

面向空气信道传播的 DCT 域音频数字水印算法

张霞¹, 常迪^{2,3}, 黄倩², 杨婉怡², 郭伟²

(1. 江汉大学物理与信息工程学院, 湖北武汉 430056;

2. 武汉大学计算机学院, 湖北武汉 430072;

3. 武汉大学国家多媒体软件工程技术研究中心, 湖北武汉 430072)

摘要:提出了一种基于双 DCT 变换的音频数字水印算法, 通过选取首次 DCT 变换后的 DCT 域低频系数, 对其再次进行 DCT 变换, 得到双 DCT 域系数. 通过修改双 DCT 域系数, 实现水印的嵌入, 并且嵌入量较大. 实验结果显示, 该算法隐藏效果好, 对 30% 的重采样攻击和 MP3 压缩攻击具有很好的鲁棒性. 而且, 该算法能够适用于面向空气信道传播的场景, 误码率可控制在 1.2%~6.5% 之间.

关键词:数字水印; AD/DA 变换; DCT 变换; 同步码

中图分类号: TP391 **文献标识码:** A doi:10.3969/j.issn.0253-2778.2011.07.012

An audio digital watermarking algorithm in DCT domain for air-channel transmitting

ZHANG Xia¹, CHANG Di^{2,3}, HUANG Qian², YANG Wanyi², GUO Wei²

(1. School of Physics & Information Engineering, Jiangnan University, Wuhan 430056, China;

2. School of Computer, Wuhan University, Wuhan 430072, China;

3. National Engineering Research Center for Multimedia Software, Wuhan University, Wuhan 430072, China)

Abstract: An audio sonic watermarking algorithm was proposed based on double DCT transform. The low-frequency coefficients after the first DCT transform were chosen for the second DCT transform. Using the method of modifying the coefficients in the double DCT domain, watermarking was embedded and in large quantities. Experimental results indicate that the proposed algorithm has a good hidden effect, and it has very good robustness, especially to 30% resample and MP3 compression attacks. Moreover, it can be applied to transmitting situation via air channel, with bit error ranging from 1.2% to 6.5%.

Key words: digital watermarking; AD/DA conversion; DCT transform; synchronization signal

0 引言

随着信息技术的飞速发展, 数字媒体原创者的

版权和经济利益如何得到保护, 数字音频是否安全可靠, 由此引发的信息安全问题已成为日益严重的社会问题. 数字水印技术作为隐蔽通信和音频信息

收稿日期: 2011-05-04; 修回日期: 2011-07-04

基金项目: 国家科技重大专项(2010ZX03004-003-03), 国家自然科学基金重点项目(60832002), 湖北省自然科学基金(2009CDB222, 2010CDB08602), 湖北省教育厅科学技术研究计划指导性项目(B20114505), 教育部博士点基金(20090141110054), 中央高校基本科研业务费专项资金(6081012), 武汉大学博士生自主科研项目(20082160101000050, 20102110101000099)资助.

作者简介: 张霞(通讯作者), 女, 1964年生, 副教授. 研究方向: 数字水印, 信息安全, 信息隐写. E-mail: zx1964@163.com

安全保护的主要技术手段,正得到广泛研究与应用.数字音频水印的主要性能指标包括:不可觉察性、稳健性和嵌入量.三者相互牵制,相互依存又互相矛盾.在实际应用中往往需要根据具体需求在三者之间寻求一个恰当的平衡.

当前,空气信道传播领域的音频盗版产品主要是从影院、音乐厅等公共场所经由偷录等方式而产生的,这使得通过空气传播的音频水印具有很大的研究价值.但是在空气传播中,除了经过复杂的 A/D, D/A 变换外,音频水印还将受到空气噪声、播放源音量及距离、录音设备所在方位等影响.此外,音频作品还可能受到重采样攻击、格式转换攻击(如 MP3 压缩)等攻击.这些问题都给水印算法性能的提高造成很大的技术困难.

早期的音频数字水印都是基于时间域的水印,例如一种最简单的时间域嵌入方法—最低有效位方法(LSB),此算法利用原始数据的最低几位嵌入水印,简单易行,但是鲁棒性很差,无法经受一些信号处理的操作.变换域算法通过修改音频的变换系数来隐藏水印,主要常用的变换有离散余弦变换(DCT)、离散傅里叶变换(DFT)、改进离散余弦变换(MDCT)、离散小波变换(DWT)等,还有一些其他的变换方法.变换域算法具有许多空域水印技术不具备的优点,最突出的一点是其稳健性得到加强.

在面向重采样攻击方面,文献[1]提出的算法能很好地抵抗重采样攻击,然而这个水印算法却不能抵抗加性噪声攻击,例如 MP3 压缩和低通滤波攻击.当然有些算法(例如文献[2]中提出的算法)能够抵抗加噪攻击,但是却不能抵抗重采样攻击.文献[3]利用音频的低频子带中的极大值点来定位水印的嵌入区域,在定位点前后取长度为 4096 样本的音频段进行 FFT 变换,并交换 DFT 域的低频子带系数来嵌入水印.这样水印信息仅仅嵌入到音频信号的小部分样本中,即定位点前后的样本中,不是均匀分布在整个音频信号之中.文献[4]还通过在音频不同部分重复嵌入同一水印来降低水印的误码率,以抵抗很强的几何变换攻击.文献[5]利用边信息来解决基于 Resample 模式下的攻击问题,由于边信息的引入而大大降低了水印检测的计算量.

面向数字信道传播方面的音频数字水印的研究相对比较成熟,但由于音频信号在空气中传播会受到很多方面的攻击,因此,相比于数字信道方面,对空气传播方面的音频水印的研究则要少得多.最早

德国的 Steinbach 等^[6]开展了研究,通过设定 5~400 cm 等多个不同间距,同时使用了 4 种不同的麦克风,研究了 5 种音频类型的水印技术,但实验容量太小,实用性不够;随后,日本的 Tachibana 等^[7]研究将水印实时地隐藏到公共环境如音乐演奏会的音乐之中,能够在 30 s 音乐片段内隐藏 64 bit 的消息,但提取效果一般.在直通电缆传播方面,项世军等^[8]采用了三段能量比值方法,嵌入的是一串 32 bit 信息,虽然提取效果比较好,但由于实验容量太小,实用性不够,且对同步技术有较高的要求.

可见,在面向公共音频传播方面,如何既能提高隐藏效果又能增大容量,仍然是音频水印算法要解决的一个难题^[9].本文中,我们提出了一种面向空气传播的,且能够有效地抵抗多方面攻击的音频水印算法.

1 空气信道传播对音频水印的影响

数字音频信号经过空气信道传播过程时,如图 1 所示,将会受到许多的影响,主要需要考虑有以下几个方面问题:

- ① D/A 处理过程中的噪声;
- ② 音频能量幅度的改变,这要求水印算法能够抵抗音量的大范围变化;
- ③ 空气信道背景噪声与能量的变化;
- ④ A/D 处理过程中的量化失真与放大器引入的噪声;
- ⑤ 空气传播时,由于音频录制时刻的不确定性,所以信号中必须设有同步信号;
- ⑥ 转换过程具有一定的滤波特点,可滤除较高频率信号.

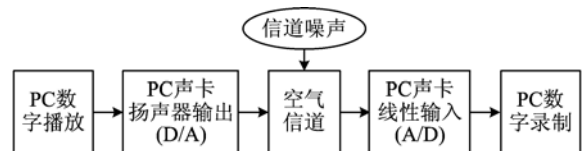


图 1 音频在空气信道传播的实验模型

Fig. 1 Audio signal's transmitting model via air channel

2 算法基本原理及流程

系统总体算法流程如图 2 所示,其核心算法分为 3 部分:水印嵌入算法、同步码嵌入算法,以及水印提取算法.以下,笔者将分别介绍各部分算法,并从整体上对算法进行分析.

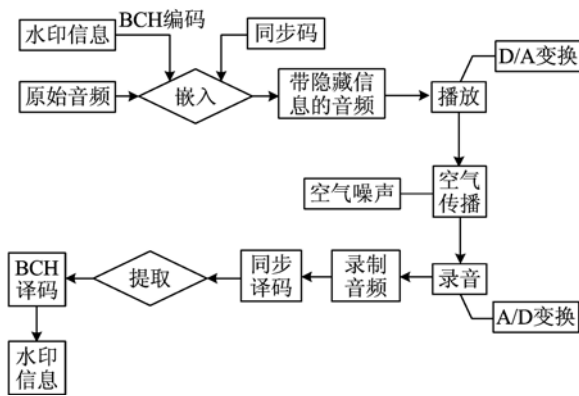


图 2 总体算法流程图

Fig. 2 Diagram of watermarking algorithm

2.1 水印嵌入算法

将水印信息进行 BCH 编码, 然后基于双 DCT 变换, 将编码后的水印信息嵌入到音频信号的变换域中, 并进行逆双 DCT 变换, 形成带水印信息的音频信号.

音频水印算法的分类方式较多, 从实现角度看可分为时域算法和变换域算法, 基于变换域的方法更具稳健性. 基于变换域嵌入的水印算法的核心思想, 是要寻找一个特定的变换域, 使得嵌入水印前后的原始信息, 在时域上的变化很小, 而在特定的变换域上变化较大, 且攻击对该变换域中的系数影响较小. 经大量实验分析, 作者推导出双 DCT 域能够有效抵抗空气噪声攻击及其他普适性攻击, 因此本文中的算法选择的是双 DCT 域. 嵌入算法的总体流程如图 3 所示.

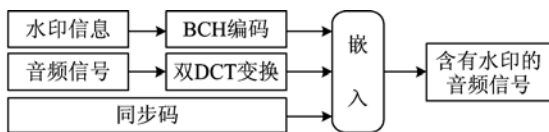


图 3 水印嵌入流程图

Fig. 3 Diagram of embedding algorithm

DCT 变换(discrete cosine transform, 离散余弦变换)是与傅里叶变换相关的一种变换, 它类似于离散傅里叶变换, 但是只使用实数. 本文中所采用的 DCT 变换公式如式(1)所示.

$$f_m = \sum_{k=0}^{n-1} x_k \cos \left[\frac{\pi}{n} m \left(k + \frac{1}{2} \right) \right] \quad (1)$$

本文中所采用的 IDCT 变换(inverse discrete cosine transform, 逆离散余弦变换)公式如式(2)所示.

$$f_m = \frac{1}{2} x_0 + \sum_{k=1}^{n-1} x_k \cos \left[\frac{\pi}{n} m \left(k + \frac{1}{2} \right) \right] \quad (2)$$

将原始水印信息 $data(i), 0 \leq i \leq \text{length}(data)$ 进行 BCH 编码, 得到编码后的水印信息 $code(i), 1 \leq i \leq \text{length}(code)$, 将其嵌入到音频的双 DCT 域中, 再进行逆双 DCT 变换, 即可得到带水印音频. 具体过程如下:

对原始音频信号进行分帧, 每 2 048 个采样点为一帧, 记为 N , 对 N 做首次 DCT 变换, 即可得到 N 的频域系数, 记作 N' , 如图 4 所示.

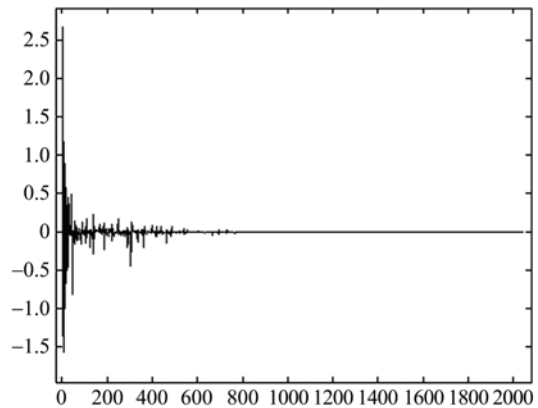


图 4 N' 的系数图

Fig. 4 Coefficients of N'

设定水印嵌入系数 S, T , 根据采样率 FS , 计算出 $S' = \lfloor \frac{2048 \cdot S}{FS} \rfloor, T' = \lfloor \frac{2048 \cdot T}{FS} \rfloor$, 其中, $\lfloor x \rfloor$ 表示对 x 进行取整. 取出 N' 中的部分系数: $P = N'(i), S' \leq i \leq T'$, 对 P 进行第 2 次 DCT 变换, 得到 P' , 如图 5 所示.

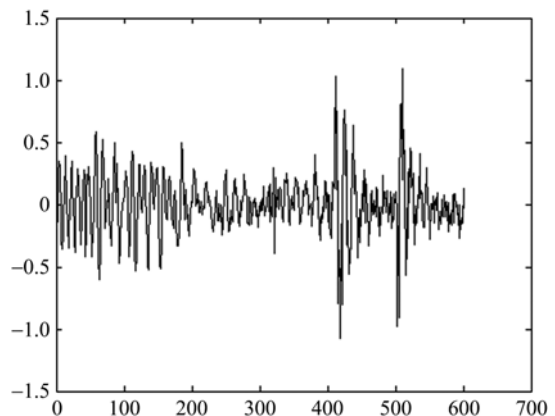


图 5 P' 的系数图

Fig. 5 Coefficients of P'

水印嵌入调整系数为 λ , 水印嵌入分割点 $\text{mid} =$

$\left\lfloor \frac{S'-T'}{2} \right\rfloor$, 经编码后的水印信息为 $code(i)$, 具体水印嵌入策略如下:

①若 $code(i)=0$, 则如式(3)修改系数, 结果如图 6 所示.

$$P'(j) = \frac{P'(j)}{\lambda}, 1 \leq j \leq mid \quad (3)$$

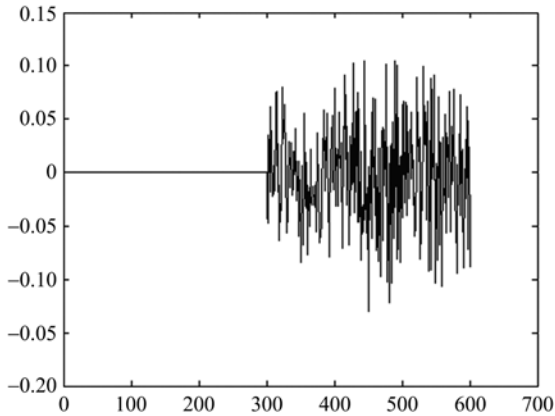


图 6 水印为“0”的双 DCT 域系数图

Fig. 6 Watermark “0” coefficients in double DCT domain

②若 $code(i)=1$, 则如式(4)修改系数, 结果如图 7 所示.

$$P'(j) = \frac{P'(j)}{\lambda}, mid + 1 \leq j \leq S' - T' \quad (4)$$

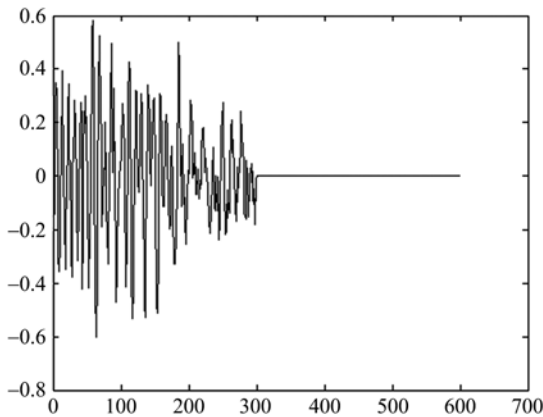


图 7 水印为“1”的双 DCT 域系数图

Fig. 7 Watermark “0” coefficients in double DCT domain

2.2 同步码嵌入算法

同步在通信系统中占有十分重要的地位. 同步码的性能好坏, 将直接影响通信系统的通信质量高低, 直接影响通信能否正常进行. 下文提出一种能够有效应用于空气信道传播的同步码算法.

经大量实验发现, 空气背景噪声的频率较低, 而且能量也很低, MATLAB 仿真频谱图如图 8 所示.

为了避免空气噪声的干扰, 选择在原始音频信号的中频频域嵌入同步信号, 使其能够抵抗空气背景噪声干扰. 如图 9 所示, 该算法在原始音频信号中植入少量中频同步信号.

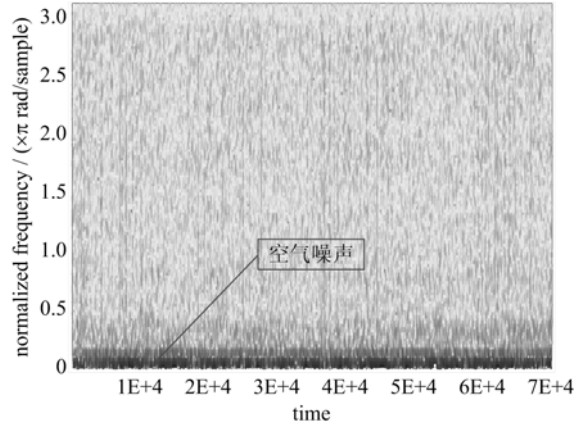


图 8 MATLAB 仿真空气噪声频谱图

Fig. 8 Simulated spectrum in MATLAB

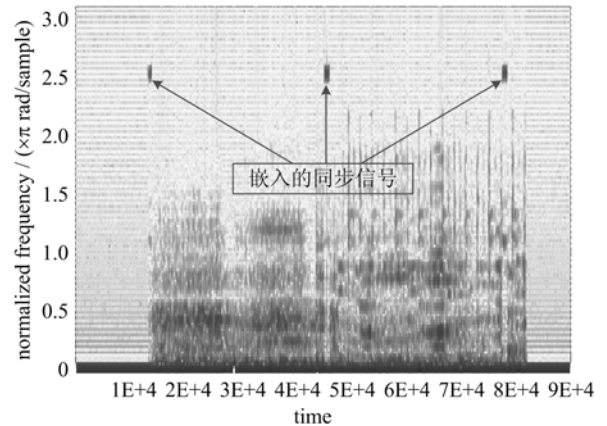


图 9 嵌入同步信号的频谱图

Fig. 9 Spectrum with embedded synchronous code

在空气传播中, 声波的能量会受距离、方位等因素的影响, 在一定程度上减弱. 如果嵌入的中频同步信号的能量值是绝对的, 那么同步信号不仅易被发现、抹除, 并且在提取时, 由于受到空气传播时距离、方位的影响, 不存在一个绝对的阈值作为同步信号的提取标准, 继而无法提取水印.

我们所采用的是一种循环相对能量同步码算法. 每个循环体的开始的 2 帧嵌入同步信号, 接下来嵌入一段水印信息, 这样构成了一个循环体, 将此循环体重复嵌入, 其隐藏信息结构如图 10 所示.

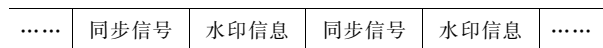


图 10 隐藏信息结构示意图

Fig. 10 The framework of embedded information

相对能量同步码算法能够在不同播放录音延时情况下,自动控制所嵌入的同步信号的强度,使得同步信号不易被发现、抹除,并能有效抵抗空气噪声干扰.并且在同步译码时,根据信号中低频和中频的能量的比值,能准确提取同步信息.具体同步码嵌入策略如下:

首先对 2 帧音频信号做快速傅里叶变换(FFT),计算出能量的绝对值,即幅频图,如图 11 所示.取其低频部分,求得能量的最大值,分别为 \max_1 , \max_2 .计算出用于修改同步信号能量值的参考系数 $\max_E_1 = \frac{\max_1}{1.8}$, $\max_E_2 = \frac{\max_2}{1.8}$.

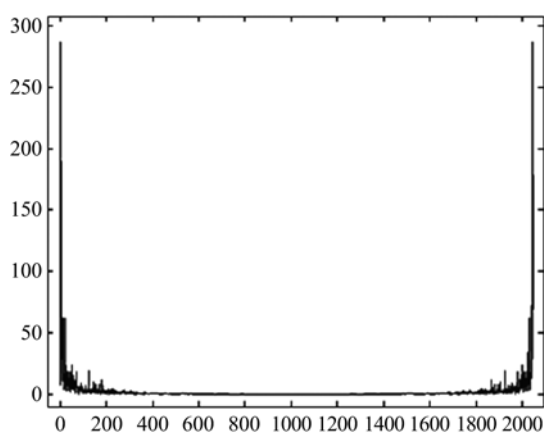


图 11 原始信号 FFT 变换

Fig. 11 FFT spectrogram of original signal

接着取中频部分的 8 个能量值 $E(i)_1, E(i)_2, 0 \leq i < 8$, 计算出 8 个点的最大值 \max'_1 和 \max'_2 , 然后根据式(5), (6), 对 $E(i)'_1, E(i)'_2$ 进行修改.

$$E(i)'_1 = \frac{E(i) \cdot \max_E_1}{\max'_1} \quad (5)$$

$$E(i)'_2 = \frac{E(i) \cdot \max_E_2}{\max'_2} \quad (6)$$

由于傅里叶变换后的能量分布具有对称性,所以在与这 8 个点对称的部分取 8 个点,作相同的嵌入操作,嵌入后 FFT 频谱图如图 12 所示.最后,将频域信号做快速逆傅里叶变换(IFFT),得到嵌入同步信息的时域信号.

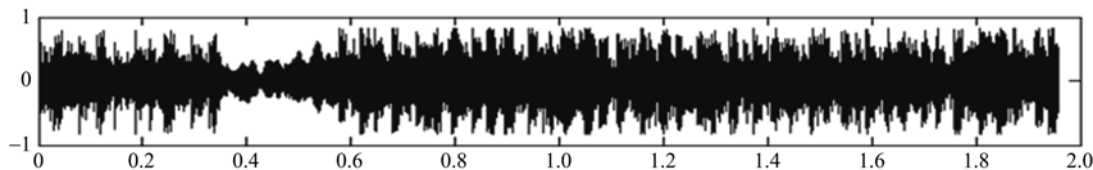


图 13 原始音频信号

Fig. 13 The oscillogram of original audio signal

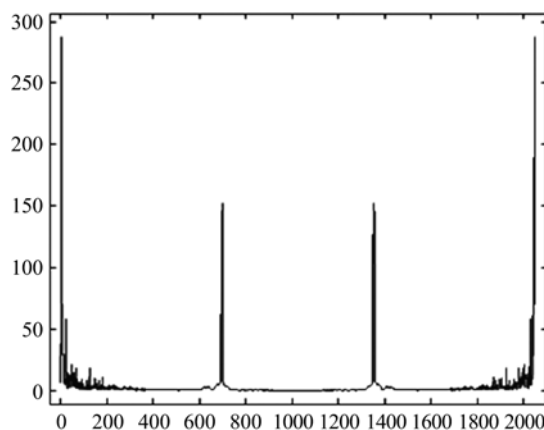


图 12 嵌入同步信号后的 FFT 变换

Fig. 12 FFT spectrogram with embedded sync signal

2.3 嵌入算法关键性能分析

此处,我们将对算法的 SNR(信噪比)、嵌入量等关键性能做一定的分析,使该算法更具可行性.

2.3.1 SNR 分析

在音频信号中嵌入水印,不可避免地对原始音频信号有所影响. SNR(signal to noise ratio, 信噪比)是目前音频水印不可察觉性的重要性能指标,将水印信号看作加性噪声. SNR 一般以分贝(dB)为单位,信噪比越高表示水印的不可察觉性越好. SNR 值计算表示为式(7)所示.

$$\text{SNR} = -10 \log_{10} \left\{ \frac{\|F - F'\|^2}{\|F\|^2} \right\} \quad (7)$$

式中, F 为原始音频信号, F' 为含有隐藏信息的音频信号.

国际留声联盟要求,在嵌入水印前后,信号的 SNR 值需大于 20 dB. 作者对 20 个音频样本进行了测试,测试结果显示,该算法的平均 SNR 值大于 21.63 dB,其中,最低 SNR 为 20.31 dB,最高 SNR 为 24.76 dB.

原始音频信号如图 13 所示,含有隐藏信息的音频信号如图 14 所示.由图 13、图 14 易见,音频信号的波形基本一致,在听觉上几乎不造成影响.

2.3.2 嵌入量分析

嵌入量是指单位时间内嵌入到音频载体信号中

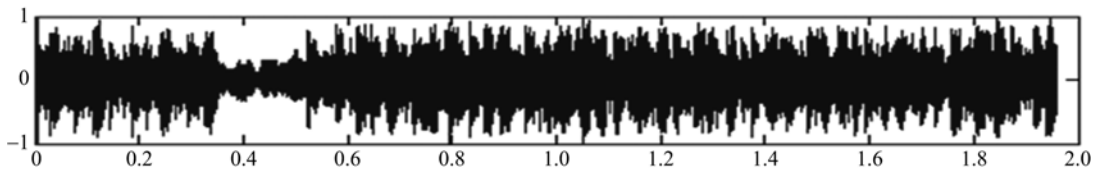


图 14 含有隐藏信息的音频信号

Fig. 14 The oscillogram of audio signal with watermarking

的比特数,单位用 bps 表示. 该算法的嵌入量表达式如式(8)所示,

$$B = \frac{R}{k \cdot (S' - T') \cdot L} /(\text{bps}) \quad (8)$$

式中, R 表示信号的采样率; L 为用于水印嵌入的单 DCT 变换域的系数段长度; S', T' 为双 DCT 系数的地址参数; k 为单 DCT 变化域系数和双 DCT 变化域系数的长度关系比值. 实验测试结果显示, 以 44.1 kHz 采样率的信号为例, 帧长为 2 048 个点, 本算法嵌入量为 21.53 bps.

2.4 水印提取算法

水印提取算法流程如图 15 所示. 首先对音频进行同步译码, 然后进行双 DCT 变换(具体流程详见节 2.1 水印嵌入算法), 通过比较双 DCT 域的系数提取嵌入信息, 将其进行 BCH 译码, 最后获得水印信息.

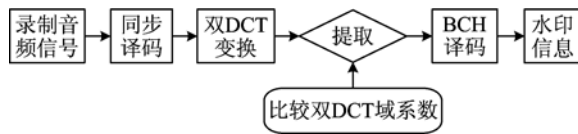


图 15 水印提取算法流程图

Fig. 15 Diagram of extraction algorithm

2.4.1 同步译码

设定阈值 std_mul, 以一帧为单位, 做 FFT 变换后取绝对值, 取中频段的 8 个点 $E(i), i=0, 1, \dots, 8$, 计算平均值 av_E, 然后计算低频部分能量最大值 max_E.

①若满足 $\frac{\max_E}{\text{av_E}} < \text{std_mul}$, 则找到信号中的一个同步信息, 并记录. 由于同步信号采用循环嵌入的方式(如图 10 所示), 因此我们跳过接着的一段水印信息, 再次进行同步, 直至信号的结束.

②若满足 $\frac{\max_E}{\text{av_E}} \geq \text{std_mul}$, 则将同步的起始点向后偏移 32 个采样点, 再次进行同步, 直到信号结束.

2.4.2 水印检测算法

在确定了待检测音频的同步信号后, 可按照嵌入水印时的分段规则将音频信号分段, 再对分段了的信号分帧, 并进行双 DCT 变换, 记双 DCT 域的系数为 $P'(i), 1 \leq i \leq \text{length}(P')$, 水印分割点记为 $\text{mid} = \lfloor \frac{\text{length}(P')}{2} \rfloor$, 如式(9)、式(10)所示提取隐藏信息.

$$E_0 = \sum_{i=1}^{\text{mid}} P'(i) \quad (9)$$

$$E_1 = \sum_{i=\text{mid}+1}^{\text{length}(P')} P'(i) \quad (10)$$

- ①若 $E_0 < E_1$, $\text{code}[i] = 0$;
- ②若 $E_0 > E_1$, $\text{code}[i] = 1$.

此时, 得到隐藏信息序列 $\text{code}[i], 0 \leq i < \text{length}$, 在该序列进行 BCH 纠错译码后, 得到水印信息 $\text{data}(i), 1 \leq i \leq \text{length}(\text{data})$.

3 性能测试与分析

通过在不同测试环境下, 对不同测试样本进行测试, 选取系数 $\text{MUL} = 1.8, \text{EPS} = 0.2, \text{std_mul} = 1.6$, 其测试结果如下.

3.1 测试环境及样本

测试环境如表 1 所示.

表 1 测试环境信息

Tab. 1 Test environment

硬件		播放设备 录音设备 声卡	漫步者 PT R201T08 得胜 T&S SM-8B 电容麦克风 ALC655
软件		播放软件 录音软件	Windows Media Player 10 Cool Edit Pro 2.1
距离	音箱到麦克风距离为 0.3, 1.0, 1.5 m 等		
音量	20%, 60%, 100% 等		
方位	麦克风正对音箱为 0°, 偏移 0°, ±30°, ±60°, ±90°		
环境	实验室(背景噪声约 20~30 dB)		

我们选用了采样率为 44.1 kHz 的各种音频样本, 如表 2 所示.

表 2 测试所用的音频文件

Tab. 2 Test samples

测试样本	长度	特点
Pop. wav	1min32s	频谱的分布较广, 音乐伴奏力度适中
Folk. wav	1min16s	信号前半段主要为低频, 后半段主要为高频
Song1. wav	1min41s	歌曲——男生演唱, 音乐伴奏较轻
Song2. wav	1min58s	歌曲——女生演唱, 音乐伴奏适中
Song3. wav	1min24s	交响曲——混有重低音, 背景音乐力度强

3.2 不同播录设备间距离、播放音量的影响

为了测试音箱到麦克风距离、播放音量大小等因素对误码率的影响, 我们取音箱到麦克风之间的距离分别为 0.3, 1.0, 1.5 m 等值, 音量分别取 20%, 60%, 100% 等值的情况下测试算法的误码率, 结果如表 3、表 4 所示。

测试表明: 在较合适的播放音量和适当的播录

表 3 不同距离下算法误码率统计(without BCH)

Tab. 3 The error rate in different situation without BCH

测试样本	音量与距离								
	volume: 20%			volume: 60%			volume: 100%		
	0.3 m	1.0 m	1.5 m	0.3 m	1.0 m	1.5 m	0.3 m	1.0 m	1.5 m
Pop. wav	3%	4%	6%	3%	5%	4%	6%	4%	3%
Flok. wav	5%	3%	2%	4%	3%	3%	6%	5%	4%
Song1. wav	3%	3%	4%	5%	6%	2%	2%	5%	3%
Song2. wav	3%	2%	5%	4%	3%	4%	4%	6%	3%
Song3. wav	5%	6%	4%	2%	4%	2%	3%	4%	2%

表 4 不同距离下算法误码率统计(with BCH)

Tab. 4 The error rate in different situation with BCH

测试样本	音量与距离								
	volume: 20%			volume: 60%			volume: 100%		
	0.3 m	1.0 m	1.5 m	0.3 m	1.0 m	1.5 m	0.3 m	1.0 m	1.5 m
Pop. wav	2%	3%	4%	2%	3%	3%	5%	3%	2%
Flok. wav	4%	2%	1%	3%	2%	2%	5%	4%	3%
Song1. wav	1%	3%	3%	5%	5%	2%	2%	3%	1%
Song2. wav	3%	1%	4%	2%	2%	3%	4%	5%	1%
Song3. wav	2%	4%	3%	1%	3%	1%	3%	4%	2%

表 5 不同方位下算法误码率变化

Tab. 5 The error rate under situation with different orientation

测试样本	角度							
	0°		±30°		±60°		±90°	
	without BCH	with BCH	without BCH	with BCH	without BCH	with BCH	without BCH	with BCH
Pop. wav	2.14%	1.12%	4.08%	2.91%	2.18%	1.02%	3.01%	1.53%
Flok. wav	3.32%	2.13%	2.26%	1.19%	4.12%	3.19%	4.35%	3.00%
Song1. wav	5.14%	4.01%	5.88%	4.67%	4.91%	3.62%	5.19%	3.92%
Song2. wav	3.91%	2.85%	4.13%	2.93%	3.29%	2.10%	4.03%	3.09%
Song3. wav	3.58%	2.29%	3.49%	1.76%	5.02%	3.83%	3.16%	2.01%

设备间距离下, 该算法可以有效降低误码率, 水印信息的平均正确提取率高达 96.6%, 对水印信息进行 BCH 纠错码可以降低误码率。

3.3 不同方位的影响

为了测试播放、录音设备之间的方位差对算法的影响, 以麦克风正对音箱为 0°, 左右各偏 30°, 60°, 90° 等(左偏移为正值, 右偏移为负值) 情况下测试算法的误码率, 结果如表 5 所示。

测试表明: 播放设备与录音设备的方位差异, 对水印的提取影响较小, 增加 BCH 纠错算法可降低误码率。

3.4 对常见信号处理攻击的鲁棒性测试

对录制信号进行常见信号处理攻击, 如 normalize, re-quantization, low pass 等, 其鲁棒性测试结果如表 6 所示。

表 6 该算法对常见攻击的误码率
Tab. 6 The error rate in the attacks of regular signal processing

测试样本	攻击类型					
	normalize		re-quantization 16→32→16 bit		low pass 11.025 kHz	
	without BCH	with BCH	without BCH	with BCH	without BCH	with BCH
Pop. wav	0.01%	0.00%	0.04%	0.02%	0.02%	0.01%
Folk. wav	0.00%	0.00%	0.07%	0.06%	0.03%	0.01%
Song1. wav	0.00%	0.00%	0.03%	0.02%	0.04%	0.02%
Song2. wav	0.01%	0.00%	0.01%	0.00%	0.14%	0.10%
Song3. wav	0.00%	0.00%	0.02%	0.00%	0.06%	0.03%

测试表明:该算法能够抵抗常见的信号处理攻击,如重量化、截止频率为 11.025 kHz 的低通滤波等攻击,增加 BCH 纠错编码能够降低误码率。

3.5 对重采样攻击的鲁棒性测试

对录制的信号进行不同系数的重采样攻击,该算法的鲁棒性测试结果如表 7 所示。

表 7 该算法对重采样攻击的误码率

Tab. 7 The error rate in the resample attack

攻击系数	误码率	
	without BCH	with BCH
-35%	15.12%	10.26%
-30%	8.91%	6.12%
-25%	6.87%	3.96%
-20%	3.72%	1.78%
-15%	1.42%	0.54%
-10%	0.81%	0.20%
-6%	0.17%	0.07%
-4%	0.08%	0.02%
-2%	0.01%	0.00%
+2%	0.01%	0.00%
+4%	0.02%	0.01%
+6%	0.14%	0.05%
+10%	0.75%	0.18%
+15%	1.54%	0.57%
+20%	2.99%	1.80%
+25%	5.97%	3.89%
+30%	9.24%	5.97%
+35%	14.93%	10.38%

测试表明:该算法能够抵抗约 30% 的重采样攻击,稳健性较好,且 BCH 纠错编码能够降低误码率,提升该算法的鲁棒性。

3.6 对其他攻击的鲁棒性测试

对录制音频进行基于 StriMark benchmark for audio 的攻击等其他攻击,该算法的鲁棒性测试结果如表 8 所示。作者对该算法还进行了 MP3 压缩攻

击,该算法同样具有较好的稳健性和鲁棒性,在此不再详细赘述。

表 8 该算法对基于 StriMark benchmark for audio 的攻击的误码率

Tab. 8 The error rate in the attacks based on StriMark benchmark for audio

攻击类型	误码率	
	without BCH	with BCH
Addbrumm_100	0.01%	0.00%
Addbrumm_10100	0.02%	0.00%
Addsinus	0.02%	0.00%
Invert	0.03%	0.00%
Stat2	3.61%	2.18%
Addbrumm_1100	0.02%	0.00%
Nothing	2.11%	0.64%
Compressor	0.01%	0.00%
Original	0.02%	0.00%
Rc-lowpass	7.46%	4.62%

测试表明:该算法对基于 StriMark benchmark for audio 的常见攻击误码率较低,具有很好的稳健性,增加 BCH 纠错编码能够进一步提升鲁棒性。

3.7 对比实验测试

以下,将该算法与文献[8]中算法相对比,其测试环境如表 9 所示。

表 9 对比实验测试环境信息

Tab. 9 Test environment of contrast experiment

硬件	播放设备	漫步者 PT R201T08
	录音设备	得胜 T&S SM-8B 电容麦克风
	声卡	ALC655
软件	播放软件	Windows Media Player 10
	录音软件	Cool Edit Pro 2.1
距离	音箱到麦克风距离为 1.0 m	
音量	60%	
方位	麦克风正对音箱为 0°,偏移±30°	
环境	实验室(背景噪声约 20~30 dB)	

对比实验的测试结果如表 10 所示.

表 10 对比实验的误码率

Tab. 10 Bit error rate of contrast experiment

测试样本	攻击类型			
	文献[8]中算法		本文算法	
	without BCH	with BCH	without BCH	with BCH
Pop. wav	3.36%	1.99%	2.04%	1.12%
Folk. wav	5.19%	4.01%	4.07%	2.86%
Song1. wav	4.43%	3.25%	3.34%	1.89%
Song2. wav	5.21%	3.72%	3.16%	2.09%
Song3. wav	4.18%	2.96%	2.27%	1.60%

测试表明,在真实的播放录音测试中,该算法在保持文献[8]中算法的优越性能的同时,有效地抵抗了抗 AD/DA 转换攻击,明显提升了抗空气噪声攻击的性能,且误码率较低,具有较好的鲁棒性和稳健性.

3.8 测试总结

该算法不仅能够有效抵抗 AD/DA 变换以及背景噪声干扰,还能抵抗空气信道传播时受播放距离、播放角度、播放音量等其他因素的干扰,并且能够抵抗常见信号处理攻击、重采样攻击、MP3 压缩攻击,以及基于 StriMark benchmark for audio 的攻击,误码率较低,具有较好的稳健性和鲁棒性,能够适用于空气信道传播的实际场景.

4 结论

通过修改双 DCT 域的系数,提出基于双 DCT 变换的音频数字水印算法.测试表明,该算法能够抵抗约 30%左右的重采样攻击,并能抵抗常见信号处理攻击、MP3 压缩攻击,以及基于 StriMark benchmark for audio 的攻击.此外,该算法还能够有效抵抗 AD/DA 变换以及背景噪声干扰,有效抵抗空气信道传播时受播放距离、播放角度、播放音量等其他因素的干扰.测试结果显示,该算法在真实的播放录音测试中,误码率较低,具有较好的鲁棒性和稳健性.该算法能够适用于面向空气信道传播的实际场景,为面向空气信道传播的音频信息的安全性,提供了更好的保障;为面向空气信道传播的数字水印研究领域,提供了新的发展方向.

致谢 本文得到武汉大学国家多媒体软件工程技术研究中心老师和博士生的大力帮助和支持.在此,对给予过帮助和支持的人表示最诚挚的谢意.

参考文献 (References)

- [1] Kim W G, Lee J C, Lee W D. An audio watermarking scheme with hidden signatures[C]//Proceedings of the International Conference on Signal Processing. Beijing, 2000:250-253.
- [2] Zhang Xiaoming. Audio watermarking algorithm for public information transmission [J]. Journal of Computer Applications, 2009, 29(9): 2 323-2 326.
- [3] Guhl D, Lu A, Bender W. Echo hiding[J]. Lecture Notes in Computer Science, 1996, 1174 :295-315.
- [4] Wu Shaoquan, Huang Jiwu, Huang Daren. DWT-based audio watermarking with self-synchronization [J]. Chinese Journal of Computers. Beijing, 2004, 27(3):365-370.
吴绍权,黄继武,黄达人.基于小波变换的自同步音频水印算法[J].计算机学报,2004,27(3):365-370.
- [5] Beauget S, van der Veen M, Lemma A. Informed detection of audio watermark for resolving playback speed modifications [C]// Proceedings of the Multimedia and Security Workshop. New York: ACM, 2004: 117-123.
- [6] Steineach M, Lang A, Dittmann J, et al. Audio watermarking quality evaluation: Robustness to DA/AD processes [C]// Proceedings of the International Conference on Information Technology: Coding and Computing. Washington: IEEE Press, 2002: 100-103.
- [7] Tachibana R. Audio watermarking for live performance [J]. Proc SPIE, 2003,5020: 32-43.
- [8] Xiang Shijun, Huang Jiwu, Wang Yongxiong. An audio watermarking algorithm against the D/ A and A/ D conversions [J]. Chinese Journal of Computers, 2006, 29(2): 308-316.
项世军,黄继武,王永雄.一种抗 D/A 和 A/D 变换的音频水印算法 [J]. 计算机学报,2006, 29 (2): 308-316.
- [9] Lei Yun, Liu Jian, Yan Bo, et al. Narrowband-Based Audio Watermarking with Self-Synchronization [J]. Chinese Journal of Computers, 2008, 31(7): 1 283-1 290.
雷赟,刘建,严波,等.窄带的自同步音频水印算法 [J]. 计算机学报,2008, 31(7): 1 283-1 290.