鲁棒的混合域音频信息隐藏算法

由守杰[1],柏 森[1,2],曾 辉[3]
(1.重庆通信学院,重庆 400035;2.重庆大学,重庆 400044;
3.西藏军区77668部队,山南 851200)

摘 要:提出了一种基于离散小波变换(DWT)和离散余弦变换(DCT)的音频信息隐藏算法。对载体音频整体进行小波分解,将其划分成若干频带,然后根据人耳听觉的频率掩蔽效应,选择对人耳听觉不敏感的频带所对应的小波系数,将小波系数分段进行离散余弦变换,将水印嵌入到DCT直流系数上。实验表明,嵌入水印后的音频文件不仅具有良好的不可感知性,而且对诸如噪声、低通滤波、重采样、回声和Mp3压缩等的攻击具有很强的鲁棒性,并能抵抗一定程度的样点裁剪攻击,算法的鲁棒性和不可感知性达到了很好的平衡。

关键词:信息隐藏;音频盲水印;DWT;DCT;鲁棒性
中图分类号:TP391 **文献标识码**:A **文章编号**:1673-629X(2008)03-0169-04

Robust Audio Information Hiding Algorithm Based on DWT and DCT

YOU Shou-jie[1], BAI Sen[1,2], ZENG Hui[3]
(1. Chongqing Communication Institute, Chongqing 400035, China;
2. Chongqing University, Chongqing 400044, China;
3. 77668 Army of Tibet Military Area Command, Shannan 851200, China)

Abstract: A new audio information hiding algorithm based on the DWT and DCT domain is proposed. First the entire host audio is decomposed by wavelet to several frequency bands, then select a frequency band which is insensitivity to the human ear hearing and its corresponding wavelet coefficients to hide information. Those coefficients are separated to some segments, and at the same time convert those segments to DCT domain. The watermark can be embedded to the DCT direct current coefficients. Experimental results show that the watermarked audio has good imperceptibility and is robust against different kinds of attacks, such as samples cropping, noise adding, low-pass filtering, resampling, echo, Mp3 compression. The robustness and the imperceptibility about watermarked audio reach a good balance.

Key words: information hiding; blind audio watermarking; DWT; DCT; robustness

0 引 言

信息隐藏技术是研究将某一秘密信息隐藏于公开信息中,通过公开信息来传递秘密信息的技术。因为含有秘密信息的媒体是公开的,并且该媒体和原始媒体具有很高的相似性,使得可能的检测者难以判断该公开信息是否含有秘密信息,更难以截获秘密信息,从而达到了保证秘密信息安全传递的目的。音频信息隐藏技术的应用主要是两方面,即通过嵌入数字水印实现版权保护和利用音频文件的冗余空间嵌入信息实现隐蔽通信。网络技术和多媒体技术的飞速发展,使得音频产品的版权保护问题迫在眉睫,而以音频为载体进行隐蔽通信在诸如军事等方面有广泛的应用前景,近几年信息隐藏技术因此发展迅速。以音频为载体的经典的信息隐藏算法基本可以分为两类:时域算法和变换域算法。时域算法的主要代表有LSB算法[1],回声隐藏[2],其算法比较简单,但是鲁棒性差,当然也有一些算法具有很好的鲁棒性,如时域能量算法[3]等;变换域算法如在DFT[4],离散余弦变换DCT[5],离散小波变换DWT[6,7]等变换域中通过改变其变换域的系数来嵌入秘密信息,因此具有较好的鲁棒性。

DCT域信息隐藏算法计算量比较小,现在研究的比较多。文献[5]是对音频分段进行离散余弦变换,通过DCT噪声信号模型,定义DCT系数的噪声敏感度,

收稿日期:2007-06-20
基金项目:国家自然科学基金资助项目(6067215);重庆市自然科学基金资助项目(CSTC 2005BB2208,CSTC 2005BB2210)
作者简介:由守杰(1982-),男,山东淄博人,硕士研究生,研究方向为音频水印、音频文件密写分析;柏 森,博士后,教授,硕士研究生导师,主要研究方向为信息隐藏、掩密通信、图像处理、模式识别。

建立水印嵌入位置和嵌入水印后的音频信号的听觉感知性之间的关系，根据音频水印的不可感知性的要求选择最优的嵌入系数，然后调节水印强度来满足鲁棒性的要求，从而保持音频水印的不可感知性和鲁棒性。小波理论采用多分辨率分析的思想，非均匀地划分时频空间，为非平稳信号的分析提供了新的途径，受到了人们的普遍重视。文献[6]提出了一种基于小波变换的语音信息隐藏算法，算法是对语音分段，在各段小波变换的高频系数中嵌入密文信息，获得了很好的隐藏效果。而文献[7]充分利用了 DCT 和 DWT 的优点，提出一种混合域盲水印算法，具有很好的鲁棒性。水印信息是利用音频文件的冗余空间嵌入的，而以上大部分算法都是首先利用某种规则将音频分段，然后逐段嵌入水印，这样就将音频文件分成了含有水印部分和不含有水印部分。这种算法的缺点首先是没有充分利用整个音频文件的冗余空间；其次由于分段以后，在提取信息过程中需要有精确的同步信息才能保证较高的水印信息提取正确率，因此含有水印的音频如果受到同步攻击，同步信息将会丢失，造成水印提取正确率大幅度降低甚至完全无法正确提取。针对这些缺点，如果能充分利用整个音频文件的冗余空间，算法鲁棒性和不可感知性会达到一个更好的平衡。

文中提出的算法，是对整个音频文件进行小波分解，将水印嵌入到对人耳听觉不敏感的小波系数上，这样就更充分地利用了整个音频文件的冗余空间。实验结果表明，算法不可感知性良好，与文献[7]的混合域算法相比，文中算法鲁棒性有了较大提高，特别的，该算法还具有抵抗一定程度去同步攻击的能力。

1 音频文件小波分解和水印嵌入频带选择

人耳的听觉频率范围为 20Hz～20kHz，由人耳掩蔽效应可以知道人耳对不同频率信号具有不同的听觉敏感性[8]，由图 1 安静时人耳听阈曲线可以看出，在大约 2kHz～4kHz 的范围内，人耳对于这个频率的声音最为敏感。而在直流附近的 0Hz～150Hz 以及高频靠近 20kHz 附近的频率成分，则人耳听觉最不敏感。因此相同的水印算法，在嵌入强度相当时，选择对人耳不敏感的频率成分作为水印信息的嵌入位置比选择对人耳敏感的频率成分会有更好的不可感知性。而从鲁棒性上讲，后者由于对人耳听觉起主要影响作用，低通滤波等信号处理都会避开这些频率成分，尽量不改变或最少改变这些频率成分，因此具有更好的稳健性。对于这些不敏感的频率成分，高频部分非常容易受这些信号处理的影响，而低频部分对于低通滤波、重采样等信号处理影响比较小。为了使水印音频鲁棒性和不可感知性达到比较好的平衡，可以选择人耳听觉不敏感的低频段作为嵌入位置。

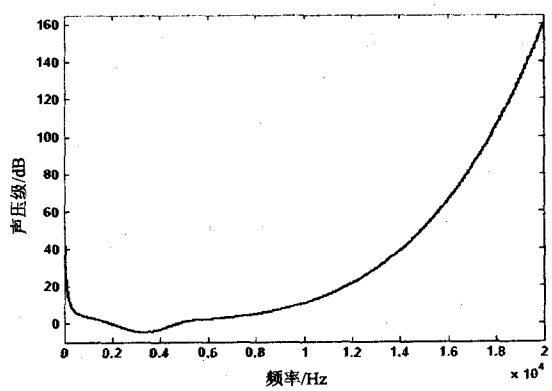

图 1 安静时人耳听阈曲线

人耳频率分辨率是非线性的，使得传统的线性信号处理方法，如傅里叶变换来模拟人耳听觉特性分析声音是比较困难的。而小波的多分辨率分析的思想可以很好分析音频等非平稳信号。利用小波变换可以将音频划分成若干个频带，以抽样速率为 8kHz 的语音为例，对语音信号进行多级小波分解，频带划分如图 2 所示。

因为语音带宽为 4kHz，进行小波变换分解级数为 5 时，则子带最小宽度为 125Hz。由人耳听阈曲线可知，人耳对在 0～125Hz 的这个频带的声音最不敏感。以上分析可知，为了使水印音频鲁棒性和不可感知性达到较好的平衡，可以选择这个频带作为水印嵌入位置。

2 水印算法

从以上分析可知，对音频文件进行小波分解，分解

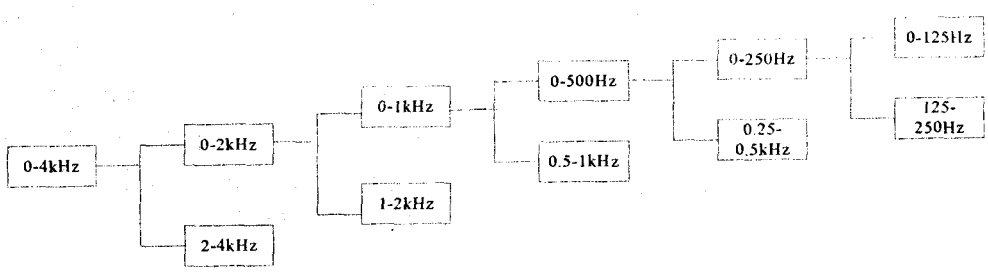

图 2 对语音信号的小波分解的频带划分示意图

到合适的频带宽度,根据人耳的听觉掩蔽效应和各种鲁棒性攻击的特点,选择小波划分的最低频的频带作为水印嵌入位置,将会使水印的鲁棒性和不可感知性达到较好的平衡。

2.1 水印信号的预处理

文中算法可以将任何一维的水印信号嵌入到载体音频中,但是为了更直观地比较原始水印信号和提取的水印信号,文中采用一幅大小为 $M_1 \times M_2$ 的二值图像 W,该图像可以表示为

$$W = [w(i,j)]_{M_1 \times M_2} \tag{1}$$

其中 $w(i,j) \in \{0,1\}$,表示第(i,j) 个像素的值。

为将二维信号嵌入到载体音频文件中,对图像进行降维操作,使其转换为一维序列。具体操作为:

$$B = (F(1),F(2),\cdots,F((i-1) \times M_2 + j),\cdots, F(M_1 \times M_2)) \tag{2}$$

其中 $F((i-1) \times M_2 + j) = w(i,j), 1 \leqslant i \leqslant M_1, 1 \leqslant j \leqslant M_2$

为了去掉该序列的相邻元素的相关性,并且进一步增加水印信息的安全性,对该序列进行随机化处理,方法如下:

$$B_{rand} = (m(k) = \mathrm{XOR}(F(k),C(k))) \tag{3}$$

其中 $C(k)$ 为一个由种子生成的 0,1 伪随机序列,种子可作为私钥,进一步保证了嵌入的信息的安全性。XOR 表示两个序列相同位置的数值进行异或。

2.2 水印嵌入算法

假设给定的载体音频信号为 $x(n), n = 0,1,\cdots, N_1 - 1$。其中 N_1 表示该载体音频信号的总采样点个数,$x(n)$ 则表示音频信号第 n 个采样点的幅值。

Step1:为了更好利用整个音频文件的冗余空间嵌入信息,不对音频文件实行分段处理,而是使用“db4”小波对载体音频序列整体进行五级离散小波变换,并提取分解的第五级低频小波系数 ca5(r),$0 \leqslant r \leqslant N_2$ 作为水印嵌入部分。小波基的选择对于算法鲁棒性的影响,实验表明“db4”小波要比“db1”小波效果要好。

Step2:根据水印信号的长度,将提取出的低频小波系数进行分段,每段长度为 N_3 个小波系数,分段总数为 $M_1 \times M_2$。小波系数分段的集合为

$$R = \{p_k(t) = \mathrm{ca5}((k-1) \times N_3 + t), 0 \leqslant t < N_3\} \tag{4}$$

对分出的每一系数段再进行离散余弦变换,变换后的 DCT 系数集合为

$$S = \{T_k(t) = \mathrm{DCT}(P_k(t)), 0 \leqslant k < (M_1 \times M_2)\} \tag{5}$$

Step3:对 $M_1 \times M_2$ 个 DCT 系数段,提取每一段的第一个系数,即直流系数,为了更方便嵌入水印信息,首先将直流系数值置为零,然后根据水印信息,按照以下规则嵌入水印信息:

$$P_k(0) = \begin{cases} T_k(0) = \Delta, m(k) = 1 \\ T_k(0) = -\Delta, m(k) = 0 \end{cases} \tag{6}$$

其中 Δ 为嵌入强度,其值大小要保证水印透明性。

Step4:将嵌入水印后的 DCT 系数段逐段进行反离散余弦变换,变换后的系数和未改变的小波系数按照原始顺序合并,得到新的低频小波系数。

Step5:利用新的低频小波系数和其它频带未改变的小波系数进行反离散小波变换,对信号进行重构,得到含有水印的音频信号。

2.3 水印提取算法

水印的提取不需要原始公开音频,属于盲水印算法,提取算法是嵌入算法的逆过程,具体步骤为:

Step1:使用“db4”小波基对待检测音频序列整体进行五级离散小波变换,并提取分解的第五级低频小波系数 ca5(r),$0 \leqslant r \leqslant N_2$。

Step2:将提取出的低频小波系数进行分段,每段长度为 N_3 个小波系数,分段总数为 $M_1 \times M_2$,对分出的每一系数段再进行离散余弦变换,并提取每一段 DCT 系数的直流分量。

Step3:根据以下规则提取水印,若该段 DCT 直流系数大于 0,则提取比特“1”;相反,提取比特“0”,将提取的比特按照顺序合并得到提取的水印序列

$$M = (m_1(1),m_1(2),\cdots,m_1(k),\cdots,m_1(M_1 \times M_2)) \tag{7}$$

Step4:利用种子产生伪随机序列 $C(k)$,将提取的水印序列与伪随机序列进行异或,得到序列 $E(k)$。

Step5:对 $E(k)$ 进行升维操作,得到水印信号

$$G = [g(i,j) = E(k)]_{M_1 \times M_2} \tag{8}$$

其中 $k = (i-1) \times M_2 + j, 1 \leqslant i \leqslant M_1, 1 \leqslant j \leqslant M_2$。

为了消除主观因素的影响,采用归一化相关系数对提取的水印和原始水印进行客观评价,归一化相关系数定义为:

$$\rho(W,G) = \frac{\sum_{i=1}^{M_1}\sum_{j=1}^{M_2} w(i,j)g(i,j)}{\sqrt{\sum_{i=1}^{M_1}\sum_{j=1}^{M_2} w^2(i,j)} \times \sqrt{\sum_{i=1}^{M_1}\sum_{j=1}^{M_2} g^2(i,j)}} \tag{9}$$

3 实验结果

在实验中,以一段长度为 10 秒,抽样速率为 44.

1kHz,16 位量化的流行英文歌曲为公开信息,在该段语音中嵌入水印,采用 32×32 的二值图像(如图 4 所示)作为水印,根据上述算法,在低频小波系数分段阶段,取 $N_3 = 8$,嵌入强度 $\Delta = 0.05$。

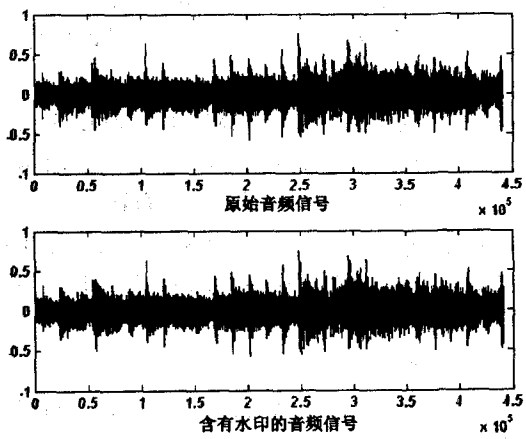

图 3 原始音频与含有水印音频时域波形

图 4 原始水印图像

图 3 是嵌入水印的音频文件和原始文件的时域波形图,从波形图上看两者相似性很高。利用主观评价方法测试算法的隐蔽性,事先不告知参与测试人哪个是原始音频和水印音频,将水印音频和原始音频播放给 5 个人听,这 5 人都辨别不出二者有何区别。尽管音频文件信噪比较低,但是由于水印嵌入点选择的是对人耳听觉很不敏感的频带,因此嵌入水印后的音频与原始音频听起来几乎没有差别。这就比较充分地利用了整个音频文件的冗余空间,在不可感知性条件下,提高了算法的鲁棒性。这也说明利用信噪比评价算法的隐蔽性有一定局限性。为了检验水印的鲁棒性,对音频水印文件做以下各项信号处理:

(1)低通滤波:低通滤波器为长度为 8 阶、截止频率为 1.5kHz 的巴特沃思低通滤波器。

(2)加高斯白噪声:均值 $\mu=0$,均方差 $\sigma=0.02$。

(3)回响:将原始信号的延时拷贝叠加到原始信号之上,延时时间为 400ms,延时信号幅度为原始信号 10%。

(4)重采样:将音频文件用 22050Hz 重采样,再重新用 44100Hz 重新采样恢复原始音频。

(5)Mp3 压缩:将含水印音频压缩为 Mp3 文件,然后恢复成与原始音频格式相同的音频,压缩比分别为 11:1, 22.1:1。

(6)重量化:将含水印音频从 16bit 量化为 8bit,再重新量化为 16bit。

(7)同步攻击:在含有水印音频中随机选取若干个采样点,然后裁剪掉。裁剪的样点个数分别是 100、300、500 个。

根据(9)式分别计算各种攻击下提取的水印图像与原始水印图像的相关系数,结果如表 1 所示。

从实验结果看,文中提出的混合域音频信息隐藏算法能抵抗噪声、低通滤波、重采样、回声和 Mp3 压缩等的攻击,特别是在抗 Mp3 压缩攻击方面,压缩比达到 22.1:1 时,几乎还能完全正确提取水印,表现出了非常强的鲁棒性,与文献[7]提出的混合域水印算法相比,鲁棒性有了比较明显的提高。另外水印文件的不可感知性也比较好,这样就使鲁棒性和不可感知性达到了比较好的平衡。

4 结 论

提出了一种新的混合域信息隐藏算法,使用小波将音频文件整体进行分解,将其划分成若干频带,利用人耳听觉系统的频率掩蔽效应分析确定了水印的嵌入频带,即选择对听觉不敏感的低频带作为嵌入频带。然后对所选频带小波系数进行分段 DCT 变换,将水印嵌入到每段 DCT 直流系数上,这样就比较充分利用了整个音频文件的冗余空间,使鲁棒性和不可感知性达到了比较好的平衡。实验表明,文中算法能够抵抗噪声、低通滤波、重采样、回声和 Mp3 压缩等的攻击,与

表 1 鲁棒性攻击检测结果及算法比较

算法	攻击方法	未攻击	低通滤波	加白噪声	重量化	重采样	Mp3 (11:1)	Mp3 (22:1)	裁剪样点 (100)	裁剪样点 (300)	裁剪样点 (500)	回响
文中	相似系数	1	1	1	1	1	0.9986	0.9926	0.9351	0.8333	0.7663	0.9926
	提取出的水印图像											
文献[7]	相似系数	1	0.9308	0.9605	0.9913	0.7901	0.9701	0.8074	未提	未提	未提	未提
	提取出的水印图像								未提	未提	未提	未提

(下转第 176 页)

息数据包后，PeerIDS 包装器实例将把附于其中的 Snort 规则添加到本地 Snort 系统的配置文件中(通常是 snort.conf，在这里起到的是活动工作集列表 Wa 的作用)并回应发送者一条确认信息。随后再通过执行命令 snort -d -c snort.conf 重新启动本地 Snort 系统。若因为某种原因运行中的 Snort 进程占用了超出配额的资源，则 PeerIDS 包装器将根据其超出的比例把 snort.conf 中的部分规则移至一个起到 Ws 的作用的本地文件如 suspended.conf 中并重启 Snort 系统以降低其负载。随后在执行 PUSH 操作时，PeerIDS 包装器实例把 suspended.conf 中的一条已挂起的 Snort 规则装载入一个 Workload 数据包发送至一个同伴并在收到该同伴的确认信息后删除该规则的本地拷贝。值得注意的是在入侵检测对等网中，各 PeerIDS 实例所包装的 Snort 系统所执行的规则应相互正交而没有重叠。为做到这一点可以先清空各 Snort 安装的规则集，当 PeerIDS 对等网建立起来后，全部的入侵检测规则可以通过任一 PeerIDS 实例注入对等网中并在其中自动分布。同样地，为检测新种类的网络入侵也只需把相应的 Snort 规则添加至网络中任何一个 Snort 进程的配置文件 snort.conf 中，系统的自动负载平衡功能将会把该规则移动到合适的 PeerIDS 实例上执行。

PeerIDS 系统运行所需的参数如：状态检测模块两次操作间的时间间隔、收到拒绝消息直到再次发起同一操作的时间间隔、同伴间通信的超时设置等均和系统运行的网络及软硬件环境相关，在 PeerIDS 的实现中这些值将是可配置的。对于文中多次提及的 PeerIDS 实例上的执行资源进行定量的一种很自然的方法是为其入侵检测引擎的 CPU 占用率设置上限(超过即 FEDUP)和下限(不足即 HUNGRY)以表明其运行时的负载。这两个参数值的设置取决于运行 PeerIDS 实例的计算机的性能，可以通过在启动时运行一段特定的性能测试程序并根据其运行结果由 PeerIDS 系统自动设置。

IP Head	UDP Head	Source pid	A piece of Snort rule

图 3 一种用于 Snort 包装器的 Workload 消息

6 结 论

通过把网络入侵检测工作的各正交子集分布到点对点对等网络中，PeerIDS 系统具有较高的可靠性。同时由于整合了众多的 PeerIDS 实例，系统提供了强大的入侵检测性能。PeerIDS 系统的可扩展性主要表现在：只要遵守与系统中运行的入侵检测 Workload 相一致的接口，对新种类网络攻击进行检测的工作将很容易加入到系统中来。

为提高系统的通讯效率，未来拟在 PeerIDS 的通讯中采用认证机制来强化其安全性并将通过实现名声(Reputation)机制和在 Peer List 中根据各 Peer 的 Reputation 和 Activeness 应用优先级列表，具体仍待后续研究。

参考文献：

[1] 董晓梅，王丽娜，于 戈. 分布式入侵检测系统综述[J]. 计算机科学，2002，29(3)：16－19.

[2] 彭志豪，李冠羽. 分布式入侵检测系统研究综述[J]. 微电子学与计算机，2006，23(9)：191－196.

[3] 蔡 晟，王泽兵，冯 雁，等. 基于 Super－Peer 的对等网络研究[J]. 计算机应用研究，2004，21(6)：258－260.

[4] 王伟平，罗 熹，王建新. SACS：一种可扩展的匿名通信系统[J]. 小型微型计算机系统，2007，28(2)：237－242.

(上接第 172 页)

文献[7]算法比较，鲁棒性有了明显的提高。

文中提出的信息隐藏算法不能抵抗同步攻击和 D/A，A/D 转换操作，这是下一步算法研究改进的重点。

参考文献：

[1] Cooperman M, Moskowitz S. Steganographic Method and Device[P]. USA: [s.n.], 1997.

[2] Kim H J, Choi Y H. A novel echo hiding scheme with backward and forward kernels[J]. IEEE Trans. Circuits Syst. Video Technol., 2003, 13: 885－889.

[3] Lie Wen－Nung, Chang Li－Chun. Robust and high－quality time－domain audio watermarking based on low－frequency amplitude modification[J]. IEEE Transactions on Multimedia, 2006, 8(1): 46－59.

[4] Seok Jong won, Hong Jin woo, Kim Jin woong. A novel audio watermarking algorithm for copyright protection of digital audio[J]. ETRI Journal, 2002, 24(3): 181－189.

[5] 马翼平，韩纪庆. DCT 域音频水印：嵌入对策和算法[J]. 电子学报，2006，34(7)：1260－1264.

[6] 何 琴，邹华兴，白 剑. 基于小波变换的语音信息隐藏算法[J]. 计算机应用研究，2005(12)：118－119.

[7] 王向红，赵 红，崔永瑞. 一种新的混合域自适应数字音频水印算法[J]. 小型微型计算机系统，2006，27(2)：316－319.

[8] 韩纪庆，张 磊，郑铁然. 语音信号处理[M]. 北京：清华大学出版社，2004：27－30.