

一种有效的典型波形块内插方法¹

鲍长春

(北京工业大学电子工程系 北京 100022)

摘 要 本文将基音周期波形 (PCW) 在时间方向上逐个相位对齐, 提出了一种新的线性预测残差 (LPR) 域典型波形 (PW) 提取和块内插方法。该方法可以避免内插后的 PCW 分界处的不连续而引起的听觉上的失真, 同时也使得 PW 分界的选择不至于太严格。借助于这种 PW 提取和块内插方法, 本文实现了一种 2.4 kb/s 线性预测编码方案, 计算机模拟结果表明, 该方案得到的重建语音与 4.8 kb/s 码激励线性预测算法接近。

关键词 语音编码, 线性预测, 波形内插, 矢量量化

中图分类号 TN912.3

1 引 言

与传统的合成分析语音编码方法 (如 CELP 算法^[1]) 相比, Kleijn 博士提出的典型波形内插 (PWI, Prototype Waveform Interpolation) 方法^[2-5]不再由波形匹配过程重建信号。这是因为当比特率减少时, 波形匹配的准确性下降, 致使重建语音有噪声特征, 这说明在低比特率客观 SNR 不能作为重建语音的理想测度, 而主观评价显得更重要。研究表明, 保持语音信号周期性的正确程度对感性质量非常重要^[6,7]。

PWI 方法就是力图在低比特速率重新产生浊音语音信号的周期性水平, 以达到改善语音感性质量的目的。浊音语音由一个高水平的周期性表征, 使得辨别单个基音周期成为可能。PWI 编码器的基本思想是以一定的时间间隔提取一个有代表性的基音周期波形 (PCW, Pitch-Cycle Waveform) 作为典型波形 (PW, Prototype Waveform), 传送描述这个 PW 的信息, 通过联合 PW 形状的线性内插和 PW 的长度重建语音。换句话说, 重建语音由 PCW 的减样序列获得。这个重建方法产生了 PCW 的平滑变化, 表达了浊音语音最普通的波形动态。

在低比特率语音编码器中, 回声和蜂鸣声是影响语音质量的重要因素, 这些影响与 PCW 的波形动态密切相关, 即与 PCW 随时间渐变的规律相联系, PWI 必须使得 PCW 的波形动态自适应。这种自适应需要长时信号改变比 (LTSCR, Long-Term Signal-to-Change Ratio) 和短时信号改变比 (STSCR, Short-Term Signal-to-Change Ratio) 两方面的制约, STSCR 用于测量相邻 PCW 间的信号改变比 (SCR), LTSCR 用于测量间隔为 20-30ms 的 PCW 间的 SCR。这种波形动态自适应将削弱两类失真, 第一类为 PW 的不准确量化引起的 PCW 形状大的起伏而产生的回声; 第二类为当用太多的周期性恢复包含大量送气声的语音信号时出现的音调人工产物。这种波形动态自适应过程在编码器中占据了很大的计算量, 使得实时实现编码算法困难。另外, 还需额外的比特传送 STSCR 和 LTSCR 信息, 增加了比特率。

实际上, 要想获得高质量语音, 在激励信号中, 插值方法不应产生不连续性, 而应有一个平滑的基音轮廓和平滑渐变的 PCW 形状。本文实验发现, 当块内插的基音轨迹使得 PCW 在 t 方向上最大程度平滑时, 即使得逐个 PCW 间相对对齐时, 可使得块内插产生的 PCW 分界处不连续性感觉失真最小。为此, 本文提出了一种新的 PW 提取和块内插 (blockwise interpolation)

¹ 1997-12-10 收到, 1999-02-27 定稿

本文工作得到国家自然科学基金 (69872005)、北京市跨世纪人才工程专项基金和北京市科技新星基金资助

方法。由于残差信号的大部分感性重要的信息位于基音脉冲附近,故本文在 LPR 域完成 PW 的提取、插值,这样做有两个优点,一是可以避免内插后的 PCW 分界处的不连续而引起的听觉上的失真;二是使得 PW 分界的选择不至于太严格。

借助于上述 PW 提取和内插方法,本文实现了一种 2400 bit/s 线性预测编码 (LPC, Linear Predictive Coding) 方案,该方案假定共振峰结构和激励信号独立变化,即在 PW 插值前,将共振峰结构和激励信号分离。激励波形和它的基音周期为一方面,共振峰结构为另一方面,这两个方面独立插值,然后利用插值的共振峰结构和激励信号重建语音信号。非正式主观试听表明,得到的重建语音与 4800 bit/s CELP 算法得到的重建语音接近。

2 PCW 个数的确定

令分析帧长为 $FRAME_SIZE$, P_{m-1} 和 P_m 分别为第 $m-1$ 帧和第 m 帧估计的基音周期, $p(k)$ 表示第 m 帧中第 k 个 PCW 的基音周期, M 为第 m 帧中包含的 PCW 个数。插值间隔开始于基音周期为 P_{m-1} 的第 0 个 PCW, 结束于基音周期为 P_m 的第 M 个 PCW。图 1 给出了确定 PCW 个数的示意图 (以 $M=3$ 为例)。

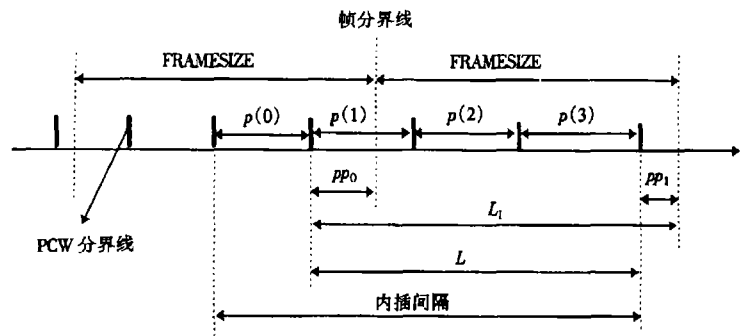


图 1 确定 PCW 个数的示意图 (以 $M=3$ 为例)

对于语音开始的第一帧, 即 $m=1$ 时, 取 $P_{m-1} = P_m$, 则 $M = \lfloor FRAME_SIZE / P_m \rfloor$, 这里 $\lfloor \cdot \rfloor$ 表示求最小整数, 由此得到用 PCW 序号 k 线性内插的基音周期为

$$p(k) = P_m, \quad k = 0, 1, \dots, M. \quad (1)$$

第一帧中 M 个 PCW 所包含的样点个数为

$$L = M \times P_m. \quad (2)$$

第一帧中去除 M 个 PCW 余留的样点个数为

$$pp = FRAME_SIZE \% P_m, \quad (3)$$

上式中符号 $\%$ 表示求余运算。

对于 $m \neq 1$, 令 pp_0 为第 $m-1$ 帧中去除它所包含的全部 PCW 后余留的样点个数, 则第 m 帧中实际确定 PCW 个数的样点数为

$$L_1 = FRAME_SIZE + pp_0. \quad (4)$$

考虑到基音的平滑变化及基音突变的影响, 令相邻帧基音的绝对误差为

$$\Delta = |P_m - P_{m-1}|. \quad (5)$$

如果 $\Delta > \varepsilon_{\max}$, 则限制第 $m-1$ 帧的基音为

$$P'_{m-1} = \begin{cases} P_{m-1} + \varepsilon_{\max}, & P_m > P_{m-1}; \\ P_{m-1} - \varepsilon_{\max}, & P_m < P_{m-1}. \end{cases} \quad (6)$$

如果 $\Delta \leq \varepsilon_{\max}$, 则第 $m-1$ 帧的基音保持不变, 即 $P'_{m-1} = P_{m-1}$.

上面的 $\varepsilon_{\max}$ 表示相邻帧基音变化的最大容许值, 按照 Sunberg 给定的数据^[8], 基频改变的最大速率为每毫秒 1%, 对于 25ms 的帧长, 最大频率改变将是 25%, 则时域中对应的范围为 $0.75P_m \sim 1.25P_m$, 由此可得 $\varepsilon_{\max}=25(\text{个样点})$.

这时用 PCW 序号线性内插的基音周期为

$$p'(k) = \frac{M' - k}{M'} P'_{m-1} + \frac{k}{M'} P_m, \quad k = 0, 1, \dots, M', \quad (7)$$

则第 m 帧中 M' 个 PCW 所包含的样点个数为

$$L = \sum_{l=1}^{M'} p'(l) = M' P'_{m-1} + \frac{1}{2}(M' + 1)[P_m - P'_{m-1}]. \quad (8)$$

依据 (8) 式, 依次取 $M' = 1, 2, 3, \dots$, 直到 $L > L_1$ 为止, 则最后实际得到的 PCW 个数为 $M = M' - 1$.

类似于 (7) 式和 (8) 式可以得到实际插值的基音周期 $p(k)$ 、第 m 帧中 M 个 PCW 所包含的样点个数 L 及第 m 帧中去除 M 个 PCW 余留的样点个数 $pp_1 = L_1 - L$.

3 PW 的提取

PW 的提取运作在 LPR 上, 在残差域提取特征波形的主要理由是周期扩展过程在特征波形形状中不易产生人工产物. 对于浊音语音而言, 防止周期扩展引发的明显的不连续性尤其重要, 为了使这种不连续性最小, 应使 PCW 分界附近的信号功率最小, 即不连续性位于信号低能区, 这就可以避免内插后的 PCW 分界处的不连续而引起的听觉上的失真. 由于残差信号的大部分感性重要的信息位于基音脉冲附近, 而 LPR 域通常已清晰地定义了基音脉冲和低功率区域, 所以, PW 的提取过程特别适合这个区域.

一旦从 LPR 信号中确定了 PCW, 则在 t 方向上, PCW 的平滑必须最大, 将当前得到的 PCW 与原来得到的 PCW 对齐能完成这种最大程度的平滑, 递归地使 PCW 间的相关最大可完成这种对齐.

令 $u(i, \tau)$, $0 \leq i \leq M-1$ 和 $u(j, \tau)$, $1 \leq j \leq M$ 分别对应于第 i 个和第 j 个 PCW, 这里 $u(i, \tau)$, $0 \leq i \leq M-1$ 定义在区间 $[-p(i)/2, p(i)/2]_{0 \leq i \leq M-1}$ 上, $u(j, \tau)$, $1 \leq j \leq M$ 定义在区间 $[-p(i)/2, p(i)/2]_{1 \leq j \leq M}$ 上. 通过添零定义扩展的位于中心的 PCW 为

$$\tilde{y}(k, \tau) = \begin{cases} u(k, \tau), & -p(k)/2 \leq \tau \leq p(k)/2; \\ 0, & \text{其他.} \end{cases} \quad (9)$$

这里 $k = 0, 1, 2, \dots, M$ 。“ $\sim$ ”表示扩展的基音周期波形 $\tilde{y}(k, \tau)$ 以原点为中心, 这个中心化处理并不意味着相继的 PCW 的特征 (如基音脉冲) 被对齐。

令 $y(k, \tau)$ 为与原来的 PCW 适当对齐的 PCW。将 $\tilde{y}(k+1, \tau)$ 移位一距离 ξ 使得失真测量 $D[y(k, \tau), \tilde{y}(k+1, \tau - \xi)]$ 最小可完成对齐过程:

$$\xi = \arg \min_{\xi'} D[y(k, \tau), \tilde{y}(k+1, \tau - \xi')]. \quad (10)$$

然后得到

$$y(k+1, \tau) = \tilde{y}(k+1, \tau - \xi). \quad (11)$$

上述过程递归进行, 直至 $k+1 = M$ 结束。这里 (10) 式中的 $D[\cdot, \cdot]$ 为以循环移位为基础的 PCW 间的互相关。

在残差域中, 仅仅在基音脉冲附近 (大多数能量集中处) 的 PCW 的表达与 PWI 方法的性能至关重要。所以分界的准确位置并不重要, 只要将 PCW 所包含的基音脉冲标志点置于其中心就足够了。基于这个理由, 为防止上述对齐过程中 PCW 包含的基音脉冲的漂移, 在上述对齐过程之前, 可将当前的 PCW 循环移位, 再与一个位于中心的三角窗相乘, 三角窗函数定义如下:

$$w(n) = 1 - \frac{2|n|}{N}, \quad n = -\frac{N}{2}, \dots, -1, 0, 1, \dots, \frac{N}{2}. \quad (12)$$

实际计算中 N 取 4 ~ 10 之间的偶数, 这样, 对齐前就可以保证基音脉冲位于 PCW 的中间, 实验表明这在 PCW 的基音周期里产生的误差确实小。

上述 PCW 间的对齐完成之后, 取第 M 个对齐的 PCW 作为要编码和传送的 PW, 至此完成 PW 的提取过程。

4 PW 的块内插

PW 的内插在译码端进行, 令 $\text{pcw}_0^{(Q)}(j)$, $0 \leq j < P_{m-1}$ 为前一帧已经对齐和量化的 PCW, $\text{pcw}_1^{(Q)}(j)$, $0 \leq j < P_m$ 为当前帧对齐和量化的 PCW。当译码器收到 P_{m-1} 和 P_m 后, 首先按照第 2 节介绍的算法重新确定 PCW 的个数和内插的基音周期 $p(k)$, 在此过程中, 由 (5) ~ (6) 式可得到插值所用的前一帧的基音周期为 P'_{m-1} 。

借助于 PCW 序号 k 的线性插值可计算中间的 PCW, 块内插算法分为如下两种情况进行:

(1) 如果 $P_m \geq P'_{m-1}$, 则对应前一帧 P'_{m-1} 的 PCW 的中点位置为

$$\text{mid}_0 = \lfloor P'_{m-1}/2 \rfloor, \quad (13)$$

对应当前帧的第 k 个 PCW 的中点位置为

$$\text{mid}_1^{(k)} = \lfloor P(k)/2 \rfloor, \quad 1 \leq k \leq M. \quad (14)$$

上述二者之差为

$$y_s^{(k)} = \text{mid}_1^{(k)} - \text{mid}_0, \quad 1 \leq k \leq M. \quad (15)$$

依照 (9) 式添零将前一帧已经对齐和量化的 PCW 扩展到长度为 $p(k)$:

$$E_{\text{pcw}_0}^{(Q)}(k, j) = \begin{cases} 0, & 0 \leq j < y_s^{(k)}; \\ \text{pcw}_0^{(Q)}(j - y_s^{(k)}), & y_s^{(k)} \leq j \leq y_s^{(k)} + P'_{m-1}; \\ 0, & y_s^{(k)} + P'_{m-1} \leq j < p(k). \end{cases} \quad (16)$$

则前向扩展的 PCW 残差为

$$\text{res}_0(k, j) = \frac{E_{\text{pcw}_0}^{(Q)}(k, j) * (M - k)}{M}, \quad 0 \leq j \leq p(k), \quad 1 \leq k \leq M. \quad (17)$$

为计算后向扩展 PCW 残差, 需重新确定 PCW 的中间位置, 对应当前帧 P_m 的 PCW 的中点位置为

$$\text{mid}_1 = \lfloor P_m/2 \rfloor, \quad (18)$$

而对应当前帧的第 k 个 PCW 的中点位置为

$$\text{mid}_0^{(k)} = \lfloor P(k)/2 \rfloor, \quad 1 \leq k \leq M. \quad (19)$$

上述二者中心位置距离为

$$y_s^{(k)} = \text{mid}_1 - \text{mid}_0^{(k)}, \quad 1 \leq k \leq M, \quad (20)$$

将当前帧已经对齐和量化的 PCW 截断到长度为 $p(k)$

$$T_{\text{pcw}_1}^{(Q)}(k, j) = \text{pcw}_1^{(Q)}(j + y_s^{(k)}), \quad 0 \leq j \leq p(k), \quad 1 \leq k \leq M, \quad (21)$$

则获得后向扩展的 PCW 残差为

$$\text{res}_1(k, j) = \frac{T_{\text{pcw}_1}^{(Q)}(k, j) * k}{M}, \quad 0 \leq j \leq p(k), \quad 1 \leq k \leq M. \quad (22)$$

将前向和后向扩展的 PCW 残差叠接相加, 即可得到最后的 PCW 残差为

$$\text{res}(k, j) = \text{res}_0(k, j) + \text{res}_1(k, j), \quad 0 \leq j \leq p(k), \quad 1 \leq k \leq M. \quad (23)$$

(2) 如果 $P_m < P'_{m-1}$, 则对应前一帧 p'_{m-1} 的 PCW 的中点位置为

$$\text{mid}_1 = \lfloor P'_{m-1}/2 \rfloor, \quad (24)$$

而对应当前帧的第 k 个 PCW 的中点位置为

$$\text{mid}_0^{(k)} = \lfloor P(k)/2 \rfloor, \quad 1 \leq k \leq M. \quad (25)$$

上述二者中心位置距离为

$$y_s^{(k)} = \text{mid}_1 - \text{mid}_0^{(k)}, \quad 1 \leq k \leq M. \quad (26)$$

令前一帧已经对齐和量化的 PCW 截断到长度为 $p(k)$

$$T_{\text{pcw}_0}^{(Q)}(k, j) = \text{pcw}_0^{(Q)}(j + y_s^{(k)}), \quad 0 \leq j \leq p(k), \quad 1 \leq k \leq M, \quad (27)$$

则获得前向扩展的 PCW 残差为

$$\text{res}_0(k, j) = \frac{T_{\text{pcw}_0^{(Q)}}(k, j) * (M - k)}{M}, \quad 0 \leq j \leq p(k), \quad 1 \leq k \leq M. \quad (28)$$

为计算后向扩展 PCW 残差, 需重新确定 PCW 的中间位置, 对应当前帧 P_m 的 PCW 的中点位置为

$$\text{mid}_0 = \lfloor P_m/2 \rfloor, \quad (29)$$

而对应当前帧的第 k 个 PCW 的中点位置为

$$\text{mid}_1^{(k)} = \lfloor P(k)/2 \rfloor, \quad 1 \leq k \leq M. \quad (30)$$

上述二者中心位置距离为

$$y_s^{(k)} = \text{mid}_1^{(k)} - \text{mid}_0, \quad 1 \leq k \leq M. \quad (31)$$

依照 (9) 式添零, 令当前帧已经对齐和量化的 PCW 扩展到长度为 $p(k)$,

$$E_{\text{pcw}_1^{(Q)}}(k, j) = \begin{cases} 0, & 0 \leq j < y_s^{(k)}; \\ \text{pcw}_1^{(Q)}(j - y_s^{(k)}), & y_s^{(k)} \leq j \leq y_s^{(k)} + P_m; \\ 0, & y_s^{(k)} + P_m \leq j < p(k); \end{cases} \quad (32)$$

则后向扩展的 PCW 残差为

$$\text{res}_1(k, j) = \frac{E_{\text{pcw}_1^{(Q)}}(k, j) * k}{M}, \quad 0 \leq j \leq p(k), \quad 1 \leq k \leq M. \quad (33)$$

利用 (23) 式将前向和后向扩展的 PCW 残差叠接相加, 即可得到最后的 PCW 残差。

级联上面得到的开始于 t_0 的 PCW 残差波形可获得激励波形 $e(t)$:

$$e(t) = \sum_{k=1}^M \text{res}(k, t - t_k), \quad t_1 \leq t \leq t_M. \quad (34)$$

将当前区间的最后一个 PCW 重新置为初始值, 可进行下一个区间的插值。

5 译码端 PW 的对齐

PW 的提取过程在每个修正点提供一个傅里叶级数表达, 在低比特率, 由于量化误差, 在译码端这些相继的 PW 未被适当对齐, 即两个 PW 的类似特征不出现在相位 φ 的类似值处, 这需要重新进行对齐处理, 以便内插。上述对齐既可在时域完成, 也可在频域完成。本文在频域完成对齐过程。

在频域对齐过程中要考虑谐波数偶而地加倍或减半情况, 为了进行对齐, 较短的 PW 应重复整数倍, 以便在对齐前与较长的 PW 长度匹配。对具有较少谐波数的 PW 而言, 这等于在原来的谐波间插入零幅度谐波。

由于所提取的 PW 是周期的, 所以 PW 可用傅里叶级数表达。令前一帧量化和对齐的 PW 表示为

$$u(t_{m-1}, \varphi) = \sum_{k=0}^K [C_k^{m-1} \cos(k\varphi) + D_k^{m-1} \sin(k\varphi)]. \quad (35)$$

当前帧量化和未对齐的 PW 表示为

$$u(t_m, \varphi) = \sum_{k=0}^K [C_k^m \cos(k\varphi) + D_k^m \sin(k\varphi)]. \quad (36)$$

这里求和数 K 是基音周期和信号带宽的函数, 信号带宽由采样语音的 Nyquist 频率决定, 其中 $D_0(t) = 0$. 将 $u(t_m, \varphi)$ 相位移位 ξ , 使得 $u(t_m, \varphi - \xi')$ 和 $u(t_{m-1}, \varphi)$ 间的互相关 $\text{Cor}(\xi')$ 最大可完成对齐.

$$\text{Cor}(\xi') = E < u(t_{m-1}, \varphi) u(t_m, \varphi - \xi') >. \quad (37)$$

将 (35) 式和 (36) 式代入 (37) 式得

$$\begin{aligned} \text{Cor}(\xi') = \frac{1}{2} \sum_{k=0}^K [C_k^{m-1} C_k^m + D_k^{m-1} D_k^m] \cos(k\xi') \\ + [C_k^{m-1} C_k^m - D_k^{m-1} D_k^m] \sin(k\xi'), \end{aligned} \quad (38)$$

(38) 式右边为以傅里叶级数系数表示的互相关.

取 $\xi = \arg \max_{\xi'} \text{Cor}(\xi')$, 则对齐的 PCW 的余弦和正弦系数为

$$\left. \begin{aligned} AC_k^m &= C_k^m \cos(k\xi) - D_k^m \sin(k\xi), \\ AD_k^m &= C_k^m \sin(k\xi) + D_k^m \cos(k\xi). \end{aligned} \right\} \quad (39)$$

6 一种 2400 bit/s 编码方案

运用上面的 PCW 块内插方法, 本文研究了一种 2400 bit/s 语音编码方案, 比特分配如表 1 所示.

表 1 2400bit/s 语音编码方案的比特分配 (帧速率为 40Hz)

参数	LSP	基音	PW 功率	PW 谱幅度	总比特数 (bit/s)
帧比特分配	25	7	5	8+8+7	2400

图 2 给出了编码端原理框架. LP 预测器为 10 阶, 用自相关方法进行分析. 用 PCW 序号 k 对量化的线谱频率 (LSF) 进行线性内插. LSF 系数以 40 Hz 速率提取, 用 25bit 进行分裂矢量量化^[10]. 基音检测采用归一化互相关方法^[11]. 对于 8 kHz 采样率, 用一个样点解析提取基音周期. LPR 用 LP 逆滤波器获得, PW 以 40Hz 的速率在 LPR 上提取, 然后确定 PW 的平均功率, 用其对 PW 归一化. PW 的平均功率以 40Hz 速率提取, 用 5bit 对其进行标量量化. 归一化谐波幅度分裂为三个频带, 运用变维矢量量化 (VDVQ) 技术^[12] 对它们分别进行矢量量化, 这三个频带的谐波范围分别为 1~15, 16~37 和 38~66, 维数分别为 15、22 和 29. 采用 GLA (Generalized Lloyd Algorithm) 算法^[9] 分别迭代设计体积为 256、256 和 128 的谱形状矢

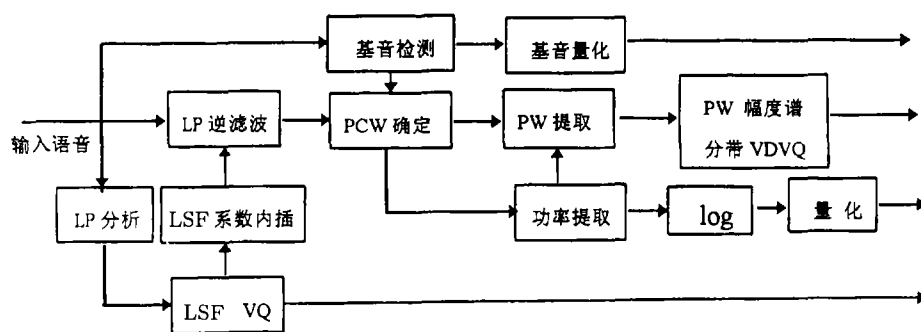


图 2 编码器原理框图

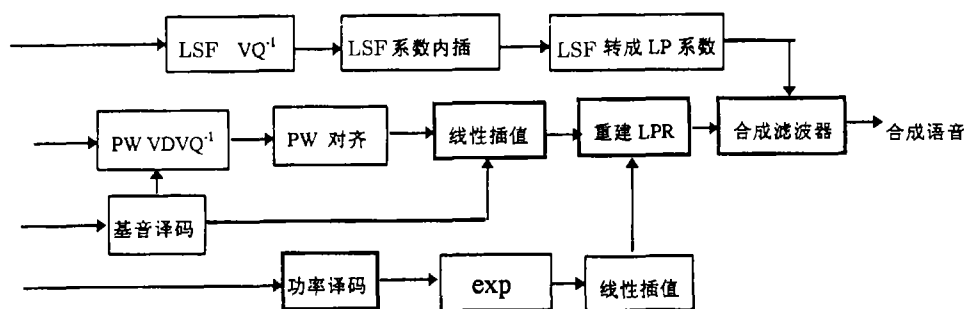


图 3 译码器原理框图

量码书。译码端根据接收到的基音译码值和三个码本的码字序号，分别将它们扩展到 $K_{\max}$ 维，然后将码字序号所对应的三个 $K_{\max}$ 维码矢量分别转换为 K_1 、 $K_2 - K_1$ 和 $K_3 - K_2$ 维矢量，将这些谐波矢量级联即可得到 K 维 PW 谱形状矢量，完成译码。其中 $K_1 = \lfloor 15P_m/2K_{\max} \rfloor$ 为第一个谐波分界点， $K_2 = \lfloor 37P_m/2K_{\max} \rfloor$ 为第二个谐波分界点， $K_3 = \lfloor 66P_m/2K_{\max} \rfloor$ 为第三个谐波分界点， $K_{\max}$ 为最大谐波数。

7 结 论

传统的 PWI 算法仅仅用于产生浊音语音信号，而在清音段要用其它方法对其编码。为了避免清浊误判带来的清浊过渡处语音波形的不连续性，清音段语音同样使用本文建议的 PW 块插值方法。当 PW 未量化时，由这种内插方法获得的重建语音非常接近于原始语音，当以 920 bit/s 对 PW 编码时，即整个编码速率为 2400 bit/s 时，非正式主观试听表明，得到的重建语音与 4800 bit/s CELP 算法得到的重建语音接近。

参 考 文 献

- [1] Schroeder M R, Atal B S. Code-excited linear predictive(CELP): High quality speech at very low bit rates. in Proc. ICASSP, 1985, 937-940.
- [2] Kleijn W B. Analysis-by-synthesis speech coding based on relaxed waveform-matching constraints: [Ph.D.thesis]. Delft University of Technology, Delft, The Netherlands: 1991.
- [3] Kleijn W B. Continuous representations in linear predictive coding, Proc. ICASSP'91, Toronto: 1991, 501-504.

- [4] Kleijn W B, Granzow W. Waveform Interpolation in Speech Coding. in the Speech and Audio Coding for Wireless and Network Applications, edited by B. S. Atal, Vladimir Cuperman and Allen Gersho, 1993, 111-118.
- [5] Kleijn W B. Encoding speech using prototype waveforms. IEEE Trans. on Speech and Audio Process. 1993, 1(4): 386-399.
- [6] Shoham Y. Constrained-excitation coding of speech at 4.8kb/s. in Advances in Speech Coding edited by B S. Atal, Vladimir Cuperman and Allen Gersho, 1991, 339-348.
- [7] Granzow W, Atal B S, Paliwal K K, Schroeter J. Speech coding at 4kb/s and lower using single-pulse and stochastic models of excitation. Proc. ICASSP'91, 1991, 217-220.
- [8] Sundberg J. Maximum speed of pitch changes in singers and untrained subjects, Journal of Phonetics, 1979, 7(1): 71-79.
- [9] Gersho A, Gray R M. Vector Quantization and Signal Compression. Kluwer Academic Publisher, 1991.
- [10] 鲍长春, 等. 线谱频率参数的分裂矢量量化. 电子科学学刊, 1998, 20(4), 508-513.
- [11] 鲍长春, 等. 基于归一化互相关函数的基音检测算法. 通信学报, 1998, 19(10): 27-31.
- [12] 鲍长春, 等. 低速率声码器中残差谱的变维矢量量化, 电子学报, 1999, 21(1): 7-13.

AN EFFICIENT BLOCKWISE INTERPOLATION ALGORITHM OF PROTOTYPE WAVEFORM

Bao Changchun

(Dept. of Electronic Engineering, Beijing Polytechnic University, Beijing 100022)

Abstract A new method of prototype waveform extraction and blockwise interpolation in Linear Predictive Residual(LPR) domain was suggested by aligning phase of Pitch-Cycle Waveform(PCW) along the time axis one by one. This method may avoid the perceptual distortion due to discontinuities near boundaries of interpolated pcw. At the same time, it made the boundary selection of Prototype Waveform(PW) not too strict. A novel 2.4 kb/s Linear Predictive Coding(LPC) model was established based on this extraction and blockwise interpolation of PW. The computer simulation results show that the reconstructed speech quality with this model closed to Code-Excited Linear Predictive(CELP) algorithm at 4.8 kb/s.

Key words Speech coding, Linear prediction, Waveform interpolation, Vector quantization

鲍长春: 男, 1965 年生, 教授, 博士, 中国电子学会高级会员, 研究兴趣为数字信号处理, 语音压缩编码.