

# 基于多相关 HMT 模型的 DT CWT 域数字水印算法

王向阳<sup>1</sup> 牛盼盼<sup>1</sup> 杨红颖<sup>1</sup> 李丽<sup>1</sup>

**摘要** 本文以双树复数小波变换 (Dual-tree complex wavelet transform, DT CWT) 及隐马尔科夫树 (Hidden Markov tree, HMT) 理论为基础, 提出了一种基于 Weibull 向量 HMT 模型的 DT CWT 域数字音频盲水印算法. 原始数字音频首先进行 DT CWT, 然后利用局部信息熵刻画音频内容特征并据此确定出重要 DT CWT 系数段, 进而将水印信息乘性嵌入到重要 DT CWT 高频系数幅值内. 水印检测时, 首先根据 DT CWT 系数幅值的边缘分布及系数间的多种相关性 (包括子带内、尺度间、分解树间等相关性), 构造出 Weibull 混合向量 HMT 统计模型, 并估计出其统计模型参数; 然后, 利用局部最大势能 (Locally most powerful, LMP) 检验理论构造出局部最优检测器 (Locally optimum decoder, LOD) 以盲提取水印信息. 仿真实验结果表明, 本文算法可以较好地获得不可感知性、鲁棒性、水印容量之间的良好平衡, 其总体性能优于现有同类音频水印算法.

**关键词** 音频水印, 向量隐马尔科夫树, Weibull 混合模型, 局部信息熵, 双树复数小波变换, 局部最优检测器

**引用格式** 王向阳, 牛盼盼, 杨红颖, 李丽. 基于多相关 HMT 模型的 DT CWT 域数字水印算法. 自动化学报, 2021, 47(12): 2857-2869

**DOI** 10.16383/j.aas.c190075

## A Blind Watermark Decoder in DT CWT Domain Using Weibull Distribution-Based Vector HMT Model

WANG Xiang-Yang<sup>1</sup> NIU Pan-Pan<sup>1</sup> YANG Hong-Ying<sup>1</sup> LI Li<sup>1</sup>

**Abstract** In this paper, we propose a blind audio watermark decoder in dual-tree complex wavelet transform (DT CWT) domain, wherein the Weibull distribution-based vector hidden Markov tree (HMT) model is used. In the proposed watermarking approach, the DT CWT is firstly performed on the original host audio, then the significant DT CWT coefficient segments are determined according to local information entropy, and finally the watermark data is embedded into the significant high-frequency coefficient amplitudes in the DT CWT domain. At the watermark receiver, DT CWT highpass coefficient amplitudes are firstly modeled by employing the Weibull distribution-based vector HMT Model, where both the local statistical properties and various dependencies of the DT CWT coefficients are captured. Then the parameters of the Weibull distribution-based vector HMT model are estimated on the highpass coefficients of digital audio using the maximum likelihood estimation (MLE). And finally, by employing locally most powerful test and the Weibull distribution-based vector HMT model, a blind local optimum decoder (LOD) is developed. We conduct extensive experiments to evaluate the performance of the proposed blind watermark decoder, in which encouraging results validate the effectiveness of the proposed technique.

**Key words** Audio watermarking, vector hidden Markov tree (HMT), Weibull mixture model, local information entropy, dual-tree complex wavelet transform (DT CWT), locally optimum decoder (LOD)

**Citation** Wang Xiang-Yang, Niu Pan-Pan, Yang Hong-Ying, Li Li. A blind watermark decoder in DT CWT domain using Weibull distribution-based vector HMT model. *Acta Automatica Sinica*, 2021, 47(12): 2857-2869

收稿日期 2019-01-30 录用日期 2019-08-08

Manuscript received January 30, 2019; accepted August 8, 2019

国家自然科学基金 (61472171, 61701212), 中国博士后科学基金 (2018T110220), 辽宁省教育厅科学研究经费项目 (面上项目) (LJKZ0985), 辽宁省自然科学基金 (2019-ZD-0468) 资助

Supported by National Natural Science Foundation of China (61472171, 61701212), Project Funded by China Postdoctoral Science Foundation (2018T110220), Scientific Research Project of Liaoning Provincial Education Department (LJKZ0985), and Natural Science Foundation of Liaoning Province (2019-ZD-0468)

本文责任编辑 刘成林

Recommended by Associate Editor LIU Cheng-Lin

1. 辽宁师范大学计算机与信息技术学院 大连 116029

1. School of Computer and Information Technology, Liaoning

作为传统加密方法的有效补充手段, 数字水印 (Digital watermarking) 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, 近年来已引起人们的高度重视, 并成为国际学术界研究的一个热点<sup>[1-3]</sup>. 依据水印信息所附载的数字媒体形式, 人们习惯上将数字水印划分为图像水印、音频水印、视频水印、文本水印等. 本文重点关注数字音频水印技术.

理论上讲, 数字水印技术应同时满足三项基本

Normal University, Dalian 116029

要求<sup>[4]</sup>, 即不可感知性、鲁棒性、水印容量. 其中, 不可感知性是指嵌入的水印信息不应该影响人对载体的感知效果; 鲁棒性则是指在对含水印载体进行常规信号处理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区域或时间内嵌入的水印比特数, 显然水印容量越大越好. 不可感知性、鲁棒性、水印容量是衡量一个数字水印算法优劣的最重要指标, 而这三方面却又存在着固有的相互矛盾关系, 三者之间的最佳平衡也成为数字水印算法所共同追求的目标<sup>[4]</sup>.

数字水印技术发展到今天, 已有大量不同的数字水印算法. 但遗憾的是, 现有绝大多数数字水印方案往往着重关注数字水印的某一单项技术要求 (不可感知性或鲁棒性或水印容量), 并相继提出了不可感知水印、鲁棒水印、大容量水印等方案<sup>[5-8]</sup>, 而没有全面充分地考虑不可感知性、鲁棒性、水印容量三者之间的最佳平衡问题, 从而严重影响了数字水印的隐藏效果. 也就是说, 现有绝大多数数字水印技术均无法有效同时满足数字水印所需的三项基本要求. 因此, 可有效保持不可感知性、鲁棒性、水印容量之间良好平衡的高性能数字水印算法研究仍然是一项富有挑战性的工作.

一般说来, 为全面提高数字水印的不可感知性、鲁棒性和水印容量, 应该充分结合人类视觉掩蔽特性与载体自身统计特性而进行数字水印信息的嵌入与检测<sup>[9]</sup>. 基于统计模型的变换域乘性水印较好地体现了上述思想, 为有效解决不可感知性、鲁棒性、水印容量之间良好平衡问题提供了可能的解决方向<sup>[9]</sup>. 基于统计模型的变换域乘性水印工作原理为: 在水印嵌入环节, 利用简单的乘性策略调制原始载体信号, 以保证水印嵌入强度与载体信号强度成一定比例变化, 让人更不容易察觉, 从而最大程度地平衡鲁棒性和不可感知性; 在水印检测环节, 结合能够充分体现载体自身特性的多尺度变换特性 (多分辨率性、能量聚集性等), 有效利用变换系数的统计特性而构造数字水印检测器, 以检测和提取数字水印信息.

近年来, 基于统计模型的高性能数字水印算法研究日益受到关注, 并已广泛应用于数字图像领域<sup>[10-12]</sup>, 但在数字音频领域内的应用研究却显得相对滞后, 且仅有的统计模型音频水印方案检测性能并不理想<sup>[13-17]</sup>. Akhaee 等<sup>[14]</sup> 首先将水印信息直接嵌入到原始载体的音频样本内, 然后利用 Gaussian 分布和 Laplacian 分布对含水印音频样本进行统计建模, 同时结合极大似然方法估计统计模型参数并构造数字水印检测器. 由于该算法估计模型参数

时需要原始载体部分信息, 属于半盲方案, 故不利于实际应用. Majoul 等<sup>[15]</sup> 结合广义高斯分布 (Generalized Gaussian distribution, GGD) 与最大后验 (Maximum a posteriori, MAP) 参数估计理论, 提出了一种基于统计模型的空间域音频水印方案. 然而, 其数字水印检测器的抗攻击能力较弱. 唐鑫等<sup>[16]</sup> 利用高斯混合模型 (Gaussian mixture model, GMM) 刻画数字音频的 DCT (Discrete cosine transform) 域 Mel 频率倒谱系数 (Mel-scale frequency cepstral coefficients, MFCC) 统计特性, 并使用变分贝叶斯方法估计 GMM 模型参数, 同时结合最大似然检验理论构造了数字水印检测器. Akhaee 等<sup>[17]</sup> 以样本累加统计 (Cumulative sums statistics) 理论与中心极限定理为基础, 提出了一种与载体音频信号分布近似无关的极大似然数字水印检测算法, 但该检测算法的时间复杂度很高.

本文以双树复数小波变换 (Dual-tree complex wavelet transform, DT CWT) 及隐马尔科夫树 (Hidden Markov tree, HMT) 理论为基础, 提出了一种基于 Weibull 向量 HMT 模型的 DT CWT 域数字音频盲水印算法. 由于算法结合局部信息熵自适应确定了数字水印嵌入位置、建立了可全面反映 DT CWT 系数局部统计特性与多种相关性的 Weibull 混合向量 HMT 统计模型、设计了适合于非高斯与小样本的局部最优检测器, 故可获得较好的工作性能.

本文结构安排如下: 第 1 节全面分析了 DT CWT 系数幅值的统计特性; 第 2 节给出了基于 Weibull 混合分布的向量 HMT 模型; 第 3 节和第 4 节叙述了数字水印的嵌入和提取过程; 第 5 节给出了本算法的实验结果及对比结果; 第 6 节对全文进行了总结.

## 1 DT CWT 系数幅值统计特性分析

Kingsbury 等<sup>[18]</sup> 首次提出了双树复小波变换 (Dual-tree complex wavelet transform, DT CWT) 概念, Selesnick 等<sup>[19]</sup> 进一步提出了 DT CWT 的分解与重构算法. 由于 DT CWT 具有近似平移不变、数据冗余有限、可完全重构、计算效率高等优良特性, 故本文采用 DT CWT 进行数字音频信号分解, 并在此基础上利用 DT CWT 系数幅值进行数字水印嵌入与水印检测器构造.

### 1.1 DT CWT 系数幅值的边缘分布分析

DT CWT 是在 DWT 基础上进行了新的改进, 由原来只使用一个滤波树变为两个, 对于一维音频

信号而言, 使用两个滤波树将获得双倍数量的 DWT 小波系数. 由于 DT CWT 拥有近似的平移不变特性, 当输入信号中存在小幅度偏移时, 各个尺度下 DT CWT 系数幅值的能量分布几乎保持不变. 然而实验结果表明, 绝大多数 DT CWT 高频系数幅值接近于零, 即其体现了明显的高峰拖尾特点. 同时, 其峰度值远大于高斯分布的峰度值 3, 表明 DT CWT 高频系数幅值具有明显的非高斯分布特性. 韦伯 (Weibull) 分布<sup>[20]</sup> 是上世纪 40 年代 Wallodi Weibull 提出的一种具有封闭形式的概率密度函数 (Probability density function, PDF), 其 PDF 曲线拖尾厚重, 冲激特性明显. 对于随机变量, Weibull 分布的概率密度函数一般形式为

$$f_{a,b}(x) = \frac{a}{b} \left(\frac{x}{b}\right)^{a-1} \exp\left(-\left(\frac{x}{b}\right)^a\right) \quad (1)$$

其中,  $x$  为随机变量, 且有  $0 < x < \infty$ ;  $a$  与  $b$  分别表示 Weibull 分布中的形状与尺度参数, 并且  $a > 0$ ,  $b > 0$ . 大量研究表明, Weibull 分布非常适合用来建模具有高尖峰、长拖尾分布特征的随机变量, 故在包括图像检索、数字水印在内的诸多统计建模中得到了广泛应用.

鉴于 DT CWT 高频系数幅值分布具有明显的高峰拖尾特点, 而 Weibull 分布能够更加准确地描述出这种拖尾性质, 因此本文采用 Weibull 分布, 对 DT CWT 高频系数幅值进行建模并构造数字水印检测器.

## 1.2 DT CWT 系数幅值的相关性分析

近年来, 研究者普遍关注并利用了变换域系数幅值的高尖峰重拖尾非高斯性, 但不同程度忽略了变换系数幅值之间的强依赖性 (多种相关性). 事实上, DT CWT 系数幅值间存在着子带内、尺度间、分解树间等多种相关性. 其中, 子带内相关性是指在同一分解子带内, 每个 DT CWT 系数幅值与其附近邻域内的 DT CWT 系数幅值密切相关. 尺度间相关性是指某一分解子带内的 DT CWT 系数幅值与其上一尺度对应的 DT CWT 系数幅值密切相关. 分解树间相关性是指不同分解树 (即树 A 和树 B) 相同位置上的 DT CWT 系数幅值间密切相关. 为说明 DT CWT 系数幅值间所存在的子带内、尺度间、分解树间等多种相关性, 以下采用 Chi-plot 工具<sup>[21]</sup> 分析了 DT CWT 域系数幅值间的相互依赖关系. 图 1 给出了数字音频的 DT CWT 域系数幅值子带内、尺度间、分解树间 Chi-plot 图. 不难看出, 图 1 所示 3 幅 Chi-plot 图中的大多数点都比较集中且较大地偏离水平线, 其说明 DT CWT 域系

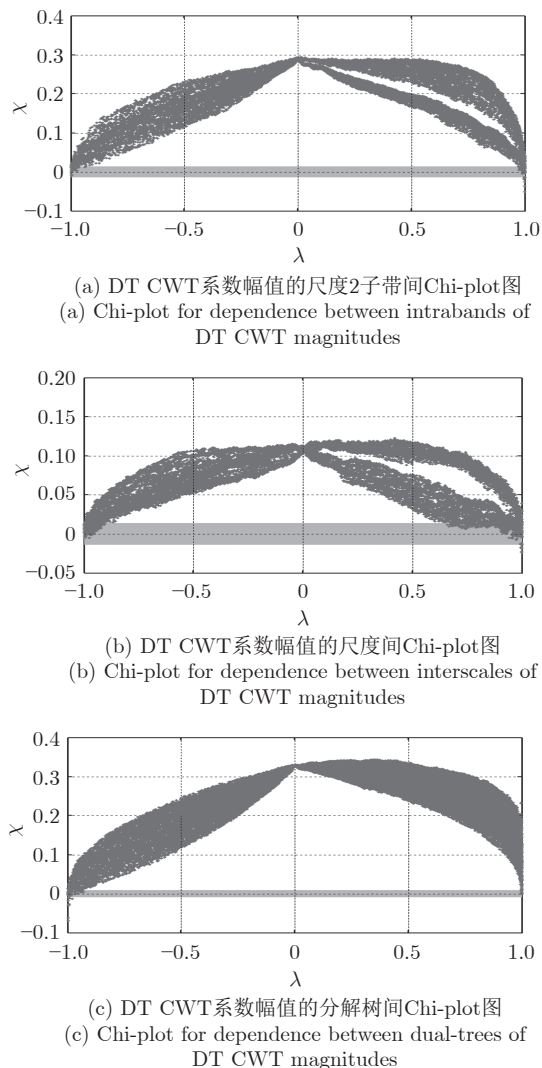


图 1 DT CWT 域系数幅值的子带内、尺度间、分解树间 Chi-plot 图

Fig. 1 Chi-plot to illustrate the different degrees of dependence between intraband, interscale and dual-tree, DT CWT coefficient pairs

数幅值之间存在较强的子带内、尺度间、分解树间相关特性.

## 2 基于 Weibull 混合分布的 DT CWT 域向量 HMT 模型

Crouse 等<sup>[22]</sup> 所提出的小波域 HMT 模型是基于树结构的概率图模型, 其优势在于能够利用简单树型结构有效捕获内部尺度之间的从属特点, 进而全面描述小波系数在尺度间、尺度内的统计相关性. 对于数字音频的 DT CWT 而言, 不仅其 DT CWT 系数幅值与它的父节点幅值同样存在树状结构特点, 而且其不同分解树相同位置上的 DT CWT 系



数幅值密切相关. 受小波域 HMT 思想启发, 本文根据 DT CWT 系数幅值的分布特性与 DT CWT 系数幅值的多种相关性, 建立了如图 2 所示的基于 Weibull 混合分布的 DT CWT 域向量 HMT 模型.

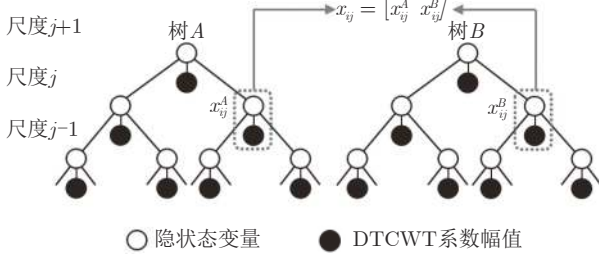


图 2 DT CWT 域向量 HMT 模型

Fig. 2 Vector HMT model in DT CWT domain

在图 2 所示的 DT CWT 域向量 HMT 模型中, 实心点表示 DT CWT 系数幅值, 空心点表示 DT CWT 系数幅值所处的状态. 为叙述方便, 作如下约定: 在具有  $M$  种状态 ( $m = 1, 2, \dots, M$ , 本文选取  $M = 2$ ) 的 HMT 结构中,  $i$  表示任一结点,  $j$  表示尺度,  $x_{ij}$  表示该结点的 DT CWT 系数幅值,  $S_{ij}$  表示该结点系数幅值所对应的隐状态,  $p(i)$  表示该结点的父结点, 且该结点所在树结构的根结点为  $N_{p(i)}$ , 其系数幅值所对应的隐状态为  $S_{p(i)}$ . 鉴于在  $S_{ij} = m$  条件下,  $x_{ij}$  满足 Weibull PDF, 因此有尺度参数  $a_{ij}^m$  和形状参数  $b_{ij}^m$ , 则 DT CWT 系数幅值的边缘分布可定义成如下形式

$$f_{a,b}(x_{ij}) = \sum_{m=1}^M P_{ij}^m \frac{b_{ij}^m}{a_{ij}^m} \left( \frac{x_{ij}}{a_{ij}^m} \right)^{b_{ij}^m-1} \times \exp \left( - \left( \frac{x_{ij}}{a_{ij}^m} \right)^{b_{ij}^m} \right) \quad (2)$$

其中,  $\sum_{m=1}^M P_{ij}^m = 1$ ;  $A_{ij}$  为状态转移概率矩阵, 用于刻画各 DT CWT 系数幅值之间的相互关系,  $A_{ij}$  定义形式如下

$$A_{ij} = \begin{bmatrix} p_{ij}^{1 \rightarrow 1} & p_{ij}^{1 \rightarrow 2} & \dots & p_{ij}^{1 \rightarrow M} \\ p_{ij}^{2 \rightarrow 1} & \dots & \dots & p_{ij}^{2 \rightarrow M} \\ \vdots & \vdots & \ddots & \vdots \\ p_{ij}^{M \rightarrow 1} & \dots & \dots & p_{ij}^{M \rightarrow M} \end{bmatrix}_{M \times M} \quad (3)$$

其中,  $P_{ij}^{m \rightarrow m'}$  代表状态转移概率, 其中父系数处于  $m'$  状态, 子系数处于  $m$  状态, 且  $m' = 1, 2, \dots, M$ ,  $j = 1, 2, \dots, J$ . 其父结点与子结点间的状态转移概率可表示为

$$P(S_{ij} = m) = \sum_{m'} P(S_{p(i)} = m') \times P(S_{ij} = m | S_{p(i)} = m') \quad (4)$$

上式的含义是, 已知父结点  $p(i)$ , 且其隐状态变量  $S_{p(i)}$  的值为  $m'$ , 与之相对应的子结点  $i$ , 其隐状态变量  $S_{ij}$  的值为  $m$  的条件概率.

不难看出, 以上 DT CWT 域 HMT 结构仅仅利用 Weibull 混合模型, 描述与刻画了 DT CWT 系数幅值的非高斯性与尺度相关性.

为有效捕获 DT CWT 域系数幅值的分解树间相关性, 从而显著增强 DT CWT 系数幅值的统计建模能力, 可以进一步将分解树  $A$  和  $B$  上位于相同位置  $i$  和尺度  $j$  的 DT CWT 系数幅值  $x_{ij}^A$  和  $x_{ij}^B$  组成 DT CWT 系数幅值向量  $x_{ij} = [x_{ij}^A, x_{ij}^B]^T$  (参见图 2 虚线框部分), 并利用协方差矩阵  $C_{ij}^m$  来描述  $x_{ij}^A$  和  $x_{ij}^B$  之间的相关性, 进而得到 DT CWT 域向量 HMT 模型.

实际应用中, 为减少 DT CWT 域向量 HMT 模型的参数数量 (即降低复杂度), 本文将同一尺度  $j$  下的所有结点近似看作含有相等的统计量, 即有  $A_{ij} = A_j$ ,  $P_{ij}^m = P_j^m$ ,  $a_{ij}^m = a_j^m$ ,  $b_{ij}^m = b_j^m$ ,  $C_{ij}^m = C_j^m$ . 于是, 基于 Weibull 混合分布的 DT CWT 域向量 HMT 的位于  $j$  尺度上的 DT CWT 系数幅值的状态概率分布函数可定义为式 (5) (见本页下方).

由式 (5) 可以看出, 基于 Weibull 混合分布的 DT CWT 域向量 HMT 模型由以下参数组成

$$\Theta = \{p_1, A_2, \dots, A_J; a_j^m, b_j^m, C_j^m\},$$

$$\forall j = 1, 2, \dots, J; m = 1, 2, \dots, M \quad (6)$$

其中,  $p_1$  为初始状态概率;  $A_2, \dots, A_J$  为状态转移概率矩阵;  $a_j^m$  和  $b_j^m$  为 Weibull PDF 的尺度参数和形状参数;  $C_j^m$  表示用于刻画 DT CWT 域系数幅值分解树间相关性的协方差矩阵.  $J$  表示分解尺度数,  $M$  表示状态数. 对基于 Weibull 混合分布的 DT CWT 域向量 HMT 模型参数, 本文拟采用 HMT-EM 算法迭代进行最优估计, 其具体计算过程如下:

$$f_{a,b}(x_{ij}) = \sum_{m=1}^M \frac{P_j^m \frac{b_j^m}{a_j^m} \left( \frac{(x_{ij})^T (C_j^m)^{-1} (x_{ij})}{a_j^m} \right)^{b_j^m-1} \exp \left( - \left( \frac{(x_{ij})^T (C_j^m)^{-1} (x_{ij})}{a_j^m} \right)^{b_j^m} \right)}{\sqrt{|\det(C_j^m)|}} \quad (5)$$

**步骤 1.** 初始化模型参数集  $\Theta = \Theta_0$ , 并令迭代计数器  $l = 0$ ;

**步骤 2.** E 步骤: 对参数集  $\Theta^l$  的隐状态变量  $S$ , 计算其概率质量函数  $P(S_{ij}|x, \Theta^l)$  与  $P(S_{ij}, S_{\rho(i)}|x, \Theta^l)$ ;

**步骤 3.** M 步骤: 更新 HMT 模型的参数  $P(S_{ij} = m' | S_{\rho(i)} = m)$ ,  $a_{ij}^m$ ,  $b_{ij}^m$ ,  $c_{ij}^m$ ;

**步骤 4.** 令  $l = l + 1$ , 如果参数集  $\Theta$  收敛则停止, 否则返回步骤 2.

### 3 数字水印嵌入

本文选取 DT CWT 域分解树  $A$  的第 2 尺度高频子带作为水印嵌入区, 并将数字水印信息嵌入到局部信息熵较大的 DT CWT 高频系数段内, 以期同时提高数字水印的不可感知性和鲁棒性.

假设原始数字音频信号为  $A = \{a(i), 1 \leq i \leq L\}$ . 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  $W = \{w(i, j), 1 \leq i \leq I, 1 \leq j \leq J\}$ . 其中,  $w(i, j) \in \{0, 1\}$  代表二值水印图像的第  $i$  行、第  $j$  列象素值. 则数字水印嵌入过程可描述如下.

#### 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:  $W = \{w_k\} (k = 1, 2, \dots, N)$ ,  $N = I \times J$ , 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, 以得到分解树  $A$  和分解树  $B$ . 其中, 分解树  $A$  和分解树  $B$  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, 本文选取分解树  $A$  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 系数段确定

信息熵是系统有序化程度的一个度量, Shannon 将其定义为离散随机事件的出现概率. 本质上讲, 信息熵是一种特征的统计形式, 它反映了信源中所含有信息量的多少. 近年来, 信息熵理论已广泛应用于物理、数学、统计、计算机等众多领域. 在

信号处理方面, 信息熵可以估测信号所含有的平均信息量, 普遍认为信号中熵值越大的区域, 需要获取的信息量就越大, 这种局部区域通常被认作是显著性局域 (即重要区域).

鉴于数字音频样本 (值) 具有相互独立性, 而信息熵能够从显著性角度估计特征位置及其周围信息, 进而获得更多的数字音频关键内容, 故本文引入局部信息熵描述数字音频的内容特征, 并据此确定出适合水印嵌入的重要 DT CWT 系数段. 具体步骤如下:

**步骤 1.** 构造局部 DT CWT 系数段. 所谓局部 DT CWT 系数段, 是指以分解树  $A$  尺度 2 高频子带中的每个 DT CWT 高频系数为标识, 从尺度 2 高频子带中分割出的一部分子系数. 为了减少计算量, 本文以每个 DT CWT 高频系数为标识, 选取出其前面  $D/2$  个高频系数和后面  $D/2$  个高频系数组成局部 DT CWT 系数段.

**步骤 2.** 计算每个 DT CWT 系数段的局部信息熵. 首先计算出局部系数段中 DT CWT 系数幅值  $x(i)$  ( $i = 1, \dots, D$ ) 在整个局部系数段中所出现的概率, 即

$$p_i = \frac{x(i)}{\sum_{i=1}^D x(i)} \quad (7)$$

然后计算出 DT CWT 系数段的局部信息熵, 即

$$H = - \sum_{i=1}^n p_i \log_2 p_i \quad (8)$$

由上述定义知, 局部信息熵与相应 DT CWT 系数的邻域系数幅值密切相关, 其反映了局部系数段内 DT CWT 系数幅值的离散程度. 对于局部信息熵较大的 DT CWT 系数段, 其 DT CWT 系数幅值离散小, 说明 DT CWT 系数幅值分布较均匀. 反之, 对局部信息熵较小的 DT CWT 系数段, 其 DT CWT 系数幅值散度大, 说明 DT CWT 系数幅值分布较不均匀. 局部信息熵表征了局部音频所含信息量的多少. 局部信息熵越大的 DT CWT 系数段所含信息量越多, CWT 系数幅值分布无序程度越低. 反之, 局部信息熵越小的 DT CWT 系数段所含信息量越少, 相应的 DT CWT 系数幅值分布无序程度越高.

**步骤 3.** 选取重要局部 DT CWT 系数段. 选取局部信息熵较大的前  $N$  个局部 DT CWT 系数段, 用以水印嵌入与检测.

#### 3.4 数字水印嵌入

设  $A_k$  表示原始的重要 DT CWT 系数段  $A_k =$

$\{x_{k1}, x_{k2}, \dots, x_{kD}\}$ ,  $A'_k$  表示含水印的重要 DT CWT 系数段  $A'_k = \{y_{k1}, y_{k2}, \dots, y_{kD}\}$ , 则数字水印嵌入策略可表示为

$$y_{ki} = \begin{cases} x_{ki}(1 + \alpha), & \text{若 } w_k = 1 \\ x_{ki}, & \text{若 } w_k = 0 \end{cases}, \quad x_{ki} \in A_k \quad (9)$$

其中,  $\alpha$  为水印嵌入强度,  $x_{ki}$  和  $y_{ki}$  分别表示原始 DT CWT 高频系数幅值和含水印 DT CWT 高频系数幅值,  $i = 1, 2, \dots, D, k = 1, 2, \dots, I \times J$ .

重复上述数字水印嵌入步骤, 直到分解树  $A$  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

本文将局部最大势 (Locally most power, LMP) 检验理论为基础, 结合基于 Weibull 混合分布的 DT CWT 域向量 HMT 模型, 首次构造出可有效提取水印信息的多相关局部最优水印检测器 (Locally optimum decoder, LOD). 数字水印检测可以看作二元假设问题, 在此假设下,  $H_1$  和  $H_0$  分别表示 DT CWT 系数幅值嵌入 “1” 及嵌入 “0” 时的假设, 可以表示为

$$\begin{aligned} H_1 : y_{ki} &= x_{ki} \cdot (1 + \alpha), & \text{若 } w_k &= 1 \\ H_0 : y_{ki} &= x_{ki}, & \text{若 } w_k &= 0 \end{aligned} \quad (10)$$

其中,  $\alpha$  为水印嵌入强度,  $w_k$  表示水印信息位,  $i \in A'_k$ , 且  $A'_k$  为含水印 DT CWT 系数段,  $x_{ki}$  和  $y_{ki}$  分别表示原始 DT CWT 高频系数幅值和含水印 DT CWT 高频系数幅值.

对于含水印 DT CWT 系数段  $A'_k$  中的系数幅值  $y_{ki}$ , 可根据最大似然统计理论, 将水印检测器设计转化为如下决策表达式

$$\prod_{i \in A'_k} f_Y(y_{ki} | w_k = 1) \begin{matrix} H_1 \\ > \\ < \\ H_0 \end{matrix} \prod_{i \in A'_k} f_Y(y_{ki} | w_k = 0) \quad (11)$$

其中,  $f_Y(\cdot)$  表示含水印系数幅值在两种假设下的条件概率.

分别对上式两边取自然对数, 并经过移项后可得到如下对数似然比, 即 LOD 检测器势函数

$$L_k(y) = \sum_{i \in A'_k} \ln \frac{f_Y(y_{ki} | w_k = 1)}{f_Y(y_{ki} | w_k = 0)} \begin{matrix} H_1 \\ > \\ < \\ H_0 \end{matrix} 0 \quad (12)$$

由局部最大势 (LMP) 检验理论知, LOD 检测器将在水印嵌入强度  $\alpha = 0$  时, 使其势函数 (检测概率) 的斜率达到最大<sup>[23]</sup>, 于是可得到 LOD 检测统计量

$$L_k(y) = \sum_{i \in A'_k} \ln \frac{f_Y(y_{ki} | w_k = 1)}{f_Y(y_{ki} | w_k = 0)} \begin{matrix} H_1 \\ > \\ < \\ H_0 \end{matrix} 0 \quad (13)$$

其中,  $g_1, g_0$  分别表示嵌入函数  $y_1(x) = x(1 + \alpha)$  和  $y_0(x) = x$  的反函数.

鉴于数字水印信息属于弱信号, 且该弱信号不会改变原始 DT CWT 高频系数幅值的统计特性, 即可认为含水印 DT CWT 高频系数幅值仍符合 Weibull 向量 HMT 分布. 于是, 将基于 Weibull 混合分布的 DT CWT 域向量 HMT 的 PDF (即式 (5)) 代入 LOD 检测统计量 (即式 (13)), 可得到式 (14) (见本页下方).

其中,  $(g_1)_{ki} = \frac{y_{ki}}{1 + \alpha}, (g_0)_{ki} = y_{ki}, \mathbf{y}_{ki} = [y_{ki}, y_{ki}]^T, C_p^m$  为  $2 \times 2$  的协方差矩阵,  $a_q^m$  为尺度参数,  $b_q^m$  为形状参数,  $P_q^m$  为初始状态概率,  $m = 1, 2, \dots, M, q$  表示最粗尺度,  $D$  表示重要 DT CWT 系数段长度. 在本文中, 所有这些参数均采用 HMT-EM 算法迭代进行最优估计.

于是, 可以按照如下方式, 从第  $k$  个含水印 DT CWT 系数段  $A'_k$  中提取出第  $k$  个水印信息位, 即

$$\hat{w}_k = \begin{cases} 1, & Z_k(y) \geq T_k \\ 0, & Z_k(y) < T_k \end{cases} \quad (15)$$

$$LOD_k(y) = D \ln \frac{1}{1 + \alpha} + \sum_{i \in A'_k} \ln \frac{\sum_{m=1}^M \frac{P_q^m b_q^m}{a_q^m} \left\{ \left( \frac{(g_1)_{ki}^T (C_q^m)^{-1} (g_1)_{ki}}{a_q^m} \right)^{b_q^m - 1} \exp \left( - \left( \frac{(g_1)_{ki}^T (C_q^m)^{-1} (g_1)_{ki}}{a_q^m} \right)^{b_q^m} \right) \right\} \theta_{(g_1)_{ki}}}{\sum_{m=1}^M \frac{P_q^m b_q^m}{a_q^m} \left\{ \left( \frac{(g_0)_{ki}^T (C_q^m)^{-1} (g_0)_{ki}}{a_q^m} \right)^{b_q^m - 1} \exp \left( - \left( \frac{(g_0)_{ki}^T (C_q^m)^{-1} (g_0)_{ki}}{a_q^m} \right)^{b_q^m} \right) \right\} \theta_{(g_0)_{ki}}} \quad (14)$$

其中,  $Z_k$  见式 (16) (见本页下方).

$$T_k = D \ln(1 + \alpha) \quad (17)$$

## 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:

**步骤 1.** 对含水数字音频做二级 DT CWT 分解, 并选取分解树  $A$  的尺度 2 高频子带用于水印提取.

**步骤 2.** 利用分解树  $A$  和  $B$  的尺度 2 和尺度 1 高频子带的 DT CWT 系数幅值, 构造基于 Weibull 混合分布的 DT CWT 域向量 HMT 模型并利用 EM 算法进行模型参数估计, 同时采纳第 4.1 节中的方法构造局部最优数字水印检测器.

**步骤 3.** 采用类似于第 3.3 节的工作步骤, 在分解树  $A$  的尺度 2 高频子带中确定重要 DT CWT 系数段.

**步骤 4.** 利用步骤 2 所构造的局部最优数字水印检测器, 从重要 DT CWT 系数段中提取出水印信息位.

**步骤 5.** 重复步骤 4, 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

$$W^* = \{w_k^*, k = 1, 2, \dots, N\} \quad (18)$$

**步骤 6.** 对  $W^*$  做升维与逆 Arnold 变换, 即可获得最终的二维水印图像

$$\hat{W} = \{\hat{w}(i, j), 1 \leq i \leq I, 1 \leq j \leq J\} \quad (19)$$

## 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 [5, 7-8, 17] 进行了对比. 实验中, 本文选取了 10 段 5 种不同风格的数字音频信号进行性能测试, 包括 Popular、Classical、Speech、Rock、Jazz 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  $32 \times 32$  像素的二值图像. 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 MATLAB 7.12.0.

### 5.1 实验参数选取

由数字水印嵌入原理知, 局部 DT CWT 系数段长度 ( $D$ ) 将影响整个算法的工作性能. 为合理选取局部 DT CWT 系数段长度, 本文利用 10 段 5 种不同风格的数字音频信号, 给出了实验参数  $D$  与算法平均工作性能的关系, 即在不同局部 DT CWT 系数段长度下, 整个水印系统的平均峰值信噪比 (Peak signal to noise ratio, PSNR) (dB)、平均比特出错概率 (Bit error ratio, BER) (%)、平均水印嵌入时间 (s) 以及平均水印提取时间 (s) 等工作性能, 如表 1 所示.

表 1 测试结果表明, 当局部 DT CWT 系数段长度  $D$  设定为 60 时, 可以获得较好的水印算法工作性能.

### 5.2 本文算法的工作性能测试

本文首次结合局部最大势检验与 DT CWT 域 Weibull 混合向量 HMT 统计建模理论, 构造了可有效提取数字水印信息的多相关局部最优水印检测器 (Locally optimum decoder, LOD). 这里, 本文对 LOD 水印检测器工作性能进行了评测, 并与常用的最大似然 (Maximum likelihood, ML) 检测器进行了对比. 实验中, 首先从测试音频内任意选取一个局部重要 DT CWT 系数段, 并将水印信息位 “1” 或 “0” 嵌入到其 DT CWT 系数幅值中. 然后分别计算每个 DT CWT 系数幅值的 LOD、ML 检测响应值及检测阈值. 由数字水印检测原理知, 如果该 DT CWT 系数段的累积响应值大于累积阈值 (即响应曲线与阈值曲线所围成面积大于 0), 则判断该 DT CWT 系数段嵌入了水印信息位 “1”; 否则 (即响应曲线与阈值曲线所围成面积小于 0), 判断该系数段嵌入了水印信息位 “0”. 同时, 响应曲线与阈值曲线所围成面积 (绝对值) 越大, 则说明检测响应效果越好, 即水印检测器工作能力越强.

$Z_k =$

$$\sum_{i \in A'_k} \ln \frac{\sum_{m=1}^M \frac{P_q^m b_q^m \left( \frac{(g_1)_{ki}^T (C_q^m)^{-1} (g_1)_{ki}}{a_q^m} \right)^{b_q^{m-2}} \exp \left( - \left( \frac{(g_1)_{ki}^T (C_q^m)^{-1} (g_1)_{ki}}{a_q^m} \right)^{b_q^m} \right) (C_q^m)^{-1} (g_1)_{ki} (g_1)_{ki} + (g_1)_{ki}^T}{a_q^{m2} \sqrt{|\det(C_q^m)|}}}{\sum_{m=1}^M \frac{P_q^m b_q^m \left( \frac{(g_0)_{ki}^T (C_q^m)^{-1} (g_0)_{ki}}{a_q^m} \right)^{b_q^{m-2}} \exp \left( - \left( \frac{(g_0)_{ki}^T (C_q^m)^{-1} (g_0)_{ki}}{a_q^m} \right)^{b_q^m} \right) (C_q^m)^{-1} (g_0)_{ki} (g_0)_{ki} + (g_0)_{ki}^T}{a_q^{m2} \sqrt{|\det(C_q^m)|}}} \quad (16)$$



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 1 024 位)

Table 1 Performance of the watermarking system in different coefficient lengths (watermark capacity is 1 024 bits)

DT CWT系数段长度	平均PSNR (dB)	平均BER (%)	平均水印嵌入时间 (s)	平均水印提取时间 (s)
40	48.45	0.29	0.45	8.42
50	47.96	0.09	0.62	9.57
60	47.28	0.00	0.71	10.67
70	46.34	0.00	1.21	11.53

图 3 给出两种假设下, LOD 检测器和 ML 检测器的水印提取结果. 图 3 评测结果表明, 本文所提出的 LOD 检测器及目前普遍使用的 ML 检测器均能准确提取出水印信息, 但 LOD 检测器的灵敏度明显高于 ML 检测器. 因为 LOD 检测响应曲线与阈值曲线所围成面积 (绝对值) 远大于 ML 检测响应曲线与阈值曲线所围成面积.

表 2 利用 10 段 5 种不同风格的数字音频信号, 给出了水印容量与算法平均工作性能的关系, 即不同水印容量 (包括 64 位、256 位、1 024 位及 4 096 位) 下, 整个水印系统的平均 PSNR (dB)、平均 BER (%)、平均水印嵌入时间 (s) 以及平均水印提取时间 (s) 等工作性能. 从表 2 实验数据可以得出, 随着水印容量增大, PSNR 不断下降, BER 与嵌入提取时间持续增加. 考虑到原则上应该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  $32 \times 32$  时, 水印提取效果达到了比较好的平衡, 因此本文选取了  $32 \times 32$  像素的二值图像水印.

为了更加直观地显示数字水印的感知透明性, 图 4 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, 本文还采用 PSNR 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, 计算出最终的客观差异等级 (Objective difference grade, 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.

从图 4 不难看出, 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 4(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, 说明本算法具有较好的不可感知性.

为验证本文算法的鲁棒性, 这里对含水印音频进行了一系列攻击实验, 包括重新量化 (Requantization)、重新采样 (Resampling)、叠加噪声 (Additive of Gaussian noise)、添加回声 (Echo addition)、低通滤波 (Lowpass filtering)、MP3 压缩 (MP3 compression) 等常规信号处理, 及随机剪切 (Random cropping)、幅值改变 (Amplitude scaling)、抖动 (Jittering) 等去同步攻击. 图 5 给出了本文算法的鲁棒性能测试结果, 包括所提取出的数字水印、BER (%) 等.

### 5.3 不同算法的工作性能对比

为充分客观地评价本文算法的有效性, 以下对本文算法与文献 [5, 7-8, 17] 的水印检测性能进行了对比. 表 5~8 给出了 10 段 5 种不同风格的数字音频信号的平均检测性能. 对比实验中, 采用了相同大小的数字水印容量 (1 024 位).

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信号处理 (如重新量化、重新采样、叠加噪声、添加回声、低通滤波、MP3 压缩等) 及部分去同步攻击 (如随机剪切、幅度缩放、抖动等), 较好地解决了不可感知性、鲁棒性、水印容量之间的良好平衡问题. 这是因为本文算法: 1) 引入局部信息熵描述数字音频的内容特征, 并据此自适应确定了适合水印嵌入的重要 DT CWT 系数段 (位置); 2) 结合 DT CWT 系数幅值分布及多种相关特性, 建立了预测能力更强的基于 Weibull 混合分布的 DT CWT 域向量 HMT 统计模型, 并有效估计了统计模型参数; 3) 以局部最大势能检验理论为基础, 首次构造了适合于非高斯与小样本的可有效提取水印信息的多相关局部最优水印检测器.

## 6 结束语

本文以双树复数小波变换 (Dual-tree complex wavelet transform, DT CWT) 及隐马尔科夫树 (Hidden Markov tree, HMT) 理论为基础, 结合 DT CWT 系数幅值分布及多种相关特性, 提出了



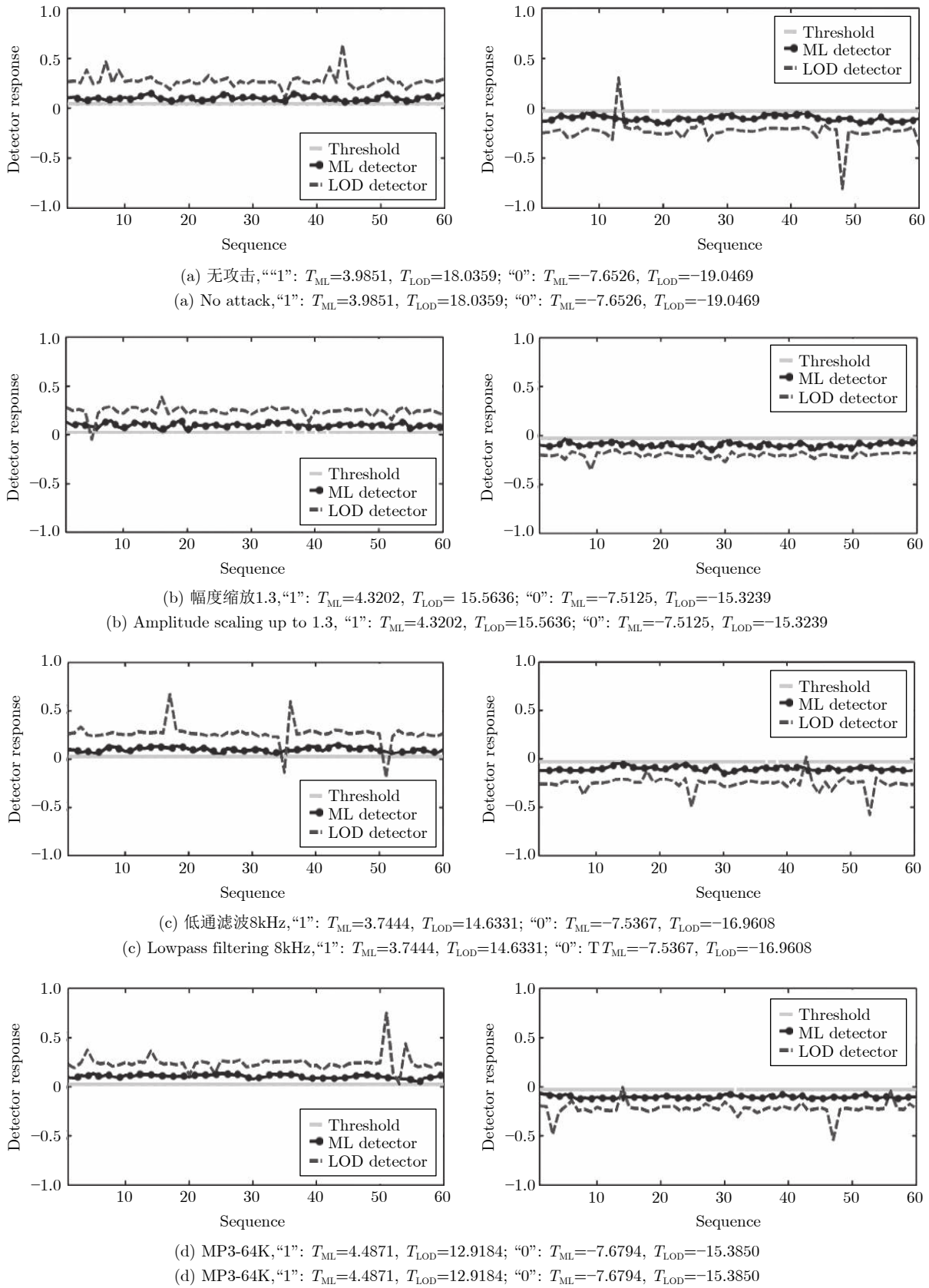


图 3 两种数字水印检测器 (LOD 和 ML) 的检测响应结果

Fig.3 The test results of LOD watermark detector and ML watermark detector under various attacks

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 60)

Table 2 Performance of the watermarking system in different watermark capacities (coefficient length is 60)

水印容量 (bit)	PSNR (dB)	BER (%)	水印嵌入时间 (s)	水印提取时间 (s)
$8 \times 8$	49.23	0.00	0.47	9.53
$16 \times 16$	48.16	0.00	0.68	10.24
$32 \times 32$	47.28	0.00	0.71	10.67
$64 \times 64$	44.34	0.78	1.35	15.46

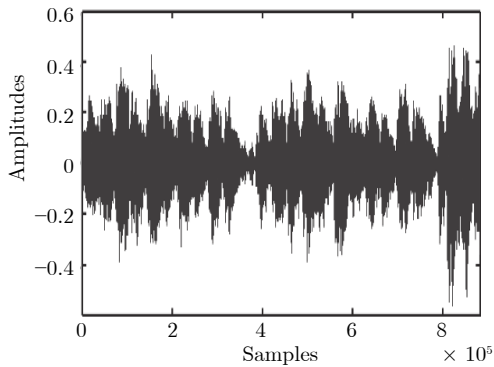
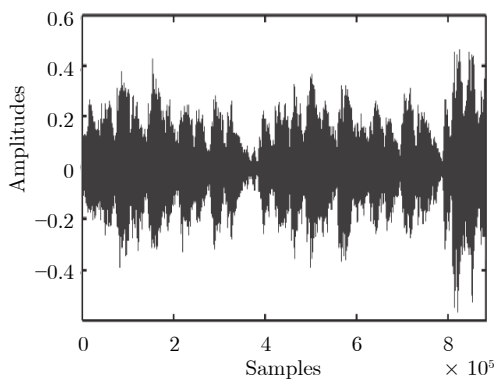
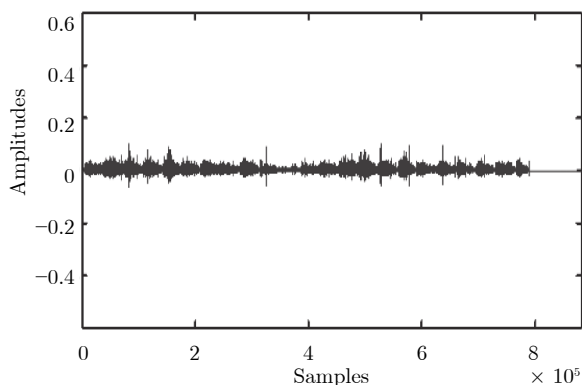
(a) 原始音频与原始水印  
(a) The waveform of original audio(b) 含水印音频与提取的水印  
(b) The waveform of watermarked audio(c) 差值  
(c) The waveform of difference audio

图 4 原始音频、含水印音频和差值音频波形图

Fig.4 The waveform of original audio, watermarked audio, and difference audio

表 3 客观听觉测试区分度 ODG  
Table 3 Objective difference grades

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

Table 4 Objective evaluation of perceived transparency and robustness

音频类型	ODG	BER (%)
Popular	-0.61	0.00
Jazz	-0.48	0.03
Rock	-0.68	0.16
Speech	-0.35	0.01
Classical	-0.79	0.08

一种基于 Weibull 向量 HMT 模型的 DT CWT 域数字音频盲水印算法. 该算法利用局部信息熵刻画数字音频内容并依此自适应确定水印嵌入位置, 同时将水印信息乘性嵌入到 DT CWT 高频系数幅值内. 进行数字水印检测时, 首先根据 DT CWT 系数幅值的非高斯分布特性及 DT CWT 系数的子带内、方向间、尺度间等多种相关特性, 建立起基于 Weibull 混合分布的 DT CWT 域向量 HMT 统计模型并估计出其模型参数, 然后根据局部最大势能检验理论, 利用基于 Weibull 混合分布的 DT CWT 域向量 HMT 统计模型参数, 构造出多相关局部最优数字水印检测器并盲提取水印信息. 仿真实验结果表明, 本文算法不仅具有较好的不可感知性, 而且能够抵抗大部分常规信号处理 (如重新量化、重新采样、叠加噪声、添加回声、低通滤波、MP3 压缩等) 及部分去同步攻击 (如随机剪切、幅度缩放、抖动等), 较好地解决了不可感知性、鲁棒性、水印容量之间的良好平衡问题, 其总体性能优于现有同类方案.

Attack types	No attack	Requantization (16 bit - 8 bit - 16 bit)	Resampling (11 025 Hz)	Resampling (8 000 Hz)	Resampling (6 000 Hz)	Additive of Gaussian noise	Echo addition (1 s delay, 10% decay)	Audio equalization	MP3 compression (256 kbps)
Extracted watermark									
BER (%)	0	0	0	0	0	0	0.39	0	0
PSNR(dB)	45.9823	43.9186	22.7800	22.9953	27.6388	42.9922	26.6733	41.5723	43.8910
Attacks	MP3 compression (128 kbps)	MP3 compression (112 kbps)	MP3 compression (64 kbps)	MP3 compression (48 kbps)	Lowpass filtering (8 kHz)	Lowpass filtering (6 kHz)	Lowpass filtering (4 kHz)	Lowpass filtering (3 kHz)	Lowpass filtering (2 kHz)
Extracted watermark									
BER (%)	0	0	0	0.13	0	0	0	0	0.49
PSNR (dB)	41.3166	41.3283	34.7130	13.5525	30.0974	29.2364	28.6896	26.4917	23.9212
Attacks	Cropping (front) 1 s	Cropping (front) 2 s	Cropping (front) 3 s	Cropping (front) 4 s	Cropping (front) 5 s	Cropping (front) 6 s	Amplitude scaling up to 1.5	Amplitude scaling up to 1.4	Amplitude scaling up to 1.3
Extracted watermark									
BER (%)	0	0	0	0	0	0.38	0	0	0
Attacks	Amplitude scaling up to 1.2	Amplitude scaling up to 1.1	Amplitude scaling down to 0.9	Amplitude scaling down to 0.8	Amplitude scaling down to 0.7	Jittering (1/100 000)	Jittering (1/50 000)	Jittering (1/10 000)	Jittering (1/5 000)
Extracted watermark									
BER (%)	0	0	0	0	0	0	0	0.59	1.46

图 5 本文算法的鲁棒性能测试结果

Fig.5 The watermark detection results for various attacks

表 5 不同算法的平均检测性能对比 (本文算法和文献 [8])

Table 5 Comparison of average detection performance of different schemes (our scheme and [8])

攻击类型	本文算法	文献 [8]				
		DWT-RDM-W+DC	DWT-RDM-W	DWT-ROM	DWT-LQIM	DWT-norm
Resampling (22 050 Hz)	0.000	0.000	0.000	0.000	0.000	0.000
Requantization (16 bit - 8 bit - 16 bit)	0.000	0.000	0.000	0.000	0.001	0.000
Amplitude scaling down to 0.85	0.000	0.000	0.000	0.000	0.000	73.723
Additive of Gaussian noise	0.000	0.002	0.006	0.003	0.006	0.000
Lowpass filtering (4 kHz)	0.000	0.219	0.262	0.290	0.376	0.184
Echo addition (50 ms delay, 5% decay)	0.021	0.026	0.041	0.041	0.187	0.142
MP3 compression (128 kbps)	0.000	0.000	0.000	0.000	0.000	0.002
MP3 compression (64 ps)	0.000	0.152	0.169	0.211	0.256	0.258

表 6 不同算法的平均检测性能对比 (本文算法和文献 [5])

Table 6 Comparison of average detection performance of different schemes (our scheme and [5])

攻击类型	本文算法		文献 [5]			
	Classical	Popular	Classical $N=8$	Popular $N=4$	Classical $N=8$	Popular $N=4$
Resampling (22 050 Hz)	0.00	0.00	0.00	0.29	0.00	0.71
Resampling (11 025 Hz)	0.00	0.00	0.00	1.22	0.00	1.06
Resampling (8 000 Hz)	0.02	0.02	0.01	1.25	0.02	1.20
Lowpass filtering (3 kHz)	0.00	0.35	24.03	27.32	26.06	23.85
MP3 compression (128 kbps)	0.00	0.00	0.11	0.13	0.08	0.08
MP3 compression (112 kbps)	0.00	0.00	0.13	0.11	0.06	0.12
MP3 compression (96 kbps)	0.00	0.00	1.01	2.07	1.06	1.40
MP3 compression (80 kbps)	0.01	0.00	1.57	3.65	1.16	2.50

表 7 不同算法的平均检测性能对比 (本文算法和文献 [7])

Table 7 Comparison of average detection performance of different schemes (our scheme and [7])

攻击类型	-30 dB WSR		-25 dB WSR		-20 dB WSR	
	文献 [7]	本文算法	文献 [7]	本文算法	文献 [7]	本文算法
No Attack	0.14	0.00	0.04	0.00	0.00	0.00
MP3 compression (64 kbps)	32.12	10.35	25.12	0.83	17.11	0.04
MP3 compression (128 kbps)	24.22	0.33	19.54	0.04	12.01	0.01
Resampling (24 kHz)	0.21	0.00	0.12	0.00	0.01	0.00
Resampling (16 kHz)	8.17	0.03	6.35	0.02	3.23	0.00
Additive of Gaussian noise (30 dB)	13.07	0.10	11.06	0.09	9.04	0.01
Lowpass filtering (12 kHz)	0.41	0.13	0.22	0.00	0.04	0.00
Amplitude scaling down to 0.7	0.43	0.41	0.31	0.01	0.05	0.00

表 8 不同算法的平均检测性能对比 (本文算法和文献 [17])

Table 8 Comparison of average detection performance of different schemes (our scheme and [17])

攻击类型	本文算法			文献 [17]		
	Classical	Popular	Speech	Classical	Popular	Speech
Additive of Gaussian noise (22 dB)	0.00	0.00	0.00	0.00	0.00	0.14
Lowpass filtering (8 kHz)	0.00	0.00	0.00	1.23	0.72	0.54
Highpass filtering (50 Hz)	0.00	0.00	0.00	0.00	0.05	0.45
Echo addition (50 ms delay, 40% decay)	0.21	0.34	0.41	1.56	0.70	3.23
Requantization (16 bit-8 bit-16 bit)	0.00	0.00	0.00	0.00	0.00	0.00
Resampling (11 025 Hz)	0.00	0.00	0.00	0.00	0.00	0.06
Resampling (6 000 Hz)	0.00	0.00	0.00	0.00	0.00	0.15
Amplitude scaling up to 1.3	0.00	0.00	0.00	0.00	0.00	0.00
Amplitude scaling down to 0.7	0.00	0.00	0.00	0.00	0.00	0.00
MP3 compression (64 kbps)	0.00	0.00	0.00	0.02	0.00	0.06
MP3 compression (48 kbps)	0.11	0.01	0.54	0.13	0.00	1.25

## References

- 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*, 2018, **28**(5): 1047–1055
- Xiong Xiang-Guang. A zero watermarking scheme with strong robustness in spatial domain. *Acta Automatica Sinica*, 2018, **44**(1): 160–175  
(熊祥光. 空域强鲁棒零水印方案. 自动化学报, 2018, **44**(1): 160–175)
- Asikuzzaman M, Pickering M R. An overview of digital video watermarking. *IEEE Transactions on Circuits and Systems for Video Technology*, 2018, **28**(9): 2131–2153



- 4 Shih F Y. *Digital Watermarking and Steganography: Fundamentals and Techniques*. Boca Raton: CRC Press, 2017.
- 5 Chen S T, Huang H N. Optimization-based audio watermarking with integrated quantization embedding. *Multimedia Tools and Applications*, 2016, **75**(8): 4735–4751
- 6 Jiang Chuan-Xian, Yang Tie-Jun, Dong Ming-Gang, Cheng Xiao-Hui, Li Zhi. A reversible image watermarking algorithm using linear space hiding model. *Acta Automatica Sinica*, 2014, **40**(10): 2324–2333  
(姜传贤, 杨铁军, 董明刚, 程小辉, 李智. 基于线性空间隐藏模型的可逆图像水印算法. 自动化学报, 2014, **40**(10): 2324–2333)
- 7 Hwang M J, Lee J S, Lee M S, Kang H G. SVD-based adaptive QIM watermarking on stereo audio signals. *IEEE Transactions on Multimedia*, 2018, **20**(1): 45–54
- 8 Hu H T, Chang J R, Hsu L Y. Windowed and distortion-compensated vector modulation for blind audio watermarking in DWT domain. *Multimedia Tools and Applications*, 2017, **76**(24): 26723–26743
- 9 Barni M, Bartolini F, DeRosa A, Piva A. Optimum decoding and detection of multiplicative watermarks. *IEEE Transactions on Signal Processing*, 2003, **51**(4): 1118–1123
- 10 Cui Han-Guo, Liu Jian-Xin, Li Zheng-Min. STL model watermarking algorithm based on pyramid technique. *Acta Automatica Sinica*, 2013, **39**(6): 852–806  
(崔汉国, 刘健鑫, 李正民. 基于金字塔技术的STL模型数字水印算法. 自动化学报, 2013, **39**(6): 852–806)
- 11 Etemad S, Amirmazlaghani M. A new multiplicative watermark detector in the contourlet domain using t locationscale distribution. *Pattern Recognition*, 2018, **77**: 99–112
- 12 Amini M, Sadreazami H, Ahmad M O, Swamy M N S. A channel-dependent statistical watermark detector for color images. *IEEE Transactions on Multimedia*, 2019, **21**(1): 65–73
- 13 Hua G, Huang J W, Shi Y Q, Goh J, Thing V L L. Twenty years of digital audio watermarking — a comprehensive review. *Signal Processing*, 2016, **128**: 222–242
- 14 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
- 15 Majoul T, Raouafi F, Jaidane M. An improved scheme of audio watermarking based on turbo codes and channel effect modeling. In: *Proceedings of the 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Prague, Czech Republic: 2011. 353–356
- 16 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  
(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)
- 17 Akhaee M A, Sahraeian S M E. Scaling-based watermarking with universally optimum decoder. *Multimedia Tools and Applications*, 2015, **74**(15): 5995–6018
- 18 Kingsbury N G. The dual-tree complex wavelet transform: A new technique for shift invariance and directional filters. In: *Proceedings of the 8th IEEE Digital Signal Processing Workshop*, Bryce Canyon, USA: 1998. 120–131
- 19 Selesnick I W, Baraniuk R G, Kingsbury N C. The dual-tree complex wavelet transform. *IEEE Signal Processing Magazine*, 2005, **22**(6): 123–151
- 20 Kwitt R, Uhl A. Lightweight probabilistic texture retrieval. *IEEE Transactions on Image Processing*, 2010, **19**(1): 241–253
- 21 Fisher N I, Switzer P. Chi-plots for assessing dependence. *Bio-*

*metrica*, 1985, **72**: 253–265

- 22 Crouse M S, Nowak R D, Baraniuk R G. Wavelet-based statistical signal processing using hidden Markov models. *IEEE Transactions on Signal Processing*, 1998, **46**(4): 886–902
- 23 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**(2): 2372–2384



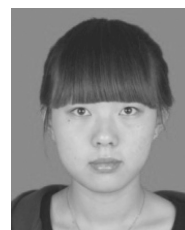
**王向阳** 辽宁师范大学计算机科学与技术学院教授. 主要研究方向为网络信息安全与智能多媒体计算. 本文通信作者. E-mail: wxy37@126.com  
(**WANG Xiang-Yang** Professor at the Multimedia and Information Security Laboratory, School of Computer and Information Technology, Liaoning Normal University. His current research interest covers network information security, and intelligent multimedia computing. Corresponding author of this paper.)



**牛盼盼** 博士, 辽宁师范大学计算机与信息学院副教授. 主要研究方向为网络信息安全与图像处理. E-mail: niupanpan3333@163.com  
(**NIU Pan-Pan** Ph.D., associate professor at the School of Computer and Information Technology, Liaoning Normal University. Her research interest covers network information security and image processing.)



**杨红颖** 辽宁师范大学计算机科学与技术学院教授. 主要研究方向为多媒体信息安全与图像处理. E-mail: yhy65@126.com  
(**YANG Hong-Ying** Professor at the School of Computer and Information Technology, Liaoning Normal University. Her current research interest covers multimedia information security and image processing.)



**李丽** 辽宁师范大学计算机科学与技术学院硕士研究生. 主要研究方向为信息隐藏与数字水印. E-mail: workprocess0003@163.com  
(**LI Li** Master student at the School of Computer and Information Technology, Liaoning Normal University. Her current research interest covers information hiding and digital watermarking.)