

基于迭代结构的联合信源信道 GSM 语音解码

周琳 吴镇扬

(东南大学无线电工程系 南京 210096)

摘要: 由于算法复杂度和延迟的限制, 语音编码参数仍然保留了一定的冗余信息, 这些冗余量表现为参数的非等概分布和时间相关性。该文在比特级上利用信源编码参数的冗余信息在软输入、软输出信道解码和信源参数估计之间进行迭代运算, 并将此联合解码算法运用于 GSM EFR 系统。仿真结果表明此算法降低了编码参数的 BER, 提高了解码语音的信噪比, 具有一定的可行性和有效性。

关键词: 迭代算法, 软输入软输出解码器, 参数估计

中图分类号: TN911.22 **文献标识码:** A **文章编号:** 1009-5896(2005)06-0884-04

Iterative Source-Channel Decoding for GSM

Zhou Lin Wu Zhen-yang

(Dept. of Radio Engineering, Southeast University, Nanjing 210096, China)

Abstract In speech signal transmission system, residual redundancy is still left in the parameters after source encoding. The redundancy includes non-uniform distribution and time correlation of subsequent frame. Exploiting the redundancy on bit-level, the paper is proposed iterative algorithm based on soft-in/soft-out channel and parameter estimation, and applies the algorithm to GSM EFR system. The result shows this method reduces the BER and improves SNR of decoded speech.

Key words Iterative decoding, Soft-in/soft-out decoder, Parameter estimation

1 引言

数字通信系统由信源编码(压缩)和信道编码(误差控制)组成。一般情况下, 这两部分独立设计, 分别实现其功能。为了在一定的复杂度和延迟条件下提供更好的通信系统性能, 联合信源信道编解码方案逐步成为研究的热点。联合编解码设计主要分为: (1) 针对信道特性的信源编码, 重点为发送端的信源参数量化编码。由于信源编码考虑了编码参数的抗噪声性能, 一般而言不再进行信道编码, 比较常用是信道最佳矢量量化^[1]。(2) 针对信源编码参数冗余量的信道信源联合解码, 重点为接受端的信道解码过程, 如信源控制信道解码^[2,3]以及利用信道软输出信息的参数估计^[4]。本文基于迭代思想提出一种信源信道联合解码方案, 基本概念来源于 Turbo-code^[5]的迭代过程。本算法由软输入软输出信道解码和信源参数估计模块两部分组成。信源编码参数的冗余量包括参数的非等概分布和参数的记忆性; 信道编码引入的可控冗余用于检错、纠错, 因此信道、信源解码可交换使用这两类冗余信息, 进行迭代运算, 在不改变信源、信道编码结构的前提下, 提高系统的整体性能。仿真实验以 GSM EFR(Global

System for Mobile communications: Enhanced Full Rate)语音编码为例, 利用参数的冗余量进行迭代解码。本文将首先讨论 GSM EFR 语音编码算法以及编码参数的冗余量; 第 3 节详细介绍基于迭代过程的联合解码系统框架; 最后给出实验结果和分析。

2 GSM EFR 语音编码和冗余量

GSM EFR 语音编码^[6]是由欧洲 ETSI 提出的 12.2kbit/s 语音压缩算法。以线性预测为基础, GSM EFR 用一组线谱对参数(LSF 系数)和激励信号表示一帧语音, 其中激励信号用自适应码书和固定码书进行近似。5 组线谱对系数采用分裂矩阵量化, 每一组矢量分别为 7, 8, 9, 8, 6bit 编码; 一帧分为 4 个子帧, 每一子帧的自适应码书增益和固定码书增益分别用 4, 5bit 量化; 9bit 用于量化第 1, 3 子帧的自适应码书延迟, 第 2, 4 子帧码书延迟采用 4bit 的差值编码; 每一子帧的固定码书索引值量化需要 35bit, 这样共 244bit 用于量化编码一帧参数。

各个参数对信道噪声的灵敏度依次为^[7]: LSF、固定码书增益、自适应码书延迟、自适应码书增益以及固定码书索

引;除了LSF系数,激励信号中对合成音质的重要性依次为自适应码书参数、固定码书参数以及接收端线性滤波器零输入响应^[8]。基于以上不同编码参数的特性,我们主要分析5组LSF系数的高3位比特、自适应码书增益、固定码书增益的冗余量。

若 $\{X_i; i=1,2,\dots\}$ 表示一离散信源,取值集 χ 有限,则信源的熵定义为

$$H_X = \lim_{n \rightarrow \infty} H(X_n | X_1, \dots, X_{n-1}) \quad (1)$$

$$H(X_n | X_1, \dots, X_{n-1}) = S - \sum_{\{x_1, \dots, x_n\} \in \chi^n} P(X_1 = x_1, \dots, X_n = x_n) \cdot \log_2 [P(X_n | X_1 = x_1, \dots, X_{n-1} = x_{n-1})] \quad (2)$$

H_X 表示在对 $\{X_i\}$ 编码时不引入任何误差所需的最小编码比特数;也就是说,如果 $\{X_i\}$ 以 R 的码率进行编码,那么 $\rho = R - H_X$ 就是编码所产生的冗余量。

为了计算每一参数的熵,GSM EFR编码器首先对语音文件进行编码(语音文件总计272737帧,90min,语音数据库由The Centre for Speech Technology Research, University of Edinburgh提供)以得到LSF和增益参数;同时假设每一参数为一阶平稳马尔可夫过程,这样 $P(X_n | X_1, \dots, X_{n-1}) = P(X_n | X_{n-1})$,文献[9]提供GMTD算法计算各参数的概率转移矩阵,利用式(2)得到各参数的熵率和冗余量。

计算结果见表1。总的冗余量 ρ_T 分为两个部分: $\rho_T = \rho_D + \rho_M$,其中 ρ_D, ρ_M 分别是由参数的非等概分布和帧间相关性引入的冗余。参数共占51bit,冗余信息为11.66bit,约为22.86%。由于LSF参数采用了一阶MA预测模型,对残差信号进行量化,因此由帧间相关性引入的冗余量要明显小于两个增益参数;同时固定码书增益的冗余量最多,为1.41bit,约占编码比特的28.6%。

表1 LSF及码书增益参数的冗余信息

参数	码率(R)	熵(H_X)	ρ_T	ρ_D	ρ_M
LSF1	3	2.2418	0.7582	0.5630	0.1953
LSF2	3	2.1855	0.8145	0.5338	0.2807
LSF3	3	2.4514	0.5486	0.4305	0.1181
LSF4	3	2.5436	0.4564	0.2948	0.1616
LSF5	3	2.4165	0.5835	0.2902	0.2933
自适应码书增益	4	3.2940	0.7060	0.2153	0.4907
固定码书增益	5	3.5821	1.4179	0.6360	0.7819

3 迭代算法结构

图1所示为迭代算法的结构框图,发送端信源比特采用卷积码信道编码,基于迭代的联合解码由软输入软输出维特比(Viterbi)信道解码(SOVA)和参数估计两部分组成。类似于Turbo-Code两个解码器之间的迭代运算,我们在信道解码、参数估计之间引入迭代过程。

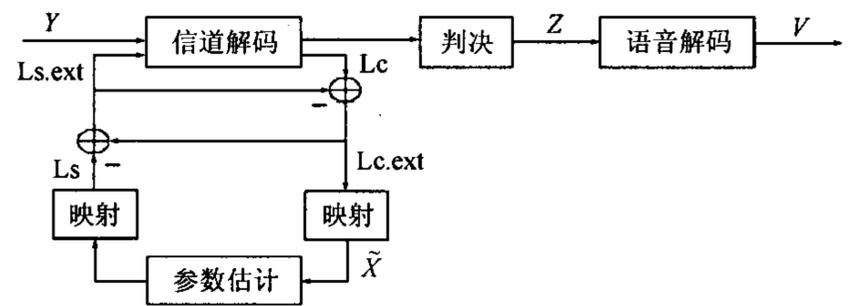


图1 迭代解码系统框图

假设参数 x_i 为一比特矢量: $x_i = \{x_i(1), x_i(2), \dots, x_i(M)\}$,其中 $x_i(m) \in \{1, -1\}$ 。

基于不同参数对合成语音音质的影响,并且考虑到参数高位比特的重要性要大于低位比特,因此在参数估计模块,我们只计算5组LSF系数高6位比特、两个码书增益高4位比特的先验信息,作为信道解码的先验信息,其它参数忽略估计过程,因此这里 x_i 表示LSF系数或两个码书增益; M 为量化比特数(LSF参数 $M=6$,码书增益 $M=4$); i 为时间标号。

基于迭代的联合解码过程如下:维特比信道解码根据信道输出 y 和先验信息 $Ls.ext$ (初始值为0)计算信息比特 $x_i(m)$ 的似然值 $Lc^{[10]}$, Lc 反映了每一信息比特的硬判决值和可信度信息:

$$Lc(x_i(m)) = \log \frac{P(x_i(m)=0)}{P(x_i(m)=1)} = L(x_i(m) | y) + Ls.ext(x_i(m)) \quad (3)$$

由式(3)可知,SOVA算法中,每一信息比特的似然值由两部分组成:信道数据 y 提供的似然值;先验信息提供的似然值 $Ls.ext$,因此由信道数据 y 单独提供的似然值:

$$L(x_i(m) | y) = Lc.ext(x_i(m)) = Lc(x_i(m)) - Ls.ext(x_i(m)) \quad (4)$$

则参数比特的硬判决值 $\tilde{x}_i(m)$ 为

$$\tilde{x}_i(m) = \text{sgn}(Lc.ext(x_i(m))) \quad (5)$$

从信道解码到参数估计的‘映射’模块将每一信息比特的似然值转换为概率值:

$$\left. \begin{aligned} Pc.ext(x_i(m)=0 | y) &= \frac{e^{Lc.ext(x_i(m))}}{1 - e^{Lc.ext(x_i(m))}} \\ Pc.ext(x_i(m)=1 | y) &= \frac{1}{1 - e^{Lc.ext(x_i(m))}} \end{aligned} \right\} \quad (6)$$

参数估计模块根据本帧信道解码的硬判决值 \tilde{x}_i 计算由参数的非等概分布和前一帧已解码参数 z_{i-1} 由帧间相关性提供的当前时刻信源参数比特的概率分布 $P_s(\tilde{x}_i(m))$:

$$P_s(x_i(m) | \tilde{x}_i(1), \tilde{x}_i(2), \dots, \tilde{x}_i(M), z_{i-1}) = \frac{1}{C} P_{c.ext}(x_i(m) | \tilde{x}_i(m)) \cdot P_s(x_i(m) | \tilde{x}_i(1 \setminus m), z_{i-1}) \quad (7)$$

其中, C 为归一化因子, $\tilde{x}_i(1 \setminus m) = \{\tilde{x}_i(1), \dots, \tilde{x}_i(m-1), \tilde{x}_i(m+1), \dots, \tilde{x}_i(M)\}$ 。

式(7)中各个概率值计算如下:

$$P_{c.ext}(x_i(m) | \tilde{x}_i(m)) = \begin{cases} P_{c.ext}(x_i(m) | y), & x_i(m) = \tilde{x}_i(m) \\ 1 - P_{c.ext}(x_i(m) | y), & x_i(m) \neq \tilde{x}_i(m) \end{cases} \quad (8)$$

$$P_s(x_i(m) | \tilde{x}_i(1 \setminus m), z_{i-1}) = \sum_{x_i(1 \setminus m)} P_s(x_i(m), x_i(1 \setminus m) | \tilde{x}_i(1 \setminus m), z_{i-1}) = \sum_{x_i(1 \setminus m)} \frac{P_s(x_i | z_{i-1})}{P_s(x_i(1 \setminus m) | z_{i-1})} \cdot P_{c.ext}(x_i(1 \setminus m) | \tilde{x}_i(1 \setminus m)) \quad (9)$$

信道解码提供了参数比特级的概率分布, 由于参数各比特在信道编码前进行了比特交织, 若同一参数比特相隔的比特距离超过卷积码的约束长度, 则参数概率为各比特概率值的乘积:

$$P_{c.ext}(x_i(1 \setminus m) | \tilde{x}_i(1 \setminus m)) = \prod_{j \neq m} P_{c.ext}(x_i(j) | \tilde{x}_i(j))$$

这样式(7)转化为

$$P_s(x_i(m) | \tilde{x}_i(1), \tilde{x}_i(2), \dots, \tilde{x}_i(M), z_{i-1}) = \frac{1}{C} P_{c.ext}(x_i(m) | \tilde{x}_i(m)) \cdot \sum_{x_i(1 \setminus m)} \frac{P_s(x_i | z_{i-1})}{P_s(x_i(1 \setminus m) | z_{i-1})} \cdot \prod_{j \neq m} P_{c.ext}(x_i(j) | \tilde{x}_i(j)) \quad (10)$$

由式(10), 等式右边第1项为信道数据 y 提供的 $x_i(m)$ 概率分布; 第2项考虑参数的非等概分布, 利用 x_i 参数中其它比特 $x_i(1 \setminus m)$ 结合其信道信息联合提供了关于 $x_i(m)$ 的概率信息, 同时考虑了参数的帧间相关性, 利用前一帧已解码参数 z_{i-1} 提供的条件转移概率 $P_s(\cdot | z_{i-1})$ 。由信道编码可知, 第1项概率值由信息比特 $x_i(m)$ 以及相邻比特通过信道编码结合信道特性共同决定, 由于在信道编码前, 信源编码比特根据其其对解码音质的影响进行了重排, 类似于 Turbo-Code 中的比特交织, 参数 x_i 各个比特相隔一定的比特距离, 这样由信源参数冗余提供的外信息 $P_s(x_i(m) | \tilde{x}_i(1 \setminus m), z_{i-1})$ 与信道提供的信息 $P_{c.ext}(x_i(m) | \tilde{x}_i(m))$ 相互独立, 进行迭代解码有助于提高解码性能。

从参数估计到信道解码的‘映射’将参数比特的概率值转化到对数似然值:

$$L_s(x_i(m)) = \log \frac{P_s(x_i(m) = 0 | \tilde{x}_i(1), \tilde{x}_i(2), \dots, \tilde{x}_i(M), z_{i-1})}{P_s(x_i(m) = 1 | \tilde{x}_i(1), \tilde{x}_i(2), \dots, \tilde{x}_i(M), z_{i-1})} = \log \frac{P_{c.ext}(x_i(m) = 0 | \tilde{x}_i(m))}{P_{c.ext}(x_i(m) = 1 | \tilde{x}_i(m))} + \log \frac{P_s(x_i(m) = 0 | \tilde{x}_i(1 \setminus m), z_{i-1})}{P_s(x_i(m) = 1 | \tilde{x}_i(1 \setminus m), z_{i-1})} = L_{c.ext}(x_i(m)) + L(x_i(m) | \tilde{x}_i(1 \setminus m), z_{i-1}) \quad (11)$$

则由参数估计模块根据信源参数统计特性提供的关于 $x_i(m)$ 比特的外信息为

$$L_{s.ext}(x_i(m)) = L_s(x_i(m)) - L_{c.ext}(x_i(m)) = L(x_i(m) | \tilde{x}_i(1 \setminus m), z_{i-1}) \quad (12)$$

迭代过程结束后, 参数比特最终的硬判决结果为

$$z_i(m) = \text{sgn}(L_c(x_i(m))) \quad (13)$$

参数 z_i 进行语音解码, 得到合成语音, 同时反馈到参数估计模块, 作为下一帧解码的先验信息。

4 仿真结果分析

仿真中采用 GSM EFR 语音编码, 2.6min 的语音作为测试文件, 约为 7828 帧。AWGN 信道, BPSK 调制, 信道编码为 GSM [11] 提供的卷积码, 码率为 $1/2$, 生成多项式为 (23, 33)。一帧参数 BER 以及各个参数 BER 作为评定系统性能指标, 同时考虑解码参数以及解码语音的信噪比。

图2-图4是 LSF、自适应码书增益和固定码书增益的 BER 比较。图5为一帧编码比特的 BER 比较。BER 均有不同程度的降低。

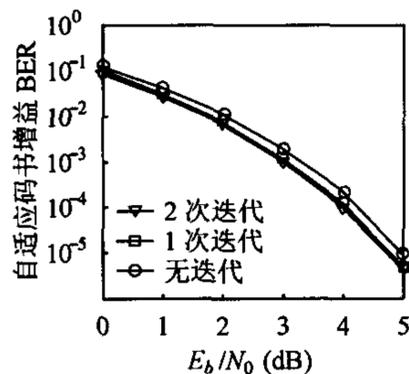


图2 自适应码书增益 BER 与迭代次数的关系

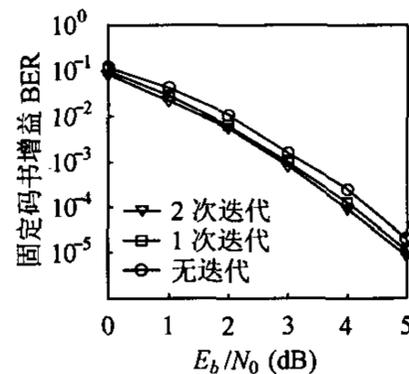


图3 固定码书增益 BER 与迭代次数的关系

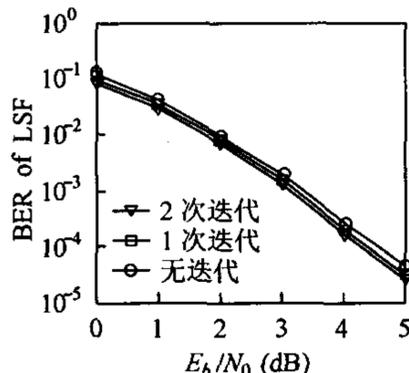


图4 LSF 系数 BER 与迭代次数的关系

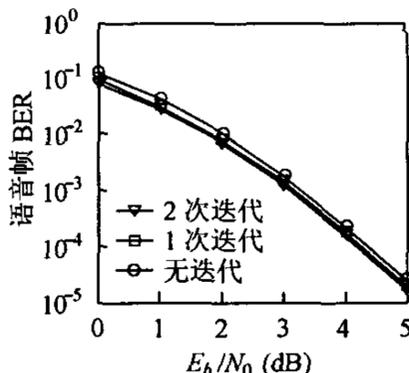


图5 一帧参数 BER 与迭代次数的关系

图6-图8为各参数的信噪比比较,图9为解码语音信噪比比较。由于参数的冗余量有所区别,因此迭代运算对参数信噪比改善的效果也有所不同,固定码书增益的冗余量最多,其BER和信噪比的改善最为明显。在所示信道信噪比 E_b/N_0 范围内,固定码书增益信噪比改善了4.9-2.7dB(2次迭代),2.8-1.9dB(1次迭代),而自适应码书增益和LSF系数解码后的信噪比分别提高了约3.7-1.7dB(2次迭代),2.3-1.3dB(1次迭代)以及2.6-1.5dB(2次迭代),2.0-1.2dB(1次迭代)。

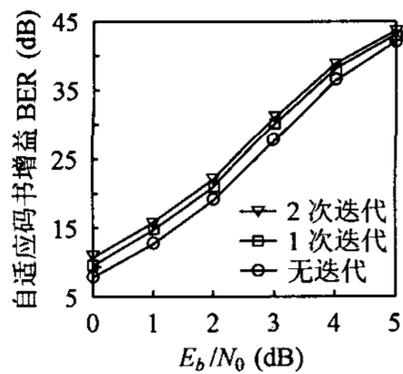


图6 自适应码书增益信噪比与迭代次数的关系

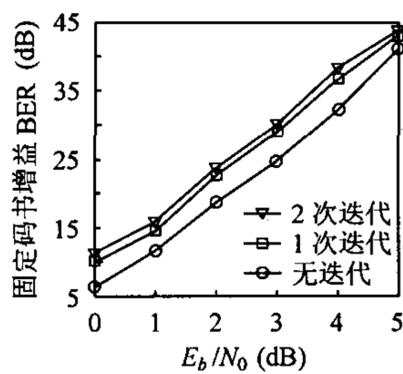


图7 固定码书增益信噪比与迭代次数的关系

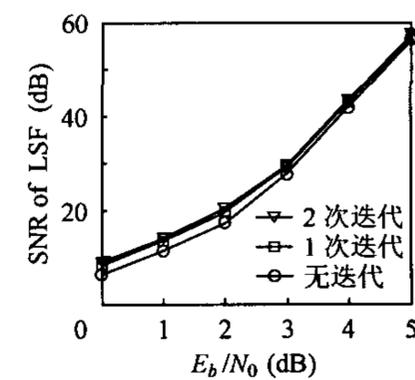


图8 LSF系数信噪比与迭代次数的关系

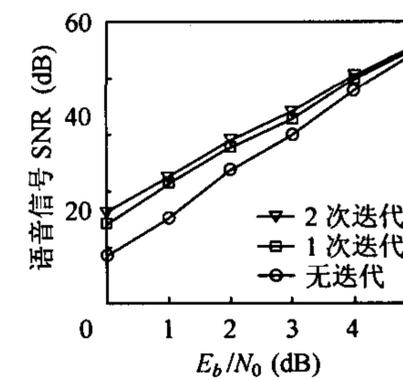


图9 解码语音信噪比与迭代次数的关系

由于自适应码书增益、固定码书增益和LSF系数对解码音质的重要性比较大,因此迭代解码后,解码语音信噪比有了明显提高,分别为3.5-0.7dB(2次迭代),2.5-0.5dB(1次迭代)。

仿真中我们发现第1,2次迭代效果比较明显,从第3次迭代开始对BER以及信噪比改善的效果可以忽略不计。这是由于迭代算法中性能提高的主要因素为信源编码参数冗余量提供了有关编码比特的额外信息,作为先验信息提高信道解码的性能。由式(12)可知,外信息由 $Ls.ext(x_i(m)|\tilde{x}_i(1:m), z_{i-1})$ 提供,式(10)中的条件转移概率是根据训练语音库计算而来的统计量,在上一帧解码参数 z_{i-1} 已知条件下,迭代次数不会影响转移概率值 $P_s(|z_{i-1})$ 的变化;而 $Pc.ext(x_i(1:m)|\tilde{x}_i(1:m))$ 是由信道输出数据 y 决定,随着迭代次数的增加,此概率值与先验信息的相关性也随之增加,对信道解码性能的改善也就逐步减少。

5 结束语

利用语音编码参数的非等概分布和时间相关性提供的冗余信息,我们在信道、信源之间引入迭代解码的概念。LSF和码书增益参数加入外信息的计算,并将其作为先验信息对信道解码进行调整。仿真结果表明降低了BER,并较大幅度的提高了解码参数和解码语音的信噪比。

参考文献

- [1] Farvardin N, Vaishampayam V. Optimal quantizer design for noisy channels: an approach to combined source-channel coding. *IEEE Trans. on Info. Theory*, 1987, 33(6): 827-838.
- [2] Hagenauer J. Source-controlled channel decoding. *IEEE Trans. on Communications*, 1995, 43(9): 2449-2457.
- [3] Lu X B, Koh S N. Hybrid symbol-and bit-based source-controlled channel decoding for MELP speech parameters. *IEEE Communication Letters*, 2003, 7(8):394-396.
- [4] Fingscheidt.T, Vary.P. Softbit speech decoding: a new approach to error concealment. *IEEE Trans.on Speech Audio Processing*, 2001, 9(3): 240-251.
- [5] Berrou C, Glavieux A, Thitimajshima P. Near Shannon limit error-correcting coding and decoding: Turbo-code. ICC-93, Geneva, Switzerland, May 1993, 2: 1064-1070.
- [6] GSM Recommendation 06.11. Enhanced Full Rate Speech Trans-coding.
- [7] Cox R V, Kleijn W B, Kroon P. Robust CELP coders for noisy backgrounds and noisy channels.ICASSP-89, Glasgow, Scotland, May 1989, 2: 739-742.
- [8] Kroon P, Atal B S. Strategies for improving the performance of CELP coders at low bit rates. ICASSP-88, New York, USA, April 1988, 1: 151-154.
- [9] Berchtold A. Estimation of the mixture transition distribution model. <http://www.andreberchtold.com/software.html>
- [10] Hagenauer J, Hoehner P. A Viterbi algorithm with soft-decision outputs and its applications. GLOBECOM-89, Dallas, Texas, USA, November 1989, 3:1680-1686.
- [11] GSM Recommendation 05.03. Digital Cellular Telecommunications System: channel coding.

周琳: 女, 1978年生, 博士生, 研究方向为信源信道联合编码及在语音传输中的应用

吴镇扬: 男, 1949年生, 教授, 博士生导师, 研究领域为通信系统中的信号处理, 智能信号处理、音频、视频信号处理等。