

DCT 域音频水印: 水印算法和不可感知性测度

温 泉, 王树勋, 年桂君
(吉林大学通信工程学院, 吉林长春 130025)

摘 要: 针对音频信号, 提出了一种水印算法, 该算法利用 DCT 频谱中的 DC 直流系数数据正负性不易改变的特性来嵌入水印信息, 通过实验结果说明该水印算法可以在鲁棒性与不可感知性之间达到平衡. 另外, 提出了一种结合心理声学模型的客观评价含水印信号失真度的方法, 通过实验结果和理论上的分析都说明这种不可感知性的测度可以体现人类听觉的主观特性, 是一种比较合适的客观评价方法.

关键词: 数字水印; 离散正弦变换 (DCT); 心理声学模型; 不可感知性

中图分类号: TN911.73 **文献标识码:** A **文章编号:** 0372-2112 (2007) 09-1702-04

Audio Watermarking in DCT Domain: Algorithm and Measurement of Imperceptibility

WEN Quan, WANG Shu xun, NIAN Gu r jun
(Information Department, Jilin University, Changchun, Jilin 130025, China)

Abstract: An audio watermarking algorithm is presented. Robustness of the algorithm is based on the steady positive negative sign of the DC coefficient in DCT domain. The result of simulation indicates that the algorithm has achieved balance between robustness and imperceptibility. At the same time, an objective measurement associated with psychoacoustic model for measuring the distortion of watermarked audio signal is presented in this paper. The result of experiment and theory analysis indicate that the measurement of imperceptibility matching the human auditory characteristics is an appropriate objective evaluating method.

Key words: audio watermarking; discrete cosine transformation; psychoacoustic model; imperceptibility

1 引言

为了解决数字产品盗版问题, 人们想了很多办法, 其中的数字水印技术公认为很有潜力解决这一棘手的问题. 在实际应用中, 水印技术要达到两个基本要求: 鲁棒性和不可感知性. 不过, 水印技术的这两个特性是一对矛盾, 在增加水印鲁棒性的同时不可避免地将减弱水印的不可感知性. 如何在鲁棒性和不可感知性之间取得一个平衡, 是研究水印算法中的一大难点.

为了解决这个难题, 人们提出了大量的水印算法. 这些算法, 按照嵌入域的不同, 可以分为两类: 空域/时域水印算法和变换域水印算法. 而在变换域嵌入水印可以得到比空域或时间域嵌入水印更好的鲁棒性, 目前在水印算法中常用的变换域有 DCT 域^[1,2]、DFT 域^[3]以及 DWT 域^[4]. 这里选择在 DCT 域内嵌入水印, 通过调整音频信号 DCT 变换后的 DC 直流系数的正负性, 以嵌入二值序列形式的水印信息, 通过理论分析以及实验结果可以说明设计的音频水印算法在鲁棒性和不可感知性之

间取得了很好的平衡.

音频数字水印的研究中经常利用心理声学模型求取信号在频域的掩蔽阈值, 这个掩蔽阈值一般用于整形水印信号使得水印算法具有较好的不可感知性^[3]. 本文从另外一个角度运用了心理声学模型, 将掩蔽阈值用于客观评价含水印信号的质量, 提出了一种符合人类听觉主观特性的不可感知性测度.

接下来第二部分介绍音频信号的水印嵌入和提取算法, 并分析采用 DC 直流系数嵌入水印的原因, 理论上说明了算法的鲁棒性; 第三部分提出了结合心理声学模型的不可感知性测度 DMOPM; 第四部分给出实验结果及分析, 从实验结果上验证了前面理论分析得出的结论.

2 水印算法

配合嵌入的 96bit 长的二值水印序列 $WM(i) = \{1, -1\}$, $i = 1, 2, \dots, 96$, 将一段音乐作品截取了其中 4915200 个采样点作为原始音频信号 $x(t)$. 以 512 点作

为一帧信号, 原始音频信号 $x(t)$ 恰好可分为 9600 帧.

嵌入水印过程需要分别对每一帧信号进行 DCT 正反变换, 为讨论方便, 对每一帧信号的 DCT 变换后系数引入下面的记号来表示:

$$F(n) = \{F_i(n) | i = 1, 2, \dots, 9600\},$$

$$n = 1, 2, \dots, 512 \quad (1)$$

其中 $F(1)$ 为 DC 直流系数, 余者均为 AC 交流系数.

1.2 水印嵌入算法

水印算法的嵌入步骤如下:

第一步: 将 $F(1)$ 分为 96 段序列, 每段序列长度为 100, 记为 $FDC_i(k)$, $i = 1, \dots, 96$, $k = 1, \dots, 100$.

第二步: 调整 $FDC_i(k)$ 中正负数的个数嵌入水印信息.

令 $FDC_i(k)^+$ 表示 $FDC_i(k)$ 中正数的个数, 而 $FDC_i(k)^-$ 表示 $FDC_i(k)$ 中负数的个数, 同时 D 表示水印嵌入的强度.

如果 $WM(i) = 1$, 则修改 $FDC_i(k)$ 中数据的正负性, 使其满足下式:

$$FDC_i(k)^- - FDC_i(k)^+ > D \quad (2)$$

反之, 如果 $WM(i) = -1$, 则修改 $FDC_i(k)$ 中数据的正负性, 使其满足下式:

$$FDC_i(k)^+ - FDC_i(k)^- > D \quad (3)$$

为了尽量不影响加水印后音频信号的质量, 从 $FDC_i(k)$ 绝对值最小的数开始修改. 将调整后的 $FDC_i(k)$ 记为 $FDC_i(k)^w$.

第三步: 对调整后的 DCT 系数进行逆变换得到时域的含水印音频信号 $x(t)^w$.

2.2 水印提取算法

水印提取的过程是嵌入过程的逆过程, 提取步骤如下:

第一步: 对嵌入过程中得到的含水印信号 $x(t)^w$ 分为同样的 9600 帧, 然后对每帧信号做 DCT 变换得到 $F(n)^w$.

第二步: 提取 $F(n)^w$ 中的 DC 系数 $F(1)^w$ 分为 96 段, 每段 100 个数, 记作 $FDC_i(k)_w$, $i = 1, 2, \dots, 96$, $k = 1, 2, \dots, 100$.

第三步: 计算 $FDC_i(k)_w$ 中数据正数个数 $FDC_i(k)_w^+$ 和负数个数 $FDC_i(k)_w^-$, 然后按下面的伪代码进行水印的提取:

```
IF  $FDC_i(k)_w^+ > FDC_i(k)_w^-$  THEN  $EW(i) = -1$ 
ELSE  $EW(i) = 1$ 
```

其中的 EW 表示提取出来的水印序列. 最后将 EW 和原始水印 WM 比较计算水印提取的误码率 BER.

2.3 水印算法鲁棒性分析

观察如下的 DCT 正变换公式:

$$y(k) = w(k) \sum_{n=1}^N x(n) \frac{\pi(2n-1)(k-1)}{2N},$$

$$k = 1, \dots, N \quad (4)$$

其中, N 是信号序列 $x(n)$ 的长度, $w(k)$ 定义如下:

$$w(k) = \begin{cases} \frac{1}{\sqrt{N}}, & k = 1 \\ \sqrt{\frac{2}{N}}, & 2 \leq k \leq N \end{cases}$$

在式(4)中, 令 k 为 1 可以得到 DC 系数的表达式如下:

$$y(1) = \frac{1}{\sqrt{N}}(x(1) + x(2) + \dots + x(N)) \quad (5)$$

又因为三角函数的性质 $-1 \leq \cos x \leq 1$, 所以 $y(1)$ 即 DC 直流系数的绝对值一般比 AC 交流系数的绝对值要大很多, 而绝对值越大的数其正负特性更加不容易被改变. 要使攻击后的音频信号的 DC 系数的正负性改变, 则相应的攻击强度要比使 AC 系数改变正负性的攻击强度大得多. 所以, 修改 DC 系数的正负性特征可以得到取得更好的鲁棒性.

3 不可感知性测度

人类的听觉是有限敏感的, 因此在听力阈值以下的声音是听不到的, 而且在不同的频率下灵敏度是不一样的, 也就是说不同的频率有不一样的掩蔽阈值. 心理声学模型是针对人类这种频率选择敏感听觉特性进行模拟的数学模型. 根据心理声学模型求取一帧音频信号在其频谱上的全局掩蔽阈值的详细算法步骤见参考文献[3].

根据心理声学模型的掩蔽阈值, 在此提出一种测量含水印信号的感知质量的方法. 计算得到的量度值记为 DMOPM (Distortion Measurement On Psychoacoustic Model), 按如下公式计算:

$$DMOPM = \sum_{i=1}^N [Tnorm_o(i) - Tnorm_w(i)]^2 / N \quad (6)$$

其中, $Tnorm_o(i)$ 表示原始音频信号 $x(t)$ 的掩蔽阈值向量, $Tnorm_w(i)$ 是含水印信号 $x(t)^w$ 的掩蔽阈值向量, N 是掩蔽阈值向量的长度.

基于心理声学模型的掩蔽阈值是一种人类听觉的主观特性的数学近似, 可以视作人类听觉的主观特性在数值上的映射. 那么, 水印嵌入引起的主观听觉上的变化也会反映在掩蔽阈值数值的变化上, 而式(6)的计算值恰恰正是原始音频信号和含水印音频信号之间掩蔽阈值的客观差异程度, 所以 DMOPM 的量度值可以从客观上较准确地反映含水印信号在主观听觉上的失真程度.

4 仿真实验及分析

实验中采用了三种不同素材的原始音频信号, 分

别来自于“贝多芬第三交响曲”、“贝尔格小提琴协奏曲”和“周杰伦—双节棍”,分别以测试信号 A、B、C 表示,信号的参数均为:采样频率 44.1kHz、量化精度 16bit/采样,采样点数 4915.2k.采用的仿真软件为 matlab 7.0.1.

4.1 水印鲁棒性

使用 Cool Edit 对含水印信号进行剪切攻击,以不同剪切长度随机选取剪切区域的方式做了五次实验,实验结果见表 1,表中以误码率(BER)来表示检测水印的效果.

表 1 抵抗剪切攻击实验结果

音乐作品	剪 切 长 度				
	5 秒	10 秒	15 秒	20 秒	25 秒
A	1.04%	2.08%	3.13%	5.21%	7.29%
B	1.04%	2.08%	4.17%	5.21%	7.29%
C	2.08%	3.12%	4.17%	6.25%	9.38%

从表 1 的实验数据看来,算法对于剪切攻击有很好的鲁棒性.

然后,将本文算法和经典的 SMM 音频水印算法^[6]进行比较,SMM 的嵌入参数设定为:帧长 100,调制因子 P 为 0.004,检测阈值为 0.2.需要注意的是,SMM 方法的帧长不能选择太长,否则影响水印系统的不可感知性.

使用标准水印测试软件 Audio Stirmark V3.0.1,对两种算法进行了同样条件下的攻击测试,测试结果分别见表 2 和表 3.

表 2 Audio Stirmark 测试本文算法实验结果

音乐作品	攻击类别	Echo	Snooth	Amplify	Flipsample	Addsinus
A		0	0	0	0	4.16
B		1.04	0	0	0	0
C		1.04	0	0	0	0

表 3 Audio Stirmark 测试 SMM 算法实验结果

音乐作品	攻击类别	Echo	Snooth	Amplify	Flipsample	Addsinus
A		44.8	3.13	1.04	16.7	33.3
B		44.8	3.13	1.04	16.7	33.3
C		37.5	12.5	1.04	16.7	5.21

从表 2 中看出在各种攻击下,水印提取的误码率(BER)大部份都是 0,即全部正确提取.表 3 中多数情况下的 BER 相对表 2 要大很多,可见本文算法的鲁棒性是令人满意的,而且在一定程度上比经典的 SMM 音频水印算法更加鲁棒.

针对 MP3 压缩攻击,比较了在 DC 系数上和在 AC 系数上嵌入水印的鲁棒性.首先,分别利用 DC、F(51)、F(91)、F(201)、F(301)、F(401) 嵌入水印(嵌入强度一律为 D=50),然后对各个含水印的音频信号进行压缩

率不同的 MP3 编码转换(采用软件 Power MP3 WMA Converter 1.0),然后对 MP3 压缩后的信号提取其中的水印.需要注意的是从 WAV 格式转换到 MP3 格式的时音频信号的长度产生了变化,对于这种情况,在提取的时候剪切了采样样本的开始部分的静音信号以保持和嵌入时原始音频信号同样的长度.这个比较实验采用的测试信号是流行音乐“双节棍”测试音频信号,比较结果见表 4.

表 4 不同嵌入位置的水印鲁棒性比较

压 缩 率	嵌 入 位 置				
	DC	F(91)	F(201)	F(301)	F(401)
128kbps	0	6.25%	14.58%	56.25%	51.04%
96kbps	1.04%	5.21%	20.83%	59.38%	42.71%
64kbps	0	7.29%	33.33%	50%	41.67%
48kbps	0	8.33%	37.5%	47.92%	45.83%

从表 4 的结果可以发现采用 DC 系数的算法比采用 AC 系数的水印算法在抵抗 MP3 压缩攻击上具有更好的表现.

4.2 水印不可感知性实验

为了实验水印的不可感知性,分别在嵌入强度 D 为 5、10、15、25、35、55、75、95 的情况下,对测试信号 A、B、C 嵌入了完全相同的水印信息.

首先采用常用的主观评价方法 MOS^[7]测量本文算法的不可感知性,测试对象有 10 个人,男性女性各占一半,其中一半有语音信号处理的专业背景.把在 8 次不同的嵌入强度下的得到的含水印信号 MOS 打分记录在了表 5.

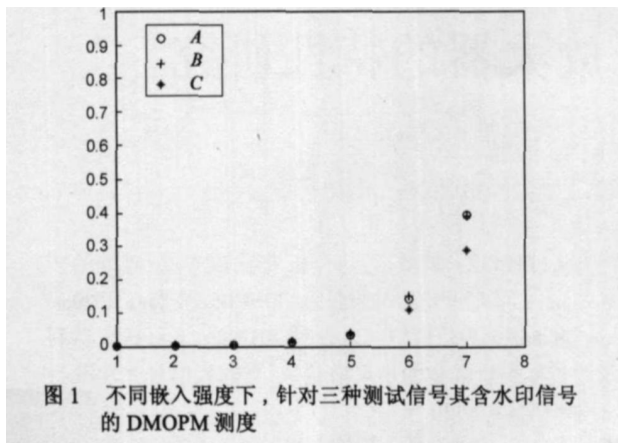
表 5 本文算法的 MOS 测试结果

音乐作品	周杰伦— 双节棍	贝多芬 第三交响曲	贝尔格小提 琴协奏曲
嵌入强度 D			
5	5	5	5
10	5	5	5
15	5	5	5
25	5	5	5
35	4.9	5	5
55	4.8	4.9	4.9
75	4.3	4.5	4.6
95	4.1	4.3	4.4

表 5 中右栏所示的 MOS 值是 10 个被测试者打分的平均值.注意鲁棒性实验中采用的嵌入强度 55 的 MOS 打分值是在高质量级别(4.5~5),这说明水印算法在鲁棒性已经足够的同时保证了水印的不可感知性.

然后,采用 DMOPM 计算含水印音频信号不可感知性测度,实验结果如图 1 所示.

图中的‘o’、‘+’、‘*’符号分别表示的是测试信号 A、B、C 的 DMOPM 测度结果,这里为了比较,8 次嵌入强度不同下的 DMOPM 测度值按其中的最大值做了归



一化处理。

在前面几次嵌入强度不大的实验情况下, 主观上感觉不到含水印信号的差异, 这从 MOS 测试结果也说明了嵌入强度小的情况下含水印信号质量很好。同样的, 从图 1 中也看出三种测试信号的 DMOPM 的测量值在前 4 个变化很小, 从第 5 个开始才有了一些变化。实验嵌入强度变化大(D 每次变化 20) 的时候, DMOPM 测度相应的变化幅度也比较大。总之, 实验结果说明了 DMOPM 的测度比较符合人类主观听觉系统特性。

5 总结和讨论

提出了一种 DCT 域的鲁棒性水印算法, 这种算法通过利用 DC 直流系数正负性不易改变的特性使得水印系统得到了很好的鲁棒性。通过实验说明了水印算法在鲁棒性和不可感知性之间取得了良好的均衡效果。另外, 结合心理声学模型的掩蔽阈值提出了一种更为适合的衡量水印系统不可感知性的测度方法 DMOPM。理论分析和实验结果都说明了 DMOPM 的合理性, 是一种符合人类主观听觉特性的客观评价方法。

参考文献:

- [1] 牛少影, 伍宏涛. 抗打印扫描数字水印算法的鲁棒性[J]. 中山大学学报, 2005(43): 1-4.
Shao zhang Niu, Hong tao, Wu. Robustness investigation of digital watermarking algorithm for print and scan process[J]. Acta Scientiarum Naturalium Universitatis Sunyatseni, 2005(43): 1-4. (in Chinese)
- [2] 黄继武, Yun Q SHI, 程卫东. DCT 域图象水印: 嵌入对策和算法[J]. 电子学报, 2000, (4): 57-60.
Ji wu, Huang, Yun Q SHI, Wei dong Cheng. Image watermark

ing in DCT: an embedding strategy and algorithm[J]. Acta Electronica Sinica, 2000, (4): 57-60. (in Chinese)

- [3] R A Garcis. Digital Watermarking of Audio Signals Using a Psychoacoustic Auditory Model and Spread Spectrum Theory [D]. American: Music Engineering Technology, University of Miami, 1999.
- [4] 刘红梅, 黄继武, 肖自美. 一种小波变换域的自适应视频水印算法[J]. 电子学报, 2001(12): 1656-1660.
Hong mei, Liu, Jr wu, Huang, Zhi mei Xiao. An adaptive video watermarking algorithm in wavelet domain[J]. Acta Electronica Sinica, 2001(12): 1656-1660. (in Chinese)
- [5] 孙圣和, 陆哲明, 牛夏牧. 数字水印技术及应用[M]. 北京: 科学出版社, 2004. 11. 440-442.
Sheng he Sun, Zhe ming Lu, Xiu mu Niu. Application and Technology of Digital Watermarking[M]. Beijing: Science Process, 2004. 11. 440-442. (in Chinese)
- [6] LI X, YU H. Transparent and robust audio data hiding in cepstrum domain[A]. IEEE International Conference on Multimedia and Expo(ICME) [C]. NY, USA, 2000, 3. 397-399.

作者简介:



温 泉 男, 1977 年生. 2005 年毕业于吉林大学通信工程学院, 获得工学博士学位, 现为吉林大学通信工程学院测控系讲师. 研究兴趣方向为信息安全、面向对象视窗编程. 已发表学术论文 15 篇, 获 2 项发明专利, 3 项实用新型专利. 做为主要人员参加多项国家、省级项目, 其中“交互式无线定位电子白板”获吉林省科技进步一等奖. E-mail: wenquan@jlu.edu.cn

王树勋 男, 1946 年 3 月生于吉林省长春市, 1970 年毕业于清华大学电机系, 1995 年被评为中国机械部科技专家, 1993 年获得国务院颁发的政府特殊津贴, 1999 年被 IEEE 总部聘为 IEEE IVEC'99 大会主席. 曾获国家发明奖三等奖一项, 机械工业部科技进步奖二、三等各一项, 吉林省科技进步奖一等奖一项, 发表论文近三百篇, 撰写著作三部. E-mail: shuxurwang@yahoo.com.cn



年桂君 女, 1973 年生, 吉林大学通信工程学院博士研究生, 吉林大学物理电子学院讲师, 研究方向为信息安全, 抗打印扫描水印技术、音视频水印技术. 发表论文 6 篇, 获发明专利 1 项. E-mail: niangj@jlu.edu.cn