基于 EMD 的音频水印算法研究

唐 鑫[1],马兆丰[1,2],钮心忻[1],杨义先[1]
(1.北京邮电大学 信息安全中心,北京 100876; 2.北京国泰信安科技有限公司,北京 100086)

摘 要: 提出了一种基于经验模态分解(Empirical Mode Decomposition,EMD)的音频水印算法,选择 EMD 分解得到的冗余信号分量作为水印的嵌入位置,并证明了冗余信号分量按照提出的算法嵌入水印以后仍然是冗余信号分量,从而为水印嵌入提取提供了理论基础.通过粒子群优化算法求解出适用于 EMD 分解的最优音频水印嵌入强度,按此强度嵌入水印,可以同时满足水印的健壮性和不可感知性.仿真结果表明,使用计算出的最优嵌入强度嵌入水印,嵌入水印后的音频信号在受到大部分攻击的情况下可以确保水印的不可见性和健壮性.

关键词: 经验模态分解; 粒子群优化; 音频数字水印; 嵌入强度
中图分类号: TP309 **文献标识码:** A **文章编号:** 0372-2112 (2015)07-1457-08
电子学报 URL: http://www.ejournal.org.cn **DOI:** 10.3969/j.issn.0372-2112.2015.07.032

Research on Audio Watermarking Algorithm Based on EMD

TANG Xin[1], MA Zhao-feng[1,2], NIU Xin-xin[1], YANG Yi-xian[1]
(1. *Information Security Center, Beijing University of Posts and Telecommunications, Beijing* 100876, *China*;
2. *Beijing National Security Science and Technology Co. Ltd, Beijing* 100876, *China*)

Abstract: An audio watermarking algorithm based on EMD(Empirical Mode Decomposition) was proposed. Residual signal components decomposed from EMD were selected to embed watermark by the proposed algorithm and the watermarked components are proved to be still residual signal components, which provides the theoretical basis for the watermark embedding and extraction. The optimal audio watermark embedding strength in EMD was solved by particle swarm optimization algorithm. The robustness and imperceptibility of watermarking can be satisfied according to this embedding strength. The simulation results show that using the optimal embedding strength to embed watermark can guarantee the imperceptibility and robustness of the watermark in most attack conditions.

Key words: EMD (empirical mode decomposition); particle swarm optimization; audio digital watermarking; embedding strength

1 引言

近年来,随着音频数字版权需求的日益加大,音频数字水印作为音频数字版权的一种重要保护手段,在学术界引起了越来越大的关注.音频水印需要满足三大基本要求,即水印的鲁棒性、不可见性和嵌入容量,也就是说,水印需要在保持健壮性的前提下,以尽可能不被感知的方式嵌入在音频载体中,并且在给定音频载体中嵌入的水印量要尽可能大,在需要证明版权归属的时候能够快速提取嵌入的水印.

为了满足这些要求,近年来学术界提出了许多新型的音频水印算法,它们可以分为时域方法和变换域方法两大类.通常时域方法鲁棒性较差,尤其在面对一些攻击的时候,水印检测的性能不是很理想,具有代表性的时域方法有文献[1,2].变换域方法通常首先将音频载体信号做离散余弦变换(Discrete Cosine Transform,DCT)、离散小波变换(Discrete Wavelet Transform,DWT)等变换,在变换后的信号中完成水印的嵌入[3],这样的方法具有较强的鲁棒性.尤其是基于小波变换的方法[4~6],可以有效地提高水印嵌入容量.

由于在基于小波变换的方式中,小波基函数是固定的,所以它们不能很好地匹配所有的实信号,并且基于小波变换的方式搜索同步码的效率也不高[7].为了解决这些问题,学术界提出了基于 EMD 的水印算法[7~10],EMD 不需要预先选定基函数,并且是一种数据自适应的分解方法[9].文献[7]提出了一种基于 EMD 的音频水

收稿日期:2013-12-04;修回日期:2014-04-18;责任编辑:郭游
基金项目:国家自然科学基金(No.60803157,No.90812001,No.61272519,No.61170271)

印算法,将水印嵌入在分解出的低频固有模态函数(Intrinsic Mode Functions, IMF)分量的极值上,再把所有的IMF分量累加求和得到嵌入水印后的音频信号.然而,文献[7]并没有考虑到EMD的数据自适应性带来的问题,由于EMD分解具有数据自适应性,不同的数据分解出的IMF分量的数量不同,嵌入水印前与嵌入水印后的音频信号做EMD分解,得到的IMF分量的数量和极值的位置也会改变,具体视嵌入强度而定,如果嵌入强度选取不合适,嵌入水印后的音频分解出的IMF分量的数量和极值的位置发生变化,就会导致发生检测错误.文献[8]提出了一种基于EMD的图像水印算法,将水印嵌入在EMD分解出的冗余信号分量上,并通过实验验证了该方法对于高斯噪声和JPEG(Joint Photographic Experts Group)压缩的健壮性.然而,文献[8]对嵌入位置的选择缺乏理论依据,并且文献[7,8]都采用人为给定的方式确定水印的嵌入强度,没有去考虑水印嵌入强度与水印的不可见性和鲁棒性之间的关系.如果水印的嵌入强度较大,则水印能够取得较强的鲁棒性,但是其不可见性会受到影响;而如果嵌入强度过小,则水印的鲁棒性不够,在网络传输和一些攻击手段的干扰下无法保证水印的完整性和水印检测的准确率,所以选择合适的嵌入强度是一个很重要的问题.

针对上述问题,本文提出了一种基于EMD的智能音频水印算法,从理论上分析了如何选择水印的合适嵌入位置,并且同时考虑水印的健壮性与不可感知性,采用粒子群优化算法求得能同时满足水印健壮性与不可感知性要求的最佳嵌入强度值,通过实验验证了本文所提算法对于加性高斯白噪声、滤波、重采样、MP3(Moving Picture Experts Group Audio Layer Ⅲ)压缩和重量化等常见音频信号攻击情况下具有较强的鲁棒性.

2 EMD分解理论基础

EMD是一种以完全自适应的方式分析线性或非线性系统中非静态信号的工具[9].在介绍EMD之前,首先给出固有模态函数IMF的定义.

一个分量称为IMF分量,当且仅当其同时满足以下两个条件[8,9]:

(1)在整个分量里,极值点的数量和过零点的数量必须相等或者至多相差1;

(2)在任意一点,由局部极大值和极小值经过三次样条插值得到的上包络线和下包络线的均值为0.

给定一个信号$x(t)$,对其进行EMD分解的执行过程可以概括为如下步骤[9]:

Step 1 找到信号$x(t)$的所有局部极大值点和局部极小值点,分别采用三次样条插值曲线连接起来,得到上、下包络线;

Step 2 对于每一个时间点t,计算上下包络的均值$m(t)$,并从原始信号中逐点减去均值,得到$h(t)=x(t)-m(t)$;

Step 3 根据IMF的定义判断$h(t)$是否是一个IMF,如果不是,则将$h(t)$作为一个待分解的信号,重复以上步骤,直到分解出的信号是一个IMF;

Step 4 用原始信号减去分解出的IMF,得到$r(t)=x(t)-\text{IMF}$.重复以上过程,对$r(t)$进行分解,直到分解不出新的IMF时,终止分解过程.

综合以上过程,给定一个信号$x(t)$,通过EMD分解可以得到一系列固有模态函数IMF和冗余信号$r_C(t)$之和,如式(1)所示:

$$x(t)=\sum_{i=1}^{C}\text{IMF}_i(t)+r_C(t) \tag{1}$$

在式(1)中,这C个IMF具有零均值,并且彼此之间近似正交.

3 基于EMD的音频水印算法

本文提出的基于EMD的音频水印算法首先对原始音频信号做EMD分解,考虑到水印嵌入的健壮性要求,将水印信号嵌入在冗余信号分量上.根据前文的分析可知,水印的嵌入强度直接影响到水印的不可见性和鲁棒性,对于音频水印来说,需要在保持健壮性的前提下尽可能少的影响音质,即保证水印的不可感知性.水印的不可感知性可以通过人耳的试听测试来评分[10],为了避免评分的主观倾向性,本文采用峰值信噪比(PSNR, Peak Signal to Noise Ratio)[11,12]作为不可感知性的量化指标.由于水印的嵌入强度是影响音质的主要因素,所以嵌入强度与PSNR值之间存在联系.本文使用粒子群优化算法求解最优的水印嵌入强度,采用PSNR与水印检测准确度加权求和的形式作为粒子群算法的适应度函数,使得在求解出的强度下嵌入水印,能够同时确保水印的健壮性与不可感知性.

3.1 水印嵌入

在嵌入水印之前,我们首先确定水印的嵌入位置,本文考虑将水印嵌入在EMD分解得到的冗余信号分量上,主要是基于以下两点原因:

(1)根据上文的分析可知,EMD是从原始信号中分解出有限个IMF分量,剩下的冗余信号分量不满足IMF的两个条件,将水印嵌入在冗余信号分量上,只要保证嵌入水印后的冗余信号分量仍然不满足IMF的两个条件,则合成以后再进行EMD分解,不会从嵌入水印后的冗余信号分量中产生新的IMF分量,水印信息仍然保留在冗余信号分量中.

(2)文献[8]通过实验验证了将水印嵌入在图像载体的冗余信号分量中,对于高斯噪声和JPEG压缩的健

壮性.文献[10]将水印嵌入在音频载体的冗余信号分量上,验证了水印对于加性高斯白噪声和 MP3 压缩的健壮性.实验结果证明了将水印嵌入在冗余信号分量上可以具有较强的鲁棒性.

在嵌入水印之前,首先我们需要对音频文件进行分帧处理,对每一帧音频信号分别做 EMD 分解,取每次分解得到的冗余信号分量作为水印的嵌入分量,每一帧嵌入 1 比特的水印信号.本文考虑嵌入水印后音频信号的 PSNR 值以及检测水印的准确度 AOW(Accuracy of watermark Detection),将它们加权求和的形式作为适应度函数,采用粒子群算法选择合适的嵌入强度.

本文嵌入水印的思想是把每一帧的冗余信号分量替换掉,当水印序列为空的情况下,冗余信号分量不发生变化;当水印序列非空的情况下,嵌入公式如式(2)所示.

$$r_C(t)^* = \begin{cases} (r_C(t) - \min_t(r_C(t))) \cdot S, & \text{若 } w(i) = 1 \\ (r_C(t) - \max_t(r_C(t))) \cdot S, & \text{若 } w(i) = 0 \end{cases} \tag{2}$$

在式(2)中,水印的嵌入是对于每一个音频帧而言的,$r_C(t)$表示嵌入水印前的冗余信号分量,$r_C(t)^*$表示嵌入水印后的冗余信号分量.由 EMD 分解的定义可知,$r_C(t)$并不是一个 IMF 分量,为了确保水印提取成功,下面我们给出命题及证明过程.

命题　嵌入水印后得到的 $r_C(t)^*$ 仍然是一个冗余信号分量.

证明　按照式(2)的方法嵌入水印,如果嵌入的水印比特是 1,则找出原音频帧冗余信号分量的最小值 $\min_t(r_C(t))$,由于 $r_C(t)$上每一个时间点 t 对应的值都大于等于$\min_t(r_C(t))$,所以 $r_C(t) - \min_t(r_C(t))$得到的值一定大于等于 0,乘以正的嵌入强度 S,得到的 $r_C(t)^*$上每一个时间点 t 对应的值都是非负的.按照前文所述的 IMF 需要满足的条件,分别由局部极大值和局部极小值三次样条插值得到的上下包络线在每一个时间点 t 的均值需要为零,然而,由于 $r_C(t)^*$中不存在负值,所以该条件不能满足,因此嵌入 1 时,$r_C(t)^*$仍然是一个冗余信号分量.

同理,在嵌入水印比特是 0 的时候,得到的 $r_C(t)^*$在每一个时间点 t 对应的值都非正,同样不满足 IMF 的条件,因此嵌入 0 时,$r_C(t)^*$也是一个冗余信号分量.

综上可知,嵌入水印后的 $r_C(t)^*$是一个冗余信号分量.

证毕

如前文所述,本文需要通过选定合适的嵌入强度 S,在水印的健壮性与不可感知性之间达到平衡.在本文提出的方法中,我们将水印信号的健壮性用检测出的水印的准确度来表征,不可感知性通过嵌入水印后的音频信号的 PSNR 值来表示,即水印检测的准确度越高,则音频水印越健壮,嵌入水印后的音频信号的 PSNR 值越大,则音频水印的不可感知性就越好.

假定音频帧的长度为 L,用于嵌入水印的音频信号包含 N 帧,水印信号的长度是 $NW(NW \leqslant L)$,PSNR[11]的定义如下:

$$\mathrm{PSNR} = 10 \cdot \log_{10} \frac{L \cdot N \cdot \max\limits_{1 \leqslant i \leqslant N, 1 \leqslant j \leqslant L} (x(i,j))^2}{\sum\limits_{i=1}^{N} \sum\limits_{j=1}^{L} (x(i,j) - y(i,j))^2} \tag{3}$$

在式(3)PSNR 的定义中,$x(i,j)$与 $y(i,j)$,$1 \leqslant i \leqslant N$,$1 \leqslant j \leqslant L$ 分别表示嵌入前和嵌入后的音频信号对应采样点,如式(3)可见,若音频信号嵌入水印后改变相对较小,则 PSNR 值就会较大,此时水印的不可感知性较好.

检测水印的准确度 AOW 定义为检测出的水印与原水印信号相同的位数除以水印信号的长度,给出其数学表达式如下:

$$\mathrm{AOW} = \frac{NW - \sum\limits_{i=1}^{NW} | w'(i) - w(i) |}{NW} \tag{4}$$

为了同时考虑水印的健壮性和不可感知性与嵌入强度之间的关系,我们将粒子群算法的适应度函数定义如下:

$$fitness = r \cdot \mathrm{PSNR} + \mathrm{AOW} \tag{5}$$

在式(5)中,r 表示一个缩放因子,由于 AOW∈[0,1],而 PSNR 的取值显然大于 1,二者直接相加,则适应度函数的取值几乎不受 AOW 的影响,无法准确评判水印的健壮性.所以需要选一个合适的 r,来平衡 PSNR 与 AOW 的关系.由于 PSNR 的取值越大说明水印的不可见性越好,AOW 的取值越大说明水印的健壮性越好,所以,适应度的值越大越好.

根据粒子群算法的速度更新公式,定义对一个粒子来说水印嵌入强度 S 变化速度的更新公式如下[12]:

$$\begin{aligned} v^{(t+1)} = {} & w \cdot v^{(t)} + c_1 \cdot rand_1 \cdot (p^{(t)} - S^{(t)}) \\ & + c_2 \cdot rand_2 \cdot (p_g^{(t)} - S^{(t)}) \end{aligned} \tag{6}$$

在式(6)中,$v^{(t)}$和 $v^{(t+1)}$分别表示在第 t 次迭代和第 $t+1$ 次迭代中水印嵌入强度的更新速度,w 表示惯性权重,表示速度更新的过程中保持原来速度的系数,$p^{(t)}$和 $p_g^{(t)}$ 分别表示粒子自己的历史最优强度和全局最优强度,c_1 和 c_2 分别表示粒子跟踪自己历史最优强度和跟踪全局最优强度的权重系数,$rand_1$ 和 $rand_2$ 是在[0,1]区间均匀分布的随机数.定义水印嵌入强度的更新公式如下:

$$S^{(t+1)} = S^{(t)} + \alpha \cdot v^{(t+1)} \tag{7}$$

式(7)中的 α 表示速度的约束因子,用来控制水印

嵌入强度每一步更新的幅度,嵌入强度的初始值用 $S^{(0)}$ 来表示.在本文的算法中,通过设置最大迭代次数 MaxIter 来控制水印嵌入强度迭代更新的停止,即一旦算法的迭代次数达到了 MaxIter,则迭代停止,此时的嵌入强度即为最优的水印嵌入强度.

3.2 水印提取

在水印的嵌入过程中,我们对音频帧做 EMD 变换以后仅改变了冗余信号分量的值,并且我们已经证明,嵌入水印以后的 $r_C(t)^*$ 仍然是一个冗余信号分量,即嵌入水印以后不会产生新的 IMF 分量.因此在水印提取的过程中,我们对嵌入水印以后的音频信号做 EMD 变换,然后找出冗余信号分量.根据式(2)嵌入公式及上文分析可知,嵌入 1 时,得到的 $r_C(t)^*$ 中每一个时间点 t 对应的值都是非负的,嵌入 0 时,每一个时间点对应的值都是非正的,所以,参考文献[8],我们据此提出水印的盲检测公式如式(8)所示:

$$w'(i)=\begin{cases}1, & \text{若 } \mathrm{sum}_t(r_C(t)^*)>0\\ 0, & \text{若 } \mathrm{sum}_t(r_C(t)^*)<0\end{cases} \tag{8}$$

如式(8)所示,在水印提取时我们评判的标准是将不同时间点 t 的 $r_C(t)^*$ 值累加起来,若大于 0,则认为提取出的水印是 1;若小于 0,则认为提取出的水印是 0.

4 仿真实验

4.1 仿真环境及实验方法

在这个部分,我们实现了本文提出的水印算法,并进行了一些仿真实验,来对算法进行检测,验证本文算法求解得到的最优水印嵌入强度的效果.我们选取长度为 20s 的单声道音频信号作为测试集,测试音频为 44.1kHz 采样,16bit 量化.仿真实验的参数配置如表 1 所示.

表 1 仿真实验参数配置

参数符号说明	参数取值
分帧长度	64
缩放因子 r	0.0167
惯性权重 w	0.7298
粒子跟踪自己历史最优强度的权重系数 c_1	1.4962
粒子跟踪全局最优强度的权重系数 c_2	1.4962
约束因子 α	0.1
最大迭代次数 MaxIter	50

仿真实验分为四个部分,实验一从两个方面验证嵌入水印后的 $r_C(t)^*$ 是否仍然是冗余信号分量.首先比较嵌入水印前后的 $r_C(t)$ 和 $r_C(t)^*$,观察 $r_C(t)^*$ 是否符合冗余信号分量的条件,如果是,则验证了 $r_C(t)^*$ 是冗余信号分量.接下来考虑将嵌入水印后的 $r_C(t)^*$ 与原 IMF 分量合成为含水印的音频帧,再对该音频帧重新进行分解,得到新的 IMF 分量以及冗余信号分量,比较嵌入水印前后的分解结果,验证嵌入水印以后是否对 EMD 分解的结果产生了影响;实验二计算不同嵌入强度下音频信号的 PSNR 值和水印的检测准确度 AOW,比较适应度的值,评价本文提出方法的有效性;实验三分析了按照本文提出的方法计算出的水印嵌入强度嵌入水印,嵌入水印后的音频信号受到不同攻击后,水印的不可见性和健壮性;实验四将本文算法的效果和文献[7,13]做了比较,分析评价文献[7,13]的方法在固定条件下与本文方法的性能比较.

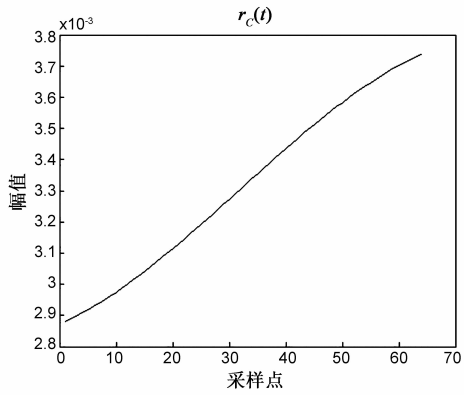
4.2 冗余信号分量验证

在这部分实验中,我们首先验证原冗余信号分量 $r_C(t)$ 在嵌入水印后得到的 $r_C(t)^*$ 是否仍然是冗余信号分量.根据上文的证明可知,验证 $r_C(t)^*$ 是冗余信号分量,即验证其不满足 IMF 的条件,根据式(2)所示的嵌入公式及上文的分析,如果嵌入水印后得到的 $r_C(t)^*$ 在所有时间点 t 的取值都是非正或非负的,则认为 $r_C(t)^*$ 仍然是冗余信号分量.因此,我们随机选取一帧音频信号,分别考虑嵌入 0 和嵌入 1 两种情况,嵌入水印前的 $r_C(t)$ 图像和嵌入 0、嵌入 1 后的 $r_C(t)^*$ 图像分别如图 1(a),(b),(c)所示.

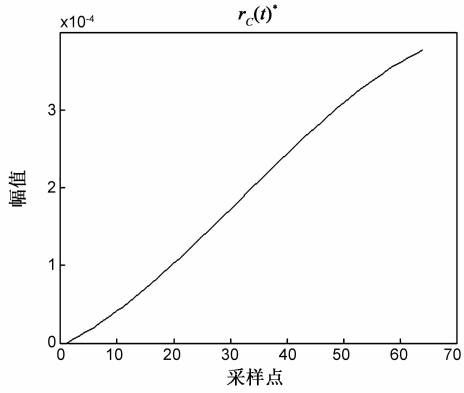
如图 1 所示,嵌入水印比特 1 以后,$r_C(t)^*$ 曲线对应的幅值都大于 0,因此,将所有采样点对应的幅值累加起来,得到的值大于 0,符合上文的分析结果;同理,观察嵌入水印比特 0 得到的 $r_C(t)^*$ 曲线,对应的幅值都小于 0,所有的幅值累加求和,得到的值小于 0,符合上文分析结果,因此嵌入水印后得到的 $r_C(t)^*$ 是冗余信号分量.

接下来的实验将嵌入水印后的 $r_C(t)^*$ 与原 IMF 分量合成为含水印的音频帧,对其重新进行 EMD 分解.这部分实验中,我们考虑对同一个音频帧在未嵌入水印、嵌入水印比特 1 和嵌入水印比特 0 三种情况下进行 EMD 分解的情况,三种情况的分解结果分别如图 2(a),(b),(c)所示.

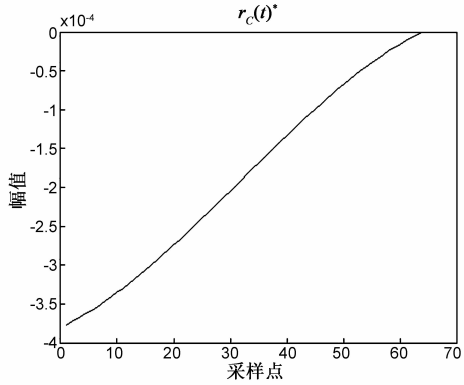
由图 2 可以看出,对于同一个音频帧而言,未嵌入水印时与嵌入水印比特 1、嵌入水印比特 0 两种情况下得到的 IMF 分量的数量是相同的,并且这些分解出的 IMF 分量的形态几乎完全一致,所不同的就是三种情况下分解出的冗余信号分量,而后两种情况下得到的冗余信号分量性质与前一部分实验得到的结果是一致的.这部分实验验证了水印信息并不会改变分解出的 IMF 分量,也就是说嵌入在冗余信号分量的水印信息,在含水印音频信号 EMD 分解后,仍然存在于冗余信号分量中.



(*a*) 嵌入水印前的冗余信号分量$r_C(t)$曲线图

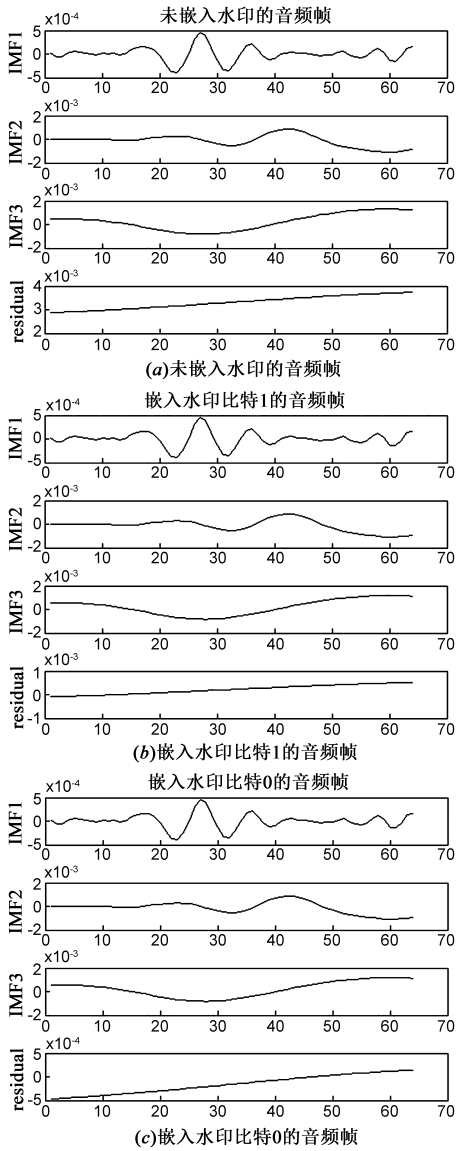


(*b*) 嵌入水印比特1后的$r_C(t)^*$曲线图



(*c*) 嵌入水印比特0后的$r_C(t)^*$曲线图

图1 嵌入水印前后的冗余信号分量曲线图



(*a*)未嵌入水印的音频帧

(*b*)嵌入水印比特1的音频帧

(*c*)嵌入水印比特0的音频帧

图2 嵌入水印前后的音频帧EMD分解比较图

4.3 适应度评价

在这部分实验中,我们首先考虑水印的嵌入强度与 PSNR 取值的关系,以及强度与水印检测准确度 AOW 的关系,给出各自的曲线图,分别如图 3,图 4 所示;接下来的实验记录了全局最优水印强度值以及适应度值随着迭代次数变化的情况,曲线图分别如图 5,图 6 所示.

如图 3 所示,随着嵌入强度取值在[0.1,1]区间上变化,PSNR 取值首先缓慢变大,接下来以较快速度变小.由于按照式(2)嵌入水印,我们仅改变音频信号中的冗余信号分量,所以按照式(3)计算PSNR,嵌入水印

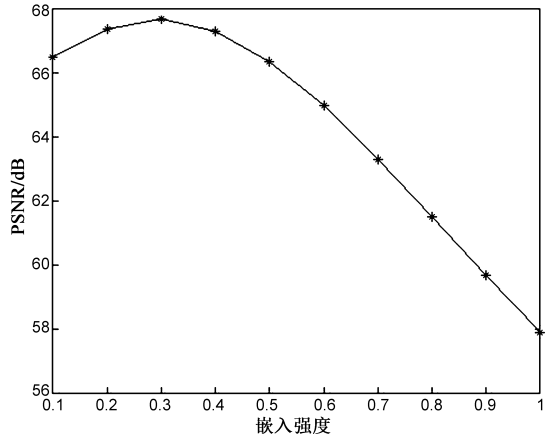


图3 语音信号PSNR值与嵌入强度的关系

前后的采样点 $x(i,j)$与 $y(i,j)$,$1\leqslant i\leqslant N$,$1\leqslant j\leqslant L$,它们的差别就在于冗余信号分量部分.根据嵌入强度的不同,$\sum_{i=1}^{N}\sum_{j=1}^{L}(x(i,j)-y(i,j))^2$ 的取值也会发生变化,总体的趋势是随着嵌入强度变大,$x(i,j)$与 $y(i,j)$差异变大,$\sum_{i=1}^{N}\sum_{j=1}^{L}(x(i,j)-y(i,j))^2$ 取值变大,所以PSNR 呈下降趋势.

在图 4 中可以看到,检测水印的准确度随着嵌入强度的增大逐渐提高.由图 3,图 4 可知,PSNR 与 AOW 随嵌入强度变化的趋势是相反的,根据式(5)定义的适应度函数,使用粒子群算法求解出最优的水印嵌入强度.

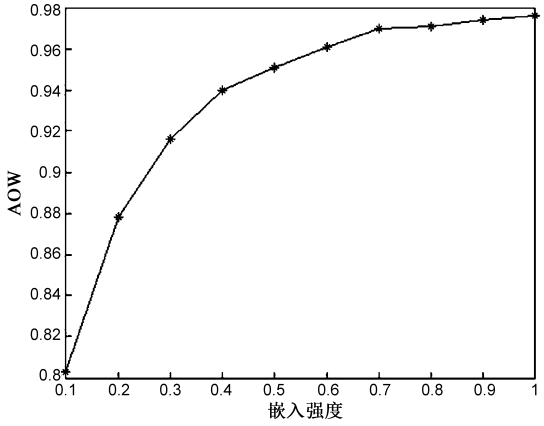


图4 水印检测准确度AOW与嵌入强度的关系

我们首先将音频信号作为水印嵌入的载体,设定初始的嵌入强度 $S^{(0)}$为一个随机的正值,然后粒子群算法开始迭代,我们记录下每一次迭代过程中的全局最优水印嵌入强度值,如图 5 所示.

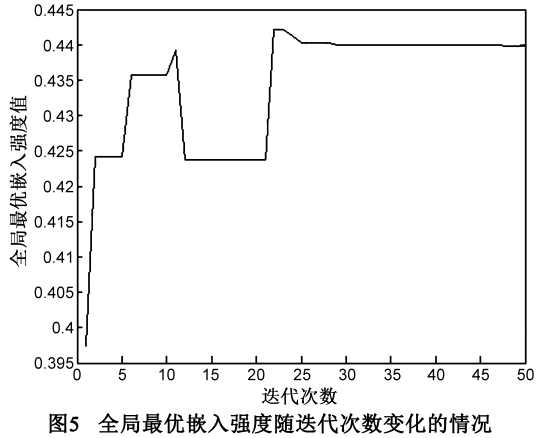


图5 全局最优嵌入强度随迭代次数变化的情况

在图 5 中,算法的整个迭代过程里,全局最优嵌入强度值大小会有起伏,迭代达到三十步以后,嵌入强度值的大小变化不大,基本趋于稳定.在第 12 步到 22 步迭代过程中,全局最优嵌入强度值出现下降,然而总体保持增大趋势.

图 6 的曲线描述了适应度值随迭代次数变化的情况,由于每一步迭代只记录全局最优嵌入强度下的适应度值,所以适应度值保持增加.从图中可以看出,从第 30 步迭代开始,适应度值增加缓慢,基本保持平稳.

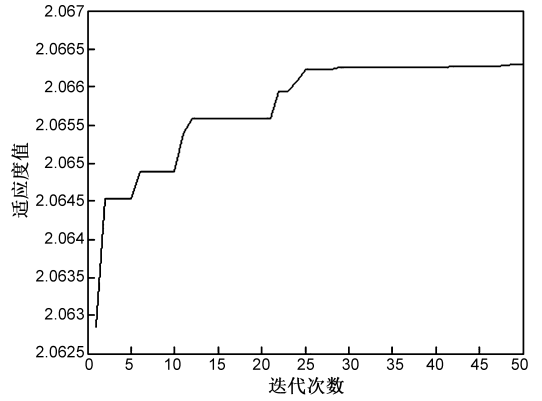


图6 适应度值随迭代次数变化的情况

4.4 抗攻击条件下水印的不可见性和健壮性比较

这部分的实验取一段长度为 20s 的音频信号片段,随机生成长度为 1000 比特的水印信号,根据式(2)嵌入水印,并且采用本文所提算法计算出最优水印嵌入强度,按此强度嵌入水印,考虑嵌入水印的音频信号受到攻击干扰的情况,分别考虑了如表 2 所示的 18 种攻击手段,对于受到攻击干扰后的含水印音频信号,计算它们的 PSNR 值和水印检测准确率 AOW,得到的结果用以验证水印的不可见性和健壮性.结果如表 2 所示.

根据表 2 的数据可以看出,采用本文所提算法计算出最优嵌入强度,并在冗余信号分量上嵌入水印,对于 compressor、dynnoise、exchange、extrastereo_ 70、flippsample、lsbzero、rc_lowpass、smooth、stat1 等攻击而言,由于它们主要是修改局部样本点,诸如交换样本点、最低比特位清零、滤波等操作,并不会在根本上改变 EMD 分解的结果,因此在这些攻击下,PSNR 值能达到 60dB 以上,水印检测准确度 AOW 能达到 0.95 以上,水印的不可见性和健壮性都能得到保证;normalize 攻击对于采样点的幅值做了较大的等比例变化,因此 PSNR 值较低,然而由于 EMD 分解每一次迭代都是减去上下包络的均值,所以幅值的等比例变化不会改变 EMD 分解的结果,水印信息仍然保存在冗余信号分量里,所以在 normalize 攻击下,AOW 值能够达到 0.981;在 fft_hlpass、rc_highpass 这些攻击情况下,对原始音频进行了高滤波.由于本文算法将水印嵌入在音频信号的低频部分,所以水印信息丢失较多,因此 AOW 值不高;cutsamples 和 zeroremove 下,去除了一些样本点,因此出现同步错误,造成 AOW 较低,fft_invert 对样本点的幅值方向做了翻转,所以造

成 PSNR 和 AOW 值都较小.总的看来,本文的方法在大多数攻击条件下能够同时较好地确保水印的不可感知性和健壮性.

表 2　抗攻击条件下水印的不可见性和健壮性比较

攻击类型	PSNR(dB)	AOW	攻击类型	PSNR(dB)	AOW
addbrumm_10100	7.564	0.512	fft_invert	29.529	0.021
addnoise_900	56.756	0.601	flippsample	64.350	0.958
compressor	66.813	0.982	lsbzero	66.814	0.982
cutsamples	36.094	0.509	normalize	27.848	0.981
dynnoise	65.676	0.968	rc_highpass	55.547	0.747
echo	50.182	0.786	rc_lowpass	66.598	0.978
exchange	65.131	0.951	smooth	66.835	0.961
extrastereo_70	66.812	0.982	stat1	66.444	0.967
fft_hlpass	56.468	0.705	zeroremove	36.677	0.505

4.5　水印性能比较

这部分实验从嵌入水印后音频信号的 PSNR、水印检测准确率 AOW 以及在 128Kb/s、64Kb/s 的 MP3 压缩下的检测准确率等几个方面比较了本文的水印算法与文献[7,13]的水印算法的性能,我们取同一段音频信号作为水印的嵌入载体,比较嵌入水印后的音频信号在达到相同 PSNR 的情况下的检测准确率情况,具体比较结果如表 3 所示.

表 3　水印性能比较

算法	PSNR(dB)	AOW	MP3(128Kb/s)	Mp3(64Kb/s)
本文算法	66.81	0.983	0.979	0.960
文献[7]	66.03	0.870	0.854	0.832
文献[12]	66.21	0.717	0.690	0.674

由于本文所提算法综合考虑水印的不可见性和健壮性,如表 3 所示,采用本文算法计算得出,PSNR 值为 66.81dB 时,AOW 值能达到 0.983,我们比较文献[7,13]的算法,通过调整合适的嵌入强度,使得它们各自嵌入水印后的 PSNR 值达到 66dB 左右,可以看到,在此情况下,两者的水印检测准确率都低于本文得到的结果,并且在 128Kb/s 和 64Kb/s 的 MP3 压缩下,本文的方法检测性能更优.文献[7]基于 EMD 分解,将水印嵌入在低频 IMF 分量的极值上,然而由于缺乏理论依据,无法确保嵌入水印后的音频信号进行 EMD 分解,水印信息依然在低频 IMF 分量上,并且算法没有考虑水印的不可见性,因此相同 PSNR 条件下,检测准确率低于本文方法;文献[12]考虑了水印的不可见性和健壮性,将水印嵌入在 DWT 域,实验结果也表明了,基于 EMD 分解的嵌入方法相比传统方法,具有更好的性能.

5　结束语

本文提出了一种基于 EMD 的音频水印算法,选择冗余信号分量作为水印的嵌入位置,并证明了嵌入水印后的 $r_C(t)^*$ 仍然是冗余信号分量,从而为水印嵌入提取提供了理论基础.本文的方法考虑了水印的健壮性和不可感知性,求解出同时满足这两个要求的最优嵌入强度.仿真实验结果表明,按照此强度嵌入水印,在含水印的音频信号受到大多数攻击干扰的情况下,能够保证水印的健壮性和不可感知性,并且本文的算法相比传统算法,具有更好的性能.在下一步的工作中,我们将考虑水印信号的同步问题.

参考文献

[1] Baghel R, Sharma P. Increasing robustness of LSB audio steganography using a novel embedding method[A]. National Conference on Security Issues in Network Technologies (NCSI-2012)[C], Gwalior: IEEE Press, 2012.1 – 5.

[2] 杨榆,雷敏,钮心忻,等.基于回声隐藏的 VDSC 隐写分析算法[J].通信学报,2009,30(2):83 – 88.
Yang Y, Lei M, Niu X X, et al. VDSC steganalysis algorithm of echo hiding[J]. Journal on Communications, 2009, 30(2): 83 – 88. (in Chinese)

[3] 张金全,王宏霞.DCT 和 DWT 域音频水印幅值变化规律分析[J].电子学报,2013,41(6):1193 – 1197.
Zhang J Q, Wang H X. Analysis on law of distortion of audio signal for embedding watermark in DCT and DWT[J]. Acta Electronica Sinica, 2013, 41(6): 1193 – 1197. (in Chinese)

[4] Chen S T, Wu G D, Huang H N. Wavelet-domain audio watermarking scheme usingoptimization-based quantisation[J]. IEEE Transactions on Signal Processing, 2013, 4(6): 720 – 727.

[5] Wu S, Huang J, Huang D, et al. Efficiently self-synchronized audio watermarking for assured audio data transmission[J]. IEEE Transactions on Broadcasting, 2005, 51(1): 69 – 76.

[6] Bhat K V, Sengupta I, Das A. An adaptive audio watermarking based on the singular value decomposition in the wavelet domain[J]. Digital Signal Processing, 2010, 20(6): 1547 – 1558.

[7] Khaldi K, Boudraa A. Audio watermarking via EMD[J]. IEEE Transactions on Audio, Speech, and Language Processing, 2013, 21(3): 675 – 680.

[8] Bi N, Sun Q, Huang D, et al. Robust image watermarking based on multiband wavelets and empirical mode decomposition[J]. IEEE Transactions on Image Processing, 2007, 16(8): 1956 – 1966.

[9] Huang N E, Shen Z, Long S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis[J]. Proceedings of the Royal Society of London (Series A: Mathematical, Physical and Engineering Sciences), 1998, 454(1971): 903 – 995.

[10] Wang L, Emmanuel S, Kankanhalli M S. EMD and psychoacoustic model based watermarking for audio[A]. 2010 IEEE International Conference on Multimedia and Expo (ICME)[C]. Suntec City: IEEE Press, 2010. 1427 – 1432.

[11] Wang Y R, Lin W H, Yang L. An intelligent watermarking method based on particle swarm optimization[J]. Expert Systems with Applications, 2011, 38(7): 8024 – 8029.
[12] 杨维,李岐强.粒子群优化算法综述[J].中国工程科学,2004,6(5):87 – 94.
[13] Peng H, Zhang Z, Wang J, et al. Audio watermarking framework using multi-objective particle swarm optimization[J]. International Journal of Innovative Computing, Information and Control, 2013, 9(7): 2789 – 2800.

作者简介

唐 鑫 男,1987 年出生于江苏南京,北京邮电大学博士生,主要研究方向为数字版权管理、数字内容安全、数字水印等.
E-mail: etangxin@gmail.com

钮心忻 女,1963 年出生于浙江湖州,北京邮电大学教授、博士生导师,主要研究方向为数字水印、信息隐藏、隐写分析.

马兆丰 男,1974 年出生于甘肃镇原,博士,北京邮电大学讲师,主要研究方向为数字版权管理、数字内容安全、计算机网络安全.

杨义先 男,1961 年出生于四川盐亭,北京邮电大学教授、博士生导师,主要研究方向为密码学、计算机网络与信息安全.