



基于采样值排序的音频可逆隐写算法

余 恒,王让定,严迪群,张雪垣

(宁波大学 信息科学与工程学院,浙江 宁波 315211)

摘 要:针对已有音频可逆隐写算法失真较大的问题,提出一种基于采样值排序的改进音频可逆信息隐写算法。将音频采样值序列划分为固定大小的采样块,对采样块内部的采样值进行升序排列。计算每个采样块的复杂度,根据复杂度确定采样块是否能够用于嵌密。对于能够嵌密的块,通过其复杂度等级确定最优预测采样值,将该采样值与其他采样值相减得到预测误差,依据预测误差值的大小决定执行嵌密操作或移位操作。在EBU-SQAM标准测试集上的实验结果表明,在相同嵌入容量的条件下,相较于DE算法和PEE算法,该算法的SNR值较高,具有高保真的特性。

关键词:采样值排序;隐写算法;音频;复杂度等级;嵌密操作

开放科学(资源服务)标志码(OSID):



中文引用格式:余恒,王让定,严迪群,等.基于采样值排序的音频可逆隐写算法[J].计算机工程,2021,47(1):123-128,138.

英文引用格式:YU Heng, WANG Rangding, YAN Diqun, et al. Reversible audio steganography algorithm based on sample value ordering[J]. Computer Engineering, 2021, 47(1): 123-128, 138.

Reversible Audio Steganography Algorithm Based on Sample Value Ordering

YU Heng, WANG Rangding, YAN Diqun, ZHANG Xueyuan

(Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo, Zhejiang 315211, China)

[Abstract] To address the large distortion suffered by the existing reversible audio steganography algorithms, this paper proposes an improved reversible audio steganography algorithm based on Sample Value Ordering (SVO). In the algorithm, the complexity of each sample block is calculated, which determines whether the sample block can be used for embedding. As for the to-be-embedded blocks, the optimal predicted sample value is determined by its complexity level, and the prediction error is obtained from the optimal predicted sample value minus other sample values. Based on the prediction error value, perform the embedding or shift operation. The experimental results on the EBU-SQAM standard test set show that the Signal-to-Noise Ratio (SNR) value of the proposed algorithm is significantly improved compared with DE, PEE and other algorithms under the same embedding capacity. It also has higher fidelity.

[Key words] Sample Value Ordering (SVO); steganography algorithm; audio; complexity level; embedding operation

DOI: 10.19678/j.issn.1000-3428.0056619

0 概述

信息隐藏是指在人的感官无法察觉的情况下,将秘密信息嵌入到各种数字媒体中,如图像、音频和视频等。在嵌密结束后,传统信息隐藏载体会产生不可逆的失真,这被称为不可逆信息隐藏,在一些特殊场合,这种失真是不被允许的^[1]。为了解决该问题,研究人员提出了可逆信息隐藏技术^[2],其特点是

在秘密信息被提取完毕后,原始载体能够被完整恢复并且没有任何失真,这种特性使得该技术在一些如图像处理、对抗样本等^[3]重要领域有着广阔的应用前景。

2003年,TIAN等人^[4]提出一种基于差值扩展(Difference Expansion, DE)的可逆信息隐藏方法,其首先对连续的2个像素值做差,然后将差值转化成二进制形式,最后将密信比特追加到二进制差值的

基金项目:国家自然科学基金(U1736215, 61672302, 61901237);浙江省自然科学基金(LY17F020010, LY20F020010);浙江省移动网应用技术重点实验室开放基金(F2018001)。

作者简介:余 恒(1995—),男,硕士研究生,主研方向为信息安全、音频信息隐藏技术;王让定,教授、博士;严迪群,副教授、博士;张雪垣,硕士研究生。

收稿日期:2019-11-18 **修回日期:**2019-12-28 **E-mail:** 1811082201@nbu.edu.cn

最低位上。该方法操作简单,但是相邻像素之间的差值在扩展之后失真较大。THODI等人^[5]在DE的基础上,提出利用预测误差来替代相邻像素的差值进行扩展嵌密的方法,实验结果表明,预测误差扩展(Prediction Error Expansion, PEE)相较DE具有更低的失真,因此,PEE引起了研究人员的广泛关注^[6-7]。

上述方法的本质都是借助差值、预测误差等进行扩展嵌密,对载体的修改较大,在解决失真问题上的效果较直方图移位(Histograms Shift, HS)方法差。TSAI^[8]将预测误差与HS相结合,在增加一部分嵌密容量的前提下有效降低了失真。OU等人^[9]又在TSAI的基础上利用多个直方图修改(Multiple Histograms Modification, MHM)来进行嵌密,弥补了TSAI在嵌密容量上的不足。2013年,LI等人^[10]提出一种像素值排序(Pixel Value Ordering, PVO)的可逆隐写框架,其将一副图像分成若干个固定大小的像素块,然后对像素块中的像素值进行排序,通过第二大(小)值来预测最大(小)值,最后将密信嵌入到预测误差值为“1”或“-1”的像素值中。由于该框架只对载体进行了少量修改,因此隐写后的载体具有高保真的特点。PENG等人^[11]在PVO的基础上提出改进的像素值排序(IPVO),其增加预测误差“0”作为嵌密条件,有效提高了嵌密容量。WENG等人^[12]对像素块内的预测机制进行研究,增加了像素块内生成的预测误差个数,提高了块内像素的利用率。

目前,关于可逆隐写的研究工作大多集中在图像领域^[13-15],而较少关注以音频为载体的可逆隐写算法。严迪群等人^[16]将DE框架应用在音频载体上,证实了DE框架的通用性。WANG^[17]结合音频的特点,提出一种基于改进的预测误差扩展和直方图移位的新可逆音频水印方案,与严迪群等人方法相比,其进一步提高了嵌密容量与信噪比。XIANG^[18]将音频时域采样值按照位置的奇、偶分成2个集合,提出一种复杂的非因果预测算法来计算预测误差,然后将密信以扩展的方式嵌入到预测误差中,该方法的预测性能是目前较优的。文献^[19]在XIANG算法的基础上对容量控制进行优化,在高嵌入率的情况下其性能得到一定提升。

当前关于音频可逆隐写的研究大都围绕PEE^[20-21]展开,限制这类算法性能提升的主要因素包括预测精度和嵌密机制2个。尽管XIANG将预测精度提升到了一个新高度,但是其仍然利用预测误差的二进制扩展进行嵌密,这种方法对载体的修改幅度较大,难以获得高保真的嵌密效果。

本文提出一种针对音频载体的PVO可逆隐写算法。将音频时域序列分成若干个大小相同的采样块,利用音频时域相关的特点对每个采样块的复杂度进行评估,根据复杂度来确定采样块是否能够进行嵌密。此外,由于采样块的大小直接影响嵌密效率,本文进一步利用采样块的复杂度来确定最优预测采样值,以此提高采样块的嵌密效率。

1 基于IPVO的可逆隐写算法

PENG等人^[11]提出的基于IPVO的可逆隐写算法,将载体图像划分为若干个大小为 n 的非重叠像素块,并将这些像素块内的像素值按照升序进行排列,得到一个有序集合 $\{x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(n)}\}$ 。其中, $x_{\sigma(1)} \leq x_{\sigma(2)} \leq \dots \leq x_{\sigma(n)}$, $x_{\sigma(i)}$ 表示有序集合中的第 i 个像素值, $\sigma(i)$ 表示 $x_{\sigma(i)}$ 在原始像素块中的位置。

将像素值 $x_{\sigma(2)}$ 与 $x_{\sigma(n-1)}$ 作为预测像素值,分别用来预测 $x_{\sigma(1)}$ 与 $x_{\sigma(n)}$ 的大小。对于 $x_{\sigma(1)}$,通过式(1)得到预测误差 PE_{\min} :

$$PE_{\min} = x_s - x_t \quad (1)$$

其中:

$$\begin{cases} s = \min(\sigma(1), \sigma(2)) \\ t = \max(\sigma(1), \sigma(2)) \end{cases} \quad (2)$$

然后利用式(3)将密信比特 $b \in \{0, 1\}$ 嵌入到预测误差中得到 PE'_{\min} :

$$PE'_{\min} = \begin{cases} PE_{\min} + b, PE_{\min} = 1 \\ PE_{\min} + 1, PE_{\min} > 1 \\ PE_{\min} - b, PE_{\min} = 0 \\ PE_{\min} - 1, PE_{\min} < 0 \end{cases} \quad (3)$$

再利用式(4)得到 $x_{\sigma(1)}$ 的预测值 $x'_{\sigma(1)}$:

$$x'_{\sigma(1)} = x_{\sigma(2)} - |PE'_{\min}| = \begin{cases} x_{\sigma(1)} - b, PE_{\min} = 1 \\ x_{\sigma(1)} - 1, PE_{\min} > 1 \\ x_{\sigma(1)} - b, PE_{\min} = 0 \\ x_{\sigma(1)} - 1, PE_{\min} < 0 \end{cases} \quad (4)$$

最后得到含密像素块 $Y = (y_1, y_2, \dots, y_n)$, 其中, $y_{\sigma(1)} = x'_{\sigma(1)}$, 对于任意 $i \neq \sigma(1)$, $y_i = x_i$ 。

在解密端,首先根据式(5)计算出 PE'_{\min} :

$$PE'_{\min} = x'_s - x'_t \quad (5)$$

其中, s 和 t 的计算方法与嵌密时一致。然后根据得到的 PE'_{\min} 与 $x'_{\sigma(1)}$ 对 $x_{\sigma(1)}$ 进行恢复:

$$x_{\sigma(1)} = \begin{cases} x'_{\sigma(1)} + PE'_{\min} - 1, PE'_{\min} \in \{1, 2\} \\ x'_{\sigma(1)} + 1, PE'_{\min} > 2 \\ x'_{\sigma(1)} - PE'_{\min}, PE'_{\min} \in \{-1, 0\} \\ x'_{\sigma(1)} + 1, PE'_{\min} < -1 \end{cases} \quad (6)$$

同时提取出嵌入的密信比特 b :

$$b = \begin{cases} PE'_{\min} - 1, PE'_{\min} \in \{1, 2\} \\ -PE'_{\min}, PE'_{\min} \in \{-1, 0\} \end{cases} \quad (7)$$

与最小值 $x_{\sigma(1)}$ 的嵌密操作相似,通过式(8)将 $x_{\sigma(n)}$ 修改为 $x'_{\sigma(n)}$ 从而对最大值 $x_{\sigma(n)}$ 进行嵌密:

$$x'_{\sigma(n)} = \begin{cases} x_{\sigma(n)} + b, PE_{\max} = 1 \\ x_{\sigma(n)} + 1, PE_{\max} > 1 \\ x_{\sigma(n)} + b, PE_{\max} = 0 \\ x_{\sigma(n)} + 1, PE_{\max} < 0 \end{cases} \quad (8)$$

其中:

$$PE_{\max} = x_u - x_v \quad (9)$$

$$\begin{cases} u = \min(\sigma(n), \sigma(n-1)) \\ v = \max(\sigma(n), \sigma(n-1)) \end{cases} \quad (10)$$

在解密端,首先计算出含密的预测误差 PE'_{\max} :

$$PE'_{\max} = x'_u - x'_v \quad (11)$$

其中, u 和 v 的计算方式与嵌密时一致。然后根据得到的 PE'_{\max} 与 $x'_{\sigma(n)}$ 对 $x_{\sigma(n)}$ 进行恢复:

$$x_{\sigma(n)} = \begin{cases} x'_{\sigma(n)} - PE'_{\max} + 1, PE'_{\max} \in \{1, 2\} \\ x'_{\sigma(n)} - 1, PE'_{\max} > 2 \\ x'_{\sigma(n)} + PE'_{\max}, PE'_{\max} \in \{-1, 0\} \\ x'_{\sigma(n)} - 1, PE'_{\max} < -1 \end{cases} \quad (12)$$

同时提取出嵌入的密信比特 b :

$$b = \begin{cases} PE'_{\max} - 1, PE'_{\max} \in \{1, 2\} \\ -PE'_{\max}, PE'_{\max} \in \{-1, 0\} \end{cases} \quad (13)$$

PENG 等人根据排序后的相邻像素具有高度相关的特性,利用第二小(大)的像素去预测最小(大)的像素。将密信比特嵌入到符合条件的预测误差中,对于不符合要求的预测误差,通过加1或者减1来进行移位,以此保证解密的正确性。该方法对载体的修改较少,符合高保真的要求。然而,当分块大小 n 增大时,其仅预测最小(大)值可能会对其他像素值造成浪费。本文对此进行改进,并将该方法应用于音频载体上。

2 音频可逆隐写算法

本文引入采样块复杂度的概念,对采样块复杂度等级进行评估,得出最优预测采样值,利用最优预测采样值实现采样块的高效利用。

2.1 采样块复杂度

一些无损格式的音频采样值范围通常较大,在对音频载体进行分块后,如果不对采样块加以区分,会造成不必要的失真。因此,本文引入采样块复杂度 Δ 来衡量一个采样块内部采样值的变化幅度。首先利用事先设定的复杂度阈值 C_c 将采样块分成复杂块 ($\Delta > C_c$) 与平滑块 ($\Delta \leq C_c$),然后在进行嵌密操作时,对平滑块进行嵌密,对复杂块直接跳过。

实际上,音频的采样值是时域相关的,一个采样值的大小与其前后的采样值之间有着高度相关性。因此,本文通过与采样块相邻的前一个采样值 x_p 与后一个采样值 x_q 来计算该采样块复杂度。在给定密信长度的情况下,通过设定不同的阈值 C_c 可将密信自适应地嵌在不同的位置上。

2.2 最优预测采样值

对于一个大小为 n 的有序采样块 $X = \{x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(n)}\}$, 当 $n=3$ 时,用于预测的采样值被唯一确定,此时采样块的嵌密效率也最高,即由3个采样值组成的采样块在理想状态下最多能嵌入两位密信。当 n

的值逐渐增大时,按照传统的方式选取用于预测的2个采样值,位于这2个采样值之间的采样值会被浪费,这样显然不合理。为此,本文设计一种确定最优预测采样值的机制,从而充分利用块内所有可能嵌密的采样值。

首先,将采样块的复杂度分为 T 个等级, $1 \leq T \leq m$, 其中, $m = \lfloor (n-1)/2 \rfloor$, $\lfloor \cdot \rfloor$ 表示向下取整。对于平滑块,满足条件 $(T-k) \times \frac{C_c}{T} \leq \Delta \leq (T-k+1) \times \frac{C_c}{T}$ 的 k 即为该块的实际复杂度等级,最优预测采样值为 $x_{\sigma(k+1)}$ 与 $x_{\sigma(n-k)}$, 其中, $k \in [1, T]$ 。通过 $x_{\sigma(k+1)}$ 来预测比其小的所有采样值 $\{x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(k)}\}$, 通过 $x_{\sigma(n-k)}$ 来预测比其大的所有采样值 $\{x_{\sigma(n-k+1)}, x_{\sigma(n-k+2)}, \dots, x_{\sigma(n)}\}$ 。当 T 确定时, $\{x_{\sigma(T+2)}, x_{\sigma(T+3)}, \dots, x_{\sigma(n-T-1)}\}$ 在嵌密前后保持不变,将其记作 x_{in} 。最后,将2.1节的 x_p, x_q 与 x_{in} 相结合以计算采样块的复杂度:

$$\Delta = \sqrt{\frac{(x_p - \mu)^2 + (x_q - \mu)^2 + (x_{in} - \mu)^2}{n-2T}} \quad (14)$$

其中, μ 表示 x_p, x_q, x_{in} 的均值。

3 算法描述

3.1 嵌密算法

嵌密算法描述如下:

输入 长度为 N 的音频载体 S , 长度为 h 的密信比特

输出 一段含密音频 S'

步骤1 对音频载体 $S = \{x_1, x_2, \dots, x_N\}$ 按照大小为 n 进行分块,得到采样块 $S_i = \{x_{i+n \cdot (i-1)+1}, x_{i+n \cdot (i-1)+2}, \dots, x_{i+n \cdot i}\}$, 其中, $x_{i+n \cdot (i-1)}$ 、 $x_{i+n \cdot i+1}$ 分别表示与子块 S_i 相邻的前一个采样值、后一个采样值, $1 \leq i \leq \lfloor \frac{N-1}{n+1} \rfloor$ 。

步骤2 依次遍历所有采样块,对当前子块的采样值按照升序排列,得到有序的采样块 $X_i = \{x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(n)}\}$, 其中, $x_{\sigma(1)} \leq x_{\sigma(2)} \leq \dots \leq x_{\sigma(n)}$, $\sigma(i)$ 表示 $x_{\sigma(i)}$ 在原始采样块 S_i 中的位置。

步骤3 通过给定的复杂度等级 T , 得到子块内部用于计算复杂度的采样值 $\{x_{\sigma(T+2)}, x_{\sigma(T+3)}, \dots, x_{\sigma(n-T-1)}\}$, 将其与步骤1得到的 $x_{i+n \cdot (i-1)}$ 、 $x_{i+n \cdot i+1}$ 相结合, 通过式(14)来计算 S_i 的复杂度 Δ 。

步骤4 利用预先设定的复杂度阈值 C_c 对当前子块 S_i 进行判断, 如果 $\Delta > C_c$, 说明当前子块 S_i 属于复杂块, 不做任何操作并将该采样块的采样值按照原始顺序放回 S 中, 然后转到步骤2; 如果 $\Delta \leq C_c$, 说明该块属于平滑块, 进入步骤5。

步骤5 由2.2节得到最优预测采样值 $x_{\sigma(k+1)}$ 与 $x_{\sigma(n-k)}$, 利用式(1)、式(9)来预测 $\{x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(k)}\}$ 与 $\{x_{\sigma(n-k+1)}, x_{\sigma(n-k+2)}, \dots, x_{\sigma(n)}\}$, 得到 k 个 PE_{\min} 与 k 个 PE_{\max} , 最后通过式(4)、式(8)将密信嵌入载体中。

步骤6 判断密信是否嵌入完毕, 如果还有密信

未嵌入,则转到步骤2;如果密信嵌入完毕,进入步骤7。

步骤7 嵌密结束,得到含密音频 S' 。

3.2 辅助信息

在嵌密过程中,由于采样值发生了改变,因此可能会产生溢出问题。为了解决该问题,本文建立一个位置映射图 L_M 来标记可能产生溢出的采样块。16 bit量化的音频采样值范围在 $[-32\ 768, 32\ 767]$ 之间,因此,如果当前块内有采样值等于 $-32\ 768$ 或者 $32\ 767$,则令 $L_M(i)=1$;反之, $L_M(i)=0$ 。在嵌密结束后,将得到的位置映射表 L_M 进行无损压缩得到 L_s 。实际上,当分块大小确定后, L_M 的大小也被唯一确定。将压缩后的位置图 L_s 与分块大小 n 、复杂度阈值 C_z 、密信长度 h 、辅助信息结束标志 E_G 依次连接组合成辅助信息 \mathcal{L} ,将 \mathcal{L} 以 LSB 替换的方式嵌入音频的开头部分。同时,为了保证可逆,把被 \mathcal{L} 替换的部分与密信一起嵌入载体中。辅助信息的各部分所占比特大小如表1所示。

表1 辅助信息各部分所占比特大小

Table 1 Bit size of each part of auxiliary information bit

辅助信息	大小
分块大小	4
复杂度阈值	8
密信长度	20
结束标志	8

3.3 密信提取与载体恢复

在接收到含密音频后,根据辅助信息结束标志 E_G 提取出开头部分的辅助信息,依次得到位置映射表 L_s 、分块大小 n 、密信长度 h ,再将 L_s 解压缩得到原始的 L_M 。解密过程具体如下:

输入 含密音频 S'

输出 原始音频 S , 长度为 h 的密信

步骤1 对含密音频载体 S' 按照大小为 n 进行分块,得到采样块 $S'_i = \{x'_{i+n \cdot (i-1)+1}, x'_{i+n \cdot (i-1)+2}, \dots, x'_{i+n \cdot i}\}$, 其中, $x'_{i+n \cdot (i-1)}$ 、 $x'_{i+n \cdot i+1}$ 分别表示与子块 S'_i 相邻的前一个采样值、后一个采样值, $1 \leq i \leq \left\lfloor \frac{N-1}{n+1} \right\rfloor$ 。

步骤2 依次遍历所有采样块,对当前块 S'_i 的采样值按照升序排列,得到有序的采样块 $X'_i = \{x'_{\sigma(1)}, x'_{\sigma(2)}, \dots, x'_{\sigma(n)}\}$, 其中, $x'_{\sigma(1)} \leq x'_{\sigma(2)} \leq \dots \leq x'_{\sigma(n)}$, $\sigma(i)$ 表示 $x'_{\sigma(i)}$ 在原始采样块 S'_i 中的位置。

步骤3 通过给定的复杂度等级 T , 得到采样块内部用于计算复杂度的采样值 $\{x'_{\sigma(T+2)}, x'_{\sigma(T+3)}, \dots, x'_{\sigma(n-T-1)}\}$, 将其与步骤1得到的 $x'_{i+n \cdot (i-1)}$ 、 $x'_{i+n \cdot i+1}$ 相结合, 通过式(14)来计算采样块 S'_i 的复杂度 Δ' 。由于用于计算复杂度的采样值在嵌密前后保持不变,因此 $\Delta = \Delta'$, 从而保证了在解密时不会发生差错。

步骤4 利用预先设定的复杂度阈值 C_z 对当前子块 S'_i 进行判断。如果 $\Delta' > C_z$, 说明当前子块 S'_i 属于复杂块, 不做任何操作并将该采样块的采样值按照原始顺序放回 S' 中, 然后转到步骤2; 如果 $\Delta' \leq C_z$, 说明该块属于平滑块, 进入步骤5。

步骤5 由2.2节得到最优预测采样值 $x'_{\sigma(k+1)}$ 与 $x'_{\sigma(n-k)}$, 利用式(5)、式(11)来预测 $\{x'_{\sigma(1)}, x'_{\sigma(2)}, \dots, x'_{\sigma(k)}\}$ 与 $\{x'_{\sigma(n-k+1)}, x'_{\sigma(n-k+2)}, \dots, x'_{\sigma(n)}\}$, 得到 k 个 PE'_{\min} 与 k 个 PE'_{\max} , 最后通过式(7)、式(13)提取密信, 利用式(6)、式(12)恢复载体。

步骤6 判断密信是否提取完毕, 如果提取完毕, 则进入步骤7; 否则, 转到步骤2。

步骤7 解密结束, 得到长度为 h 的密信。

在解密后, 根据辅助信息结束标志 E_G , 将提取出的密信分成2个部分, 前一部分是原始载体头部的一段LSB序列, 将其进行替换得到完整的原始载体, 后一部分即为密信。

4 实验结果与分析

4.1 实验设置

分块大小直接影响算法的性能, 本文将分块大小 n 设置为 $\{3, 4, \dots, 11\}$, 选取信噪比(SNR)值最高的分块大小用作最终嵌密。实验采用的载体为公开的EBU-SQAM标准音频测试集^[22], 其包括70段长度不一的音序列。其中, (1、2)为噪音信号, (3~7)为人工信号, (8~43)为单一乐器音乐, (44~48)为纯人声音乐, (49~54)为人的语音, (55~60)为独奏乐, (61~64)为声乐与管弦乐, (65~68)为管弦乐, (69、70)为流行音乐。采样率为44.1 kHz, 量化位数为16 bit。

为了简化操作, 统一将载体设置为单声道。密信是随机生成的0和1, 共分为50 000 bit、100 000 bit 2种长度。对比算法采用音频领域较具代表性的DE^[16]和PEE^[18]2种算法, 在Matlab R2018a仿真软件上进行实验。

4.2 评价指标

SNR与客观差异等级(ODG)常被用来评价隐写后的音频质量。SNR的单位是分贝(dB), 其值越高, 代表隐写后音频的失真越小、质量越高。SNR的计算公式为:

$$SNR = 10 \lg \left(\frac{\sum_{n=0}^{N-1} x^2(n)}{\sum_{n=0}^{N-1} [\bar{x}(n) - x(n)]^2} \right) \quad (15)$$

其中, $\bar{x}(n)$ 表示含密音频, $x(n)$ 表示原始音频, N 表示采样点个数。

ODG是一种音频质量客观评价指标, 将参考信号和测试信号分别经过心理声学模型处理后, 通过感知模型对参考信号与测试信号之间的声学特征差

异进行综合评估,得到一个评价分数,即为 ODG。ODG 的取值范围一般在 $[-4,0]$ 之间,值越大表示音频的感知质量越高。

4.3 结果分析

隐写前后的音频波形在视觉上已无法分辨,原始音频波形如图 1 所示,隐写后的波形(50 000 bit)如图 2 所示,隐写前后音频波形的差值如图 3 所示。

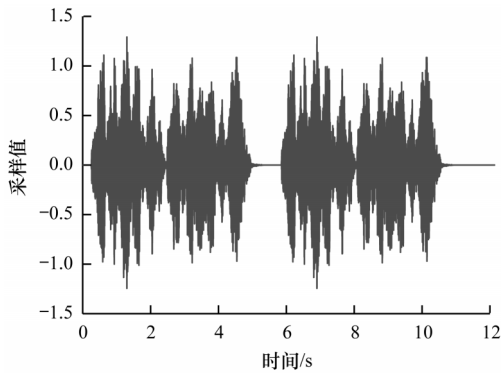


图1 原始音频波形

Fig.1 Original audio waveform

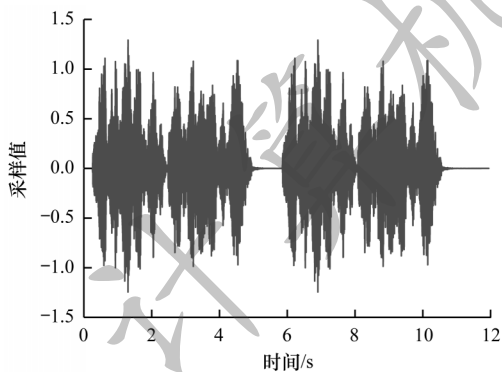


图2 隐写后的音频波形

Fig.2 Audio waveform after steganography

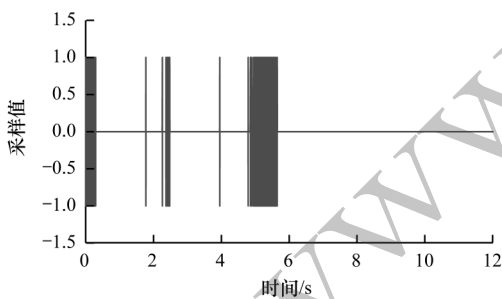


图3 隐写前后音频波形的差值

Fig.3 The difference of audio waveform before and after steganography

将图 1 与图 2 进行对比可以看出,隐写前后音频在波形图上没有明显变化。从图 3 可以看出,本文方法对采样值的修改幅度在 $[-1,1]$ 之间,对载体的影响很小。为了进一步验证算法的优越性,将不同长度密信下的 SNR 值与 ODG 值进行对比,结果如表 2~表 5 所示,其中,最优结果加粗表示。

表 2 密信长度为 50 000 bit 时的平均 SNR 值

Table 2 The average SNR value when the secret message length is 50 000 bit

音频序列	SNR/dB		
	DE 算法	PEE 算法	本文算法
1、2	41.58	71.62	82.36
3~7	36.97	91.59	94.14
8~43	37.87	77.19	83.29
44~48	34.72	75.77	82.71
49~54	29.79	73.70	80.80
55~60	32.75	71.78	79.64
61~64	43.16	76.92	85.72
65~68	47.99	77.42	85.16
69、70	35.82	71.81	81.05
1~70	37.38	77.04	83.66

表 3 密信长度为 50 000 bit 时的平均 ODG 值

Table 3 The average ODG value when the secret message length is 50 000 bit

音频序列	ODG		
	DE 算法	PEE 算法	本文算法
1、2	-0.152	0.021	0.014
3~7	-0.837	0.006	0.011
8~43	-0.404	0.015	0.018
44~48	-0.497	0.005	0.005
49~54	-0.695	0.002	-0.005
55~60	-0.616	0.005	0.013
61~64	-0.392	0.001	0.004
65~68	-0.230	0.011	0.011
69、70	-0.424	-0.005	0.004
1~70	-0.467	0.010	0.012

表 4 密信长度为 100 000 bit 时的平均 SNR 值

Table 4 The average SNR value when the secret message length is 100 000 bit

音频序列	SNR/dB		
	DE 算法	PEE 算法	本文算法
1、2	36.82	70.50	79.46
3~7	33.99	88.82	91.20
8~43	33.32	75.58	80.68
44~48	29.94	72.77	79.79
49~54	26.18	70.24	78.89
55~60	28.48	68.41	76.93
61~64	34.37	74.74	83.39
65~68	45.61	73.37	81.45
69、70	31.22	71.40	78.50
1~70	32.90	74.80	81.00

表5 密信长度为100 000 bit时的平均ODG值

Table 5 The average ODG value when the secret message length is 100 000 bit

音频序列	ODG		
	DE算法	PEE算法	本文算法
1、2	-0.258	0.023	0.010
3~7	-0.754	0.005	0.004
8~43	-0.535	0.014	0.019
44~48	-0.623	0.016	0.007
49~54	-0.835	0.004	-0.005
55~60	-0.798	0.008	0.012
61~64	-0.480	0.007	0.004
65~68	-0.314	0.008	0.010
69、70	-0.596	-0.004	0.005
1~70	-0.583	0.011	0.012

从表2、表4可以看出,当密信长度为50 000 bit时,本文算法的平均SNR值相比DE算法、PEE算法分别提高了46.28 dB、6.62 dB;当密信长度增加到100 000 bit时,本文算法的平均SNR值相比DE算法、PEE算法分别提高了48.10 dB、6.20 dB。对于9种不同的音频,经过本文算法隐写后的SNR值都有了明显提高,验证了本文算法具有高保真的特性。

从表3、表5可以看出,当密信长度为50 000 bit时,本文算法的平均ODG值相比DE算法、PEE算法分别提高了0.479、0.002;当密信长度达到100 000 bit时,本文算法的平均ODG值相比DE算法、PEE算法分别提高了0.595、0.001。对于数据集集中的9种音频,本文算法的ODG值均能保持一个良好的水平,并且整体上较对比算法有一定幅度的提高。

通过上述对SNR与ODG 2个指标的综合分析可以得出:本文算法能够很好地保证音频的感知质量,且对于不同种类的音频具有一定的普适性,在保真度上优于现有的DE算法、PEE算法。

5 结束语

DE、PEE等隐写算法将密信比特以二进制位扩展的方式嵌入到预测误差值中,对载体的修改幅度较大。为此,本文提出一种基于SVO的音频可逆信息隐写算法,该算法对采样值的修改幅度在 $[-1, 1]$ 之间,具有高保真的特点。同时,本文算法引入采样块复杂度的概念,根据采样块的复杂度实现密信的自适应嵌入。实验结果表明,相对DE算法和PEE算法,该算法的SNR值和ODG值较高。在嵌密过程中,采样块的大小固定不变,这在一定程度上影响了算法的嵌密容量,因此,设计更合理的分块策略以提高嵌密容量将是下一步的研究方向。

参考文献

- [1] SHI Yunqing, LI Xiaolong, ZHANG Xinpeng, et al. Reversible data hiding: advances in the past two decades [J]. IEEE Access, 2016, 4: 3210-3237.
- [2] SHI Y Q, ANSARI N, SU W, et al. Reversible data hiding [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16(3): 354-362.
- [3] HOU Dongdong, ZHANG Weiming, LIU Jiayang, et al. Emerging applications of reversible data hiding [C]// Proceedings of the 2nd International Conference on Image and Graphics Processing. New York, USA: ACM Press, 2019: 105-109.
- [4] TIAN Jun. Reversible data embedding using a difference expansion [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(8): 890-896.
- [5] THODI D M, RODRIGUEZ J J. Expansion embedding techniques for reversible watermarking [J]. IEEE Transactions on Image Processing, 2007, 16(3): 721-730.
- [6] NEDELCO T, IORDACHE R, COLTUC D. Three stages prediction-error expansion reversible watermarking [C]// Proceedings of European Signal Processing Conference. Washington D. C., USA: IEEE Press, 2014: 2455-2459.
- [7] MEHTA J H, KELKAR V. Comparison of reversible watermarking using prediction error expansion and prediction error expansion considering region of interest for medical images [C]// Proceedings of International Conference for Convergence in Technology. Washington D. C., USA: IEEE Press, 2017: 766-770.
- [8] TSAI P, HU Y C, YEH H L. Reversible image hiding scheme using predictive coding and histogram shifting [J]. Signal Process, 2009, 89(6): 1129-1143.
- [9] OU Bo, ZHAO Yao. High capacity reversible data hiding based on multiple histograms modification [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2020, 30(8): 2329-2342.
- [10] LI Xiaolong, LI Jian, LI Bin, et al. High-fidelity reversible data hiding scheme based on pixel-value-ordering and prediction-error expansion [J]. Signal Processing, 2013, 93(1): 198-205.
- [11] PENG Fei, LI Xiaolong, YANG Bin. Improved PVO-based reversible data hiding [J]. Digital Signal Processing, 2014, 25: 255-265.
- [12] WENG Shaowei, SHI Yunqing, HONG Wen, et al. Dynamic improved pixel value ordering reversible data hiding [J]. Information Sciences, 2019, 489: 136-154.
- [13] HE Wenguang, XIONG Gangqiang, WENG Shaowei, et al. Reversible data hiding using multi-pass pixel-value-ordering and pairwise prediction-error expansion [J]. Information Sciences, 2018, 467: 784-799.
- [14] WU Haorui, LI Xiaolong, ZHAO Yao, et al. Improved reversible data hiding based on PVO and adaptive pairwise embedding [J]. Real Time Image Processing, 2019, 16: 685-695.
- [15] WANG Weiqing, YE Junyong, WANG Tongqing, et al. A high capacity reversible data hiding scheme based on right-left shift [J]. Signal Processing, 2018, 150: 102-115.
- [16] YAN Diqun, WANG Rangding. Audio lossless information hiding algorithm based on difference expansion [J]. Computer Engineering and Applications, 2008, 44(9): 98-100. (in Chinese)

严迪群,王让定. 基于差值扩展的音频无损信息隐藏算法[J]. 计算机工程与应用, 2008, 44(9): 98-100.

(下转第138页)

~~~~~  
(上接第 128 页)

- [17] WANG Fei, XIE Zhaoxin, CHEN Zuo. High capacity reversible watermarking for audio by histogram shifting and predicted error expansion [EB/OL]. [2019-10-10]. <https://www.hindawi.com/journals/tswj/2014/656251/>.
- [18] XIANG Shijun, LI Zihao. Reversible audio data hiding algorithm using noncausal prediction of alterable orders [EB/OL]. [2019-10-10]. <https://asmp-urasipjournals.springeropen.com/track/pdf/10.1186/s13636-017-0101-9>.
- [19] BOBEICA A, DRAGOI I C, CACIULA I, et al. Capacity control for prediction error expansion based audio reversible data hiding [C]//Proceedings of International Conference on System Theory, Control and Computing. Washington D. C. , USA; IEEE Press, 2018; 810-815.
- [20] NISHIMURA A. Reversible audio data hiding using linear prediction and error expansion [C]//Proceedings of International Conference on Intelligent Information Hiding and Multimedia Signal Processing. Washington D. C. , USA; IEEE Press, 2011; 318-321.
- [21] CHOI K, PUN C, CHEN C L P. Application of a generalized difference expansion based reversible audio data hiding algorithm [J]. Multimedia Tools and Applications, 2013, 74; 1961-1982.
- [22] EBU committee; sound quality assessment material recordings for subjective tests [EB/OL]. [2019-10-10]. <https://tech.ebu.ch/publications/sqamed>.

编辑 吴云芳