结合密钥和随机标准正交基的音频伪装方案

邵京津,邵利平,任平安
(陕西师范大学 计算机科学学院,陕西 西安 710119)

摘 要:传统基于 Tangram 的音频伪装方法所采用的变换模型为仿射变换模型,变换精度低且不满足基本的正交关系,从而无法保证秘密音频与公开音频之间的拟合精度,同时当分段变换音频为恒值序列时,需添加随机扰动以保证变换后音频的恢复质量,由此会降低信道传输音频的听觉质量。针对此问题,提出一种结合密钥和随机标准正交基的音频伪装方法。首先对秘密音频和公开音频分段,利用密钥构造随机标准正交基;其次通过秘密音频小段在随机标准正交基上的投影来对秘密音频小段进行表达,从中选取包括均值幅值较大的前 k 个投影系数,并记录对应的索引位置;再次通过 EMD-q 密写方法嵌入到对应的公开音频小段中形成信道公开传输音频;最后通过信道公开传输音频提取的变换参数结合密钥重构秘密音频。实验表明,所提方法可充分利用随机标准正交基重构不同精度的秘密音频,且随着选取的幅值系数增多,恢复的秘密音频质量也越来越好,同时所述策略严格依赖于密钥,只有掌握正确密钥的用户才能进行高精度的重构。

关键词:施密特正交化;标准正交基;音频伪装;拟合;变换
中图分类号:TP309.7 **文献标识码**:A **文章编号**:1673-629X(2020)01-0087-07
doi:10.3969/j.issn.1673-629X.2020.01.016

Audio Camouflage Scheme Combining Key and Random Orthonormal Basis

SHAO Jing-jin, SHAO Li-ping, REN Ping-an
(School of Computer Science, Shaanxi Normal University, Xi'an 710119, China)

Abstract: The transformation model adopted in traditional tangram based audio camouflage method is the affine transformation model, where its transformation accuracy is low and doesn't meet the basic orthogonal relationship. It cannot be guaranteed the fitting accuracy between the secret audio and the public audio in affine transformation model. At the same time when the segment of transforming audio is constant value sequence, it must be added random disturbance to ensure the quality of audio recovery, which will reduce the audio's quality transmitted in channel. To solve these problems, an audio camouflage method combining key and random orthonormal basis is proposed. The secret audio and public audio are firstly segmented, and a key is used to construct the random orthonormal basis. Secondly, the secret audio segment is effectively expressed by the different projections of secret audio segment on the random orthonormal bases. The first k projection coefficients with larger amplitude included mean value are selected and their corresponding index positions are also recorded. Then the EMD-q steganography is employed to embed them into the corresponding public audio segment to form the public audio transmitted in the channel. Finally, the secret audio can be reconstructed by the transformation parameters extracted from the transmitted audio and the key. Experiment shows that the proposed method can make full use of the random orthonormal bases to reconstruct the secret audio with varying precision degrees, and the more larger amplitude coefficients are selected, the higher quality of the recovered secret audio will be. Simultaneously, the strategy described relies strictly on the key and only the user who masters the correct key can get the recovered secret audio in high quality.

Key words: Schmidt orthogonalization; orthonormal basis; audio camouflage; fitting; transformation

1 概 述

针对图像音频信息安全,人们已提出多种安全保护方法,如加密[1-3]、分存[4]、密写[5]和伪装等。相对于其他图像音频保护方法,伪装是将机密图像音频伪

收稿日期:2019-02-13 **修回日期**:2019-06-14 **网络出版时间**:2019-09-24
基金项目:国家自然科学基金(61100239);陕西省自然科学基金(2011JQ8009,2016JM6065);中央高校基本科研业务费支持项目(GK201402036,GK201703057)
作者简介:邵京津(1993-),女,硕士研究生,研究方向为音频图像伪装;邵利平,博士,副教授,研究生导师,CCF 会员(11901M),通信作者,研究方向为数字图像音频置乱、加密、密写、水印、隐匿、分存、伪装和欺骗等;任平安,副教授,研究生导师,研究方向为计算机网络安全。
网络出版地址:http://kns.cnki.net/kcms/detail/61.1450.TP.20190924.1534.004.html

装成有意义的非机密图像音频,从而在传输时不易引起攻击者的注意,减少潜在攻击的可能性。

Tangram 方法,也称七巧法或中国拼图方法[6-7],是一种典型的基于变换的图像伪装方法。但传统 Tangram 方法在匹配过程中需全局搜索,计算复杂度高,等距变换数量少,制约了将密图转变为公开图像的匹配精度。

为降低 Tangram 方法的搜索时间和加快编码速度,文献[8]添加了块均化操作,在减小计算代价的同时也减少了等距变换数量,导致匹配精度降低。为进一步降低搜索代价,提高编码速度,文献[9]将三角剖分用于对图像三角区域近似重建,尽管免除了全局搜索,但也降低了密图重构精度。文献[10-12]利用 2 维双尺度矩形映射确定密图子块和公开图像子块间的对应关系,通过直接最小 2 乘法匹配避免了全局匹配,编码代价远低于传统 Tangram 方法,但依然仅提供有限的等距变换,导伪装图像视觉质量不高。

Tangram 方法也被进一步拓展为音频伪装方法。文献[13-14]分别通过图像子块和音频小段序列构造旋转向量来增加正交等距变换的数量,在提高匹配精度和公开载体伪装质量的同时保证了秘密信息的重构精度,但等距变换所能提供的元素组合十分有限且非全部组合,因此无法找到所有向量元素组合的最优解,同时由于向量旋转所产生的多个等距变换向量都需进行最小 2 乘匹配,从而提高了计算代价。为减少等距变换数量,文献[15]引入排序线拟合使得排序后公开音频元素和秘密音频划分的小段序列变化趋势保持一致,在避免高昂匹配代价的同时,也提高了拟合匹配精度。

但无论是标准 Tangram 算法[6-7],还是 Tangram 的改进算法[8,10-15],这些方法所基于的变换模型都是仿射变换模型。对于图像,仿射变换模型只有均值块和差异块,对于音频,只有均值向量和差异向量,且无论是均值块和差异块还是均值向量和差异向量都不满足基本的正交关系,导致变换精度低,不能有效地保证拟合精度,从而无法有效地进行信道欺骗和保证秘密图像音频的重构精度。除此以外,对于恒值块或恒值序列,文献[13-15]需添加扰动来改善匹配性能,由此进一步限制了仿射变换模型的变换精度。对于文献[9],只能对密图三角剖分区域进行近似重构,其实际应用价值较小。

针对以上问题,文中首先将秘密音频和公开音频划分为同等数量的小段序列,利用密钥来构造随机标准正交基;其次通过求取秘密音频小段序列在随机标准正交基上的投影来对秘密音频小段序列进行充分有效的线性表达,从中选取幅值和能量较大且包含均值的前 k 个投影系数来表达秘密信息并记录对应的索引位置;再次将选定的投影系数和索引位置序列通过 EMD-q 嵌入方法嵌入到与之对应的公开音频小段序列中,从而形成信道公开传输音频;最后通过公开传输音频提取的变换参数并结合密钥来对秘密音频重构。同现有方法相比,所提方法可实现秘密音频不同精度的重构并严格依赖于用户密钥,只有掌握正确密钥的用户才能进行高精度的重构。

2 所提方法

文献[6-15]基于的仿射变换模型仅包含均值向量和差异向量,且不满足基本的正交关系,导致拟合精度较低;文献[9]仅能对密图三角剖分区域进行近似重构,导致其实际应用价值较小;文献[13-15]添加扰动来改善恒值块或恒值序列的匹配性能,导致变换精度降低。针对上述问题,文中将秘密音频小段与密钥产生的随机标准正交基进行拟合投影,记录幅值较大系数及其对应的位置索引为变换参数,再通过 EMD-q 密写方法将变换参数嵌入到公开音频小段中,不仅解决了均值向量和差异向量不满足基本的正交关系所带来的拟合精度较低的问题,还避免了恒值块和恒值序列加噪处理,且所提方法可根据实际需要选取多个正交基对秘密音频进行表达,从而可实现秘密音频不同精度的重构且所提方法严格依赖于密钥,只有特定用户才能对秘密音频进行高精度重构。

2.1 嵌密阶段

记秘密音频和公开音频分别为长度为 $l \cdot n$ 和 $l \cdot N$ 的 r 位音频序列 $\boldsymbol{S}=(s_i)_{l\cdot n}$ 和 $\boldsymbol{P}=(p_i)_{l\cdot N}$,即 $s_i,p_i \in \{-2^{r-1},\cdots,0,\cdots,2^{r-1}-1\}$,将 $\boldsymbol{S}$ 和 $\boldsymbol{P}$ 分别划分为长度为 n 和 N 大小的小段序列,记为 $\boldsymbol{S}_u=(s_i^u)_n$ 和 $\boldsymbol{P}_u=(p_i^u)_N$,其中 $u=0,1,\cdots,l-1$。由密钥 Key 生成随机序列 $\boldsymbol{G}=(g_u)_l$,初始化 $u=0$。

将 g_u 作为密钥,生成 $n\times n$ 维随机矩阵 $\boldsymbol{X}'$,将 $\boldsymbol{X}'$ 的第 0 行所有元素置为 1,按式 1 对 $\boldsymbol{X}'$ 进行行标准施密特正交化,即将 $\boldsymbol{X}'$ 的每一行转换为标准正交向量:

$$\boldsymbol{X}=\text{Schimidt}(\boldsymbol{X}') \tag{1}$$

其中,函数 Schimidt( ) 为行标准施密特正交化函数。

记 $\boldsymbol{X}$ 对应的 n 个行依次为 $\boldsymbol{X}_f=(x_i^f)$,$f=0,1,\cdots,n-1$,按式 2 计算 $\boldsymbol{S}_u$ 在标准正交基 $\boldsymbol{X}_f$ 上的投影 $\hat{\alpha}_f$,$f=0,1,\cdots,n-1$。

$$\hat{\alpha}_f=\min_{\beta_f}\|\boldsymbol{S}_u-\beta_f\boldsymbol{X}_f\|_2^2 \tag{2}$$

其中,“ $\|\ \|_2$ ”为 2 范数,系数 $\hat{\alpha}_f$ 可按式 3 进行计算。

$$\hat{\alpha}_f=\sum_{i=0}^{n-1}x_i^f\cdot s_i^u \tag{3}$$

从 $\hat{\alpha}_f, f=0,1,\cdots,n-1$ 中可找到除 $\hat{\alpha}_0$ 的前 $k-1$ 个最大幅值系数,记为 $\alpha_0,\alpha_1,\cdots,\alpha_{k-2}$,通过长度为 $k-1$ 的索引序列 $\boldsymbol{R}=(r_i)_{k-1}, r_i \in \{0,1,\cdots,n-1\}$ 来记录 $\alpha_0,\cdots,\alpha_{k-2}$ 在 $\hat{\alpha}_f, f=0,1,\cdots,n-1$ 上对应的行索引,然后将 $\hat{\alpha}_0$ 记为 α_{k-1}。

记 l_r 为 r_i 对应的 q 进制数序列长度,由式 4 确定:

$$l_r = \lceil \log_q n \rceil \tag{4}$$

将 $\alpha_i, i=0,1,\cdots,k-1$ 进行 q 进制数表示的具体方法是:将 α_{k-1} 按式 5 转换为 10 进制数 α''_{k-1},l_m 是其对应的 q 进制数序列长度,由式 6 确定:

$$\alpha''_{k-1} = \lceil \alpha_{k-1}/\sqrt{n} \rceil + 2^{r-1} \tag{5}$$

其中,"⌈ ⌋"为四舍五入取整函数。

$$l_m = \lceil \log_q 2^r \rceil \tag{6}$$

将 $\alpha_i, i=0,1,\cdots,k-2$ 借助 $\alpha^i_{\text{sign}},\alpha^i_{\text{pow}},\alpha^i_{\text{int}}$ 进行近似表示,如式 7 所示,然后转换为 q 进制数表示:

$$\alpha_i \cong (-1)^{\alpha^i_{\text{sign}}} \times \alpha^i_{\text{int}} \times 10^{-\alpha^i_{\text{pow}}-2} \tag{7}$$

其中,α^i_{sign} 为 α_i 的符号部分;$\alpha^i_{\text{pow}},\alpha^i_{\text{int}}$ 分别为 α_i 的幂次和有效数字。

$\alpha^i_{\text{sign}},\alpha^i_{\text{pow}},\alpha^i_{\text{int}}$ 具体的确定方法如式 8 ~ 式 10 所示。

$$\alpha^i_{\text{sign}} = \begin{cases} 0, \alpha_i \geqslant 0 \\ 1, \alpha_i < 0 \end{cases} \tag{8}$$

$$\alpha^i_{\text{pow}} = \begin{cases} q-1 & 0 \leqslant |\alpha_i| < 1 \\ q-2 & 10^0 \leqslant |\alpha_i| < 10^1 \\ \cdots & \cdots \\ 0 & 10^{q-2} \leqslant |\alpha_i| < 10^{q-1} \end{cases} \tag{9}$$

$$\alpha^i_{\text{int}} = \begin{cases} \lfloor \alpha_i \times 10^{q+1} + \frac{1}{2} \rfloor & 0 \leqslant |\alpha_i| < 1 \\ \lfloor \alpha_i \times 10^{q} + \frac{1}{2} \rfloor & 10^0 \leqslant |\alpha_i| < 10^1 \\ \cdots & \cdots \\ \lfloor \alpha_i \times 10^{2} + \frac{1}{2} \rfloor & 10^{q-2} \leqslant |\alpha_i| < 10^{q-1} \end{cases} \tag{10}$$

其中,α^i_{sign} 可用 1 位 q 进制数进行存储,其中 0 对应为正数,1 对应为负数,记对应的长度为 l_{sign};α^i_{pow} 为 1 位 q 进制数,记其对应的长度为 l_{pow};α^i_{int} 直接映射为 q 进制数,然后用 l_{int} 位 q 进制数进行存储。l_{pow} 满足的约束如式 11 所示:

$$l_{\text{pow}} \leqslant \lfloor (N - (l_r + l_{\text{sign}} + l_{\text{int}}) \times (k-1) - l_m)/(k-1) \rfloor \tag{11}$$

在满足式 11 约束的前提下,可按式 12 将 $\boldsymbol{R}=(r_i)_{k-1}$ 和 q 进制数表示的 $\alpha_i, i=0,1,\cdots,k-1$ 嵌入到 $\boldsymbol{P}_u=(p^u_i)_N$ 中,从而将 $\boldsymbol{P}_u$ 转换为 $\boldsymbol{P}'_u=(p'^u_i)_N$,其中 $i=0,1,\cdots,k-2$,$\boldsymbol{P}_u[0,l_m-1]$,$\boldsymbol{P}'_u[0,l_m-1]$ 分别表示 $\boldsymbol{P}_u$,$\boldsymbol{P}'_u$ 索引位置位于 0 到 l_m-1 之间的元素所构成的元素序列。

$$\begin{aligned}
&\boldsymbol{P}'_u[0,l_m-1] = \text{EMD}-q(\alpha''_{k-1}, \boldsymbol{P}_u[0,l_m-1]) \\
&\boldsymbol{P}'_u[l_m+il_r, l_m+(i+1)l_r-1] = \\
&\text{EMD}-q(r_i, \boldsymbol{P}_u[l_m+il_r, l_m+(i+1)l_r-1]) \\
&\boldsymbol{P}'_u[l_m+(k-1)l_r+il_{\text{sign}}, l_m+(k-1)l_r+(i+1)l_{\text{sign}}-1] = \\
&\text{EMD}-q(\alpha^i_{\text{sign}}, \boldsymbol{P}_u[l_m+(k-1)l_r+il_{\text{sign}}, l_m+(k-1)l_r+(i+1)l_{\text{sign}}-1]) \\
&\boldsymbol{P}'_u[(k-1)(l_r+l_{\text{sign}})+l_m+il_{\text{pow}}, (k-1)(l_r+l_{\text{sign}})+l_m+(i+1)l_{\text{pow}}-1] = \\
&\text{EMD}-q(\alpha^i_{\text{pow}}, \boldsymbol{P}_u[(k-1)(l_r+l_{\text{sign}})+l_m+il_{\text{pow}}, (k-1)(l_r+l_{\text{sign}})+l_m+(i+1)l_{\text{pow}}-1]) \\
&\boldsymbol{P}_u[l_m+(k-1)(l_r+l_{\text{sign}}+l_{\text{pow}})+il_{\text{int}}, l_m+(k-1)(l_r+l_{\text{sign}}+l_{\text{pow}})+(i+1)l_{\text{int}}-1] = \\
&\text{EMD}-q(\alpha^i_{\text{int}}, \boldsymbol{P}'_u[l_m+(k-1)(l_r+l_{\text{sign}}+l_{\text{pow}})+il_{\text{int}}, l_m+(k-1)(l_r+l_{\text{sign}}+l_{\text{pow}})+(i+1)l_{\text{int}}-1])
\end{aligned} \tag{12}$$

其中,函数 EMD-q() 即为 EMD-q 全方位扩展嵌入函数,其执行的功能是在长度为 d 的 10 进制数所构成的元素序列中,通过对每个元素进行 $\pm q/2$ 范围内的调整嵌入 d 位 q 进制数,其具体定义见文献[15]。

将所有的 $\boldsymbol{P}'_u, u=0,1,\cdots,l-1$ 连接在一起,即可得到信道传输音频 $\boldsymbol{P}'=(p'_i)_{l\cdot N}$。

2.2　恢复阶段

当接收到 $\boldsymbol{P}'=(p'_i)_{l\cdot N}$ 时,可将 $\boldsymbol{P}'$ 划分成长度为 N 的小段序列 $\boldsymbol{P}'=(\boldsymbol{P}_u)_l$,$\boldsymbol{P}_u=(p^u_i)_N$,并进一步从 $\boldsymbol{P}_u$ 中提取出隐藏的变换参数 $R=(r_i)_{k-1}$ 和 q 进制数表示的 $\alpha_i, i=0,1,\cdots,k-1$,其具体方法是:

首先按式 13 从 $\boldsymbol{P}_u=(p^u_i)_N$ 中提取出隐藏的变换参数 α''_{k-1}、$R=(r_i)_{k-1}$:

$$\begin{aligned}
&\alpha''_{k-1} = \text{EMD}-q^{-1}(\boldsymbol{P}_u[0,l_m-1]) \\
&r_i = \text{EMD}-q^{-1}(\boldsymbol{P}_u[l_m+il_r, l_m+(i+1)l_r-1])
\end{aligned} \tag{13}$$

其中,EMD-q^{-1}() 为 EMD-q() 对应的恢复函数,用于从长度为 d 的 10 进制数所构成的元素序列中提取出 d 位 q 进制数,其具体定义见文献[15]。

其次按式 14 将 α''_{k-1} 转变为 α_{k-1}:

$$\alpha_{k-1} = (\alpha''_{k-1} - 2^{r-1}) \cdot \sqrt{n} \tag{14}$$

最后由式 15 中提取出 $\alpha^i_{\text{pow}},\alpha^i_{\text{int}},\alpha^i_{\text{sign}}, i=0,1,\cdots,k-2$ 并按式 7 计算 $\alpha_0,\alpha_1,\cdots,\alpha_{k-2}$:

$$\alpha^i_{\text{sign}} = \text{EMD}-q^{-1}(\boldsymbol{P}_u[l_m+(k-1)l_r+il_{\text{sign}}, l_m+(k-1)l_r+(i+1)l_{\text{sign}}-1])$$

$$\alpha_{\text{pow}}^{i} = \text{EMD} - q^{-1}(\boldsymbol{P}_u[l_m + (k-1)(l_r + l_{\text{sign}}) + il_{\text{pow}}, l_m + (k-1)(l_r + l_{\text{sign}}) + (i+1)l_{\text{pow}} - 1])$$

$$\alpha_{\text{int}}^{i} = \text{EMD} - q^{-1}(\boldsymbol{P}_u[l_m + (k-1)(l_r + l_{\text{sign}} + l_{\text{pow}}) + il_{\text{int}}, l_m + (k-1)(l_r + l_{\text{sign}} + l_{\text{pow}}) + (i+1)l_{\text{int}} - 1]) \tag{15}$$

当得到了 $\alpha_0, \alpha_1, \cdots, \alpha_{k-1}$ 和 $\boldsymbol{R} = (r_i)_{k-1}$，就可进一步利用密钥 Key 生成随机序列 $\boldsymbol{G} = (g_u)_l$ 生成的标准正交矩阵 $\boldsymbol{X}$ 来重构秘密音频，其具体方法是：

首先由密钥 Key 生成随机序列 $\boldsymbol{G} = (g_u)_l$，将 g_u 作为密钥，生成 $n \times n$ 的随机矩阵 $\boldsymbol{X}'$，将 $\boldsymbol{X}'$ 的第 0 行置为 1，对 $\boldsymbol{X}'$ 按式 1 进行行标准正交化后得到 $\boldsymbol{X}$。

其次利用序列 $\boldsymbol{R} = (r_i)_{k-1}$ 从矩阵 $\boldsymbol{X}$ 中选取对应的行 $\boldsymbol{X}_{r_i}$，得到 $k-1$ 个 1 维序列，记为 $\boldsymbol{Y}_i, i = 0, 1, \cdots, k-2$，并将矩阵 $\boldsymbol{X}$ 中 $\boldsymbol{X}_0$ 记为 $\boldsymbol{Y}_{k-1}$。

最后按式 16 将 $\boldsymbol{Y}_i, \alpha_i, i = 0, 1, \cdots, k-1$ 转换为秘密音频对应的小段序列 $\boldsymbol{S}_u$：

$$\boldsymbol{S}_u = \left[\sum_{i=0}^{k-1} \alpha_i \cdot \boldsymbol{Y}_i\right] \tag{16}$$

3 完整的结合密钥和随机标准正交基的音频伪装与恢复算法

结合第 2 节的工作，以下给出完整的结合密钥和随机标准正交基的音频伪装与恢复算法，记为算法 1 和算法 2。

算法 1：结合密钥和随机标准正交基的音频伪装算法。

(1)将长度为 $l \cdot n$ 的秘密音频 $\boldsymbol{S}$ 和长度为 $l \cdot N$ 的公开音频 $\boldsymbol{P}$ 分别划分为长度为 n 和 N 的小段序列 $\boldsymbol{S}_u$ 和 $\boldsymbol{P}_u$，由密钥 Key 生成随机序列 $\boldsymbol{G} = (g_u)_l$，初始化 $u = 0$；

(2)将 g_u 作为密钥，生成 $n \times n$ 的随机矩阵 $\boldsymbol{X}'$，将 $\boldsymbol{X}'$ 的第 0 行所有元素置为 1，然后按式 1 对 $\boldsymbol{X}'$ 进行行标准施密特正交化后得到 $\boldsymbol{X}$；

(3)对 $\boldsymbol{X}$ 的行 $\boldsymbol{X}_f = (x_i^f)_n, f = 0, 1, \cdots, n-1$，按式 2 计算 $\boldsymbol{S}_u$ 在标准正交基 $\boldsymbol{X}_f$ 上的投影 $\hat{\alpha}_f$；

(4)从 $\hat{\alpha}_f, f = 0, 1, \cdots, n-1$ 中找到除 $\hat{\alpha}_0$ 的前 $k-1$ 个最大幅值系数 $\alpha_0, \alpha_1, \cdots, \alpha_{k-2}$，通过长度为 $k-1$ 的索引序列 $\boldsymbol{R} = (r_i)_{k-1}$ 来记录 $\alpha_0, \cdots, \alpha_{k-2}$ 在 $\hat{\alpha}_f, f = 0, 1, \cdots, n-1$ 上对应的行索引，然后将 $\hat{\alpha}_0$ 记为 α_{k-1}；

(5)将 r_i 通过 l_r 位 q 进制数进行表达；

(6)按式 7 将 $\alpha_i, i = 0, 1, \cdots, k-1$ 进行 q 进制数表示；

(7)将 $\boldsymbol{R} = (r_i)_{k-1}$ 和 q 进制数表示的 $\alpha_i, i = 0, 1, \cdots, k-1$ 按式 12 嵌入到 $\boldsymbol{P}_u = (p_i^u)_N$ 中，从而将 $\boldsymbol{P}_u$ 转换为 $\boldsymbol{P}_u' = (p_i'^u)_N$，置 $u = u + 1$；

(8)反复执行第 2 步到第 7 步，直至 $u = l$，然后将所有的 $\boldsymbol{P}_u', u = 0, 1, \cdots, l-1$ 依次连接作为信道传输音频 $\boldsymbol{P}'$。

算法 2：结合密钥和随机标准正交基的音频恢复算法。

(1)输入接收到的信道传输音频 $\boldsymbol{P}' = (p_i')_{l \cdot N}$，将 $\boldsymbol{P}'$ 划分成长度为 N 的小段序列 $\boldsymbol{P}' = (\boldsymbol{P}_u)_l$，$\boldsymbol{P}_u = (p_i^u)_N$，由密钥 Key 生成随机序列 $G = (g_u)_l$，初始化 $u = 0$；

(2)将 g_u 作为密钥，依次生成 $n \times n$ 的随机矩阵 $\boldsymbol{X}'$，将 $\boldsymbol{X}'$ 的第 0 行置为 1，对 $\boldsymbol{X}'$ 按式 1 进行行标准正交化后记为 $\boldsymbol{X}$；

(3)从 $\boldsymbol{P}_u$ 中按式 13 和式 15 提取出隐藏变换参数 $\boldsymbol{R} = (r_i)_{k-1}$ 和 q 进制数表示的 $\alpha_i, i = 0, 1, \cdots, k-1$，并进一步将其转换为 10 进制数 $\alpha_i, i = 0, 1, \cdots, k-1$，然后利用这些恢复出的变换参数按式 16 重构秘密音频小段序列 $\boldsymbol{S}_u$，置 $u = u + 1$；

(4)重复第 2 到 3 步，直至 $u = l$，然后将所有的小块 $\boldsymbol{S}_u, u = 0, 1, \cdots, l-1$ 依次拼接作为解密后的秘密音频 $\boldsymbol{S}$ 输出。

同文献[6-8,10-15]采用的仿射变换模型相比，所提方法的变换精度更高，且随着选取的幅值系数增多，恢复的秘密音频的听觉质量也越来越好；所提方法也避免了文献[13-15]添加扰动导致的变换精度降低，且所提方法严格依赖于密钥，只有特定用户才能对秘密音频进行高精度的重构。

4 实验与结果分析

以下对所提策略进行实验验证，操作系统为 Windows 10，CPU 为 Intel(R) Core(TM) i5-6600 4 核 CPU，内存为 8.00 GB，编码语言为 JAVA jdk1.8.0_65。

测试音频由百度随机搜索的 WAV 音频通过 Audacity 软件转码得到，为采样频率为 44 100 Hz，单声道 16 位波形音频：告白气球和倾尽天下，并依次编号为 A、B。采用信噪比(SNR)和方差(σ)来衡量音频差异和听觉质量。其中 SNR 按式 17 计算，σ 按式 18 计算。

$$\text{SNR} = 10\lg \frac{\sum_{i=0}^{d-1} v_i^2}{\sum_{i=0}^{d-1} (v_i - v_i')^2} \tag{17}$$

$$\sigma = \sqrt{\frac{1}{d-1} \sum_{i=0}^{d-1} (v_i - v_i')^2} \tag{18}$$

其中，d 为音频长度；v_i, v_i' 分别为原音频变动前后对应的第 i 个采样值。

为验证所提方法的有效性，首先将秘密音频和公开音频(见图 1)截取对应长度的采样数据按算法 1 进

行嵌入,并按算法 2 恢复秘密音频。

(a)秘密音频 A:告白气球

(b)公开音频 B:倾尽天下

图 1　实验测试音频

图 2 给出了实验测试音频波形图,与之对应的实验参数和测试结果衡量如表 1 所示。

表 1　不同划分小段序列长度实验测试参数和实验测试结果

编号	k	小段长度	嵌密听觉质量			重构听觉质量		
			嵌密音频	SNR/dB	方差	重构音频	SNR/dB	方差
1	3	8	图 2(a)	68.26	1.39	图 2(e)	18.73	1 429.81
2	5	9	图 2(b)	66.83	1.36	图 2(f)	21.56	920.17
3	8	10	图 2(c)	64.58	1.32	图 2(g)	28.83	397.86
4	9	11	图 2(d)	63.54	1.31	图 2(h)	27.33	473.25

(a)嵌密音频 1

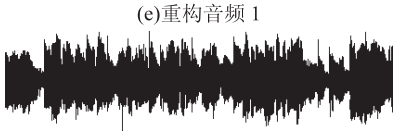
(b)嵌密音频 2

(c)嵌密音频 3

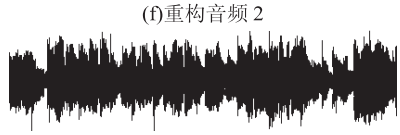
(d)嵌密音频 4

(e)重构音频 1

(f)重构音频 2



(g)重构音频 3



(h)重构音频 4

图 2　不同划分小段序列长度实验测试音频

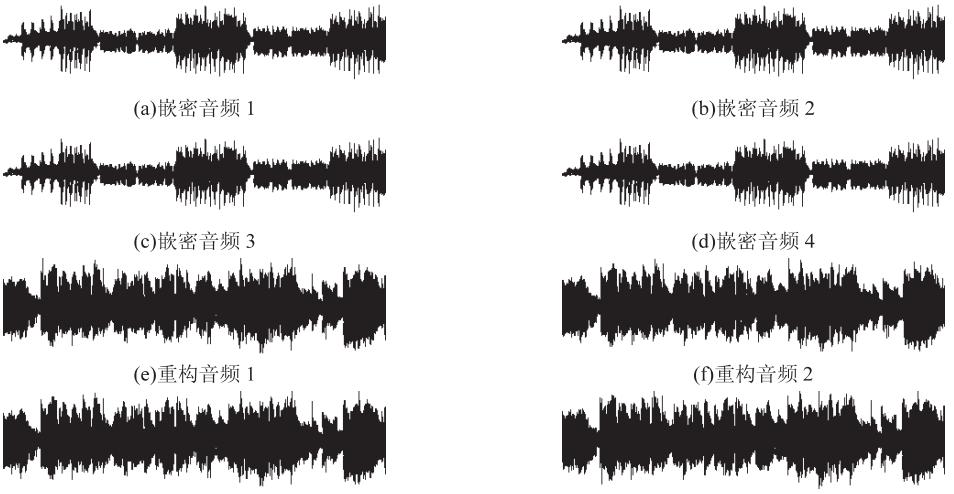
从表 1 和图 2 可看出,通过文中方法可将音频 A 嵌入到音频 B 中,且嵌入参数的伪装频音与原始公开音频波形图只有细微差别,并且从伪装音频中提取参数恢复出的秘密音频与原秘密音频的波形图也极为相似。

从表 1 可看出,伪装音频的质量相对于原公开音频的值信噪比大于 63 dB,方差在 1.3 左右,说明伪装音频质量极好,且只嵌入了恢复秘密音频的参数,减少了数据嵌入量。而恢复音频相对于原秘密音频,编号 1 和 2 的信噪比为 20 dB 左右,方差在 1 000 左右,编号 3 和 4 的信噪比在 27 dB 左右,方差在 400 左右,恢复音频依然具有较好的听觉质量,几乎听不到杂音。原因是文中对秘密音频小段进行拟合的是随机标准正交基,保证向量严格正交,避免了投影存在交叠,从而拟合精度更高,因此提高了秘密音频的重构精度。

图 3 给出了不同 k 取值对应的不同精度重构的测试音频波形图,与之对应的实验参数和测试结果如表 2 所示。

表 2　重构精度实验测试参数和实验测试结果

编号	k	小段长度	嵌密听觉质量			重构听觉质量		
			嵌密音频	SNR/dB	方差	重构音频	SNR/dB	方差
1	2	8	图 3(a)	69.01	1.39	图 3(e)	15.17	1 780.20
2	4	8	图 3(b)	67.43	1.37	图 3(f)	21.35	989.77
3	6	9	图 3(c)	66.41	1.35	图 3(g)	24.15	683.29
4	8	9	图 3(d)	64.56	1.32	图 3(h)	32.62	257.41



(a)嵌密音频 1 (b)嵌密音频 2

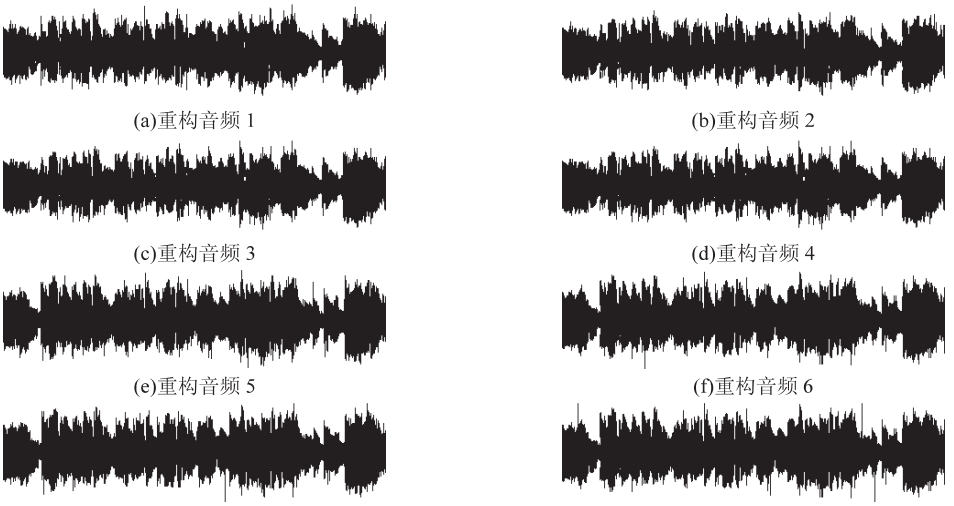
(c)嵌密音频 3 (d)嵌密音频 4

(e)重构音频 1 (f)重构音频 2

(g)重构音频 3 (h)重构音频 4

图 3 重构精度实验测试音频

从图 3 和表 2 可以看出,通过算法 1 和算法 2 可正确进行音频伪装和恢复,对于 k 的不同取值,音频可进行不同精度的重构,且 k 值越大,恢复的秘密音频精度也越来越高。从表 2 中编号 1、2 可看出,对于同等长度划分的小段序列,用以重构的标准正交基越多,即 k 取值越大,重构精度越高,方差越小,编号 4 中重构精度有 32 dB,远高于文献[15]20 dB 左右。对于信道中传输的嵌密公开音频,当 k 值越大时,嵌密音频的听觉质量下降。原因是文中选取了更多的基用于重构,从而造成了更多参数的嵌入,由此嵌密音频的信噪比随 k 值增大而降低,但其仍保持在 65 dB 左右,方差也保持在 1.3 左右,仍具有良好的听觉质量。因此,文中在提高秘密音频重构精度的同时,还有效平衡了嵌密公开音频的听觉质量。



(a)重构音频 1 (b)重构音频 2

(c)重构音频 3 (d)重构音频 4

(e)重构音频 5 (f)重构音频 6

(g)重构音频 7 (h)重构音频 8

图 4 不同密钥重构精度实验测试音频

表 3 不同密钥测试参数和实验测试结果

编号	k	错误密钥重构质量				正确密钥重构质量			
		错误密钥	重构音频	SNR/dB	方差	正确密钥	重构音频	SNR/dB	方差
1	4	243	图 4(a)	7.47	4 659.32	1 569	图 4(e)	22.35	989.77
2	6	16 784	图 4(b)	7.53	4 624.70	6 321	图 4(f)	26.05	462.07
3	7	32 873	图 4(c)	7.55	4 616.18	34 267	图 4(g)	29.74	358.69
4	8	32	图 4(d)	7.56	4 611.51	565	图 4(h)	31.48	293.51

图 4 给出了对于不同密钥对重构精度影响的实验测试音频波形图,与之对应的实验参数和测试结果如表 3 所示。

从图 4 和表 3 可以看出,只有掌握正确密钥才能

对秘密音频实现高精度的重构。对于错误密钥,重构的密音频信噪比仅在 7.5 dB 左右,方差高达 4 600,说明恢复音频听觉质量极低,具有大量杂音,且当密钥错误时,不管 k 取值多少,都无法正确高精度地恢复秘密音频。而对于正确的密钥,信噪比为 27 dB 左右,说明恢复的秘密音频听觉质量较高,几乎听不到杂音。且在密钥正确的前提下,可根据需求选取不同 k 值,实现不同精度的秘密音频重构。

5　结束语

传统基于 Tangram 的音频伪装方法变换精度低且不满足基本的正交关系,从而无法保证秘密音频与公开音频之间的拟合精度,同时当分段变换音频为恒值序列时,需添加扰动以保证变换后音频的恢复质量,由此会降低信道传输音频质量。为避免上述问题,提出一种结合密钥和随机标准正交基的音频伪装方法。首先对秘密音频和公开音频分段,利用密钥构造随机标准正交基;其次通过秘密音频小段在随机标准正交基上的投影来对秘密音频小段进行表达,从中选取包括均值幅值较大的前 k 个投影系数,并记录对应的索引位置;再次通过 EMD-q 密写方法嵌入到对应的公开音频小段中形成信道公开传输音频并通过信道公开传输音频提取变换参数和通过密钥重构秘密音频。实验表明,所提方法可充分利用随机标准正交基重构不同精度的秘密音频,且随着选取的幅值系数增多,恢复的秘密音频质量也越来越高,同时所述策略严格依赖于密钥,只有掌握正确密钥的用户才能进行高精度的重构。

参考文献:

[1] LIU Hongjun, KADIR A, LI Yanling. Audio encryption scheme by confusion and diffusion based on multi-scroll chaotic system and one-time keys[J]. Optik-International Journal for Light and Electron Optics, 2016, 127(19): 7431-7438.

[2] BOUSLEHI H, SEDDIK H. Innovative image encryption scheme based on a new rapid hyperchaotic system and random iterative permutation[J]. Multimedia Tools & Applications, 2018, 77(23): 30841-30863.

[3] LIMA J B, NETO E F D S. Audio encryption based on the cosine number transform[J]. Multimedia Tools and Applications, 2016, 75(14): 8403-8418.

[4] 邵利平,乐志芳. 基于 DCT 的多门限渐进秘密图像分存方案[J]. 信息网络安全,2018(3):54-62.

[5] 张　洋,邵利平,任平安. 免基向量 EMD(n,m)模型及其在图像密写上的应用[J]. 计算机辅助设计与图形学学报,2018,30(8):1490-1504.

[6] 丁　玮. 数字图像信息安全的算法研究[D]. 北京:中国科学院计算技术研究所,2000.

[7] 齐东旭. 画图的数学[M]. 北京:科学出版社,2009: 113-122.

[8] 吴　军,吴秋新. 一种基于七巧板游戏的数字图像信息伪装方法[J]. 计算机应用,2004,24(6):125-128.

[9] 余建德,宋瑞霞,齐东旭. 基于数字图像三角形剖分的信息伪装算法[J]. 计算机研究与发展,2009,46(9): 1432-1437.

[10] 邵利平,李苑梦. 基于 Tangram 算法和 2 维双尺度矩形映射的图像伪装及重构方法:中国,CN201410404838.7[P]. 2014-11-12.

[11] 李苑梦,邵利平. 一种基于改进 Tangram 算法的数字图像伪装方法[C]//全国信息隐藏暨多媒体信息安全学术大会. 武汉:中国电子学会通信学分会,2015.

[12] 邵利平. 数字图像置乱技术[M]. 北京:科学出版社,2016: 200-210.

[13] 邵利平,李苑梦,谢贤文. 基于分块序列的数字图像伪装及重构方法:中国,CN201510239140.9[P]. 2015-08-12.

[14] 邵利平,谢贤文,李苑梦. 基于分段序列的数字音频伪装及重构方法:中国,CN201510239139.6[P]. 2015-08-19.

[15] 谢贤文,邵利平. 结合字典序和排序线性拟合的音频隐藏方法[J]. 小型微型计算机系统,2017,38(12):2658-2667.