

CDMA 2000移动通信系统语音编码算法研究及仿真^{*}

孙浩海¹, 张雄伟², 邹 霞²

(1. 解放军理工大学通信工程学院研究生²队, 江苏 南京 210007;

2. 解放军理工大学通信工程学院电子信息工程系)

摘 要: 全面介绍了3GPP2公布的CDMA 2000移动通信系统中的可选择模式语音编码算法(SMV), 并着重讨论了SMV算法对语音进行编码和解码过程中的EX-CELP算法、语音分类处理、自适应码本贡献时域去加重和编码速率选择等关键技术。仿真结果表明, 它在灵活调整平均编码速率的同时可以提供优良的话音质量, 满足第三代移动通信系统对语音编码高质量和变速率的要求, 能够使通话服务质量与信道利用率之间达到较为理想的均衡。

关键词: 语音编码; 可选择模式声码器; 扩展码激励线性预测

中图分类号: TN912.3 **文献标识码:** A **文章编号:** CN32-1289(2005)01-0026-06

Research and Simulation of Speech Coding Algorithm of CDMA 2000 Mobile Communication System

SUN Hao-hai¹, ZHANG Xiong-wei², ZOU Xia²

(1. Postgraduate Team² ICE, PLAUST, Nanjing 210007, China;

2. Department of Electronic Information Engineering ICE, PLAUST)

Abstract: The Selectable Mode Vocoder(SMV) algorithm has been chosen by 3GPP2 as the standard for CDMA 2000. In this paper, the whole speech coding algorithm was introduced, with the emphasis on the key techniques throughout the procedure of speech coding and decoding, such as EX-CELP, temporal deemphasising of LTP contribution, speech classification, rate decision, etc. The result of simulation shows that, SMV can provide excellent speech quality with the flexibility of operating at various ABR(Average Bit Rate), and it is able to resolve the problem of tradeoff between the speech quality and CDMA system capacity in an effective way.

Key words: speech coding; SMV; EX-CELP

随着移动通信用户数量和数据业务量的飞速增长, 高效利用频带资源已经成为未来移动通信系统设计的热门话题。对于语音通信业务来讲, 采用变速率多模式语音编码是解决频带利用率问题的好办法。目前, 国际上已有的适合于下一代CDMA通信系统的变速率语音编码方案主要有IS-96C, EVRC和SMV。它们均工作于8.55 kbps, 4.0 kbps, 2.0 kbps和0.8 kbps, 根据输入语音类别的不同采用不同速率的语音编码, 从而达到降低平均速率的目的。IS-96C是最早提出的此类编码方案, 但是它只满足了速率上的要求, 其合成语音质量明显不及现存的其他移动通信系统(如GSM)。随后出现的EVRC编码方案改善了IS-96C的语音质量, 但是它的实际编码速率大都集中在8.55 kbps, 其他速率所占比重非常少, 并且根本没有用到2.0 kbps编码, 因此速率变化不灵活, 平均编码速率较高, 同时也不能控制调节其平均编码速率。为了解决这些问题, 满足CDMA系统对语音编码高质量和变速率的要求, 3GPP2制定了可选择模式声码器(SMV)算法标准^[1]。

(C)1994-2021 China Academic Journal Electronic Publishing House. All rights reserved. <http://www.cnki.net>

^{*} 收稿日期: 2004-09-21; 修回日期: 2004-10-19

作者简介: 孙浩海(1978-), 男, 硕士生。

SMV 算法针对不同类别的输入语音采用灵活变速率的编码,编码速率与IS-96C 和EVRC 兼容,而平均编码速率大大降低,同时其合成语音质量达到甚至超过了EVRC。它的另一个明显优势在于能够通过模式选择来控制调节平均编码速率,适应当前的信道要求。当无线信道拥挤的时候,声码器可以工作于节省模式,降低平均编码速率,提高信道容量;当信道条件宽松时,又可以工作于标准模式,提高编码速率,为用户提供更高质量的话音。本文全面介绍了SMV 算法的原理。

1 SMV 算法简介

SMV 语音编码算法可以对语音进行全速率8.55 kbps,半速率4.0 kbps,1/4 速率2.0 kbps 和1/8 速率0.8 kbps 四种速率的编码,并能根据外界提供的模式(mode) 选择信息,调节4 种速率在整个编码过程中所占比例,从而改变平均编码速率(ADR),按照ADR 从高至低可以分为模式0(与EVRC 兼容)、模式1(标准模式)、模式2(节省模式)、模式3(扩容模式)、以及模式4、5(最高编码速率为4.0 kbps)。SMV 算法对输入语音的分析与合成是分帧进行的,语音帧长为20 ms,即在以8000 Hz 的频率进行采样的系统中,每帧长度为160 个抽样点,并分为4 个子帧。采用激励-滤波模型对语音信号编码,将LPC 参数及激励参数转换成比特流进行传输,当对语音信号进行8.55 kbps 速率的编码时,每一帧语音信号用171 个比特的码字来合成,当对语音信号进行4.0 kbps,2.0 kbps 以及0.8 kbps 速率的编码时,每一帧语音信号则分别用80 个、40 个和16 个比特的码字来合成。不同速率的语音编码受速率判决算法(RAD) 控制,它根据每帧语音的特征和模式(mode) 的选择来决定当前帧的编码速率。

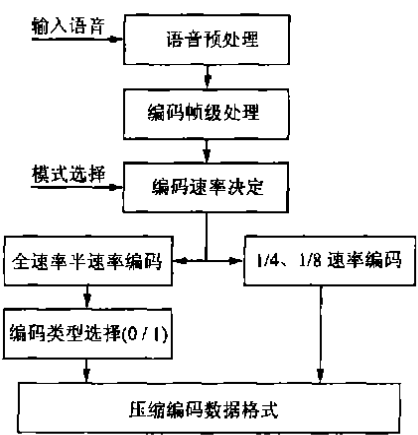


图1 SMV 算法框图

SMV 算法的框图如图1 所示。对输入语音的预处理包括高通滤波、噪声抑制和自适应斜滤波几个环节。编码帧级处理包括对语音信号的LPC 分析、开环基音检测、信号修正以及语音分类,RAD 算法根据以上得到的特征参数和模式选择信息,决定该帧的编码速率。2.0 kbps 和0.8 kbps 速率编码专门用来表示平稳清音或背景噪音,采用随机噪声模型,其LPC 滤波器的激励源由随机数产生器提供。8.55 kbps 和4.0 kbps 编码专门针对浊音信号,基于扩展码本激励线性预测(EX-CELP) 算法,并根据浊音是否平稳,分别对语音信号进行类型(type) 为1 或0 的编码。最后将LPC 参数及激励参数等转换成比特流格式^[2]。

2 EX-CELP 编解码算法

2.1 编码

EX-CELP 在具体实现的时候,吸取了传统CELP 的特点,分帧进行LPC 分析,用LPC 参数构造合成滤波器,并用自适应码本和固定码本,分别逼近语音的长时基音信号和经过短时、长时预测后的残差信号,从中搜出最佳码本矢量,乘以各自的增益后相加即得到激励信号源。在此基础上,它又采用了开环与闭环相结合的分析算法对不同类型的语音信号采取最恰当的处理方法和算法结构,以达到最佳的编码效果,提高算法效率^[3]。

为了能对不同的语音信号进行的变速率的语音编码,以达到降低平均编码速率的目的,采用多级判决的方法对每帧语音信号进行分类,一共分成无声或背景噪音、平稳清音、非平稳清音、话音起始、非平稳浊音以及平稳浊音5 个类别。首先对输入的原始语音信号分帧进行的LPC 分析,得到的LPC 参数即可用于产生加权语音,生成LSP(线谱对) 参数以及语音的准确分类,接下来利用话音激活检测(VAD) 算法将输入信号分为

无声和话音两类,然后通过有限状态机模型把话音分成清音、话音起始、浊音三类,并把清音区分成平稳和非平稳清音,最后通过对加权语音进行信号修正的方案,得到高质量的修正语音和准确的基音预测增益,其中具有稳定基音预测增益的语音帧被判为平稳浊音,其余的则为非平稳,从而完成了语音分类工作。

根据算法要求,对无声或背景噪声帧和平稳清音帧采用 2.0 kbps 或 0.8 kbps 速率的编码。其余类型则根据所选择的模式进行 8.55 kbps 或 4.0 kbps 速率编码,过程如图 2 所示,两种速率的编码方案大体相同,均以子帧为单位进行,这时应根据语音帧是否为平稳浊音,将语音分为 type 1 和 type 0 语音帧。非平稳的 type 0 语音帧采用传统的 CELP 编码方式,每一帧分成子帧处理,8.55 kbps 和 4.0 kbps 速率编码分别分为 4 个子帧和 2 个子帧,采用闭环分析的方法依次对自适应码本和固定码本进行搜索,并对自适应码本增益(基音预测增益)和固定码本增益进行 2 维联合矢量量化。type 1 语音帧代表平稳的浊音帧,对其进行处理时采用了高效的闭环与开环相结合的分析算法(COLA),将 8.55 kbps 和 4.0 kbps 速率编码分别划分为 4 个和 3 个子帧,由于平稳浊音具有稳定的基音轮廓,在子帧处理开始之前,就可以用开环分析的方法计算得到所有子帧基音增益,作为自适应码本的增益,并对它们进行 3 维或 4 维的联合矢量量化。在此基础上固定码本的搜索可以通过闭环分析得到最佳码矢,然后对所有子帧的固定码本增益进行 3 维或 4 维的联合矢量量化。

算法中对固定码本的搜索较传统的 CELP 也有了较大的改进,也采用开环与闭环相结合的分析方法。为提高码本搜索的性能,算法采用了对自适应码本贡献时域去加重的方法来产生一个具有良好感知特性的目标信号:

$$T_{FCB}(n) = T_{total}(n) - \lambda(R_p) * G_p * F_{AC}(n) \tag{1}$$

其中 $T_{FCB}(n)$ 是用于固定码本搜索的目标信号, $T_{total}(n)$ 是整个目标信号, G_p 是基音增益, $F_{AC}(n)$ 是经过滤波的自适应码本信号。 $\lambda(R_p)$ 就是去加重因子,其中的 R_p 是 $T_{total}(n)$ 和 $F_{AC}(n)$ 的归一化的自相关函数:

$$R_p = \langle T_{FCB}(n), F_{AC}(n) \rangle \tag{2}$$

当相关程度高,信号波形极为匹配时,去加重因子的值也会接近 1, 否则去加重因子的值会随相关程度的减小而变小。通过这一做法,去除两个码本中的冗余,更好的发挥两个码本的作用。

固定码本被分成几个子码本,每个子码本都代表了不同类型的语音信号,有脉冲子码本,还有随机高斯子码本,算法在进行码本搜索时,先根据前面得到的语音分类、背景噪声等信息选中最为匹配的子码本,然后用闭环分析的方法,利用高效的分析合成搜索算法在子码本中找出最佳码矢量。搜索基于最小均方误差准则,即寻找最匹配码矢量使得下式取最小值。

$$\frac{(B * C^T)^2}{C * \Phi * C^T} \tag{3}$$

其中 B 是通过加权合成滤波的目标信号, Φ 是相关矩阵, C 为码矢量。

解码过程部分借鉴了传统 CELP 的基本思想,首先从接收比特流中提取解码所需的参数序号,解码这些

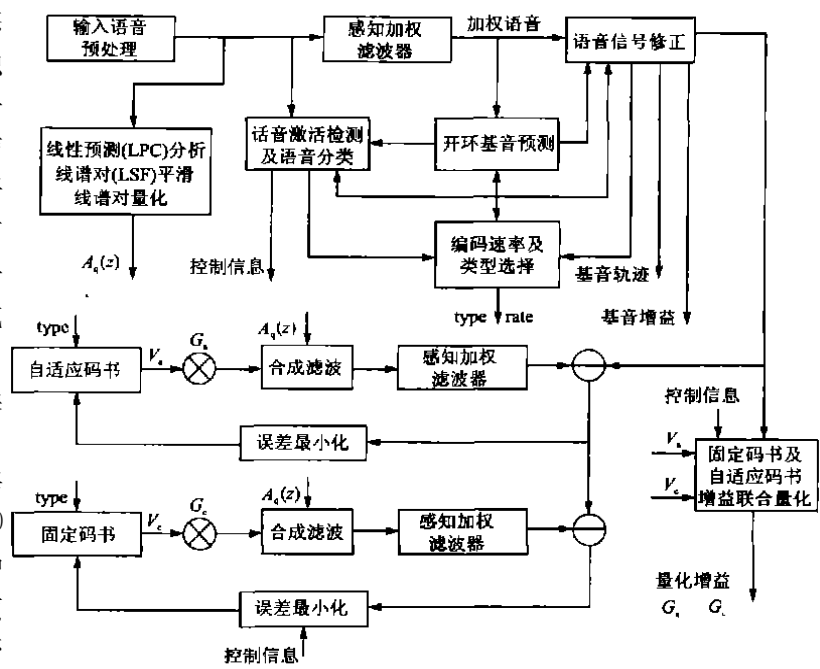


图2 EX-CELP 算法8.55 kbps 和 4.0 kbps 编码流程图

序号可以得到20 ms 语音帧所对应的编码参数。这些参数包括LSP 参数、8.55 kbps 和4.0 kbps 速率编码类型type、基音周期、自适应和固定码本矢量、自适应和固定码本增益。然后将每帧的LSP 参数进行内插转换为LP 滤波器系数,得到LP 合成滤波器,即声道模型,2.0 kbps 和0.8 kbps 速率编码的激励由随机数发生器提供,8.55 kbps 和4.0 kbps 速率解码流程如图3 所示,由自适应和固定码本分别乘以各自的增益并相加构成激励,激励信号经过LP 合成滤波即可得到重构语音,再通过长时和短时滤波等后置处理进行降噪处理,得到高质量的合成语音。

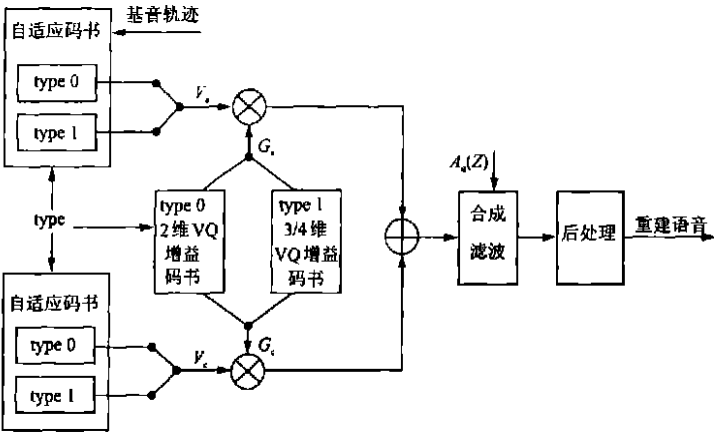


图3 EX-CELP 算法 8.55 kbps 和 4.0 kbps 解码流程图

与前面的编码过程相对应,对8.55 kbps 和4.0 kbps 速率的解码也要分为type 0 和type 1 来分别处理。对于type 0 语音帧,每一子帧的固定码本矢量根据参数序号从脉冲子码本和高斯子码本中挑选出来,而子帧的自适应码本矢量则取根据基音周期值确定,整个语音帧的增益来自type 0 语音帧的2 维矢量量化增益码本;对于type 1 语音帧,由于它具有稳定的基音轮廓和预测增益,因此处理方式与type 0 有所不同,每一子帧的自适应码本矢量由内插的基音轮廓确定,两个码本的增益分别从各自的矢量量化增益码本中得到。

合成滤波器的传递函数为:

$$H(z) = \frac{1}{A(z)} = \frac{1}{1 + \sum_{i=1}^{10} a_i * z^{-i}}$$

(4)

其中 a_i 为解码得到的10 阶LPC 系数。

表1 SMV 算法参数编码的比特分配

编码速率	8.55 kbps		4.0 kbps		2.0 kbps	0.8 kbps
类型(type)	0	1	0	1		
LSP 参数	27	25	21	21	27	11
能量					12	5
类型标志	1	1	1	1		
基音周期	26	8	14	7		
固定码本序号	88	120	30	39		
码本增益	28	16	14	12		
标志位	1	1			1	
总计	171	171	80	80	40	16

3 SMV 比特分配方案

SMV 算法参数编码的比特分配如表1 所示。对语音信号进行8.55 kbps、4.0 kbps、2.0 kbps 以及0.8 kbps 速率的编码时,每一帧语音信号则分别用171 个、80 个、40 个和16 个比特的码字来合成。2.0 kbps 和0.8 kbps 速率编码无需基音周期、码本矢量和码本增益等信息,只需要为能量和LSP 参数分配一定的比特即可,从而降低了编码速率。8.55 kbps 和4.0 kbps 速率编码中用一个类型标志比特来区分type 0 和type 1 语音帧,由于type 1 语音帧是稳定的浊音帧,具有非常稳定的基音轮廓和预测增益,用较少的比特即可代表基音周期和增益,因此能够分配更多的比特给固定码本,增大了固定码本,从而提高合成语音的质量。

4 仿真结果

SMV 算法的特色和优点在于其能够在灵活改变编码速率的同时保证较好的语音质量,我们用一段4 s

的语音文件对算法性能进行仿真测试,语音文件包括男声和女声各 2 s。首先利用语音分类算法对每一帧语音进行分类,如图 4 所示,分成 0 无声或背景噪音、1 平稳清音、2 非平稳清音、3 话音起始、4 平稳浊音、5 非平稳浊音六个类别,然后用速率选择算法根据语音类别和外界模式信息决定每帧语音的编码速率。图 4 和图 5 给出了 SMV 在各种模式下的编码速率比较,算法可以通过对不同模式的选择来灵活调节平均编码速率。其中模式 0 与 EVRC 算法兼容,它的平均编码速率明显高于其他模式,模式 1 为标准模式,模式 2,3 为节省模式,另外还有模式 4,5,其最高编码速率为 4.0 kbps。

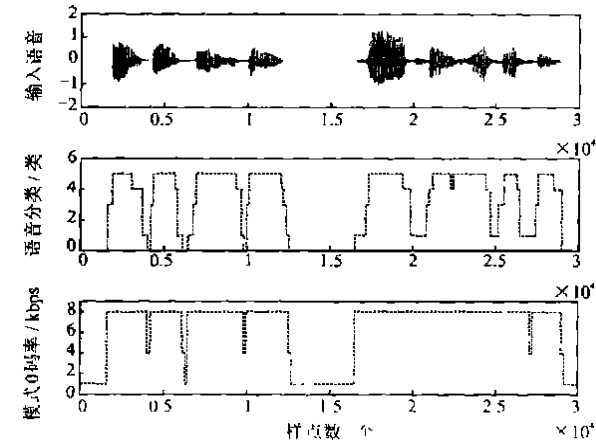


图 4 测试语音分类及模式 0 下的编码速率选择

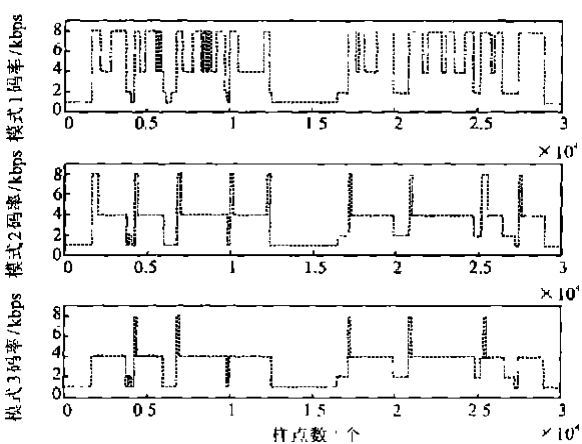


图 5 模式 1,2,3 下的编码速率选择

表 2 给出了编码算法对测试语音在各种模式编码时的码率分布表,及利用定点 C 代码仿真出的各模式下语音编码算法的算法平均复杂度和处理所有语音帧时出现的算法最大复杂度,这里用百万次加权操作/秒 (WMOPS) 作为复杂度的衡量单位。

表 2 各模式下码率分布及算法复杂度

编码模式 (mode)	码率分布表(%)				平均编码 速率(bps)	平均复杂度 (WMOPS)	最大复杂度 (WMOPS)
	全速率	半速率	1/4 速率	1/8 速率			
0	67.56	3.17	0.00	29.27	6989.2	19.63	27.48
1	44.15	13.90	10.24	31.71	5531.7	18.33	27.34
2	17.32	41.22	6.59	34.87	4217.5	18.31	27.45
3	14.63	44.63	6.10	34.64	4109.2	18.43	27.48
4	0.00	70.98	0.00	29.02	3755.1	19.81	27.78
5	0.00	59.51	9.27	31.22	3453.6	18.64	27.72

对于声码器音质的评价,我们采用了语音质量感知评价 PESQ(Perceptual Evaluation of Speech Quality) 算法对 SMV 算法各个模式以及 EVRC 算法的音质进行客观评估。PESQ 算法是目前国

际上公认的最权威、最有效的声码器音质评估算法。进行评估时,利用 PESQ 算法对原始语音和经过编解码后的合成语音进行电平调整、滤波、时间对准和听觉变换之后,分别提取两路信号的参数,综合其时频特性,得到 PESQ 分数,然后将这个分数映射到主观平均意见分(MOS 分),分值的范围在-0.5~4.5 之间,分数越高,音质越好。经过对多段语音进行多次测试,得到平均 PESQ 分数,如表 3 所示,可以看出算法改变编码速率的同时仍能提供比较好的语音质量。与 EVRC 相比较,SMV 算法降低了平均编码速率,但语音质量仍优于 EVRC。

表 3 EVRC 与 SMV 的编码速率及 PESQ 分数比较

	EVRC	模式 0	模式 1	模式 2	模式 3	模式 4	模式 5
平均速率 kbps	7.1	6.9	5.6	4.2	4.1	3.7	3.5
PESQ 分数	3.11	3.46	3.29	3.20	3.12	3.07	3.01

5 结束语

本文全面介绍了CDMA 2000 移动通信系统可选择模式声码器算法(SMV)，它吸取了传统CELP 以及以往的变速率编码算法的特点，借鉴了其中码本激励、线形预测、感知加权滤波、矢量量化、以及根据语音分类进行变速率编码的思想。同时它采用了特有的开环与闭环相结合的分析方法，提高了码本搜索效率，改善了合成语音质量，并通过灵活的速率选择算法根据语音分类以及模式信息，选择最恰当的算法结构，调节平均编码速率。仿真结果证明，SMV 算法调整平均编码速率的灵活性以及合成语音的质量，优于以往的变速率算法，能够使通话质量与信道利用率之间达到较为理想的均衡，顺应了未来移动通信系统的发展要求。

参考文献：

[1] Gao Y, Shlomot E, Benyassine A, et al. The SMV algorithm selected by TIA and 3GPP2 for CDMA applications[A]. IEEE International Conference on Acoustics, Speech and Signal Processing, Vol. 2[C]. Salt Lake City: ICASSP, 2001. 709-712.

[2] 3GPP2 C.S0030-0 V3.0-2004 Selectable mode vocoder service option for wideband spread spectrum communication systems[S].

[3] 张雄伟, 陈亮, 杨吉斌. 现代语音处理技术及应用[M]. 北京: 机械工业出版社, 2003.

(上接第25页)

3 结束语

本文分析了电荷泵锁相频率综合器的系统组成及工作原理，采用LMX2370 芯片进行双路宽频带射频锁相频率合成器的设计、制作，实验结果表明该频率合成器相噪低、频谱纯度高，做到了小型化、模块化、低功耗，并自成体系可方便的与应用系统连接。该频率源可应用于卫星通信、GSM 和无线局域网等通信系统。

参考文献：

[1] Behzad R. RF microelectronics[M]. 北京: 清华大学出版社, 2003. 247-289.

[2] 高凤辉. 卫星通信软件化终端中频模块的研究和实现[D]. 南京: 解放军理工大学通信工程学院, 2004.

[3] Roland E B. Phase-locked loops design simulation and applications[M]. 北京: 清华大学出版社, 2003.

[4] Fredrikj. PLL performance, simulation and design [EB/OL]. Second edition. <http://www.national.com/appinfo/wireless/files/DeansBook-4-01.pdf>. 2003-12-20.