基于随机位置选择和矩阵编码的语音信息隐藏方法

吴志军* 李常亮 李 荣

(中国民航大学电子信息与自动化学院 天津 300300)

摘 要：针对低速率语音编码问题，该文提出基于G.723.1编码标准的信息隐藏算法。在基音预测编码过程中，通过控制闭环基音周期(自适应码本)的搜索范围，该文结合随机位置选择方法(RPS)和矩阵编码方法(MCM)，实现秘密信息的嵌入，在语音编码过程中实现了信息的隐藏。RPS方法的采用降低了载体码字之间的关联性，MCM方法的采用降低了载体的改变率。实验结果证明，该文算法下PESQ恶化率平均值最大为1.63%，隐蔽性良好。

关键词：低速率语音编码；G.723.1编码标准；基音预测；闭环基音周期

中图分类号：TN912.3 **文献标识码**：A **文章编号**：1009-5896(2020)02-0355-09

DOI: 10.11999/JEIT181163

Speech Information Hiding Method Based on Random Position Selection and Matrix Coding

WU Zhijun LI Changliang LI Rong

(*College of Information Engineering and Automation, Civil Aviation University of China, Tianjin* 300300, *China*)

Abstract: For the low-rate speech encoding problem, an information hidden algorithm based on the G.723.1 coding standard is proposed. In the pitch prediction coding process, by controlling the search range of the closed-loop pitch period (adaptive codebook), combined with the Random Position Selection (RPS) method and the Matrix Coding Method (MCM), the secret information is embedded, which is implemented in the speech coding process. The adoption of the RPS method reduces the correlation between the carrier code-words, and the adoption of the MCM method reduces the rate of change of the carrier. The experimental results show that the average PESQ (Perceptual Evaluation of Speech Quality) deterioration rate under the algorithm is 1.63%, and the concealment is good.

Key words: Low rate speech coding; G.723.1 coding standard; Pitch prediction; Closed loop pitch period

1 引言

隐写(steganography)技术是信息隐藏的重要分支[1]，它利用感觉器官的不可感知性和多媒体数字信号本身的冗余性，将秘密信息隐藏在公开的多媒体载体中，从而达到保密通信的目的[2]。基于动态流媒体的信息隐藏技术已经成为研究的热点[3]，最具有代表性的是基于IP语音(Voice over IP, VoIP)的隐藏技术研究[4]。VoIP中有两种类型的语音编码器：波形编码器(例如G.711, G.726)和声码器(例如G.723, G.729, iLBC)。相对于基于原始语音信号量化值的波形编码器，声码器多采用合成-分析法的线性预测编码方法，节省大量带宽资源[5]。

基于声码器的语音信息隐藏方法根据嵌入位置可分为3类：第1类是利用LPC(Linear-Predictive Coding)合成滤波器进行信息隐藏[6–9]；第2类是利用基音预测器进行信息隐藏[10,11]；第3类是属于编码后的隐藏算法[12–15]。上述第2类方法利用基音预测器进行信息隐藏，隐蔽性极高。文献[10]利用基音预测器进行信息隐藏，其能够保持较好的合成语音质量。文献[16]发现文献[10]所提出的隐藏方法将导致压缩语音流中相邻语音帧的自适应码本的关联

收稿日期：2018-12-18；改回日期：2019-07-20；网络出版：2019-09-20
*通信作者： 吴志军 zjwu@cauc.edu.cn
基金项目：国家自然基金委员会与中国民航局联合基金(U1933108)，天津市 自然科学基金(17JCZDJC30900)，天津市教委科研计划项目 (2019KJ117), 2018 年中央高校基本科研业务费项目(3122018D007, 3122018C003)

Foundation Items: The Joint Funds of National Natural Science Foundation of China and Civil Aviation Administration of China (U1933108), The Key Program of Natural Science Foundation of Tianjin (17JCZDJC30900), The Scientific Research Project of Tianjin Municipal Education Commission (2019KJ117), The Fundamental Research Funds for the Central Universities of China (3122018D34007, 3122018C003)

特性发生改变，并以此为设计隐藏分析的关键线索，设计了隐藏检测器，最高检测率达到100%。因此，针对G.723.1低速率编解码器标准，如何进一步提高基于基音预测器信息隐藏方法的隐蔽性是本文研究的动机和目标。

本文结构如下：第2节研究基于基音预测过程的信息隐藏方法；第3节介绍本文方法的实验验证；第4节对本文算法的性能指标实验结果进行分析；第5节总结全文。

2 基于G.723.1编码标准的基音调制信息隐藏算法

G.723.1编码器对每帧240个样本点进行分析，一帧语音定义为$\boldsymbol{S}[M]=\{s[n]|n=0,1,\cdots,239\}$。其中，每帧分为4个具有60个样本点的子帧。每帧输入信号$\boldsymbol{S}[M]$转换为感知加权信号$\boldsymbol{F}[M]=\{f[n]|n=0,1,\cdots,239\}$。由$\boldsymbol{F}[M]$进行两次开环基音周期估计，分别针对前后两个子帧。采用交叉关联判断标准$C_{\mathrm{OL}}(j)$最大化方法用来估计开环基音周期，表达式为

$$C_{\mathrm{OL}}(j)=\frac{\left(\sum_{n=0}^{119}f[n]\cdot f[n-j]\right)^2}{\sum_{n=0}^{119}f[n-j]\cdot f[n-j]},\quad 18\le j\le 142 \tag{1}$$

最大化交叉关联$C_{\mathrm{OL}}(j)$的指数j被选定作为适当的开环基音周期。每一子帧的闭环基音周期定义为L_i，$i=1,2,3,4$。开环基音周期定义为$L_{\mathrm{OL}(i)}$，$i=0,1$。$L_{\mathrm{OL}(1)}$和$L_{\mathrm{OL}(2)}$分别代表前后两子帧的开环基音周期。闭环基音周期L_i的值位于开环基音周期$L_{\mathrm{OL}(i)}$的附近。将L_i的搜索范围$M_i(i=0,1,2,3)$分为奇数组M_{odd}和偶数组M_{even}，如式(2)与式(3)。定义秘密比特流为$\boldsymbol{B}=[b_0,b_1,b_2,\cdots]$，假设第$i(i=0,1,2,3)$子帧所要嵌入秘密比特位为$b_i$。当$b_i=0$时，基音预测器在$M_{\mathrm{even}}$中搜索最佳延迟；当$b_i=1$时，在$M_{\mathrm{odd}}$中搜索最佳延迟。

$$M_{\mathrm{odd}}=\{m_k \mod 2=1, m_k\in M_i\} \tag{2}$$

$$M_{\mathrm{even}}=\{m_k \mod 2=0, m_k\in M_i\} \tag{3}$$

文献[10,11]隐藏秘密信息的基本思想即对基音预测器进行改进，根据待嵌入的秘密信息位对基音的搜索范围进行调整，从而达到保密通信的效果。本文采用这种基本的隐藏思想，同时结合随机位置选择方法(RPS)[17]和矩阵编码方法(MCM)[18]，提高了传输透明性，降低了载体改变率。

2.1 基音调制的改进

文献[10]采用G.723.1编码器进行基音调制信息的隐藏，假设秘密信息比特位为$S=\{s_0,s_1,s_2,\cdots\}$，第n帧的第$i(i=0,1,2,3)$要嵌入的秘密比特位为s_i，嵌入0时在偶数组搜索最佳基音延迟，嵌入1时则在奇数组搜索最佳基音延迟。这样就通过对基音搜索范围的调整，完成了秘密信息的嵌入，达到了保密通信的效果。

由于语音信号存在局部周期性，因此子帧存在周期性，那么子帧基音预测所得的值应该是相同的，也就是说，子帧的自适应码本参量之间具有关联性。文献[16]通过实验证明在G.723.1压缩语音码流中，相邻帧第2子帧与第4子帧之间具有极强相关性，帧内第2子帧与第4子帧之间具有极强相关性，并基于此构建码书关联网络，得到对隐藏敏感的特征向量，采用SVM分类器实现了对文献[10]算法的有效检测。文献[16]所构造的SVM分类器忽略了非相邻帧子帧之间的关联性，对第2子帧的依赖性太强。据此，本文将第1, 3, 4子帧选为载体子帧。

2.2 隐藏位置的选择

隐藏位置的选择采用随机位置选择方法(RPS)，以随机的方式来选择隐藏位置。假设原始载体语音总共有L帧，需要承载秘密信息的帧总共有l帧。定义嵌入率为t，那么$t=l/L$，其中，$0\le t\le 1$。本文根据嵌入率$t(t\in[0,1])$来选择载体语音帧。

编码器对模拟输入信号滤波，采样，量化。然后，编码器将连续语音信号样点$y[n]$分割为帧长为240个样点的帧f_i。令嵌入率t为阈值，由此来判断f_i是否用来隐藏秘密信息。定义$\boldsymbol{F}=\{f_1f_2\cdots f_L\}$为原始语音帧系列，$\boldsymbol{F}^*=\{f_1^*f_2^*\cdots f_L^*\}$为载密语音帧系列。

令密钥K作为伪随机数产生器的种子，使其产生一系列随机数字$\lambda=\{\lambda_i|\lambda_i\in[0,1],i=1,2,\cdots,L\}$。$\lambda$为发送方和接收方共享。其中秘钥的传输采用Diffie-Hellman方案进行传输。在发送端，如果$\lambda_i\le t$，那么f_i用来隐藏秘密信息；否则，f_i不用来隐藏秘密信息。最终，实现了隐藏位置的选择。

2.3 嵌入效率的控制

对于每一帧载体语音，本文采用矩阵编码的方法增加嵌入效率。每一子帧闭环基音周期的搜索区间为$\boldsymbol{M}=\{M_i|i=1,2,\cdots,N\}$，$M_i=\{M_{\mathrm{odd}},M_{\mathrm{even}}\}$。根据MCM方法，秘密信息$\boldsymbol{S}=\{s_1,s_2,\cdots,s_k\}$被嵌入进载体语音帧中，其中$N=2^k-1$。嵌入过程如下

(1)对于每一帧$f[m]$，在M中搜索每子帧对应的闭环基音周期$L=\{L_i|L_i\in M_i,i=1,2,\cdots,N\}$。

(2)根据L_i的奇偶性判断其属于哪个搜索区间，以此确定L_i的状态$\boldsymbol{T}=\{t_i|t_i=0\text{或}1;i=1,2,\cdots,N\}$。$t_i$的确定根据如下规则

$$t_i = \begin{cases} 1, & \mod(L_i, 2) == 1 \\ 0, & \mod(L_i, 2) == 0 \end{cases} \tag{4}$$

(3)将t_i的下标采用二进制编码，并将二进制编码序列定义为一个矢量$\underset{\sim}{\boldsymbol{B}}_i = (b_{i1}b_{i2}\cdots b_{ik})^{\mathrm{T}}$。$i$ 的表达式如式(5)。编码矩阵$\underset{\sim}{\boldsymbol{R}}$由这些矢量构成，具体定义如式(6)所示

$$i = \sum_{j=1}^{k} b_{ij} \cdot 2^{(j-1)}, b_{ij} = 0\text{或}1 \tag{5}$$

$$\underset{\sim}{\boldsymbol{R}} = (\underset{\sim}{\boldsymbol{B}}_1\ \underset{\sim}{\boldsymbol{B}}_2\ \cdots\ \underset{\sim}{\boldsymbol{B}}_N) = \begin{bmatrix} b_{11} & b_{21} & \cdots & b_{N1} \\ b_{12} & b_{22} & \cdots & b_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ b_{1k} & b_{2k} & \cdots & b_{Nk} \end{bmatrix} \tag{6}$$

(4) 对于$\underset{\sim}{\boldsymbol{R}}$中每一行，计算$y_j = \begin{cases} 0, s_j = \oplus_{i=1}^{N}(t_i \cdot b_{ij}) \\ 1, s_j \neq \oplus_{i=1}^{N}(t_i \cdot b_{ij}) \end{cases}$ $1 \le j \le k, \oplus_{i=1}^{N}$代表连续异或操作。

(5) 计算表达式

$$Y = \sum_{j=1}^{k} y_j \cdot 2^{j-1} \tag{7}$$

如果$Y = 0$，没有L_i需要重新搜索最佳值；反之，第i个闭环基音周期需要在$M_i = \{M_{\text{odd}}, M_{\text{even}}\}$的另一个子区间重新搜索。在最多不超过一个闭环基音周期需要重新搜索的情况下，完成了k bit秘密信息在N个搜索区间中的嵌入。

相应的秘密信息的提取算法如下

(1)提取每帧语音对应的闭环基音周期$L^* = \{L_i^* | L_i^* \in M_i; i = 1, 2, \cdots, N\}$，确定其状态$T^* = \{t_i^* | t_i^* = 0\text{或}1;\ i = 1, 2, \cdots, N\}$。$t_i^*$的选择根据如下规则

$$t_i = \begin{cases} 1, & \mod(L_i, 2) == 1 \\ 0, & \mod(L_i, 2) == 0 \end{cases} \tag{8}$$

(2) 根据如下规则恢复秘密比特信息的每一位：$s_j = \oplus_{i=1}^{N}(t_i^* \cdot b_{ij}), 1 \le j \le k$。

3　实验配置及其环境

在局域网中设计信息隐藏的实验环境，搭建发送、接收平台，如图1所示。图1中，实验拓扑包括发送端、接收端两部分。其中，发送端完成秘密信息的嵌入，接收端完成相应的提取。发送端包含3个模块：秘密信息嵌入模块、G.723.1编码器和加密模块；接收端包含3个模块：秘密信息提取模块、G.723.1语音解码器和解密模块。

采用本文提出的语音信息隐藏算法，将语音隐藏和提取功能分别实现在发送端和接收端。语音隐藏和提取的工作流程如图2所示。

发送端和接收端的各个模块均实现在台式PC机上，它们的配置如表1所示。

语音样本库采用文献[12]中使用的数据集，数据集包含时长41 h中文语音样本和72 h英文语音样本。中文语音样本包含中文男声(Chinese Speech Man, CSM)、中文女声(Chinese Speech Woman, CSW)，英文语音样本包含英文男声(English Speech Man, ESM)和英文女声(English Speech Woman, ESW)。

4　实验结果及其分析

隐藏算法的性能主要从隐蔽性、实时性和抗检测性3个方面来衡量。本文根据这3个指标对隐藏算法做性能评估，并将测试结果与文献[10]基音调制信息隐藏算法进行比较，验证本文算法的可行性。

4.1　隐藏/提取算法的实现

根据本文提出的基音调制信息隐藏方法，在图 1所示实验环境中，依据图2算法流程，对隐藏算法进行了仿真实现。

发送端隐藏算法的实现

(1) 对秘密信息(明文)进行预处理，得到二进制秘密信息，并对二进制秘密信息进行加密得到密文。

(2) 编码器将载体语音分为一系列语音帧$f[i]$。

(3) 通过基音预测器计算每一子帧的初始闭环基音周期L_i。

(4) 发送端和接收端共享密钥K。令密钥K作为伪随机数发生器的种子，发送端和接收端产生相同的伪随机数序列$\lambda = \{\lambda_i | \lambda_i \in [0, 1], i = 1, 2, \cdots, L\}$。

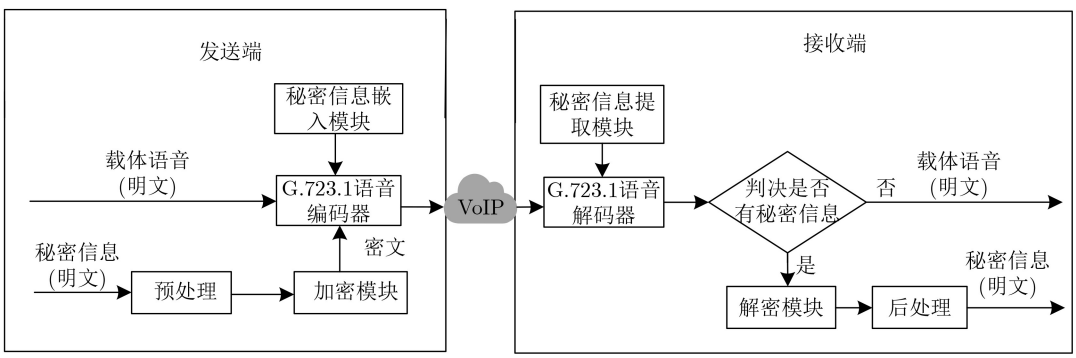

图 1 实验拓扑图

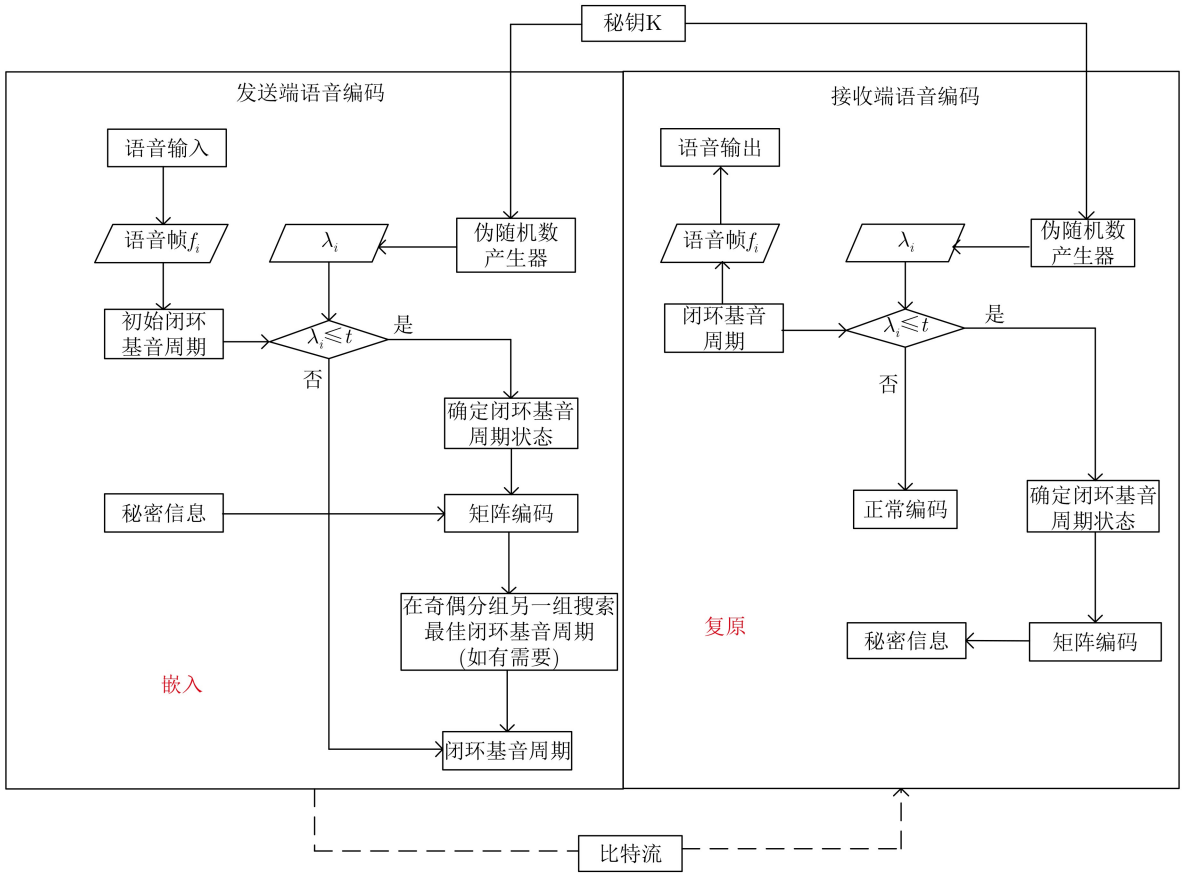

图 2 隐藏算法实现框图

表 1 发送方、接收方台式PC配置情况

处理器	内存	声卡	系统
Intel(R) Core(TM) i5-4590 CPU @ 3.30 GHz	4 GB	Realtek High Definition Audio	Windows 7专业版 Service Pack 1

若$\lambda_i \leq t$，语音帧f_i用来作为秘密信息载体；反之，语音帧f_i不用来作为秘密信息载体。

(5) 若语音帧f_i用来作为秘密信息载体，则应用上节所提出的MCM方法进行秘密信息的嵌入；反之，则进行正常编码。

(6) 发送语音比特流至网络。

接收端提取算法的实现

(1) G.723.1语音编码器从网络中接收语音比特流。

(2) 发送端和接收端共享密钥K，接收端与发送端产生相同的伪随机数序列。

(3) 若$\lambda_i \leq t$，则根据本文提出的矩阵编码算法进行秘密信息的提取；反之，则进行正常解码。

(4) 对提取的秘密信息进行解密及后处理，最终得到秘密信息(明文)。

4.2 隐藏性能测试与分析

隐蔽性是衡量隐秘信息在载体信息中不可感知能力的一个重要指标。本文针对隐蔽性从两方面进行分析：(1)针对语音时域、频域波形图和语谱图进行分析；(2)利用客观测试对语音质量进行评估。选择不同发音人的多个语音片段组成语音样本库，所用语音片段样本包含4种，分别是中文男声、中文女声、英文男声、英文女声。每种语音库分别包括100条3 s长的语音和100条10 s长的语音，这4种样本总共包括800条语音。

4.2.1 时频域角度的测试与分析

选取3 s中文女声1条，分别采用本文隐藏算法和文献[10]隐藏算法对其进行秘密信息嵌入。对正常编解码后的语音和嵌入秘密信息的语音进行时频域和语谱图的仿真，其仿真结果如图3所示。图3(a)，图3(b)，图3(c)为原始语音分析图；图3(d)，图3(e)，图3(f)为采用文献[10]隐藏算法的合成语音分析图；图3(g)，图3(h)，图3(i)为采用本文隐藏算法的合成语音分析图。

从图3看出，未载密语音和载密语音在时域、频域和语谱上无明显差别。表明，文献[10]和本文隐藏算法对语音质量影响较小。对于图3(a)，图3(d)，图3(g)，在采样点5000左右，图3(g)更加接近于

图3(a)；对于图3(b)，图3(e)，图3(h)，图3(e)和图3(h)均接近于图3(b)；对于图3(c)，图3(f)，图3(i)，图3(f)，图3(i)也比较接近于图3(c)。仿真结果表明，本文提出的隐藏算法相对于文献[10]在一定程度上对语音质量影响较小。

4.2.2 客观语音质量的测量与分析

本文选用文献[19]中的客观语音质量评估方法对语音质量进行评估。采用G.723.1编码器，分别结合本文算法和文献[10]中的隐藏算法对语音样本进行编解码，并对重构语音进行质量评估。图4和图5描述了本文隐藏算法下载密语音与未载密语音样本的PESQ(Perceptual Evaluation of Speech Quality)值。其中，蓝实线代表载密语音的PESQ值；红虚线代表未载密语音的PESQ值。由图4和图5可以看出：对于时长为3 s和10 s的载体语音，载密语音与未载密语音的PESQ差异很小，表明本文提出的隐藏算法对语音PESQ值有着微乎其微的影响。

表2描述了本文隐藏算法下载密重构语音样本

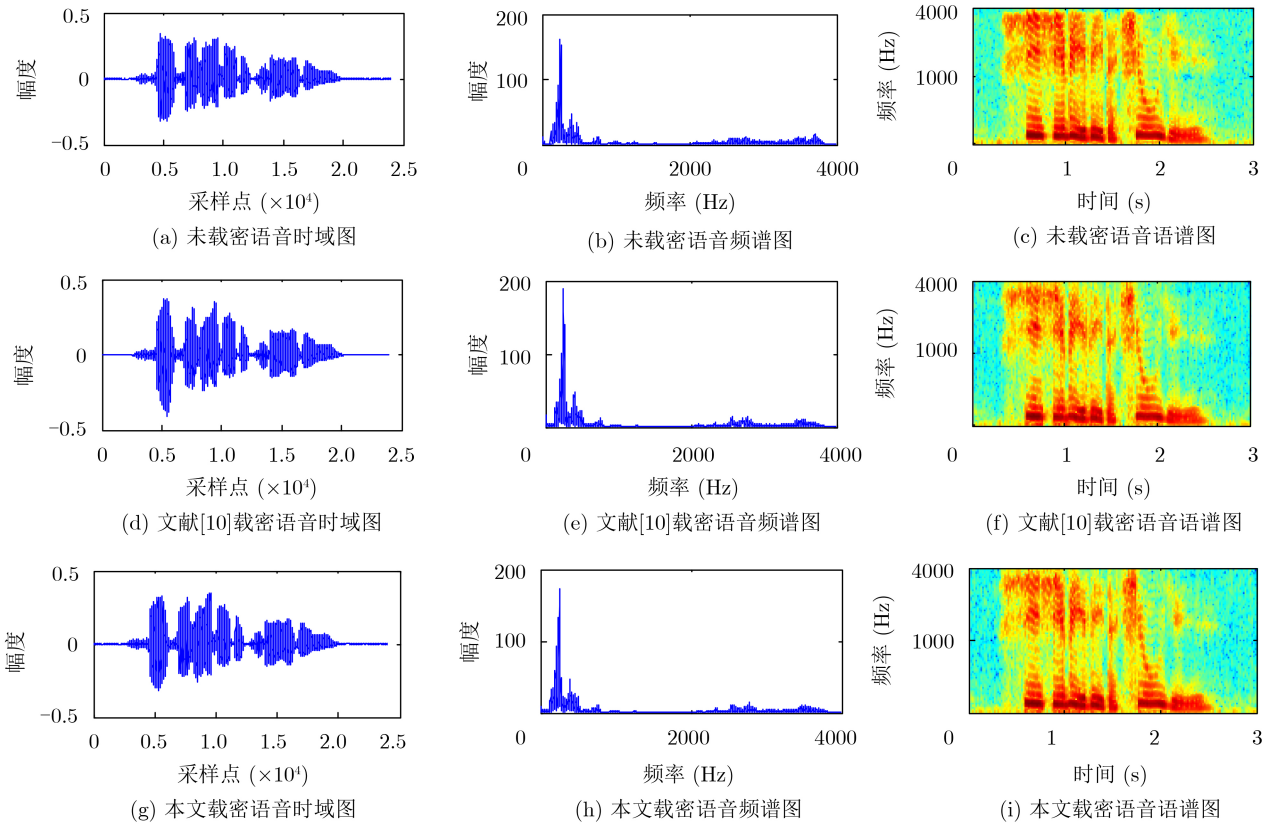

图 3 载密语音与未载密语音对比图

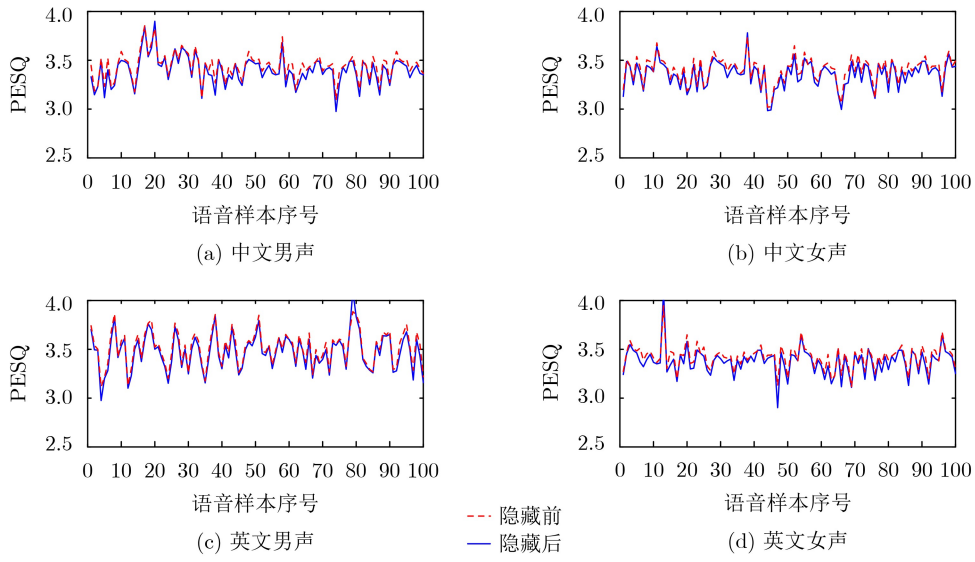

图 4 时长为3 s样本的PESQ值对比

与未载密重构语音样本的PESQ值。本文对表2从两方面进行分析：(1) 对于3 s, 10 s语音样本，其PESQ平均恶化率和PESQ最大恶化率均在标准误差范围之内，表明本文所提出的语音信息隐藏算法对语音质量有着微乎其微的影响；(2)男声样本的PESQ恶化率均小于女声样本，说明此算法对女声的影响大于对男声的影响，这可能是由于女声的基音频率有着更宽的范围。

表3描述了本文隐藏算法与文献[6]及文献[10]隐藏算法的PESQ值对比结果。对于3 s语音样本，PESQ值恶化率分别为–0.96%, –0.96%和–0.91%，本文隐藏算法比文献[6]和文献[10]均低0.05%；对于10 s语音样本，PESQ值恶化率分别为–0.89%, –0.94%和–0.82%，本文隐藏算法分别比文献[6]和文献[10]低0.07%和0.12%。实验结果表明：(1)本文隐藏算法对语音质量的影响低于文献[6]和文献[10]的隐藏算法；(2)随着时长的增加，本文算法语音质量的优势进一步凸显。

4.3 算法实时性分析

语音信息隐藏算法的实时性一般采用语音帧的处理时间来衡量。在3 s语音库中，分别选择50条中文男声、50条中文女声、50条英文男声和50条英

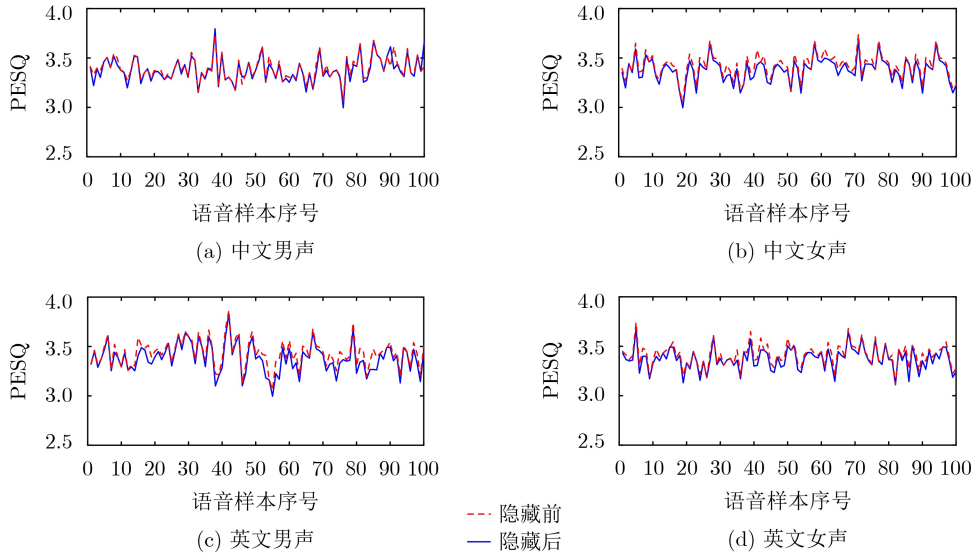

图 5 时长为10 s样本的PESQ值对比

表 2 本文隐藏算法的PESQ统计值

	载密语音样本				未载密语音样本				PESQ改变率(%)			
	3 s语音样本											
	CSM	CSW	ESM	ESW	CSM	CSW	ESM	ESW	CSM	CSW	ESM	ESW
平均值	3.4237	3.3440	3.4165	3.4750	3.4429	3.3995	3.4259	3.5193	–0.55	–1.63	–0.27	–1.20
最大值	3.8999	3.7815	4.0995	4.1015	3.8262	3.7406	3.8863	3.9458	1.93	1.09	5.49	3.94
最小值	2.9756	2.9615	2.9756	2.9022	3.0930	3.1701	3.1160	3.1265	–3.80	–6.58	–4.51	–7.17
	10 s语音样本											
	CSM	CSW	ESM	ESW	CSM	CSW	ESM	ESW	CSM	CSW	ESM	ESW
平均值	3.4095	3.3528	3.3990	3.3822	3.4218	3.3900	3.4102	3.4362	–0.37	–1.09	–0.33	–1.50
最大值	3.7954	3.7102	3.7614	3.8714	3.7406	3.6783	3.7406	3.8632	1.47	0.87	2.12	0.21
最小值	2.9965	2.9113	3.1003	2.9965	3.0590	3.0180	3.1123	3.0590	–2.00	–3.53	–1.20	–2.00

表 3 隐藏算法的PESQ统计对比(%)

隐藏算法	3 s语音样本					10 s语音样本				
	CSM	CSW	ESM	ESW	均值	CSM	CSW	ESM	ESW	均值
文献[6]隐藏算法	–0.49	–1.05	–0.93	–1.37	–0.96	–0.62	–1.44	–0.29	–1.22	–0.89
文献[10]隐藏算法	–0.59	–1.63	–0.28	–1.35	–0.96	–0.52	–1.42	–0.35	–1.47	–0.94
本文隐藏算法	–0.55	–1.63	–0.27	–1.20	–0.91	–0.37	–1.09	–0.33	–1.50	–0.82

文女声。对于每条语音，分别采用本文算法、文献[6]算法和文献[10]算法隐藏秘密信息，并统计语音处理时长，最终统计计算得到每帧处理时间。实验结果如表4所示。

从表4可以看出，(1)文献[6]隐藏算法的帧处理时间最长，原因是在寻找临近码字时，该算法要计算一遍所有码字同此码字的欧式距离。(2)本文隐藏算法实时性均低于文献[6]和文献[10]，原因为本文算法引入矩阵编码算法，有的自适应码本不需要重新进行搜索，大大节约了处理时间。

表 4　语音帧处理时间统计对比(ms)

隐藏算法	CSM	CSW	ESM	ESW	均值
文献[6]隐藏算法	7.43	8.48	8.24	8.94	8.27
文献[10]隐藏算法	7.02	7.96	7.89	8.74	7.90
本文隐藏算法	6.67	7.64	7.55	8.32	7.56

4.4　算法抗检测性对比

文献[16]针对文献[10]和文献[11]基音调制信息隐藏算法提出了有效的语音隐藏分析，最大检测率分别达到99.0%和100%。因此，本文采用文献[16]的检测方法对语音信息隐藏算法进行隐藏对比分析。数据集包括中文男声、中文女声、英文男声、英文女声，每种语音样本各包含500个语音片段，总共2000个语音片段。其中，每个语音片段时长6 s。

4.4.1　不同时长下的抗检测性分析

为了测试本文隐藏算法抵抗基于码书关联网络检测算法的能力，令嵌入率 $t=1$，数据集时间长度为0.1～1.0 s(步长为0.1 s), 1.0～6.0 s(步长为1.0 s)。实验获得的检测率结果如表5所示。

表 5　在满嵌入率下两种隐藏算法检测率(%)

隐藏方法	语音种类	样本时长(s)														
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	2.0	3.0	4.0	5.0	6.0
文献[10]隐藏算法	英文	85.40	88.00	88.50	89.25	90.10	91.45	91.40	92.40	92.95	93.70	96.20	96.95	97.15	97.65	97.35
	中文	86.80	88.65	90.20	90.50	91.20	92.25	93.10	94.25	94.70	94.05	96.80	97.20	98.15	97.75	97.95
文献[11]隐藏算法	英文	87.65	90.11	90.81	91.65	92.89	94.00	94.15	95.01	95.20	96.64	97.20	98.70	100.00	100.00	100.00
	中文	90.70	92.95	94.00	94.55	95.79	96.83	97.16	98.22	98.30	98.89	99.16	99.99	100.00	100.00	100.00
本文隐藏算法	英文	43.20	46.43	46.95	47.33	48.39	49.51	49.44	50.36	51.64	51.93	54.67	55.68	55.92	56.04	55.94
	中文	45.51	46.91	47.56	47.48	48.72	49.70	49.53	52.14	52.37	52.87	55.06	55.86	56.09	56.35	55.63

从表5可以得出：(1)随着样本时长增加，3种隐藏算法的检测率均不断增加并逐渐趋于稳定，原因为更长的序列提供了更多的码字相关性，因此可以更准确地构建码书关联网络。(2)当采用文献[10]隐藏算法时，最大检测率为98.15%，最小值为85.40%；采用文献[12]隐藏算法时，最大检测率为100%，最小值为97.65%；采用本文隐藏算法时，最大检测率为56.35%，最小值为43.20%。说明本文算法抗检测性较文献[10]和文献[11]有一定优势。

图6给出文献[10]隐藏算法与本文隐藏算法的平均检测率与语音片段时长的关系图。

从图6可以得出：(1)随着语音时长增加，隐藏检测正确率也随之增加，最终趋于稳定；(2)本文隐藏方法在任何时长下的平均检测率均低于文献[10]，且最大值仅为56.35%。因此本文隐藏算法抗检测性远远优越于文献[10]所提出的方法。

4.4.2　不同嵌入率下的抗检测性分析

取时长为6 s的中、英文样本，令嵌入率为10%～100%(步长为10%)，实验结果如表6。

由表6可得：(1)对于3种隐藏算法，中文检测率在大多数嵌入率下高于英文检测率。原因可能是中文有412个音节，相对于英文更容易构建码书关

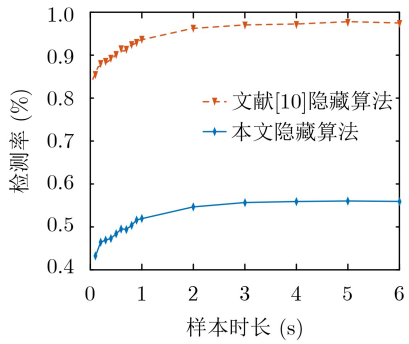

图 6 压缩语音码流在不同样本时长下的平均检测率

联模型；(2)在任何一种嵌入率状态下，本文算法检测率远低于文献[10,11]，说明本文隐藏算法的抗检测性较文献[10,11]具有一定优势。

图7给出了文献[10]隐藏算法和本文隐藏算法下嵌入率与检测率的关系图。

由图7可以得出：(1)随着嵌入率的增加，隐藏算法的检测率也随之增加。其中，在嵌入率低于30%时，检测速率增加较快；(2)本文隐藏算法检测率远远低于文献[10]隐藏算法的检测率，说明本文算法在抗检测性方面优越于文献[10]所提出的算法。

表 6 在不同嵌入率下3种隐藏算法检测率(%)

隐藏方法	语音种类	嵌入率(%)									
		10	20	30	40	50	60	70	80	90	100
文献[10]隐藏算法	英文	50.61	57.64	75.53	80.29	82.81	85.63	86.35	90.58	94.32	97.35
	中文	51.93	58.49	75.28	80.56	82.49	85.27	87.92	92.56	95.63	97.95
文献[11]隐藏算法	英文	53.94	60.89	78.66	83.49	85.91	88.69	89.35	93.74	97.25	100.00
	中文	54.53	61.67	79.55	84.88	86.77	89.55	91.83	95.66	98.55	100.00
本文隐藏算法	英文	16.32	17.52	30.87	39.48	41.16	44.15	45.62	48.12	52.61	55.94
	中文	15.22	17.35	30.56	39.52	41.74	44.16	44.59	49.31	52.38	55.63

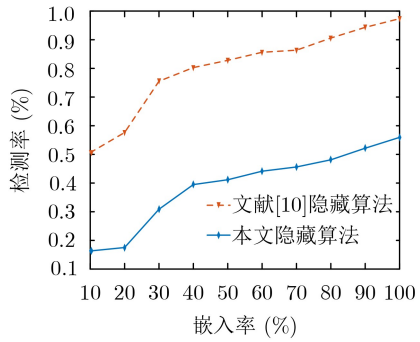

图 7 压缩语音码流在不同嵌入率下的平均检测率

5 结论

本文针对基于基音预测过程的信息隐藏方法，提出了有效的改进方法。根据码字之间的关联特性选择载体子帧，并且在基于基音预测的信息隐藏过程中引入了随机位置选择算法和矩阵编码方法，增加了嵌入透明度，减小了载体改变率，提高了载密比特流的抗检测性。实验结果表明，相对于目前性能优越的信息隐藏算法[10]，本文算法语音质量更高，针对基于码书关联网络的支持向量机的检测更具有优越性。

参 考 文 献

[1] PETITCOLAS F A P, ANDERSON R J, and KUHN M G. Information hiding - a survey[J]. *The IEEE*, 1999, 87(7): 1062–1078. doi: 10.1109/5.771065.

[2] 丁琦, 平西建. 基于脉冲位置参数统计特征的压缩域语音隐写分析[J]. 计算机科学, 2011, 38(1): 217–220. doi: 10.3969/j.issn.1002-137X.2011.01.051.

DING Qi and PING Xijian. Steganalysis of compressed speech based on statistics of pulse position parameters[J]. *Computer Science*, 2011, 38(1): 217–220. doi: 10.3969/j.issn.1002-137X.2011.01.051.

[3] DITTMANN J, HESSE D, and HILLERT R. Steganography and steganalysis in voice-over IP scenarios: Operational aspects and first experiences with a new steganalysis tool set[J]. *SPIE*, 2005, 5681: 607–618.

[4] TIAN Hui, SUN Jun, CHANG C C, *et al*. Detecting bitrate modulation-based covert voice-over-IP communication[J]. *IEEE Communications Letters*, 2018, 22(6): 1196–1199. doi: 10.1109/LCOMM.2018.2822804.

[5] TIAN Hui, SUN Jun, CHANG C C, *et al*. Hiding information into voice-over-IP streams using adaptive bitrate modulation[J]. *IEEE Communications Letters*, 2017, 21(4): 749–752. doi: 10.1109/LCOMM.2017.2659718.

[6] XIAO Bo, HUANG Yongfeng, and TANG Shanyu. An approach to information hiding in low bit-rate speech stream[C]. 2008 IEEE Global Telecommunications Conference, New Orleans, USA, 2008: 1–5.

[7] TIAN Hui, LIU Jin, and LI Songbin. Improving security of quantization-index-modulation steganography in low bit-rate speech streams[J]. *Multimedia Systems*, 2014, 20(2): 143–154. doi: 10.1007/s00530-013-0302-8.

[8] CHIANG Y K, TSAI P, and HUANG Fenglong. Codebook partition based steganography without member restriction[J]. *Fundamenta Informaticae*, 2008, 82(1/2): 15–27.

[9] LI Songbin, JIA Yizhen, and KUO C C J. Steganalysis of QIM steganography in low-bit-rate speech signals[J]. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2017, 25(5): 1011–1022. doi: 10.1109/TASLP.2017.2676356.

[10] HUANG Yongfeng, LIU Chenghao, TANG Shanyu, *et al*. Steganography integration into a low-bit rate speech

codec[J]. *IEEE Transactions on Information Forensics and Security*, 2012, 7(6): 1865–1875. doi: 10.1109/TIFS.2012.2218599.

[11] LIU C, BAI S, and HUANG Y. An information hiding algorithm in G. 729a based on pith prediction[C]. The 10th National Academic Conference on Information Hiding and Multimedia Information Security, Beijing, China, 2012: 15–18.

[12] LIN Zinan, HUANG Yongfeng, and WANG Jilong. RNN-SM: Fast steganalysis of VoIP streams using recurrent neural network[J]. *IEEE Transactions on Information Forensics and Security*, 2018, 13(7): 1854–1868. doi: 10.1109/TIFS.2018.2806741.

[13] YANG Wanxia, TANG Shanyu, LI Miaoqi, *et al*. Steganalysis of low embedding rates LSB speech based on histogram moments in frequency domain[J]. *Chinese Journal of Electronics*, 2107, 26(6): 1254–1260. doi: 10.1049/cje.2017.09.026.

[14] WU Zhijun, Gao Wei, and YANG Wei. LPC parameters substitution for speech information hiding[J]. *The Journal of China Universities of Posts and Telecommunications*, 2009, 16(6): 103–112. doi: 10.1016/S1005-8885(08)60295-2.

[15] HUANG Yongfeng, TANG Shanyu, and YUAN Jian. Steganography in inactive frames of VoIP streams encoded by source codec[J]. *IEEE Transactions on Information Forensics and Security*, 2011, 6(2): 296–307. doi: 10.1109/TIFS.2011.2108649.

[16] 李松斌, 贾已真, 付江云, 等. 基于码书关联网络的基音调制信息隐藏检测[J]. 计算机学报, 2014, 37(10): 2107–2117. doi: 10.3724/SP.J.1016.2014.02107.
LI Songbin, JIA Yizhen, FU Jiangyun, *et al*. Detection of pitch modulation information hiding based on codebook correlation network[J]. *Chinese Journal of Computers*, 2014, 37(10): 2107–2117. doi: 10.3724/SP.J.1016.2014.02107.

[17] TIAN Hui, JIANG Hong, ZHOU Ke, *et al*. Transparency-orientated encoding strategies for voice-over-IP steganography[J]. *The Computer Journal*, 2012, 55(6): 702–716. doi: 10.1093/comjnl/bxr111.

[18] WESTFELD A. F5-A steganographic algorithm: High capacity despite better steganalysis[C]. The 4th International Workshop on Information Hiding, Pittsburgh, USA, 2001: 289–302.

[19] ITU-T. ITU-T P.862 Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs[S]. Geneva, Switzerland: International Telecommunications Union, 2001.

吴志军：男，1965年生，教授，博士生导师，研究方向为网络和信息安全.

李常亮：男，1993年生，硕士生，研究方向为信息安全.

李　荣：女，1995年生，硕士生，研究方向为信息安全.