

# 线性预测编码(LPC)语音合成器的设计\*

李永明 陈弘毅 于政

(清华大学微电子学研究所 北京 100084)

**摘要** 本文提出一种线性预测编码(LPC)语音合成器。它主要是由可编程 12 极点格型滤波器, 二极点低通滤波器及 12 位的 D/A 转换器组成, 内部并带有与外部 CPU 的接口电路。门级的逻辑模拟和 C 编译程序的模拟结果说明, 该合成器可在低数据率下合成出高质量语音。它可用于 LPC 语音分析合成系统和 LPC 语音合成芯片 IC 的开发。

EEACC: 2570, 1290, 6130

## 1 前言

近年来, 数字信号处理(DSP)技术进步很快, 特别是 VLSI 的数字电路, 从处理速度和规模上都有了很高的水平, 极其复杂的系统在单片上实现已成为现实, 在语音的抽样速率下, 信号处理工作所要求的大量计算也可以实时实现, 各种语音合成技术都在不断地发展。LPC 语音合成技术, 以它的低数据率、低复杂度、低成本, 受到特别的重视, 六十年代后期发展起来的 LPC 语音分析方法可有效地估计基本语音参数, 基音、共振峰, 谱、声道面积函数, 可以对语音的基本模型给出精确的估计, 而且计算速度较快<sup>[1,2,4]</sup>。线性预测语音分析合成技术是基于频域的分析合成方法。基于语音参数分析、合成方法的 LPC 语音技术, 以它的低数据率( $<1.8\text{ kbit/s}$ ), 便于窄带传输和大量语音信息的存储得到广泛重视, 使该技术可应用于通信、银行服务、玩具产品、教学辅助设备等等, 应用开发可延伸到各个领域。

LPC 的编码方法, 在低码率语音编码方法中是好的, 具有相当的可懂度<sup>[5,6]</sup>, 但在语音的自然度方面则较差。为此, 人们提出了残余误差激励线性预测编码(RELPC), 多脉冲激励线性预测编码(MPELP), RPE(Regular-Pulse Excitation), CELP(Code-book Excited Linear Predictive Coding)等方法。但这些方法大大增加了计算的复杂度和计算量, 码率也相应提高, 对集成语音合成器来讲, 有着较高的难度<sup>[3]</sup>。我们提出的语音合成器设计方案, 仍保留 LPC 编码方法最初的优点, 采用了音调同步插值及音调控制激励信号产生的方法, 并且采用了与语音分析阶段的予加重相适应的去加重技术, 改善了语音质量。运用这些方法实施电路设计至晶体管级, 并作了初步验证, 由电路设计的模拟结果和计算机仿真效果

\* 国家自然科学基金资助项目

李永明 男, 1945 年生, 副教授, 从事大规模模拟集成电路设计和教学, 语音信号处理方面集成电路的设计等方面研究

陈弘毅 男, 1942 年生, 教授, 博士生导师, 主要从事大规模集成电路设计, 数字信号处理等方面研究

于政 男, 1969 年生, 清华大学微电子学专业硕士, 从事语音合成功器研究

1996 年 1 月 11 日收到初稿, 1996 年 6 月 4 日收到修改稿

比较表明,所设计的LPC语音合成器可以在1.6kbit/s数据率下,产生出达到商品应用价值的语音质量.

## 2 电路设计考虑

LPC语音压缩技术应用于语音合成,其基本思想是来源于源-滤波器模型结构.由白噪声序列和周期脉冲序列构成的激励信号,经选通、放大并通过时变数字滤波器(由语音参数控制的声道模型),从而再获原语音信号,这种参数编码的语音合成器的框架见图1.

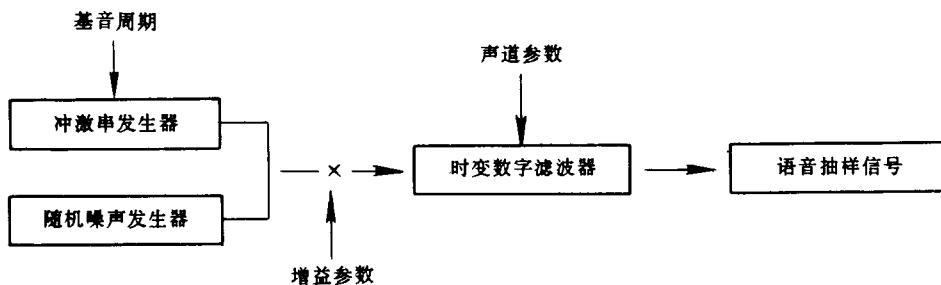


图1 语音合成简化模型框图

我们采用了这简化模型进行语音合成器的设计.模型参数有浊音/清音判决,浊音的音调周期及滤波器参数 $K_i$ 等.合成器的数字滤波器采用了12极点的格型滤波结构(LPC-12),是全极点滤波器,来源于线性预测误差滤波器的格型结构<sup>[7,8]</sup>.采用格型滤波结构实现语音合成,不仅有其直接的物理意义,可以保证系统的稳定性,而且易于VLSI的实现.LPC语音合成器采用简化模型,可以大大减小激励信号的存储量,但与严格的分析过程是不一致的,结果音质下降.因而,近年来有许多提高音质的方法,但计算复杂,码率高.因此,我们在降低数据率,简化硬件系统与提高音质之间采用了折衷处理.

其一,激励模型上仍沿用简化激励模型,白噪声应用于清音,浊音采用准周期脉冲,不使用单一的脉冲,而使用脉冲组,相当于码本激励的一组最优“码本”.合成时根据音调插值后的低四位来选择相应脉冲组,实验表明,这种激励方法在数码率不增加的情况下,对音质有着明显的改善.

其二,原始语音信号中,由于人类发声时口唇辐射而引起语音信号的高频衰减,在机理模型中含有低通滤波作用.在进行语音分析时,为了使LPC合成模型与声管模型相一致,提高参数提取精度,在LPC分析系统中加入予加重处理,常用予加重传递函数为: $H(z)=1-az^{-1}$ , $a$ 为予加重系数,这代表了人类发声时口唇辐射模型的影响,实验表明,在大多数情况下,取 $a=0.93$ ,可较多地去除口唇辐射的影响.所以在合成器的设计中要设置去加重,去掉分析时予加重的影响.恢复原始语音,用一个两极点数字低通滤波器来实现传输函数为: $y(i)=C_1x(i)+C_1y(i-2)-C_2y(i-1)-C_2x(i-1)$ ,其中 $x(i)$ 表示格型滤波器当前输出, $y(i)$ 为当前低通输出, $x(i-1)$ 表示为格型滤波器前一采样点的输出, $y(i-1)$ , $y(i-2)$ 为低通滤波器输出的前两个样点信号, $C_1,C_2$ 为低通滤波器系数.低通滤波器的系数若选择适当,则去加重的效果就最佳.

在语音合成技术中,参数插值是一重要的处理方法,特别是在LPC语音合成器的设计中,由于硬件复杂度受到限制,而产生截去误差,这时插值技术的引入则更为重要。有时两帧参数之间变化会很大,若不进行插值处理,音质会很差,一般要加入帧间的平滑处理。参数插值通常采用线性插值方法,该方法简单,可较好地近似人类发声过程的反射系数及音源参数的平稳特性;而且保证了参数不会因为插值而超出其数值范围,确保了系统的稳定性。插值过程对各种LPC参数均适用。这样,帧间的参数不再跳变,具体插值算法为: $a'_i(t) = a_{i+1}(a_{i+1} - a_i) \times t/T_f, 0 \leq t \leq T_f$ ,该式是在第*i*帧合成时参数的插值公式, $a_i$ 和 $a_{i+1}$ 是第*i*、*i*+1帧的原参数值, $a'_i(t)$ 是第*i*帧内时刻*t*的参数插值, $T_f$ 是帧长。参数插值时插值步距的选择尤为关键,本文设计的语音合成器的插值步距由音调控制,即进行音调同步插值。合成时,在音调处理器中对音调进行减计数,当音调参数减至某一值时,合成器即发出插值申请,外部CPU根据这一申请进行插值运算,然后及时将结果参数送至合成器。实验表明,在波谷进行的音调同步插值丰富了语音基频谐波,由于插值在能量最低处进行,可使语音插值噪声降到最低水平。

根据抽样定理,若模拟信号频带受限,抽样率至少为2倍Nyquist频率。但语音信号为非固定的频带信号。就浊音而言,一般情况下频率大于4kHz时,其峰值会下降40dB以上;而清音,高频大于8kHz时峰值却无明显下降。因而要精确表示语音,抽样率应大于20kHz,但这样会造成在线性分析中计算量太大。因此,通常采用的做法是:在抽样前使原始语音信号通过一个4kHz锐截止模拟滤波,而采样率就可降为8kHz或10kHz,本文的ASIC采用了10kHz采样率。

根据阿塔尔(ATAL)的经验公式阶数  $P = f_s(r=2\sim 5)$ , ( $f_s$ :采样率)我们选用线性预测阶数为12。

### 3 电路结构及工作原理

一个完整的语音合成系统应包括CPU,语音库ROM, RAM及语音合成器<sup>[6]</sup>。LPC语音合成器应当实现以下功能:(1)激励信号的产生;(2)参数的分离与译码;(3)合成过程的参数插值;(4)格型滤波器的递推运算及结果输出;(5)合成信号的去加重处理。

这里设计的LPC语音合成器完成上述功能的(1)、(4)、(5)三部分,由一般通用微处理器来实现功能中(2)、(3)两部分。该ASIC的设计中含合成器和微处理器之间接口电路。合成一个样点,12阶格型滤波的乘加运算量加上低通滤波的运算量,10kHz采样率时ASIC总的乘加运算量在20万次/秒以上。图2为合成器的总体结构图。

语音合成过程如下:首先,来自CPU的INIT信号初始化合成器,然后在CPU控制下分离,译码语音参数,并进行插值参数运算,将它送入合成器的RAM内。当这些参数准备好后,CPU通过接口电路又分别向控制电路和音调处理器送数。控制电路启动格型滤波器工作,格型滤波器按一定时序从RAM中读出能量, $K_i$ 及 $C_1, C_2$ 等参数,从激励发生器中取激励,运算后产生的采样信号进入D/A转换器(DAC)。最后,DAC输出信号是脉宽调制信号,在片外经滤波平滑为语音信号。同时,控制电路产生INT1信号向CPU申请换帧,CPU接受此信号后即调用换帧子程序进入下一帧信号处理,参数插值申请信号INT2由音调处理器产生。

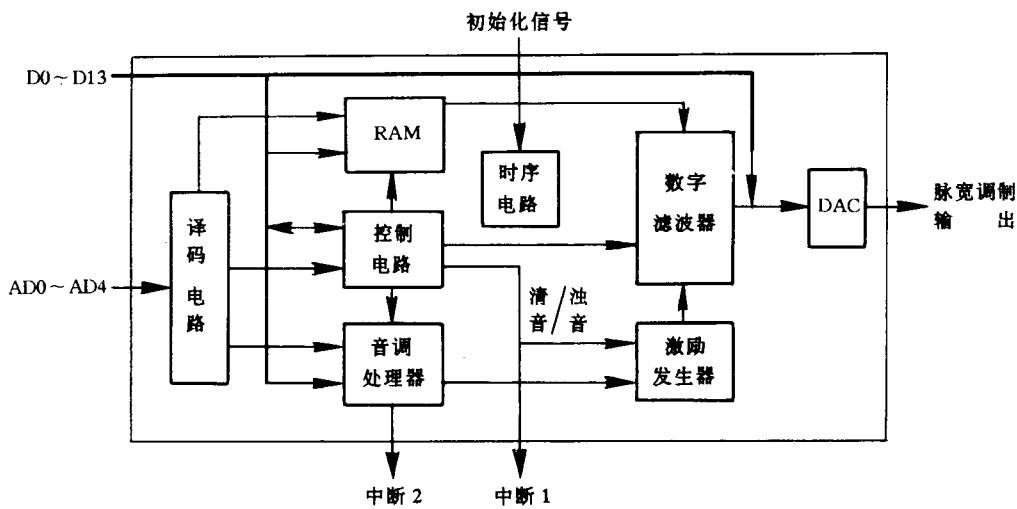


图 2 LPC 语音合成器功能块示意图

下面对合成器的部分功能块作简要说明：

1. 时序电路：来自 CPU 初始化信号进入初始化电路后，产生系列内部初始化信号，使电路启动，片内时钟频率为 4.8MHz. 该电路产生片内工作时序系列。
2. RAM：共 16 个单元，暂存 CPU 送入的能量( $E$ )、 $K_i$  及低通参数( $C_1$ 、 $C_2$ ). 通过数据总线(D0~D13)与外部相连，通过译码器按地址 AD0~AD4 选通 RAM，并写入数据. 然后，在合成时序控制下，RAM 的“00→0F”存储单元轮流选通并适时向数字滤波器提供数据.
3. 数字滤波器及其工作：

(1) 图 3 为格型滤波器原理图.

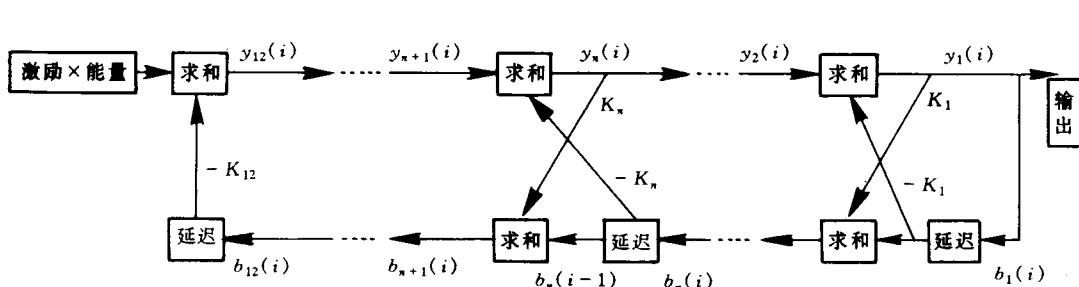


图 3 12 阶格型滤波原理图

(2) 数字滤波器结构.

从原理图可看到，12 级格型滤波器进行的运算是非常大的. 这里采用了波茨编码原理设计的 14 位高精度“乘加器”，实时地实现  $14 \times 12\text{bit}$  的乘法运算和加法运算，完成 12 级的格型滤波和二级数字低通滤波运算，简化了硬件结构. 移位寄存器是实现流水线操作的关键部分，在合成当前采样点信号时，要用到前一采样点的反向预测误差信号  $b_{12}(i-1), \dots, b_1(i-1)$  及前面合成样点的值. 一个  $14 \times 14$  位的移位寄存器阵列可实现上述目标，用于存

储上述数据适应流水线运算操作；而正向预测误差则为高阶逐步推算至低阶与格型滤波方向相同，用暂存寄存器即可实现。这样，“乘加器”、移位寄存器及其时序控制电路构成流水线工作的数字滤波器。

(3) 一个采样点合成过程：第一步取激励( $R$ )；第二步实现  $E$  与  $R$  相乘，第三步进行音调运算，产生下一个样点激励地址，并进行第 12 阶格型滤波  $Y_{12}(i) = E \times R - K_{12} \times b_{12}(i-1)$ ；第四步完成其余 11 阶格型滤波，从原理图可知其滤波方程为：

$$\begin{aligned}y_n(i) &= y_{n+1}(i) - K_n \times b_n(i-1) \\b_{n+1}(i) &= b_n(i-1) + K_n \times y_n(i)\end{aligned}$$

第五步实现二极点的低通滤波运算。

## 4 验证

该电路的设计工作做至器件级，电路的总元件数为 1.2 万器件。按 CMOS  $2\mu\text{m}$  N 阵工艺设计，为验证电路设计是否合理，目前采用了以下方法：(1) 利用 LPC 语音分析合成系统对新设计思想逐个调试比较，并得到确认；(2) 用 C 语言编译程序模拟 ASIC 中各功能块，用软件进行全系统仿真，产生的语音质量是好的，相应的数据率为  $1.6\text{kbit/s}$ ；(3) ASIC 硬件设计验证，是通过全电路逻辑模拟进行。输入参数为 LPC 语音分析合成系统分析出的参数，工艺参量参考标准  $2\mu\text{m}$  CMOS 工艺。表 1 数据为相应的连续产生的一些数据，可看到它们的差别很小，说明 ASIC 硬件设计是可接收的。

表 1 部分数据比较

C 程序运行结果	逻辑模拟运行结果	差 别
-0.001609	-0.001953	0.000344
0.009112	0.009277	-0.000165
0.016805	0.016602	0.000203
0.023774	0.023926	-0.000152
0.031945	0.031738	0.000207
0.036290	0.036621	-0.000331
0.040806	0.041016	-0.000209
0.045648	0.045410	0.000238
0.051575	0.051758	-0.000183
0.050774	0.050781	-0.000007
0.051243	0.051270	-0.000027
0.049930	0.049805	0.000125
0.045105	0.044922	0.000183
0.042908	0.042480	0.000428

## 5 结论

这里所设计的 LPC 语音合成器，经验证表明，现有电路设计是可行的。它具有下面的特点：LPC 参数插值计算和激励函数的选取是通过音调处理器加以控制；浊音激励采用一组优化“码本”；LPC 参数和激励函数可以根据用户要求灵活选择；电路接口和电路控制简

单；可根据不同用户实际情况配以专用微处理器，便可较方便地实现语音合成系统的单片集成化。

## 参 考 文 献

- [1] H. Witten, "Principles of Computer Speech", London, New York 1982.
- [2] L. R. Rabiner and R. W. Schafer, "Digital Processing of Speech Signals", Prentice-Hall, Inc., 1978.
- [3] 丁俊明, "LPC 语音合成技术的研究与应用", 清华大学硕士论文, 1994.
- [4] F. Fallside, "Computer Speech Processing", Prentice/Hall International (UK) LTD., 1985.
- [5] P. Italiano, G. Ponte and M. Sartori, IEEE Trans. on Consumer Electronics, 1985, **CE-31**, NO. 3.
- [6] B. Fette, D. Harrison, D. Olson *et al.*, IEEE J. of SSC, 1983, SC-18(1): 23.
- [7] F. I. Itaknra and S. Saito, "A Statistical Method for Estimations of Speech Spectral Density and Formant Frequencies", Elec. and Commu., in Japan, 1970, **53-A**(1):36~43.
- [8] J. Makhoul, IEEE Trans. Acoustics, Speech and Signal Proc, 1977, **ASSP-25**(5): 423~428.

## Design on LPC Speech Synthesizer

Li Yongming, Chen Hongyi and Yu Zheng

(Institute of Microelectronics, Tsinghua University, Beijing 100084)

Received 11 January 1996, revised manuscript received 4 June 1996

**Abstract** A dedicated LPC speech synthesizer is presented. It mainly acts as a programmable 12-pole lattice filter with a 2-pole low-pass filter, a 12bit D/A converter and a built-in interface to external CPU. The gate level logic simulation and the results of C programs show that the synthesizer can synthesize high-quality speech at low data rate. It can be used in LPC analyze - synthesize systems and developments of LPC speech synthesizer IC.

**EEACC:** 2570, 1290, 6310