

# 一种基于 DWT 和 DCT 域的音频信息隐藏算法\*

柏玉保, 柏 森, 暴晋飞, 尤春艳  
(重庆通信学院, 重庆 400035)

**摘要:** 提出了一种基于离散小波变换(DWT)和离散余弦变换(DCT)的音频信息隐藏的新算法。首先,对载体音频信号整体进行小波分解,将其低频小波系数分段后进行离散余弦变换;其次根据人耳听觉的频率掩蔽效应,选择出对人耳听觉容差最大的直流系数组成序列并分段;最后通过修改各段统计特征来嵌入水印。实验表明,嵌入水印后的语音信号不仅具有良好的不可感知性,而且对诸如加噪、低通滤波、重采样、重量化、回声、MP3 压缩、样点裁剪、时域线性延伸和缩短等的攻击具有很强的鲁棒性。

**关键词:** 信息隐藏; 音频盲水印; 离散小波变换; 离散余弦变换

中图分类号: TP391

文献标识码: A

## An audio information hiding algorithm based on DWT and DCT

BAI Yu Bao, BAI Sen, BAO Jin Fei, YOU Chun Yan  
(Chongqing Communication Institute, Chongqing 400035, China)

**Abstract:** An algorithm based on DWT and DCT is presented to embed watermark in cover audio signal. At first, the host audio is decomposed by wavelet transform and its low frequency wavelet coefficients after separated into segments are converted to DCT domain. Then, select out the DC coefficients which is insensitive to the human audio system and has the biggest hearing margin, and at the same time convert these coefficients to a sequence and separate into segments. Finally, embed watermark according to its statistical feature in DC coefficients. Experimental results show that the watermarked audio has good imperceptibility and is robust against different kinds of attacks, such as noise adding, low-pass filtering, resampling, equalization, echo, MP3 compression, resample cropping, temporal domain scaling.

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

随着计算机网络通信技术的发展,人们借助计算机、数字扫描仪等电子设备可以方便、迅速地将数字信息传输到所期望的目的地。然而,当人们享受网络技术带来的便利的同时,不得不面对随之而来的负面影响。例如,盗版者通过简单的操作就可以获得与正版一样的复制品,并以此获取暴利。近年来,音频水印发展非常迅速,音频数据的版权保护也显得越来越重要。通过在音频载体中嵌入水印,可以实现盗版确认、使用跟踪等功能。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<sup>[1]</sup>、相位编码<sup>[2]</sup>和回声隐藏<sup>[3]</sup>,这些算法比较简单,但是鲁棒性差,但也有些算法具有很好的鲁棒性

(如时域能量算法<sup>[4]</sup>等)。目前研究的方向主要集中在 DFT 变换域<sup>[5]</sup>、离散余弦变换 DCT<sup>[6-7]</sup>、离散小波变换 DWT<sup>[8]</sup>及其他正交域。DWT 是信号处理中一种非常有力的工具,它是一种信号的时间-频率分析方法,在时域、频域内都具有表征信号局部特性的能力,具有良好的鲁棒特性等。参考文献[8]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。DCT 域信息隐藏算法计算量比较小,现在研究得比较多。参考文献[7]就是利用人耳听觉的频率掩蔽效应分析了 DCT 系数的听觉容差,发现直流系数具有最大的听觉容差。通过改变 DCT 直流系数来嵌入水印,鲁棒性和不可感

\* 基金项目:重庆市自然科学基金(2008BB2296)

知性都达到很好的平衡。而参考文献[9]提出了一种基于 DCT 和 DWT 域的音频信息隐藏算法。该算法充分利用了离散小波变换的多分辨率特性和离散余弦变换的能量压缩能力,具有很好的鲁棒性。参考文献[10]充分利用了 DCT 和 DWT 的优点,提出 1 种混合域盲水印算法,具有很好的鲁棒性。参考文献[11]则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,都具有非常好的不可感知性和强鲁棒性。

本研究结合了 DCT 和 DWT 域的算法,将数字水印隐藏到另一个载体音频中。在实现时,首先对载体音频进行小波变换,获取低频小波系数,然后适当分段并进行离散余弦变换得到直流系数,将得到的直流系数按顺序组合成一组向量,对其选取适当的长度进行分段,最后根据其统计特征嵌入水印<sup>[12]</sup>。实验结果表明,该算法不可感知性良好,与参考文献[9-11]的混合域算法相比,本文算法鲁棒性有了较大提高。

### 1 水印嵌入算法

假设载体语音为  $X=(x_1, x_2, \dots, x_L)$ 。其中  $L$  为载体语音信号的采样点数,  $x_i$  为第  $i$  个采样点的幅度值。

#### (1) 离散小波变换

用“db2”小波对载体语音  $X$  进行离散小波变换(DWT):  $[S, H]=DWT(X)$ , 其中  $S$  为低频分量, 记为  $S=(s_1, s_2, \dots, s_k)$ ,  $k$  为低频分量长度(低频系数的个数)。对低频分量  $S$  进行分段, 每段长度为  $N$ , 得到  $S_1, S_2, \dots, S_M$ , 如下:

$$S = (\underbrace{s_1, s_2, \dots, s_N}_{s_1}, \underbrace{s_{N+1}, \dots, s_{2N}}_{s_2}, \dots, \underbrace{s_{(M-1)N+1}, \dots, s_{MN}}_{s_M})$$

其中,  $M$  为分段数且  $M \leq \lfloor k/N \rfloor$ ,  $S_i$  为第  $i$  段低频系数。

#### (2) 离散余弦变换

对  $S_1, S_2, \dots, S_M$  分别进行 DCT 变换:  $D_i = DCT(S_i)$  为第  $i$  段低频系数 DCT 变换后的系数, 且  $D_i = (d_i(1), d_i(2), \dots, d_i(N))$ 。提取每段的直流系数组成向量  $C$ , 记为  $C = (d_1(1), d_2(1), \dots, d_M(1))$ 。将  $C$  进行分段, 每段长度取  $T$  (不足的用零补齐), 记为

$$R = (\underbrace{d_1(1), d_2(1), \dots, d_T(1)}_{R_1}, \underbrace{d_{T+1}(1), d_{T+2}(1), \dots, d_{2T}(1)}_{R_2}, \dots, \underbrace{d_{(n-1)T+1}(1), \dots, d_M(1), 0, \dots, 0}_{R_n})$$

其中,  $n = \lceil M/T \rceil$ ,  $\lceil \cdot \rceil$  表示向上取整。在本文中,  $T$  的取值范围对算法的鲁棒性很重要, 实验表明“ $T=10$ ”时效果比较好。

#### (3) 采用统计特征的方法嵌入水印

用种子“seed”产生一段二进制伪随机序列作为水印信息, 记为  $W = \{m(i) | i=1, 2, \dots, \lceil M/T \rceil\}$ , 其中  $m(i) \in \{0, 1\}$ , 表示序列的第  $i$  个值。

在本文中, 根据参考文献[12], 找到一种对攻击不敏感的统计特征, 该特征值就是通过计算上述每段直流系数向量  $R_i$  求得的平均值。实验表明, 该统计特征值对一般的信号处理(如低通滤波、加噪声、重量化、重采样等)具有稳健性。

根据如下步骤修改每段直流系数向量  $R_i$  进行水印信息的嵌入。

$$\textcircled{1} R_i' = \frac{1}{T} \sum_{j=1}^T d_{(i-1)T+j}(1) \quad (1)$$

② 如果  $W(i)=1$ , 则

$$R_i' = R_i' + \Delta * abs(R_i') \quad (2)$$

如果  $W(i)=0$ , 则

$$R_i' = R_i' - \Delta * abs(R_i') \quad (3)$$

其中,  $\Delta$  表示嵌入强度系数, 该值越大, 提取水印鲁棒性越强, 但衡量不可感知性的信噪比越弱, 反之则相反。

#### (4) 反离散余弦变换

将嵌入水印后的 DCT 直流系数段逐个和未改变的其他 DCT 系数按照原始顺序合并, 进行反离散余弦变换, 得到新的 DCT 系数。

#### (5) 反离散小波变换

利用新得到的 DCT 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 水印提取算法

由于水印嵌入过程中使用了统计特征的方法, 所以在信息提取时不需要原始载体语音信号, 属于盲提取算法。提取过程如下:

(1) 离散小波变换。使用“db2”小波基对待检测语音信号序列整体进行小波变换, 并提取低频小波系数, 对其按嵌入时的长度值进行分段。

(2) 离散余弦变换。对每段低频小波系数进行 DCT 变换, 提取其各段的直流系数, 并按顺序组合成向量, 同样对其按嵌入时的适当长度值进行分段。

(3) 水印提取。计算各段直流系数的均值。如果该段的均值大于“0”, 则提取出 bit 为“1”; 而该段的均值小于“0”, 则提取出 bit 为“0”。

(4) 将从各个分段中提取的水印信息按顺序进行合并, 得到完整的水印信息。

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:

$$S(w, g) = \frac{\sum_{i=1}^l w(i)g(i)}{\sqrt{\sum_{i=1}^l w^2(i)} \times \sqrt{\sum_{i=1}^l g^2(i)}} \quad (4)$$

### 3 实验结果

在实验中, 以一段长度为 16 s, 抽样速率为 8 kHz, 位速为 128 Kb/s, 16 位量化的语音信号为公开信息, 在

网络与通信 Network and Communication

该段语音中嵌入水印,采用种子“seed”产生一段二进制伪随机序列作为水印信息,根据上述算法,在低频小波系数分段阶段,取  $N=64$ ,小波基为“db2”;对直流系数组成的向量,取  $T=10$ ,嵌入强度系数  $\Delta=3.5$ 。

图 1 是原始文件和嵌入水印的音频文件时域波形图,从波形图上可以看到,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哪个是水印音频,将水印音频和原始音频播放给 5 个人听,一致辨别不出两者有何区别;用归一化相关系数和信噪比对音频文件进行客观评价,归一化相关系数如公式(4)所示。

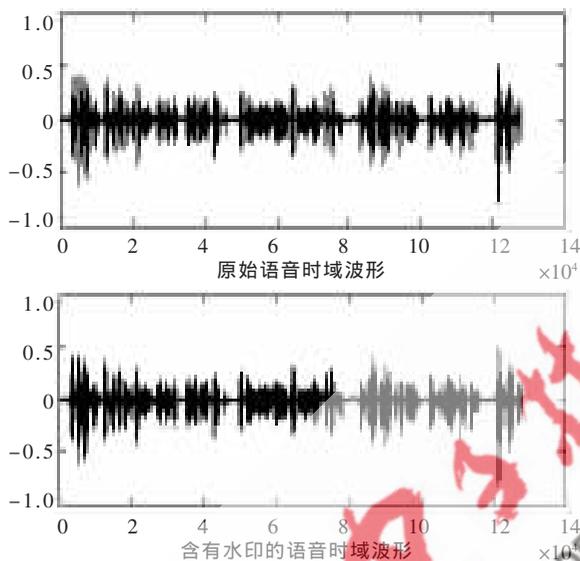


图 1 原始音频与含有水印音频时域图形

信噪比公式为:

$$SNR = 10 \log \left( \frac{\sum_{n=1}^L x(n)^2}{\sum_{n=1}^L (x(n) - x'(n))^2} \right) \quad (5)$$

其中,  $x(n)$  为原始音频,  $x'(n)$  为嵌入水印后的音频。通过计算,嵌入水印信息前后音频文件的整体信噪比为 21.779 8 dB, 相关系数为 0.996 7。这说明该算法无论在主观和客观上都具有很好的不可感知性。

为了检验水印的鲁棒性,对音频水印文件做以下信号处理。

(1)低通滤波。低通滤波器为长度为 9 阶、截止频率为 1.5 kHz 的巴特沃思低通滤波器。

(2)加高斯白噪声。均值  $u=0$ , 均方差  $\sigma=0.01$ 。

(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 400 ms,延时信号幅度为原始信号 10%。

(4)重采样。将音频文件用 6 kHz 重采样,再重新用 8 kHz 重新采样恢复成原始音频。

(5)MP3 压缩。将含水音频压缩为 MP3 文件,然后恢复成与原始音频格式相同的音频。

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

(7)同步攻击:在含有水印音频中随机选取若干个采样点,然后进行裁剪。裁剪的样点个数分别是 100、300、500 个;在原始音频中每 0.05 s 分别连续删除 1、2、3 个点,共分别删除 100、300、498 个采样点。

(8)时域线性压缩。用 Cool Edit Pro 2.0 软件将含密音频压缩 1%,即从 16 s 线性缩短到 15.98 s。

(9)时域线性延伸。用 Cool Edit Pro 2.0 软件将含密音频延伸 1%,即从 16 s 线性延伸到 16.02 s。

(10)音调改变。不改变信号速度而改变音调频率。用 Cool Edit Pro 2.0 软件将含密音频进行音调的改变,分别将音调增加为原始音调频率的 102%(Ritio=102%)和减少为原始音频频率的 98%(Ritio=98%)。

(11)DA/AD 转换。利用 Winamp 音乐播放器播放原始音频,利用 Windows 自带的录音机通过线路输出方式进行录音。

根据归一化相关系数定义分别计算各种攻击下提取水印的相关系数,结果如表 1 所示。

本文提出的基于离散小波变换(DWT)和离散余弦变换(DCT)的混合域音频信息隐藏算法,首先使用小波将

表 1 鲁棒性攻击检测结果及算法比较

攻击类型	归一化相关系数		
	参考文献[10]	参考文献[11]	本文算法
未攻击	1	1	1
低通滤波	0.930 8	1	1
高斯白噪声	0.960 5	1	1
回响	未提	0.992 6	1
低通滤波+高斯白噪声+回响	未提	未提	1
重采样	0.790 1	1	1
重量化	0.991 3	1	1
MP3 压缩	0.970 1(11:1)	0.998 6(11:1)	0.972 5(3:1)
	0.807 4(22:1)	0.992 6(22:1)	1(2:1)
样点裁剪	未提	0.935 1(100)	0.972 0(100)
		0.833 3(300)	0.954 2(300)
		0.766 3(500)	0.921 3(500)
		0.955 3(每 400 点裁剪 1 点)	0.929 2(每 400 点裁剪 2 点)
		0.836 5(每 400 点裁剪 3 点)	0.836 5(每 400 点裁剪 3 点)
时域线性伸长(+1%)	未提	未提	0.972 0
时域线性压缩(-1%)	未提	未提	0.981 5
音调增加(+2%)	未提	未提	0.897 2
音调减小(-2%)	未提	未提	0.945 6
DA/AD 转换	未提	未提	0.596 6

音频文件整体进行分解,选择对听觉不敏感和抗噪抗低通的低频小波系数部分作为嵌入频带;然后适当分段并进行离散余弦变换得到直流系数,将得到的直流系数组合成一组向量,对其选取适当的长度进行分段,最后根据其统计特征来嵌入水印,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。能抵抗噪声、低通滤波、重采样、重量化、回声、MP3压缩、样点裁剪、时间伸长和压缩等的攻击,特别是在抗样点裁剪、时域线性伸长与压缩、音调改变攻击方面,相似度都达到0.9以上,几乎能完全正确提取水印,表现出了非常强的鲁棒性。从实验结果来看,本文所提算法与参考文献[10,11]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。

然而这种信息隐藏算法的缺点在于:一是不能抵抗同步攻击以及D/A、A/D转换操作方面比较弱;二是算法的信息容量小。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[M]. USA: Patent, 1997.
- [2] BENDER W, GRUHL D, MORIMOTO N, et al. Techniques for data hiding[J]. IBM System Journal, 1996, 35: 313-336.
- [3] GRUHL D, BENDER W, LU A. Echo hiding [C]. Proceedings of the 1st International Workshop on Information Hiding. [S.l.]: Springer, 1996: 295-315.
- [4] WEN Nung Lie, LI Chun Chang. Robust and high-quality time-domain audio watermarking based on low-frequency amplitude modification[J]. IEEE Transactions on Multimedia, 2006, 8(1): 46-59.
- [5] 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.
- [6] 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.
- [7] 由守杰, 柏森, 曹巍巍. 鲁棒的DCT域音频盲水印算法[J]. 重庆邮电大学学报(自然科学版), 2008, 20(2): 206-211.
- [8] 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, 22(12): 118-119.
- [9] 王先春, 郭杰荣, 胡惟文, 等. 基于DCT和DWT域的音频信息隐藏算法[J]. 计算机工程与设计, 2008, 29(16): 4389-4391.
- [10] 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.
- [11] 由守杰, 柏森, 曾辉. 鲁棒的混合域音频信息隐藏算法[J]. 计算机技术与发展, 2008, 18(3): 169-172.
- [12] LI Wei, XUE Xiang Yang, LI Xiao Qiang, et al. A novel feature-based robust audio watermarking for copyright protection [C]. Proceedings of the International Conference on Information Technology: Computers and Communications (ITCC'03) 2003 IEEE, 2003.

(收稿日期: 2009-09-04)

#### 作者简介:

柏玉保, 男, 1982年生, 硕士研究生, 主要研究方向: 语音信息隐藏, 数字音频水印。

柏森, 男, 1963年生, 博士后, 教授, 主要研究方向: 信息隐藏, 掩密通信, 图像处理, 模式识别。

暴晋飞, 男, 1982年生, 硕士研究生, 主要研究方向: 语音信息隐藏, 数字音频水印。