

# 协方差斜格法的LPC参数提取及其FPGA实现

范炳远, 方建安

(东华大学信息科学与技术学院, 上海 201620)

**摘要:** 分析语音信号声道特征参数提取问题, 针对自相关法的缺陷, 提出声道特征参数提取的改进算法。介绍其运算步骤和流程, 考虑FPGA适于短期开发及高速性的优点, 设计Finite State Machine来控制复杂运算操作及对寄存器的频繁访问。利用Cyclone EP1C6 FPGA实现语音信号声道特征参数提取算法。

**关键词:** 语音信号; 声道模型; 特征参数; 协方差斜格法

## LPC Parameters Extraction of Covariance Lattice Method and Its Implementation on FPGA

FAN Bing-yuan, FANG Jian-an

(College of Information Science and Technology, Donghua University, Shanghai 201620)

**【Abstract】** This paper analyzes the problem of sound track characteristic parameters' extraction of speech signal. According to the disadvantage of self-correlation, an improved algorithm of the sound track characteristic parameters' extraction of speech signal is presented to introduce its operational procession. In view of the high speed and suit for short term exploitation of FPGA, it designs the Finite State Machine so as to control the complicated operations and frequent visit to the registers for the FPGA design principles. The algorithm is implemented successfully into the single Altera Cyclone EP1C6 FPGA to fulfill the algorithm of the sound track characteristic parameters' extraction of speech signal.

**【Key words】** speech signal; sound track model; characteristic parameters; covariance lattice method

### 1 概述

语音信号处理是语言学与数字信号处理技术相结合的交叉学科, 主要包括3个研究领域: 语音编码, 语音合成和语音识别<sup>[1]</sup>。语音信号处理步骤如下: (1)对输入语音信号进行预处理, 通过反混叠滤波消除工频信号的干扰; (2)进行数字化和特征参数提取, 用反映语音信号特点的若干参数代表语音; (3)根据不同任务, 采取不同的处理方法。本文研究的特征参数提取主要针对语音编码领域。

经过70年的发展, FPGA已有百万门级, 通过不断更新、优化产品架构和生产工艺, 实现了更多逻辑单元、更高性能、更低单位成本和功耗。FPGA采用流水线技术设计系统, 可极大提高系统工作速度。

随着可用IP核数量的增加, FPGA可以完成许多DSP功能。如滤波器、相关器、正弦/余弦发生器、存储器、数学功能等, 从而支持各种应用。随着微电子技术的发展, 越来越多的语音信号数字处理系统需要以实时方式或在线方式工作, 特别在嵌入式系统设计中, 对系统的硬件环境要求更高。本文采用一款高性能低成本的FPGA——Altera公司的Cyclone EP1C6器件。Cyclone架构由垂直结构的逻辑单元(LE)、嵌入式存储块和锁相环(PLL)周围环绕着的I/O单元(IOE)、RAM块组成, 逻辑容量为5980个LE<sup>[2]</sup>。

### 2 语音信号的声道模型

可以认为声道是一个谐振腔, 共振峰是这个腔体谐振时产生的, 因此, 也称声道模型为共振峰模型。由于人耳听觉的柯蒂氏器官的毛细胞按频率感受排列其位置, 因此共振峰

模型被广泛应用。实践表明用前3个共振峰可以代表一个元音, 对辅音或鼻音通常须使用5个以上共振峰<sup>[3]</sup>。

一般可用全极点模型来刻画共振峰特性, 即

$$v(z) = \frac{1}{1 - \sum_{i=0}^p a_i z^{-i}} \quad (1)$$

其中,  $p$ 为全极点滤波器的阶, 一般取共振峰的2倍;  $a_i$ 为声道模型的参数, 它随声道的调音运动不断变化, 声道惯性使这些参数变化的速度受到限制, 一般在10ms~30ms的间隔内, 认为这些参数保持不变, 这是语音信号短时分析的理论依据之一。

### 3 线性预测参数的协方差斜格法提取

由于自相关法需要进行加窗处理, 因此参数精度很差。理论上自相关法可以保证系统的稳定性, 但在实际应用中, 有限字长的影响造成自相关函数计算精度不够, 导致病态矩阵, 降低了系统稳定性。

由于斜格法无须用窗函数对语音信号进行加窗, 且能保证系统稳定性, 因此较好解决了精度与稳定性的矛盾。但斜格法无法得到推广的原因是其运算量太大, 一般为自相关法的4倍以上, 不适于硬件的实现。

本文改进传统斜格法, 将协方差法与斜格法相结合。此方法保证了系统稳定性及精度, 降低了运算量。

**作者简介:** 范炳远(1983-), 男, 硕士, 主研方向: 语音编码, 语音识别, 现场可编程门阵列; 方建安, 教授、博士、博士生导师

**收稿日期:** 2007-10-12 **E-mail:** qbx-d-boy@163.com

### 3.1 协方差斜格法的原理

在 Levinson-Durbin 算法中, 当递推进行到第  $i$  阶时, 可得第  $i$  阶的预测系数  $a_j^{(i)}$  ( $j=1,2,\dots,i$ ), 可以定义一个第  $i$  阶的线性预测(Linear Predictive, LP)误差滤波器, 即

$$A^{(i)}(z) = 1 - \sum_{j=1}^i a_j^{(i)} z^{-j} \quad (2)$$

利用

$$a_j^{[i]} = a_j^{[i-1]} + k_i a_{i-j}^{[i-1]}, \quad 1 \leq j \leq i-1 \quad (3)$$

可得

$$A^{(i)}(z) = A^{(i-1)}(z) - k_i z^{-i} A^{(i-1)}(z^{-1}) \quad (4)$$

根据信号第  $i$  阶预测误差  $e^{(i)}(n)$ , 可得其 Z 变换为

$$E^{(i)}(z) = A^{(i-1)}(z)S(z) - k_i z^{-i} A^{(i-1)}(z^{-1})S(z) = E^{(i-1)}(z) - k_i z^{-i} B^{(i-1)}(z) \quad (5)$$

根据  $E^{(i)}(z)$  和  $B^{(i)}(z)$ , 可得正向预测误差信号  $f^{(i)}(n)$  和反向预测误差  $b^{(i)}(n)$  如下:

$$f^{(i)}(n) = s(n) - \sum_{j=1}^i a_j^{(i)} s(n-j) \quad (6)$$

$$b^{(i)}(n) = s(n-i) - \sum_{j=1}^i a_j^{(i)} s(n-i+j) \quad (7)$$

推导可得  $f^{(i)}(n)$  和  $b^{(i)}(n)$  的递推关系式, 即

$$\begin{cases} f^{(i)}(n) = f^{(i-1)}(n) - k_i b^{(i-1)}(n-1) \\ b^{(i)}(n) = b^{(i-1)}(n-1) - k_i f^{(i-1)}(n) \\ f^{(0)}(n) = b^{(0)}(n) = s(n) \end{cases} \quad (8)$$

### 3.2 协方差斜格法的求解

在求解  $a_i$  时, 先求得  $k_i$  值, 再利用式(12)计算  $a_i$ , 本文对斜格法的改进实际上是改写  $F, B, C$  的表达式, 使其成为协方差  $c(i, j)$  的函数形式, 可得

$$F^{(i)}(n) = E[(f^{(i)}(n))^2] = \sum_{k=0}^i \sum_{j=0}^i a_k^{(i)} a_j^{(i)} c(k, j) \quad (9)$$

$$B^{(i)}(n-1) = E[(b^{(i)}(n-1))^2] = \sum_{k=0}^i \sum_{j=0}^i a_k^{(i)} a_j^{(i)} c(i+1-k, i+1-j) \quad (10)$$

$$C^{(i)}(n) = E[(f^{(i)}(n)b^{(i)}(n-1))] = \sum_{k=0}^i \sum_{j=0}^i a_k^{(i)} a_j^{(i)} c(k, i+1-j) \quad (11)$$

根据协方差的时间平均估计, 利用得到的  $c(i, j)$ , 可以进行  $F, B, C$  的计算, 从而极大减小运算量。

本文采用几何平均斜格法求解 LP 参数, 可得

$$k_i = \frac{\sum_{n=0}^{239} f^{(i-1)}(n)b^{(i-1)}(n-1)}{\sqrt{\sum_{n=0}^{239} [f^{(i-1)}(n)]^2 \sum_{n=0}^{239} [b^{(i-1)}(n-1)]^2}} \quad (12)$$

运用 Cauchy-Schwarz 不等式, 可以证明  $-1 < k_i < 1$ 。此方法确定的反射系数可保证系统稳定性。由于没有引入窗函数, 因此系统误差小于自相关法。可以利用式(3)得到最终的 LP 参数  $a_i$ 。

## 4 声道特征参数的量化

由上述方法得到的 LP 参数不是语音编码中需要提取的特征参数, 经矢量量化后的 LP 参数, 即线性预测编码参数(Linear Predictive Coding, LPC)才是语音编码所需声道特征参数。

由于 LSP 参数和语音信号的信号谱包络的峰联系更紧密, 语音信号的频谱具有明显声学特性并对外界环境变化有一定顽健性, 因此更适于进行量化。需要将  $\{a_i\}$  转换成 LSP 参数。

先将  $q_i$  转换为  $\omega_i$ , 即  $\omega_i = \arccos(q_i)$ ,  $i = 1, 2, \dots, 10$ 。本文

用与 5 个人的随机谈话中产生的码本, 并通过 LBG 算法进行训练。本文采用二级分裂矢量进行量化。第 1 级是一个 10 维 VQ, 使用 128 个(7 bit)码字的码书  $L1$ 。第 2 级是一个 10 维 VQ, 使用 2 个 5 维码书按分裂的 VQ 实现, 每个码书( $L2$  和  $L3$ )都包含 32 个(5 bit)码字。

在量化过程中, 每个  $\omega_i$  由 2 个码书的和得到, 即

$$\hat{\omega}_i = \begin{cases} L1_i(L1) + L2_i(L2) & i = 1, 2, \dots, 5 \\ L1_i(L1) + L3_{(i-5)}(L3) & i = 6, 7, \dots, 10 \end{cases} \quad (13)$$

为了避免在量化的 LP 合成滤波器中出现尖锐的共振, 系数  $\hat{\omega}_i$  应该进行调整, 使相邻系数有最小距离  $J$ 。进行 2 次调整过程, 第 1 次使  $J=0.0012$ , 第 2 次使  $J=0.0006$ 。调整后用  $\hat{\omega}_i^{(m)}$  表示,  $\hat{\omega}_i^{(m)}$  表示当前第  $m$  帧消除共振后的 LSF 系数。最终调整后的 LSF 系数  $\hat{\omega}_i^{(m)}$  应该是以前的量化器输出  $\hat{\omega}_i^{(m-k)}$  和现行量化器输出  $\hat{\omega}_i^{(m)}$  的加权和, 即

$$\hat{\omega}_i^{(m)} = (1 - \sum_{k=1}^4 \hat{p}_{i,k}) \hat{\omega}_i^{(m)} + \sum_{k=1}^4 \hat{p}_{i,k} \hat{\omega}_i^{(m-k)}, \quad i = 1, 2, \dots, 10 \quad (14)$$

其中,  $\hat{p}_{i,k}$  是可切换 MA 预测器的系数。

量化后的 LSF 参数要转换回 LP 参数  $\hat{a}_i, i = 1, 2, \dots, 10$ 。

## 5 FPGA 实现

声道特征参数提取的一个重要特点是整个算法由多个递推过程实现, 在整个递推过程中, 全部参与运算量的初值、中间值和最终值都小于 1<sup>[4]</sup>, 因此, 很适于采用定点运算的硬件系统来实现。本文采用 FPGA 实现声道特征参数提取算法。

由于此算法中存在减法运算, 因此在设计中采用补码形式, 所有运算都采用定点运算方式来实现。FPGA 的设计应遵循面积和速度的平衡与互换原则。面积与速度指标是设计质量评价的终极标准<sup>[5]</sup>。本文综合考虑面积和速度问题, 采用流水线技术并调用 Altera 宏功能运算模块(LPM), Altera 提供的 LPM 均基于 Altera 器件的结构做了优化设计, 确保了面积与速度性能最优。协方差斜格法部分采用的运算单元包括一个 16 位的定点补码乘法器、一个 16 位的定点除法器、一个 32 位的定点加法器。在电路中设计了 3 个 16 位 11 字的存储器 R、A 和 An。其中, R 存储  $F^{(i)}(n), i = 1, 2, \dots, 10$ ; An 存储正  $B^{(i)}(n)$ ; A 存储  $C^{(i)}(n)$ 、采用双 RAM 结构和同时取数同时运算的并行处理技术。用计数器 counter( $i$ ) 计数误差滤波器的阶数。用计数器 counter( $j$ ) 计数  $i$  阶的第  $j$  个误差滤波器的系数。对 LP 系数的量化和插值部分采用牛顿迭代法进行 LP-LSP 参数的转换, 由于量化和插值需要 6 组 10 维的线性预测参数, 因此设计 6 个 RAM 用以存储这些参数, 利用线性预测分析中的运算器和计数器 counter( $i$ ) 实现最终算法。

## 6 实验仿真结果

声道特征参数提取算法用 Altera Cyclon EP1C6 FPGA 实现。先使用 VHDL 语言对算法进行描述, 然后用 ModelSim 将算法编码成 RTL 级 VHDL, 并进行编译和仿真, 用 Aletra QuarterS5.0 进行综合, 最后将编译综合产生的 SOF 文件下载到目标器件 Cyclon EP1C6Q240C8。目标器件的资源利用情况如表 1 所示。

表 1 CYCLON EP1C6Q240C8 的资源利用率

Resources	Number used	Maximum available	Percentage used/(%)
logic elements	2 013	5 980	33.6
pins	112	185	60.5
memory bits	1 826	92 160	2.0

仿真的时间周期是 28 ns, 为了验证算法的正确性, 将由

随机函数产生的 16 位二进制源码序列作为算法的输入信号。表 2 对自相关分析的 FPGA 运算结果与协方差斜格法的 FPGA 运算结果进行比较, 自相关法的误差为 1.362 1, 协方差斜格法的误差为 0.823 1。表 3 是量化后的 LPC 系数。

表 2 FPGA 实现结果的比较

线性预测系数	自相关法的结果	协方差斜格法的结果
$a_1$	-0.410 7	-0.392 3
$a_2$	0.121 8	0.100 1
$a_3$	-0.061 2	0.037 8
$a_4$	0.032 8	0.012 4
$a_5$	0.057 4	0.051 0
$a_6$	-0.062 9	-0.037 1
$a_7$	-0.071 2	-0.029 9
$a_8$	0.091 8	0.135 2
$a_9$	-0.032 4	-0.026 3
$a_{10}$	0.029 8	0.031 1

表 3 FPGA 实现量化后 LPC 系数的结果

声道特征参数	FPGA 实现结果	声道特征参数	FPGA 实现结果
$\hat{a}_1$	-0.389 9	$\hat{a}_6$	-0.019 3
$\hat{a}_2$	0.099 3	$\hat{a}_7$	-0.008 9
$\hat{a}_3$	0.079 4	$\hat{a}_8$	0.242 0
$\hat{a}_4$	0.005 1	$\hat{a}_9$	-0.023 4
$\hat{a}_5$	0.050 0	$\hat{a}_{10}$	0.042 3

由表 2 可以看出, 提取的 LP 参数都小于 1。采用协方差斜格法降低了系统误差, 在转换到 LSP 参数后保证了所有根都在单位圆上。量化后的 LSF 参数满足相应滤波器稳定性的要求。

系统的时序仿真如图 1 所示。

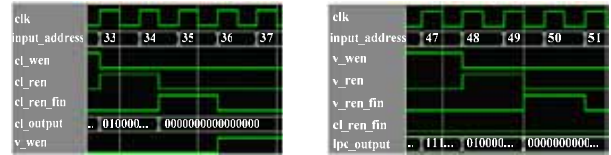


图 1 系统的时序仿真

图 1 系统的时序仿真

整个系统在 24 MHz 晶振作用下, 通过对读/写信号的操作, 得到最终结果。

为了验证结果的正确性, 在 Matlab 语言程序中计算各功能模块和系统。各功能模块和系统时序仿真结果表明, 仿真输出与在 Matlab 语言程序输出的结果完全相同。因此, 本文设计的 LPC 参数提取达到了预期目标。

## 7 结束语

本文研究了线性预测编码中声道特征参数提取的 FPGA 实现问题。在设计中将该算法与 FPGA 技术结合, 充分利用 FPGA 作为一种快速、高效的硬件平台在数字信号处理领域所具有的独特优势, 实现了语音信号声道特征参数的快速提取, 为语音信号的进一步研究、特别是在低速率编码领域的研究打下了基础。

## 参考文献

- [1] 普罗奇斯. 数字信号处理: 原理、算法与应用[M]. 3 版. 北京: 电子工业出版社, 2004.
- [2] 吴继华, 王 诚. Altera FPGA/CPLD 设计[M]. 北京: 人民邮电出版社, 2005.
- [3] ITU-T. ITU-T Recommendation G.729: Coding of Speech at 8kbits/s Using Conjugate-Structure Algebraic Code Excited Linear Prediction[Z]. 1996.
- [4] Thomas F Q. 离散时间语音信号处理: 原理与应用[M]. 北京: 电子工业出版社, 2004.
- [5] Mobini N, Vahdat B, Radfar M H. An FPGA Based Implementation of G.729. Circuits and Systems[C]//Proceedings of the IEEE International Symposium on Circuits and Systems. Kobe, Japan: IEEE Press, 2005.

(上接第 263 页)

进行审批。通过查看明细可以了解采购计划的详细信息, 符合实际则审批通过, 不符合实际可拒绝审批。

## 5 结束语

本文提出基于特征映射的自定义审批模型, 并将其应用于物资供应系统, 实践证明它是一个简单、实用、有效的柔性系统, 审批应用方案实现了审批流程、审批节点、审批角色和人员的自定义功能, 基本解决了企业组织、人员、审批流程频繁变化与信息系统相对不变的矛盾, 提高了企业审批工作的效率, 增加了信息系统的稳定性。

(上接第 268 页)

## 参考文献

- [1] 包起帆, 张文渊, 徐 盛, 等. 有源 RFID 的性能及其应用场景[J]. 微型电脑应用, 2006, 22(4): 1-3.
- [2] Culler D. TinyOS: Operating System Design for Wireless Sensor Networks[Z]. (2006-01-21). <http://www.sensormag.com/sensors/content/printContentPopup.jsp?id=324975#>.

## 参考文献

- [1] 何鹤立, 倪小平, 盛步云, 等. workflow管理系统的柔性技术[J]. 计算机工程, 2004, 30(6): 63-66.
- [2] 张 繁, 蔡家楣. 电子政务系统中动态 workflow技术的应用[J]. 计算机工程, 2003, 29(12): 72-75.
- [3] 何 青. 一个基于TBAC的审批业务 workflow模型[J]. 山东大学学报: 工学版, 2006, 36(4): 120-124.
- [4] 韩月娟, 赵 蕾, 吕 强, 等. 一个基于SOFM的柔性化审批应用方案的设计[J]. 计算机工程, 2006, 32(21): 275-277.
- [5] Ultra Low Power 125kHz 3D Wake-up Receiver with RSSI (ATA5282 Datasheet)[Z]. (2005-02-03). <http://www.atmel.com>.
- [6] 125kHz Transmitter IC for TPMS (ATA5275 Datasheet)[Z]. (2005-02-03). <http://www.atmel.com>.
- [7] LF Wake-up Demonstrator ATAK5275-83[Z]. (2005-02-03). <http://www.atmel.com>.