

文章编号:1004-4736(2008)02-0091-03

基于多分辨分解的乐音水印算法实现

徐国庆,张彦铎,王海晖

(武汉工程大学计算机科学与工程学院,湖北 武汉 430074)

摘要:利用小波变换在时变信号的低频分解特性,提出了一种针对基音低频特性的水印嵌入算法,该水印具有良好的隐蔽性和对各种音频压缩算法的适应性,对于保护音乐作品知识产权具有较高的应用价值。

关键词:乐音;水印;小波变换;鲁棒性

中图分类号:TP 301.6 文献标识码:A

0 引言

数字水印技术在图像版权信息保护中得到了比较广泛的应用,各种水印嵌入方法也被诸多学者提出并验证。随着互联网技术的普及,原创音乐版权的保护成为一个较为突出和棘手的问题。原创音乐在网络上的肆意盗链和下载甚至篡改其中的旋律片段等违反版权法的行为层出不穷,如何有效地保护音乐知识产权,如何对违法行为进行有力的取证成为一个热点问题。

1 乐音信号特性分析

从技术角度来看,乐音信号是一维时变离散信号,而图像是二维离散信号,两者在结构上有较大的区别。重要的是,乐音有很多独有特性,诸如基音、泛音、旋律、时长,因此,乐音水印的研究应结合乐音的信号特性,并充分利用乐音信号在变换域的一些本质特征。乐音是由单音按照时间先后规律构成的,一段连续的乐音中包含诸多的单乐音音符。乐音音符音调、响度、音色特征由乐音音符的物理特性基频、振幅及倍频所确定,乐音的这种特性使其能够用物理的方法进行分析和测量。

单乐音音符的音调(即音高)是由这个音符的基频决定的;一个单音的响度(即音强)是这个单音的空气振动到达人耳处的能流。音色由单乐音音符频谱(即各谐波成份比例)构成决定。对于弦鸣乐器,系统的固有频率为:

$$f_n = \frac{n\pi}{l} \sqrt{\frac{T_0}{\rho A}}, (n=1, 2, 3, \dots)$$

当 $n=1$ 时

$$f_1 = \frac{\pi}{l} \sqrt{\frac{T_0}{\rho A}}$$

称为基频,由它所确定的音称为弦的基音,决定了音高。较高频率

$$f_n (n=2, 3, 4, \dots)$$

确定的音称为泛音,它是基音的整数倍,弦乐器如小提琴、钢琴、吉他等的音调和响度通常以基音为主。对于人的歌唱乐音,可以用语音的声管模型加以描述。按照十二平均律,相邻乐音的音高(频率)符合:

$$f_2/f_1 = \sqrt[12]{2}$$

十二音平均律为乐音音符的定量刻画提供了可靠的科学手段,这为诸多乐器发音的研究提供了一个统一的度量。乐音各频率成分中,按照能量分布,占主要部分的是基频成分,由基频的整数倍频所构成的泛音成分决定了音色的甜润程度。所以在乐音中嵌入水印,从变换域来看,应尽可能地在基频部嵌入。

2 二进离散小波变换

小波变换^[1]是应用较为广泛的信号处理手段之一,它可以很好地刻画信号时频突变信息。

设 $l(x)$ 为一维输入信号,记:

$$\phi_{jk}(x) = 2^{-j/2} \phi(2^{-j}x - k),$$

$$\psi_{jk}(x) = 2^{-j/2} \psi(2^{-j}x - k),$$

这里 $\phi(x)$ 与 $\psi(x)$ 分别称为定标函数与子波函数, $\{\phi_{jk}(x)\}$ 与 $\{\psi_{jk}(x)\}$ 为二个正交基函数的集合。记 $P_0 t = t$, 在第 j 级上的一维离散小波变换,通过正交投影 P_{jt} 与 Q_{jt} 将 $P_{j-1} t$ 分解为:

$$P_{j-1} t = P_{jt} \mid Q_{jt} = \sum_k c_k^j \phi_{jk} \mid \sum_k d_k^j \psi_{jk}$$

其中:

$$c_k^j = \sum_{n=0}^{2^j-1} h(n) c_{2k+1}^{j-1}$$

收稿日期:2007-10-08

基金项目:湖北省自然科学基金(3005ABA232)。

作者简介:徐国庆(1974-),男,江苏徐州人,讲师,硕士,研究方向:乐音识别。

$$d_k^j = \sum_{n=0}^{r-1} g(n) c_{2k+1}^{j-1}$$

$$(j=1,2,\dots,L,k=0,1,\dots,N/2^j-1)$$

这里, $\{h(n)\}$ 与 $\{g(n)\}$ 分别为低通与高通系数, 它们由基函数 $\{\phi_{jk}(x)\}$ 与 $\{\psi_{jk}(x)\}$ 来确定, r 为权系数的长度. $\{C_n^0\}$ 为信号的输入数据, N 为输入信号的长度, L 为所需的级数. 小波分解把信号分解为低频部分的近似信号和高频部分的噪声信号, 由于噪声主要集中在高频部分, 而有效能量主要集中在低频部分, 所以对原始信号实施小波变换可以在进行信号高频滤波的同时, 得到低频有效系数.

Mallat^[2] 在 MRA 的基础上给出了小波系数快速分解的金字塔算法, 并构造了用于小波分解和重构的高、低通滤波器组, 大大简化了小波系数的计算.

3 基于低频基音特性的乐音水印方法

利用二进离散小波变换, 将原始乐音信号进行频带分解, 按照文献[3,4]对乐音进行基音求解, 由于多声部乐音的基音频率不单一, 本文采用在低频频带内的能量最高的基频分量作为嵌入位置, 使用这种嵌入方法的优点是隐蔽性强, 并且具有最佳的抗高频滤波攻击, 具体的过程为:

(1) 使用二值图像作为水印, 将水印图像在时域上进行信息置乱, 以达到加密效果. 所使用的函数为:

$$P_N(k): \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ k & k \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \bmod N$$

其中, (x, y) 和 (x', y') 为图像的像素点在变换前后的位置, 每个像素点经过函数 $P_N(k)$ 变换之后, 若原始水印图像从 $n \times n$ 矩阵变换为一 $N \times N$ 矩阵.

(2) 将原始乐音作频带分解, 并求出低频的基音最大分量系数, 选择一个起始位置 P_1 , 长度为 $N \times N$ 的一维系数矩阵, 其大小与上一步水印图像变换后的矩阵的元素个数相等.

(3) 在选择的系数矩阵中嵌入水印信息, 由于是在一维信号中嵌入二维水印, 所以对于水印信息采用逐行嵌入方法, 与乐音低频基音系数进行叠加.

在运用二进小波变换进行乐音的低频分解时, 关于小波分解尺度的选择, 可以按照^[4]提供的公式求出:

$$j = \left\lceil \log_2 \left(\frac{f_s}{f_0} \right) \right\rceil$$

其中, j 为小波分解尺度, f_s 为采样频率, f_0 为基音频率, $[x]$ 表示不超过 x 的最大整数. 取 $f_s = 44\,100 \text{ Hz}$, $f_0 = 450 \text{ Hz}$, 得到 $j = 6$. 经过 $j = 6$ 尺度的小波变换后的低频分量带宽为 $44\,100/2^6 = 689.0625 \text{ Hz}$, 考虑到乐音中含有的高音乐器, 可以适当加宽此低频带, 经过实验, 分解尺度为 4 时低通上限频率为 $2\,756.25 \text{ Hz}$, 这个频带包含一般人声乐音以及乐器等高音的频带, 能量较为集中. 分解以后, 乐音仍然保持较高的可听度.

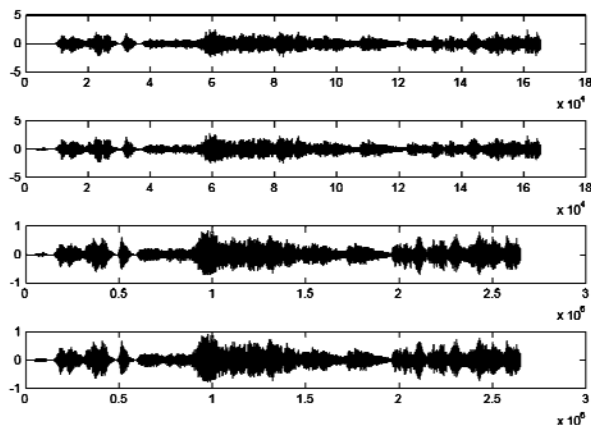


图 1 原始乐音及含水印乐音的 DWT 低频分解

Fig. 1 The original music and the watermarked music with DWT decomposition

水印的提取过程如下:

(1) 将原始乐音进行离散小波频带分解, 对于标准 CD 采样 $44\,100 \text{ Hz}$ 信号, 分解尺度为 4, 得到低频近似系数 P ; 并确定低频最大的 $N \times N$ 个系数的位置 $S(i)$.

(2) 将待测乐音进行同尺度分解, 提取同位置低频近似系数 P' .

(3) 在 $S(i)$ 位置计算 P, P' 之绝对值差 ϵ , 并与门限 δ 比较, 比较的结果记入水印图像的相应位置:

$$\text{Bin}(i, j) = \begin{cases} 1, & \epsilon > \delta \\ 0, & \epsilon < \delta \end{cases}$$

(4) 按照加密步骤(1)的逆运算, 提取 $n \times n$ 的水印二值图像.

4 水印攻击实验

乐音的常见攻击方法为: 噪声污染、降噪、滤波, 针对这几种攻击手段分别进行水印鲁棒性实验.

乐音嵌入水印后没有经过任何攻击, 从其中提取的水印如图 2 所示.

图 3 为含水印乐音被噪声污染后提取的水印, 从图中可以直观看出乐音噪声污染已经十分严重, 并且严重影响了乐音的清晰度. 但水印信息依然清晰可辨.

图1为对被噪声污染的乐音进行一级离散小波变换,并将高频信号置零后进行小波逆变换信号重构,实现低通滤波消噪后提取的水印信息。



图2 原始水印及提取的水印

Fig. 2 The original and the distilled watermark

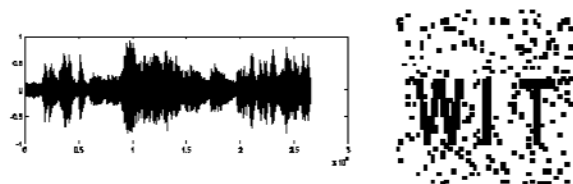


图3 原始乐音中混入白噪声及提取的水印

Fig. 3 Watermark distilled from noised music

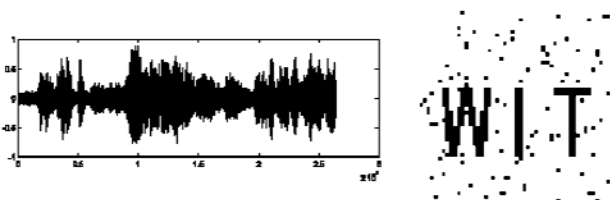


图4 乐音降噪处理及提取的水印

Fig. 4 Watermark distilled from denoised music

5 结 语

基于基频特性的乐音水印方法根据乐音的本质属性,由于乐音的大部分能量和特性集中在基频和少数几个倍频段,所以通过选择基音频率分量作为嵌入水印的区域可以实现水印的隐蔽性和良好的抗高频滤波的能力,从实验结果来看,在维持乐音的可听质量的情况下,提出的水印算法对于几种常见的乐音攻击均具有较好的稳健性,可以作为提高乐音反盗版和抗压缩、滤波等攻击的参考方法。

参考文献:

- [1] Ingrid Daubechies. Orthonormal bases of compactly supported wavelets[J]. Comm. Pure Appl Math, 1988,11(41): 909-996.
- [2] Mallat S. 信号处理的小波导引[M]. 杨力华译. 北京:机械工业出版社,2003. 150~220.
- [3] 徐国庆,杨 丹. 基于 daubechies 小波的基音检测[J]. 武汉化工学院学报,2005,27(4):55-57.
- [4] 徐国庆,杨 丹. 乐音识别方法及应用[J]. 计算机应用,2005,25(4):968-970.

Approach of the watermark embed method based on the multi-resolution of the music

XU Guo-qing, ZHANG Yan-duo, WANG Hai-hui

(School of Computer Science and Engineering, Wuhan Institute of Technology, Wuhan 430074, China)

Abstract: Fundamental frequency is an important parameter in speech analysis. This paper studies elaborates a new watermark embed method. It uses discrete wavelet transform on the analysis of the fundamental frequency of the music. The experimental results show that the watermark embed through this approach is inaudible and robust against noise. It can be used as an effect copyright protected method of music masterpieces.

Key words: music sound; audio watermarking; DWT; robustness

本文编辑:陈晓苹