对音频文件的处理使得它有所改变，但是并不能因此怀疑该文件的真实有效性，例如：为了保存或传输方便而压缩原文件。因此，可以根据实际的应用场合设定不同的阈值 τ，在 $p=\sum_{i=1}^{n}\text{tamper}(i)\geq\tau$ 时，认为对原文件的改动破坏了其真实可信性。

5　结束语

本文通过对 MPEG1 心理声学模型 II 进行改进，使之适用于小波包域，对音频信号进行自适应小波包分解。在满足算法复杂度和时间分辨率的要求下，音频信号进行接近临界频带的子带划分。本文算法得到的子带掩蔽阈值比 MPEG1

标准和小波分解得到的子带掩蔽阈值有所提高，因而可更加灵活地控制每个子带嵌入的水印量。采用量化小波包系数的方法自适应地嵌入水印，根据水印和篡改检测函数确定音频文件的可信性。从实验结果来看，本文的脆弱水印算法对多种音频处理都十分敏感，可判断音频作品是否被篡改，分析被篡改的程度。

本文算法也存在一些局限性，包括不能区分篡改的种类，也没能利用音频信号的一些统计特性，特征量来衡量音频文件的可信性。今后的工作将结合音频信号的感知特性和统计特性，进一步完善脆弱水印算法，从而对音频媒体的真实可信性给予更准确的判断。

参 考 文 献

[1] Lu Chunshien, Liao Hongyuan, Chen Lianghua. Multipurpose audio watermarking. in Proc.15th International Conference on Pattern Recognition, Barcelona Spain, Sep. 2000, 3: 3286 – 3289.

[2] Kundur D, Hatzinakos D. Digital watermarking for telltale tamper proofing and authentication. *Proc. of the IEEE*, 1999, 87(7): 1167 – 1180.

[3] Podilchuk C I, Delp E J. Digital watermarking: algorithms and applications. *IEEE Signal Processing Magazine*, 2001, 18(4): 33 – 46.

[4] Gordy J D, Bruton L T. Performance evaluation of digital audio watermarking algorithms, Proceedings of the 43rd IEEE Midwest Symposium on Circuits and Systems, Aug. 2000, 1: 456 – 459.

[5] Pan D Y. Digital audio compression. *Digital Technical Journal*, 1993, 5(2): 1 – 14.

[6] ISO/IEC IS11172-3, Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s—Part 3: Audio. 1992.

[7] Pan D. A tutorial on MPEG/audio compression. *IEEE Multimedia*, 1995, 2(2): 60 – 74.

全笑梅： 女， 1975 年生，博士生，主要从事数字水印、音频压缩方面的研究.

张鸿宾： 男， 1944 年生， 教授，博士生导师，主要从事模式识别、图像处理、数字水印和多媒体数据认证等领域的研究.