新的语音信号统一 VBR 编码方法

杨 震,郑宝玉
(南京邮电学院 165 信箱,江苏南京 210003)

摘 要: 本文提出一种两级语音信号编译码新方法 - EMSVBR 系统,输入信号经语音活动性检测后,经两级编码器进行压缩. 其核心编码器基于混合编码技术,增强编码器基于小波分带的 SBC 技术,系统的码流是分层嵌入式的,系统码率变化既利用了语音的突发性,又可根据网络容量或信道特性变化而变化,涵盖了目前几乎所有语音编码标准的码率,并且新系统的解码语音质量,高于同样码率下的单一编码标准的质量. 这种语音 VBR 编译码方法,尤其适合于未来 IP 和 ATM 网络中的语音通信.

关键词: 语音信号处理; 变比特率编码; 小波; 比特分配
中图分类号: TN912. 3 **文献标识码**: A **文章编号**: 0372-2112 (2002) 01-0049-05

A New Consolidated Speech VBR Coding Method

YANG Zhen, ZHENG Bao-yu
(Nanjing University of Posts & Telecommunications, P. O. Box 165, Nangjing, Jiangsu 210003, China)

Abstract: We present in this paper a new two-level speech coding method-EMSVBR system, which consists of two subsystems. The core subsystem is based on hybrid coding technique and the subordinate subsystem is based on SBC technique that utilizes wavelet packet to construct SBC filter banks. Bit streams from this EMSVBR codec are layered and its bit rate is variable in wide ranges with the available capacity of networks or with the change of channel characteristic. This new codec demonstrates better performances than other algorithms if they operate at the same rate and therefore is suitable to be used in the cases of VOIP and VOA.

Key words: speech signal processing; VBR coding; wavelet; bit allocation

1 引言

语音信号压缩的主要目的有三个: ①节省通信传输时所需的信道容量, ②节省信号存储时所需的存储空间, ③减少冗余度,有利于保密和军事通信. 众所周知,传统的实时通信方式都是基于电路交换的,并且电路的容量也是固定的. 基于这一事实,绝大多数开发出来的语音压缩国际、地区和国家标准,都是固定速率编码系统,输出 CBR(Constant Bit Rate)码流. 当然,由于传输的特殊需要,也研究和颁布了 VBR(Variable Bit Rate)编码标准,比如, ITU-T 的 G. 727 嵌入 ADPCM 算法; CTIA 的 IS-96 QCELP 算法. 但对比众多的语音压缩标准和算法, VBR 编码国际标准可谓凤毛鳞角,且用途有限. 不过,进入九十年代以来, VBR 编码逐渐引起了人们的广泛注意,出现了许多学术论文[1~7],其主要原因是:

(1) 语音信号固定速率的压缩技术,已使编码系统码率降到了 4kb/ s 以下,进一步压缩码率并保持解码语音质量,越来越困难. 而未被充分利用的语音信号的突发特性和人类交谈的方式特性,将是导致语音信号压缩码率再次大大下降的主要方法.

(2) 通信领域的传输体系正发生重大的变革. 由于计算机技术和网络技术的大发展,因特网和基于 IP 技术的通信网络的发展如火如荼; 高速通信网络如 ATM 等成为了国际标准,它们的研究和发展大大推动了多媒体通信、信息高速公路的飞速发展,而新一代的计算机网络和通信网络中,信息是分组传输的,并且网络中资源的复用是统计方式,如果信源输出是与信源信息量变化有关的 VBR 码流,则更能提高网络的有效利用率.

(3) 不同的通信系统,可分配用于话音通信的网络容量和对通信质量的要求,是不相同的. 因此,人们研究和开发了许多压缩算法和系统,它们有不同的码率和不同的解码声音质量,尽管保证了各类通信系统的正常运行,然而也给标准的制定、不同系统的互相联接、互相通信,带来了很大困难. 如果能研究和制定出一个统一的、具有 VBR 输出的编译码系统标准,使之在不同的场合工作在不同的速率上,那么,由于不同系统采用的标准相同、信息压缩原理相同,它们的互相联接、互相通信,将大大简化.

事实上, D. M. Alley 等作者在其关于 VOA 问题的文章[8]中已经提出,将来的通信网络中,语音信号编码系统必须具有下列四种技术: VBR 能力、分层编码能力、声音活动检测 SAD

收稿日期: 1999-05-10; 修回日期: 2001-08-10
基金项目: 邮电科研基金预研项目(科字 1998 第 19 号); 江苏省科技发展基金(00KJD510012)

能力和回波抵消能力. 通信终端必须具有以上技术的原因是:“变比特率利用了话音的突发性,大大提高了网络的效率,降低了通信费用;分层编码适合了可变的信道容量及接收终端的不同驱动能力;SAD 和回波抵消则更经济地使用传输容量,节省了硬件代价;VBR 还可用于拥塞控制,当网络拥塞时,可以通过信令使编码器降低速率 ……”. 在以上四种技术中,前面的几类实际上都要求信源产生 VBR 码流,因此作者将前三种均归入变速率编码 VBR 系统中,区别仅仅在于它们产生 VBR 码流的目的和方法不一样.

2 VBR 语音编码的分类及新一代编码器应具有的功能

2.1 VBR 语音编码的分类

VBR 语音压缩,以往一般认为就是利用语音中的静音(Silence),其它都是采用 CBR 编码中的具体技术,尚未对它进行详细研究,不过,具体的、针对语音信号特性的 VBR 编码压缩方法,90 年代以来出现了许多,如文献[1、3～7]等.

作者认为,VBR 语音压缩编码,可以根据其用途和技术分成下列四大类:

(1) 与信源特性相关的 VBR 编码 ——SCDVBR(Source Characteristic Dependent VBR)

语音是时变的信号,在语音通信中,不但说话人有静默期和说话停顿期,即使是说话期内的信号,所含信息量也不同,而且人耳对不同的信号还有不同的听觉响应. 所以,合理的语音编码系统应是与信源特性相关的 VBR 编码. 仅仅利用 SAD 技术的 VBR 语音编码,很显然是 SCDVBR 的一个特例.

(2) 与网络容量相关的 VBR 编码 ——NCDVBR(Network Capacity Dependent VBR)

在这一类系统中,纯粹从传输的角度出发,希望信源输出进入网络的信息速率,有一个层次结构,可以随网络容量的时变而选择不同层的速率. 这是基于分组传输原理、统计复用信道资源的一类网络具有的要求.

(3) 与信道特性相关的 VBR 编码 ——CCDVBR(Channel Characteristic Dependent VBR)

编码系统根据对信道状态信息(比如信噪比)的估计,来调整输出速率. 这一类系统多出现在移动和卫星通信信道上,信道状态变化较快,所需的纠错码也应相应变化,信源输出也就必须随之变化. 欧洲新的 GSM 系统中的 AMR (adaptive multi rate) 标准[7]就基于这一设想.

(4) 混合控制型 VBR 编码 ——HCVBR (Hybrid Control VBR)

同时具有上述两种以上速率调节机制的 VBR 编码系统,定义为混合控制型 VBR 系统. 显然,理想化的系统是同时考虑了信源信息量、网络容量和信道特性的 HCVBR 系统.

2.2 新一代通信网络对 VBR 语音编码的要求

ITU-T 于 96 年颁布了 ATM 和 IP 网络中,进行多媒体通信时的 H.310、H.321 (ATM 网) 标准和 H.323 (IP 网络和局域网络) 标准. 其中,对声音媒体编码的要求可以用图 1 来示意.

据此可知,ATM 网络和 IP 网络对语音信号传输速率的要求,变动范围约在 4～64 kb/s 之间,不过,如图中所示,根据现有技术,做到这一点需要采用 5～6 种编码算法. 显然,要求一个终端具有图 1 中所有的语音压缩技术,是不可能的. 因此,如果要求 IP 和 ATM 网络中的语音终端,能够与网络中各种速率的终端通信,则需要设计一种新的语音编码器.

3 两级主次结构型嵌入式 VBR 系统-EMSVBR 编码新方案

如果要使一个 VBR 编码器,满足网络中对语音信号传输的各种可能速率要求,其速率不但要考虑到语音的突发性,并且要考虑到网络容量的时变,从而在语音活动期,输出码率可调范围也要从低速率一直到高速率,则在低速率工作时,必须是基于混合编码技术的,才有可能保证良好的语音编译码质量. 如果当码率提高,作者设想将低速率编译码系统产生的误差信号,也采用某种方法编码,它的解码信号逐步补偿进入原来低速率时的解码信号中,用以提高重建语音质量. 这一设想需要解决误差信号如何编译码、又如何逐步补偿,以及与原解码信号叠加的问题.

首先分析编码系统的误差信号. 图 2(a)～(d)是一段含三男一女声音的时域波形及通过 G.723.1 (6.3kb/s 码率)、CELP(5.69kb/s 码率)、ADPCM(32kb/s 码率)三种编译码器后,所产生的误差信号波形图.

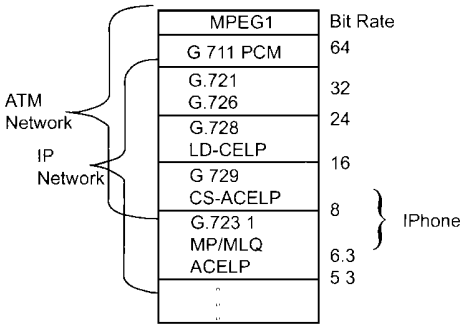

图 1 TM/IP 网络中传输声音信号采用的编码标准及码率

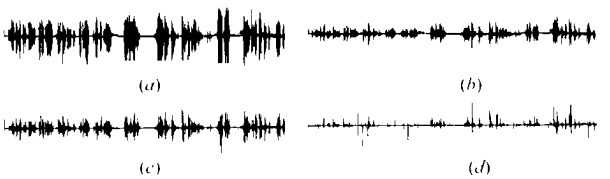

图 2 (a) 原始语音波形;(b) G.723.1 系统误差信号波形;(c) CELP 系统误差信号波形;(d) G.721 ADPCM 系统误差信号波形

它们的编码主观质量,采用基于人耳感觉特性的 Bark 域失真的 MOS 分估计法,计算得 MOS 分分别为:3.59、3.48、4.08. 使人感到意外的是,对 G.723.1 和 CELP 系统而言,尽管已有不错的编解码语音主观质量,但从图 2(b)(c)图形中,可以清晰看到,系统误差中仍含有明显的原始语音信息,并且,误差语音竟然还能大致听懂,这说明误差信号中仍有相当大的相关性. 不过,图 2(d)告诉我们,高码率的 ADPCM 系统产

生的误差信号,与原始语音的关系,并不密切,经频谱分析也可以证明,它的误差信号谱,已比较平坦.所有这些分析说明了:语音低速率混合编码系统,其解码语音与原始语音间的误差仍能被进一步利用.据此,图 3 提出一种新的两级主次结构型嵌入式 VBR 系统-EMSVBR(Embedded Main and Subordinate VBR)编码方案.

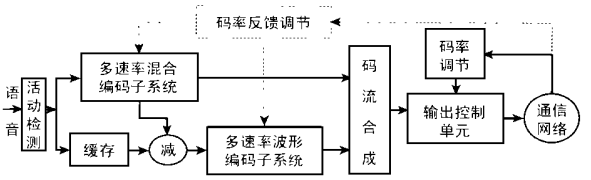

图 3　新的码率大范围可调、高质量语音编码器-EMSVBR

系统首先进行语音活动性检测,利用 E_n(短时能量)、Z_n(过零率)、R_{a_T}(自相关)、pe_n(LPC 预测误差能量)参数,进行计算(这些参数除过零率外,均被后续的混合编码子系统所利用,故几乎未增加计算量).活动和非活动两类的分类边界函数,由最小错误概率 Bayes 分类器决定,Atal 和 Rabiner 证明,对高斯信源,等效于判断下面分类函数 $f_r(\cdot)$ 的最小值:

$$f_r(X) = \frac{1}{2}(X - M_r)^T \Phi_r^{-1}(X - M_r) - \frac{1}{2}\ln|\Phi_r| - \ln(P_r) \quad (1)$$

Φ_r 是第 r 类的协方差矩阵,$M_r = \{m_{ir}\}$,$i = 1 \cdots d$,是第 r 类 d 个特征的均值,$X = \{x_i\}$ 则是某帧输入信号的 d 个特征参数,是第 r 类的概率.

如判断为活动期(talkspurt),输入语音首先经过多速率混合编码子系统编码,产生低码率的码流,该多速率混合编码器,可以采用多速率分层 CELP 系统,也可直接采用 G.723.1 系统.由于现代的混合编码器,都内含解码器,所以在该编码器完成工作的同时,它的解码语音也产生了.一段输入语音信号,一方面输入系统中上支路低码率编码子系统编码,同时这一段信号也被系统中下支路缓存器存储,因此在低码率系统编码结束,误差信号也就产生了.图 3 中下支路的编码子系统,完成对误差信号的编码,以实现整个编码系统码率大范围的可调,并相应地使解码质量逐步提高,采用何种波形编码技术因而成为重要问题.

由于语音的听觉质量,主要取决于它的频率特性,尤其是幅度频率特性,因此如果要使编译码系统的解码语音,在码速率调整后,仍保持良好的质量,则合理的变速率系统的设计,应与输入语音的频率特性有关,所以图 3 中误差信号波形编码器,采用基于小波包分带的子带编码法.首先将输入信号用小波包滤波器分成不同的子带信号,每个子带信号再用波形编码器编码,其原理图如图 4 所示.现在面临下面三个问题:(1)子带划分数目如何确定?子带波形用什么编码法?(2)波形编码系统的码流,如何与混合编码系统的码流合并,构成速率大范围可调 VBR 系统?(3)波形编码系统的加入,是否使得整个系统的解码质量随码率的增加而增加?

既然混合编码系统产生的误差信号仍具有原始语音的特征,那么就仍能期望有效的波形编码方法 ADPCM,对误差信号的子带信号,也产生比较好的编码效果,所以采用嵌入式 ADPCM 来处理子带信号.至于子带的数目,取决于速率调节的需要,比如四带划分(如图 4),则子带信号抽样频率为 2kHz,于是波形编码速率的调节最小台阶为 2kb/s.

对第二个问题的回答,由图 4 中的码流排列位置控制来解决.换言之,各子带信号编码码流的排列,按照对应子带信号的重要性来顺序排放,重要性由小波包分带能量代价函数来决定,具体计算根据下面序列的对数熵:

$$M(\{s_i\}) = \sum_i \log|s_i|^2 \quad (s_i \neq 0) \quad (2)$$

传输时根据网络容量,从重要的码流位置处开始输出,容量不够时,最不重要的子带码流被抛弃,依次类推,因此它是与 G.727 系统类似的嵌入式的码字排列结构.当然,图 3 系统中混合编码系统产生的码流,始终排在比误差波形编码码流更重要位置上.那么图 3 中通过码率反馈调节,分配给误差信号波形编码系统不同码率时,各子带如何分配比特呢?显然这对解码语音质量也至关重要,本系统采用如下策略:

(1)误差语音能量小于门限 T(实验中取 0.001),则基带占有所有比特

(2)按各带能量大小分配比特

(3)第二步分配结果若因四舍五入,导致各带分得比特数目之和小于总比特数,则差额分给能量最大的子带

(4)若以上分配结果导致某一带比特数超过 5 比特/每样值,则比特数目取 5,多余的比特分给其相邻的子带

(5)总比特数超过 12(对应误差信号编码码率为 24kb/s)时,考虑各子带均分比特

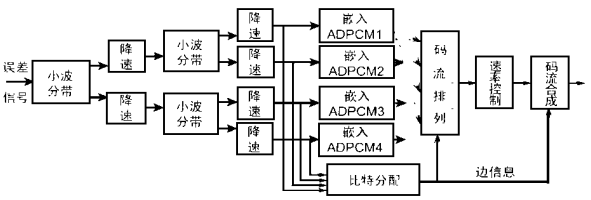

图 4　基于小波包滤波器分带的嵌入式 VBR 波形编码子系统

显然,整个图 3 编码系统的码率变化范围,最低速率由多速率混合编码系统确定,最高速率等于多速率混合编码系统的最高速率与多速率波形编码系统的最高速率之和,因此仅从速率上讲,已经可以包含目前几乎所有语音编码标准中的码率范围.

下面讨论解码语音质量问题.首先要考虑的是,误差波形编码系统的解码信号,如何与混合编码系统的解码信号叠加,这取决于小波包分带/合成系统的特性.

采用小波包滤波器进行分带/合成的原因在于,与 Fourier 变换不同的是,小波变化函数中引入了尺度和平移因子,可以分析信号的时频局部特征,并且借助于尺度因子的变化,容易构成信号的多分辨率分析(MRA),这种 MRA 方法更适合于象语音这样的时变信号[9].另外,由于可以针对不同信号设计不同的小波基,通常只要较短的滤波器就能获得良好的分带效果.

语音信号的频谱分布是随时间不断变化的,我们希望在

整个频带内都能产生各种分辨率的空间划分,所以,本文用正交小波包[9]来处理语音,由于小波包分解时对高频部分也进行了逐级分解,可得到比小波分析更好的频率局部化信息,便于在 ATM 和 IP 网络中,根据需要和频带资源,选择合理的信号分辨率进行编码.

图 3 系统中误差信号经过上面讨论的、由图 4 所示的分带/编码/解码/合成(图 4 只示出前二步,相反的后二步是它的逆系统)运算后,与原来的误差信号位置上发生了什么变化,变化了多少是将解码误差信号叠加到混合解码系统输出上去的关键.

命题:采用小波包滤波器进行信号的分解/合成时,输入和输出信号位置的变化最小等于下式:

$$P=\sum_{k=1}^{m}2^{k-1}\times(L-1) \tag{3}$$

其中,$m=\log2(N)$,N 是分带数,L 是紧支撑正交小波长度,等于滤波器长度. 下面以二带的划分来证明.

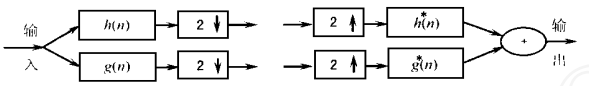

图 5 二带小波分带/合成系统

图 5 即为二带小波分带/合成系统原理图,按照 Mallat 金字塔算法,可以继续构成四带、八带等的分带/合成系统. 其中 $h(n)$ 是平滑滤波器(低通),$g(n)$ 是差分滤波器(高通),且根据小波包多分辨率理论:

$$g(n)=(-1)^n h(-n+1) \tag{4}$$

再根据信号的完全重构理论定理,$h(n)$ 和 $g(n)$ 的对偶系统 $h^*(n)$,$g^*(n)$ 须满足:

$$H^*(Z)=H(Z^{-1})\qquad G^*(Z)=G(Z^{-1}) \tag{5}$$

$H^*(Z)$,$G^*(Z)$ 是 $h^*(n)$,$g^*(n)$ 的 Z 变换. 本文中小波平滑滤波器 $h(n)$,是根据有限支集规范正交小波基构造法,采用二进 Daubechis 小波基,计算得到的 FIR 型滤波器,取 20 阶. 然而,单个滤玻器的脉冲响应 $h(n)$ 一般不对称,故不是一个线性相位滤波器,因此给信号滤波后输入输出位置关系的确定带来困难. 不过,如果从图 5 中的分带/合成整体过程看,如令图 5 中系统输入为 $s(n)$,其 Z 变换为 $S_{\text{input}}(Z)$,则容易得到,系统输出 $S_{\text{output}}(Z)$ 为:

$$\begin{aligned}S_{\text{output}}(Z)=&\frac{1}{2}S_{\text{input}}(Z)[H(Z)H^*(Z)+G(Z)G^*(Z)]+\\&\frac{1}{2}S_{\text{input}}(-Z)[H(-Z)H^*(Z)+G(-Z)G^*(Z)]\end{aligned} \tag{6}$$

式中 $H(Z)$,$G(Z)$ 分别是图 5 中平滑滤波器 $h(n)$、差分滤波器 $g(n)$ 的 Z 变换.

1977 年,在讨论 SBC 系统中的分解/合成滤波器时,Esteban 和 Galand 提出,如令:

$$G(Z)=H(-Z)\qquad H^*(Z)=H(Z)\qquad G^*(Z)=-H(-Z) \tag{7}$$

那么式(6)中第二项,即包含由于信号重抽样引起的失真项,可以抵消. 由于此时高通滤波器 $g(n)$ 的频率响应,与低通滤波器 $h(n)$ 的频率响应,在整个频率轴上互补,且关于四分之一抽样频率对称,所以这样的高低通滤波器,称为 QMF. 不过,很难找到根据(7)形成的 QMF,使(6)式第一项中的括号内的失真,也完全消失. 1985 年和 1986 年,由 Smith,Barnwell 和 Vetterli 等人,提出了如下的另一类正交镜像滤波器[10]:

$$\begin{aligned}&G(Z)=Z^{-1}H(-Z^{-1})\qquad H^*(Z)=H(Z^{-1})\\&G^*(Z)=H^*(Z^{-1})=ZH(-Z)\end{aligned} \tag{8}$$

容易证明,此时的 QMF,同样消除了式(6)的第二项,并且可以找到满足式(8)的 FIR,使式(6)式第一项中的括号内的失真,完全消失,符合式(8)特性的系统又称为共轭正交镜像滤波器(CQMF).

容易看出,满足式(4)、(5)的图 5 系统,与满足式(8)的 CQMF 系统是相同的. 于是,我们有这样的结论:小波 MRA 分析/合成系统,正是 CQMF.

不过,CQMF 系统,并不一定能构成信号的小波 MRA,因为在信号的小波 MRA 中,对 $h(n)$ 和 $g(n)$,以及它们的对偶系统,有下列限制条件:

(1) $\sum_n h(n)|n|^{\partial}<\infty$,对有限的 ∂ (9)

(2) $\sum_n h(n-2i)h(n-2j)=\delta_{ij}$ (10)

(3) $\sum_n h(2n)=\sum_n h(2n+1)=2^{-1/2}$ (11)

(4) $HG^*=0$, $HH^*+GG^*=I$ (12)

这是 CQMF 系统不一定具备的.

在考虑了物理可实现性后(因果性),对图 5 中合成滤波器(即式(5)),取:

$$h^*(n)=h(L-1-n)\qquad g^*(n)=g(L-1-n) \tag{13}$$

将(8)以及(13)式,代入(6),有:

$$\begin{aligned}S_{\text{output}}(e^{jw})=&S_{\text{input}}(e^{jw})[H(e^{jw})H(e^{-jw})+G(e^{jw})G(e^{-jw})]\\&e^{-jw(L-1)}\end{aligned} \tag{14}$$

根据式(12),右边方括号内等于 1. 这表明,采用小波包滤波器进行信号二带分解时,经过分解再合成后,信号无失真,但位置后移了 $L-1$ 点. 当采用四带分解和合成时,显然从二带分解的结果出发,经过后一级四带分解,再经过四带合成,位移仍是 $L-1$ 点,不过,因为四带分解和合成时,系统工作频率,比之二带划分系统,降低了一半,因而此时的 $L-1$ 点位置偏移,对应于二带时的 $2(L-1)$ 点偏移,所以四带时总的位置移动,等于 $L-1+2(L-1)$. 以此类推,可证明更多带分解时,命题也成立.

作者将 G.723.1 系统,作为图 3 中混合编码系统,实验了 EMSVBR 编译码系统的性能,结果如表一,其中分类函数(1)式中的参量预先由 3000 帧试验信号确定,信号属于叙述型语音,检测得到活动期占 89.6 %.

整个系统码率,在语音活动期,等于 G.723.1 混合编码系统码率(此处是用的 MP-MLQ 方案,为 6.3kb/s),加上表中误差信号编码码率,所以试验的码率范围是 6.3kb/s～38.3kb/s,也可以继续增加码率,一直做下去,不过从性能上看,已经没有太大的必要了. 此处不妨比较一下对同样的声音,采用 G.726 ADPCM 系统编码的质量:对于 32kb/s 系统,MOS = 4.084,SNR = 26.78dB;对于 24kb/s 系统,MOS = 3.9236,SNR =

21.73dB.可见,本文新 VBR 系统,从主观质量上而言,在速率约 14kb/s 时,具有 ADPCM 系统 24kb/s 速率时的性能;在速率约 22kb/s 时,具有 ADPCM 系统 32kb/s 时性能.同样的比较,得到的结论是:本文方法,在约 10kb/s 码率时,具有 GSM13kb/s 系统的性能;在同样码率时,本文法与 G.728 LD-CELP 系统性能相仿.如果考虑到本文 EMSVBR 系统,具有的语音活动检测能力,显然在质量相仿时,本文法平均码率远低于上述系统码率.

表 1　本文新的语音 EMSVBR 系统主客观性能

波形编码系统速率	0 kb/s (0bit)	2 kb/s (1bit)	4 kb/s (2bit)	8 kb/s (4bit)	12 kb/s (6bit)	16 kb/s (8bit)	20 kb/s (10bit)	24 kb/s (12bit)	32 kb/s (16bit)
MOS 分	3.591	3.413	3.617	3.963	4.022	4.101	4.151	4.188	4.294
SNR (dB)	7.28	10.81	14.30	18.01	18.31	19.61	20.23	21.69	30.45

说明:测试语音长 14.61 秒,含三男一女声音,MOS 分根据基于人耳听觉误差的 Bark 域谱失真测度计算得到

图 6 显示了图 3 系统在不同码率下,对一段结构比较复杂的语音信号的编码性能.可以看到,随着码率的增加,解码语音信号的细节结构不断改善.

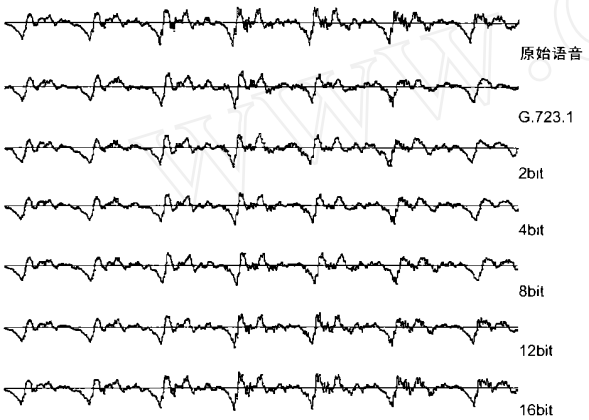

图 6　新的基于混合和波形编码结合技术的 VBR 系统性能(从上至下逐渐增加波形编码码率)

4　结论

本文提出一种两级语音信号编译码新方法-EMSVBR 系统.这种语音编码新方法,可以在良好解码质量条件下,产生涵盖目前几乎所有的语音编码系统速率,并且该系统码率,既随输入特性的变化而变化(语音活动检测),又可以根据网络容量的变化而增减,更适合基于分组传输的通信系统.所以作者认为:结合混合编码和波形编码两种技术的语音压缩方法,是构成 VBR 系统,适应网络容量变化的理想方案,且将成为未来统一语音编译码系统标准的基础.

参考文献:

[1] S McClellan J D Gibson. Variable-rate CELP based on subband flatness [J]. IEEE Transactions on Speech and Audio Processing, 1997, 5(2): 120-130.

[2] K Kondo, M Ohno. Packet speech transmission on ATM networks using a variable rate embeded ADPCM coding scheme [J]. IEEE Transactions on Communications, 1994, 42(2): 243-247.

[3] Eric W M Yu & Cheung-Fat Chan. Variable bit rate MBELP speech coding via V/UV distribution dependent spectral quantization [C], Proceedings of ICC, Montreal, Canada, 1997, 2: 1607-1610.

[4] A Shen, B Tang, A Alwan, G Pottie. A robust variable-rate speech coder [C], Proceedings of ICASSP, Detroit, Michigan, U.S.A., Vol. 1, 1995: 249-252.

[5] L Zhang, T Wang, V Cuperman. A CELP variable rate speech codec with low average rate [C], Proceedings of ICASSP, Munich, Germany, Vol. 1, 1997: 735-738.

[6] B Tang, A Shen, A Alwan, G Pottie. A perceptually based embeded subband speech coder [J]. IEEE Transactions on Speech and Audio Processing, 1997, 5(2): 131-139.

[7] H Ito, M Serizawa, K Ozawa, T Nomura. An adaptive multi-rate speech codec based on MP-CELP coding algorithm for ETSI AMR standard [C], Proceedings of ICASSP, Seattle, Washington, USA, May, Vol. 1, 1998: 137-140.

[8] D M Alley et al. Audio services for an asynchronous transfer mode network [J]. BT Technol. J., 1995, 13(3): 81-91.

[9] 徐佩霞,孙功宪. 小波分析与应用实例 [M]. 合肥:中国科学技术大学出版社, 1996: 4-18.

[10] M J Smith, T P Barnwell. Exact reconstruction technique for tree-structured subband coder [J]. IEEE Trans. on ASSP, 1986, 34(3): 434-441.

作者简介:

杨　震　男.1961 年 11 月生于江苏省苏州市.1999 年毕业于上海交通大学通信与信息系统专业,获博士学位,现为南京邮电学院教授.主要研究方向为:信号处理,语音编码与识别,IP 与 ATM.已发表论文 30 多篇.

郑宝玉　男.1945 年 12 月生于福建省闽候.南京邮电学院教授、博导;上海交通大学兼职教授、博导.主要研究方法为:智能信号处理、通信信号处理、语音信号处理.已发表论著多部,文章 70 多篇.