

一种新的 8~ 32kb/s 宽带嵌入式变速率语音编码方法

刘泽新, 鲍长春, 贾懋

(北京工业大学语音与音频信号处理研究室, 北京 100022)

摘要: 本文基于 ACELP 和 TCX 编码技术, 提出了一种 8~ 32kb/s 五层宽带嵌入式变速率语音编码方法, 其中, 前三层采用 ACELP 实现了 8kb/s、12kb/s 和 16 kb/s 的嵌入式编码, 后两层采用 TCX 技术实现了 24 kb/s 和 32 kb/s 嵌入式编码. 实验结果表明, 该嵌入式语音编码方法的质量在纯净语音、办公室噪声和层间转换方面接近于 ITU-T G. VBR 的 TOR 要求.

关键词: 语音编码; 嵌入式语音编码; ACELP 编码; TCX 编码

中图分类号: TN912.3 **文献标识码:** A **文章编号:** 0372-2112(2008)05-1013-06

A New 8~ 32 kb/s Wideband Embedded Variable Bit-Rate Speech Coding Algorithm

LIU Ze-xin, BAO Chang-chun, JIA Mao-shen

(Speech and Audio Signal Processing Lab, Beijing University of Technology, Beijing 100022, China)

Abstract: A five layers wideband embedded variable bit rate speech coding algorithm at 8~ 32kb/s was proposed in this paper based on algebraic code excited linear prediction (ACELP) and transform coded excitation (TCX) speech coding techniques. The first three layers at 8kb/s, 12kb/s and 16 kb/s are embedded by ACELP coding model, while the last two layers at 24 kb/s and 32 kb/s are embedded with TCX technique. The test result indicates that the quality of the embedded speech coding algorithm with five layers is close to Terms of Reference (TOR) for ITU-T G. VBR codec for clean speech, speech with office noise and switching between layers.

Key words: speech coding; embedded speech coding; ACELP coding; TCX coding

1 引言

在以分组交换为基础的 IP 网络上进行语音分组传输时, 可能会出现分组丢失或分组延时的情况, 由于这两种情况的存在, 使得在语音接收端可能因为码流信息丢失, 无法合成高质量的连续语音信号. 目前现有的单速率、多速率语音编码标准无法解决该问题.

为减小码流信息丢失对合成语音质量的影响, 从语音编码的角度考虑, 可以对原始语音进行嵌入式编码, 从而获得具有嵌入式结构的码流.

在嵌入式的码流中, 低码率的码元包含(嵌入)在高码率的码元中. 也就是说, 一个完整的嵌入式码流可以逐层分解成几个不同级别的码流, 这些码流的码率逐次递减, 但依然都能表示原始语音信号的主要参数, 只是在不同程度上损失了一些细节. 当网路容量足够时, 网路中传输高速率码流, 在接收端可以恢复较高的或事先设定的任意速率的语音质量; 当遇到不同程度的网路拥塞时, 根据信道编码协议, 码流中非核心码元被逐层丢

弃, 在接收端恢复低码率质量较差的语音, 但总能保证语音的连续性.

因此, 嵌入式的码流结构不仅可以有效地解决由于分组丢失所引起的合成语音质量下降的问题, 而且可以提供多种编码速率, 以适应不同种类的通信终端^[1,2].

嵌入式变速率语音编码作为目前语音信号处理领域最前沿的课题之一, 诸多专家和学者积极对此展开研究, 也已取得了一些初步成果, 如: 用 ACELP 编码技术实现 12.65~ 27.85kb/s 的三层嵌入式变速率语音编码算法^[3], 用 ACELP 和 MDCT 编码技术实现 12.2~ 16kb/s 的两层混合的嵌入式编码结构和用 VSELP 技术实现的多阶嵌入式编码结构等^[3,5]. 2006 年 ITU-T 制定了新一代语音编码标准 G.729.1, 它采用了嵌入式 CELP 编码技术、时域频带扩展技术(TDBWE)以及被称作时域混叠抵消(TDAC)的预测变换编码技术, 实现了码流的嵌入式结构^[6,7]. 目前, ITU-T 的另一项嵌入式变速率语音编码标准 G. VBR 也在紧张地制定过程中. 本文正是为竞标此项国际语音编码标准而做的研究工作.

2 算法原理

本文提出的是一种 8~ 32kb/s 的五层宽带嵌入式语音编码算法, 运用了 ACELP 和 TCX 的联合编码技术, 处理帧长为 20ms。该编码算法根据输入语音采样率的不同, 可以编解码采样率为 16kHz 和 8kHz 的宽窄带语音。图 1 给出了编码器原理框图。

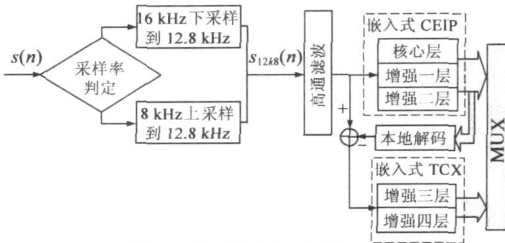


图 1 编码器原理框图

2.1 预处理

预处理包括采样率转换和高通滤波。采样率转换的目的是为了降低整个编解码算法的计算复杂度, 将输入原始语音的 16kHz 或 8kHz 采样率下采样或上采样为内部采样率 12.8kHz。采样率转换公式为:

$$x_{out}(n) = \sum_{i=0}^{k-1} h(iL + \langle nM \rangle_L) \cdot x_{in}(\lfloor nM/L \rfloor - i) \quad (1)$$

其中 L 为插值倍数, M 为抽取倍数, $K = N/L$, N 为滤波器 $h(n)$ 的长度, $\langle nM \rangle_L$ 表示对 nM/L 求模, $\lfloor nM/L \rfloor$ 表示对 nM/L 向下取整。

高通滤波是为了去除不必要的低频成分, 截止频率为 50Hz, 其传递函数为:

$$H_{hl}(z) = \frac{0.982910156 - 1.965820313z^{-1} + 0.982910156z^{-2}}{1 - 1.965820313z^{-1} + 0.966308593z^{-2}} \quad (2)$$

2.2 嵌入式 ACELP 编码

嵌入式 ACELP 编码技术主要应用在核心层、增强一层和增强二层上。

预处理后的语音信号 s_{re_p} 被送入到嵌入式 ACELP 编码器中, 该编码器能够产生 8~ 16kb/s 的嵌入式码流。其核心层由改进的 ACELP 编码算法构成, 增强一层和增强二层用额外的代数码书来实现。

核心层编码在标准 ACELP 编码基础上做了三点改进: (1) ISP (Immittance Spectral Pairs) 参数量化采用各维非等系数帧间 MA (Moving Average) 预测分裂矢量量化方法^[8], 在比特数相同的情况下, 此方法有更小的谱失真; (2) 限制自适应码书增益, 即限制自适应码书的贡献, 使代数码书有一定的准周期性, 此技术加速了丢失帧后正常帧的正确恢复; (3) 采用全搜索方法进行代数码书的搜索。

在本文提出的算法中, 由于 ACELP 编码是三层的嵌入式编码, 各层感知加权合成滤波器的历史状态不同, 因此每一层都有与之相对应的目标矢量。核心层、增

强一层、增强二层的的第一级目标矢量分别记作 $x_n(n)$ 、 $x_{n_12}(n)$ 、 $x_{n_16}(n)$ 。

由于当前帧自适应码书是在基音延迟处 (整数基音延迟和分数基音延迟) 内插过去的激励得到的, 又由于每一层有各自的激励信号, 所以, 核心层、增强一层和增强二层也有各自的自适应码书 $v_n(n)$ 、 $v_{n_12}(n)$ 、 $v_{n_16}(n)$ ^[9]。

搜索增强一层的代数码书之前, 将增强一层和增强二层的的第一级目标矢量分别减去其各自的自适应码书贡献和核心层的代数码书贡献, 得到用于搜索增强一层代数码书的第二级目标矢量 x_{n2} 和增强二层的第二级目标矢量 x_{n2_16}

$$\begin{cases} x_{n2}(n) = x_{n_12}(n) - g_p \cdot y_{1_12}(n) - g_c \cdot y_2(n) \\ x_{n2_16}(n) = x_{n_16}(n) - g_p \cdot y_{1_16}(n) - g_c \cdot y_2(n) \end{cases} \quad (3)$$

其中, $y_{1_12}(n) = v_{n_12}(n) * h(n)$, $y_2(n) = c(n) * h(n)$, $y_{1_16}(n) = v_{n_16}(n) * h(n)$, $h(n)$ 为感觉加权合成滤波器的脉冲响应, $c(n)$ 为核心层代数码书, g_p 和 g_c 分别为核心层的自适应码书和代数码书增益。

由 x_{n2} 和感觉加权合成滤波器的脉冲响应 $h(n)$, 通过公式

$$Q_k = \frac{(x_{n2} \cdot Hc_k)^2}{c_k^2 H^2 Hc_k} \quad (4)$$

搜索增强一层代数码书 $c_1(n)$ 及增益 g_{c1} 。其中, c_k 为要搜索的代数码书, H 为由 $h(n)$ 组成的下三角 Toeplitz 矩阵。

将增强二层的第二级目标矢量减去增强一层代数码书的贡献, 得到搜索增强二层代数码书的第三级目标矢量 x_{n3} :

$$x_{n3}(n) = x_{n2_16}(n) - g_{c1} \cdot y_3(n) \quad (5)$$

其中 $y_3(n) = c_1(n) * h(n)$ 。

由 x_{n3} 搜索得到增强二层代数码书 $c_2(n)$ 及增益 g_{c2} 。增强一、二层代数码书搜索采用的是 Focused-Search 快速算法^[10]。

ACELP 编码中激励的生成原理如图 2 所示。

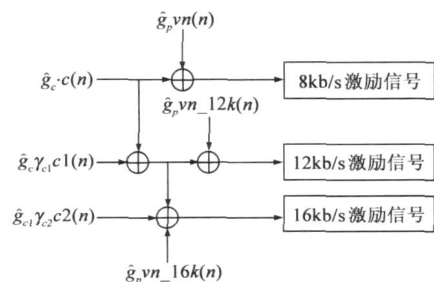


图 2 嵌入式 ACELP 编码中激励的生成框图

前三层的自适应码书和代数码书搜索完毕后, 为计算下一子帧的目标矢量, 对各层滤波器状态进行独立更

新. 嵌入式 ACELP 编码比特分配如表 1 所示.

表 1 嵌入式 ACELP 编码比特分配表

核心层					
参数	第 1 子帧	第 2 子帧	第 3 子帧	第 4 子帧	总比特
ISP 索引					46
基音延迟	8	5	8	5	26
码书索引	16	16	16	16	64
增益索引	6	6	6	6	24
总比特					160
增强一层					
第一脉冲	5	5	5	5	20
第二脉冲	5	5	5	5	20
第三脉冲	5	5	5	5	20
轨迹标志	2	2	2	2	8
增益索引	3	3	3	3	12
总比特					80
增强二层					
增强二层比特分配和增强一层相同					

在本文的嵌入式 ACELP 编码中, 分五子帧进行自适应码书和代数码书搜索, 并在本地解码出五子帧采样率为 12.8kHz 的 16kb/s 合成语音 s_{yn_16} . 做五子帧搜索的目的是为增强三层和增强四层的 TCX 编码做准备, 因为在 TCX 编码中, 要用到将来帧的样点. 编码时仅仅对前四子帧的参数进行编码, 实际写入码流的还是四个子帧的信息, 并不将第五子帧的信息写入码流.

2.3 嵌入式 TCX 编码

TCX 编码是一种变换域编码. 主要思想是把输入的时域信号变换到频域, 将所得的频谱系数用格型矢量量化的方法进行量化, 得到要传输的参数^[11]. 其编码原理如图 3 所示.

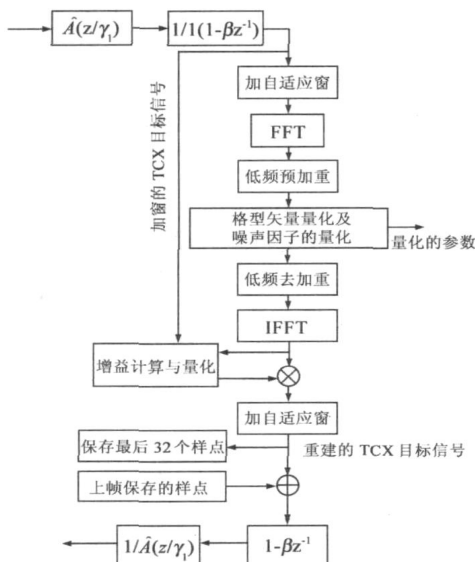


图 3 TCX 编码原理框图

2.3.1 增强三层 TCX 编码

预处理后的原始语音 s_{re_sp} 减去本地解码出的合成语音 s_{yn_16} 将得到的差值 T_{CX_24} 作为增强三层要处理的目标信号, 对此目标信号做 TCX 编码.

在做前三层的嵌入式 ACELP 编码时, 为了变采样 (上采样和下采样) 滤波器和加窗求解线性预测系数的需要, 当前帧语音比预处理后的语音延迟了 52 个样点, 如图 4 所示.

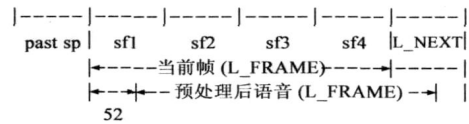


图 4 样点延迟示意图

为了保证预处理后的语音 s_{re_sp} 和本地解码出的 16kb/s 合成语音 s_{yn_16} 相减时样点相对应, 在做减法前, 将预处理后的语音也做 52 个样点的延迟, 记为 s_{new_sp} . 分别取 s_{new_sp} 和 s_{yn_sp} 内的前 288 个样点做减法, 得到增强三层要处理的 288 个样点目标信号 T_{CX_24} , 如式 (6) 所示:

$$T_{CX_24}(n) = p_{new_sp}(n) - s_{yn_16}(n) \quad (6)$$

在 TCX 编码中, 感觉加权滤波器的传递函数为:

$$w_1(z) = \frac{\hat{A}(z/\gamma)}{1 - \beta_1 z^{-1}} \quad (7)$$

式中, $\hat{A}(z/\gamma)$ 是量化的线性预测滤波器的一种加权形式, 分母 $1 - \beta_1 z^{-1}$ 是一阶常系数多项式, 系数 $\beta_1 = 0.8$. 输入信号通过感觉加权滤波器, 得到感觉加权信号 x_{n1} . 由于在解码端对解码出的信号要进行去感觉加权操作, 因此, 本传递函数的分子部分采用了线性预测滤波器的加权形式, 其系数为量化后的线性预测系数; 同时, 因为此时处理的是差值信号, 所以对分母的系数 β_1 进行了调整.

TCX 编码处理的目标信号长度为 288 个样点, $9 \times 2^5 = 288$, 所以, 对加自适应窗的感觉加权信号 x_{n1} 做以 9 为基底的 FFT 变换, 得到 288 个样点的傅氏系数^[12]. 取其前 144 个有效复数值的实部和虚部, 组成一个包含 288 个元素的实数序列, 即当前帧的谱序列 X .

$$X = \{re[0], im[0], re[1], im[1], \dots, re[N/2-1], im[N/2-1]\} \quad (8)$$

式中, $re[]$ 和 $im[]$ 分别表示实部和虚部, $N = 288$.

从谱序列 X 的第一个元素值开始, 将连续的每 8 个元素值分成一个模块, 共有 36 个模块, 记第 k 个模块为 B_k .

$$B_k = \{re[4k], im[4k] \dots re[4k+3], im[4k+3]\}, \quad k = 0, 1, 2, \dots, 35 \quad (9)$$

这样做的目的是为了可以利用 8 维格型矢量量化

方法来量化谱序列 X 。为了最小化低频感觉失真,还要对谱序列进行低频自适应加重处理。

在格型矢量量化中,输入到量化器矢量的元素越大,量化相应码字时用的比特数就越多。所以,为了保证量化码字用的比特数在事先给定的比特数范围内,量化前将输入矢量除以一个估计的全局增益因子 g 。

对比例缩放后的矢量进行 RE8 格的球型矢量量化(格型矢量量化的一种),得到要输出参数(谱序列和舒适噪声因子)的索引值和量化后的谱序列 $\hat{X}^{[13,14]}$ 。

对量化后的谱序列 \hat{X} 做低频去加重和 IFFT 变换,得到量化的感觉加权信号 \hat{x} 。

感觉加权信号量化前做了比例缩放,所以,量化后要进行还原,乘以一个量化后的全局增益因子 \hat{g} 。因为量化前估计的全局增益因子 g 不能保证量化前的加权信号 x 和量化后加权信号 \hat{x} 之间有最大的相关,所以,通过式(11)重新求得使 x 和 \hat{x} 有最大相关的最优全局增益 g' :

$$g' = \frac{\sum_{n=0}^L x(n) \hat{x}(n)}{\sum_{n=0}^L \hat{x}(n) \hat{x}(n)} \quad (10)$$

为了减少量化带来的误差,对幅度还原后的加权信号 \hat{x} 重新加了自适应窗,为了去除加自适应窗对信号造成的影响,还要对信号的过渡段采用混叠相加技术。

最后,对量化后的感觉加权信号 \hat{x} 去感觉加权,得到量化的增强三层目标信号 q_{CX_24} 。

将 q_{CX_24} 和本地解码出的采样率为 12.8kHz 的 16kb/s 合成语音相加,得到本地解码出的采样率为 12.8kHz 的 24kb/s 合成语音 S_{yn_24} 。

$$S_{yn_24}(n) = q_{CX_24}(n) + S_{yn_16}(n) \quad (11)$$

2.3.2 增强四层 TCX 编码

将量化前增强三层的目标信号 T_{CX_24} 减去其量化后的值 q_{CX_24} 得到的差值作为增强四层的目标信号 T_{CX_32} , 如式(12)所示。

$$T_{CX_32}(n) = T_{CX_24}(n) - q_{CX_24}(n) \quad (12)$$

为了满足嵌入式 TCX 编码的需要,对增强四层的目标信号 T_{CX_32} 做 TCX 编码时,对感觉加权滤波器的系数也进行了调整, $\beta_2 = 0.88$ 。此时,感觉加权滤波器的传递函数为

$$W_2(z) = \hat{A}(z/\gamma)/(1 - \beta_2 z^{-1}) \quad (13)$$

其他各步操作和增强三层的处理方式相似。因为增强四层编码后,不需要在编码端再本地解码出增强四层的目标信号,所以,对全局增益量化以后,对增强四层的编码就结束了。

嵌入式 TCX 编码比特分配如表 2 所示。

表 2 嵌入式 TCX 编码比特分配表

参数	模式	
	增强三层	增强四层
格型矢量量化索引值	150	150
舒适噪声因子	3	3
全局增益	7	7

至此,整个编码过程结束,得到五层嵌入式结构的码流,输入到网络中进行传输。

2.4 解码算法原理

图 5 为解码器结构原理框图。

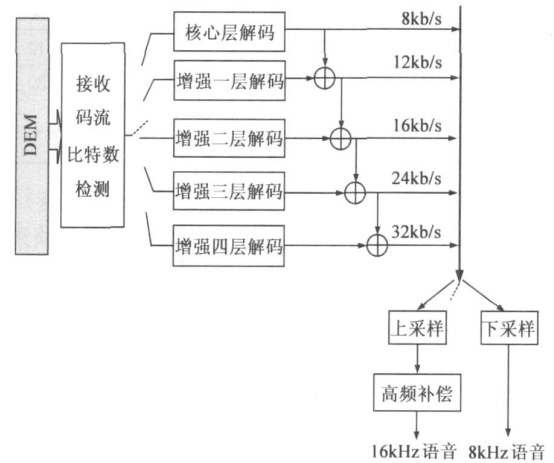


图 5 解码器结构框图

在解码前首先进行接收码流比特数检测:

(1) 当解码端每帧接收到比特数大于 160bits 时,解码出核心层的基本参数(LPC 系数, 自适应码书和代数码书及它们的增益);

(2) 判断接收到的比特数是否大于或等于 240bits, 如果是,则继续解码出增强一层代数码激励及其增益; 否则,根据第(1)步解码出核心层基本参数得到 8kb/s 的激励信号 e_{xc} , 将激励信号通过合成滤波器得到 8kb/s 的合成语音。

$$e_{xc}(n) = \hat{g}_p v(n) + \hat{g}_c c(n) \quad (14)$$

(3) 判断接收到的比特数是否大于或等于 320bits, 如果是,则继续解码出增强二层代数码激励及其增益; 否则,根据第(1)步解码出核心层基本参数和第(2)步解码出的增强一层代数码书及其增益,得到 12kb/s 的激励信号 e_{xc_12} , 将激励信号通过合成滤波器得到 12kb/s 的合成语音。

$$e_{xc_12}(n) = \hat{g}_p v_{12}(n) + \hat{g}_c c(n) + \hat{g}_{c1} c(n) \quad (15)$$

(4) 不管接收到的比特数是否大于或等于 480bits, 都先根据第(1)步解码出核心层基本参数、第(2)步解码出的增强一层代数码书及其增益和第(3)步解码出的增

强二层代数码书及其增益, 得到 16kb/s 的激励信号 e_{xc_16} , 将激励信号通过合成滤波器得到 16kb/s 的合成语音 s_{yn_16} , 然后再进行判断, 如果接收到的比特数大于或等于 480bits, 则根据增强三层格型矢量量化的参数及全局增益参数的索引值和前面解码出的线性预测系数一起得到采样率为 12.8kHz 的增强三层解码信号 q_{CX_24} , 否则, 输出解码出的 16kb/s 的合成语音。

$$e_{xc_16}(n) = \hat{g}_p v_{n_16}(n) + \hat{g}_c c(n) + \hat{g}_{c1} c_1(n) + \hat{g}_{c2} c_2(n) \quad (16)$$

(5) 判断接收到的比特数是否大于或等于 640bits, 如果是, 则根据增强四层格型矢量量化的参数及全局增益参数的索引值和前面解码出的线性预测系数一起得到采样率为 12.8kHz 的增强四层解码信号 q_{CX_32} , 将采样率为 12.8kHz 的 24kb/s 合成语音和解码出的增强四层信号相加, 得到 12.8kHz 采样的 32kb/s 合成语音。

$$s_{yn_32}(n) = q_{CX_32}(n) + s_{yn_24}(n) \quad (17)$$

否则, 将采样率为 12.8kHz 的 16kb/s 合成语音与增强三层解码信号相加, 得到 12.8kHz 采样的 24kb/s 合成语音 s_{yn_24} 。

$$s_{yn_24}(n) = q_{CX_24}(n) + s_{yn_16}(n) \quad (18)$$

(6) 将 12.8kHz 采样的合成语音调整至所需的采样率输出。如果输出采样率为 8kHz 的合成语音, 则直接进行下采样, 然后输出; 如果输出采样率为 16kHz 的合成语音, 则先进行上采样, 得到采样率为 16kHz 带宽为 [50Hz~6400Hz] 的语音, 然后再对上采样后的语音信号进行高频补偿, 输出频带范围为 50Hz~7000Hz、采样率为 16kHz 的合成语音。

由于嵌入式编码器输出的是嵌入式结构的码流, 所以, 在解码端可以根据接收到的码流大小, 解码出不同速率的合成语音。当网路容量足够时, 接收到全码流, 可以解码出最高码率的语音, 也可以根据不同终端的要求, 解码出不同码率的合成语音; 当遇到网路不同程度的拥塞时, 即使在网络传输时发生了分组丢失现象, 只要在解码端接收到大于某一门限的码流, 就可以解码出连续的语音, 保证语音质量。并且, 在保证语音质量没有明显下降的情况下, 可以进行任意层间的转换。

3 本编码算法的性能分析

根据 ITU-T 的客观测试语音质量的标准测试流程, 测试纯净语音在电平为 -26dB 下, 本文提出的 G.VBR 候选编码器和 G.722.2 编码标准相应码率之间的 MOS 对比如表 3 所示:

在办公室环境下, 测试本文提出的 G.VBR 候选编码器和 G.722.2 编码标准相应码率之间的 MOS 对比如

表 4 所示:

表 3 G.VBR 和 G.722.2 相应码率的 MOS 对比

汉语	G.VBR	8kb/s	16kb/s	24kb/s	32kb/s
		3.65767	3.8765	3.96567	4.0088
G.722.2	8.85kb/s	12.65kb/s	15.85kb/s	23.85kb/s	
	3.593	3.8835	3.96067	4.0733	
英语	G.VBR	8kb/s	16kb/s	24kb/s	32kb/s
		3.61317	3.89033	3.89733	3.9227
G.722.2	8.85kb/s	12.65kb/s	15.85kb/s	23.85kb/s	
	3.6135	3.90067	3.97467	4.1041	

表 4 背景噪声下 G.VBR 和 G.722.2 之间的 MOS 分对比

汉语	G.VBR	8kb/s	16kb/s	24kb/s	32kb/s
		3.63	3.926	4.1047	4.1643
G.722.2	8.85kb/s	12.65kb/s	15.85kb/s	23.85kb/s	
	3.7048	3.9422	4.0407	4.1617	
英语	G.VBR	8kb/s	16kb/s	24kb/s	32kb/s
		3.5947	3.92517	4.0918	4.131
G.722.2	8.85kb/s	12.65kb/s	15.85kb/s	23.85kb/s	
	3.676	3.9337	4.0447	4.1397	

测试纯净语音在电平为 -26dB 下, G.VBR 候选编码器五层之间快速转换时的 MOS 分和 G.722.2 编码标准在码率为 8.85kb/s 时的 MOS 分对比如表 5 所示:

表 5 G.VBR 层间转换和 G.722.2 相应码率的对比

汉语		英语	
G.VBR	G.722.2	G.VBR	G.722.2
3.87233	3.593	3.80567	3.6135

由以上结果可以看出, 本编码算法, 在纯净语音下, 每层都能达到不错的语音质量, 而且语音质量逐层提高。同时在背景噪声下和层间转换时也都能保证相当好的语音质量。

4 结论

本文基于 ACELP 和 TCX 编码技术, 提出了一种五层的宽带嵌入式变速率语音编码方法, 本编码方法可以有效地减少网络传输中分组丢失造成的影响, 适用于现代网络传输。实验结果表明, 本编码方法在纯净语音、背景噪声和层间转换时都取得了不错的效果, 而且作为国际上的四家 G.VBR 候选编码器之一提交给 ITU-T。

参考文献:

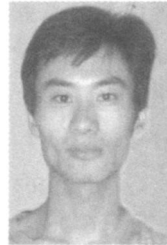
- [1] 鲍长春. 低比特率数字语音编码基础[M]. 北京: 北京工业大学出版社, 2001.
- [2] 王炳锡, 王洪. 变速率语音编码[M]. 西安: 西安电子科技大学出版社, 2004. 202-271.
- [3] Kyung Jin Byun, Ik Soo Eo, Hee Bum Jeong, Minsoo Hahn. An embedded ACELP speech coding based on the AMR-WB codec[J]. ETRI Journal, 2005, 27(2): 231-234.
- [4] Sean A Ramprasad. A tow stage hybrid embedded speech/aur

- dio coding structure[A]. IEEE ICASSP 98[C]. Seattle, Washington, USA, 1998. 337– 340.
- [5] A L Guyader, E Boursicaut. Embedded wideband VSELP speech coding with optimized codebooks[A]. IEEE Workshop on speech coding[C]. Quebec, 1993. 15– 16.
- [6] ITU-T TD260 (PLEN/16). Draft new Recommendation G. 729. 1 (ex G. 729EV): G. 729 Based Embedded Variable Bit rate Coder: An 8 32 kbit/ s Scalable Wideband Speech and Audio Coder Bitstream Interoperable with G. 729[R]. Geneva, ITU-T, 2006.
- [7] ITU-T TD33 (WP3/16). Terms of Reference (ToR) for the G. 729 Based Embedded Variable Bit Rate (G729EV) Extension to the ITU-T G. 729 Speech Codec[R]. Geneva, ITU-T, 2004.
- [8] 鲍长春, 李海婷, 等. 导抗谱频率参数的矢量量化方法及装置[P]. 中国专利: 200710003 193. 6, 2007.
- [9] 鲍长春, 范睿, 等. 语音编码中的自适应码书更新方法、系统及装置[P]. 中国专利: 200710084292. 1, 2007.
- [10] ITU-T. Coding of speech at 8kbit/s using conjugate structure algebraic code excited linear prediction (CS ACELP) [S]. (03/96).
- [11] 3GPP TS 26. 290 V6. 3. 0 Extended Adaptive MultiRate Wideband (AMR-WB+) codec [S]. 2005– 06.
- [12] H V Sorensen, D L Jones, M T Heideman, C S Burrus. Real valued fast fourier transform algorithm[J]. IEEE Trans on Signal Processing. 1987, 35(6): 849– 863.
- [13] M Xie, J P Adoul. Embedded algebraic vector quantization (EAVQ) with application to wideband audio coding [A].

IEEE ICASSP 96[C]. Atlanta, 1996. 240– 243.

- [14] 孙圣和, 陆哲明. 矢量量化技术及应用[M]. 北京: 科学出版社, 2002. 221– 231.

作者简介:



刘泽新 男, 1982 年生于山东菏泽, 硕士研究生. 研究方向为语音信号处理、宽带变速率语音编码等.



鲍长春 男, 1965 年生于内蒙古赤峰市, 博士, 教授、博士生导师, 国际语音通信学会(ISCA) 会员, 中国电子学会理事, 信号处理学会委员, 《通信学报》编委会副主任委员会、《信号处理》和《数据采集与处理》编委. 主要研究方向为语音与音频编码. E mail: chchbao@bjut. edu. cn



贾懋 男, 1982 年生于河北张家口, 博士研究生. 研究方向为语音、音频混合分层编码等.