

基于共生特性的 G.729A 基音调制信息隐藏的检测

贾已真², 李松斌¹, 蒋雨欣¹, 戴琼兴¹, 邓浩江³

(1. 中国科学院声学研究所南海研究站, 海南海口, 570105; 2. 海南大学信息科学技术学院, 海南海口 570228;
3. 国家网络新媒体工程技术研究中心, 北京 100190)

摘要: 提出了一种 G.729A 自适应码本分组基音调制信息隐藏的检测算法. 对语音码流的分析发现, 通过基音预测进行信息隐藏将改变相邻语音帧中基音周期估计值的共生特性. 通过量化这种共生特性, 并经过 PCA (Principal Component Analysis, 主成分分析) 降维获得对隐写检测敏感的特征向量. 最后基于特征向量和 SVM (Support Vector Machine, 支持向量机) 构建隐写检测器. 对不同语音样本数据集的检测表明, 当语音长度在 2s 及以上时, 检测正确率均超过 96%. 此检测算法是一种有效的压缩域信息隐藏检测方法.

关键词: 基音调制信息隐藏; 共生特性; 隐写检测; 主成分分析; 支持向量机

中图分类号: TP309

文献标识码: A

文章编号: 0372-2112 (2015)08-1513-05

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2015.08.007

G.729A Pitch Modulation Information Hiding Detection Based on Symbiotic Characteristics

JIA Yi-zhen², LI Song-bin¹, JIANG Yu-xin¹, DAI Qiong-xing¹, DENG Hao-jiang³

(1. Haikou Laboratory, Institute of Acoustics, Chinese Academy of Sciences, Haikou, Hainan 570105, China;
2. College of Information Science and Technology, Hainan University, Haikou, Hainan 570228, China;
3. National Network New Media Engineering Research Center, Beijing 100190, China)

Abstract: A detection algorithm of pitch modulation information hiding in G.729A low bit-rate speech codec is proposed. The analysis of speech stream showed that pith modulation information hiding would change the pitch symbiotic characteristics of adjacent speech frames. We designed a model to quantify these pitch symbiotic characteristics for steganalysis. However, the dimension of quantitative feature vector of pitch symbiotic characteristics was too high, so PCA (Principal Component Analysis) was employed to reduce the dimension of the feature vector. Finally, we built a pitch modulation information hiding detector based on the dimension reduced feature vector and SVM (Support Vector Machine) classifier. Experiments on different speech datasets show that the proposed steganalysis algorithm is very effective: the accuracy is more than 96% when speech length equals to or is greater than 2s. So this paper gives an effective method for compression domain information hiding detection.

Key words: pitch modulation information hiding; symbiotic characteristics; steganography detection; principal component analysis (PCA); support vector machine (SVM)

1 引言

随着 IP 网络语音通信技术的高速发展, 以 VoIP (Voice over Internet Protocols) 通信中的语音码流作为载体的信息隐藏日益受到重视. 从现有的文献看, 基于低速率语音编码器的信息隐藏方法根据嵌入位置的不同可分为三类: 其一是利用 LPC 合成滤波器进行信息隐藏^[1,2]; 其二是通过修改压缩语音码流中的某些特定编码元素进行信息隐藏^[3,4]; 以及利用基音预测器进行信

息隐藏^[5,6].

利用基音预测器进行信息隐藏, 本质上是修改基音周期的估计过程即修改编码器中基音延迟码元的搜索过程. 虽然基音延迟的搜索范围会被调整, 但编码器仍能在一定范围内搜索有限条件的最优值, 这就使其能够保持较好的合成语音质量. 而且, 由于合成分析过程的存在, 搜索过程被修改引入的附加失真会在后续过程中得到补偿. 因此这类方法引入的失真极小, 具有很高的隐蔽性.

在 VoIP 语音信息隐藏检测研究方面, 见诸报道的一些检测方法大多是在非压缩域提取语音的隐写检测特征^[7~9]. 这些方法并不能有效的检测压缩语音码流中的信息隐藏. 其原因是, 压缩域的信息隐藏在解码的语音信号中引入很小的附加失真, 因此在非压缩域很难获得能够区分是否隐写的特征. 最近, 李等人在文献[10]中提出了一种针对低速率编码器 G.723.1 的 QIM 隐写检测算法, 这种算法从压缩域提取隐写检测特征, 达到了很好的检测效果, 为信息隐藏的检测提供了一种新的思路.

基于自适应码本分组的基音调制信息隐藏方法, 利用基音预测本身就存在的误差, 具有极高的隐蔽性, 对其隐写分析困难极大, 迄今尚未有相应的隐写分析方法见诸报道. 现有的一些语音隐写分析方法^[7~10]也不能直接用于该方法的检测. 为此, 本文拟对该种方法进行隐写分析, 并借鉴文献[10]的思想, 在压缩域提取用于隐写检测的特征.

2 基音调制信息隐藏的介绍

基音预测技术是低速率语音编码器中普遍采用的技术, 在几个 ITU-T 定义的标准中都得到了应用, 如 G.723.1 以及 G.729. 最近, 文献[6]针对 G.729A 低速率语音编码器, 提出了一种基于自适应码本分组的基音调制信息隐藏方法. 该方法在搜索最佳基音周期时将原始自适应码本进行分组处理: 当子帧具有整数精度延迟时, 将基音周期整数部分的搜索范围 $U_{\text{inte}} = [t_{\min}, t_{\max}]$ 调整为 U'_{inte} :

$$U'_{\text{inte}} = \{x | x \in U_{\text{inte}}, x \equiv B \pmod{2}\} \quad (1)$$

其中 B 为即将嵌入的秘密信息比特, 其值为 0 或 1, mod 为求模算法. 式(1)表示当嵌入秘密比特为 0 时在偶数组搜索最佳基音周期, 嵌入 1 时在奇数组搜索.

如果基音周期为分数精度, 将分数部分的搜索范围 $U_{\text{frac}} = [T - f, T + f]$ 调整为 U'_{frac} :

$$U'_{\text{frac}} = \begin{cases} [T - f, T + f], & B = 1 \\ [T], & B = 0 \end{cases} \quad (2)$$

其中 T 为基音周期的整数部分, $f \in (0, 1)$ 为分数精度延迟. 式(2)表示当嵌入秘密比特为 0 时在整数范围搜索最佳基音周期, 嵌入 1 时在分数范围搜索最佳基音周期. 这样就通过对基音预测的结果进行微调, 完成了秘密信息的嵌入.

3 隐写检测算法

语音信号存在局部周期性, 特别是浊音音素对应的语音片段. 从图 1 浊音音素“o”对应的语音片段可以清楚的看到这一点. 通常浊音音素的发音时间在 30 ~ 50ms 左右, 而在 G.729A 编码时进行基音预测的子帧时

长为 5ms. 因此如果相邻子帧语音信号正好是周期性重复的信号(例如, 图 1 中浊音音素“o”明显是一个重复了 4 次的周期性信号), 那么这些相邻子帧进行基音预测所得的值应该是相同的. 本文称这种相邻子帧基音延迟的等值性为基音延迟的共生特性.

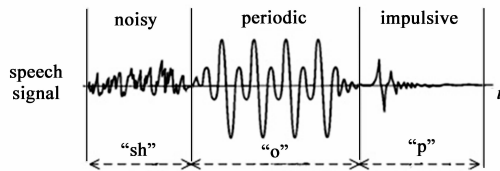


图1 语音信号的周期性

基于自适应码本分组进行信息隐藏时, 基音延迟的取值将依据当前嵌入的秘密比特是 0 还是 1, 将其原来的取值重新调整, 这将不可避免的导致相邻帧基音延迟的取值发生改变, 从而导致相邻帧基音延迟取值的共生特征被破坏. 因此, 可以利用这种共生特征进行隐写检测. 接下来, 我们首先证明相邻子帧的基音延迟具有共生特性, 其后给出基于这一关键线索的隐写检测算法.

3.1 基音延迟的共生特性

在 G.729A 进行基音预测时, 每帧分为两个子帧即第 1 子帧和第 2 子帧, 其基音延迟整数部分分别用 T_1 和 T_2 表示. 为了量化分析基音延迟的共生特性, 定义如下统计量:

$$M(a, b) = Fm(T_k(i) = a, T_k(i+1) = b), \quad (3)$$

$$\text{其中 } k \in \{1, 2\}, i = \{1, 2, \dots\}$$

其中, $M(a, b)$ 表示 G.729A 码流中, 第 i 帧中第 k 子帧的基音延迟 $T_k(i)$ 取值为 a , 且第 $i+1$ 帧中第 k 子帧的基音延迟 $T_k(i+1)$ 取值为 b 的共生基音对 (a, b) 数量. 使用 2000 个 10 秒长度的 G.729A 压缩语音片段组成数据集对 $M(a, b)$ 进行统计. 第 1 子帧的基音延迟 T_1 的统计结果如图 2 所示. 从图中可以看出 $M(a, b)$ 的峰值出现在 $b = a$ 时; 基音延迟 T_2 的统计结果与 T_1 的情况完全一致. 这就证实了相邻子帧基音延迟具有共生特性.

为了简化算法的复杂度, 以下将使用编码以后的基音延迟参量 P_1 和 P_2 来测试语音片段的共生特性强度. G.729A 编码器工作在 10ms 的语音帧上, 自适应码本搜索每一子帧(5ms)作一次. 第 1 子帧基音延迟用 8bit 编码, 第 2 子帧基音延迟用 5bit 编码, 编码后分别用 P_1, P_2 表示^[11].

$$P_1 = \begin{cases} 3(\text{int}(T_1) - 19) + \text{frac} - 1 \\ (\text{int}(T_1) - 85) + 197 \end{cases} \quad (4)$$

$$P_2 = 3(\text{int}(T_2) - t_{\min}) + \text{frac} + 2 \quad (5)$$

式(4)中, $\text{int}(T_1)$ 和 frac 分别是第 1 子帧基音延迟的整

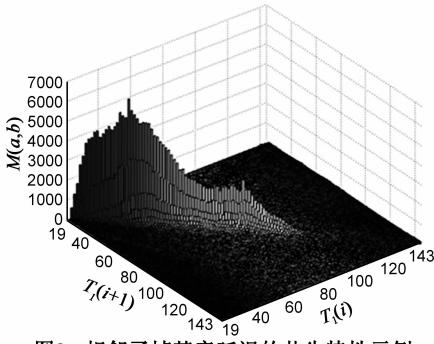


图2 相邻子帧基音延迟的共生特性示例

数和分数部分. 式(5)中, $\text{int}(T_2)$ 和 frac 分别是第 2 子帧基音延迟的整数和分数部分, t_{\min} 是由基音延迟 T_1 得到的^[11], $\text{frac} = [-1, 0, 1]$. 因为隐写破坏了基音延迟的共生特性, 因此编码后的基音延迟参量 P_1 和 P_2 的共生特性强度越强, 则对隐写检测越敏感. 为了衡量 P_1 和 P_2 的共生特性强度, 定义共生强度指数如下:

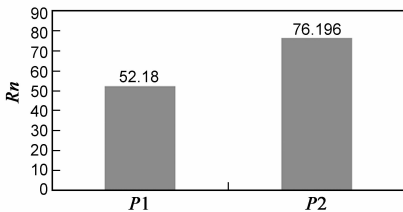
$$R_n = \sum_{i=0}^{k-2} F_n(i),$$

$$\text{其中 } F_n(i) = \begin{cases} 1, & P_n[i] = P_n[i+1] \\ 0, & P_n[i] \neq P_n[i+1] \end{cases}, n \in \{1, 2\} \quad (6)$$

式(6)中, R_1 和 R_2 分别代表 P_1 和 P_2 的共生特性强度. $P_n[i]$ 代表语音片段中第 i 帧中第 n 子帧的基音延迟 P_n . k 表示当前处理的语音片段共包含的语音帧的个数, 用式(7)计算. 式(7)中 t 为语音片段时长, t' 为每帧的时长, 在 G.729A 中,

$$k = t/t' \quad (7)$$

t' 等于 10ms. 用 2000 段随机选择的语音样本测试 P_1 和 P_2 的共生特性强度, 结果如图 3 所示.

图3 基音延迟 P_1 和 P_2 的共生特性强度

可见, P_2 的共生特性要强于 P_1 , 因此下面将对基音延迟 P_2 的共生特性进行量化, 以获取对隐写检测敏感的特征向量.

3.2 共生特性的量化

共生特性表示的是相邻帧基音延迟取值之间的关系, 而条件概率表示在某一事件发生的条件下, 另一事件发生的概率, 因此可以用条件概率来量化这种共生特性. 本文选用相邻帧基音延迟的条件概率建立量化统计模型, 获得其有效的量化表示特征. 定义相邻帧基音延迟取值的条件概率如下:

$$P(P_2^{i+1} = b | P_2^i = a) = \frac{P(P_2^i = a, P_2^{i+1} = b)}{P(P_2^i = a)}$$

$$0 \leq i \leq k-2$$

$$a, b \in [0, 31] \quad (8)$$

式(8)中, $P(P_2^{i+1} = b | P_2^i = a)$ 表示相邻两帧中第 i 帧的基音延迟 P_2 取值为 a 的条件下, 第 $i+1$ 帧的 P_2 取值为 b 的概率. a 与 b 取值都在 0 至 31 之间, 因此条件概率为 1024 维. 可以计算出每段语音样本的特征 $F = \{p_1, p_2, p_3, \dots, p_{1024}\}$. 假设使用语音样本数为 N , 至此我们可以获得一个 $N \times 1024$ 维的特征数据集.

高维数据中包含了大量的冗余, 在训练样本容量一定的前提下, 特征维数的增加将使得样本统计特性的估计变得更加困难. 这样可能导致分类器的推广能力或泛化能力降低, 呈现所谓的“过学习”或“过训练”的现象. 为了消除冗余, 避免出现“过学习”的情况, 本文对 $N \times 1024$ 维的特征数据集进行降维.

本文选用主成分分析 (Principal Component Analysis, PCA) 对提取到的高维特征进行降维. PCA 的目的是寻找在最小均方意义下最能够代表原始数据的投影方法, 实际上就是求得这个投影矩阵, 用高维的特征乘以这个投影矩阵, 便可以将高维特征的维数下降到指定的维数. 本文经过实验发现, 将高维特征降至 100 维时可以获得对隐写检测较为敏感的特征向量, 因此本文用 PCA 将上述特征数据集降至 100 维. 对于上述得到的特征数据集, 它可以建立一个 $N \times 1024$ 的样本矩阵 M , M 用公式(9)表示如下:

$$M = \begin{pmatrix} p_{1,1} & p_{1,2} & p_{1,3} & \cdots & p_{1,1024} \\ p_{2,1} & p_{2,2} & p_{2,3} & \cdots & p_{2,1024} \\ p_{3,1} & p_{3,2} & p_{3,3} & \cdots & p_{3,1024} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{N,1} & p_{N,2} & p_{N,3} & \cdots & p_{N,1024} \end{pmatrix} \quad (9)$$

其中, $p_{i,j}$ 表示第 i 个样本的第 j 维特征. 第 j 维的均值 \bar{p}_j 用下式计算:

$$\bar{p}_j = \frac{\sum_{k=1}^N p_{k,j}}{N}, \quad (10)$$

$$\text{其中 } 1 \leq j \leq 1024$$

第 m 维与第 n 维的协方差定义如下:

$$\text{cov}(p_m, p_n) = \frac{\sum_{k=1}^N (p_{k,m} - \bar{p}_m)(p_{k,n} - \bar{p}_n)}{N-1},$$

$$\text{其中 } m, n \in [1, 1024] \quad (11)$$

上式中, p_m 和 p_n 分别表示 M 矩阵中第 m 列和第 n 列, \bar{p}_m 和 \bar{p}_n 表示 M 矩阵中第 m 列和第 n 列的均值. 接下来求这个样本的协方差矩阵, 如下式所示.

$$\mathbf{c} = \begin{pmatrix} \text{cov}(p_1, p_1) & \text{cov}(p_1, p_2) & \text{cov}(p_1, p_3) & \cdots & \text{cov}(p_1, p_{1024}) \\ \text{cov}(p_2, p_1) & \text{cov}(p_2, p_2) & \text{cov}(p_2, p_3) & \cdots & \text{cov}(p_2, p_{1024}) \\ \text{cov}(p_3, p_1) & \text{cov}(p_3, p_2) & \text{cov}(p_3, p_3) & \cdots & \text{cov}(p_3, p_{1024}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \text{cov}(p_{1024}, p_1) & \text{cov}(p_{1024}, p_2) & \text{cov}(p_{1024}, p_3) & \cdots & \text{cov}(p_{1024}, p_{1024}) \end{pmatrix} \quad (12)$$

这样就得到一个 1024×1024 维的协方差矩阵, 然后求出这个协方差矩阵的特征值和特征向量, 应该有 1024 个特征值和特征向量. 我们根据特征值的大小, 取前 100 个特征值所对应的特征向量, 构成一个 1024×100 的特征矩阵. $N \times 1024$ 的样本矩阵乘以这个 1024×100 的特征矩阵, 就得到了一个 $N \times 100$ 维的降维之后的样本矩阵, 即用于隐写检测的特征向量, 用 \mathbf{t} 表示.

3.3 基于机器学习的隐写检测

隐写检测的目标是判断一段 G.729A 压缩语音是否存在基音调制信息隐藏. 假设有一段未知是否隐写的压缩语音片段 S , 从语音片段 S 中提取到的特征向量为 \mathbf{t} , 则隐写检测过程可以表示如下:

$$y = f(\mathbf{t}) \quad (13)$$

函数 f 即为隐写检测器, 其输出结果 y 等于 +1 时, 表示 S 存在隐写; y 等于 -1 时, 表示 S 不存在隐写. 显然函数 f 是一个二值分类器, 隐写检测过程实质上是分类过程. 对于分类问题, 基于机器学习分类方法是当前主流, 因此本文也将采用这种方法. 鉴于支持向量机(SVM)分类器在小样本学习时, 具有优良特性, 本文选用它作为分类器. 隐写检测过程分为以下 4 个步骤: (1) 获取原始语音片段, 经过 G.729A 编码器编码, 得到未隐写的压缩语音样本; 使用文献[6]对上述原始语音片段进行隐写, 得到隐写的压缩语音样本; (2) 提取上一步骤所获得的两类样本的特征向量 \mathbf{t} , 标记每个向量的类别; (3) 使用上述得到的已标记类别的特征向量对 SVM 分类器进行训练, 得到分类器 f ; (4) 使用分类器 f 对未知类别样本进行隐写检测: 提取未知是否隐写的语音片段的特征向量 \mathbf{t} , 使用分类器 f 对其分类, 分类器的输出即为检测结果.

4 实验结果

本文采用文献[6]给出的方法作为隐写算法, 针对 G.729A 低速率语音编码器进行了本文隐写检测算法的性能测试, 并与文献[9]给出的隐写检测方法进行了比较. 为了阐明本算法具有较好的普适性, 本文从互联网随机搜索了不同发音人的语音片段组成语音样本库. 语音样本库包含 5 种数据集, 分别是中文男声 (Chinese Man, CM), 包含 500 个语音片段; 中文女声 (Chinese Woman, CW), 包含 533 个语音片段; 英文男声 (English

Man, EM), 包含 818 个语音片段; 英文女声 (English Woman, EW), 包含 824 个语音片段; 第 5 种为上述四类样本的混合 (Hybrid).

我们将没有信息隐藏的压缩语音片段称为未隐写类 (C 类), 否则称之为隐写类 (S 类). 对语音样本库中 CM 数据集中的每个原始语音片段使用 G.729A 编码器编码, 获得其对应的 C 类压缩语音片段; 对原始语音片段再次使用 G.729A 编码器编码, 并在基音预测过程中采用文献[6]所述方法嵌入秘密信息, 获得其对应的 S 类压缩语音片段. 选择 400 个 C 类及其对应的 S 类语音片段组成训练样本, 剩余的组成测试样本. 同理可得到 CW、EM、EW 和 Hybrid 数据集的训练样本和测试样本的数量, 如表 1 所示.

表 1 训练样本及测试样本数量

	CM	CW	EM	EW	Hybrid
训练样本	400	400	600	600	2000
测试样本	100	133	218	224	675

用本文所述方法提取每种数据集的训练样本和测试样本的特征向量, 接着用训练样本的特征向量对分类器进行训练, 用训练好的分类器对测试样本进行分类并计算分类准确率. 同时, 我们还评估了语音片段时长对隐写检测结果的影响. 表 2 给出了用本文检测方法对 5 种数据集在 7 种时长下的隐写检测结果.

表 2 本文隐写检测结果

	CM (%)	CW (%)	EM (%)	EW (%)	Hybrid (%)
0.5s	87.50	82.70	88.76	90.17	89.48
1s	85.50	91.35	93.80	93.30	93.92
2s	96.00	96.99	97.24	98.43	98.14
4s	99.50	100.00	100.00	100.00	99.85
6s	100.00	100.00	99.77	100.00	100.00
8s	100.00	100.00	100.00	100.00	99.77
10s	100.00	100.00	100.00	100.00	100.00
平均值	95.50	95.86	97.08	97.41	97.31
最大值	100.00	100.00	100.00	100.00	100.00
最小值	85.50	82.70	88.76	90.17	89.48

文献[6]使用了文献[9]的隐写分析算法对其隐写算法的性能进行了评估. 文献[9]基于二阶差分衍生的梅尔倒谱系数 (Derivative Mel-Frequency Cepstral Coefficients, DMFCC) 特征及 SVM 构建了一种通用隐写分析方法, 能够对多种隐写算法达到较好检测效果. 因此将它与本文提出的隐写检测方法进行了对比. 使用相同的语音样本库, DMFCC 方法的隐写检测结果如表 3 所示.

可见, 用 DMFCC 方法对 5 种数据集最大检测正确

率为 67.5%, 最小值为 50.89%。用本文提出的隐写检测方法, 最大准确率为 100%, 最小值为 82.7%。在语音码流长度大于等于 2s 时, 本文方法的隐写检测准确率均超过 96%。并且, 检测正确率随着语音片段长度的增加而增大。因此本文所提出的隐写检测方法优于文献[9]的方法, 可有效实现对基音调制信息隐藏的检测。

表 3 DMFCC 隐写检测结果

	CM(%)	CW(%)	EM(%)	EW(%)	Hybrid(%)
0.5s	51.50	53.76	53.67	50.89	53.24
1s	52.00	56.02	53.44	56.25	54.59
2s	54.50	55.26	54.59	57.81	56.15
4s	57.00	57.89	55.50	61.83	58.37
6s	56.50	58.27	55.28	60.94	57.48
8s	60.50	61.28	56.19	60.04	59.56
10s	67.50	63.16	54.13	62.50	62.15
平均值	57.07	57.95	54.69	58.61	57.36
最大值	67.50	63.16	56.19	62.50	62.15
最小值	51.50	53.76	53.44	50.89	53.24

5 结束语

针对 G.729A 自适应码本分组的基音预测信息隐藏, 提出了一种有效的检测方法。该方法发现通过基音预测进行信息隐藏将改变相邻语音帧中基音延迟取值的共生特性, 因此通过量化这种分布特性, 获得对隐写分析敏感的特征向量。然后对特征向量进行降维并通过机器学习构造隐写检测器, 实现了在压缩域对基音调制信息隐藏的有效检测。

参考文献

- [1] Xiao B, Huang Y, Tang S. An approach to information hiding in low bit-rate speech stream [A]. Proceedings of the 2008 IEEE Global Telecommunications Conference [C]. New Orleans, LA, USA: IEEE, 2008. 1 - 5.
- [2] Tian H, Liu J, Li S. Improving security of quantization-index-modulation steganography in low bit-rate speech streams [J]. Multimedia Systems, 2014, 20(2): 143 - 154.
- [3] Liu L, et al. Perceptually transparent information hiding in G.729 bit stream [A]. Proceedings of the 4th International Conference on Intelligent Information Hiding and Multimedia Signal Processing [C]. Harbin, China, 2008. 406 - 409.
- [4] Huang Y, Tang S, Yuan J. Steganography in inactive frames of VoIP streams encoded by source codec [J]. IEEE Transactions on Information Forensics and Security, 2011, 6(2): 296 - 307.

- [5] Huang Y, Liu C, Tang S, Bai S. Steganography integration into a low-bit rate speech codec [J]. IEEE Transactions on Information Forensics and Security, 2012, 7(6): 1865 - 1876.
- [6] 刘程浩, 柏森, 黄永峰, 等. 一种基于 G.729a 基音预测的信息隐藏算法 [A]. 第十届全国信息隐藏暨多媒体信息安全学术大会论文集 [C]. 北京, 2012. 15 - 18.
Liu C, Bai S, Huang Y, et al. An information hiding algorithm in G.729a based on pitch prediction [A]. The tenth National Academic Conference Proceedings of Information Hiding and Multimedia Information Security [C]. Beijing, China, 2012. 15 - 18. (in Chinese)
- [7] Koçal Osman Hilmi, Yürüklü Emrah, Avcıbas İsmail. Chaotic-type features for speech steganalysis [J]. IEEE Transactions on Information Forensics and Security, 2008, 3(4): 651 - 661.
- [8] Huang Y, Tang S, Zhang Y. Detection of covert voice-over internet protocol communications using sliding window-based steganalysis [J]. IET Communications, 2011, 5(7): 929 - 936.
- [9] Liu Q, et al. Temporal derivative-based spectrum and mel-cepstrum audio steganalysis [J]. IEEE Transactions on Information Forensics and Security, 2009, 4(3): 359 - 368.
- [10] Li S, Tao H, Huang Y. Detection of QIM steganography in G.723.1 bit stream based on quantization index sequence analysis [J]. Zhejiang Univ-Sci C (Comput & Electron), 2012, 13(8): 624 - 634.
- [11] ITU. ITU-T Recommendations G.729. Coding of Speech at 8 kbit/s Using Conjugate-Structure Algebraic-Code-Excited Linear-Prediction (CS-ACELP) 2007 [S/OL]. <http://www.itu.int/rec/T-REC-G.729-199603-S/en.1996-03>.

作者简介



贾已真 女, 1989 年出生, 河南商丘人。2012 年毕业于海南大学信息科学技术学院, 获得学士学位。现为海南大学硕士研究生, 研究方向为网络与流媒体技术。

E-mail: JYZ231@126.com



李松德(通信作者) 男, 1981 年出生, 福建诏安人。中科院博士, 清华大学博士后。现为中国科学院声学研究所南海研究站副研究员。研究方向为多媒体信息处理、信息隐藏等。

E-mail: lisb@dsp.ac.cn