

4.2Kb/s 码激励线性预测语音编码方案的实时实现¹

鲍长春 戴逸松* 樊昌信

(西安电子科技大学信息所 西安 710071)

*(吉林工业大学电子系 长春 130025)

摘 要 本文表述了一个 4.2Kb/s 码激励线性预测语音编码方案在单个 DSP 上的实时实现; 提出了一个减少自适应码本搜索复杂性的算法; 讨论了 A 参数转变为 LSP 参数的求解方法; 描述了该编码方案在具有一片 TMS320C30 的商用开发系统上的实时实现过程。

关键词 语音编码, 线性预测, 矢量量化

中图分类号 TN912.3

1 引 言

目前, 码激励线性预测 (CELP) 编码技术^[1]是人们最感兴趣的一种算法, 许多中、低速率语音编码方案都利用了它的基本思想, 诸如 16Kb/s 低延时码激励线性预测 (LD-CELP) 编码方案^[2], 8Kb/s 矢量和激励线性预测 (VSELP) 编码方案^[3], 4.8Kb/s CELP 编码方案^[4], 即便在 2.4Kb/s 或以下编码速率上也融进了其思想。由此可以看到 CELP 编码算法的潜在生命力。CELP 编码又称之为矢量激励编码, 它被看成是一帧语音的两级矢量量化。自适应码本首先用于表征语音信号的长时周期性, 然后得到的残差信号用一个固定的随机码本矢量量化, 这两种码本的搜索算法都基于包含感性加权的合成分析过程。这里的感性加权函数利用了人耳的掩蔽效应, 用于改进主观语音质量。CELP 编码算法有其致命的弱点, 即运算量庞大, 具体体现在两级码本的搜索上, 一般的数字信号处理 (DSP) 芯片很难实时实现, 故在近三、四年时间里, 许多研究者都把精力集中在实时实现上, 以达到实用化的目的。本文也不例外, 在研究了国内外有关 CELP 编码算法后, 提出了一些简化算法, 在中国科学院声学研究所十室研制的 TMS320C30-A 型商用开发板上实时实现了 4.2Kb/s CELP 编码方案, 为今后实用化奠定了基础。

2 基本 CELP 结构

本文所要考虑的基本 CELP 结构与文献 [4] 所表述的 CELP 算法相一致, 所以这里仅给出我们算法的三个基本功能。

(1) 短时延谱预测 使用了一个 10 阶线性预测滤波器, 分析窗为 Hamming 窗, 窗长为 30ms, 使用自相关法求得线性预测器系数, 帧间搭接 15ms。求得的预测器系数转化为线谱对 (LSP) 参数, 对每个激励子帧, LSP 线性内插以便形成一个中间参数组。

¹ 1995-01-16 收到, 1995-06-24 定稿

(2) 自适应码本激励 在这级里, 通过滤波—子帧选择到的和激励信号的过去样值, 形成一个 7.5ms 的语音子帧, 这个和激励信号为自适应码本激励信号与随机码本激励信号的加权和, 权为各自码本的增益, 自适应码本体积为 128 个码字, 奇子帧全搜索, 偶子帧增量搜索, 增量搜索范围为 32。

(3) 随机码本激励 通过滤波一个已选择到的三元值 (-1, 0, +1) 固定随机码本矢量, 获得一个 7.5ms 的残差信号矢量 (即原语音与由自适应码本激励得到的合成语音的差)。随机码本体积为 128 个码字, 码字间叠接 1。

3 自适应码本搜索

自适应码本的搜索过程实质等同于一阶基音预测器的闭环预测过程。一般地, 基音延迟 M 的范围为 20–147, 对应于自适应码本的 128 个码字, 故自适应码本可由 147 位移位寄存器 (SR) 构成, 结构如图 1。

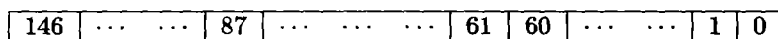


图 1 自适应码本结构图

当两级码本搜索完毕后, 两个码本优化激励码字的加权和从右端推入移位寄存器, 移位寄存器中旧的内容被移出。令 N 为子帧长, 下面分两种情况导出快速算法。

(1) $M > N - 1$ 情况 此时 $59 < M < 148$, 码字间存在着叠接 1 的关系, 这时第 M 个码矢量 AC_M 的第 n 个元素为

$$AC_M(n) = SR(M - 1 - n), \quad 0, 1, \dots, N - 1, \quad M = 60, 61, \dots, 147. \quad (1)$$

由此可以得到激励矢量与加权合成滤波器的脉冲响应 $h(n)$ 的卷积递推关系:

$$\begin{aligned} AT_M(n) &= \sum_{k=0}^n AC_M(n)h(n-k) = \sum_{k=0}^n SR(M-1-k)h(n-k) \\ &= SR(M-1)h(n) + \sum_{k=0}^{n-1} SR(M-1-1-k)h(n-1-k) \\ &= SR(M-1)h(n) + AT_{M-1}(n-1), \quad 1 \leq n \leq N-1 \end{aligned} \quad (2)$$

$$AT_M(0) = SR(M-1)h(0), \quad n = 0. \quad (3)$$

由 (2), (3) 式可知, 对于 $M = 61 - 147$, 计算每个 $AC_M(n)$ 仅需 N 次乘法。

(2) $19 < M < 60$ 情况 第 M 个码矢量 AC_M 的第 n 个元素为

$$AC_M(n) = SR(M - 1 - n \% M), \quad n = 0, 1, \dots, N - 1, \quad M = 20 - 59, \quad (4)$$

其中符号 % 表示求余。这时用上述类似方法得到卷积递推关系为

$$AT_M(0) = SR(M - 1), \quad n = 0; \quad (5)$$

$$AT_M(n) = SR(M - 1)h(n) + AT_{M-1}(n - 1), \quad 0 < n < M; \quad (6)$$

$$\begin{aligned} AT_M(n) &= SR(M - 1)h(n) + AT_{M-1}(n - 1) + AT_M(n - M) - AT_{M-1}(n - M), \\ &M - 1 < n < N. \end{aligned} \quad (7)$$

由 (5), (6), (7) 式可知, 对于 $20 < M < 60$, 计算全部 $AT_M(n)$ 仅需 $N(N+1)/2 + (60-21)(N-1)$ 次乘法, 而硬性卷积则需 $(60-20)N(N+1)/2$ 次乘法。

通过上述递推快速算法和在偶子帧进行增量搜索, 大大降低了计算量, 编码语音质量无感性降低。

4 随机码本搜索

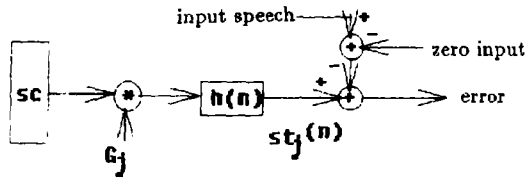


图 2 随机码本搜索框图

随机码本搜索如图 2 所示, 搜索过程基于下式最大:

$$E_{\max} = \left\{ \sum_{n=0}^{N-1} s'(n)st_j(n) \right\}^2 / \sum_{n=0}^{N-1} st_j^2(n), \quad (8)$$

其中 $s'(n)$ 为加权原语音减掉零输入后的语音信号, $st_j(n)$ 为随机激励通过加权合成滤波器的零状态响应。

随机码本中码字矢量由三元值 $(-1, 0, +1)$ 组成, 稀疏性为 90% (即 90% 的零元素), 码字间叠接 1, 码本体积为 L , 码字长度为 N , 所需存储空间为 $L-1+N$ 。文献 [5, 6] 提出了一种卷积运算的反序递推算法, 计算量比正序法下降一半, 由于篇幅所限, 仅给出卷积项 $st_j(n)$ 和能量项 $E_j = \sum_{n=0}^{N-1} st_j^2(n)$ 的递推公式:

$$st_j(n) = sc(L-1-j)h(n) + st_{j-1}(n-1), \quad 1 \leq n \leq N-1; \quad 1 \leq j \leq L-1; \quad (9)$$

$$st_j(0) = sc(L-1-j)h(0), \quad n=0, \quad 1 \leq j \leq L-1; \quad (10)$$

$$E_j = E_{j-1} - st_{j-1}^2(N-1) + st_j^2(0) + sc^2(L-1-j) \sum_{n=1}^{N-1} h^2(n) + 2sc(L-1-j) \sum_{n=1}^{N-1} h(n)st_{j-1}(n-1), \quad 1 \leq j \leq L-1. \quad (11)$$

通过上述简化递推算法, 随机码本搜索的复杂性大大减小, 计算量比较见文献 [5]。

5 线谱对 (LSP) 参数的求解

由线性预测器的传递函数 $A(z) = \sum_{i=0}^{10} \alpha_i z^{-i}$, $\alpha_0 = 1$ 可构造下面两个多项式:

$$P(z) = A(z) + z^{-11}A(z^{-1}), \quad (12)$$

$$Q(z) = A(z) - z^{-11}A(z^{-1}). \quad (13)$$

显然有

$$A(z) = (P(z) + Q(z))/2 \quad (14)$$

$P(z)$ 是对称多项式, $Q(z)$ 是非对称多项式, 通过如下变换可得两个对称多项式:

$$P'(z) = P(z)/(1+z^{-1}) = 1 + p_1 z^{-1} + \dots + p_1 z^{-9} + z^{-10}, \quad (15)$$

$$Q'(z) = Q(z)/(1-z^{-1}) = 1 + q_1 z^{-1} + \dots + q_1 z^{-9} + z^{-10}. \quad (16)$$

由于 $P'(z)$, $Q'(z)$ 在单位圆上有零值, 所以求 $P'(z)$, $Q'(z)$ 的根 $\{\exp(j\omega_i)\}$ 的问题, 就变成了求如下两个多项式根的问题。

$$P_a(x) = 32x^5 + 16p_1x^4 + 8(p_2 - 5)x^3 + 4(p_3 - 4p_1)x^2 + 2(p_4 - 3p_2 + 5)x + p_5 - 2p_3 + 2p_1, \quad (17)$$

$$Q_a(x) = 32x^5 + 16q_1x^4 + 8(q_2 - 5)x^3 + 4(q_3 - 4q_1)x^2 + 2(q_4 - 3q_2 + 5)x + q_5 - 2q_3 + 2q_1. \quad (18)$$

这里 $x_i = \cos(\omega_i)$, $\omega_i = 2\pi f_i/f_s$, $-1 \leq x_i \leq 1$, $0 < \omega_i < \pi$, f_s 为采样频率。

本文将 $[0, \pi]$ 区间化分为 314 份, 采用二分法求方程 $P_a(x)$, $Q_a(x)$ 的根。首先, 求 $P_a(x)$ 的 5 个根, 其次根据 LSP 的交错性 $\omega_0 < \omega_1 < \dots < \omega_9$, 可求得 $Q_a(x)$ 的另外 5 个根, 从而得到 10 个 LSP 参数 f_0, f_1, \dots, f_9 。由于要用到 cosine 函数, 而 TMS820C30 开发板没有满足此要求的库函数, 故本文根据最小二乘曲线拟合法得到一近似拟合多项式, 方法简介如下:

设给定 n 个点 x_i , $1 \leq i \leq n$ 和相应的函数值 y_i , $1 \leq i \leq n$, 我们要求出 $m-1$ 次最小二乘拟合多项式:

$$P(x) = \sum_{i=0}^{m-1} a_i x^i. \quad (19)$$

设拟合多项式为正交多项式 $q_i(x)$, $j = 1, 2, \dots, m$ 的线性组合:

$$p_m(x) = \sum_{i=1}^m c_i q_i(x), \quad (20)$$

而 $q_i(x)$, $j = 1, 2, \dots, m$ 可由以下递推公式得到

$$\left. \begin{aligned} q_1(x) &= 1, & q_2(x) &= (x - \alpha_2), & \dots, \\ q_{j+1}(x) &= (x - \alpha_{j+1})q_j(x) - \beta_j q_{j-1}(x), & j &= 2, 3, \dots, m. \end{aligned} \right\} \quad (21)$$

设 $d_j = \sum_{i=1}^n q_j^2(x_i)$, $j = 1, 2, \dots, m$, 则

$$\left. \begin{aligned} \alpha_j &= \sum_{i=1}^n x_i q_j^2(x_i) / d_j, & j &= 1, 2, \dots, m-1, \\ \beta_j &= d_j / d_{j-1}, & j &= 1, 2, \dots, m-1. \end{aligned} \right\} \quad (22)$$

上面得到的多项式 $\{q_j\}_{j=1,2,\dots,m}$ 是相互正交的, 根据最小二乘原理, 可得

$$c_j = \sum_{i=1}^n y_i q_j(x_i) / d_j, \quad j = 1, 2, \dots, m. \quad (23)$$

最后就可以化成一般的 $m-1$ 次多项式 (19) 式。

在实际计算中, 为防止运算溢出, x_i 用

$$x'_i = x_i - \bar{x}, \quad i = 1, 2, \dots, n \quad (24)$$

代替, 其中 $\bar{x} = \sum_{i=1}^n x_i/n$ 。此时拟合多项式变为

$$p_m(x) = \sum_{i=0}^{m-1} \alpha_i (x - \bar{x})^i. \quad (25)$$

本文对于区间 $[0, \pi]$, 求 $\cos(x)$ 时, 从 $x_1 = 0$ 开始, 取步长为 0.01, 数据点为 314 个, $m = 6$, 得

$$\begin{aligned} \bar{x} &= 1.565, & a_0 &= 0.0057938651, & a_1 &= -0.9997553991, & a_2 &= -0.0028772816, \\ a_3 &= 0.1658261584, & a_4 &= 0.0002157037, & a_5 &= -0.0075749635. \end{aligned}$$

6 4.2Kb/s CELP 编码方案的实时实现

表 1 给出了 4.2Kb/s CELP 编码方案的比特分配。

我们使用了中国科学院声学研究所研制的 TMS320C30A 型开发板, 该开发板上的单片处理器 TMS320C30 能完成 3300 万次浮点运算, 该开发板具有 512K 字节, 零等待高速存贮器,

它和主机构成 128k×32 位双寻址存贮

器, 高速处理板上有一片声频信号接口芯片 TLC32044C, 它具有 14 位 A/D 和 14 位 D/A

和可编程的输入输出模拟滤波器。加载程序为全汇编手工优化程序, 达到了实时效率。

95% 的中间变量分配空间在 RAM0 及 RAM1 中, 以便提高运算速度。根据我们所使用的

开发系统的结构, 我们的编码方案分为以下三个相互独立的过程:

(1) 传送过程 从输入缓冲区里读一帧语音和产生对应于输出编码语音的比特流。

(2) 接收过程 利用接收到的比特流合成语音, 合成语音写入到输出缓冲区里。

(3) 输入 / 输出过程 管理输入输出缓冲区。每隔 0.125ms 有一个内部时钟对处理器产生一次中断, 当中断发生时, 在中断程序里读入一个来自 A/D 转换器的样值, 并将其放到输入缓冲区里, 之后, 一个来自输出缓冲区的样值写入到 D/A 转换器。输入输出缓冲区设置在 0 等待高速存贮器里, 分别设两个缓冲区。

7 结束语

本文成功地实时实现了 4.2Kb/s CELP 编码方案, 得到了相当好的语音质量, 为今后实用化奠定了基础, 还有一些工作有待于去完成, 诸如程序的进一步优化, 算法的改进, 搭制硬件等。

参 考 文 献

- [1] Schroeder M R, Atal B S. Code-excited linear prediction(CELP): High quality speech at very low bit rates. in Proc. IEEE Int. Conf. ASSP, Tampa, FL: 1985, 937-940.
- [2] Cuperman V, Gersho A, Pettigrew R. Low-delay vector excitation coding of speech at 16Kb/s. IEEE Trans. on COM., 1992, COM-40(1): 129-139.

- [3] Gerson I, Jasiuk M A. Vector sum excited linear prediction(VSELP) speech coding at 8Kbps. in Proc. IEEE Int. Conf. ASSP, New Mexico, USA: 1990, 461-464.
- [4] Campbell J P, Welch V C, Tremain T E. An expandable error-protected 4800bps CELP coder(U. S. Federal standard 4800bps voice coder). in Proc. IEEE Int. Conf. ASSP, Glasgow, Scotland: 1989, 735-738.
- [5] Bao Changchun, *et al.* Improved populating code-excited linear predictive coder. The Journal of China Universities of Posts and Telecommunications. 1994, 1(1), 45-51.
- [6] 鲍长春, 等. 码激励线性预测语音编码快速算法及评价. 电子科学学刊, 1995, 17(1): 70-79.

A REAL-TIME IMPLEMENTATION OF 4.2Kb/s CELP SPEECH CODING

Bao Changchun Dai Yisong* Fan Changxin

(*Information Science Institute, Xidian University, Xi'an 710071*)

**(Dept. of Electronic Eng., Jilin University of Technology, Changchun 130025)*

Abstract This paper presents a real-time implementation of 4.2Kb/s CELP speech coding on single DSP. An algorithm reducing search complexity for adaptive codebook is suggested, the solving method, which changes a parameter into LSP parameter, is discussed. The real-time implementation procedure of this coding on a commercial development board with a single TMS320C30 is discribed.

Key words Speech coding, Linear prediction, Vector quantization

鲍长春：男，1965年生，副教授，博士后，中国电子学会高级会员，从事语音压缩编码的研究工作。

戴逸松：男，1936年生，教授，博士生导师，从事弱信号检测和数字信号处理等方面的教学和研究工作。

樊昌信：男，1931年生，教授，博士生导师，中国电子学会会士，IEE会士，IEEE会士，研究领域为语音和图象信号处理、B-ISDN等。