

用于MP3音频压缩的水印嵌入技术研究

黄昊 郭立 李琳

(中国科学技术大学电子科学与技术系 合肥 230027)

摘要: MP3已经成为数字音频的主流压缩编码。该文从实验出发探讨了MP3编解码对时域和频域两方面影响,基于此提出了一种MP3编解码的水印算法:针对时域变化采用载体信号和重采样划分来实现嵌入与提取同步,通过量化相邻子帧低频能量比来嵌入信息以消除频域失真。并基于嵌入帧平均信噪比和提取准确率估计给出了该算法的参数参考值。实验证明该算法不仅能够用于抗低码率低采样高压缩比的MP3编解码,且具有良好的不可感知性和嵌入容量,同时还进行了AAC编解码测试。

关键词: MP3; 音频水印; 同步机制; 低频能量比; 量化调制

中图分类号: TP391

文献标识码: A

文章编号: 1009-5896(2008)11-2678-06

Audio Watermarking for MP3 Coding with Low Bitrate and Low Sampling Frequency

Huang Hao Guo Li Li Lin

(Dept. of Electronic Science and Technology, University of Science and Technology of China, Hefei 230027, China)

Abstract: MP3 coding has become the most popular coding method for audio signals. The distortions of MP3 coding from time and frequency domain based on experiments are explored. An audio watermarking method is proposed: besides synchronization by extracting "host signal", secret bit is embedded by quantizing the proportion of low frequency energy of two sequential frames. Then reference parameter values by evaluation of each embedding frame were suggested. The experiment results show the imperceptability and the robustness against MP3 with a low bit rate and low sampling frequency and against AAC coding, also with a satisfying capacity.

Key words: MP3; Audio watermarking; Synchronization; Proportion of low frequency energy; Quantization modulation

1 引言

随着数字音频技术的发展,水印技术得到越来越多的研究与关注^[1]。由于人类听觉系统(HAS)极为灵敏,音频感知冗余较小,水印同时满足隐蔽性和鲁棒性条件的困难较大,音频水印的研究相比较于图像水印更具有挑战性。

MP3(MPEG Layer3)音频压缩算法因其出色的压缩率和音质,随着网络的普及已经成为数字音频的主流压缩方式,适用于MP3编解码的音频水印算法对于数字音频的版权保护和认证具有重要的现实意义。IFPI(International Federation of the Phonographic Industry)在其音频水印鲁棒性定义中就明确要求了对MP3压缩攻击的抵抗性能^[2]。

现有MP3音频水印算法大致可以分为两类:一类为压缩中和压缩后嵌入,其实质是通过修改MP3的编码中的感知非敏感信息来嵌入秘密信息,可以获得较好隐蔽性和较大的嵌入容量,MP3Stego隐写工具就采用了这种方法^[3]。但是这类算法只是“躲过”了有损压缩,本身并不能抵抗MP3压缩算法,鲁棒性差。严格的说,这类方法更近于“隐写”

而不是“水印”。另一类为压缩前嵌入,一般通过修改频域嵌入来获得较好的鲁棒性。尽管不少此类算法都声称其对MP3具有良好的抵抗性能,但都缺乏变采样率MP3编解码测试,对MP3攻击的测试也较简单,码率设置较高,压缩比较低,或者不说明压缩参数,无法证明其同样适用于高压缩比的MP3编码;且水印算法设计仅考虑了有损压缩带来的波形失真,嵌入与提取缺乏相应的同步机制^[4-7]。文献[8,9]虽然在音频水印系统中引入了同步信号,缺点是所用的算法对音频能量幅度的变化非常敏感^[10],不适合压缩编码。

本文采用多种MP3工具分析了编解码后信号时域和频域的变化,并以此为基础提出了一种新的压缩前嵌入的音频水印算法:通过载体信号划分来实现嵌入和提取的同步;根据编解码前后低频能量的稳定性,通过量化相邻子帧低频能量之比来嵌入水印信息,然后基于每一嵌入帧进行了不可感知性和提取准确率评价,并给出了本算法参数针对不同码率和采样率MP3编解码的参考值。最后,实验证明本水印算法不仅能抵抗多码率变采样率的MP3压缩,且具有良好的不可感知性和较高的嵌入容量,对AAC编解码也具有较弱的鲁棒性。

2 MP3 编解码对音频信号的影响

原始音频信号经过 MP3 编解码后在时域和频域都发生了改变, 时域变化主要受编码器影响, 频域失真则与压缩采用的码率和采样频率有关。实验用原始音频都为从 CD 抓取的 wav 格式音频, 44.1kHz 采样频率(下文中如不特别声明, 原始信号采样频率默认为 44.1kHz), 16bit 精度, MP3 工具包括了 Cooledit2000Pro, iTunes, MP3 Audio Converter, Winamp Pro v5.34。下文如不特别声明, MP3 编解码工具采用 Cooledit2000Pro。

2.1 时域变化

任意截取长度为 0.5s 的原始音频片段用 Cooledit2000Pro 经过 128kbps, 44.1kHz 的 MP3 编解码后音频时长从 0.5s 变为 0.574s, 前面增加 1635 个样点。这些增加的样点大部分为零值点, 少量为头尾帧边缘效应数据。MP3 编码时采用的 MDCT 为正交重叠变换, 头尾帧变换时需补零, 从而产生边缘效应, MP3 编解码后这部分数据也被加到音频中。为了表述方便, 这里把前后增加的样点分别称为前边缘信号和后边缘信号, 而把中间的样点称为有效信号, 下文同。前边缘信号长度与后边缘信号大致相等, 边缘信号长度与 MP3 编码器有关, 表 1 是不同的 MP3 编解码工具音频的前边缘信号样点个数。

经过 MP3 编解码后进行水印提取时, 由于边缘信号的加入, 嵌入位置在时域上已经改变, 须要进行重定位。若有效信号本身的开始或者结尾就为一系列零值样点, 那么经过 MP3 编解码后这些有效信号的零值样点会与边缘信号混淆, 更难确定时域的提取起始点。

2.2 频域失真

MP3 编码是一种有损压缩^[1], 将每帧样点分为等宽子

带, 每个子带根据心理声学模型计算信掩比, 再根据编码码率进行比特分配, 得到量化因子, 最后对 MDCT 系数进行量化和 Huffman 编码。若进行变采样率 MP3 压缩, 还会对信号进行减采样。

MP3 编码主要是利用声音的频域掩蔽效应, 使量化噪声处于频域的掩蔽阈值之下, 去除了音频中的感知冗余成分, 包括一些能量很小的频率成分。低频(<3kHz)部分集中了音频信号大部分能量, 可感知的频率成分最多, 虽然掩蔽阈值较大, 但信掩比(signal-to-mask ratio, SMR, 表示信号声强与最小掩蔽阈值之比)相对较大, 所以低频能量经过 MP3 编解码后失真很小。反之高频(>10kHz)部分能量很小, 可感知的频率分量也相对较少, 虽然掩蔽阈值较小, 但信掩比相对较小, 高频能量经过 MP3 编解码后失真较大。如果采用变采样率 MP3 编解码, 随着采样频率的降低, 高频部分损失会进一步增大。

随着编码码率和采样频率的降低, 压缩比增大, 高频能量失真幅度也显著增大, 而低频能量失真很小且几乎稳定。表 2 用 SNRL 表示一段交响乐音频信号经 MP3 编解码后 3kHz 以下频率能量的平均信噪比, SNRH 表示经 MP3 编解码后 10kHz 以上频率能量的平均信噪比, FFT 分析窗长 1024 点。其中对变采样 MP3 编解码信号进行了重采样, 以恢复到原始采样频率。

由表 2 可见, 经 MP3 编解码后, 10k 以上频率能量变化幅度最大可达到 33.24%, 而 3kHz 以下频率能量变化幅度最大只有 4.19%。更多实验显示, 若音频信号本身能量较大且高频成分丰富, 例如一段节奏很快的摇滚乐, 经 MP3 编码后, 10kHz 以上频率能量变化幅度最大可以超过 90%, 而 3kHz 以下频率能量变化幅度仍然不超过 4%。故音频的低频能量对 MP3 编解码具有很好的鲁棒性。

表 1 不同 MP3 编码工具的前边缘信号样点个数

| 时长(码率/采样频率) | cooledit2000pro | iTunes | MP3 Audio Converter | Winamp Pro v5.34 |
|---------------------|-----------------|--------|---------------------|------------------|
| 5s(128kbps/44.1k) | 1635 | 1154 | 1106 | 1543 |
| 10s (96kbps/32k) | 1679 | 1161 | 1105 | 1539 |
| 30s(64kbps/32k) | 1677 | 1157 | 1111 | 1545 |
| 1min(48kbps/22.05k) | 1671 | 1160 | 1110 | 1547 |
| 2min(24kbps/16k) | 1666 | 1167 | 1109 | 1556 |

表 2 经 MP3 编解码后高低频能量的变化

| 码率/采样频率(Hz) | SNRH(dB) | SNRL(dB) | 码率/采样频率(Hz) | SNRH(dB) | SNRL(dB) |
|-------------|----------|----------|-------------|----------|----------|
| 160k/32k | 24.5198 | 38.2529 | 48k/24k | 8.6524 | 31.9353 |
| 128k/32k | 24.1231 | 37.5306 | 48k/22.05k | 8.1023 | 32.1759 |
| 96k/32k | 23.9023 | 36.4089 | 24k/24k | 7.0128 | 32.3283 |
| 64k/32k | 17.8682 | 32.9295 | 24k/22.05k | 8.9630 | 31.9171 |
| 48k/32k | 9.4727 | 31.7354 | 24k/16k | 9.2652 | 32.8147 |

3 水印原理与算法

本文针对音频信号经过 MP3 编解码后的时域变化，在嵌入端和提取端都进行载体信号划分来实现同步；根据低频能量经编解码后变化很小的特点，通过量化相邻子帧低频能量比来嵌入水印信息。然后对每一嵌入帧进行不可感知性评价和提取准确率估计，并给出了本算法参数针对不同码率和采样频率 MP3 编解码的参考值。

3.1 嵌入与提取同步

同步不仅需要提取时除去 MP3 编解码增加的边缘信号，还要避免边缘信号与原始信号本身头尾部分可能的安静段相混淆。

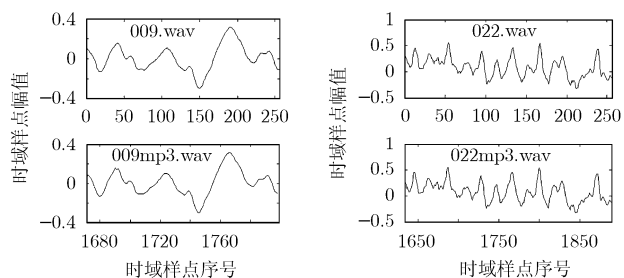
由于边缘信号大部分样点为零，其非零样点的幅值也远小于有效信号样点幅值的平均值，且原始信号本身的能量过小部分也不适合进行水印嵌入。故水印在嵌入和提取前都采用同样方法划分出载体信号部分，嵌入和提取都只对载体信号进行，以实现水印嵌入和提取的同步。载体信号的起始点定义为从信号第一个样点向后搜索到的第一个幅值不小于所有样点幅值平均值 1/3 倍的样点；载体信号的截止点定义为从信号最后一个样点向前搜索到的第一个幅值不小于所有样点幅值平均值 1/3 倍的样点。

如果嵌入水印的音频信号在提取前经过了变采样率的 MP3 编解码，水印提取前还需要对载体信号重采样，恢复到原始采样频率。重采样得到的载体信号与原始载体信号相比发生轻微的线性拉伸，频域嵌入水印可以忽略拉伸影响。(采用 11.025kHz 的 MP3 编码，拉伸幅度 < 0.05%)

采用以上方法先对音频信号 009.wav 和 022.wav 划分出载体信号。图 1 显示了经 MP3 编解码后得到的载体信号的前 256 点波形与原载体信号前 256 点波形的对比情况。

3.2 水印嵌入

音频信号能量短时平稳，且音频信号能量主要集中在低



(a) 009.wav 为原始音频划分出的载体信号前 256 点波形。对原始音频进行 24kbs/22.05kHz 的 MP3 压缩，然后划分出解码后所得音频信号的载体信号，再重采样到原始音频的采样频率，最终所得载体信号的前 256 点波形如 009mp3.wav 所示

(b) 022.wav 采用 96kbs/32kHz 的 MP3 编解码后经同步处理后所得波形

图 1 载体信号划分的同步效果

频部分，故两段相邻的短时音频信号低频能量比发生剧烈变化的可能性很小。本文对任意 20 段长度超过 3min 采样频率 44.1kHz 不同类型音频的实验，对于帧长 512 点的前后两帧，3k 以下频率能量之比主要集中在 0.5~1.5 范围内，占有帧的 83.74%。由于其分布集中，故易于采用单一量化因子进行量化。根据 2.2 节所述，MP3 编解码对低频能量影响较小。因此，本文通过量化相邻子帧低频能量比来嵌入水印信息。本算法将低频能量定义为 3kHz 以下频率能量。

原始音频信号采样频率为 F_s ，先对划分出的载体信号分帧，帧长为 N ， $N = 2n$ 。将当前帧信号 S 平均分为前后两子帧信号 X 和 Y ，帧长各为 n 。

计算 $k = \lfloor n \times 3000 / F_s \rfloor$ ， $\lfloor \cdot \rfloor$ 为向下取整操作。

求子帧 X 的低频能量 E_x 。对 X 做 FFT 得到 n 个 FFT 系数 $C_x = c_{x1}, c_{x2}, \dots, c_{xn}$ ， $E_x = \sum_{i=1}^k c_{xi}^2$ 。同样求子帧 Y 的低

频能量 E_y ，设 E_y 的 FFT 系数为 $C_y = c_{y1}, c_{y2}, \dots, c_{yn}$ 。

计算当前帧前后子帧低频能量比 $r = E_y / E_x$ 。

根据当前帧嵌入比特 b (0 或者 1) 对 r 进行量化得到 \hat{r} ：

$$\hat{r} = \left\lfloor \text{round} \left(\frac{r - b}{2} \right) \right\rfloor \times 2q + bq, \text{ 若计算得 } \hat{r} = 0, \text{ 则设}$$

$\hat{r} = 2q$ 。其中 b, r 不能同时为 0。

其中 $\text{round}()$ 为四舍五入函数， q 为预先设定的量化因子。

即如果 $b = 0$ ，则将 r 量化为距离最近的 q 的偶数倍值；如果 $b = 1$ ，则将 r 量化为距离最近的 q 的奇数倍值。为了防止量化后出现 $\hat{r} = 0$ ，一律将 $\hat{r} = 0$ 替换为 $\hat{r} = 2q$ 。

计算 $a = \sqrt{\hat{r}/r}$ ，进而算得谱系数变化比例 $d = (a - 1) / (a + 1)$ 。然后对子帧信号 X 的 FFT 系数 C_x 和子帧信号 Y 的 FFT 系数 C_y 进行缩放：

$$\hat{C}_x = (1 - d)C_x, \text{ 即 } \hat{c}_{xi} = (1 - d)c_{xi}; \hat{C}_y = (1 + d)C_y, \text{ 即 } \hat{c}_{yi} = (1 + d)c_{yi}.$$

最后对 \hat{C}_x 做 IFFT 得到嵌入后的前一子帧信号 \hat{X} ，对 \hat{C}_y 做 IFFT 得到嵌入后的后一子帧信号 \hat{Y} 。 \hat{X} 与 \hat{Y} 合成为嵌入后的当前帧信号 \hat{S} 。

水印嵌入方案如图 2 所示。

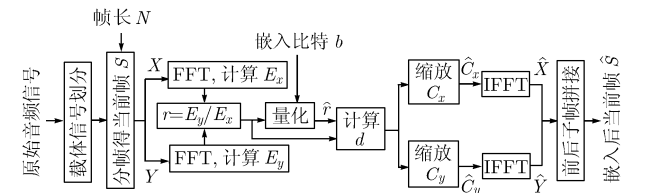


图 2 水印嵌入方案

3.3 水印提取

提取端先划分出载体信号，如果有必要则先重采样，恢复为原始采样频率。然后对载体信号以帧长 N 分帧， $N = 2n$ ，得到当前帧信号 \hat{S} ，将 \hat{S} 平均分为前后两子帧信号

\hat{X} 和 \hat{Y} , 帧长各为 n 。

采用与水印嵌入过程中相同的方法计算得到前后子帧低频能量比 \hat{r} , 则当前帧嵌入比特为 b 有

$$b = \text{round}(\hat{r}/q) \bmod 2$$

其中 $\bmod 2$ 为模 2 运算。

3.4 不可感知性评价, 提取准确率估计与参数参考值

水印算法参数包括了量化因子 q 和帧长 N , 基于不可感知性评价和提取准确率估计, 给出针对不同码率和采样频率 MP3 编解码的 q 和 N 的参考值。

(1)不可感知性评价 由嵌入算法可知, 在水印信息嵌入后, 当前帧的 X 子帧引入噪声的 FFT 谱系数为 $\text{FFT}(\hat{X} - X) = \text{FFT}(\hat{X}) - \text{FFT}(X) = \hat{C}_x - C_x = -dC_x$ 。 X 子帧信噪比为

$$\text{SNR}_x = 10 \lg \left[\frac{C \times C^T}{(-dC) \times (-dC)^T} \right] = -20 \lg |d|$$

同理可以算得 Y 子帧信噪比 $\text{SNR}_y = \text{SNR}_x = -20 \lg |d|$, 当前帧信噪比 $\text{SNR} = -20 \lg |d|$ 。 $|d|$ 值由随量化因子 q 增大而增大, 但几乎不受帧长 N 的影响。故嵌入帧信噪比随 q 增大而减小, 不可感知性随 q 增大而降低。

采用量化因子 q 对音频嵌入长度不小于 1000 的伪随机序列, 对嵌入帧信噪比求平均值可得嵌入帧平均信噪比 AFSNR, 以此来评价量化因子 q 的不可感知性。

$N = 1024$ 时 AFSNR 随 q 变化如图 3 所示。

(2)提取正确率估计 采用量化因子 q , 帧长 N 对音频嵌入长度不小于 1000 的伪随机序列, 得到量化后的每帧低频能量比 $\hat{\mathbf{R}} = \hat{r}_1, \hat{r}_2, \dots, \hat{r}_m$, m 为嵌入比特数, 再确定码率与采样频率对嵌入后音频进行 MP3 编解码, 计算编解码后每帧低频能量比 $\hat{\mathbf{R}}' = \hat{r}'_1, \hat{r}'_2, \dots, \hat{r}'_m$, 计算编解码后每帧低频能量比变化 $\Delta \hat{\mathbf{R}} = \hat{\mathbf{R}}' - \hat{\mathbf{R}}$, 即 $\Delta \hat{r}_i = \hat{r}'_i - \hat{r}_i$ 。 q 为水印嵌入后低频能量比的最小量化间隔, 只要满足 $q > 2|\Delta \hat{r}_i|$ 则可以保证水印的第 i 个比特在 MP3 编解码后被正确提取。设当前满足 $q > 2|\Delta \hat{r}_i|$ 条件的帧数为 w , 则量化因子 q , 帧长 N 的水印对当前码率与采样频率的 MP3 编解码提取正确率估计为 $\text{PA} = w/m$ 。

随着 MP3 编解码码率和采样频率的降低, $|\Delta \hat{r}_i|$ 会随之增大。 q 越大, 满足 $q > 2|\Delta \hat{r}_i|$ 条件的帧越多, 鲁棒性越好。同时 $|\Delta \hat{r}_i|$ 与帧长 N 成反比, 鲁棒性与 N 成正比。故可以通过增大 q 和 N 提高提取准确率。

采用 96kbps, 44.1kHz 的 MP3 编解码, 在 $q = 0.3$ 时 $|\Delta \hat{\mathbf{R}}|$ 的均值 ADR 随帧长 N 变化如图 4 所示。

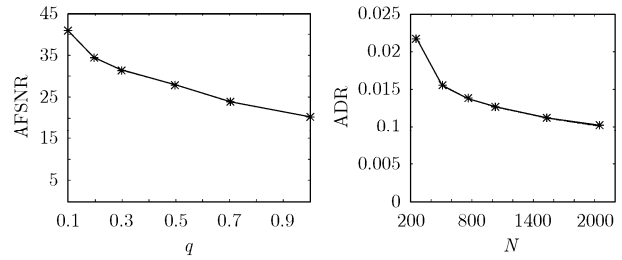


图 3 AFSNR 随 q 变化示意图 图 4 ADR 随帧长 N 变化示意图

(3) q, N 的参考值 q 决定了水印的嵌入强度, 与水印鲁棒性成正比, 与不可感知性成反比。 N 与鲁棒性成正比, 与水印信号的数据带宽成反比。

针对不同码率和采样频率的 MP3 编解码, 根据嵌入帧平均信噪比 AFSNR 和提取准确率估计 PA, 给出的 q, N 参考值见表 3。

每组参考值首先确保 q 的取值使得 AFSNR 不小于 30 dB。再根据当前 MP3 编解码的码率和采样频率以及 q , 确定 N , 保证 PA 不低于 95%。表 3 只提供了部分码率和采样频率以及对应压缩比的 q, N 参考值, 其它码率与采样频率的 q, N 可以参考进行设置, 例如针对 96kbps 码率, 32kHz 采样频率, 压缩比为 7.4:1 的 MP3 编码, 参数参考值为 $q = 0.3, N = 512$ 。

q, N 的设置也可以根据具体的应用需求做适当调整, 例如为了增大水印信道的数据带宽, 在不可感知性许可的范围内, 可以适当增大 q , 减小 N 。

4 实验结果

实验从主观感知和鲁棒性两方面对本文所提算法进行了测试。鲁棒性测试中针对不同码率和采样频率的 MP3 编解码采用不同工具, 对本算法进行了测试, 最后采用 AAC 压缩编码对本文所提算法进行了测试。

4.1 不可感知性测试

水印的不可感知性由量化因子 q 决定。采用 $q = 0.3$ 对 10 段不同类型的音频(包括交响乐, 轻音乐, 摇滚乐, 乡村乐, 流行乐, 说唱乐, 相声, 男女声独唱)进行嵌入, 水印长

表 3 q 和 N 参考值

| 码率(bps)/采样频率(Hz) | 压缩比 | q | N | AFSNR(dB) | PA(%) |
|------------------|--------|-----|------|-----------|-------|
| 112k/32k | 6.3:1 | 0.3 | 256 | 32.3404 | 98.76 |
| 80k/32k | 8.8:1 | 0.3 | 512 | 31.4676 | 99.83 |
| 56k/32k | 12.6:1 | 0.3 | 1024 | 31.2408 | 96.29 |
| 24k/16k | 29.4:1 | 0.3 | 1024 | 31.2408 | 97.14 |
| 16k/11.025k | 44.1:1 | 0.3 | 2048 | 31.8468 | 95.34 |
| 8k/8k | 88.2:1 | 0.3 | 6144 | 31.3282 | 96.72 |

表4 MP3 编解码攻击实验结果

| 码率(bps)/采样频率(Hz) | 压缩比 | q | N | DA | BW (bps) |
|------------------|--------|-----|------|-------|----------|
| 160k/44.1k(1) | 4.4:1 | 0.3 | 256 | 98.9% | 172.2656 |
| 128/44.1k(2) | 5.5:1 | 0.3 | 256 | 97.2% | 172.2656 |
| 112/32k(3) | 6.3:1 | 0.3 | 256 | 97.6% | 172.2656 |
| 96k/44.1k(4) | 7.4:1 | 0.3 | 512 | 99.0% | 86.1328 |
| 80k/32k(2) | 8.8:1 | 0.3 | 512 | 96.5% | 86.1328 |
| 64k/44.1k(2) | 11:1 | 0.3 | 1024 | 96.6% | 43.0664 |
| 56k/32k(1) | 12.6:1 | 0.3 | 1024 | 93.4% | 43.0664 |
| 48k/24k(3) | 14.7:1 | 0.3 | 1024 | 97.4% | 43.0664 |
| 32k/22.05k(3) | 22.1:1 | 0.3 | 1024 | 94.5% | 43.0664 |
| 24k/16k(3) | 29.4:1 | 0.3 | 1024 | 97.2% | 43.0664 |
| 18k/16k(1) | 39.2:1 | 0.3 | 2048 | 92.4% | 21.5332 |
| 16k/11.025k(1) | 44.1:1 | 0.3 | 2048 | 94.2% | 21.5332 |
| 8k/8k(1) | 88.2:1 | 0.3 | 6144 | 92.3% | 7.1777 |

注: 1 DA 为 10 段音频实验结果的平均值。

2 表中的(1)(2)(3)(4)表示采用的 MP3 编解码工具, (1) Cooledit2000Pro, (2)iTunes, (3)MP3Audio Converter, (4)Winamp Pro v5.34。

度不小于 1000bit, 主观感知测试邀请 5 男 2 女对比嵌入前后音频进行主观评价, 所得 70 个评价中, “感知无变化”占 97.14%, 嵌入帧信噪比平均值为 31.8658 dB, 方差为 40.4682 dB²; 采用 $q = 0.5$ 重复上述实验, “感知无变化”评价占 88.57%, 嵌入帧信噪比平均值为 26.9505 dB, 方差为 48.8409 dB²。

4.2 MP3 编解码测试

实验采用 10 段采样频率 44.1kHz 的不同类型的 wav 格式音频。采用参数 q , N 进行嵌入, 水印长度不小于 1000bit 水印, 采用不同工具进行 MP3 编解码后实际的提取准确率为 DA, 单声道水印信道带宽(嵌入容量)为 BW。实验结果如表 4 所示。

实验结果显示本文所提水印算法对 MP3 编解码具有很好的抵抗性能。且能提供较高的水印信道带宽, 当 $N \leq 512$ 时, 单声道水印信道带宽 $BW > 86$ bps, 高于 MP3stego 隐写工具(76bps)。

4.3 AAC 编解码测试

实验采用与 4.2 节中相同的实验 10 段音频嵌入长度不小于 1000bit 水印, 然后对其进行 AAC 编解码, 提取准确率的实验结果如表 5 所示。

实验结果显示, 本算法当嵌入帧的较长时($N > 2048$), 对 AAC 也具有鲁棒性, 但此时水印信道带宽较窄 ($BW < 21.5332$ bps)。

5 结束语

在实验分析基础上, 本文从时域和频域探讨了音频在 MP3 编解码过程中发生的变化, 根据低频能量的鲁棒性, 采用载体信号划分来进行嵌入和提取同步, 对相邻子帧的低频能量比进行量化来嵌入水印信息。

表5 AAC 编解码提取准确率实验结果(%)

| | N | | | |
|---------------|------|------|------|------|
| | 256 | 512 | 1024 | 2048 |
| $q=0.3$ 实验(1) | 70.1 | 82.1 | 92.4 | 97.2 |
| $q=0.3$ 实验(2) | 64.3 | 71.6 | 83.6 | 92.5 |

注: 1 表中的提取准确率为 10 段音频实验结果的平均值。

2 实验(1)中 AAC 编码不限制编码码率和频带宽度, 实验(2)中 AAC 编码限制了 32kbps 的平均码率和 16kHz 的频带宽度。

实验证明, 本水印算法对 MP3 编解码具有很强的鲁棒性, 且能提供较高的水印信道带宽, 同时采用长帧嵌入的本算法对 AAC 编解码也有较强的鲁棒性, 有较强的参考价值 and 实用价值。

本算法需要采用较长的嵌入帧才能获得对高压缩比的 MP3 和 AAC 编码算法的鲁棒性, 嵌入容量随压缩比增大而显著降低, 后续工作将力图在保证鲁棒性的前提下, 进一步提高嵌入容量。

参考文献

- [1] Arnold M. Audio watermarking: Features, applications and algorithms. In: Proceedings of IEEE International Conference on Multimedia & Expo, New York, USA, 2000, 2: 1013-1016.
- [2] Katzenbeisser S, Fabien A, Petitcolas P 编, 吴秋新, 钮心忻, 杨义先, 罗守山, 杨晓兵译, 信息隐藏技术——隐写术与数字水印. 北京: 人民邮电出版社, 2001: 104-105.
- [3] Petitcolas P. Computer Laboratory, Cambridge, MP3 Stego. <http://www.cl.cam.ac.uk/~fapp2/steganography/mp3stego/>, 1998.
- [4] Lie Wen-nung and Chang Li-chun. Robust and high quality timedomain audio watermarking subject to psychoacoustic masking. In: Proceedings of IEEE International Symposium

- on Circuits and Systems, Arizona, USA, 2002, 2: 45-48.
- [5] Seok Jongwon, Hong Jinwoo, and Kim Jinwoong. A novel audio watermarking algorithm for copy-right protection of digital audio. *ETRI Journal*, 2002, 24(3): 181-189.
- [6] Shin Seungwon, Kim Oanjin, Kim Jongweon, and Choi Jonguk. A robust audio watermarking algorithm using pitch scaling. In: *Proceedings of IEEE Workshop on Digital Signal Processing*, 2002, Vol.2: 701-704.
- [7] 刘伟, 王朔中, 张新鹏. 一种基于部分 mp3 编码原理的音频水印. *中山大学学报(自然科学版)*, 2004 年, 43 增刊(2): 26-33.
- Liu Wei, Wang Shuo-zhong, and Zhang Xin-peng. Audio watermarking based on partial MP3 encoding. *Acta Scientiarum Naturalium Universitatis Sunyatseni*, 2004, 43 Suppl.(2): 26-33.
- [8] Huang Ji-Wu, WangYong, and Shi Y Q. A blind audio watermarking algorithm with self-synchronization. In: *Proceedings of IEEE International Symposium on Circuits and Systems*, Arizona, USA, 2002, 3 : 627-630.
- [9] Wu Shao-quan, Huang Ji-wu, Huang Da-ren, and Shi Yun Q. Efficiently self-synchronized audio watermarking for assuredaudio data t ransmission. *IEEE Trans. on Broadcasting*, 2005, 51(1): 69-76.
- [10] 项世军, 黄继武, 王永雄. 一种抗AD/DA变换的音频水印算法. *计算机学报*, 2006年, 29(2): 308-316.
- Xiang Shi-jun, Huang Ji-wu, and Wang Yong-xiong. An audio watermarking algorithm against the D/A and A/D conversions. *Chinese Journal of Computers*, 2006, 29(2): 308-316.
- [11] Koukopoulos D K and Stamatiou Y C. A compressed domain watermarking algorithm for mpeg audio layer 3. *ACM Proceedings of the 2001 workshop on Multimedia and security: new challenges*, Ottawa, Ontario, Canada, 2001: 7-10.
- 黄 昊: 男, 1981年生, 博士生, 研究方向为数字声学 with 音频信息隐藏.
- 郭 立: 男, 1946年生, 教授, 博士生导师, 主要研究方向为多媒体处理与信息安全.
- 李 琳: 女, 1982 年生, 博士生, 研究方向为多媒体压缩与硬件实现.