异步模-数-模语音加密中的伪频插入置乱方法*

张海林 王育民
(西安电子科技大学,西安 710071)

摘要 本文在 L. S. Lee (1984, 1986) 异步模-数-模语音加密方案的基础上提出了自适应伪频插入置乱方法. 这种方法有利于降低密话节奏感,提高异步加密方案的保密性,且可保证良好的恢复话音质量.

关键词 语音保密;模-数-模置乱;伪频插入方法

1. 引言

80 年代中期,L. S. Lee[1,2] 提出了一类无须帧同步的模-数-模语音加密体制. 这类体制等效于细化的频带分割置乱,置乱度较高,能在异步的情况下解密得到令人满意的恢复话音质量. 故被认为是一种很实用的语音置乱方案而受到重视和研究. 但是,像其它一些模-数-模加密方法一样,这种短时的自身扰乱,基本保留了原始语音信号的短时能量分布特点,即密话中存在很强的语音节奏感,这无疑对猜测明话内容是有利的. 另外,L. S. Lee 异步方案受频谱分析法破译的威胁很大[3,4]. 针对以上问题,本文提出一种适用于异步加、解密方案的自适应伪频插空置乱算法. 这种算法有利于降低方案的密话节奏感,增加破译难度. 实验结果验证了算法的可行性.

2. 伪频插空置乱的基本思想

语音信号是非平稳的随机过程,具有很强的突发-间歇特性,这种特性的基本应用就是语音插空技术(TASI 或 DSI). 另外,强相关的语音信号经正交变换后能量发生集中分布,因而在信源编码中就有了区域抽样编码和阈值编码技术. 以上性质在模-数-模语音加密中的应用就是伪频插空置乱[4,5],即在语音信号的时空和频空状态下,插入伪语音信号参加置乱,以进一步提高模-数-模加密体制的保密度.

变换域伪频插空置乱思想可以推广到异步模-数-模加密体制中,但设计方法必须兼顾异步解密性要求. 从统计角度看,语音信号的能量大部分集中于带内低频端的几个共振峰频率(400～2500Hz)附近;高频端,特别是 3kHz 以上的能量很小. 另外,由于耳窝接收器对高频分量不敏感,允许高端分量有所失真. 据此,可以将音频带宽(300～3400Hz)由低到高分为显著段和非显著段,语音的基本信息由显著段决定,非显著段可认为具有一

1990.08.17 收到,1991.12.13 定稿.
* 国家教委博士点基金资助课题.

定的多余度。异步伪频插空置乱的基本考虑是，先对通话状态进行自适应检测。通话时，在非显著段插入一组高电平随机信号，限定最小幅度为 A_L；默话时，在显著段插入伪语音信号，非显著段插入状态编码信号。经过如此处理后再进行置乱加密。

为充分掩盖密话中节奏感，插入的伪语音信号应具有与明话信号相同的统计特征。

3. 算法设计和实现

(1) 算法设计

设 $\hat{X}=\{x_i|_{i=0}^{N-1}\}$，$\hat{Y}=\{y_i|_{i=0}^{N-1}\}$ 分别是明话和密话信号当前帧频域矢量，维数为 N。语音有效带（300～3400Hz）内的系数个数为 S，$\hat{X}$ 或 $\hat{Y}$ 中有效系数所对应的序号为 $[300N/f_s]\leqslant i\leqslant[3400N/f_s]$，$f_s$ 为抽样速率。如果选定 300～3000Hz 为显著段，3000～3400 Hz 为非显著段，则可确定显著系数集合 $\hat{X}_S=\{x_i\in\hat{X},[300N/f_s]\leqslant i\leqslant[3000N/f_s]\}$，其长度为 q，非显著系数集合为 $\hat{X}_{NS}=\{x_i\in\hat{X},[3000N/f_s]<i\leqslant[3400N/f_s]\}$，其长度为 r。显然，$S=q+r$。

加密端，首先检查当前帧显著系数集合 $\hat{X}_S$ 中系数电平，并逐个与发送门限 β_t 比较。

(a) 如果 $\hat{X}_S$ 中至少有一个系数电平 $|x_i|<\beta_t$，说明此刻处于通话状态。用一组高电平随机数替代当前帧的非显著系数集合 $\hat{X}_{NS}$，$\hat{X}_S$ 不变。插入后的 S 个混合系数之间进行置换加密。过程如图 1(a)和(b)所示。

(b) 如果 $\hat{X}_S$ 中所有系数电平 $|x_i|<\beta_t$，则认为此刻处于默话态。这时用一组类话音的强噪声替代 $\hat{X}_S$，而 $\hat{X}_{NS}$ 中系数全部置零(或低电平噪声插入)，然后在有效集合内进行置换加密。此过程示于图 1(c)和(d)。

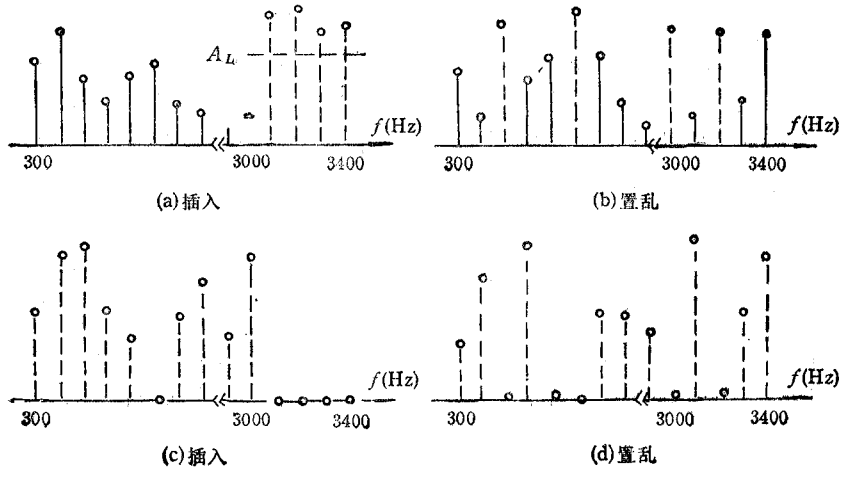

图 1 插空置乱示意图

由于插空和置乱均在有效集合中进行，故加密输出信号 Y 频带不扩展。发送门限 β_t 应由输入信号平均能量控制在每一帧内进行适当调整。

解密端，先对 $\hat{Y}$ 中有效系数进行逆置换解密，即完成由图 1(b)到(a)或由图 1(d)到(c)的过程。将非显著集合 $\hat{X}_{NS}$ 中系数逐一与接收门限 β_r 比较。

(c) 如果 $\hat{X}_{NS}$ 中多数系数电平低于接收门限 β_r，则译为默话态。这说明 $\hat{X}_S$ 中插入了

伪语音信号，将它们删除，使 $\hat{X}=0$.

(d) 如果 $\hat{X}_{NS}$ 中多数系数电平高于 β_r，则译为通话态．保留 $\hat{X}_S$ 中系数，删除 $\hat{X}_{NS}$ 中系数，使 $\hat{X}_{NS}=0$

综合输出为解密信号．β_r 与接收信噪比有关，必须慎重选择以优化(c),(d)的判决性能．

(2) 实现方法

本方法的目的就是在异步的条件下能够实现上述的伪频插空置乱算法．因此方案的实现结构与 L. S. Lee 异步频域置乱方案基本相同，只在此基础上增加伪频插入和删除单元，如图 2(a)和(b)所示．

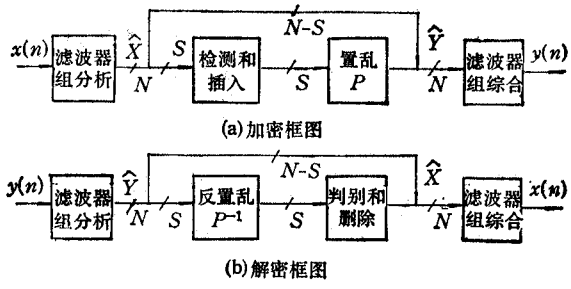

图 2　实现方法示意图

图 2 中的滤波器组分析和综合单元可用 Rabiner 方法[6]快速数字实现．整个全双工加、解密系统可用单片 DSP-TMS 32020 或 TMS 320C 25 实时实现．维数 N 可做到 64,128,256 等．

4. 性能分析和实验结果

(1) 保密性　本文的方法中去掉了语音高频端的部分多余度，而且还在时空和频空状态下插入了高电平的伪信号参加置乱．这种方法符合 Shannon 关于理想加密、扩散和混叠原则．一方面，插入的伪语音可有效地掩盖密话中的间歇和节奏感，直接干扰对密话的被动窃听(参看图 3(a)和(b)所提供的加密实验波形)；另一方面，不知道置换密钥 P 就无法消除伪语音的干扰．客观上增加了主动窃听的难度．而伪语音的插入又起到保护密钥 P 的作用，因为伪频插入置乱算法属于概率型随机加密体制，这种体制具有天然的抗击明密文对密钥的分析能力．

(2) 可靠性　如果接收译码器的判决是正确的，就可以将插入的伪信号完全消除．解密输出将得到 300～3000Hz 纯语音信号，或许还有时延和非线性相位失真[3]（参看图 3(c)和(d)提供的异步解密实验波形）．

事实上，由于异步解密和噪声的影响，第 3(1) 节中(c)和(d)的判决错误是在所难免的．设 $\hat{X}_{NS}$ 中每个系数的比较错误率为 p_e，注意到第 3(1) 节中(c)和(d)的大数判决规则，很快就可得到译码器的状态错误概率为

$$P_{de}=\sum_{i=t+1}^{r}\binom{r}{i}\cdot p_e^i(1-p_e)^{r-i} \tag{1}$$

式中 $t=[(r-1)/2]$ 是二进制重复码 $(r,1)$ 的纠错能力．比较错误率 p_e 与 A_L 和 β_r 有关．

$$p_e\leqslant P_P\cdot P(|\hat{x}_i'|\geqslant\beta_r/|\hat{x}_i|=0)+P_S\cdot P(|\hat{x}_i'|<\beta_r/|\hat{x}_i|\geqslant A_L) \tag{2}$$

$P_P, P_S=1-P_P$ 分别是默话和通话概率，汉语的 P_S 一般在 0.35～0.45．如果仅考虑高斯加性干扰，容易求得最佳的接收门限 β_r 为

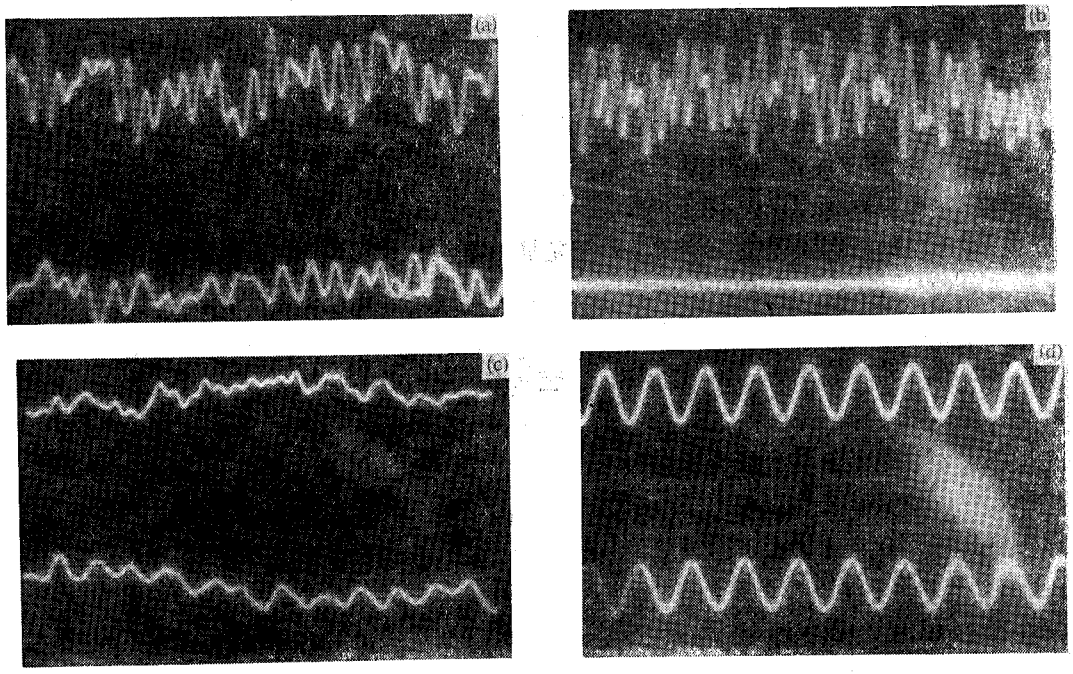

图 3 伪频插空置乱异步加、解密系统实验波形示意

(a) 通话时加密(上)、解密(下)输出
(b) 空话时加密(上)、解密(下)输出
(c) 原始语音(上)、解密(下)波形
(d) 正弦波输入(上)、解密输出(下)

$$\beta_r^* \approx \frac{\delta_N^2}{A_L} \cdot \ln (P_P/P_S) + \frac{A_L}{2} \tag{3}$$

适当地选择 β_r 和 A_L 可保证 $p_e < 10^{-2}$,使译码错误可以接受. 例如: $N = 128, r = 6$, $p_e = 10^{-2}$, 由(1)式得到 $P_{de} \approx 1.94 \times 10^{-5}$. 这意谓着平均每 10 万帧信号(平均约 6min)内恢复信号瞬时谱受到两次(16ms)污染,经过综合滤波器后将持续约 80ms (插值滤波窗函数长度为 $2 \times 5 \times N$). 这对于相关性很强的语音信号来说, 影响是足够小的. 实验结果也验证了这一点.

另外,如果在第 3(1) 节中的(c)和(d)步骤中采用欧氏距离软译码,则可靠性会更高.

5. 结束语

分析与实验表明: 本文的伪频插空置乱算法用在异步模-数-模加密体制中是可行的. 有利于提高异步模-数-模加密体制的安全性, 且保证良好的恢复语音质量. 实验还表明: 这种算法同样适用于滚动密钥异步模-数-模加密体制[4].

参 考 文 献

[1] L. S. Lee et al., *IEEE Trans. on COM*, **COM-32** (1984)4, 444—456.
[2] L. S. Lee et al., *IEEE J. of SAC.*, **SAC-4**(1986) 2, 280—287.
[3] 张海林,王育民,胡征,电子学报,**19**(1991)4,35—39。
[4] 张海林,模-数-模语音加密的理论与实现,西安电子科技大学博士学位论文,西安,1991 年.

[5] Akira, Matsunaga, *IEEE J. of SAC.*, **SAC-7** (1989)4, 540—547.
[6] L. R. Rabiner, R. W. Schafer, Digital Processing of Speech Signals, Prentice-Hall. Inc. New York, (1978).

A DUMMY SPECTRUM INSERTION ENCRYPTION METHOD USED IN ASYNCHRONOUS SPEECH SCRAMBLER

Zhang Hailin, Wang Yumin
(*Xidian University, Xi'an* 710071)

Abstract An adaptive dummy spectrum insertion method which can be applied to L. S. Lee's asynchronous speech scrambler is presented. It is concluded that the method shows high-level security and good voice quality.

Key words Speech security; A/D/A scrambler; Dummy spectrum insertion method