

基于 TETRA 的低速 ACELP 语音编码算法及其仿真

张志伟, 谢胜利

(华南理工大学 电信学院, 广东 广州 510640)

摘要: 描述了 TETRA 中的低速率语音编码算法实现过程, 并在 CCS 2.0 中对其进行了软件仿真. 由于这种低速率(4.567Kbps)的 ACELP 算法与应用于多媒体通信的变速率 G.723 标准、应用于未来个人通信的 8 Kbps 的 G.729 标准、第三代移动通信的语音压缩编码有许多相似之处, 因此对其算法仿真的研究和实时实现对第三代移动通信的语音压缩编码算法有比较重要的意义.

关键词: 语音编码; ACELP; 基音; 自适应码本; 固定码本

中图分类号: TN91 **文献标识码:** A **文章编号:** 1007-855X(2004)06-0070-05

The Simulate of Low Rate Speech Code(ACELP) Algorith Based on TETRA

ZHANG Zhǐwei, XIE Sheng2li

(School of Electronics and Communication, South China University of Technology, Guangzhou 510640, China)

Abstract: Full specification of the speech codes used in the TETRA (Trans-European Truncked Radio) system is presented. The TETRA speech code is based on the Code-Excited Linear Predictive (CELP) coding model, so the algorithm consists of Short-term prediction. Long-term prediction analysis and algebraic codebook search are carried out. Also the result of algorithm simulation is presented.

Key words: speech code; ACELP; pitch; adaptive codebook; fixed codebook

0 引言

基于 ACELP 的语音编码算法有很多种, G.729, G.728, G.723 等都 ITU 制定的是一些标准协议, 而在 TETRA 里用到的是一种改进了的低速率的 ACELP 算法. 它的速率只有 4.567 Kbps, 比其他的算法都要低, 但在语音可懂性方面却可以和速率为 8 Kbps 的 G.729 相比. 这些不同的算法标准都是基于代数码激励线性预测(ACELP)算法的. 图 1 为基本代数码激励算法的实现框图.

1 算法的详细设计

TETRA 里的语音压缩算法与其他码激励算法的不同点在于它对 LSP 系数采用了分裂矢量量化, 简化了自适应码本的搜索, 采用了代数码激励码本来简化固定码本的搜索. 它还加长了帧长(30 ms), 每一帧由 4 个子帧组成(7.5 ms), 每一帧用 137 bits 编码, 表 1 为在编码中各个编码参数的比特分配.

1.1 预处理

语音编码的输入为 16 bits 的线性 PCM 编码的数字信号, 在进行编码之前要对其进行预处理, 其中包括电平的调整是为了减少计算中溢出的可能性, 还有就是高通滤波去除里面的直流信号和低频分量. 其预处理函数的 Z 域传输函数为:



图 1 代数码的激励性预测算法原理图

Fig.1 The block diagram of ACELP

收稿日期: 2004-07-14.

第一作者简介: 张志伟(1982.4~), 男, 硕士研究生. 主要研究方向: 通信工程. E-mail: zhzwzhangvian@sina.com.

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

$$H_p(z) = \frac{1}{2} \left[\frac{1-z^{-1}}{1-Az^{-1}} \right], \quad A = 32\,735/32\,768 \quad (1)$$

1.2 线性预测与 LSP 量化

线性预测与 LSP 量化是在一整帧里进行的. 由自相关系数求出线性预测系数, 由于线性预测系数的动态范围较大, 而且语音信号波形包络对线性预测系数的变化很敏感, 因此将线性预测系数(LP) 转化成各个参数间有较强独立性和动态范围较小的线谱对(LSP) 参数来传输. 语音信号的 LSP 参数的帧间相关性较强, 对其采用分裂矢量量化可以减少量化比特数.

表 1 Tetra 声码器的比特分配

Tab. 1 Bit allocation for the TETRA code

参数	第 1 子帧	第 2 子帧	第 3 子帧	第 4 子帧	整帧
LP 滤波器					26
基音延迟	8	5	5	5	23
代数码本	16	16	16	16	64
增益量化	6	6	6	6	24
总计	30	27	27	27	137

1.2.1 线性预测

LP 窗口分析由两个不同长度的半 Hamming 窗组成, 窗口函数如下:

$$w(n) = \begin{cases} 0.54 - 0.64 \cos\left(\frac{Pn}{L_1 - 1}\right), & n = 0, \dots, L_1 - 1 \\ 0.54 + 0.64 \cos\left(\frac{P(n - L_1)}{L_2 - 1}\right), & n = L_1, \dots, L_1 + L_2 - 1 \end{cases} \quad (2)$$

32 ms 的分析窗口 (在采样频率为 8 kHz 时对应于 256 个样本) 由 $L_1 = 216$ 和 $L_2 = 4$ 组成, 窗口中的 40 个样本将会是下一帧的 (前向 40 个样本). 线性预测就是对这 32 ms 的语音段处理的, 由加窗后的语音信号 $sc(n)$ 按下式计算出自相关函数:

$$r(k) = \sum_{n=k}^{255} sc(n)sc(n-k) \quad k = 0, \dots, 10 \quad (3)$$

上式的自相关函数经过一个 60 Hz 的带宽扩展后得到修正后的自相关函数 $rc(k)$, 再由 Levinson - Durbin 算法求解方程组的线性预测系数(LP).

$$\sum_{i=1}^{10} a_i rc(|i-k|) = -rc(k) \quad k = 0, \dots, 10 \quad (4)$$

1.2.2 LSP 的量化

将 LP 参数转换成 LSP 参数:

$$F_1(z) = \frac{A(z) + z^{-11}(A(z^{-1}))}{1 + z^{-1}} = \prod_{i=1,3,5,7,9} (1 - 2q_i z^{-1} + z^{-2}) \quad (5)$$

$$F_2(z) = \frac{A(z) - z^{-11}(A(z^{-1}))}{1 - z^{-1}} = \prod_{i=2,4,6,8,10} (1 - 2q_i z^{-1} + z^{-2}) \quad (6)$$

$q_i = \cos(w_i)$, $i = 1, 2, \dots, 10$ 就是线谱对(LSP) 参数. 很显然 q_i 分布在单位圆上. 对 q_i 的量化采用分裂矢量量化, 对 10 维 LSP 矢量用 26 bits 量化, 将 10 维矢量分为三个长度分别为 3, 3, 4 的子矢量, 分别用 8 bits, 9 bits, 9 bits 进行矢量量化. 对于子帧的 LSP 系数由上一帧和当前帧的系数内插得到.

$$\begin{cases} q_1 = 0.75q_n + 0.25q_{n-1} \\ q_2 = 0.5q_n + 0.5q_{n-1} \\ q_3 = 0.25q_n + 0.75q_{n-1} \\ q_4 = q_n \end{cases} \quad (7)$$

初选值为 $q_{n-1} = \{30\,000, 26\,000, 21\,000, 15\,000, 8\,000, 0, -8\,000, -15\,000, -21\,000, -26\,000\} / 32\,768$.

1.3 自适应码本分析与基音延迟的量化编码

ACELP 算法模型中的长时预测综合滤波器是用自适应码本来实现的, 它的作用主要用来逼近语音信

号的准周期结构. 自适应码本是一个矢量码本, 当前帧的自适应码本由之前的激励码本的 T_{\max} 个元素构成, 因此长时综合滤波的目的就是要搜索出最佳的 T_{\max} 和最佳增益 g_{pit} . 算法中基音搜索范围被限定在 $[20 \sim 143]$ 之内.

1.3.1 基音延迟的分析

为了简化搜索过程, 将其分为两部分: 开环搜索和闭环搜索. 开环搜索用到的是加权的语音信号 $s_w(n)$, 加权滤波器为 $\frac{A(z/0.95)}{A(z/0.60)}$, 其中用到的是没有量化的 LP 系数. 在实现中, 为了减少搜索计算量, 激励在搜索阶段是用 LP 残余信号扩展的. 通常提高抽样率可以提高基音估计的准确性, 因此算法中在每子帧内更新一次, 并且采用分数基音估计. 在第 1 子帧内, 分数基音延迟用 $1/3$ 来量化, 范围为 $[19 \frac{1}{3} - 84 \frac{1}{3}]$, 整数延迟在 $[85 \sim 143]$. 其他三个子帧的搜索范围在 $[T_1 - 5 \frac{2}{3} - T_1 + 4 \frac{2}{3}]$, T_1 是与第 1 子帧的分数基音延迟最接近的整数.

开环估计中, 在 $[20 \sim 39]$, $[40 \sim 79]$ 和 $[80 \sim 142]$ 三个范围计算下式的最大值:

$$C_k = \sum_{j=0}^{120} s_w(2j) s_w(2j - k) \quad (8)$$

对三个最大值进行归一化得到 $R(t_i)$, 再按优先较低延迟原则选出最佳基音估计 T_{op} .

闭环搜索是以最小化原语音信号和合成语音信号之间的均方误差为原则的, 即最大化下式, $x(n)$ 是目标信号, 它是加权的输入信号减去加权合成滤波器 $H(z)W(z)$ 的零输入响应, 而 $y_k(n)$ 是过去的滤波激励信号的 k 延时, 对于分数延时用内插进行计算.

$$S_k = \frac{\sum_{n=0}^{59} x(n) y_k(n)}{\sqrt{\sum_{n=0}^{59} y_k(n) y_k(n)}} \quad (9)$$

为减少计算量, 第一子帧的闭环搜索是在开环搜索基音 T_{op} 附近 $T_{\text{op}} \pm 2$ 范围内进行的.

1.3.2 基音延迟的编码

第 1 子帧的基音延时 T_1 用 8 bits 编码, T_1 由 $\text{nint}(T_1)$ 和 $\text{frac}/3$ 组成, $\text{frac} = -1, 0, 1$, 即 $T_1 = \text{nint}(T_1) + \text{frac}/3$, 码序列 P_1 为:

$$P_1 = \begin{cases} 3 @(\text{nint}(T_1) - 19) + \text{frac} - 1 & \text{for } T_1 \in [19 \frac{1}{3}, 85], \text{frac} \in \{-1, 0, 1\} \\ (\text{nint}(T_1) - 85) + 197 & \text{for } T_1 \in \{86, 87, \dots, 143\}, \text{frac} = 0 \end{cases} \quad (10)$$

其他 3 帧由于都是以 T_1 为中心搜索的, 因此它们的编码序列 P_i ($i = 2, 3, 4$) 为:

$$P_i = 3 @(\text{nint}(T_1) - t_{\min}) + \text{frac} + 2 \quad (i = 2, 3, 4) \quad (11)$$

$t_{\min} = \text{æt} \hat{\text{a}} - 5$, 且 $t_{\max} = t_{\min} + 9$, 但 t_{\min} 和 t_{\max} 要限定在 $[20 \sim 143]$ 内.

1.4 代数码本的搜索

为了逼近语音信号中的随机部分, 一般压缩算法中用白噪声逼近. 由于代数码本为稀疏矢量结构(由 4 个大小确定且占据一定位置集的非零值交错排列而成), 且不用保存, 因此用代数码本来逼近可以减小算法复杂度.

1.4.1 代数码本的结构

每个码本最多只有四个非零的脉冲, 固定的幅度分别为 $+1.414, -1, +1, -1$ 四个, 第一个脉冲的位置用 5 bits 编码, 其他三个位置的脉冲各用 3 bits 编码, 码本可以同时移动一位到达奇数位置, 因此用 1 bit 来表示移动(shift), 还有 1 bit 用来同时将码本的符号取反. 具体分配见表 2.

1.4.2 代数码本的搜索

代数码本的搜索是在最小化加权输入语音和加权合成语音的均方误差原则上的, 用于闭环搜索的目标信号减去自适应码本的影响后就是代数码本搜索的目标信号.

$$x_2(n) = x(n) - g_p y(n) \quad n = 0, \dots, 59 \quad (12)$$

式中 $y(n) = v(n) * h(n)$, 称为被虑除的自适应码本矢量, $h(n)$ 为加权合成滤波器的冲激响应 $H(z)W(Z) = 1/A(Z/C)$.

代数码本的作用是动态地加强重要频率部分, 所用的整形滤波就是上面的加权合成滤波器, 固定码本的选择就是最大化下式:

$$S_k = \frac{C_k^2}{E_k} = \frac{(d^T \alpha_k)^2}{c_k^T S \alpha_k} \quad (13)$$

α_k 是序号为 k 的码本, H 为上三角 Toeplitz 卷积矩阵, 由 $hc(0), hc(1), \dots, hc(59)$, $d = H\alpha_2$ 是后向滤波目标矢量, $S = H^T H$.

由于代数码本只在 4 个固定位置上有值, 因此搜索的计算量不会很大. 由于语音的随机性, 需要进一步地简化搜索算法, 在进入后两个循环时, 要与提前计算的门限进行比较, 只有超过了门限才可以进入下一个循环, 并且循环的最大次数固定, 这样固定码本的搜索只有很低的百分比. 两个门限都是基于自相关系数 C 的, 前两个循环可能的最大自相关为 \max_2 , 加上第三个循环可能的最大值为 \max_3 , 门限值分别为 $k_2 \max_2$ 和 $k_3 \max_2$, 其中 $0 < k_2, k_3 < 1$ 它们用来控制搜索的百分比, 一般取 0.568. 此外还设置一个计数器, 初始值为 350, 第三个循环执行一次, 计数器减 4, 第四个循环执行一次, 计数器减 3, 当计数器为零时, 停止搜索.

1.5 增益量化

在以上的搜索中, 还会有两个重要的参数需要计算, 即自适应码本增益 g_{pit} 和代数码本增益 g_{code} :

$$g_{pit} = \frac{\sum_{n=0}^{59} x(n)y(n)}{\sum_{n=0}^{59} y(n)y(n)} \quad \text{和} \quad g_{code} = \frac{d^T c}{c^T S c} \quad (14)$$

$x(n)$ 为自适应码本的目标矢量, $y(n)$ 是最佳自适应码本 $v(n)$ 经过感知加权综合滤波的滤波矢量, c 为最佳代数码本.

根据语音信号的短时平稳性可以得知增益有较强的帧间相关性, 可以采用能量域的预测误差矢量量化编码来减少编码比特数. 每一子帧用 6 bits 编码.

这样得到的语音编码参数: LSP 参数、基音延迟周期、最佳代数码本、增益量化编码参数, 在接收端将这四种参数经过与编码相反的滤波和合成就可以得到译码信号了.

2 算法的仿真

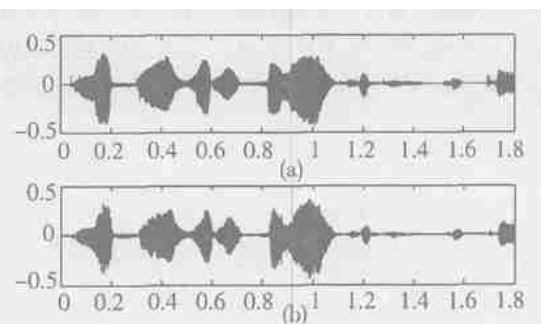
在 CCS 2.0 用 TMS320C54x 的仿真环境, 对上述算法做了软件仿真, 图 2 就是仿真结果.

比较原始信号和解码信号不难看出, 虽然这种算法不是波形编码, 但语音信号的波形还是得到了恢复. 因此在码速率要求低的情况下, 这不失为一种好的压缩算

表 2 代数码本的结构表

Tab. 2 Structure of the algebraic codebook

码本参数	脉冲位置	比特分配
幅度: 1.4142	0, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58	5
幅度: -1	2, 10, 18, 26, 34, 42, 50, 58	3
幅度: 1	4, 12, 20, 28, 36, 44, 52	3
幅度: -1	6, 14, 22, 30, 38, 46, 54	3
全局符号		1
移位标志		1



(a) 原始语音信号; (b) 为解码语音的信号

图 2 算法仿真结果

Fig. 2 Simulating result

法. 很明显这种算法最大的缺点就是延迟太大, 一帧的延迟是60 ms(包括编码与解码), 主要原因是帧长太长, 子帧7.5 ms比G. 729的子帧5 ms还长. 因此这种编码算法更加适合于低速率语音传输系统中.

参考文献:

- [1] 成德源. 8 kb/s低复杂度代数码激励线性预测编码[J]. 电路与系统学报, 1997, 2(4): 44~ 48.
- [2] Soong F K, Juang B H. Line spectrum pair(LSP) and speech data compression[J]. Proc. IEEE ICASSP, 1984.
- [3] Terrestrial Trunked Radio (TETRA); Speech codec for full-rate traffic channel; Part 2: TETRA codec, ETSI ETS 300 395-2 ed. (2): 1998~ 02.

(上接第 69 页)

比较矩阵中几个结果, 相对集中有 LivingComparison 和 TotalComparison 的散点图, foodComparison 与 TotalComparison 的散点图, 通过以上分析, 确定 LivingComparison 和 TotalComparison 为主分量.

下面分别用 gplot 和 tree 方法^{[3][6]}作以上两个主分量的散点图和树状图, 见图 5 和图 6.

图 5 中用不同颜色直观地显示不同的类, 可以看出划分为 3 类比较合理, 从而证实了前面分析结果的有效性. 图 6 中从右端开始一类分为两类, 到了最左端完全是每一个观察一类了, 而且标出了观察的客户编号, 由于客户较多不宜纵向展开, 所以显得很乱, 但基本可以看出分为 3 类比较合理.

综合以上讨论和分析可以得到结果: 根据样本数据将考察客户群分为三类.

4 结论

利用数据挖掘技术中的聚类分析, 可以把考察的客户群进行分类, 这样对不同的客户进行定位, 以便针对不同的客户采取不同的营销策略, 以及进行不同的服务内容. 在这里利用类间最小方差法的改进系统聚类分析方法对实际调查样本数据进行聚类分析, 不但可以直观地看到聚类过程, 而且根据技术分析可以得到聚类结果, 从主分量散点图和树状分类图可以清楚地判别分组情况是比较合理的, 从而证实了模型的正确性和有效性.

参考文献:

- [1] 朱明. 数据挖掘[M]. 合肥: 中国科学技术大学出版社, 2002. 129~ 162.
- [2] 张维明. 数据仓库原理与应用[M]. 北京: 电子工业出版社, 2002. 182~ 206.
- [3] 朱扬勇, 左子叶. 数据挖掘实践[M]. 北京: 机械工业出版社, 2003. 159~ 201.
- [4] 堂林燕. 数据仓库和数据挖掘技术在 ERP 中的应用[J]. 上海: 计算机工程, 2002, 28(9): 10~ 18.
- [5] 杨东龙. 客户关系管理[M]. 北京: 中国经济出版社, 2002. 106~ 138.
- [6] 樊欣. SAS 经济统计[M]. 北京: 北京希望电子出版社, 2003. 259~ 270.