

基于 DCT 系数的量化音频水印

潘洋绪 刘 猛

(河海大学 计算机与信息学院 南京 211100)

摘要: 近年许多水印算法被提出并应用到图像和视频中去,相较于人类视觉系统,听觉系统的敏感性使得音频水印的实现要难于图像视频水印。为了提高音频水印的不可见性,提出一种离散余弦变换(DCT)的量化音频水印算法,先是基于能量将原始音频分为两部分,对能量较大的部分进行分块离散余弦变换,然后对变换之后的中频系数均值采用量化索引调制(QIM)嵌入 Arnold 置乱后的二进制水印图片,保证水印的嵌入对原始音频的影响在听觉觉察范围之内。实验结果表明,算法具有很好的不可见性,并且对 MP3 压缩和噪声具有较好的鲁棒性。

关键词: 音频水印; Arnold 变换; 量化索引调制; 离散余弦变换

中图分类号: TP391 **文献标识码:** A **国家标准学科分类代码:** 510.40

Quantization audio watermarking based on DCT coefficients

Pan Yangxu Liu Meng

(Department of Computer and Information, Hohai University, Nanjing 211100, China)

Abstract: Nowadays many watermarking algorithms have been proposed and implemented for images and video. Compared to the human visual system, the sensitivity of the auditory system makes the audio watermarking more difficult to realize. In order to improve the invisibility of the audio watermarking, this paper presents an audio watermarking algorithm which is based on the quantization of discrete cosine transform (DCT), first the original audio is divided into two parts due to the energy, and do the discrete cosine transform on the larger part, then the Arnold scrambled binary watermark image is embedded into the mean value of the middle frequency coefficients using quantitative index modulation (QIM), And ensure that the embedding watermark on the original audio within the range of hearing range. Experimental results show that the algorithm has good invisibility, and have robust of MP3 compression and noise.

Keywords: audio watermarking; Arnold; quantitative index modulation; discrete cosine transform

1 引言

随着互联网技术的广泛使用,非法复制和破坏变得越来越难以控制。在利用计算机网络技术的同时,如何有效地保护版权拥有者的利益,已经成为迫在眉睫的问题,数字水印技术被认为是版权保护的有利措施,已经引起了各界学者的重视^[1-3]。

音频水印按照嵌入区域的不同可分为时域和变换域,变换域因其具有的良好特性,水印算法的效果要普遍优于时域算法。目前为止,时域音频水印算法较为成熟的有:回声隐藏法、最不重要性法(LSB)、相位编码法、扩频法;变换域水印应用较多的有离散小波变换(DWT)、离散傅里叶变换(DFT)、离散余弦变换(DCT)等^[4-5]。Yan Yang^[6]将离散余弦变换之后的单个交流系数作为水印嵌入区域,算法

具有较好的不可见性和鲁棒性。Lalitha N V^[7]为了提高水印的验证性能,将水印进行 Arnold 置乱,然后再利用 DCT 系数进行水印嵌入提取。Yin C^[8]将载体音频分为两部分,比较两部分离散余弦变换后的中频系数,选择较大的系数作为水印嵌入区域,不仅能鲁棒性较好,也可以有效地解决同步攻击问题。

由于 DCT 变换的高频系数在经过信号压缩(如 MP3 压缩)之后会产生损失,将使水印提取效果受到影响^[9],同时对于低频系数而言,虽然水印可辨,但若嵌入强度过大将会对宿主音频产生较大影响。为了更好的权衡音频水印的不可见性和鲁棒性,这里我们选择离散余弦变换后的中频系数作为水印嵌入区域,来实现水印的嵌入和提取。

2 量化索引调制理论

QIM 水印算法的基本思想是根据水印信息的不同将载体数据调制到不同的量化区间,在提取时根据量化区间的特点提取水印。这里我们采用奇偶量化索引调制^[10],以相同的步长均匀量化载体数据,将水印嵌入相邻索引,使索引号的奇偶校验位与水印信号相同。量化步长作为水印嵌入强度,影响着水印的不可见性和鲁棒性。水印嵌入方法如式(1)。

设载体数据为 $S(i)$, 水印信息为 $W(i)$, 量化步长为 Δ , 嵌入水印后的数据为 $S'(i)$ 。

$$S'(i) = \begin{cases} \left\lfloor \frac{S(i) + 0.5\Delta}{2\Delta} \right\rfloor \times 2\Delta + \frac{\Delta}{2}, & W(i) = 0 \\ \left\lfloor \frac{S(i) - 0.5\Delta}{2\Delta} \right\rfloor \times 2\Delta + \frac{3\Delta}{2}, & W(i) = 1 \\ S'(i) = S'(i) + 2 \times \Delta & S'(i) < 0 \end{cases} \quad (1)$$

式中: $\lfloor \cdot \rfloor$ 表示向下取整。

嵌入算法将二进制的水印信息调制到对应的索引区间,当载体数据量化索引值与水印信号相等时,量化区间不变;反之,将载体数据调整至相邻的量化区间。为了提高水印的鲁棒性,载体数据都被调制在量化区间的中间位置。水印提取采用最小距离方法。如下式:

$$Q(i) = \left\lfloor \frac{S'(i)}{\Delta} \right\rfloor \bmod 2 \quad (2)$$

式中: $Q(i)$ 为提取的水印数据。

3 水印嵌入提取策略和算法

3.1 音频和水印预处理

将载体音频基于能量分为两部分,能量大的部分进行水印的嵌入提取,另一部分不做变换。首先将原始音频截取成为固定长度为 L 的帧,然后再对帧进行分块操作,设块大小为 N ,计算每块的平均能量,我们选择大于帧平均能量的块来进行水印的嵌入。

$$\begin{cases} \text{do watermark } e \geq E \\ \text{do nothing else} \end{cases} \quad (3)$$

式中: $E_j = \frac{N \sum S_j^2(i)}{L}$, $e_{j,k} = \sum_N S_{j,k}^2(i)$, $S_j(i)$ 表示第 j 帧的音频数据, E_j 表示第 j 帧的平均能量, $S_{j,k}(i)$ 表示第 j 帧音频中第 k 块的音频数据, $e_{j,k}$ 表示相应块的能量。

对于 $M \times M$ 大小的图像 $R(m, n)$, 进行 Arnold 变换, 如式(4), 猫脸可以使图像白噪声化, 能量均匀分布, 同时可以通过控制变换的次数来实现对水印图像的加密。然后我们利用式(5)将置乱后的二维的图像 $R'(m, n)$ 变换成一维的水印数据 $W(i)$ 。

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \pmod{M} \quad (4)$$

$$W(i) = R'(m, n) \quad 0 \leq m, n \leq M, i = m \times M + n \quad (5)$$

3.2 水印嵌入算法

本文水印嵌入主要有两部分:对水印图片和音频文件进行预处理,在 DCT 变换域中量化嵌入水印。算法详细过程如下。

1) 对原始音频和水印图片进行预处理,如 3.1 节。

2) 对选定的块进行离散余弦变换,选择 DCT 变换后的第 2、3、4 这 3 个 AC 系数用 a_i 表示,其中 $i \in (1, 2, 3)$,取它们绝对值的均值采用式(1)的量化索引调制以嵌入水印。心理模型的听觉掩蔽效应告诉我们,强信号可以很好的隐藏弱小信号,对于弱信号即使有较大的改变,也不容易引起感知的注意。我们在此理论基础上来分别求得嵌入水印后的 AC 系数^[11]。设 u 表示这 3 个系数绝对值的均值, u' 表示调制之后的值,取 $d = u' - u$, a'_i 为嵌完水印后相应的 AC 系数。定义加权系数 p_j 。

$$p_j = \frac{\sum_{i=1, i \neq j}^3 |a_i|}{2 \sum_{i=1}^3 |a_i|} \quad i = 1, 2, 3 \quad (6)$$

式中: $\sum_{j=1}^3 p_j = 1$ 。

在确定嵌入水印后相应 AC 系数的值时采用以下方法:

$$\begin{cases} a'_i = a_i + 3p_j d & a_i \geq 0 \\ a'_i = a_i - 3p_j d & a_i < 0 \end{cases} \quad (7)$$

如(6)所示,当 a_i 的值越大时, p_j 越小,相应 a'_i 的改变就越小,当 a_i 的值越小时, p_j 越大,相应 a'_i 的改变也越大。

3) 对嵌入水印后的数据采取逆离散余弦变换(IDCT),以获得嵌入水印信号。

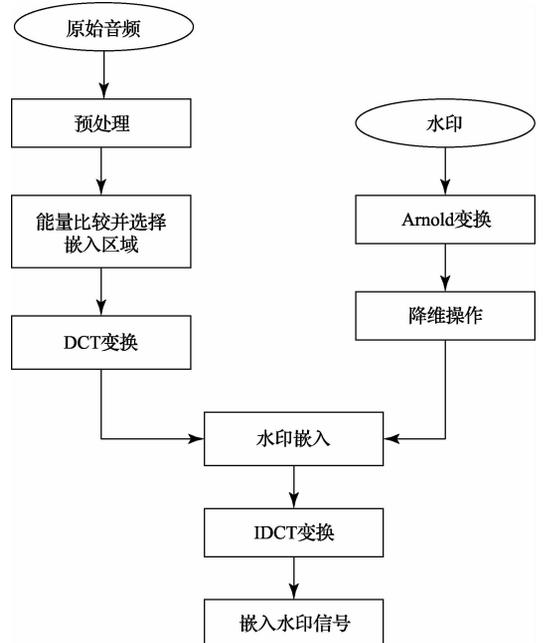


图1 水印嵌入流程图

3.3 水印提取算法

水印提取过程是水印嵌入过程的逆运算,提取水印时先进行音频信号预处理,选择提取水印区域,再根据 QIM 算法提取水印信息并恢复图片。具体步骤如下:

1)对嵌入水印音频进行 3.1 所讲的预处理,选择水印提取区域。

2)对所选择区域进行离散余弦变换,选择除第 2、3、4 三个 AC 系数计算其绝对值的均值,对计算得到的均值采用式(2)的方法来获得水印信息。

3)把获得的水印信息转换为大小为 $M \times M$ 的二维信息,并进行 Arnold 变换获得水印图片。

4 实验结果分析

通常将水印嵌入后的不可见性和鲁棒性作为评估水印算法性能的重要指标。鲁棒性是指水印抵抗各种攻击后仍能提取出水印信息能力。归一化相关系数 NC 表示提取水印图片和原始水印图片的相似性, NC 值越大,水印鲁棒性越好。

$$NC = \frac{\sum_{m=1}^M \sum_{n=1}^N R(m,n) \times R'(m,n)}{\sqrt{\sum_{m=1}^M \sum_{n=1}^N R^2(m,n)} \sqrt{\sum_{m=1}^M \sum_{n=1}^N R'^2(m,n)}} \quad (8)$$

$0 \leq NC \leq 1$

不可见性是指水印嵌入后的载体没有明显地降质现象,嵌入的水印信息无法认为的感受到。这里我们用信噪比(SNR)来表示不可见性,单位为 dB。

$$SNR(S(i), S'(i)) = 10 \log_{10} \left(\frac{\sum_{i=0}^{length} S^2(i)}{\sum_{i=0}^{length} (S(i) - S'(i))^2} \right) \quad (9)$$

实验中我们用采样率为 44.1 kHz,量化精度为 16 bit,时长 30 s 的单声道音频,水印图片为 128×128 大小的二进制图像。量化步长 Δ 取 0.08,实验仿真结果得到 NC 值为 0.9997,信噪比 SNR 的值为 66.283 9 dB。为了检测水印的不可见性,我们分别以流行、古典、民谣、乡村音乐为嵌入载体实现水印嵌入方案,实验得到的 SNR 值分别为 75.9907、67.2807、66.8184、73.8886 dB,结果显示本实验具有较好的不可见性。



图 2 原始及提取的水印图片

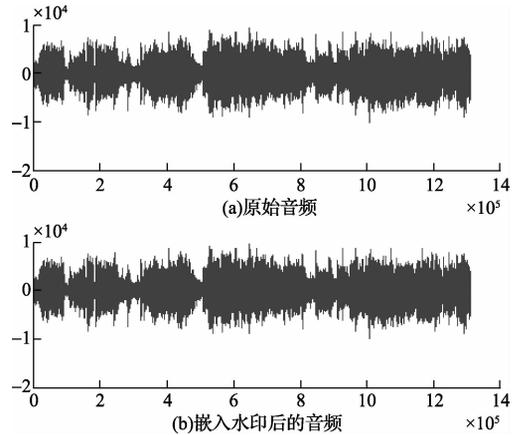


图 3 原始及嵌入水印后的音频

为了检测该算法的鲁棒性,我们对嵌入水印音频做以下处理:

- 1) 采样率转换:将包含水印的音频先降采样到 22.05 kHz,然后在升采样 44.1 kHz。
- 2) 高斯白噪声:对已嵌水印音频添加 33 dB 的高斯白噪声。
- 3) 低通滤波:巴特沃斯低通滤波器,截止频率为 10 kHz。
- 4) MP3 压缩:压缩频率为 128 kbps。

表 1 仿真结果

攻击	NC 值				
	原始	流行	古典	民谣	乡村
没有攻击	0.999 7	0.999 9	0.999 7	1	0.999 9
高斯白噪声	0.841 8	0.845 5	0.839 8	0.838 7	0.838 9
采样率转换	0.745 5	0.739 9	0.732 9	0.738 9	0.745 2
低通滤波	0.641 8	0.654 3	0.643 6	0.638 0	0.639 6
MP3 压缩	0.873 6	0.884 7	0.898 7	0.872 3	0.898 2

攻击仿真结果 NC 值如表 1。从实验数据可以看出,本实验的鲁棒性和不可见性相比较要差,算法对噪声和 MP3 压缩鲁棒性相对较好。

5 结 论

本文提出一种基于 DCT 变换域 AC 系数的均值量化水印方案,为了增加水印的不可见性,我们将音频基于能量分为两部分,在能量大的部分进行 DCT 变换,之后选取 3 个 AC 系数利用量化索引调制进行水印嵌入提取,实验结果表明,该算法具有较好的不可见性和对噪声、MP3 压缩有较好的鲁棒性。

参考文献

[1] 刘金安,金聪.基于离散小波变换的双重水印方案[J].电子测量技术,2012,35(11):45-48.

(下转第 53 页)