

# 基于复倒谱变换的数字音频水印研究

李跃强<sup>1</sup>, 孙星明<sup>2</sup>, 周天亮<sup>1</sup>

(1. 怀化医学高等专科学校公共课部, 怀化 418000; 2. 湖南大学计算机与通信学院, 长沙 410082)

**摘 要:** 提出了一种基于倒谱变换的数字音频水印算法, 把音频信号划分为包含相同采样点的若干帧, 计算指定帧的均值, 对均值大于 0 的帧实施复倒谱变换, 将复倒谱系数的均值与阈值进行比较, 结合水印序列为“0”或“1”, 采取缩小、增加或不改变复倒谱系数均值的方法, 在复倒谱系数中嵌入水印。实验证明, 该算法具有健壮性、透明性, 可以进行添加噪声、重采样、低通滤波、重新量化、音频格式转换等常见信号的处理与承受各方的攻击。

**关键词:** 音频水印; 复倒谱变换; 透明; 健壮; 盲检测

## Research on Digital Audio Watermark Based on Complex Cepstrum Transform

LI Yueqiang<sup>1</sup>, SUN Xingming<sup>2</sup>, ZHOU Tianliang<sup>1</sup>

(1. Department of Public Course, Huaihua Medical College, Huaihua 418000;

2. School of Computer and Communication, Hunan University, Changsha 410082)

**【Abstract】** This paper proposes a new algorithm of digital audio watermark based on the complex cepstrum transform. The audio signal is divided into several frames which include the same samples, the mean value of frames is calculated, the frames of mean value is greater than or equal 0 is transformed by complex cepstrum, and mean value of complex cepstrum coefficient of the frames and threshold value is compared, according to the result of comparison and combining watermark bit, adopting the way of not changing, increasing or reducing coefficient of the complex cepstrum. Watermark is embedded into the complex cepstrum coefficients. The result demonstrates that the algorithm is robust, transparent and blind against common signal processing manipulations and attacks, such as re-sample, remodulation, noise, low-pass filter and conversion etc.

**【Key words】** Audio watermark; Complex cepstrum transform; Transparent; Robustness; Blind-detection

随着计算机、网络、多媒体技术的飞速发展, 数字产品极大地方便了人们对信息的获取和交流。数字产品在给人类生活带来快捷便利的同时, 也使人们面临着诸如非法盗版以及恶意篡改等严峻的挑战, 数字产品的版权保护已经成为一个迫切需要解决的现实问题, 数字水印技术被认为是解决版权纠纷的有效方式。数字音频在数字产品中占有较大的比重, 音频数字水印正逐渐成为数字水印中的一个重要研究方向。由于人类听觉系统对声音变化的灵敏度要高于人类视觉系统对图像变化的灵敏度, 因此在音频信号中嵌入水印难度相对较大。数字音频水印算法大体上可分为: (1)时(空)域法, 直接在信号空间中叠加水印信息的方法, 如 LSB 算法、回音算法等; (2)变换域法, 为了操作上的方便, 采用数学变换的方法, 将一个域内的数字信号映射成为另一个域内的数字信号, 然后加入水印, 再逆变换为原来域中的数字信号。如离散傅里叶变换、离散余弦变换、离散小波变换、倒谱变换等。一般说来, 时(空)域算法通常计算复杂度较低, 水印比较脆弱; 变换域法计算复杂度较高, 水印有较好的健壮性, 是今后的研究的主要方向。

### 1 复倒谱变换及实现

倒谱变换是语音信号处理中一种非常有用的工具。倒谱分析是同态系统理论的基础, 在语音识别、地震检测、声纳分析、生物医学、图像处理等信号处理领域中已有成功的应用。虽然理论上倒谱变换可实现数字水印, 但研究的文献却非常少见<sup>[1-3]</sup>。用倒谱变换实现音频数字水印在透明性、健壮性方面的优势, 这是值得研究。

由于复倒谱变换保留了信号的全部信息, 因此能重建信号, 满足数字水印变换域的基本要求。而实倒谱、双相干倒谱及功率倒谱等都不含相位信息, 不能重建信号, 因此不适合用于数字水印。

复倒谱变换的计算比较复杂, 用 Matlab 信号处理工具箱中提供的复倒谱变换工具, 能给数字水印的研究带来不少方便。在使用时应注意以下两点: (1)当信号序列的均值大于或等于 0 时, 其逆变换可逆; 当信号序列的均值为负时, 其逆变换不可逆。(2)当音频信号的采样点较少时(如小于 1 024), 进行复倒谱变换时可能会产生溢出错误。经检查分析复倒谱变换函数 `cceps`, 发现一条语句为

$\log h = \log(\text{abs}(h)) + \text{sqrt}(-1) * \text{rcunwrap}(\text{angle}(h))$

在做  $\log$  运算时产生溢出而出错, 将该语句改为

$\log h = \log(\text{abs}(h + \text{eps})) + \text{sqrt}(-1) * \text{rcunwrap}(\text{angle}(h))$

即可解决溢出错误问题, 其中,  $\text{eps}$  为一非常小的浮点数 ( $2.220 4 \times 10^{-16}$ )。

### 2 算法基本思路

算法的基本思路如下:

把原始音频信号划分为包含相同采样点的若干帧, 对均值大于 0 的帧实施复倒谱变换, 得到每帧的复倒谱系数, 计

**基金项目:** 国家自然科学基金资助项目(60373062); 教育部科研基金资助重点项目(03092); 湖南省教育厅科研基金资助项目(05C020)

**作者简介:** 李跃强(1960 - ), 男, 副教授、硕士, 主研方向: 数字水印, 声音识别; 孙星明, 博士、教授、博导; 周天亮, 讲师

**收稿日期:** 2006-02-17 **E-mail:** liyueqiang@163.com

算每帧复倒谱系数的均值 $M_c$ 并与阈值 $T$ 进行比较,结合水印序列为“1”或“0”,采取不改变、缩小或增加每帧复倒谱系数均值的方法,嵌入水印。

提取水印时只要将每帧复倒谱系数的均值 $M_c$ 并与阈值 $T$ 进行比较,就能提取水印,提取属盲水印提取。

为满足音频水印的透明性、健壮性及盲检测等要求,主要采取以下几种方法来实现:

(1)间隔若干帧再嵌入水印。若将水印信号嵌入每一帧,会使人耳产生感知,因此需要间隔一定数量的帧后再嵌入下一个水印比特,这样可提高透明性。间隔的多少由音频信号的长度、水印嵌入量及透明性要求高低等因素决定,一般为几帧到几十帧。

(2)通过改变某帧复倒谱系数的均值来嵌入水印。由于同时改变各个倒谱系数的幅值而不是改变某一个复倒谱系数的幅值。

复倒谱系数整体变化的幅度通常不大,音频信号变化也不大,音频信号、水印信息二者融为一体,使得嵌有数字水印的音频信号既保持原有的音频品质又嵌入了水印,人耳很难察觉,透明性较好。

(3)通过帧复倒谱系数的均值与阈值的比较来提取水印。提取水印时是将每帧的复倒谱系数的均值与阈值进行比较,而不是某一个复倒谱系数与阈值比较。这样,即使由于干扰、攻击而改变了某一个复倒谱系数,对整个帧的复倒谱系数的均值影响不大,因此其健壮性较强。

### 3 算法实现

#### 3.1 水印的嵌入

(1)将音频信号划分帧

把原始音频信号 $A$ 划分为若干帧,每帧包含 $n(n=2^p, p=1,2,3,\dots)$ 个采样点。每帧的采样点一般为32~128个采样点,间隔若干帧(通常为几帧~几十帧)后嵌入一个水印bit。

(2)计算帧信号序列的均值

计算指定帧信号序列的均值 $A_m$ 。

$$A_m = \text{MEAN}(A_m^i) = \frac{1}{n} \sum_{i=1}^n A_m^i \quad (m=1,2,3,\dots; i=1,2,3,\dots)$$

其中, $m$ 为帧序号, $i$ 为采样点数目。

当信号序列 $A_m$ 的均值为负时不予处理。为进一步提高健壮性、透明性,对信号序列 $A_m$ 的均值虽为正但很小时也不予处理。

(3)实施复倒谱变换

对满足条件的帧实施一维复倒谱变换。

$$C_m^i = \text{CCEPS}(A_m^i) \quad (m=1,2,3,\dots; i=1,2,3,\dots)$$

得到复倒谱系数序列 $C_m^i$ 。

(4)计算复倒谱系数的均值

计算指定帧复倒谱系数有均值 $C_m$ 为

$$C_m = \text{MEAN}(C_m^i) = \frac{1}{n} \sum_{i=1}^n C_m^i$$

(5)嵌入水印

将每帧的复倒谱系数的均值 $C_m$ 与预先设定的阈值 $T$ 进行比较,结合水印序列为“1”或“0”,采取缩小、增加或不改变复倒谱系数的方法嵌入水印。嵌入水印方法如下:

1)当水印信号为“1”时,即

if  $C_m < T$  then not changing else reducing  $C_m^i$  until  $C_m < T$

2)当水印信号为“0”时,即

if  $C_m \geq T$  then not changing else increasing  $C_m^i$  until  $C_m \geq T$

(6)重构音频信号序列

对已嵌入水印的倒谱系数 $C_m^i$ 的帧实施复倒谱逆变换(ICCEPS),得到已嵌入水印的帧 $A'_m$ 为

$$A'_m = \text{ICCEPS}(C_m^i)$$

用 $A'_m$ 取代原来的帧 $A_m$ ,得到已嵌入水印的音频信号 $A'$ 。

#### 3.2 水印的提取

提取水印是嵌入水印的逆过程。将已嵌入水印的音频信号 $A'$ 按嵌入水印时的方法划分帧,对均值为负或较小的帧不予处理,对指定帧进行复倒谱变换,并计算出每帧复倒谱系数的均值 $C'_m$ 。将均值 $C'_m$ 与阈值 $T$ 比较,提取水印。

水印提取方法如下:

if  $C'_m < T$  then watermarking bit="1" else watermarking bit="0"

得到水印序列,然后将一维水印序列转换为二维图像直观地显示出来。

#### 4 仿真与检测

取采样率为22 050Hz、播放时间为2'16"859、量化精度为16bits、音频格式为WAV的数字音频文件作为样本。分别按方案1(每帧64个采样点,每间隔30帧嵌入下一个水印bit信息)、方案2(每帧128个采样点,每间隔14帧嵌入下一个水印bit信息)划分音频文件样本。将一枚尺寸为32×34、大小为198B、单色二值BMP格式的天平图案作为水印信息(见图1(a)),经降维后成为一维水印序列,将水印序列按上述方法嵌入到复倒谱系数中。

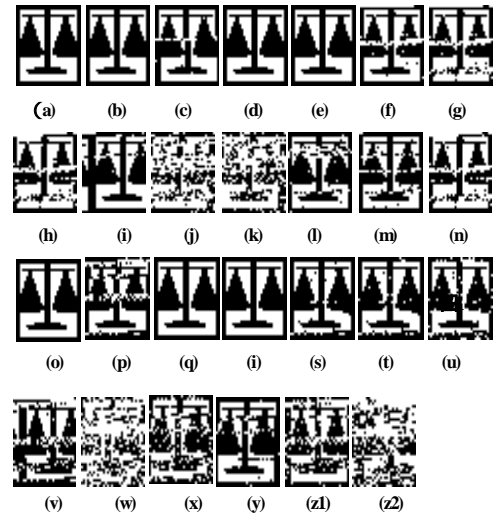


图1 仿真试验过程

#### 4.1 健壮性检测

为了检验水印的健壮性,对已嵌入水印的音频信号用常见的音频处理进行攻击仿真实验。实验表明,音频信号经音频处理、攻击后,其复倒谱系数均值会发生不同程度的变化,如:已嵌入水印的音频信号样本的复倒谱系数均值为-0.002 5;经2倍频重新采样后的复倒谱系数均值为-0.002 5;经Chebyshev型低通滤波器处理后的复倒谱系数均值为 $-6.399 2 \times 10^{-4}$ ;经椭圆型低通滤波器处理后的复倒谱系数均值为-0.001 2;经8bits重量化处理后的复倒谱系数均值为 $-4.879 4 \times 10^{-4}$ 。在提取水印时需要修改阈值 $T$ 以适应不同的情况。

图1(b)~图1(n)是将样本文件按方案1嵌入水印信息,然后用常见的音频处理方法进行操作、攻击后提取水印。图1(b)是直接从未嵌入水印的音频信号中提取的水印;图1(c)是经8bits重新量化后提取的水印;图1(d)是经24bits重新量化后提取的水印;图1(e)是经32bits重新量化后提取的水印;图1(f)是经2倍频重新采样后直接提取的水印;图1(g)是经4倍频重新采样后直接提取的水印;图1(h)是经1/2倍频重新采样后直接提取的水印;图1(i)是添加白噪声后提取的水印;图1(j)是经最大衰减为2dB的2阶Chebyshev型低通滤波器

滤波后提取的水印；图 1(k)是经最大衰减为 8dB 的 2 阶 Chebyshev 型低通滤波器滤波后提取的水印；图 1(l)是经通带内的最大衰减为 2.5dB，在阻带内的最小衰减为 6dB 的 4 阶椭圆型低通滤波器滤波后提取的水印；图 1(m)是用音频转换软件 Audiolib MP3 Converter 对已嵌入水印的音频信号经 WAV(采样率为 22 050Hz)→WAV(44 100Hz)转换后直接提取的水印；图 1(n)所示是已嵌入水印的音频信号经 WAV(2 250Hz)→MP3(44 100Hz)→WAV(2 250Hz)转换后提取的水印。需要说明的是，使用该音频转换软件进行转换后，会在音频信号起始处增加一些采样点。图 1(o)~图 1(z)将样本文件按方案 2 嵌入水印信息，然后按与上述相同的音频处理方法进行操作、攻击后提取的水印。

为消除对提取的水印图像的主观影响，可使用归一化相关系数 NC 对原始水印图像与提取水印图像的相似性作出更为客观的评价。归一化相关系数 NC 为

$$NC(W,W')=\frac{\sum_{i=1}^{m_1}\sum_{j=1}^{m_2}w(i,j)w'(i,j)}{\sqrt{\sum_{i=1}^{m_1}\sum_{j=1}^{m_2}w(i,j)^2}\sqrt{\sum_{i=1}^{m_1}\sum_{j=1}^{m_2}w'(i,j)^2}}$$

其中，W、W'分别为原始水印序列、提取的水印序列。是否存在水印的判断标准是：NC(W,W')>T，T 的取值在 0~1 之间。在实验中，T 的取值通常为 0.5。

表 1 所示是按不同帧划分音频信号，采用常见的音频处理后提取的归一化相关系数为 NC。其中，括号中的数值为阈值，w-w 表示 WAV(22 050Hz)→WAV(44100Hz)转换，w-m-w 表示 WAV(22 050Hz)→MP3(44 100Hz)→WAV(22 050Hz)转换。

表 1 常见音频处理后提取水印的相关系数 NC

方案	嵌入时 取阈值	直接 提取	8bits 重量化	24bits 重量化	32bits 重量化	×2 采样	×4 采样
64 点/帧， 间隔 30 帧	(0.200 0)	1.000 0 (0.020 0)	0.974 2 (0.020 0)	1.000 0 (0.020 0)	1.000 0 (0.020 0)	0.906 9 (0.015 4)	0.865 4 (0.010 4)
128 点/节， 间隔 14 帧	(0.005 0)	1.000 0 (0.005 0)	0.811 5 (0.005 0)	1.000 0 (0.005 0)	1.000 0 (0.005 0)	0.935 5 (0.005 1)	0.928 9 (0.004 0)
方案	/2 采样	加白 噪声	Cheb I 低通	Cheb II 低通	Ellipse 低通	w-w*	w-m-w*
64 点/帧， 间隔 30 帧	0.843 7 (0.017 9)	0.564 0 (0.020 0)	0.640 6 (0.000 1)	0.652 4 (0.009 4)	0.890 7 0.010 2	0.911 6 0.015 1	0.845 2 0.020 0
128 点/节， 间隔 14 帧	0.893 7 (0.000 6)	0.489 2 (0.005 0)	0.608 2 (0.000 1)	0.694 0 (0.000 2)	0.813 6 (0.000 1)	0.677 3 (0.005 8)	0.518 0 (0.009 1)

#### 4.2 透明性检测

对已嵌入水印的音频信号试听，在主观听觉上与原始音频比较没有明显区别，说明嵌入水印没有产生明显的降质现象，其透明性好。为消除主观影响，对嵌入水印前后的波形进行观察比较。

在方案 1 中，原始音频信号、嵌入水印后的音频信号及误差信号波形的比较如图 2 所示。在方案 2 中，原始音频信号、嵌入水印后的音频信号及误差信号波形的比较如图 3 所示。

为观察原始音频信号、已嵌入水印的音频信号的波形细节，将图 2 中误差信号较大处(采样点为 1 397 830~1 398 040)的波形放大进行详细比较，图 4 是原始音频信号、嵌入水印后的音频信号及误差信号波形详细图。可以看出，虽然嵌入水印后音频信号在采样点 1 397 830~1 398 040 区间的波形幅度整体上有一定程度的减少，但仍然较好地保持了原有波形。根据 HAS 的掩蔽特性，人耳不会察觉这细微的变化。

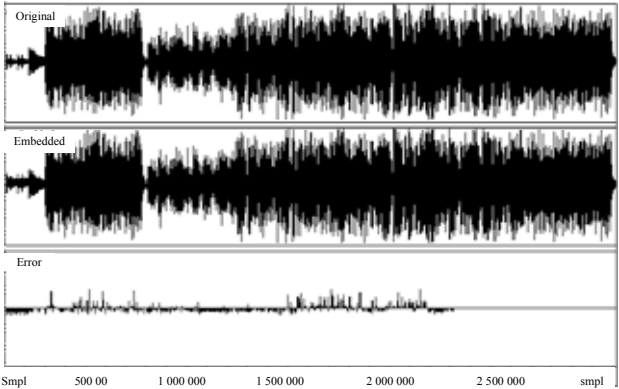


图 2 方案 1 的对比

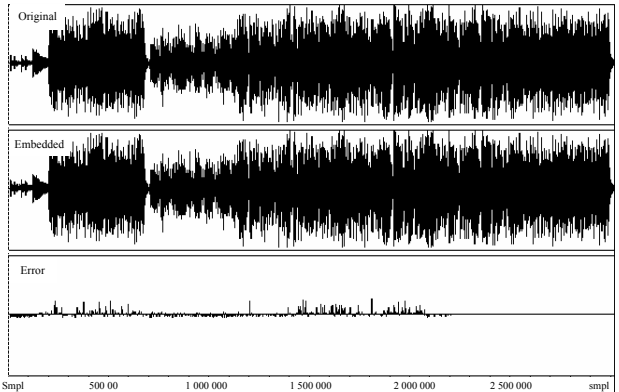


图 3 方案 2 的对比



图 4 3 种信号的波形详解

#### 4.3 阈值 T 对健壮性、透明性的影响

在嵌入水印时，选择适当的阈值非常重要，否则会对健壮性、透明性产生严重的影响。嵌入时选择的阈值是否适当，可通过比较误差信号的上部与下部是否基本平衡来判断。若误差信号的上部总体上比下部多，则健壮性会降低，需要增加阈值；反之，则透明性会降低，需要减少阈值。直到误差信号上、下部分总体上基本平衡为止。这样，透明性、健壮性均可兼顾。一般说来，每帧划分的采样点越小，阈值就越大；反之，阈值就越小。

从图 2、图 3 中可以看出，原始音频信号嵌入水印后对波形影响较小，透明性较好，对音频的影响较小。每帧的采样点数越多，其透明性就越好。从表 1 也可看出，每帧的采样点多，透明性就好，嵌入水印量就小，而健壮性并不一定就强。

#### 5 结束语

实验证明，基于复倒谱变换的数字音频水印算法与其它变换域水印算法相比，具有透明性、健壮性，能经受重新量化、重采样、添加噪声、低通滤波、音频格式转换等常见信号处理及攻击，且提取水印属盲水印提取，对倒谱域数字音

(下转第 161 页)