



(12) 发明专利申请

(10) 申请公布号 CN 104282310 A

(43) 申请公布日 2015. 01. 14

(21) 申请号 201410502396. X

(22) 申请日 2014. 09. 26

(71) 申请人 宁波大学

地址 315211 浙江省宁波市江北区风华路
818 号

(72) 发明人 王让定 羊开云 严迪群 金超
孙冉 周劲蕾

(74) 专利代理机构 宁波奥圣专利代理事务所
(普通合伙) 33226

代理人 周珏

(51) Int. Cl.

G10L 19/018(2013. 01)

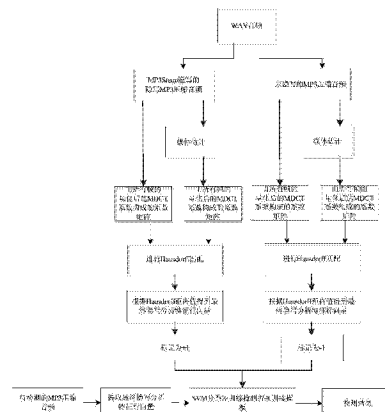
权利要求书3页 说明书8页 附图1页

(54) 发明名称

一种针对 MP3Stego 隐写后的音频的隐写检测方法

(57) 摘要

本发明公开了一种针对 MP3Stego 隐写后的音频的隐写检测方法,其先将未隐写的 MP3 压缩音频和隐写 MP3 压缩音频构成样本库;然后对每个样本进行重压缩编码,获得每个样本的载体估计;接着提取每个样本中的每帧的量化后的 MDCT 系数,获取每个样本对应的第一系数矩阵,同样提取每个载体估计中的每帧的量化后的 MDCT 系数,获取每个载体估计对应的第二系数矩阵;再通过计算对应的第一系数矩阵与第二系数矩阵中对应列之间的 Hausdorff 距离值,得到每个样本的最终隐写分析特征行向量;之后通过 SVM 分类器训练得到训练模板;最后利用训练模板对待检测的 MP3 压缩音频进行检测;优点是能很准确地确定 MP3 压缩音频是否经过 MP3Stego 隐写,尤其在隐秘信息嵌入率较低的情况下依然能获得很高的检测效率。



1. 一种针对 MP3Stego 隐写后的音频的隐写检测方法,其特征在于包括以下步骤:

①选取 N 个变化风格各不相同且未压缩的 WAV 音频,其中, $N \geq 100$; 然后利用 MP3 音频 8HZ 编码器对每个 WAV 音频进行压缩编码,得到每个 WAV 音频对应的 MP3 压缩音频;并利用 MP3Stego 音频隐写工具对每个 WAV 音频隐写不同长度、不同内容的隐秘信息,得到每个 WAV 音频对应的隐写 MP3 压缩音频;再将 N 个 MP3 压缩音频和 N 个隐写 MP3 压缩音频构成一个样本库;

②利用 MP3 音频 lame 解码器对样本库中的每个音频样本进行解压缩,得到样本库中的每个音频样本对应的 WAV 音频,在解压缩过程中,对样本库中的每个音频样本中的所有帧序号为偶数的帧的编码参数量化步长进行修正;然后利用 MP3 音频 lame 编码器对样本库中的每个音频样本对应的 WAV 音频进行压缩编码,得到样本库中的每个音频样本对应的载体估计;

③利用 MP3 音频 lame 解码器对样本库中的每个音频样本进行解压缩,提取出样本库中的每个音频样本中的每帧的 576 个量化后的 MDCT 系数,以每个音频样本中的每帧的 576 个量化后的 MDCT 系数作为一行,将样本库中的每个音频样本对应的所有量化后的 MDCT 系数构成一个第一系数矩阵,将样本库中的第 i 个音频样本对应的所有量化后的 MDCT 系数构成

的第一系数矩阵记为 X_i , $X_i = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,575} & x_{1,576} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,575} & x_{2,576} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{N_i^f,1} & x_{N_i^f,2} & \cdots & x_{N_i^f,575} & x_{N_i^f,576} \end{bmatrix}$, 其中, $1 \leq i \leq 2N$, X_i 的维

数为 $N_i^f \times 576$, N_i^f 表示样本库中的第 i 个音频样本中包含的帧的总数, $x_{1,1}$ 、 $x_{1,2}$ 、 $x_{1,575}$ 、 $x_{1,576}$ 对应表示样本库中的第 i 个音频样本中的第 1 帧的第 1 个、第 2 个、第 575 个、第 576 个量化后的 MDCT 系数, $x_{2,1}$ 、 $x_{2,2}$ 、 $x_{2,575}$ 、 $x_{2,576}$ 对应表示样本库中的第 i 个音频样本中的第 2 帧的第 1 个、第 2 个、第 575 个、第 576 个量化后的 MDCT 系数, $x_{N_i^f,1}$ 、 $x_{N_i^f,2}$ 、 $x_{N_i^f,575}$ 、 $x_{N_i^f,576}$ 对应表示样本库中的第 i 个音频样本中的第 N_i^f 帧的第 1 个、第 2 个、第 575 个、第 576 个量化后的 MDCT 系数;

同样,利用 MP3 音频 lame 解码器对样本库中的每个音频样本对应的载体估计进行解压缩,提取出样本库中的每个音频样本对应的载体估计中的每帧的 576 个量化后的 MDCT 系数,以每个载体估计中的每帧的 576 个量化后的 MDCT 系数作为一行,将样本库中的每个音频样本对应的载体估计对应的所有量化后的 MDCT 系数构成一个第二系数矩阵,将样本库中的第 i 个音频样本对应的载体估计对应的所有量化后的 MDCT 系数构成的第二系

数矩阵记为 X_i' , $X_i' = \begin{bmatrix} x'_{1,1} & x'_{1,2} & \cdots & x'_{1,575} & x'_{1,576} \\ x'_{2,1} & x'_{2,2} & \cdots & x'_{2,575} & x'_{2,576} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x'_{N_i^f,1} & x'_{N_i^f,2} & \cdots & x'_{N_i^f,575} & x'_{N_i^f,576} \end{bmatrix}$, 其中, $1 \leq i \leq 2N$, X_i' 的维

数为 $N_i^f \times 576$, N_i^f 表示样本库中的第 i 个音频样本中包含的帧的总数,亦表示样本库中的第 i 个音频样本对应的载体估计中包含的帧的总数, $x'_{1,1}$ 、 $x'_{1,2}$ 、 $x'_{1,575}$ 、 $x'_{1,576}$ 对应表示

样本库中的第 i 个音频样本对应的载体估计中的第 1 帧的第 1 个、第 2 个、第 575 个、第 576 个量化后的 MDCT 系数, $x'_{2,1}$ 、 $x'_{2,2}$ 、 $x'_{2,575}$ 、 $x'_{2,576}$ 对应表示样本库中的第 i 个音频样本对应的载体估计中的第 2 帧的第 1 个、第 2 个、第 575 个、第 576 个量化后的 MDCT 系数, $x'_{N_f^f,1}$ 、 $x'_{N_f^f,2}$ 、 $x'_{N_f^f,575}$ 、 $x'_{N_f^f,576}$ 对应表示样本库中的第 i 个音频样本对应的载体估计中的第 N_f^f 帧的第 1 个、第 2 个、第 575 个、第 576 个量化后的 MDCT 系数;

④计算样本库中的每个音频样本对应的所有量化后的 MDCT 系数构成的第一系数矩阵中的每一列与该音频样本对应的载体估计对应的所有量化后的 MDCT 系数构成的第二系数矩阵中对应列之间的 Hausdorff 距离值,将样本库中的每个音频样本对应的 576 个 Hausdorff 距离值构成一个初级隐写分析特征行向量,将样本库中的第 i 个音频样本对应的 576 个 Hausdorff 距离值构成的初级隐写分析特征行向量记为 F_i , $F_i = [hd_1 \ hd_2 \cdots hd_k \cdots hd_{575} \ hd_{576}]$, 其中, $1 \leq i \leq 2N$, hd_k 表示 X_i 中的第 k 列与 X_i' 中的第 k 列之间的 Hausdorff 距离值, $1 \leq k \leq 576$;

⑤根据样本库中的每个音频样本对应的 576 个 Hausdorff 距离值构成的初级隐写分析特征行向量,获取样本库中的每个音频样本对应的最终隐写分析特征行向量,将样本库中的第 i 个音频样本对应的最终隐写分析特征行向量记为 F_i' , $F_i' = [\mu_i \ m_i \ \sigma_i \ p_i \ s_i]$, 其中, $1 \leq i \leq 2N$, μ_i 表示 F_i 中的所有元素的均值, $\mu_i = \frac{1}{576} \sum_{k=1}^{576} hd_k$, m_i 表示 F_i 中的所有元素中的最大值, $m_i = \max(hd_1, hd_2, \dots, hd_k, \dots, hd_{576})$, $\max()$ 为取最大值函数, σ_i 表示 F_i 中的所有元素的方差, $\sigma_i = \frac{1}{576} \sum_{k=1}^{576} (hd_k - \mu_i)^2$, p_i 表示 F_i 中的所有元素的峰度系数,

$$p_i = \frac{\sum_{k=1}^{576} (hd_k - \mu_i)^4}{(\sigma_i)^2}, \quad s_i \text{ 表示 } F_i \text{ 中的所有元素的三阶中心距, } s_i = \frac{1}{576} \sum_{k=1}^{576} (hd_k - \mu_i)^3;$$

⑥对样本库中的每个音频样本对应的最终隐写分析特征行向量进行标记,对于样本库中的第 i 个音频样本,如果该音频样本为 MP3 压缩音频,则将该音频样本对应的最终隐写分析特征行向量标记为 -1,如果该音频样本为隐写 MP3 压缩音频,则将该音频样本对应的最终隐写分析特征行向量标记为 +1,其中, $1 \leq i \leq 2N$;然后将所有标记后的最终隐写分析特征行向量输入 SVM 分类器进行训练,得到训练模板;

⑦对于任意一个待检测的 MP3 压缩音频,按照步骤②至步骤⑤的过程,以相同的方式获取该 MP3 压缩音频对应的最终隐写分析特征行向量,然后利用训练模板对该 MP3 压缩音频对应的最终隐写分析特征行向量进行检测,以确定该 MP3 压缩音频是否经过 MP3Stego 音频隐写工具隐写。

2. 根据权利要求 1 所述的一种针对 MP3Stego 隐写后的音频的隐写检测方法,其特征在于所述的步骤②中对样本库中的每个音频样本中的所有帧序号为偶数的帧的编码参数量化步长进行修正的具体过程为:

② -1、将样本库中当前待处理的第 i 个音频样本定义为当前音频样本,其中, $1 \leq i \leq 2N$;

②-2、假设当前音频样本中包含有 N_i^f 帧,且帧序号从 1 开始编,并将当前音频样本中当前待处理的第 j 帧定义为当前帧,其中, $1 \leq j \leq N_i^f$;

②-3、如果当前帧的帧序号 j 为奇数,则保持当前帧的编码参数量化步长不变,然后执行步骤②-4;如果当前帧的帧序号 j 为偶数,则将前一帧的编码参数量化步长作为当前帧的编码参数量化步长,完成对当前帧的编码参数量化步长的修正,然后执行步骤②-4;

②-4、令 $j = j+1$,然后将当前音频样本中下一帧待处理的帧作为当前帧,再返回步骤②-3 继续执行,直至当前音频样本中的所有帧处理完毕,其中, $j = j+1$ 中的“=”为赋值符号;

②-5、令 $i = i+1$,然后将样本库中下一个待处理的音频样本作为当前音频样本,再返回步骤②-2 继续执行,直至样本库中的所有音频样本处理完毕,其中, $i = i+1$ 中的“=”为赋值符号。

3. 根据权利要求 1 或 2 所述的一种针对 MP3Stego 隐写后的音频的隐写检测方法,其特征在于所述的步骤④中 X_i 中的第 k 列与 X_i' 中的第 k 列之间的 Hausdorff 距离值 hd_k 的获取过程为:

④-1、将 X_i 中的第 k 列记为 $C_{i,k}$,将 X_i' 中的第 k 列记为 $C'_{i,k}$;

④-2、将 $C_{i,k}$ 中当前待处理的第 j 个元素定义为当前元素,其中, $1 \leq j \leq N_i^f$;

④-3、计算当前元素与 $C'_{i,k}$ 中的每个元素之间的欧氏距离,得到当前元素对应的 N_i^f 个欧氏距离,然后获取当前元素对应的 N_i^f 个欧氏距离中的最小欧氏距离值;

④-4、令 $j = j+1$,然后将 $C_{i,k}$ 中下一个待处理的元素作为当前元素,再返回步骤④-3 继续执行,直至 $C_{i,k}$ 中的所有元素处理完毕,得到 $C_{i,k}$ 中的每个元素对应的一个最小欧氏距离值,其中, $j = j+1$ 中的“=”为赋值符号;

④-5、获取 $C_{i,k}$ 对应的 N_i^f 个最小欧氏距离值中的最大值,然后将该最大值作为 X_i 中的第 k 列与 X_i' 中的第 k 列之间的 Hausdorff 距离值,记为 hd_k 。

一种针对 MP3Stego 隐写后的音频的隐写检测方法

技术领域

[0001] 本发明涉及一种隐写检测方法,尤其是涉及一种针对 MP3Stego 隐写后的音频的隐写检测方法。

背景技术

[0002] 隐写术是一种结合公开传播的多媒体文件的自身特点,将要传递的隐秘信息通过隐写算法隐藏到多媒体文件中的技术。隐写后的多媒体文件具有良好的不可感知性,在传输的过程中一般较难被检测出其携带的隐秘信息,只有隐写者指定的特定人群才能够从隐写后的多媒体文件中提取出隐秘信息。为防止不法分子将隐写术用于危害公共安全,需利用隐写检测技术对公开传播的多媒体文件进行检测,来判断公开传播的多媒体文件是否经过隐写。

[0003] 作为隐写载体的多媒体文件的格式主要有图像、文本、视频、音频等。音频文件因其近年来在互联网上广泛流传,已成为隐写者与隐写检测者关注的热点。MP3Stego 是目前最典型并且应用最广泛的音频隐写工具。针对 MP3Stego 隐写音频,目前已有很多检测算法,但大多数针对 MP3Stego 隐写音频的检测算法都是直接提取作为隐秘信息嵌入点的音频中的每一帧的编码块长度或者其他编码参数,通过分析这些编码参数变化特点来检测音频是否经过 MP3Stego 隐写,然而由于音频风格变化存在多样性,因此会影响编码块长度或其他编码参数的变化,从而会对隐写检测工作造成干扰,最终影响检测准确率。近年来,出现了一些对待检测音频进行消除隐写痕迹的操作如去噪,从而得到对原始音频的估计,然后将估计得到的音频与待检测音频相减从而消除音频本身风格变化的影响,提高了检测率,然而这些方法对原始音频的估计操作不能很好的消除隐写痕迹,因此得到的对原始音频的估计不够精确,导致在音频嵌入较少隐秘信息的情况下,检测效果不够理想。

发明内容

[0004] 本发明所要解决的技术问题是提供一种针对 MP3Stego 隐写后的音频的隐写检测方法,其在隐秘信息嵌入率较低的情况下,依然能够获得很高的检测效率。

[0005] 本发明解决上述技术问题所采用的技术方案为:一种针对 MP3Stego 隐写后的音频的隐写检测方法,其特征在于包括以下步骤:

[0006] ①选取 N 个变化风格各不相同且未压缩的 WAV 音频,其中, $N \geq 100$;然后利用 MP3 音频 8HZ 编码器对每个 WAV 音频进行压缩编码,得到每个 WAV 音频对应的 MP3 压缩音频;并利用 MP3Stego 音频隐写工具对每个 WAV 音频隐写不同长度、不同内容的隐秘信息,得到每个 WAV 音频对应的隐写 MP3 压缩音频;再将 N 个 MP3 压缩音频和 N 个隐写 MP3 压缩音频构成一个样本库;

[0007] ②利用 MP3 音频 lame 解码器对样本库中的每个音频样本进行解压缩,得到样本库中的每个音频样本对应的 WAV 音频,在解压缩过程中,对样本库中的每个音频样本中的所有帧序号为偶数的帧的编码参数量化步长进行修正;然后利用 MP3 音频 lame 编码器对样本

库中的每个音频样本对应的 WAV 音频进行压缩编码,得到样本库中的每个音频样本对应的载体估计;

[0008] ③利用 MP3 音频 lame 解码器对样本库中的每个音频样本进行解压缩,提取出样本库中的每个音频样本中的每帧的 576 个量化后的 MDCT 系数,以每个音频样本中的每帧的 576 个量化后的 MDCT 系数作为一行,将样本库中的每个音频样本对应的所有量化后的 MDCT 系数构成一个第一系数矩阵,将样本库中的第 i 个音频样本对应的所有量化后的 MDCT 系数

$$\text{构成的第一系数矩阵记为 } X_i, X_i = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,575} & x_{1,576} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,575} & x_{2,576} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{N_i^f,1} & x_{N_i^f,2} & \cdots & x_{N_i^f,575} & x_{N_i^f,576} \end{bmatrix}, \text{其中}, 1 \leq i \leq 2N, X_i$$

的维数为 $N_i^f \times 576$, N_i^f 表示样本库中的第 i 个音频样本中包含的帧的总数, $x_{1,1}$ 、 $x_{1,2}$ 、 $x_{1,575}$ 、 $x_{1,576}$ 对应表示样本库中的第 i 个音频样本中的第 1 帧的第 1 个、第 2 个、第 575 个、第 576 个量化后的 MDCT 系数, $x_{2,1}$ 、 $x_{2,2}$ 、 $x_{2,575}$ 、 $x_{2,576}$ 对应表示样本库中的第 i 个音频样本中的第 2 帧的第 1 个、第 2 个、第 575 个、第 576 个量化后的 MDCT 系数, $x_{N_i^f,1}$ 、 $x_{N_i^f,2}$ 、 $x_{N_i^f,575}$ 、 $x_{N_i^f,576}$ 对应表示样本库中的第 i 个音频样本中的第 N_i^f 帧的第 1 个、第 2 个、第 575 个、第 576 个量化后的 MDCT 系数;

[0009] 同样,利用 MP3 音频 lame 解码器对样本库中的每个音频样本对应的载体估计进行解压缩,提取出样本库中的每个音频样本对应的载体估计中的每帧的 576 个量化后的 MDCT 系数,以每个载体估计中的每帧的 576 个量化后的 MDCT 系数作为一行,将样本库中的每个音频样本对应的载体估计对应的所有量化后的 MDCT 系数构成一个第二系数矩阵,将样本库中的第 i 个音频样本对应的载体估计对应的所有量化后的 MDCT 系数构成的第二

$$\text{系数矩阵记为 } X_i', X_i' = \begin{bmatrix} x'_{1,1} & x'_{1,2} & \cdots & x'_{1,575} & x'_{1,576} \\ x'_{2,1} & x'_{2,2} & \cdots & x'_{2,575} & x'_{2,576} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x'_{N_i^f,1} & x'_{N_i^f,2} & \cdots & x'_{N_i^f,575} & x'_{N_i^f,576} \end{bmatrix}, \text{其中}, 1 \leq i \leq 2N, X_i' \text{ 的}$$

维数为 $N_i^f \times 576$, N_i^f 表示样本库中的第 i 个音频样本中包含的帧的总数,亦表示样本库中的第 i 个音频样本对应的载体估计中包含的帧的总数, $x'_{1,1}$ 、 $x'_{1,2}$ 、 $x'_{1,575}$ 、 $x'_{1,576}$ 对应表示样本库中的第 i 个音频样本对应的载体估计中的第 1 帧的第 1 个、第 2 个、第 575 个、第 576 个量化后的 MDCT 系数, $x'_{2,1}$ 、 $x'_{2,2}$ 、 $x'_{2,575}$ 、 $x'_{2,576}$ 对应表示样本库中的第 i 个音频样本对应的载体估计中的第 2 帧的第 1 个、第 2 个、第 575 个、第 576 个量化后的 MDCT 系数, $x'_{N_i^f,1}$ 、 $x'_{N_i^f,2}$ 、 $x'_{N_i^f,575}$ 、 $x'_{N_i^f,576}$ 对应表示样本库中的第 i 个音频样本对应的载体估计中的第 N_i^f 帧的第 1 个、第 2 个、第 575 个、第 576 个量化后的 MDCT 系数;

[0010] ④计算样本库中的每个音频样本对应的所有量化后的 MDCT 系数构成的第一系数矩阵中的每一列与该音频样本对应的载体估计对应的所有量化后的 MDCT 系数构成的第二系数矩阵中对应列之间的 Hausdorff 距离值,将样本库中的每个音频样本对应的 576 个

Hausdorff 距离值构成一个初级隐写分析特征行向量,将样本库中的第 i 个音频样本对应的 576 个 Hausdorff 距离值构成的初级隐写分析特征行向量记为 F_i , $F_i = [hd_1 \ hd_2 \cdots hd_k \cdots hd_{575} \ hd_{576}]$, 其中, $1 \leq i \leq 2N$, hd_k 表示 X_i 中的第 k 列与 X_i' 中的第 k 列之间的 Hausdorff 距离值, $1 \leq k \leq 576$;

[0011] ⑤根据样本库中的每个音频样本对应的 576 个 Hausdorff 距离值构成的初级隐写分析特征行向量,获取样本库中的每个音频样本对应的最终隐写分析特征行向量,将样本库中的第 i 个音频样本对应的最终隐写分析特征行向量记为 F_i' , $F_i' = [\mu_i \ m_i \ \sigma_i \ p_i \ s_i]$, 其中, $1 \leq i \leq 2N$, μ_i 表示 F_i 中的所有元素的均值, $\mu_i = \frac{1}{576} \sum_{k=1}^{576} hd_k$, m_i 表示 F_i 中的

所有元素中的最大值, $m_i = \max(hd_1, hd_2, \dots, hd_k, \dots, hd_{576})$, $\max()$ 为取最大值函数, σ_i 表示 F_i 中的所有元素的方差, $\sigma_i = \frac{1}{576} \sum_{k=1}^{576} (hd_k - \mu_i)^2$, p_i 表示 F_i 中的所有元素的峰度系数,

$$p_i = \frac{\sum_{k=1}^{576} (hd_k - \mu_i)^4}{(\sigma_i)^2}, \quad s_i \text{ 表示 } F_i \text{ 中的所有元素的三阶中心距, } s_i = \frac{1}{576} \sum_{k=1}^{576} (hd_k - \mu_i)^3;$$

[0012] ⑥对样本库中的每个音频样本对应的最终隐写分析特征行向量进行标记,对于样本库中的第 i 个音频样本,如果该音频样本为 MP3 压缩音频,则将该音频样本对应的最终隐写分析特征行向量标记为 -1,如果该音频样本为隐写 MP3 压缩音频,则将该音频样本对应的最终隐写分析特征行向量标记为 +1,其中, $1 \leq i \leq 2N$;然后将所有标记后的最终隐写分析特征行向量输入 SVM 分类器进行训练,得到训练模板;

[0013] ⑦对于任意一个待检测的 MP3 压缩音频,按照步骤②至步骤⑤的过程,以相同的方式获取该 MP3 压缩音频对应的最终隐写分析特征行向量,然后利用训练模板对该 MP3 压缩音频对应的最终隐写分析特征行向量进行检测,以确定该 MP3 压缩音频是否经过 MP3Stego 音频隐写工具隐写。

[0014] 所述的步骤②中对样本库中的每个音频样本中的所有帧序号为偶数的帧的编码参数量化步长进行修正的具体过程为:

[0015] ②-1、将样本库中当前待处理的第 i 个音频样本定义为当前音频样本,其中, $1 \leq i \leq 2N$;

[0016] ②-2、假设当前音频样本中包含有 N_i^f 帧,且帧序号从 1 开始编,并将当前音频样本中当前待处理的第 j 帧定义为当前帧,其中, $1 \leq j \leq N_i^f$;

[0017] ②-3、如果当前帧的帧序号 j 为奇数,则保持当前帧的编码参数量化步长不变,然后执行步骤②-4;如果当前帧的帧序号 j 为偶数,则将前一帧的编码参数量化步长作为当前帧的编码参数量化步长,完成对当前帧的编码参数量化步长的修正,然后执行步骤②-4;

[0018] ②-4、令 $j = j+1$,然后将当前音频样本中下一帧待处理的帧作为当前帧,再返回步骤②-3 继续执行,直至当前音频样本中的所有帧处理完毕,其中, $j = j+1$ 中的“=”为赋值符号;

[0019] ②-5、令 $i = i+1$,然后将样本库中下一个待处理的音频样本作为当前音频样本,

再返回步骤②-2 继续执行,直至样本库中的所有音频样本处理完毕,其中, $i = i+1$ 中的“=”为赋值符号。

[0020] 所述的步骤④中 X_i 中的第 k 列与 X_i' 中的第 k 列之间的 Hausdorff 距离值 hd_k 的获取过程为:

[0021] ④-1、将 X_i 中的第 k 列记为 $C_{i,k}$,将 X_i' 中的第 k 列记为 $C'_{i,k}$;

[0022] ④-2、将 $C_{i,k}$ 中当前待处理的第 j 个元素定义为当前元素,其中, $1 \leq j \leq N_i^f$;

[0023] ④-3、计算当前元素与 $C'_{i,k}$ 中的每个元素之间的欧氏距离,得到当前元素对应的 N_i^f 个欧氏距离,然后获取当前元素对应的 N_i^f 个欧氏距离中的最小欧氏距离值;

[0024] ④-4、令 $j = j+1$,然后将 $C_{i,k}$ 中下一个待处理的元素作为当前元素,再返回步骤④-3 继续执行,直至 $C_{i,k}$ 中的所有元素处理完毕,得到 $C_{i,k}$ 中的每个元素对应的一个最小欧氏距离值,其中, $j = j+1$ 中的“=”为赋值符号;

[0025] ④-5、获取 $C_{i,k}$ 对应的 N_i^f 个最小欧氏距离值中的最大值,然后将该最大值作为 X_i 中的第 k 列与 X_i' 中的第 k 列之间的 Hausdorff 距离值,记为 hd_k 。

[0026] 与现有技术相比,本发明的优点在于:

[0027] 1) 本发明方法根据 MP3Stego 音频隐写工具的隐写原理,提出在每个音频样本解压缩过程中对音频样本中帧序号为偶数的帧的编码参数量化步长进行修正,从而得到对隐写后音频没隐写之前的状态的估计,这种方式能有效去除隐写痕迹,因此其与待测音频结合能够消除音频本身风格变化对隐写检测结果的干扰,从而提高检测效率,特别是在音频嵌入较少隐秘信息的情况下检测效率依然很高。

[0028] 2) 本发明方法通过计算音频样本对应的系数矩阵中的每一列与音频样本对应的载体估计对应的系数矩阵中对应列之间的 Hausdorff(豪斯多夫)距离值,来获取音频样本对应的初级隐写分析特征行向量,使得本发明方法能够敏锐检测出隐写操作对待检测音频中的每帧量化后的 MDCT 系数的微小变化,从而能够有效提高检测效率。

[0029] 3) 本发明方法根据音频样本对应的初级隐写分析特征行向量,构造五个高阶统计特征,以得到音频样本对应的最终隐写分析特征行向量,这使得本发明方法能够在较少的特征维数情况下得到较高的检测率,从而提高了检测效率。

附图说明

[0030] 图 1 为本发明方法的总体实现框图。

具体实施方式

[0031] 以下结合附图实施例对本发明作进一步详细描述。

[0032] 本发明提出的一种针对 MP3Stego 隐写后的音频的隐写检测方法,其总体实现框图如图 1 所示,其包括以下步骤:

[0033] ①选取 N 个变化风格各不相同且未压缩的 WAV 音频,其中, $N \geq 100$;然后利用 MP3 音频 8HZ 编码器对每个 WAV 音频进行压缩编码,得到每个 WAV 音频对应的 MP3 压缩音频;并利用 MP3Stego 音频隐写工具对每个 WAV 音频隐写不同长度、不同内容的隐秘信息,得到每个 WAV 音频对应的隐写 MP3 压缩音频;再将 N 个 MP3 压缩音频和 N 个隐写 MP3 压缩音频构

成一个样本库。

[0034] 在实际操作时,选取的 N 个 WAV 音频的长度既可以全部相同,也可以部分相同,也可以各不相同,即对选取的 WAV 音频的长度不作要求; N 可取值为 100;利用 MP3 音频 8HZ 编码器对每个 WAV 音频进行压缩编码,实际上可以直接利用 MP3Stego 音频隐写工具对每个 WAV 音频不隐写隐秘信息,只进行压缩编码,同样可以得到每个 WAV 音频对应的 MP3 压缩音频。

[0035] ②利用 MP3 音频 lame 解码器对样本库中的每个音频样本进行解压缩,得到样本库中的每个音频样本对应的 WAV 音频,在解压缩过程中,对样本库中的每个音频样本中的所有帧序号为偶数的帧的编码参数量化步长进行修正;然后利用 MP3 音频 lame 编码器对样本库中的每个音频样本对应的 WAV 音频进行压缩编码,得到样本库中的每个音频样本对应的载体估计。

[0036] 在此具体实施例中,步骤②中对样本库中的每个音频样本中的所有帧序号为偶数的帧的编码参数量化步长进行修正的具体过程为:

[0037] ②-1、将样本库中当前待处理的第 i 个音频样本定义为当前音频样本,其中, $1 \leq i \leq 2N$ 。

[0038] ②-2、假设当前音频样本中包含有 N_i^f 帧,且帧序号从 1 开始编,并将当前音频样本中当前待处理的第 j 帧定义为当前帧,其中, $1 \leq j \leq N_i^f$ 。

[0039] ②-3、如果当前帧的帧序号 j 为奇数,则保持当前帧的编码参数量化步长不变,然后执行步骤②-4;如果当前帧的帧序号 j 为偶数,则将当前帧的前一帧的编码参数量化步长作为当前帧的编码参数量化步长,完成对当前帧的编码参数量化步长的修正,然后执行步骤②-4。

[0040] ②-4、令 $j = j+1$,然后将当前音频样本中下一帧待处理的帧作为当前帧,再返回步骤②-3 继续执行,直至当前音频样本中的所有帧处理完毕,其中, $j = j+1$ 中的“=”为赋值符号。

[0041] ②-5、令 $i = i+1$,然后将样本库中下一个待处理的音频样本作为当前音频样本,再返回步骤②-2 继续执行,直至样本库中的所有音频样本处理完毕,其中, $i = i+1$ 中的“=”为赋值符号。

[0042] 在一个音频中,每帧的编码参数量化步长控制该帧在编码不失真情况下的精度,由于利用 MP3Stego 音频隐写工具进行隐写操作会改变每帧的编码参数量化步长,因此为了去除隐写痕迹,本发明方法对帧序号为偶数的所有帧的编码参数量化步长进行修正,即对于帧序号为偶数的帧,舍弃其本身的编码参数量化步长,复制其上一帧的编码参数量化步长进行编码。

[0043] ③利用 MP3 音频 lame 解码器对样本库中的每个音频样本进行解压缩,提取出样本库中的每个音频样本中的每帧的 576 个量化后的 MDCT 系数,以每个音频样本中的每帧的 576 个量化后的 MDCT 系数作为一行,将样本库中的每个音频样本对应的所有量化后的 MDCT 系数构成一个第一系数矩阵,将样本库中的第 i 个音频样本对应的所有量化后的 MDCT 系数

构成的第一系数矩阵记为 X_i , $X_i = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,575} & x_{1,576} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,575} & x_{2,576} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{N_i^f,1} & x_{N_i^f,2} & \cdots & x_{N_i^f,575} & x_{N_i^f,576} \end{bmatrix}$, 其中, $1 \leq i \leq 2N$, X_i

的维数为 $N_i^f \times 576$, N_i^f 表示样本库中的第 i 个音频样本中包含的帧的总数, $x_{1,1}$ 、 $x_{1,2}$ 、 $x_{1,575}$ 、 $x_{1,576}$ 对应表示样本库中的第 i 个音频样本中的第 1 帧的第 1 个、第 2 个、第 575 个、第 576 个量化后的 MDCT 系数, $x_{2,1}$ 、 $x_{2,2}$ 、 $x_{2,575}$ 、 $x_{2,576}$ 对应表示样本库中的第 i 个音频样本中的第 2 帧的第 1 个、第 2 个、第 575 个、第 576 个量化后的 MDCT 系数, $x_{N_i^f,1}$ 、 $x_{N_i^f,2}$ 、 $x_{N_i^f,575}$ 、 $x_{N_i^f,576}$ 对应表示样本库中的第 i 个音频样本中的第 N_i^f 帧的第 1 个、第 2 个、第 575 个、第 576 个量化后的 MDCT 系数。

[0044] 同样, 利用 MP3 音频 lame 解码器对样本库中的每个音频样本对应的载体估计进行解压缩, 提取出样本库中的每个音频样本对应的载体估计中的每帧的 576 个量化后的 MDCT 系数, 以每个载体估计中的每帧的 576 个量化后的 MDCT 系数作为一行, 将样本库中的每个音频样本对应的载体估计对应的所有量化后的 MDCT 系数构成一个第二系数矩阵, 将样本库中的第 i 个音频样本对应的载体估计对应的所有量化后的 MDCT 系数构成的第二

系数矩阵记为 X_i' , $X_i' = \begin{bmatrix} x'_{1,1} & x'_{1,2} & \cdots & x'_{1,575} & x'_{1,576} \\ x'_{2,1} & x'_{2,2} & \cdots & x'_{2,575} & x'_{2,576} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x'_{N_i^f,1} & x'_{N_i^f,2} & \cdots & x'_{N_i^f,575} & x'_{N_i^f,576} \end{bmatrix}$, 其中, $1 \leq i \leq 2N$, X_i' 的

维数为 $N_i^f \times 576$, N_i^f 表示样本库中的第 i 个音频样本中包含的帧的总数, 亦表示样本库中的第 i 个音频样本对应的载体估计中包含的帧的总数, $x'_{1,1}$ 、 $x'_{1,2}$ 、 $x'_{1,575}$ 、 $x'_{1,576}$ 对应表示样本库中的第 i 个音频样本对应的载体估计中的第 1 帧的第 1 个、第 2 个、第 575 个、第 576 个量化后的 MDCT 系数, $x'_{2,1}$ 、 $x'_{2,2}$ 、 $x'_{2,575}$ 、 $x'_{2,576}$ 对应表示样本库中的第 i 个音频样本对应的载体估计中的第 2 帧的第 1 个、第 2 个、第 575 个、第 576 个量化后的 MDCT 系数, $x'_{N_i^f,1}$ 、 $x'_{N_i^f,2}$ 、 $x'_{N_i^f,575}$ 、 $x'_{N_i^f,576}$ 对应表示样本库中的第 i 个音频样本对应的载体估计中的第 N_i^f 帧的第 1 个、第 2 个、第 575 个、第 576 个量化后的 MDCT 系数。

[0045] ④计算样本库中的每个音频样本对应的所有量化后的 MDCT 系数构成的第一系数矩阵中的每一列与该音频样本对应的载体估计对应的所有量化后的 MDCT 系数构成的第二系数矩阵中对应列之间的 Hausdorff 距离值, 将样本库中的每个音频样本对应的 576 个 Hausdorff 距离值构成一个初级隐写分析特征行向量, 将样本库中的第 i 个音频样本对应的 576 个 Hausdorff 距离值构成的初级隐写分析特征行向量记为 F_i , $F_i = [hd_1 \quad hd_2 \cdots hd_k \cdots hd_{575} \quad hd_{576}]$, 其中, $1 \leq i \leq 2N$, F_i 的维数为 576, hd_1 表示 X_i 中的第 1 列与 X_i' 中的第 1 列之间的 Hausdorff 距离值, hd_2 表示 X_i 中的第 2 列与 X_i' 中的第 2 列之间的 Hausdorff 距离值, hd_k 表示 X_i 中的第 k 列与 X_i' 中的第 k 列之间的 Hausdorff 距离值, hd_{575} 表示 X_i 中的第 575 列与 X_i' 中的第 575 列之间的 Hausdorff 距离值, hd_{576} 表示 X_i 中的第 576 列与 X_i' 中的第 576 列之间的 Hausdorff 距离值, $1 \leq k \leq 576$ 。

[0046] 在此具体实施例中,步骤④中 X_i 中的第 k 列与 X_i' 中的第 k 列之间的 Hausdorff 距离值 hd_k 的获取过程为:

[0047] ④-1、将 X_i 中的第 k 列记为 $C_{i,k}$,将 X_i' 中的第 k 列记为 $C'_{i,k}$ 。

[0048] ④-2、将 $C_{i,k}$ 中当前待处理的第 j 个元素定义为当前元素,其中, $1 \leq j \leq N_i^f$ 。

[0049] ④-3、计算当前元素与 $C'_{i,k}$ 中的每个元素之间的欧氏距离,得到当前元素对应的 N_i^f 个欧氏距离,然后获取当前元素对应的 N_i^f 个欧氏距离中的最小欧氏距离值。

[0050] ④-4、令 $j = j+1$,然后将 $C_{i,k}$ 中下一个待处理的元素作为当前元素,再返回步骤④-3继续执行,直至 $C_{i,k}$ 中的所有元素处理完毕,得到 $C_{i,k}$ 中的每个元素对应的一个最小欧氏距离值,其中, $j = j+1$ 中的“=”为赋值符号。

[0051] ④-5、获取 $C_{i,k}$ 对应的 N_i^f 个最小欧氏距离值中的最大值,然后将该最大值作为 X_i 中的第 k 列与 X_i' 中的第 k 列之间的 Hausdorff 距离值,记为 hd_k 。

[0052] ⑤为了减少初级隐写分析特征行向量的维数,及提高检测效果,本发明方法将步骤④中得到的 576 个元素的五个高阶统计量作为最终隐写分析特征行向量,即根据样本库中的每个音频样本对应的 576 个 Hausdorff 距离值构成的初级隐写分析特征行向量,获取样本库中的每个音频样本对应的最终隐写分析特征行向量,将样本库中的第 i 个音频样本对应的最终隐写分析特征行向量记为 F_i' , $F_i' = [\mu_i \ m_i \ \sigma_i \ p_i \ s_i]$,其中, $1 \leq i \leq 2N$, F_i' 的

维数为 5, μ_i 表示 F_i 中的所有元素的均值, $\mu_i = \frac{1}{576} \sum_{k=1}^{576} hd_k$, m_i 表示 F_i 中的所有元素中的最大值, $m_i = \max(hd_1, hd_2, \dots, hd_k, \dots, hd_{576})$, $\max()$ 为取最大值函数, σ_i 表示 F_i 中的所有元

素的方差, $\sigma_i = \frac{1}{576} \sum_{k=1}^{576} (hd_k - \mu_i)^2$, p_i 表示 F_i 中的所有元素的峰度系数, $p_i = \frac{\sum_{k=1}^{576} (hd_k - \mu_i)^4}{(\sigma_i)^2}$,

s_i 表示 F_i 中的所有元素的三阶中心距, $s_i = \frac{1}{576} \sum_{k=1}^{576} (hd_k - \mu_i)^3$ 。

[0053] ⑥对样本库中的每个音频样本对应的最终隐写分析特征行向量进行标记,对于样本库中的第 i 个音频样本,如果该音频样本为 MP3 压缩音频,则将该音频样本对应的最终隐写分析特征行向量标记为 -1,如果该音频样本为隐写 MP3 压缩音频,则将该音频样本对应的最终隐写分析特征行向量标记为 +1,其中, $1 \leq i \leq 2N$;然后将所有标记后的最终隐写分析特征行向量输入 SVM 分类器进行训练,得到训练模板。

[0054] ⑦对于任意一个待检测的 MP3 压缩音频,按照步骤②至步骤⑤的过程,以相同的方式获取该 MP3 压缩音频对应的最终隐写分析特征行向量,即利用 MP3 音频 lame 解码器对该 MP3 压缩音频进行解压缩,得到该 MP3 压缩音频对应的 WAV 音频,在解压缩过程中,对该 MP3 压缩音频中的所有帧序号为偶数的帧的编码参数量化步长进行修正,然后利用 MP3 音频 lame 编码器对该 MP3 压缩音频对应的 WAV 音频进行压缩编码,得到该 MP3 压缩音频对应的载体估计,利用 MP3 音频 lame 解码器对该 MP3 压缩音频进行解压缩,提取出该 MP3 压缩音频中的每帧的 576 个量化后的 MDCT 系数,以该 MP3 压缩音频中的每帧的 576 个量化后的 MDCT 系数作为一行,将该 MP3 压缩音频对应的所有量化后的 MDCT 系数构成一个第一系数矩阵,同样利用 MP3 音频 lame 解码器对该 MP3 压缩音频对应的载体估计进行解压缩,提取出

该 MP3 压缩音频对应的载体估计中的每帧的 576 个量化后的 MDCT 系数,以该 MP3 压缩音频载体估计中的每帧的 576 个量化后的 MDCT 系数作为一行,将该 MP3 压缩音频对应的载体估计对应的所有量化后的 MDCT 系数构成一个第二系数矩阵,计算该 MP3 压缩音频对应的所有量化后的 MDCT 系数构成的第一系数矩阵中的每一列与该 MP3 压缩音频对应的载体估计对应的所有量化后的 MDCT 系数构成的第二系数矩阵中对应列之间的 Hausdorff 距离值,将该 MP3 压缩音频对应的 576 个 Hausdorff 距离值构成一个初级隐写分析特征行向量,根据该 MP3 压缩音频对应的 576 个 Hausdorff 距离值构成的初级隐写分析特征行向量,获取该 MP3 压缩音频对应的最终隐写分析特征行向量;然后利用训练模板对该 MP3 压缩音频对应的最终隐写分析特征行向量进行检测,以确定该 MP3 压缩音频是否经过 MP3Stego 音频隐写工具隐写。

[0055] 为更好地说明本发明方法的可行性和有效性,对本发明方法进行实验验证。

[0056] 一次隐写分析的结果可能会出现以下四种情况的其中一种:1) 待检测的 MP3 压缩音频为含密载体,隐写检测方法的判断结果也为含密载体,属于真阳性 (True Positive); 2) 待检测的 MP3 压缩音频为原始载体,隐写检测方法判断其为含密载体,属于假阳性,虚警 (False Positive); 3) 待检测的 MP3 压缩音频为原始载体,隐写检测方法的判断结果也为原始载体,属于真阴性 (True Negative); 4) 待检测的 MP3 压缩音频为含密载体,隐写检测方法判断其为原始载体,属于假阴性,误警 (False Negative)。

[0057] 在实验中,任意选取风格变化较大的 50 个未隐写的 MP3 压缩音频和 50 个经 MP3Stego 音频隐写工具一次隐写最小文本长度的隐秘信息后的隐写 MP3 压缩音频,然后利用本发明方法的步骤②至步骤⑤的过程,以相同的方式分别获得这 100 个音频各自对应的最终隐写分析特征行向量,再利用本发明方法的步骤⑥获得的训练模板分别对这 100 个音频进行检测,检测结果为:属于真阳性有 49 个音频,属于假阳性的有 0 个音频,属于真阴性的有 50 个音频,属于假阴性的有 1 个音频,检测正确率如表 1 所列,从表 1 中可以看出,本发明方法的检测正确率很好,充分说明了本发明方法的可行性和有效性。。

[0058] 表 1 对 50 个未隐写的 MP3 压缩音频和 50 个经 MP3Stego 音频隐写工具一次隐写后的隐写 MP3 压缩音频的检测结果的比较

[0059]

	隐写音频	未隐写音频	检测正确率	虚警率	误警率
实际结果	50	50	/	/	/
检测结果	49	51	99%	0	2%

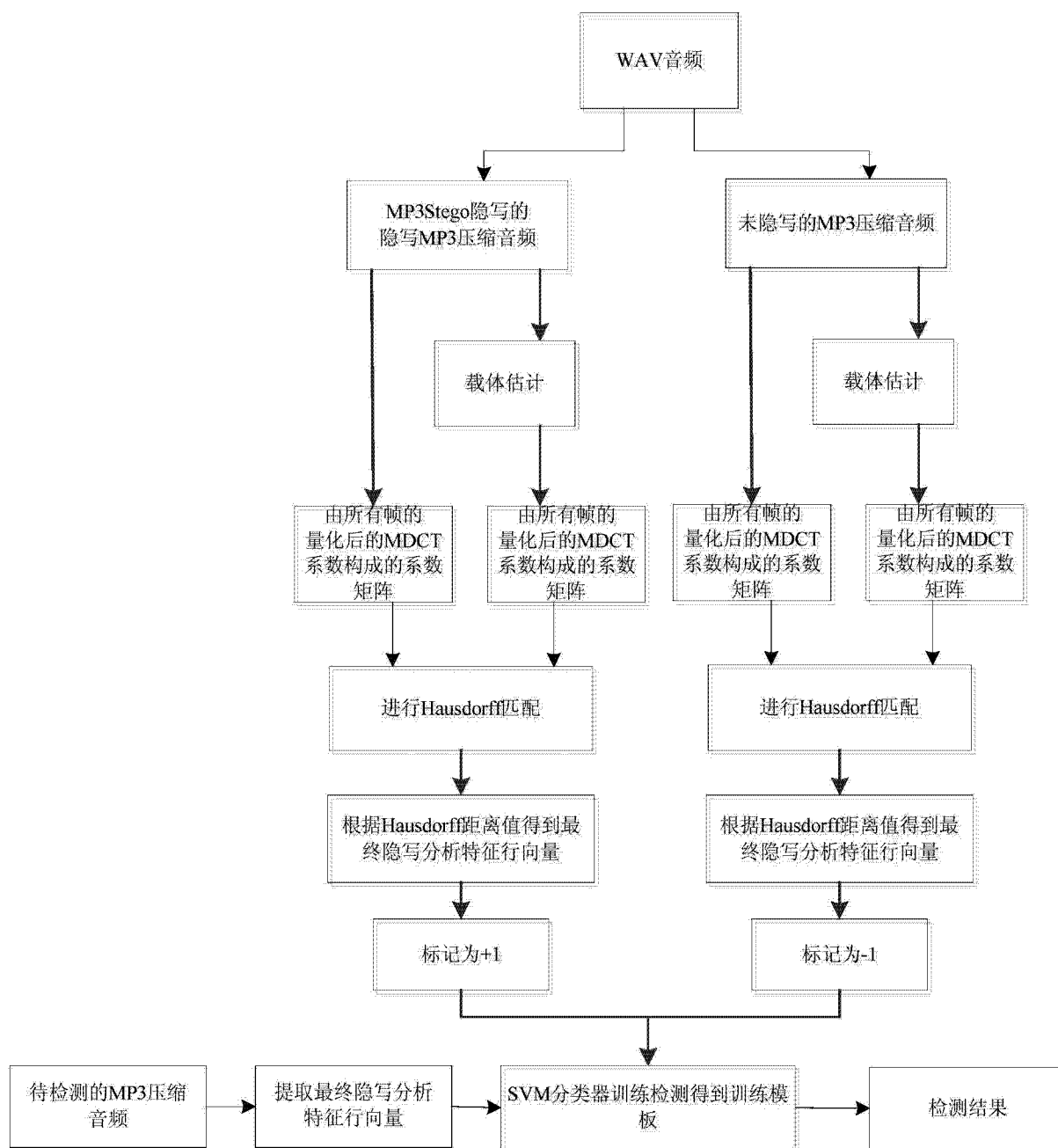


图 1