

基于二元 BKF 统计建模的双树复数小波域 数字水印检测算法

王向阳¹⁾ 李 丽¹⁾ 李海芳¹⁾ 牛盼盼^{1),2)} 王思淼¹⁾ 杨红颖¹⁾

¹⁾(辽宁师范大学计算机与信息技术学院 辽宁 大连 116029)

²⁾(大连理工大学电子信息与电气工程学部 辽宁 大连 116023)

摘 要 作为多媒体作品版权保护的有效手段,数字水印技术已成为国际学术界的研究热点之一.可保持不可感知性、鲁棒性、水印容量之间良好平衡的数字水印算法研究是一项富有挑战性的工作,而基于统计模型的变换域乘性水印思想为有效解决三者之间良好平衡问题提供了可能的解决方向.该文以双树复数小波变换(Dual-Tree Complex Wavelet Transform, DT CWT)及多元 BKF(Multivariate Bessel K Form, MBKF)分布理论为基础,提出了一种新的统计模型数字音频盲水印算法.原始音频进行 DT CWT,并选择重要的 DT CWT 高频系数嵌入水印.水印检测时,首先根据 DT CWT 系数的尺度间相关性,利用二元 BKF 分布对高频子带系数进行建模;然后结合 DT CWT 系数的子带内相关性,估计出二元 BKF 统计模型参数;最后,利用最大似然决策(Maximum Likelihood decision, ML)盲提取水印信息.仿真实验结果表明,本方法在保证良好不可感知性的同时,能够有效抵抗多种攻击,具有较好的鲁棒性.

关键词 音频水印;多元 BKF;双树复数小波变换;最大似然决策;相关性

中图法分类号 TP309

DOI 号 10.11897/SP.J.1016.2019.01112

A Blind Watermark Decoder in DT CWT Domain using Multivariate Bessel K Form Distribution

WANG Xiang-Yang¹⁾ LI Li¹⁾ LI Hai-Fang¹⁾ NIU Pan-Pan^{1),2)} WANG Si-Miao¹⁾ YANG Hong-Ying¹⁾

¹⁾(School of Computer and Information Technology, Liaoning Normal University, Dalian, Liaoning 116029)

²⁾(Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian, Liaoning 116023)

Abstract With the development of Internet technology and multimedia tools, it becomes very easy to distribute and transmit the digital media like image, audio and video. Simultaneously, the protection of intellectual property has also become more and more attentive and important for the society. Digital watermarking has been proposed and widely studied during recent years for the purposes of copyright protection and content authentication. Whilst digital watermarking can be widely applied to various digital media, this paper focuses on digital audio watermarking. Imperceptibility and robustness are two main requirements of any watermarking systems to guarantee desired functionalities, but there is a tradeoff between them from the information-theoretic perspective. Imperceptibility denotes the ability of embedding the watermarks without significantly lowering the audio quality. Robustness denotes the capability of extracting the watermarks under

收稿日期:2017-05-31;在线出版日期:2017-12-29. 本课题得到国家自然科学基金(61472171, 61272416, 61701212)、中国博士后科学基金(2017M621135, 2018T110220)资助. 王向阳, 硕士, 教授, 主要研究领域为网络信息安全、智能多媒体计算. E-mail: wxy37@126.com. 李 丽, 硕士研究生, 主要研究方向为信息隐藏与数字水印. 李海芳, 硕士研究生, 主要研究方向为数字水印、多媒体信息安全. 牛盼盼, 博士, 副教授, 主要研究方向为网络信息安全、图像处理. 王思淼, 硕士研究生, 主要研究方向为信息隐藏与数字水印. 杨红颖(通信作者), 硕士, 教授, 主要研究领域为多媒体信息安全、图像处理. E-mail: yhy_65@126.com.

various attacks. Above two requirements, respectively imperceptibility and robustness, are equally important for an audio watermarking algorithm, but there is an ambivalence between them. Improving the ability of robustness and imperceptibility at the same time has been a challenge for all audio watermarking algorithms. In order to solve the tradeoff between the imperceptibility and robustness of the watermark data, the advantage is usually taken of the statistical properties of the digital audio. Every digital audio has certain features and characteristics. In statistical modeling, it is intended to capture these characteristics using a small number of parameters. In the past decade, some statistical model-based digital audio watermarking schemes have been proposed, and it has been shown that the statistical model based multiplicative watermarking schemes are more robust and provide higher imperceptibility of the watermark than other approaches. Therefore, detection and extraction of the multiplicative watermarks have received a great deal of attention. Statistical audio modeling is mainly focused on transform domains in which the energy density has a more local structure, and the performance of a statistical model-based detection of a watermark is highly influenced by the multiresolution and compression properties of transform and the accuracy of the transform coefficients modeling. There are a number of transform and distributions that have been used in watermark detection; however, there is still scope to explore further the suitability of transform and distributions to improve the performance of watermark detectors. In this paper, a new blind detector is proposed for DT CWT (Dual-Tree Complex Wavelet Transform)-based multiplicative audio watermarking, wherein a PDF based on the multivariate BKF (Bessel K form) distribution is used. In the presented scheme, watermark data is embedded into the significant high-frequency coefficients in DT CWT domain. At the watermark receiver, DT CWT highpass coefficients are firstly modeled by employing multivariate BKF distribution according to the inter-scale coefficients dependencies, then the statistical model parameters of multivariate BKF are estimated using the intra-scale coefficients dependencies, and finally a blind watermark extraction approach is developed using the maximum likelihood decision rule. Experiments are performed to verify the imperceptibility and robustness of the proposed watermark decoder. The results show that the proposed blind watermark decoder is superior to other decoders in terms of providing a higher peak signal to noise ratio and lower bit error rate. It is also shown that the proposed decoder is highly robust against various kinds of attacks such as noise addition, MP3 compression, amplitude variation, random cropping, and jittering.

Keywords audio watermarking; multivariate Bessel K Form distribution; dual-tree complex wavelet transform; maximum likelihood decision; dependencies

1 引言

伴随着计算机技术、多媒体技术、Internet 技术的迅猛发展,数字信息(包括数字图像、数字音频、数字视频等)的传输与利用日益变得频繁与广泛。由于数字信息极易被无限制任意编辑、复制与散布,从而导致数字媒体作品的原创者蒙受巨大经济损失,数

字作品的信息安全、知识产权保护和认证等问题已变得日益突出,并成为数字世界中一个非常迫切的重要议题。传统的加密系统在数据传输过程中可以起到保护作用,但数据一旦被接受并解密,其保护作用也随之消失,同时加密的密文容易引起好事者的兴趣,激发他们的破译热情^[1]。数字水印(Digital Watermarking)作为传统加密方法的有效补充手段,是一种可以在开放的网络环境下保护版权和认

证来源及完整性的新技术,近年来已引起了人们的高度重视,并已成为国际学术界研究的一个热点^[2].

数字水印技术一般需要满足三项基本要求^[2],即不可感知性、鲁棒性、水印容量.不可感知性即是嵌入的水印不影响人对载体的视觉或听觉效果;鲁棒性则是指在对含水印载体进行常规信号攻击或几何攻击后依然能够检测出水印信息的能力;水印容量指的是单位时间内或单位区域内嵌入的水印比特数,理论上水印容量越大越好.不可感知性、鲁棒性、水印容量是衡量一个水印算法优劣的最为重要的指标,而这三方面却又存在着相互矛盾的关系,三者之间的最佳平衡也成为水印算法所共同追求的目标.截止到目前,可保持不可感知性、鲁棒性、水印容量之间良好平衡的数字水印算法研究仍然是一项富有挑战性的工作.

通常来讲,数字水印算法包含两大步骤:水印嵌入和水印提取^[3].在水印嵌入环节,可把水印算法分为加性水印^[4-5]、乘性水印^[6]和基于量化方法^[7-8]的水印算法.在宿主信号中提高水印的嵌入强度可以增强含水印信号的鲁棒性,但同时也会给宿主信号带来严重的失真效果.为了能在保证不可感知性的同时提高鲁棒性,水印嵌入的强度应由载体信号的内容来确定.相比之下,乘性水印算法较为简单,更能增强水印的隐蔽性,即水印的嵌入强度随着载体信号强度成一定比例变化,让人更不容易察觉,最大程度地平衡鲁棒性和不可感知性.近年来,变换域水印技术因其隐蔽性以及遭到攻击时更为稳定的优点,愈来愈受到重视.为充分考虑人类感知系统的掩蔽特性,乘性水印算法的设计通常在变换域内进行,如 Discrete Cosine Transform (DCT)^[9]、Discrete Wavelet Transform (DWT)^[10-11]、Ridgelet Transform (RT)^[12]、Contourlet Transform (CT)^[13]、Nonsubsampled Shearlet Transform (NSST)^[14]等.

在水印提取环节,水印检测可描述为统计决策问题,变换系数统计模型的优劣对检测效果影响很大,因此检测器的构造就显得尤为重要.对于盲水印算法,水印的检测过程并不知道原始信号,依据信号检测理论可将含水印信号看作噪声,若这些系数符合高斯分布,那么线性检测器 (Linear-Correlation (LC) detector) 即可做到最优检测.然而,实验证明,多数载体信号变换域系数均不服从高斯分布,因而基于线性相关的检测器不再是最优检测方法,那么就需要一种依据载体系数分布建立的模型来进行检测,

模型越精确,检测的可信度就越高.近年来,越来越多的统计模型数字水印算法被相继提出,主要可分为两类,第一类只能检测是否含有水印信息^[4,10,13,15-19],第二类可以检测出具体水印信息^[11,14,20-23],显然第二类算法更加符合实际需求.

截至到目前,统计模型水印技术已经广泛应用于数字图像领域,但在数字音频领域的应用却相对较少^[23-24],且仅有的统计模型音频水印方案的检测性能并不理想.本文以双树复数小波变换 (Dual-Tree Complex Wavelet Transform, DT CWT) 及多元 BKF (Multivariate Bessel K form, MBKF) 分布理论为基础,提出了一种新的统计模型数字音频盲水印算法.由于算法充分利用了 DT CWT 域系数之间的多种相关性,故可获得较好的水印检测性能.本文第 2 节主要介绍基于多元 BKF 分布的 DT CWT 域统计建模详细过程;第 3 节和第 4 节叙述数字水印的嵌入和提取步骤;第 5 节给出本算法的实验结果及对比结果;第 6 节对全文进行总结.

2 基于多元 BKF 分布的 DT CWT 域统计建模

2.1 一维 DT CWT 简介

双树复数小波变换 (DT CWT) 结合了离散小波和连续小波的共同优点,不仅计算简单,而且具有良好的重构能力及近似平移不变特性^[25].图 1 给出了一维数字音频信号的 DT CWT 结构.其中,树 a 和树 b 分别表示两个实数 DWT,树 a 生成了实数部分,而树 b 生成了虚数部分. $h_0(n)$ 和 $h_1(n)$ 分别代表上采样滤波器组的低通滤波器和高通滤波器; $g_0(n)$ 和 $g_1(n)$ 代表下采样滤波器组的低通滤波器和高通滤波器.假设两个实数 DWT 树 a 和树 b 分别用 F_h 和 F_g 表示,则 DT CWT 可以用 F 表示,其中 $F =$

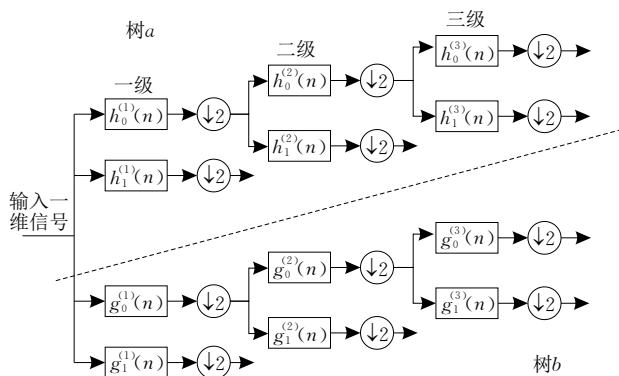


图 1 一维信号 DT CWT 结构图

$[F_h \ F_g]$. 如果 x 是输入的一维信号, 则复数的分解系数可以由 $w_h + jw_g$ 获得, 其中 $w_h = F_h x$, $w_g = F_g x$. F 的逆变换可表示为 $F^{-1} = \frac{1}{2} [F_h^{-1} \ F_g^{-1}]$.

图 2 给出了原始音频及其 DT CWT 分解第一尺度高频子带系数和第二尺度高频子带系数的分布直方图. 从图 2 可以看出, 大多数 DT CWT 高频系数接近于零, 只有少量处于中间位置的子带系数幅值较大, 即体现了其高峰拖尾特点. 二者峰值值分别为 25.9728 和 19.9376, 远大于高斯分布的峰值值 3, 其表明 DT CWT 高频系数具有明显的非高斯分布特性.

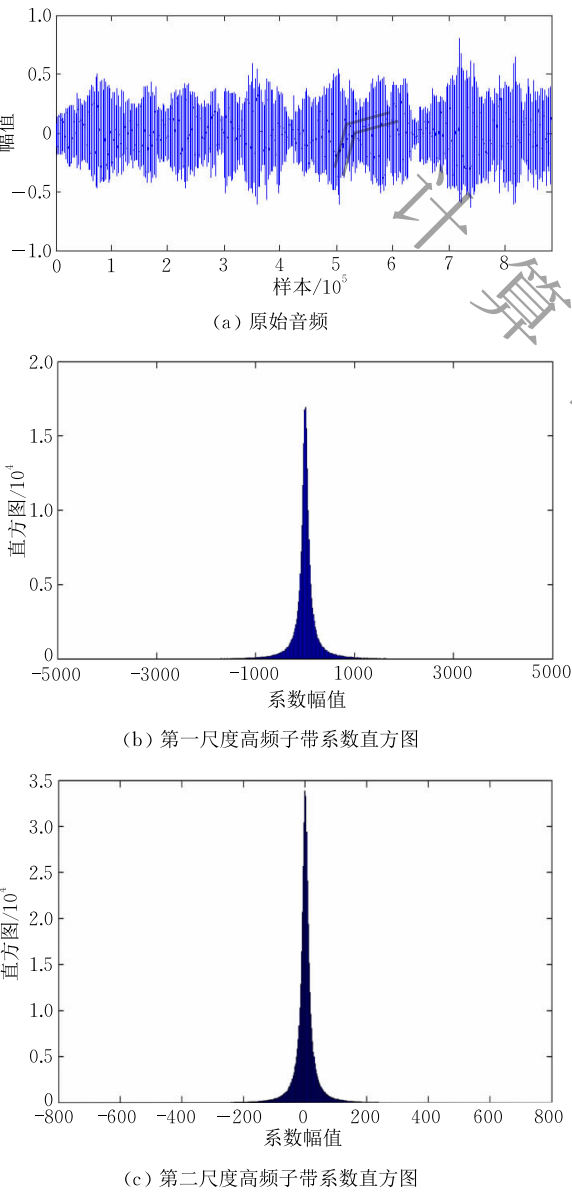


图 2 原始音频及 DT CWT 域高频子带系数分布直方图

近年来, 研究者普遍关注并利用了变换域系数的高尖峰重拖尾非高斯性, 但不同程度忽略了变换

系数之间的强依赖性(多种相关性). 事实上, 同一子带内的系数相关性及不同尺度间的系数相关性是 DT CWT 域高频子带存在的两种重要相关性. 对于每一个 DT CWT 系数 X , 这里定义在 X 附近邻域内的系数为 X 的兄弟系数 N_X , 将处于 X 上一尺度对应的系数定义为父系数 P_X , 并且每个系数有且仅有一个父系数. DT CWT 系数统计特性可以表述出系数的上述关系, 所以这里我们通过联合统计分析来分别描述 N_X 、 P_X 系数与 X 之间的依赖关系, 如图 3 所示.

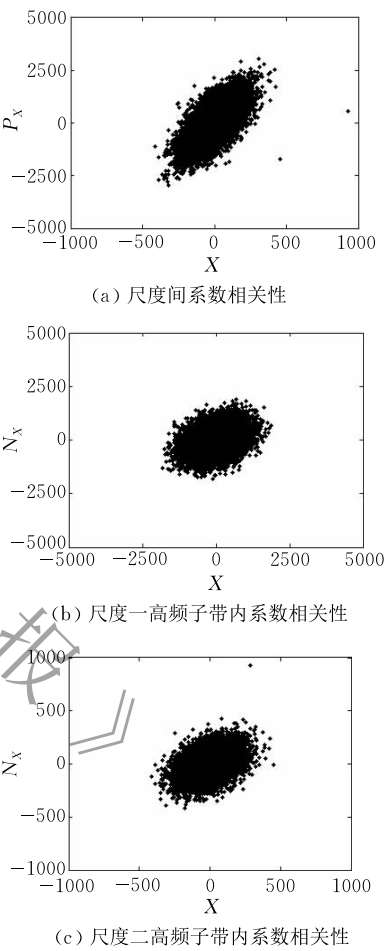


图 3 DT CWT 系数的联合概率分布图

图 3 给出的联合概率分布图有一个共同的特点, 即所有点都有向中心位置靠拢的趋势, 这说明子带内兄弟系数以及尺度间父子系数之间依赖关系很强, 如果某一系数的值是大系数, 那么它的相关系数值为大系数的概率也很大, 即说明 DT CWT 父子系数、兄弟系数间都存在明显的相关性.

2.2 多元 BKF 分布

根据图像传递生成模型, Sendur 提出了一种描述小波系数的概率密度分布—多元 BKF (Multivariate BKF) 模型, 该模型较好地体现了小波系数之间的相

关性. 定义 d 维 BKF 的概率密度函数(PDF)为^[26]

$$f_{\mathbf{X}}(\mathbf{X}; \alpha, \Sigma) =$$

$$\frac{2^{1-\alpha}}{\sqrt{(2\pi)^d |\Sigma|}} (2\alpha)^{\frac{\alpha}{2} + \frac{d}{4}} \|\mathbf{X}\|_{\Sigma^{-1}}^{\alpha - \frac{d}{2}} K_{\alpha - \frac{d}{2}}(\sqrt{2\alpha} \|\mathbf{X}\|_{\Sigma^{-1}}) \quad (1)$$

这里, $\mathbf{X} = (X_1, X_2, \dots, X_d)^T$ 表示一组随机向量, 且有 $\mathbf{X} \in R^d$, $K_{\alpha - \frac{d}{2}}$ 表示 $\alpha - \frac{d}{2}$ 阶的第一类修正的 Bessel 函数, 且有

$$\|\mathbf{X}\|_{\Sigma^{-1}} = (\mathbf{X}^T \Sigma^{-1} \mathbf{X})^{\frac{1}{2}} \quad (2)$$

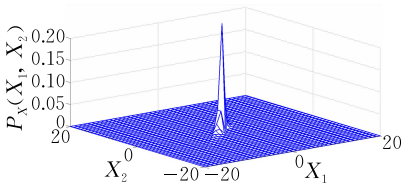
在 d 维的 BKF 分布中有两个未知参数, α 表示形状参数, Σ 表示对称的正定矩阵. 如果 $\mathbf{X} \sim$ 多元 BKF(α, Σ) 并且符合条件 $\alpha > 0$, 则多元 BKF 分布的 PDF 具有以下性质:

- (1) 当 α 趋近于无穷大时, $\mathbf{X}$ 服从于高斯分布.
- (2) $(\mathbf{X}^T \Sigma^{-1} \mathbf{X})$ 的期望可以表示为

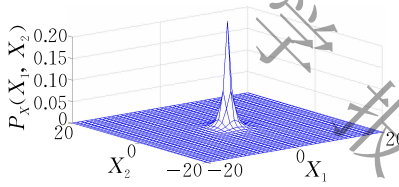
$$E[(\mathbf{X}^T \Sigma^{-1} \mathbf{X})^i] = 2^i \frac{\Gamma(\frac{d}{2} + i) \Gamma(\alpha + i)}{\alpha \Gamma(\frac{d}{2}) \Gamma(\alpha)} \quad (3)$$

- (3) 多元 BKF 的累积量可以表示为

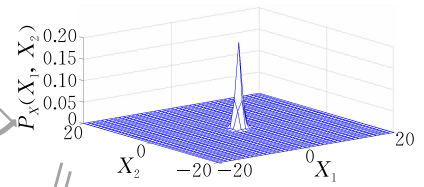
$$k_j(2i) = \alpha^{1-i} \frac{(2i)!}{2^i i!} \sum_{l=1}^d \sigma^{2i}(j, l),$$



(a) DT CWT 尺度间父子系数联合分布直方图



(b) 二元 BKF 的父子联合概率分布



(c) 二元 Cauchy 的父子联合概率分布

图 4 父子子带联合分布及二元分布拟合

为进一步比较两种二元统计模型的拟合精度, 这里选择 KL 距离^[9]来评判二元分布模型, KL 距离计算公式如下所示:

$$KL = \sum_{i=1}^n \ln(P_i/Q_i) P_i \quad (6)$$

其中, P_i 表示联合分布直方图的 bin 值, Q_i 表示在每个 bin 中由二元分布 PDF 求出的概率值. 在计算中, 要先将 P_i 和 Q_i 进行归一化处理.

实验中, 我们分别用二元 BKF 和二元 Cauchy 对 10 段数字音频的二级 DT CWT 尺度间子带建模, 计算出每种模型的平均 KL 距离, 可得出二元 Cauchy 的平均 KL 距离为 1.46, 二元 BKF 的平均 KL 距离为 0.97. 可以看出, 二元 BKF 分布可以更好地对音频 DT CWT 尺度间相关性进行建模, 因此本文利用二元 BKF 统计模型构造水印检测器.

$$i \geq 1, j = 1, \dots, d, \quad (4)$$

其中, $\sigma(j, l)$ 表示矩阵 $\Sigma^{\frac{1}{2}}$ 中位于 (j, l) 位置的值. 对于多元 BKF 统计模型中的两个参数 Σ 和 α , 本文将综合考虑误差、时间等因素, 选取矩方法与 ML 方法进行参数估计^[26].

2.3 基于多元 BKF 分布的 DT CWT 域建模

本文充分考虑 DT CWT 的尺度间相关性, 利用二元 BKF 分布对 DT CWT 尺度间父子系数进行建模, 二元 BKF 的父子联合概率分布如图 4(b) 所示. 为了能够说明二元 BKF 模型的精确性, 这里又利用多元 Cauchy 分布对 DT CWT 尺度间父子系数进行建模, 如图 4(c) 所示. 其中, 多元 Cauchy 分布的 PDF 形式如下:

$$f_{\mathbf{X}}(\mathbf{x}; n, \Sigma) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{1}{2} \pi^{\frac{1}{2}} |\Sigma|^{\frac{1}{2}} [1 + \mathbf{x}^T \Sigma^{-1} \mathbf{x}]^{\frac{n+1}{2}})} \quad (5)$$

这里, $\Gamma(\cdot)$ 表示伽马函数, Σ 表示 $n \times n$ 的协方差矩阵.

图 4(a) 给出了 DT CWT 尺度间父子系数联合分布直方图. 不难看出, 二元 BKF 和二元 Cauchy 都可以很好地刻画出父子系数的联合分布概率.

3 数字水印嵌入

本文选取 DT CWT 域树 A 与树 B 的第二尺度高频子带作为水印嵌入区, 并采用非线性乘性嵌入函数对能量较大的 DT CWT 高频系数进行修改以嵌入水印信息, 以期同时提高数字水印的不可感知性和鲁棒性. 图 5 给出了数字水印嵌入流程图.

假设原始数字音频信号为 $A = \{a(i), 0 \leq i < L\}$. 其中, L 为音频数据的个数, $a(i) \in \{0, 1, 2, \dots, (2^p - 1)\}$ 是第 i 个音频数据的幅度值, p 为表示每个数据所使用的比特数. 假设二值水印图像为 $W = \{w(i, j), 0 \leq i < I_1, 0 \leq j < J_1\}$. 其中, $w(i, j) \in \{0, 1\}$ 代表二值水印图像的第 i 行、第 j 列像素值. 则数字水印嵌入过程可描述如下.

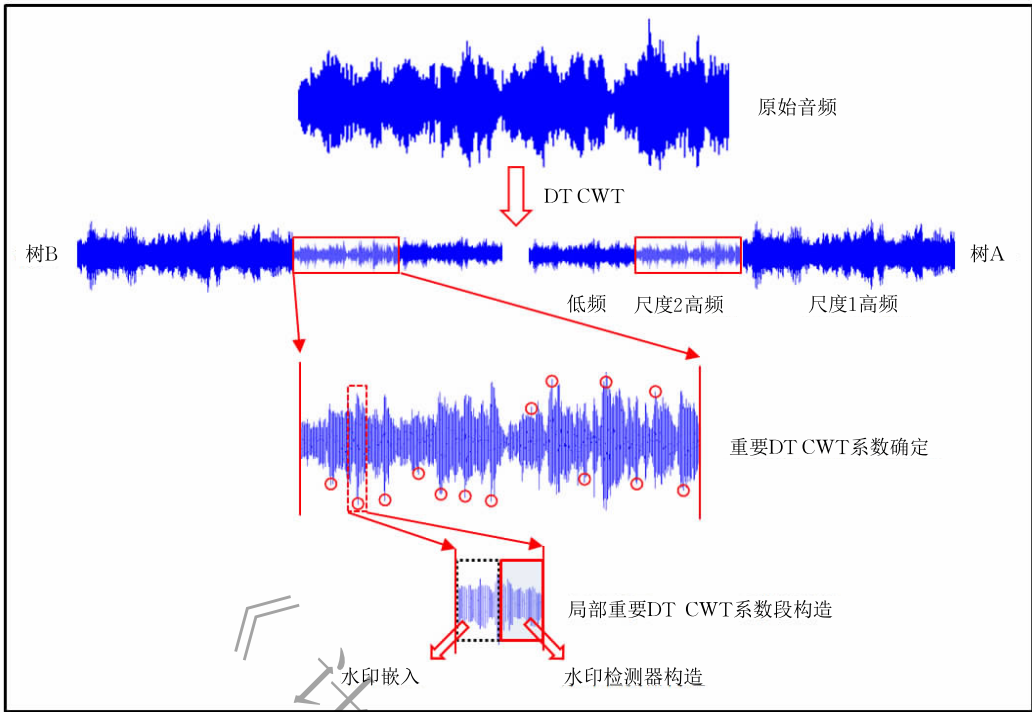


图 5 数字水印嵌入流程图

3.1 水印预处理

为消除二值水印图像的像素空间相关性,提高数字水印算法的安全性和鲁棒性,以确保含水音频某一部分受到破坏后仍能全部或部分地恢复出二值水印图像,本文首先采用了 Arnold 变换对二值水印图像进行置乱加密.接下来,再对置乱后的二值水印图像进行降维处理,即将置乱后的二值水印图像转换为一维二进制水印序列: $W = \{w(i)\} (i = 1, 2, \dots, N), N = I_1 \times J_1$, 以便于将二维的水印图像嵌入到一维的数字音频信号中.

3.2 原始音频 DT CWT

对原始载体音频做二级 DT CWT 分解,以得到树 A 和树 B.其中,树 A 和树 B 均包含有 1 个低频子带和 2 个高频子带.为了取得不可感知性和鲁棒性的良好平衡,本文选取树 A 和树 B 的尺度 2 高频子带用于水印嵌入与水印检测.

3.3 重要 DT CWT 系数确定

首先计算出尺度 2 高频子带中每个 DT CWT 系数的局部短时能量

$$\begin{aligned} E_n &= \sum_{m=0}^n [x(m)w(n-m)]^2 \\ &= \sum_{m=0}^n x^2(m)h(n-m) \\ &= x^2(n) \times h(n-m) \end{aligned} \quad (7)$$

这里,每个音频样本开长度为 n 的局部窗口, $h(n) = w^2(n)$. 本文选取的窗函数是汉宁窗 $w(n)$.

$$\begin{aligned} w(n) &= 0.5 \left(1 - \cos \left(2\pi \frac{n}{N+1} \right) \right), \\ n &= 1, 2, \dots, N \end{aligned} \quad (8)$$

进而选取局部能量较大的前 N 个 DT CWT 系数作为重要系数 $C = \{c_k, k = 1, \dots, N\}$,用以标识水印嵌入位置.需要指出的是,进行重要 DT CWT 系数确定时,必须滤除两类不适合水印嵌入的 DT CWT 系数,从而保证所确定的重要 DT CWT 系数分布均匀且各系数之间有足够空间用于水印嵌入.一类是位于音频首尾端附近的系数,即不满足

$$((p(k) - M/2) \geq 1) \cap ((p(k) + M/2 - 1) \leq L) \quad (9)$$

的 DT CWT 系数.这里, $p(k)$ 表示重要 DT CWT 系数 c_k 的原始位置, M 表示用于嵌入 1 位水印信息的局部重要 DT CWT 系数段长度, L 表示原始载体音频长度.

另一类是毗邻高局部能量 DT CWT 系数 c_{n_i} 的系数,即不满足

$$|p(n_j) - p(n_i)| \geq M \quad (10)$$

的 DT CWT 系数.其中,

$$p(n_j) \in \{p(n_i - 1), p(n_i - 2), \dots, p(n_i)\} \quad (11)$$

3.4 重要 DT CWT 系数段构造

所谓局部重要 DT CWT 系数段,是指以重要 DT CWT 系数为标识,从尺度 2 高频子带中选取的部分系数,把它们作为水印信号的嵌入和检测区域.为了减少计算量,本文以所提取的重要 DT CWT 系

数为标识,选取每个系数的前 $M/2$ 个 DT CWT 高频系数(以下简记“系数段 $B_k, k=1, \dots, N$ ”)用于水印嵌入,后 $M/2$ 个 DT CWT 高频系数(以下简记“系数段 $A_k, k=1, \dots, N$ ”)用于水印检测器构造。

3.5 数字水印嵌入

理论分析与实验结果均表明^[2],水印嵌入强度对算法工作性能有一定影响,即对于给定的载体信号而言,如嵌入强度较大,则会引起视觉失真,从而降低不可感知性;如嵌入强度较小,则会导致算法的抗攻击能力降低。也就是说,为了取得不可感知性和鲁棒性的良好平衡,水印嵌入强度应自适应于载体信号内容,即水印嵌入强度应该随着载体信号强度(值)的增大而增大。为取得不可感知性和鲁棒性的良好平衡,现有算法普遍采用线性单调函数刻画水印嵌入强度,并用于自适应调节载体信号(即嵌入水印信息)^[2,23-24]。然而,水印嵌入强度增量与载体信号改变量间并非具有简单的线性单调关系,而是呈现出明显的非线性单调关系,为此本文采用刻画能力更强的非线性单调函数,构造了自适应水印嵌入强度(函数),较好地获得了不可感知性和鲁棒性的良好平衡。

设 B_k 表示原始 DT CWT 高频系数段 $B_k = \{x_{k1}, x_{k2}, \dots, x_{k\frac{M}{2}}\}$, B'_k 表示含水印 DT CWT 高频系数段 $B'_k = \{x'_{k1}, x'_{k2}, \dots, x'_{k\frac{M}{2}}\}$, 则数字水印嵌入策略可表示为

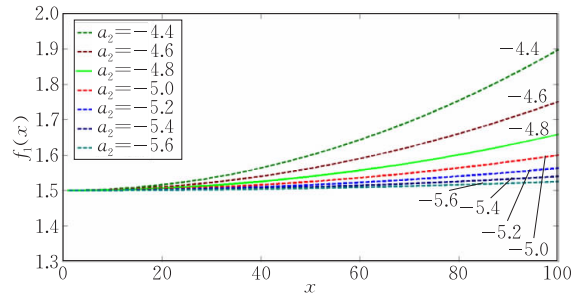
$$x'_{ki} = \begin{cases} f_1(x_{ki}) \cdot x_{ki}, & \text{如果待嵌入水印位为 1} \\ f_0(x_{ki}) \cdot x_{ki}, & \text{如果待嵌入水印位为 0} \end{cases} \quad (12)$$

其中, $i=1, 2, \dots, \frac{M}{2}, k=1, 2, \dots, N$, 本实验中水印大小 N 为 1024。水印为“1”时的嵌入函数和水印为“0”时的嵌入函数分别为

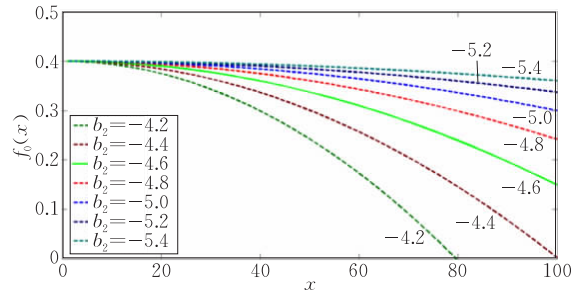
$$f_1(x) = a_1 + 10^{a_2} x^2, \quad f_0(x) = b_1 + 10^{b_2} x^2 \quad (13)$$

本文采用刻画能力更强的非线性单调函数——二次函数,构造了自适应水印嵌入强度函数,可以较好地获得不可感知性和鲁棒性的良好平衡。该自适应水印嵌入强度函数共涉及四个参数(分别是 a_1 、 a_2 、 b_1 、 b_2),而各参数值需要结合仿真实验选择确定。图 6 给出了用以刻画二次函数曲度的参数 a_2 与 b_2 的选择示意图,这里,确定参数 a_2 与 b_2 最优值的思想是择中原则,即选取参数中间值(注:指 7 条曲线整个变化曲度的中间位置值,实线位置处)。其理由为嵌入强度函数曲线图表明,函数自变量 X 值与 ΔX 的改变程度呈正比: X 值较小, ΔX 改变量较小; X 值变大时, ΔX 改变量变化程度也随着变大。基于

上述参数选取原则,本文所确定参数为 $a_1 = 1.5$, $b_1 = 0.4$, $a_2 = -4.8$, $b_2 = -4.6$ 。



(a) 水印为“1”时的嵌入函数



(b) 水印为“0”时的嵌入函数

图 6 两种数字水印嵌入强度函数

重复步骤 3.5,直到树 A 和树 B 的尺度 2 高频子带中所有局部重要 DT CWT 系数段处理完毕为止。最后,用含水印局部重要 DT CWT 系数段替换掉原始的局部重要 DT CWT 系数段,并进行逆 DT CWT,即可得到含水印数字音频信号。

4 数字水印检测

4.1 水印检测器构造

假设水印检测器端接收到的含水印系数 y_i 受到零均值的高斯白噪声 $n \sim N(0, \sigma_n^2)$, 即

$$z_i = \underbrace{f \cdot x_i}_{y_i} + n \quad (14)$$

其中, f 表示两种水印嵌入强度函数。由于 DT CWT 系数与噪声项是相互独立的,因此可获得含噪声 DT CWT 系数的 PDF

$$f_z(z) = \int_{-\infty}^{\infty} f_y(z - \tau) f_n(\tau) d\tau \quad (15)$$

为简化上式,本文利用 three-sigma 准则估计高斯 PDF

$$\Lambda_G(x) = \begin{cases} \frac{-x + 3\sigma_n}{9\sigma_n^2}, & 0 < x < 3\sigma_n \\ \frac{x + 3\sigma_n}{9\sigma_n^2}, & -3\sigma_n \leq x < 0 \\ 0, & |x| > 3\sigma_n \end{cases} \quad (16)$$

把式(16)代入式(15)中,可以得到

$$f(z) \approx \frac{\sigma_n}{4} \cdot [2 \cdot KK(z) \Lambda_G(0) + 4 \cdot KK\left(z - \frac{3\sigma_n}{2}\right) \Lambda_G\left(\frac{3\sigma_n}{2}\right) + 4 \cdot KK\left(z + \frac{3\sigma_n}{2}\right) \Lambda_G\left(-\frac{3\sigma_n}{2}\right) + KK(z + 3\sigma_n) \Lambda_G(-3\sigma_n) + KK(z - 3\sigma_n) \Lambda_G(3\sigma_n)] = \frac{1}{6} \left(KK\left(z - \frac{3\sigma_n}{2}\right) + KK(z) + KK\left(z + \frac{3\sigma_n}{2}\right) \right) \quad (17)$$

因此两种假设 $H_1(z|1)$ 和 $H_0(z|0)$ 分别可以表示为

$$H_1(z_1, z_2, \dots, z_m | 1) = \prod_{i=1}^m f(z_i | 1) = \prod_{i=1}^m \left[KK_1\left(z_i - \frac{3\sigma_n}{2}\right) + KK_1(z_i) + KK_1\left(z_i + \frac{3\sigma_n}{2}\right) \right] \quad (18)$$

$$H_0(z_1, z_2, \dots, z_m | 0) = \prod_{i=1}^m f(z_i | 0) = \prod_{i=1}^m \left[KK_0\left(z_i - \frac{3\sigma_n}{2}\right) + KK_0(z_i) + KK_0\left(z_i + \frac{3\sigma_n}{2}\right) \right] \quad (19)$$

本文通过中值方法估计出 DT CWT 子带的噪声标准偏差 $\hat{\sigma}_n$, 如下所示:

$$\hat{\sigma}_n = \frac{\text{Median}|y_{j,d}|}{0.6745} \quad (20)$$

其中, $KK_1(y)$ 和 $KK_0(y)$ 分别是在使用强度函数的情况下被计算出来, 如式(21)及式(22)

$$KK_1(\mathbf{Q}) = \frac{2^{1-\alpha}}{\sqrt{(2\pi)^2 |\mathbf{\Sigma}|}} (2\alpha)^{\frac{\alpha+1}{2}} (\mathbf{Q}_i^T \mathbf{\Sigma}^{-1} \mathbf{Q}_i)^{\frac{\alpha-1}{2}} \times K_{\alpha-1}(\sqrt{2\alpha}(\mathbf{Q}_i^T \mathbf{\Sigma}^{-1} \mathbf{Q}_i)^{\frac{1}{2}}) \quad (21)$$

$$KK_0(\mathbf{R}) = \frac{2^{1-\alpha}}{\sqrt{(2\pi)^2 |\mathbf{\Sigma}|}} (2\alpha)^{\frac{\alpha+1}{2}} (\mathbf{R}_i^T \mathbf{\Sigma}^{-1} \mathbf{R}_i)^{\frac{\alpha-1}{2}} \times K_{\alpha-1}(\sqrt{2\alpha}(\mathbf{R}_i^T \mathbf{\Sigma}^{-1} \mathbf{R}_i)^{\frac{1}{2}}) \quad (22)$$

这里,

$$\mathbf{Q} = [g_1(y1_i), g_1(y2_i)], \quad \mathbf{R} = [g_0(y1_i), g_0(y2_i)] \quad (23)$$

其中, $y1_i$ 表示 DT CWT 第一尺度高频子带不含噪声的系数, $y2_i$ 表示 DT CWT 第二尺度高频子带不含噪声的系数. 另外, 此处需要将第二尺度高频子带拓展成二倍长度, 并与第一尺度相关系数进行对应. g_1 和 g_0 分别代表嵌入函数 f_1 和 f_0 的反函数, 根据卡丹公式解三次方根的方法可推导出 g_1 和 g_0

$$g_1(y) = \omega \left[\frac{y}{2 \cdot 10^{a_2}} + \sqrt{\frac{y^2}{4 \cdot 10^{2a_2}} + \frac{a_1^3}{27 \cdot 10^{3a_2}}} \right]^{1/3} + \omega^2 \left[\frac{y}{2 \cdot 10^{a_2}} - \sqrt{\frac{y^2}{4 \cdot 10^{2a_2}} + \frac{a_1^3}{27 \cdot 10^{3a_2}}} \right]^{1/3} \quad (24)$$

$$g_0(y) = \omega \left[-\frac{y}{2 \cdot 10^{b_2}} + \sqrt{\frac{y^2}{4 \cdot 10^{2b_2}} - \frac{a_1^3}{27 \cdot 10^{3b_2}}} \right]^{1/3} + \omega^2 \left[-\frac{y}{2 \cdot 10^{b_2}} - \sqrt{\frac{y^2}{4 \cdot 10^{2b_2}} - \frac{a_1^3}{27 \cdot 10^{3b_2}}} \right]^{1/3} \quad (25)$$

对于二元 BKF 统计模型中的两个参数 $\mathbf{\Sigma}$ 和 α , 本文将采用矩估计方法对第一尺度高频子带和第二尺度高频子带 $\mathbf{\Sigma}$ 进行估计. 考虑到子带内系数相关性, 这里采用第二尺度高频子带中未嵌入水印的 DT CWT 系数通过 ML 方法估计出形状参数 α , 并进一步可通过极大似然决策规则获得最优水印检测器, 如下所示:

$$\sum_{i \in B_k} \ln \frac{KK_1\left(y_i - \frac{3\sigma_n}{2}\right) + 2KK_1(y_i) + KK_1\left(y_i + \frac{3\sigma_n}{2}\right)}{KK_0\left(y_i - \frac{3\sigma_n}{2}\right) + 2KK_0(y_i) + KK_0\left(y_i + \frac{3\sigma_n}{2}\right)} > T_k \quad (26)$$

上式基本含义为: 对于系数段 B_k 而言, 如果其嵌入“1”的概率 > 阈值 T_k , 则从系数段 B_k 提取出嵌入“0”的概率

数字水印位“1”; 如果嵌入“1”的概率 < 阈值 T_k , 嵌入“0”的概率

则从系数段 B_k 提取出数字水印位“0”. 其中, $T_k = \sum_{i \in B_k} \ln(g_1(y_i)/g_0(y_i))$, $KK_1(y)$ 和 $KK_0(y)$ 分别由式(21)和式(22)求得.

4.2 数字水印提取

本文所提出的数字音频水印算法属于盲水印算法, 其水印检测过程无需原始载体、原始水印等辅助信息. 整个数字水印提取过程如下:

(1) 对含水印数字音频做二级 DT CWT 分解, 并选取树 A 和树 B 的尺度 2 高频子带用于水印提取.

(2) 采用类似于 3.3 节和 3.4 节的工作步骤, 确定重要 DT CWT 系数, 并构造局部重要 DT CWT 系数段.

(3) 利用局部重要 DT CWT 系数段的“系数段 A_k ”, 并采纳 4.1 节方法构造数字水印检测器, 然

后从局部重要 DT CWT 系数段的“系数段 B_k ”中提取出水印信息位。

(4) 重复步骤(3), 即对每个重要 DT CWT 系数段进行水印提取, 可以得到一维水印序列

$$W_2^* = \{\omega_2^*(i)\}, i=0,1,\dots,N \quad (27)$$

(5) 对 W_2^* 做升维与逆 Arnold 变换, 即可获得最终的二维水印图像 W^*

$$W^* = \{\omega^*(i,j), 0 \leq i < I_1, 0 \leq j < J_1\} \quad (28)$$

4.3 误码概率分析

下面, 我们采用误码概率(Bit Error Probability, BEP)分析所提出水印检测器的工作性能^[11]. 在无攻击检测条件下, BEP 的计算公式可表示为

$$p_e = \frac{1}{N} \sum_{k=1}^N [P(Z_k(y) > T_k | H_0) + P(Z_k(y) < T_k | H_1)] \quad (29)$$

其中, $N = I_1 \cdot J_1$ 表示数字水印位数, $T_k = \sum_{i \in B_k} \ln(g_1(y_i) / g_0(y_i))$.

$$\begin{aligned} \mu_0 = E[Z_k(y | H_0)] &= \sum_{i \in B_k} \left[\frac{1}{2} \ln \frac{2^{1-\alpha} (2\alpha)^{\frac{\alpha+1}{2}} (\mathbf{Q}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{Q}_i)^{\frac{\alpha-1}{2}} K_{\alpha-1}(\sqrt{2\alpha}(\mathbf{Q}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{Q}_i)^{\frac{1}{2}})}{\sqrt{|\boldsymbol{\Sigma}|}} + \right. \\ &\quad \left. \frac{1}{2} \ln \frac{2^{1-\alpha} (2\alpha)^{\frac{\alpha+1}{2}} \left(\left(\frac{\mathbf{R}_i^2}{\mathbf{Q}_i} \right)^T \boldsymbol{\Sigma}^{-1} \left(\frac{\mathbf{R}_i^2}{\mathbf{Q}_i} \right) \right)^{\frac{\alpha-1}{2}} K_{\alpha-1}(\sqrt{2\alpha} \left(\left(\frac{\mathbf{R}_i^2}{\mathbf{Q}_i} \right)^T \boldsymbol{\Sigma}^{-1} \left(\frac{\mathbf{R}_i^2}{\mathbf{Q}_i} \right) \right)^{\frac{1}{2}})}{\sqrt{|\boldsymbol{\Sigma}|}} - \right. \\ &\quad \left. \ln \frac{2^{1-\alpha} (2\alpha)^{\frac{\alpha+1}{2}} (\mathbf{R}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{R}_i)^{\frac{\alpha-1}{2}} K_{\alpha-1}(\sqrt{2\alpha}(\mathbf{R}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{R}_i)^{\frac{1}{2}})}{\sqrt{|\boldsymbol{\Sigma}|}} \right] \end{aligned} \quad (31)$$

上式可简化为

$$\mu_0 = \sum_{i \in B_k} \ln \frac{\sqrt{a_i b_i}}{c_i} \quad (32)$$

其中,

$$\begin{aligned} a_i &= \frac{2^{1-\alpha} (2\alpha)^{\frac{\alpha+1}{2}} (\mathbf{Q}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{Q}_i)^{\frac{\alpha-1}{2}} K_{\alpha-1}(\sqrt{2\alpha}(\mathbf{Q}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{Q}_i)^{\frac{1}{2}})}{\sqrt{|\boldsymbol{\Sigma}|}}, \\ b_i &= \frac{2^{1-\alpha} (2\alpha)^{\frac{\alpha+1}{2}} \left(\left(\frac{\mathbf{R}_i^2}{\mathbf{Q}_i} \right)^T \boldsymbol{\Sigma}^{-1} \left(\frac{\mathbf{R}_i^2}{\mathbf{Q}_i} \right) \right)^{\frac{\alpha-1}{2}} K_{\alpha-1}(\sqrt{2\alpha} \left(\left(\frac{\mathbf{R}_i^2}{\mathbf{Q}_i} \right)^T \boldsymbol{\Sigma}^{-1} \left(\frac{\mathbf{R}_i^2}{\mathbf{Q}_i} \right) \right)^{\frac{1}{2}})}{\sqrt{|\boldsymbol{\Sigma}|}}, \\ c_i &= \frac{2^{1-\alpha} (2\alpha)^{\frac{\alpha+1}{2}} (\mathbf{R}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{R}_i)^{\frac{\alpha-1}{2}} K_{\alpha-1}(\sqrt{2\alpha}(\mathbf{R}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{R}_i)^{\frac{1}{2}})}{\sqrt{|\boldsymbol{\Sigma}|}} \end{aligned} \quad (33)$$

在 H_0 假设下, 其期望 σ_0^2 可表示为

$$\begin{aligned} \sigma_0^2 &= E[Z_k(y | H_0) - \mu_0^2] \\ &= \frac{1}{4} \sum_{i \in B_k} \left(\ln \frac{a_i}{b_i} \right)^2 \end{aligned} \quad (34)$$

由于 $Z_k(y | H_0) = -Z_k(y | H_1)$, 可以得到 $\mu_1 = -\mu_0$ 和 $\sigma_1^2 = \sigma_0^2$, 故可求得检测一位水印信息位的误码概

在无噪声检测条件下, 含水印系数都嵌入“0”的 H_0 假设下的检测概率 $Z_k(y)$ 可表示为

$$\begin{aligned} Z_k(y | H_0) &= \\ &= \frac{2^{1-\alpha} (2\alpha)^{\frac{\alpha+1}{2}} (\mathbf{Q}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{Q}_i)^{\frac{\alpha-1}{2}} K_{\alpha-1}(\sqrt{2\alpha}(\mathbf{Q}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{Q}_i)^{\frac{1}{2}})}{\sqrt{|\boldsymbol{\Sigma}|}} \\ &= \sum_{i \in B_k} \ln \frac{\sqrt{|\boldsymbol{\Sigma}|}}{2^{1-\alpha} (2\alpha)^{\frac{\alpha+1}{2}} (\mathbf{R}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{R}_i)^{\frac{\alpha-1}{2}} K_{\alpha-1}(\sqrt{2\alpha}(\mathbf{R}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{R}_i)^{\frac{1}{2}})} \sqrt{|\boldsymbol{\Sigma}|} \end{aligned} \quad (30)$$

其中,

$$\mathbf{Q} = [g_1(y1_i), g_1(y2_i)], \mathbf{R} = [g_0(y1_i), g_0(y2_i)].$$

由于数字水印是一个独立的随机序列, 且“0”和“1”取值的概率相等, $Z_k(y | H_0)$ 是大量的独立随机变量之和, 因此, $Z_k(y | H_0)$ 近似服从正态分布, 且在两种假设下, 具有有限的数学期望和方差: (μ_0, σ_0) 和 (μ_1, σ_1) . 具体地, 在 H_0 假设下, 其期望 μ_0 可以表示为

率 P_e^k , 具体表示为

$$\begin{aligned} P_e^k &= \frac{1}{2} \{P(Z_k(y) > T_k | H_0) + P(Z_k(y) < T_k | H_1)\} \\ &= \frac{1}{2} \left[1 + Q\left(\frac{T_k - \mu_0}{\sigma_0}\right) + Q\left(\frac{T_k - \mu_1}{\sigma_1}\right) \right] \\ &= \frac{1}{2} \left[1 + Q\left(\frac{T_k - \mu_0}{\sigma_0}\right) + Q\left(\frac{T_k + \mu_0}{\sigma_0}\right) \right] \end{aligned} \quad (35)$$

其中： $Q(x)=\frac{1}{\sqrt{2\pi}}\int_x^\infty \exp\left(\frac{-t^2}{2}\right)dt$. 鉴于嵌入到原始音频信号的二进制水印序列“0”或“1”具有相同概率，于是水印检测器的误码概率 BEP 为

$$P_e=\frac{1}{N}\sum_{k=1}^N P_e^k$$

(36)

5 实验结果与分析

为验证本文音频水印算法的有效性，以下给出了算法的工作性能检测、不可感知性检测和鲁棒性检测结果，同时与文献[8, 23-24]进行了对比. 实验中，我们选取了 6 段不同风格的数字音频信号进行性能测试，包括 Pop、Jazz、Rock、Blues、Speech (in Chinese)、Classic 等. 其中，每段音频信号都是波形格式 (Wave format) 的单声道音频，采样率为 44.1 kHz，量化精度为 16 bits，样本长度为 20 s. 数字水印采用大小为 32×32 的二值图像. 这里，所有测试数据均是在操作系统为 64 位 Win 7、CPU 为 4 核 Intel(R) Core (TM)i5-4590、主频为 3.30 GHz、内存为 16.0 GB 的 PC 机 HP Pro 680 G1 TW 上获得，且所使用的运算平台为 MATLAB7.12.0.

5.1 实验参数选取

下面给出各种实验参数与算法工作性能的关系，包括在不同局部重要 DT CWT 系数段长度 (M) 和不同水印容量情况下，整个水印系统的 PSNR、BER、水印嵌入时间 (s) 以及水印提取时间 (s) 等工作性能，如表 1 和表 2 所示.

表 1 不同系数段长度下整个水印系统的工作性能(水印容量为 1024 位)

局部重要系数段长度 M	PSNR	BER	水印嵌入时间/s	水印提取时间/s
100	41.88	0.17	2.21	1.52
120	40.72	0.10	2.61	1.36
140	40.37	0.00	3.41	1.92
160	40.99	0.09	3.85	1.98

表 2 不同水印容量下整个水印系统的工作性能(局部重要系数段长度为 140)

水印容量	PSNR	BER	水印嵌入时间/s	水印提取时间/s
8×8	41.88	0.17	2.21	1.52
16×16	40.72	0.10	2.61	1.36
32×32	40.37	0.00	3.41	1.92
64×64	40.99	0.09	3.85	1.98

表 1 数据表明，当局部重要 DT CWT 系数段长度 M 设定为 140 时，提取水印效果最好. 观察表 2 中数据可以得出，随着水印容量增大，PSNR 不断下

降，原则上在保证理想不可感知性前提下水印容量越大越好，而在数字水印容量为 32×32 时，水印提取效果达到了比较好的平衡，因此本文选取了 32×32 的二值图像水印.

5.2 不可感知性测试

为了更加直观地显示数字水印的感知透明性，图 7 给出了原始载体音频、含水印数字音频以及音频差值波形图. 同时，我们还采用 PSNR (Peak Signal to Noise Ratio) 和 PEAQ (Perceptual Evaluation of Audio Quality) 客观评价了原始数字音频与含水印数字音频信号之间的差别. 其中，PEAQ 评价模型首先将参考信号和失真信号通过感知声学模型模拟

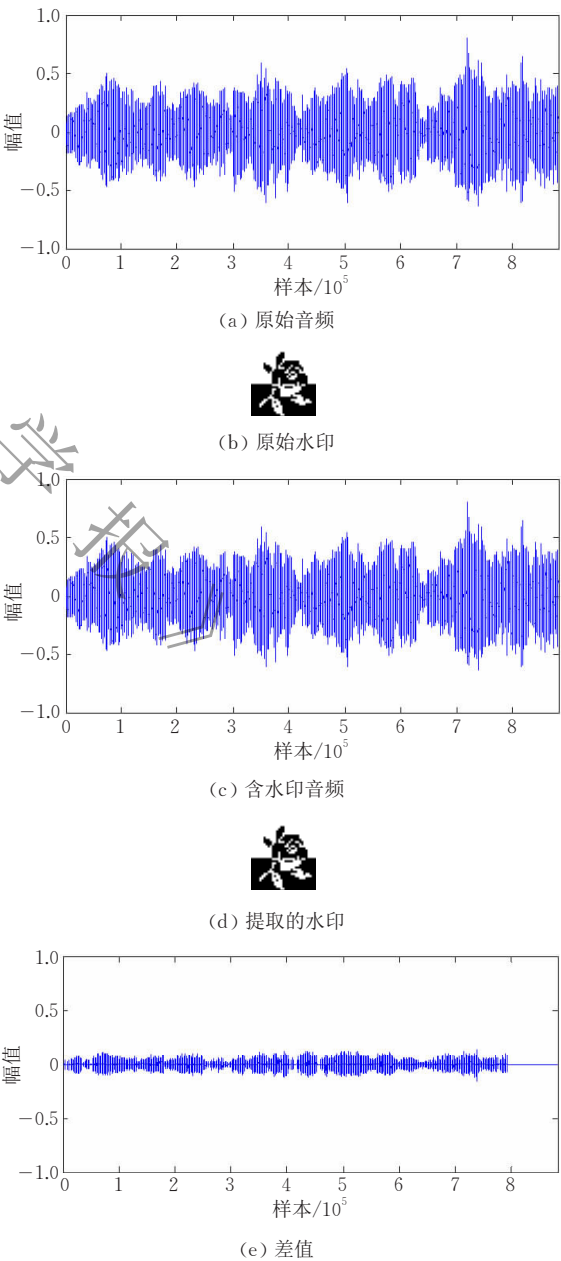


图 7 音频波形图、原始水印及提取出的水印

人对音频信号的感知,然后认知模型将感知声学模型输出值在时域和频域进行综合产生一系列模型输出变量,最后通过一个神经网络,计算出最终的客观失真等级(ODG),其含义如表 3 所示.表 4 给出了不同风格数字音频信号的感知透明性客观评价结果.

表 3 客观听觉测试区分度 ODG

等级	ODG	描述
5.0	0.0	不可感觉
4.0	-1.0	可感觉但不刺耳
3.0	-2.0	轻微刺耳
2.0	-3.0	刺耳
1.0	-4.0	非常刺耳

表 4 感知透明性与鲁棒性客观评价

音频类型	ODG	BER/%
Pop	-0.57	0.00
Jazz	-0.43	0.00
Rock	-0.72	0.19
Blues	-0.36	0.00
Speech	-0.20	0.00
Classic	-0.88	0.10

图 7 表明,在保证完全正确提取水印的情况下(原始水印与提取的水印完全一致),嵌入水印前后波形图改变较小,基本保持一致.而从图 7(e)差值波形中,更可直观地看出宿主音频在嵌入水印后,与原始音频相比变化很小,说明本算法不可感知性较好.

5.3 鲁棒性测试

为验证本文算法的鲁棒性,我们对含水印音频进行了一系列攻击实验,包括重新量化(Re-quantization)、

重新采样(Re-sampling)、叠加噪声(Noise addition)、低通滤波(Lowpass filtering)、MP3 压缩等常规信号处理,及随机剪切(Random Cropping)、幅值改变(Amplitude Variation)、抖动(Jittering)等去同步攻击.其中,

重新量化:先将音频从 16 比特量化到 8 比特,再重新量化到 16 比特.

重采样:先将采样频率为 44.1 kHz 的音频频率下降到 22.05 kHz,11.025 kHz 等,再还原为采样频率 44.1 kHz.

添加噪声:在数字音频信号中加入均值为 0、均方差为 0.01 的高斯白噪声(200).

添加回声:将原始音频的延时拷贝叠加到原始信号之上,延时时间为 50 ms,延时信号幅度为原始信号幅度的 40%.

低通滤波:采用截止频率为 8 kHz,6 kHz,4 kHz,3 kHz,2 kHz 等的低通滤波.

随机剪切:在数字音频的前面(或中间、后面)随机剪切掉 n 秒音频.

幅度缩放:数字音频载体的幅度值被放大到 150%和缩小到 50%.

抖动攻击:首先将整个音频分成 n 个音频样本的数据块,然后在每个数据块中随机复制或删除一个音频样本,接着再将数据块按原来顺序重新组合起来.

表 5 和表 6 给出了本文算法的鲁棒性能测试结果(BER).

表 5 本文算法的抗攻击能力测试(Classic 音频)

攻击类型	未攻击	重新量化(16-8)	重采样(44/11/44)	重采样(44/8/44)	重采样(44/6/44)	添加噪声
水印提取						
BER/%	0.00	0.00	0.00	0.00	2.15	0.00
攻击类型	添加回声	均衡化	低通滤波 3 kHz	低通滤波 2 kHz	高通滤波 50 Hz	MP3-64k
水印提取						
BER/%	0.00	0.00	0.00	2.54	0.00	0.00
攻击类型	随机剪切 1 秒	随机剪切 2 秒	随机剪切 4 秒	随机剪切 6 秒	幅度缩放 70%	幅度缩放 80%
水印提取						
BER/%	0.00	0.00	0.00	1.56	0.00	0.00
攻击类型	幅度缩放 120%	幅度缩放 130%	抖动(1/100000)	抖动(1/50000)	抖动(1/10000)	抖动(1/5000)
水印提取						
BER/%	0.00	0.00	0.00	0.00	1.76	5.08

表 6 不同算法的平均检测性能对比

攻击类型	文献[23]		文献[24]		文献[8]		本文算法	
	BER/%	PSNR	BER/%	PSNR	BER/%	PSNR	BER/%	PSNR
添加噪声(22 dB)	0.18	41.27	0.31	39.83	0.57	41.01	0.09	40.43
低通滤波(8 kHz)	0.56	27.56	0.07	25.83	0.41	28.02	0.08	28.98
高通滤波(50 Hz)	0.48	27.02	0.16	25.56	0.40	27.95	0.11	28.76
添加回声(40% with 50 ms)	3.52	26.04	4.31	24.93	2.04	25.94	0.89	26.79
重量化(16-8)	0.09	42.43	0.06	39.96	0.67	42.77	0.07	43.65
重采样(44/11/44)	0.09	26.87	0.14	26.03	0.08	25.87	0.06	27.64

攻击类型	文献[23]		文献[24]		文献[8]		本文算法	
	BER/%	PSNR	BER/%	PSNR	BER/%	PSNR	BER/%	PSNR
重采样(44/6/44)	0.16	26.34	0.07	26.74	0.26	26.95	0.06	27.21
幅度缩放(130%)	0.09	—	0.06	—	5.04	—	0.06	—
幅度缩放(70%)	0.11	—	0.09	—	4.47	—	0.07	—
MP3 64k	0.07	33.89	0.11	31.68	2.51	36.02	0.06	36.65
MP3 48k	1.88	31.87	0.67	31.02	5.65	34.28	0.64	34.98
MP3 32k	5.45	30.22	0.89	29.14	6.81	32.03	2.62	32.55

以上实验结果表明:本文算法不仅具有较好的不可感知性,而且能够抵抗大部分常规攻击及部分去同步攻击(如随机剪切、幅度缩放、抖动等),较好地解决了鲁棒性与不可感知性之间的平衡问题。这是因为本文算法:(1)引入了同时拥有计算简单、重构精度高、近似平移不变等诸多优良特性的双树复数小波变换;(2)采用刻画能力更强的非线性单调函数,设计了自适应水印嵌入强度函数;(3)以二元 BKF 分布理论为基础,结合变换系数间的多种相关特性,构造了预测能力更强的多相关统计模型,并有效估计了统计模型参数。

6 结 论

本文以双树复数小波变换(Dual-Tree Complex Wavelet Transform,DT CWT)及多元 BKF(Multi-variate Bessel K form,MBKF)分布理论为基础,提出了一种新的统计模型数字音频盲水印算法。该算法在 DT CWT 域采用乘性水印方法进行水印信号嵌入。在水印检测端,结合 DT CWT 系数的尺度内与尺度间相关性,利用二元 BKF 分布对 DT CWT 高频系数进行统计建模并估计出其模型参数,同时利用最大似然决策盲提取水印信息。仿真实验结果表明,本文算法不仅具有较好的不可感知性,而且具有较强的抗常规信号处理和去同步攻击的能力。

参 考 文 献

[1] Liu X L, Lin C C, Yuan S M. Blind dual watermarking for color images' authentication and copyright protection. IEEE Transactions on Circuits and Systems for Video Technology, 2016, DOI: 10.1109/TCSVT.2016.2633878

[2] Hua G, Huang J, Shi Y Q, et al. Twenty years of digital audio watermarking—A comprehensive review. Signal Processing, 2016, 128: 222-242

[3] Verma V S, Jha R K. An overview of robust digital image watermarking. IETE Technical Review, 2015, 32(6): 479-496

[4] Tang Xin, Ma Zhao-Feng, Niu Xin-Xin, Yang Yi-Xian. Blind audio watermarking mechanism based on variational Bayesian learning. Journal on Communications, 2015, 36(1): 121-128(in Chinese)

[5] 唐鑫,马兆丰,钮心忻,杨义先. 基于变分贝叶斯学习的音频水印盲检测方法. 通信学报, 2015, 36(1): 121-128

[5] Rahman S M M, Ahmad M O, Swamy M N S. A new statistical detector for DWT-based additive image watermarking using the Gauss—Hermite expansion. IEEE Transactions on Image Processing, 2009, 18(8): 1782-1796

[6] Singh J, Garg P, De A. Multiplicative watermarking of audio in DFT magnitude. Multimedia Tools and Applications, 2014, 71(3): 1431-1453

[7] Wang X Y, Shi Q L, Wang S M, et al. A blind robust digital watermarking using invariant exponent moments. AEU-International Journal of Electronics and Communications, 2016, 70(4): 416-426

[8] Ngo N M, Unoki M. Method of audio watermarking based on adaptive phase modulation. IEICE Transactions on Information and Systems, 2016, 99(1): 92-101

[9] Dong L, Yan Q, Lv Y, et al. Full band watermarking in DCT domain with Weibull model. Multimedia Tools and Applications, 2017, 76(2): 1983-2000

[10] Amini M, Ahmad M O, Swamy M N S. A new locally optimum watermark detection using vector-based hidden Markov model in wavelet domain. Signal Processing, 2017, 137: 213-222

- [11] Amini M, Ahmad M, Swamy M. A robust multibit multiplicative watermark decoder using vector-based hidden Markov model in wavelet Domain. *IEEE Transactions on Circuits and Systems for Video Technology*, 2016. DOI 10.1109/TCSVT.2016.2607299
- [12] Kalantari N K, Ahadi S M, Vafadust M. A robust image watermarking in the ridgelet domain using universally optimum decoder. *IEEE Transactions on Circuits and Systems for Video Technology*, 2010, 20(3): 396-406
- [13] Sadreazami H, Ahmad M O, Swamy M N S. Multiplicative watermark decoder in contourlet domain using the normal inverse Gaussian distribution. *IEEE Transactions on Multimedia*, 2016, 18(2): 196-207
- [14] Wang Xiang-Yang, Liu Yu-Nan, Xu Huan, et al. Blind optimum detector for robust image watermarking in nonsub-sampled Shearlet domain. *Information Sciences*, 2016, 372: 634-654
- [15] Kwitt R, Meerwald P, Uhl A. Lightweight detection of additive watermarking in the DWT-domain. *IEEE Transactions on Image Processing*, 2011, 20(2): 474-484
- [16] Bian Y, Liang S. Image watermark detection in the wavelet domain using Bessel K densities. *IET Image Processing*, 2013, 7(4): 281-289
- [17] Sadreazami H, Ahmad M O, Swamy M N S. A study of multiplicative watermark detection in the contourlet domain using alpha-stable distributions. *IEEE Transactions on Image Processing*, 2014, 23(10): 4348-4360
- [18] Bian Y, Liang S. Locally optimal detection of image watermarks in the wavelet domain using Bessel K form distribution. *IEEE Transactions on Image processing*, 2013, 22(6): 2372-2384
- [19] Khawne A, Attachoo B, Hamamoto K. Optimum watermark detection of ultrasonic echo medical images. *IEEJ Transactions on Electrical and Electronic Engineering*, 2015, 10(2): 149-156
- [20] Bhinder P, Singh K, Jindal N. Image-adaptive watermarking using maximum likelihood decoder for medical images. *Multimedia Tools & Applications*, 2018, 77(8): 10 303-10 328
- [21] Akhaee M A, Sahraeian S M E, Marvasti F. Contourlet-based image watermarking using optimum detector in a noisy environment. *IEEE Transactions on Image Processing*, 2010, 19(4): 967-980
- [22] Amirmazlaghani M, Rezghi M, Amindavar H. A novel robust scaling image watermarking scheme based on Gaussian mixture model. *Expert Systems with Applications*, 2015, 42(4): 1960-1971
- [23] Akhaee M A, Sahraeian S M E. Scaling-based watermarking with universally optimum decoder. *Multimedia Tools and Applications*, 2015, 74(15): 5995-6018
- [24] Akhaee M A, Kalantari N K, Marvasti F. Robust audio and speech watermarking using Gaussian and Laplacian modeling. *Signal Processing*, 2010, 90(8): 2487-2497
- [25] Asikuzzaman M, Alam M J, Lambert A J, et al. Imperceptible and robust blind video watermarking using chrominance embedding: A set of approaches in the DT CWT domain. *IEEE Transactions on Information Forensics and Security*, 2014, 9(9): 1502-1517
- [26] Boubchir L, Nait-Ali A, Petit E. Multivariate statistical modeling of images in sparse multiscale transforms domain// *Proceedings of the International Conference on Image Processing (ICIP 2010)*. Hong Kong, China, 2010: 1877-1880



WANG Xiang-Yang, M. S., professor. His current research interests include network information security, and intelligent multimedia computing.

LI Li, M. S. candidate. Her current research interests include information hiding and digital watermarking.

LI Hai-Fang, M. S. candidate. Her current research interests include digital watermarking, and multimedia

security.

NIU Pan-Pan, Ph. D., associate professor. Her current research interests include network information security, and image processing.

WANG Si-Miao, M. S. candidate. Her current research interests include information hiding and digital watermarking.

YANG Hong-Ying, M. S., professor. Her current research interests include multimedia information security and image processing.

Background

With the development of Internet technology and multimedia tools, it becomes very easy to distribute and transmit the digital media like image, audio and video. Simultaneously,

the protection of intellectual property has also become more and more attentive and important for the society. Digital watermarking has been proposed and widely studied during

recent years for the purposes of copyright protection and content authentication. There are three indispensable, yet contrasting requirements for a watermarking scheme: robustness, invisibility and capacity. Therefore, a watermarking scheme should provide a trade-off between these features.

In order to solve the tradeoff among the robustness, invisibility, and capacity of the watermark data, statistical properties of the host data are commonly used. Therefore, efforts in this direction have been mostly on the statistical modeling of the transform domain coefficients. There exists several works focusing on watermark decoding using the statistical properties of the transformed domain coefficients, in which the univariate statistical models are employed and only the intra-subband (neighborhood) dependency are captured. In other words, there does not exist any works in digital watermarking (especially audio watermarking) for modeling the across-scale dependencies of the transformed domain coefficients using a multivariate distribution.

In this paper, a new blind detector is proposed for DT CWT-based multiplicative audio watermarking, wherein a PDF based on the multivariate BKF distribution is used. In the presented scheme, watermark data is embedded into the significant high-frequency coefficients in DT CWT domain. At the watermark receiver, DT CWT highpass coefficients are firstly modeled by employing multivariate BKF distribution

according to the inter-scale coefficients dependencies, then the statistical model parameters of multivariate BKF are estimated using the intra-scale coefficients dependencies, and finally a blind watermark extraction approach is developed using the maximum likelihood decision rule. Experiments are performed to verify the imperceptibility and robustness of the proposed watermark decoder. The results show that the proposed blind watermark decoder is superior to other decoders in terms of providing a higher peak signal to noise ratio and lower bit error rate. It is also shown that the proposed decoder is highly robust against various kinds of attacks such as noise addition, MP3 compression, amplitude variation, random cropping, and jittering.

We have been studying the issue of digital watermarking for years. And our work have been published in the top journals and conferences in the multimedia processing field, e. g. *IEEE Transactions on Information Forensics and Security*, *IEEE Transactions on Signal Processing*, *Pattern Recognition*, *Information Sciences*, *Signal Processing*, *Journal of Visual Communication and Image Representation*, and *SCIENCE CHINA Information Sciences*.

This research is supported partially by the National Natural Science Foundation of China (Grant Nos. 61472171, 61272416, and 61701212), and China Postdoctoral Science Foundation (Nos. 2017M621135, 2018T110220).