

摘要

在通信系统中，应尽可能少占用带宽、存储空间。语音编码你能够压缩编码速率，作为。LPC 语音编码技术是一种比较简单，实用的编码技术。DSP 作为一种特别适合于数字信号处理运算的处理器，在有关音视频通信方面应用中，能够快速实现音视频编解码算法。研究音频编解码算法在 DSP 上实现具有很强的实际意义。

本文在首先介绍了语音编码的背景、种类及评价标准基础上，引出文章的讨论点---LPC 音频压缩算法，并详细介绍其理论基础和使用的关键技术。深入研究了其编解码算法的具体步骤及 C 实现，最后在 DM6437 DSP 硬件平台上实现语音数据的编解码实现，分析解码之后语音数据失真的原因，并提出了改进措施。

关键字：语音编码，线性预测编码，特征参数，CCS，

ABSTRACT

The queue phenomenon in the telecom offices is a normal issue. To improve the customers' satisfaction and to support the company changing, we should solve this problem properly. The basic goal to resolve queue problem is the appropriate tradeoff between the customers' wavy demand and the telecom offices' limited service capability.

This paper is based on the queuing theory and demand management theory. And based on the data collection and customer survey and interview, the author uses some statistical methods to reflect the actuality. Then the author finds the reason of queuing in telecom office from customers' view. Furthermore, the author analyses the real demand of the customers by sorting them into types of paying and time and price sensitivity. To follow up, three solutions had been brought forward: firstly, distributing the customers; secondly, stopping phone by different number; lastly, promoting the demand during the non- fastidious.

Key Words: audio coding, Linear Prediction Coding, speech parameters,

目录

第 1 章	绪论.....	1
1.1	选题背景与意义.....	1
1.2	语音编码技术产生及发展.....	1
1.3	语音编码的原理.....	2
1.4	语音编码算法的分类及评价标准.....	3
1.4.1	语音编码算法分类.....	3
1.4.2	语音编码算法的评价标准.....	4
1.5	课题的重点、难点及关键.....	6
1.6	论文内容与结构.....	6
第 2 章	DAVINCI 硬件与软件系统.....	8
2.1	硬件系统.....	8
2.1.1	TI TMS320DM6437 芯片介绍.....	8
2.1.2	DM6437EVM 开发板.....	10
2.2	软件环境.....	10
第 3 章	语音信号的产生模型与 LPC 编码系统.....	15
3.1	LPC 编码算法概述.....	15
3.2	语音信号的产生模型.....	16
3.2.1	激励模型.....	16
3.2.2	声道模型.....	17
3.2.3	辐射模型.....	17
3.2.4	小结.....	18
3.3	LPC 编码系统.....	19
3.4	LPC 编码具体流程.....	20
3.4.1	低通滤波.....	20
3.4.2	加窗技术.....	21
3.4.3	线性预测分析.....	21
3.4.4	基音周期估计.....	24
3.4.5	增益 Gain 估计.....	24

3.5 LPC 编码相关子函数.....	25
第 4 章 LPC 解码系统.....	26
4.1 LPC 解码概述.....	26
4.2 LPC 解码具体流程.....	27
4.2.1 增益计算.....	27
4.2.2 清/浊音激励.....	27
4.2.3 反射系数转换成预测系数.....	27
4.3 LPC 解码相关子函数.....	28
第 5 章 调试结果与失真分析.....	29
5.1 LPC 编码调试.....	29
5.1.1 PCM 语音数据.....	31
5.1.2 LPC 编码得到的线性数据流.....	33
5.2 LPC 解码调试.....	35
5.3 失真分析.....	37
5.4 LPC 改进措施.....	37
第 6 章 工作总结与展望.....	39
6.1 论文期间完成的工作.....	39
6.2 工作展望.....	39
参考文献.....	41
致谢.....	43
附录.....	44
附录一：LPC 编解码的主程序源代码.....	44
附录二：顾客调查问卷.....	44
附录三：仿真分号停机策略的 MATLAB 程序命令.....	44
附录四：分号停机策略实施前后日缴费顾客数量对比表.错误！未定义书签。	
外文资料原文.....	45
翻译文稿.....	49

第1章 绪论

1.1 选题背景与意义

我们知道，用数字方法来记录和传输信息要比用模拟方法好。因为数字信号具有抗噪声、增强音频动态范围、以及多次翻录没有信号衰减等优势。但数字声音信号要占用一个很宽的传输带宽，特别是高保真临场感的音响信号经数字化后的数字信息量非常大。因此，为了更加有效地利用宝贵的信道资源，必须对声音进行数字音频压缩。数字音频压缩技术的是对原始数字音频信号流(PCM 码流)运用适当的数字信号处理技术，在不损失有用信息量，或所引入损失可忽略的条件下，降低(压缩)其码率，称为压缩或编码。它必须具有相应的逆变换，称为解压缩或解码。

1.2 语音编码技术产生及发展

上个世纪末的战争时代[#]，由于军事保密通信的需要，贝尔电话实验室的 HomerDudley 提出并实现了在低带宽电话电报电缆上传输语音信号的通道声码器，成为语音编码技术的鼻祖。该声码器在第二次世界大战中的有效使用，为声码器技术队的应用开创了先河。此后，由于带宽和传输干扰问题急待解决，推动了数字通信网的发展，同时也推动了话音数字化编码技术的迅速发展。60 年代 CCITT（现 ITU）指定了第一个语音数字化编码标准，即 A 率或者 μ 率 PCM 编码的 G.711 标准。它将话音信号按 Nyquist 准则抽样，并量化为 64kbit/s 传输速率的数字信号。但这一简单的编码方案没有对语音信号的冗余信息做任何处理，它产生的高比特率数据占用了巨大的信道带宽和存储空间。此后，压缩编码技术得到了迅速发展，并形成了以波形匹配为目标的波形编码和以追求人的感知效果为目标的参数编码两大体系。

进入 80 年代以来，光纤传输技术的引入，为有线通信提供了巨大的信道容量，使信道带宽问题似乎得到了解决。但同时，人们对带宽有限的无线通信，信道价格昂贵的卫星通信和军用保密通信的需求不断增加。特别是智能网的提出，各种与语音应用服务相关的新业务的不断涌现，要求语音数据能够被灵活的处理，存

储, 转发和传送, 从而刺激并推动了应用与移动通信, 卫星通信, 多媒体通信, 智能网和保密通信领域的中低速率 (32kbit/s, 16kbit/s 和 8kbit/s), 甚至极低速率 (小于 2.4kbit/s) 语音编码技术的发展。到 80 年代末以后, 针对各领域的语音编码技术相继成熟, 并不断得到完善。

语音压缩编码的意义在于去处语音信息中的冗余度, 降低传输比特率或者存储空间。这就要求我们对原始语音信号进行运算处理。压缩比特率越低, 运算复杂度就越大, 然而, 语音通信往往要求同时进行, 因此, 这些算法必须在很短时间内完成。80 年代以后, 随着 VLSI 技术的发展, 专用高速 DSP 芯片的产生, 特别是通用 DSP 处理器指令效率和运算速度的不断提高, 使压缩编码算法的实时处理成为可能。也正是依托于高速 DSP 器件的硬件基础, 推动着压缩编码技术的应用进一步发展。

数字音频压缩技术称得上是目前应用最为广泛的数字技术之一, 是多媒体技术研究中的热点, 受到广播界、科技界和工业界的广泛关注, 发达国家虽然起步较早, 并已取得了一定经验, 但就目前的压缩技术来说仍不能满足市场的要求。在当前数字音频压缩技术中, 比较成功的编码是一种利用人耳心理声学特性的编码——“感知型编码(Perceptual Coding)”。它模仿了人耳对信号幅度、频率、时间的有限分辨能力, 即凡是对人耳辨别声音信号的强度、音调、方位没有贡献的部分(称为不相关部分或无关部分)不编码也不传送;对能感觉到的部分进行编码时, 也允许有较大的量化失真, 并使其处于人耳仍然感觉不到的听阈以下。简单的说感知型编码只记录那些能被人的听觉所感知的声音信号以减少数据量。

1.3 语音编码的原理

语音信号能够进行压缩主要有两个方面的原因, 第一个是语音信号本身存在冗余度, 第二则是人耳具有听觉效应^[1]。

1.1. 消除冗余度实现压缩

从信息保持的角度讲, 只有当信源本身具有冗余度, 才能对其进行压缩。根据统计分析的结果, 语音信号中存在着多种冗余度, 主要有以下几个方面:

1). 幅度非均匀分布

语音中小幅度样本比大幅度样本出现的概率要高。又由于通话必然会有间隙, 更出现了的低电平样本。此外, 实际语音信号的功率电平也趋向于出现在编码范

围的较低电平端。因此，语音信号的幅值分布是非均匀的。

2). 样本之间的相关性

对语音波形分析表明，采样数据的最大相关性存在于邻近样本之间。当采样速率为 8KHz 时，相邻采样值之间的相关系数大于 0.85，甚至在相距十个样本之间，相关系数还有 0.3 左右的数量级。如果采样速率提高，样本间的相关性将更强。

3). 周期之间的相关性

虽然电话语音信号的频率分布在 300-3400Hz 的频带内，但在特定的瞬间，某一声音却往往只是该频带内的少数频率分 I 起作用。当声音中只存在少数几个基本频率时，就会像某些振荡波形一样，在周期与周期之间存在着一定的相关性。

4). 基音之间的相关性

语音可以分为清音和浊音两种基本类型。浊音是由声带振动产生，每一次振动使一股空气从肺部流进声道。激励声道的各股空气之间的间隔称为基音周期。浊音的波形对应于基音周期的长期重复波形。因此，对语音浊音部分编码的有效方法之一是对一个基音周期波形进行编码，并以它作为其它基音段的模板。此外，电话通信中还有很大的话音间隙。

2. 利用人耳听觉效应

语音压缩编码的第二个途径是利用人耳的听觉特性。人类听觉有一个特点，就是“听觉屏蔽效应”，一个强的音能够抑制一个同时存在的弱音的听觉，利用这一性质，可以抑制与信号同时存在的量化噪声，如“噪声谱形变技术”。另外，人的听觉对语音的低频成分比较敏感，而对高频部分则不太敏感，因此引出了“子带编码技术”，；还有人的听觉对信号的相位特征不敏感，线性预测声码器利用了这一特点，不传送语音谱的相位信息，使码率能够降至 2.4kb/s 以下，仍能保持高的可懂度；后滤波技术则是利用幅度谱的适度失真来降低量化噪声对语音质量的影响。

1.4 语音编码算法的分类及评价标准

1.4.1 语音编码算法分类

实现语音压缩编码有两条途径#：波形编码和参数编码，语音编码一直沿着这两个方向发展。现在则是以综合两者的优点进行混合编码的方法居主导地位。

波形编码力图使合成语音与原始语音的波形误差最小，由于语音信号的全部信息都蕴含在原始波形里，所以这种方法编码后的合成语音质量非常好，且适应能力强，抗信道干扰性能好。所采用的压缩方法一般是基于各种有效的数学变换，通过将波形从一个域变换为另一个更易于提取特征参数的域来达到对变换后的参数进行量化编码的目的，在数学上，这实质上是一个曲线拟合或数据近似的问题。主要的有以下的编码器：(1)脉冲编码调制((PCM), (2)自适应增量调制(ADM), (3)自适应差分编码(ADPCM), (4)自适应预测编码(APC) (5)自适应子带编码(ASBC), (6)自适应变换编码(ATC)。波形编码的性能和压缩比特率决定于所用的变换方法的性能，由于语音波形的动态范围很大，目前所用的变换算子的作用又有限，因此，波形编码的比特率不能压得很低，一般在 16kbps 以上，再往下，性能就下降很快。新近蓬勃发展的小波变换，尽管具有分层的思想，“显微镜”的功能，与图象的视觉感知相吻合，但用于语音编码效果不理想，因为与人的听觉感知不相吻合。

参数编码通过对人的发声的生理过程的研究，建立一个模拟其发声的数字模型来达到提取其特征参数进行量化编码的目的，它力图使合成语音具有尽可能的懂性，保持原语音的语意，而合成语音的波形与原始语音的波形可能有相当大的差别。这种方法可以将比特率压得很低，但是质量不是很好。

1.4.2 语音编码算法的评价标准

语音编码的根本目标就是在尽可能低的编码速率下，得到尽可能高的重构语音合成质量，同时还应尽量减小编解码延时和算法复杂度。因此编码速率、合成语音质量、编解码延时以及算法复杂度这四个因素构成了评价一个语音编码算法性能的基本指标。这四个因素之间有着密切的联系，并且在不同的应用中对各方面的侧重要求也有所不同，在具体评价一种语音编码算法优劣时，需要根据具体的实际情况，综合考虑这四个因素#。

1.编码速率

编码速率直接反映了语音编码器对语音信息的压缩程度，可以用比特/秒来度量，它取决于用户对语音质量的要求，同时也受信道容量和存储介质的限制，编码速率越低，越有利于存储和传输，但相对应的编码质量也会有所下降。

2.编解码延时

编解码延时包括算法延时(语音分析所需的存储缓冲)以及计算延时(处理帧内数据所需的时间)。在实时语音通信系统中，语音编解码延时同线路传输延时一样，

对系统的通话质量有很大影响。延时过长会使通话双方都反应“迟钝”，严重时无法正常通话，延时影响通话质量的另一个可能的原因是回声干扰，传输线路如果阻抗匹配不理想会使信号反馈形成回声，当延时比较小时，回声同话机侧音及房间交混回响声相混，因而感觉不到，若往返总延时超过 100ms 时，发送端就能听到自己的回声，如果回声传输路径衰耗不够大，就会听到多次回声，从而严重影响通话质量。

3. 算法复杂度

算法复杂度同语音编码的话音质量有非常密切的关系。在同样码速率的情况下，采用复杂一些的算法将会获得更好一些的话音质量。同样，对于相同的话音质量，采用复杂一些的算法能够降低编码所需的速率。算法复杂度包括两个方面，运算复杂度和内存要求，它们影响算法在硬件上的实现，并决定了硬件实现的复杂程度、体积、功耗以及成本等。对于一些较复杂的语音编码算法，一般用每秒钟处理信号样本所需的数字信号处理指令条数来衡量其运算复杂度，可用单位百万次操作/秒(MOPS)或百万条指令/秒(MIPS)等来对算法复杂度进行描述，内存用 byte, Kbytes 等来衡量。

4. 合成语音质量

合成语音质量是衡量语音编码算法性能优劣的关键因素之一，评价合成语音的方法大致分为两类，即客观评价法和主观评价法。客观评价法是用客观测量的方法来评价合成语音质量，它建立在原始语音和合成语音的数学对比之上，常用的方法有信噪比、加权信噪比、平均分段信噪比、巴克谱失真测度、MEL 谱测度等。这些算法的特点是计算简单、结果客观，但是不能完全反映人对语音的感觉效果，对于速率低于 16kb/s 的中低速率语音编码尤为突出，所以此方法主要适用于速率较高的波形编码类型的算法。

主观评价法是在一组测试者对原始语音和合成语音进行对比试听的基础上，根据某种预先约定的尺度来对失真语音划分质量等级。常用的主观评价方法有三种：平均意见得分(MOS)、判断韵字测试(DRT)和判断满意度测量(DAM)。最常用的主观评价方法是 MOS 评分法，它采用五级评分标准，由数十名试听者在相同信道环境中试听合成语音并给予评分，然后对评分进行统计处理，求出平均得分。MOS 评分等级见表 1-1 所示。

表 1-1 MOS 评分五级标准及相应的描述该级语音质量的形容词

MOS 评分	质量级别	失真级别
--------	------	------

5	优	不察觉
4	良	刚有察觉
3	可	有察觉且稍觉可厌
2	差	明显察觉且可厌但可忍受
1	坏	不可忍受

在数字语音通信中，通常认为 MOS 评分在 4.0~4.5 分时为高质量数字化语音，达到了长途电话网的要求，接近于透明信道编码，也常称之为网络质量，MOS 评分在 3.5 分左右时称作通信质量，这时能感到重建语音质量有所下降，但不妨碍正常通话，可以满足多数话音系统的使用要求，MOS 评分在 3.0 分以下时，常称作合成语音质量，这是指一些声码器合成的语音所能达到的质量，它一般具有足够的可懂度，但自然度及讲话人的确认等方面不够好。

1.5 课题的重点、难点及关键

本课题的重难点在于选定一种语音编码算法，研究其详细的编解码过程，并用 C 语言实现。难点在于需要代码表示复杂的公式及运算过程。在 CCS 的平台下，编写好主程序，调试、链接通过生成 out 文件下载到 DSP 中，运行并实现语音数据编解码。

1.6 论文内容与结构

本文是研究一种音频编解码方案在 DSP 上实现，选定了 LPC 编解码算法，并对此进行了详细的分析和研究，论文结构如下：

第一章为绪论，重点介绍论文的选题背景意义论文的内容结构。

第二章为音频编码的系统性介绍，分类方法及发展及论文的研究课题—LPC。

第三章详细分析了 LPC 的关键技术，编码过程的具体步骤及 C 实现

第四章详细分析了 LPC 解码过程具体步骤及 C 实现

第五章介绍了本次实验的硬件平台及仿真结果

第六章对论文期间完成的工作进行总结，并指出进一步改进和研究方向。

第2章 DAVINCI 硬件与软件系统

本章主要介绍了此次实验的硬件与软件环境。硬件系统是 TI 公司的 DAVINCI 系列系列产品---TMS320DM6437；软件系统是 DSP 的开发环境---CCStudio。

2.1 硬件系统

本课题的编解码系统实现采用了 TI TMS320DM6437 芯片,并选用了 Spectrum Digital 公司生产的 DM6437 EVM 开发板进行调试。

2.1.1 TI TMS320DM6437 芯片介绍

TMS320DM6437 是一款内核为#

C64x+。DaVinci Video 400 500 600MHz 只支持小端方式

L1 SRAM 112k (32K 程序,80k 数据)

L2 SRAM 128k

ROM 64k

一个 8 位 EMIFA

一个 16/32 位 DDR2

EDMA 64 通道

Video Ports 1 input 1 output

串口: McASP IIC 1 个 HECC 1 个 McBSP 2 个 UART

接口: 32 位 PCI, VLYNQ, 10/100EMAC, 16 位 HPI

程序数据存储: Async SRAM,DDR2 SDRAM,NAND Flash

程序数据存储: Async SRAM,DDR2 SDRAM,NAND Flash



下图所示为 6437 的各个外设接口



图 3-2 DM6437EVM 的各模块图

2.1.2 DM6437EVM 开发板

DM6437 EVM 开发板的结构框图如下图所示

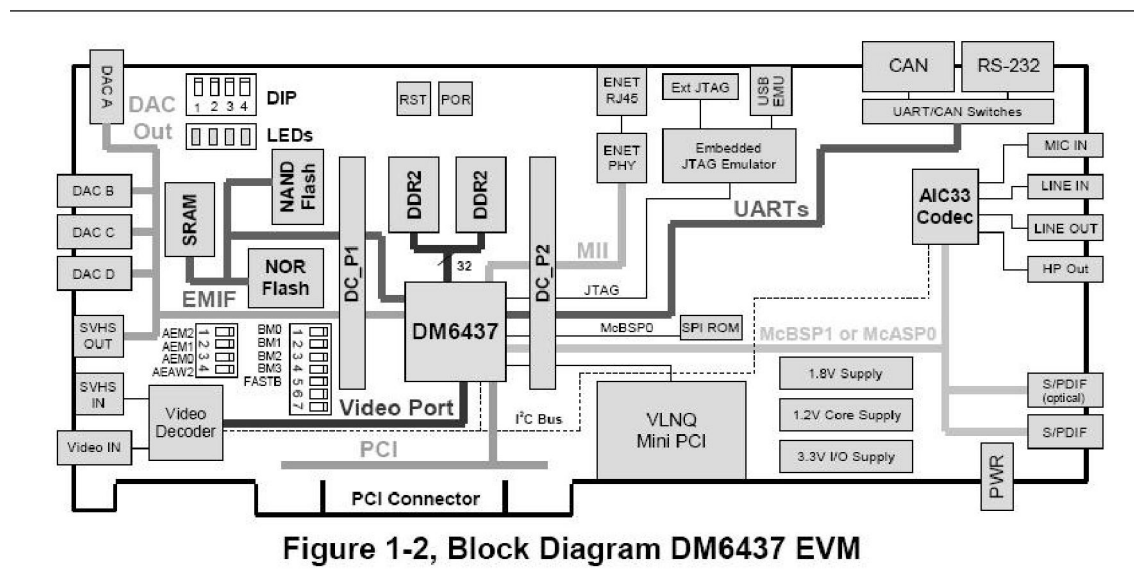


图 3-3 DM6437 EVM 开发板结构图

2.2 软件环境

本课题使用的开发环境是 Code Composer Studio，是一个完整的 DSP 集成环境。根据本次实验的 DM6437 EVM 开发版安装合适的驱动程序，使用的是硬件仿真

CCS 具有数据的图形显示根据，可绘制时域/频域波形，而用 Animate 可以实现自动刷新#。

CCS 的界面如下图所示

第2章 研究的理论基础

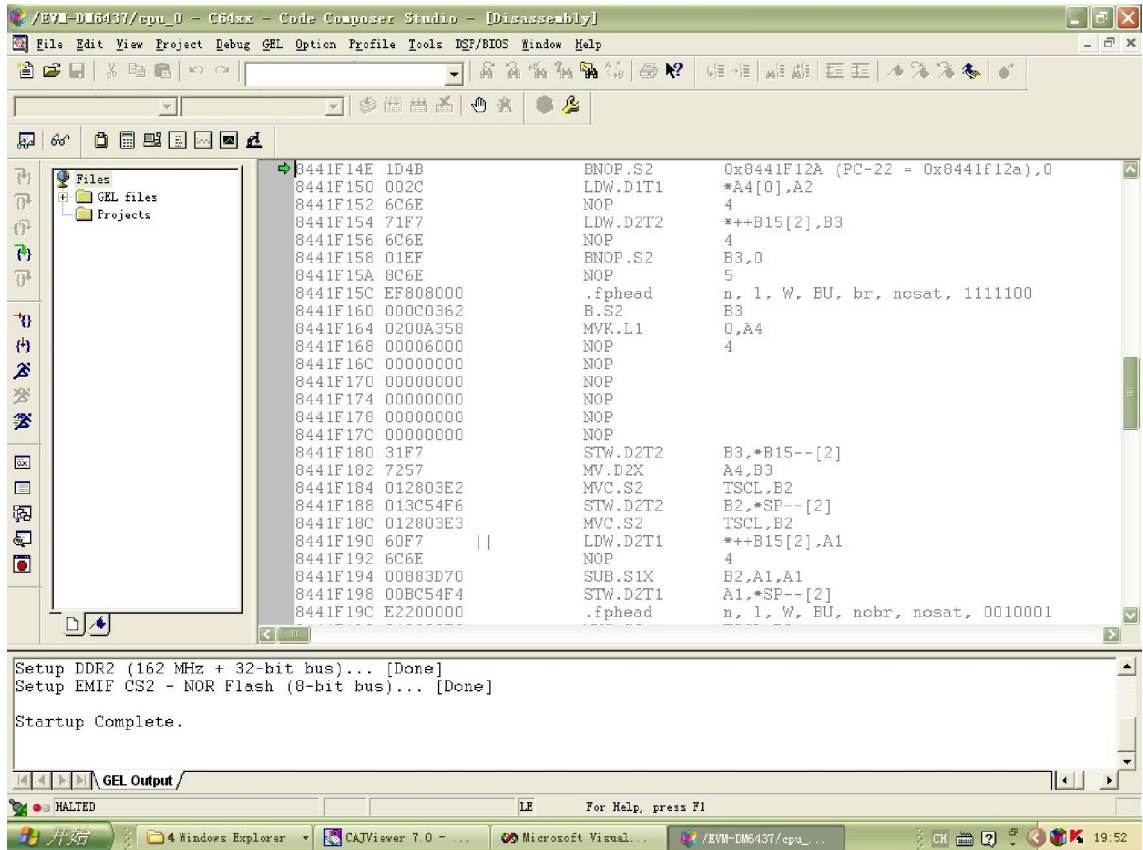


图 CCS 的界面

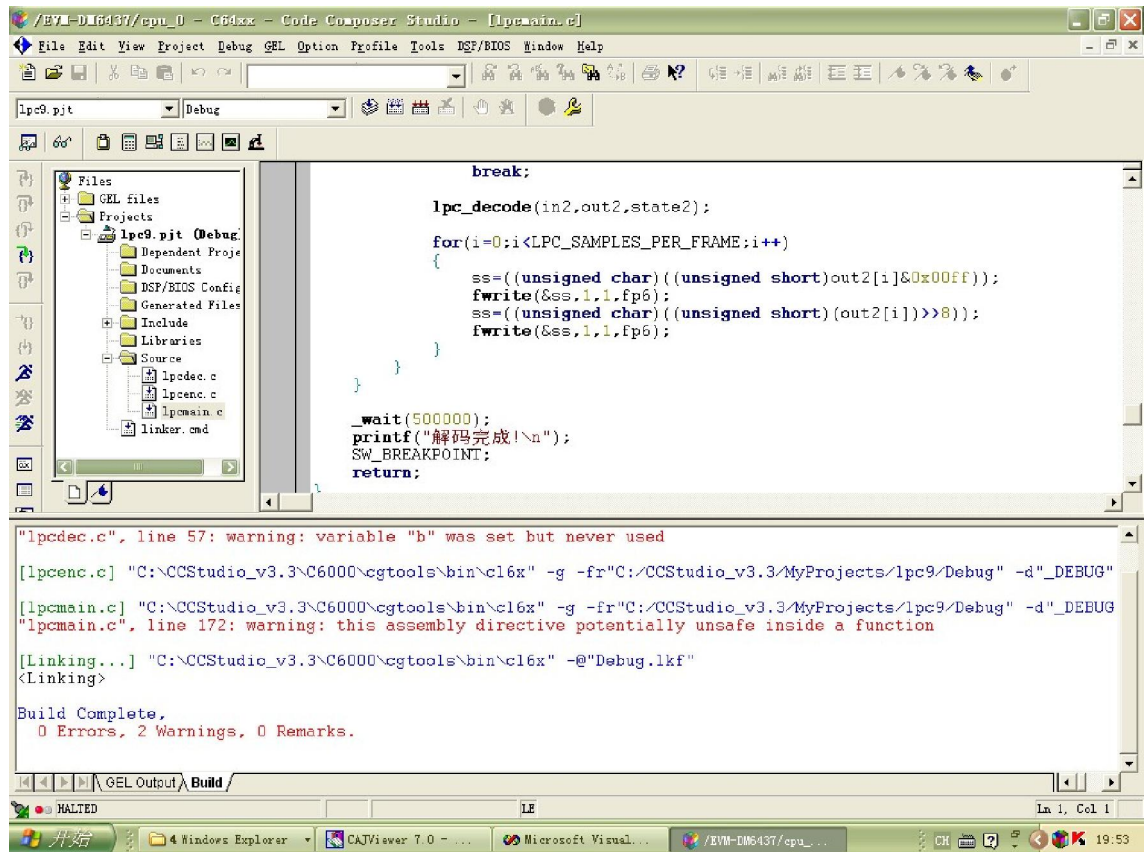


图 CCS 的开发环境及调试

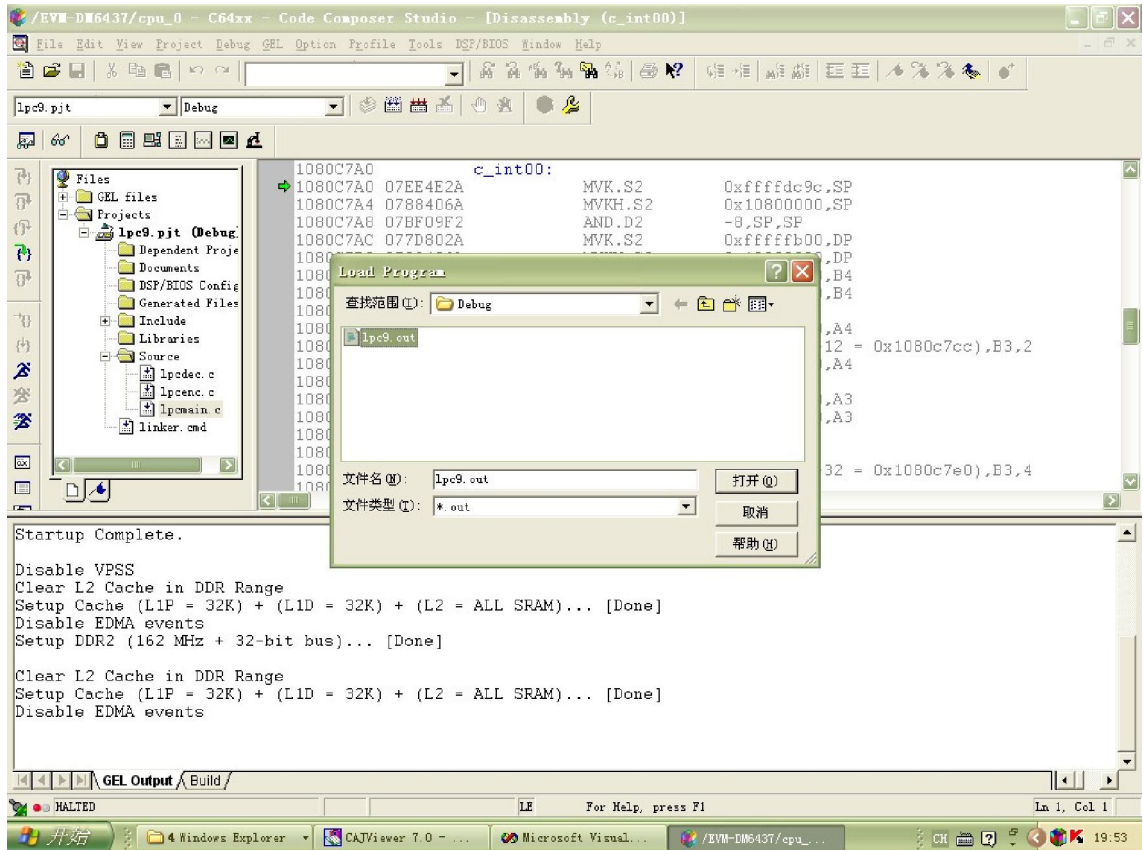
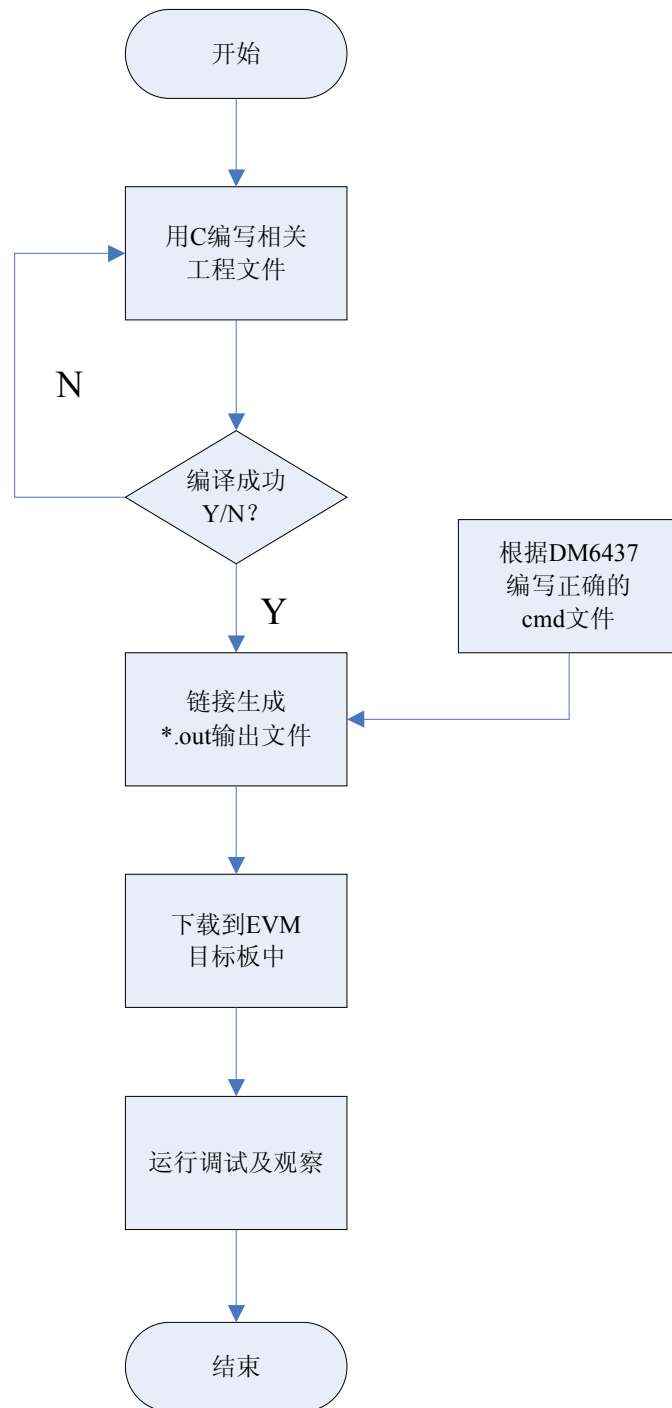


图 将生成的文件下载到 DSP 中

建立工程文件，编写 cmd 内存定位，最后在 CCS 的编译器链接生成 out 的文件，通过 JTAG 下载到 DSP 芯片中，
软件流程如下图所示



第3章 语音信号的产生模型与 LPC 编码系统

线性预测编码属于参数编码，而参数编码的基础是语音的产生模型。根据语音信号产生的数字模型对语音信号进行分析，可以得到与语音信号描述的相关信息以便编码传输。因此本章首先给出了 LPC 语音编码算法的概述，之后介绍了语音信号的产生模型及 LPC 编码的工作原理，并给出了流程及软件实现。

3.1 LPC 编码算法概述

本文选定了 LPC(Linear Prediction Coding)，即线性预测编码作为研究课题。

线性预测编码(LPC)#，属于参数编码，是基于分析和模拟人的发音器官，建立语音信号产生模型。在编码过程中，从人的语音信号中提取与语音模型有关的特征参数，将参数编码传输；在译码端中，得到的编码数据通过相应的数学模型计算，得到相应的控制参数，来合成语音。

利用这种方法能对语音信息进行大幅度的压缩，减少语音数据所占用的存储空间。LPC 分析方法已成了估计语音基本参数的主导技术之一，研究用最少存储数据来表征语音的特征参数，如音调、共振峰、频谱、发音域函数等。

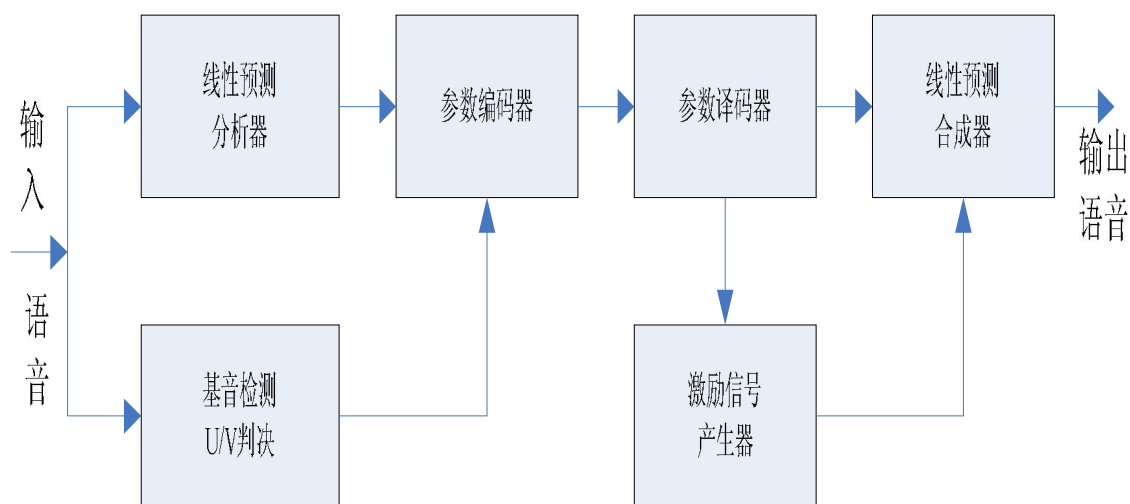


图 LPC 编解码工作原理

3.2 语音信号的产生模型

语音是由肺部的气流冲击声道，最后从嘴唇或鼻孔或同时从嘴唇和鼻孔辐射出来而形成的。语音信号的产生模型就是描述这种语音形成过程的。它由激励模型、声道模型和辐射模型组成，如图 3-1 所示#。

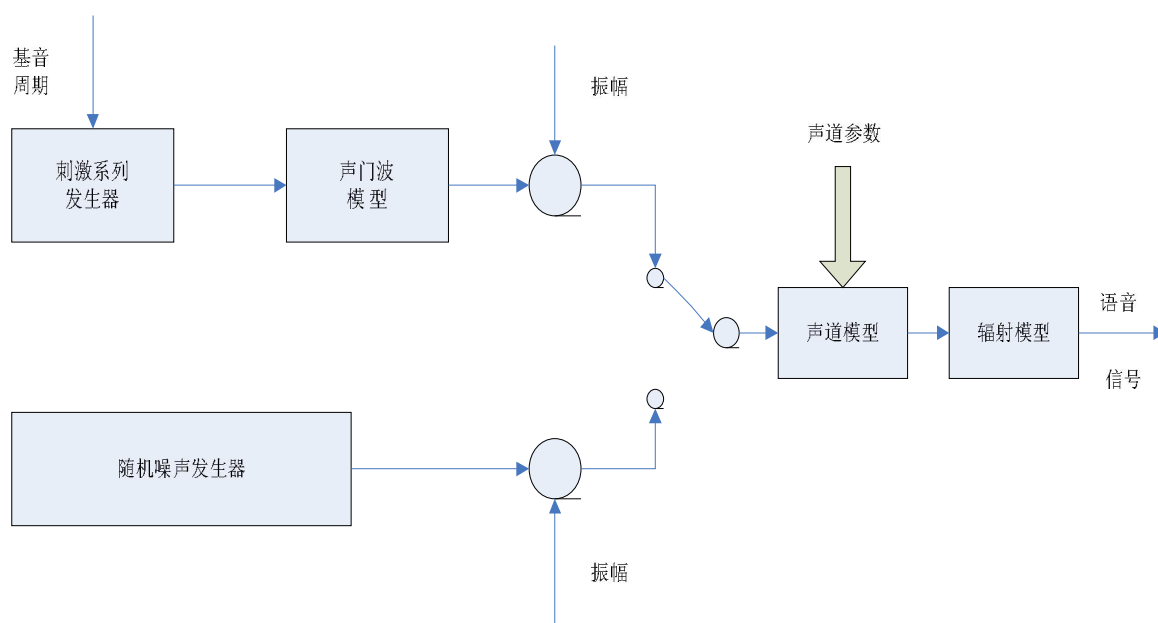


图 3-1 语音信号产生的数字模型

3.2.1 激励模型

人发出不同的音时，激励的情况也不同。简化分析可以把发音情况大致分为清音和浊音。发清音时，声带处于松弛状态，不发生振动，气流通过声门直接进入声道。激励信号类似于随机白噪声；发浊音时，气流是通过绷紧的声带，对声道进行冲击，产生振动，使声门处形成准周期的脉冲串，即类似于受周期脉冲激励。

3.2.2 声道模型

目前对声道模型有两种看法，一是视声道为多个不同截面积的声管串联而成的系统，可以称之为声管系统；二是视声道为一个谐振腔，共振峰就是这个腔体的谐振频率，称之为共振峰模型。本论文选定了共振峰模型。

共振峰的特性可以用一个全极点模型来描述，如公式（3.1）所示

$$V(z) = \frac{1}{\sum_{i=1}^p a_i z^{-i}} \quad (3.1)$$

式中： p ——全极点滤波器阶数，一般取值范围是8~12；

a_i ——声道模型参数。

3.2.3 辐射模型

声道的终端出口是口唇。从声道输出的是速度波，而语音信号是声压波，二者的倒比即为辐射阻抗，可以用来描述口唇的辐射效应及头部的绕射效应，表达式如下

$$z(\Omega) = \frac{j\Omega LR}{R + j\Omega L} \quad (3.2)$$

式中： $R = \frac{128}{9\pi^2}$ ， $L = \frac{8a}{3\pi c}$ ，此处 a 是口唇张开时的开口半径， c 是声波的传输速度。

从公式（3.2）可以看出，能量损耗正比于辐射阻抗的实部，辐射效应在低频段不明显，而在高频道较明显，因此可以用一个高通滤波器来表示辐射模型，如式（3.3）所示：

$$R(z)=1-rz^{-1} \quad (3.3)$$

式中： r 是描述辐射效应的一个系数，数值大小近似于 1。

3.2.4 小结

根据上面所述可见，语音信号的产生的完整模型，可以用三个子模型串联起来表示，并可以得出其传递函数为：

$$H(z)=U(z)V(z)R(z) \quad (3.4)$$

由式（3.4）中可以看出，通过分析出三个

由此我们得到了语音信号产生的数字模型，在本课题中，为了便于线性预测分析，可以将产生模型简化，其如下图所示#

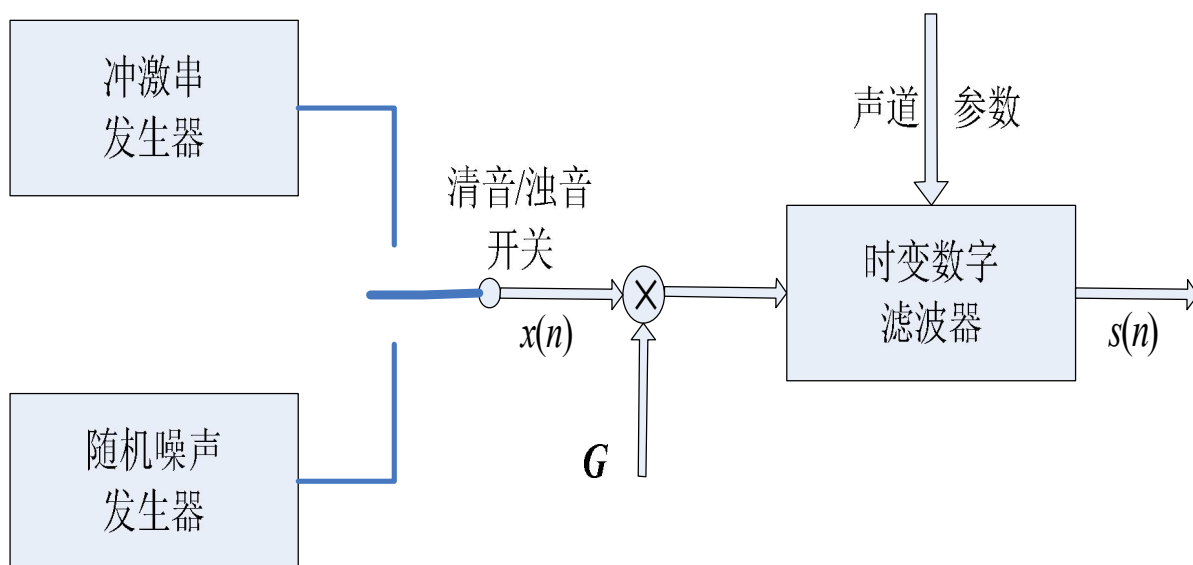


图 简化的语音产生模型

实际上，在线性预测分析中，传递函数 $H(z)$ 可以求得：

$$H(z) = \frac{1}{G} \frac{S(z)}{E(z)} = \frac{1 - \sum_{j=1}^M b_j z^{-j}}{1 - \sum_{i=1}^p a_i z^{-i}} \quad (3.5)$$

可以简单地用一个全极点模型表示：

$$H(z) = \frac{G}{1 - \sum_{i=1}^p a_i z^{-i}} \quad (3.6)$$

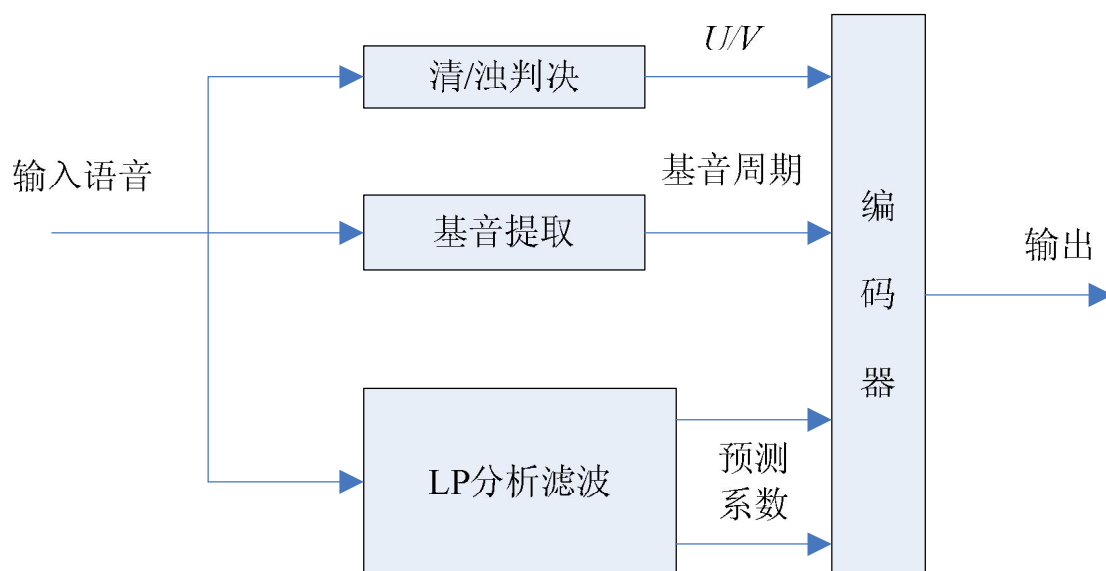
在时间域上#，

$$s(n) = Gx(n) + \sum_{i=1}^p a_i s(n-i) \quad (3.7)$$

以上即为 LPC 编解码进行的基础。

3.3 LPC 编码系统

LPC，采用了线性预测分析合成原理，对模型参数和激励参数进行编码传输，其编码工作原理如下图所示#。



根据上一小节对语音信号产生模型的分析，可以得到语音的谱包络、基音周期以及清浊音判决等参数。对得到的参数进行编码和传输，在译码端，进行对应的操作，就可以解码出语音信号。

3.4 LPC 编码具体流程

语音是由肺部的气流冲击声道，最后从嘴唇或鼻孔或同时从嘴唇和鼻孔辐射
LPC 编码的关键是获得相关的语音参数

3.4.1 低通滤波

原始语音经过一个锐截止的低通滤波器，以将输入语音限制在 100~3600Hz 的频率范围内

在本程序中，对使用编码 `lpc_encoder_state *st` 使用子函数 `init_lpc_encoder_state(lpc_encoder_state *st)` 进行运算即可完成。

3.4.2 加窗技术

语音信号具有短时平稳特性，因此将语音分成一帧一帧来进行分析。分帧采用的是用一段可移动的有限长度窗口来对语音信号产生加权实现。加窗之后的语音信号的表达式为

$$s_w(n) = s(n)w(n) \quad (3.8)$$

本课题使用的窗函数为 Hamming 窗，其表达式为

$$w(n) = \begin{cases} 0.54 - 0.46 \cos\left(\frac{2n}{N-1}\pi\right) & 0 \leq n \leq N-1 \\ 0 & \text{others} \end{cases} \quad (3.9)$$

Hamming 窗具有。在这里 $N=160$

3.4.3 线性预测分析

构成了一个线性系统方程组，高效的方法使用 Levinson-Durbin 递推算
法求得 LP 系数，

上一小节加窗之后的语音，其自相关函数为

$$r(k) = \sum_{n=k}^{239} s_w(n)s_w(n-k) \quad (3.10)$$

线性预测的关键是求解下列线性方程组#

$$\begin{cases} r(j) - \sum_{i=1}^p a_i r(j-i) = 0, & 1 \leq j \leq p \\ r(0) - \sum_{i=1}^p a_i r(i) = E_p \end{cases} \quad (3.11)$$

其运算步骤如下#

$$r_w(l) = \sum_{n=0}^{N-l-1} s_w(n)s_w(n+1), 0 \leq l \leq p \quad (3.12)$$

$$E_0 = r_w(0) \quad (3.13)$$

$$k_1 = \frac{r_w(l)}{r_w(0)} \quad (3.14)$$

$$a_1^{(l)} = k_1 \quad (3.15)$$

$$E_1 = (1 - k_1^2) E_0 \quad (3.16)$$

$$i = 2 \quad (3.17)$$

$$k_i = \frac{r_w(i) - \sum_{j=1}^{i-1} a_j^{(i-l)} r_w(i-j)}{E_{i-l}} \quad (3.18)$$

$$a_i^{(l)} = k_l \quad (3.19)$$

$$a_j^{(i)} = a_j^{(i-l)} - k_i a_{i-j}^{(i-l)} \quad (3.20)$$

$$E_i = (1 - k_i^2) E_{i-l} \quad (3.21)$$

$$m < p? \quad (3.22)$$

若为真，则令 $i=i+1$ ，并转入分成（3.19）执行；

若为假，则所求的结果为 $a_i = a_i^{(p)} \quad 1 \leq i \leq p$ 。

利用递推算法求得了预测系数之后

在实际情况下，线性预测滤波器的预测系数的微小变化都会造成滤波器极点位置很大变化，因此不适于直接进行量化编码和传输#。线性预测分析法求得是一个全极点的系统函数，而一个递归数字滤波器可以采取不同的结构实现，通过推演参数，可以得到其他有不同的在编程上使用了

线性预测系数 $\{a_i\}$ 可以通过一下递推方法求得反射系数 $\{k_i\}$ #，

初值设定

$$a_j^{(p)} = a_j \quad 1 \leq j \leq p \quad (3.23)$$

$$k_i = a_j^{(i)} \quad (3.24)$$

$$a_j^{(i-1)} = a_j^{(i)} + \frac{a_j^{(i)} a_{(i-j)}^{(i)}}{1 - k_i^2} \quad (3.25)$$

反射系数的取值范围为 $-1 \leq k_i \leq 1$ ，是保证相应的系统函数稳定的充分必要条件。

3.4.4 基音周期估计

基音周期分析也是在语音短时处理技术的处理的。选取了基于采取短时自相关函数的音周期语音估计。

3.4.5 增益 Gain 估计

增益 G 的均方根值 RMS 可以通过下面公式计算

$$RMS = \sqrt{\frac{1}{M} \sum_{i=1}^M S_i^2} \quad (3.26)$$

式中： S_i -----经过预加重的数字语音；

M -----分析帧的长度，在本程序中为 160。

3.5 LPC 编码相关子函数

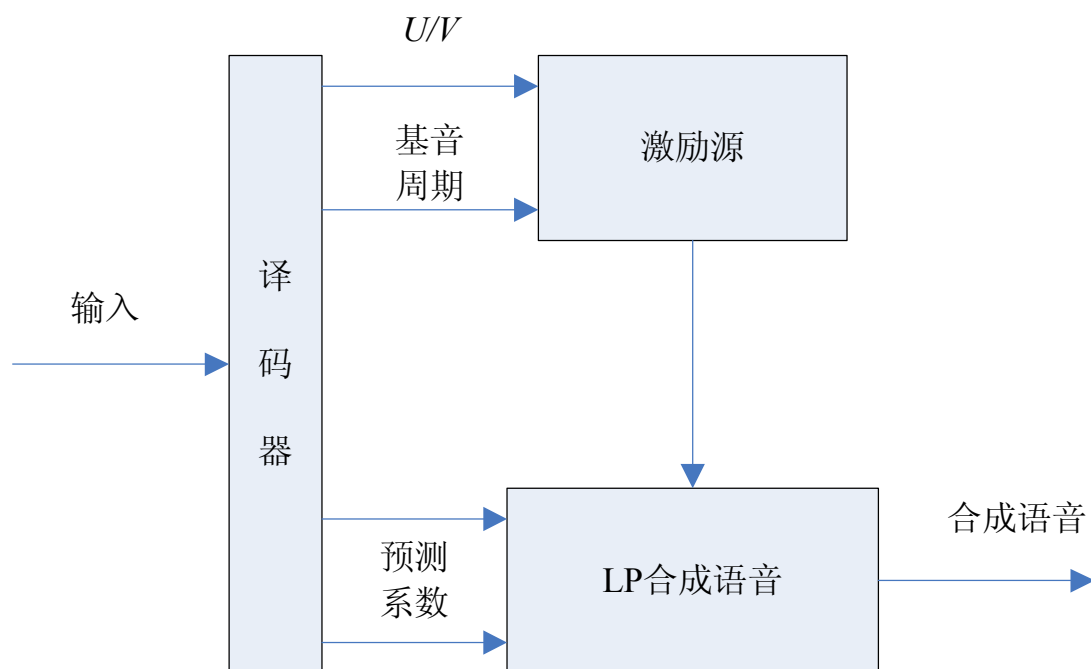
语音是由肺部的气流冲击声道，最后从嘴唇或鼻孔或同时从嘴唇和鼻孔辐射出来而形成的。语音信号的产生模型就是描述这种语音形成过程的。它由激励模型、声道模型和辐射模型组成，如图 3-1 所示#。

第4章 LPC 解码系统

上一章 LPC 编码之后会得到一系列描述语音信号的参数。LPC 解码系统将得到的参数在相同的语音信号产生模型上合成原始语音信号。本章将讲述 LPC 分析解码系统。

4.1 LPC 解码概述

LPC 编码端通过 LP 分析提取了声道滤波器、清浊音和基音相关参数，LPC 解码端得到这些参数，再重新合成语音信号，其原理框图如图所示



由图可以得知，对于解码端，预测系数是用于

4.2 LPC 解码具体流程

由上节的语音是由肺部的气流冲击声道，最后从嘴唇或鼻孔或同时从嘴唇和鼻孔辐射出来而形成的

4.2.1 增益计算

语音是由肺部的气流冲击声道，最后从嘴唇或鼻孔或同时从嘴唇和鼻孔辐射出来而形成的。

4.2.2 清/浊音激励

使用了一个简单的清、浊音判决器，对每一帧语音只进行一次清、浊判决。若为清音，则以白噪声作为该帧的激励信号；若为浊音，则以基音周期间隔的脉冲系列作为该帧的激励信号。

关键所在即为根据得到的信息对该帧语音进行清浊判决，

4.2.3 反射系数转换成预测系数

在解码端得到的是反射系数，由反射系数求出线性预测系数步骤如下#

$$a_j^{(i)} = k_i \quad (3.27)$$

$$a_j^{(i)} = a_j^{(i-1)} + k_i a_{i-j}^{(i-1)} \quad 1 \leq j \leq i-1 \quad (3.28)$$

最终可得

$$a_j = a_j^p \quad 1 \leq j \leq p \quad (3.29)$$

4.3 LPC 解码相关子函数

语音是由肺部的气流冲击声道，最后从嘴唇或鼻孔或同时从嘴唇和鼻孔辐射出来而形成的。

第5章 调试结果与失真分析

本章首先给出了 LPC 编解码算法的调试结果,之后对解码之后的波形存在较大的失真现象分析了原因,并提出了改进措施。

5.1 LPC 编码调试

此次实验的 LPC 编码调试是将 PCM 语音数据编码成二进制比特流。LPC 编解码系统的主程序流程如下图 5-1 所示

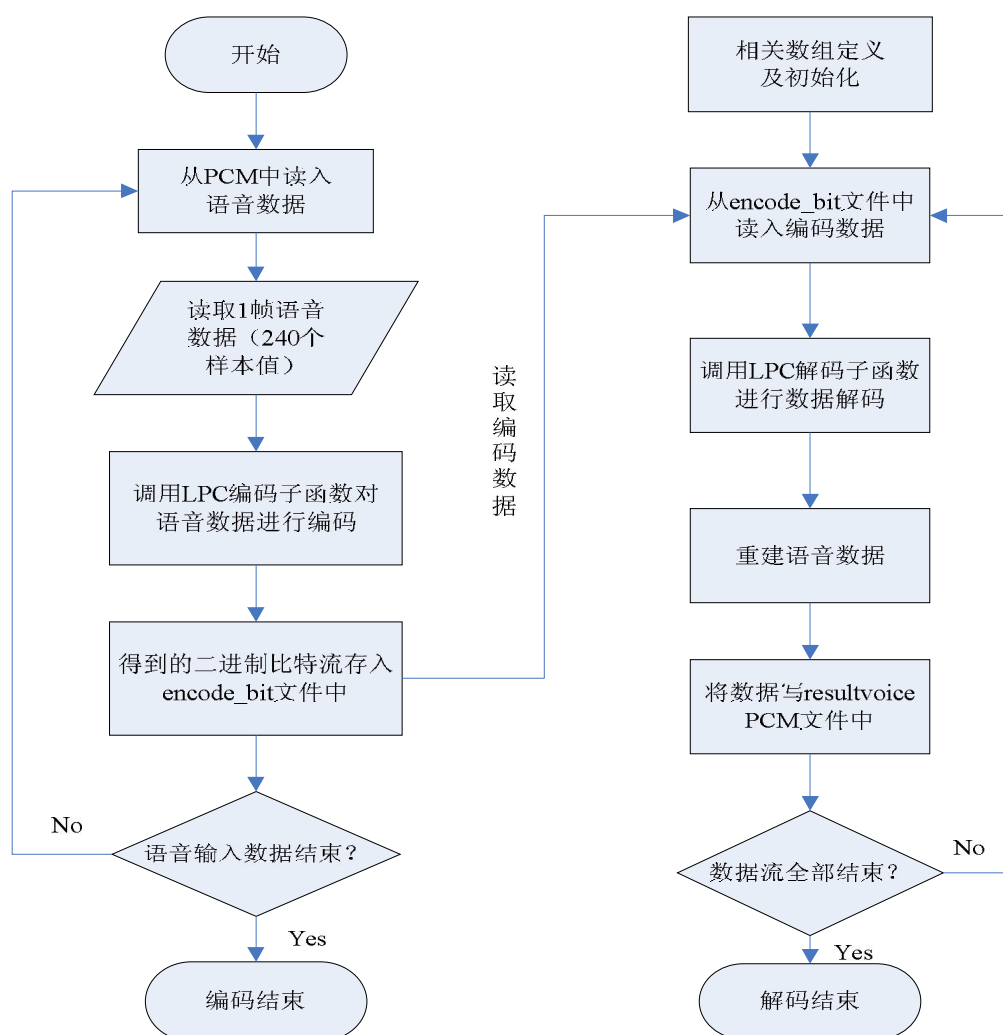


图 主程序流程图

实验的原始语音数据是用 Cool Edit Pro 2.0 在电脑上的录音，采样率为 8kb/s，单声道，采样精度（即量化比特数为 16bits）。语音信号为中文慢速状态下说的“语音编码算法”，其时域波形如图 5-1 所示。

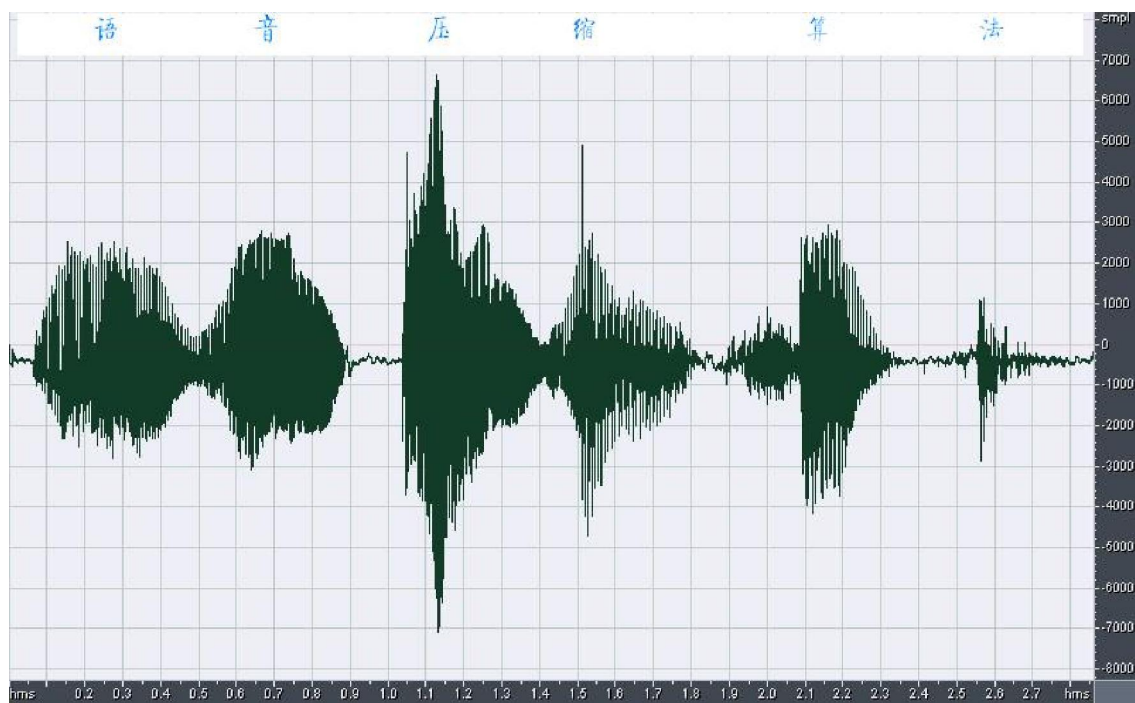


图 5-1 语音信号的时域波形

图 5-1 所示横轴为时间，纵轴为语音信号的幅度。由图可以很直观的看到语音信号的能量起伏，很容易分辨出语句“语音编码算法”中每一个字的文字。由于时间轴压缩得太短，无法看到具体的细节，取中间空隙与下一个字之间的时域波形放大，得到的局部细节展开图如图 5-2 所示。

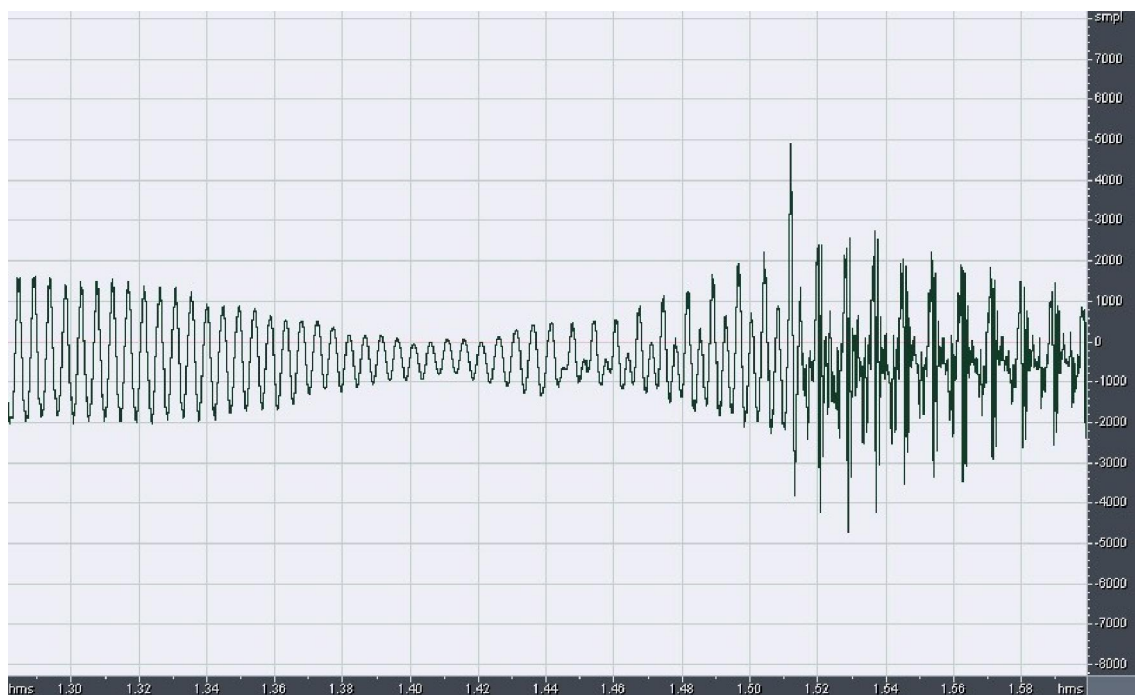


图 5-2 时域波形的局部细节展开图

上图中左边为噪声语音，右边为语音信号。

5.1.1 PCM 语音数据

声音经过麦克风，转换成一连串电压变化的信号，将这声音波形的模拟信号进行 AD 转换，量化得到一系列的信号即为 PCM 语音数据。PCM 语音信号的缺点是体积太大，因此传输时需要加以压缩。PCM 的文本形式如图 5-3 所示。

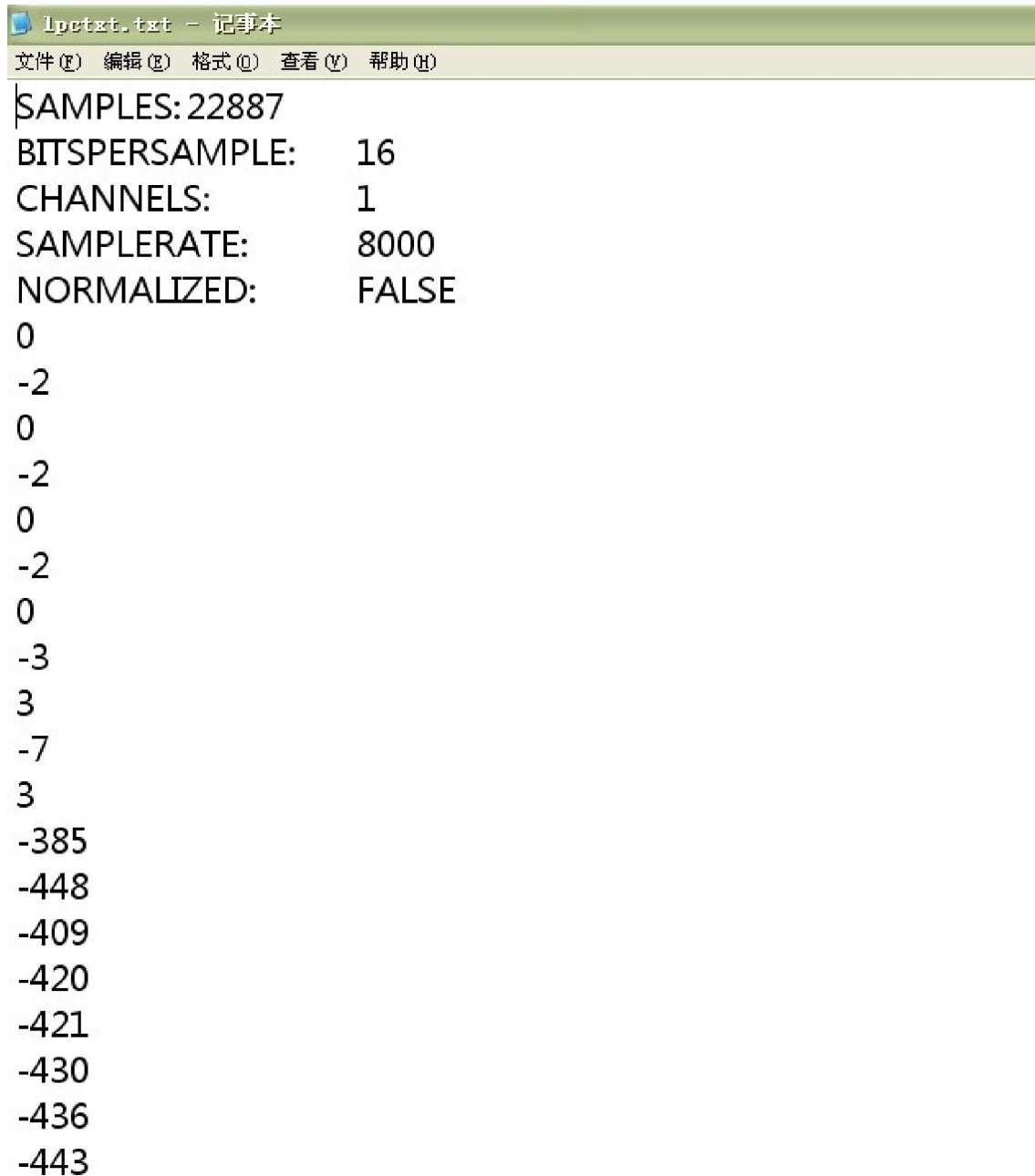


图 5-3 PCM 的文本形式

通过 CCS 中从 PCM 语音数据中取得数据，如图 5-4 所示。

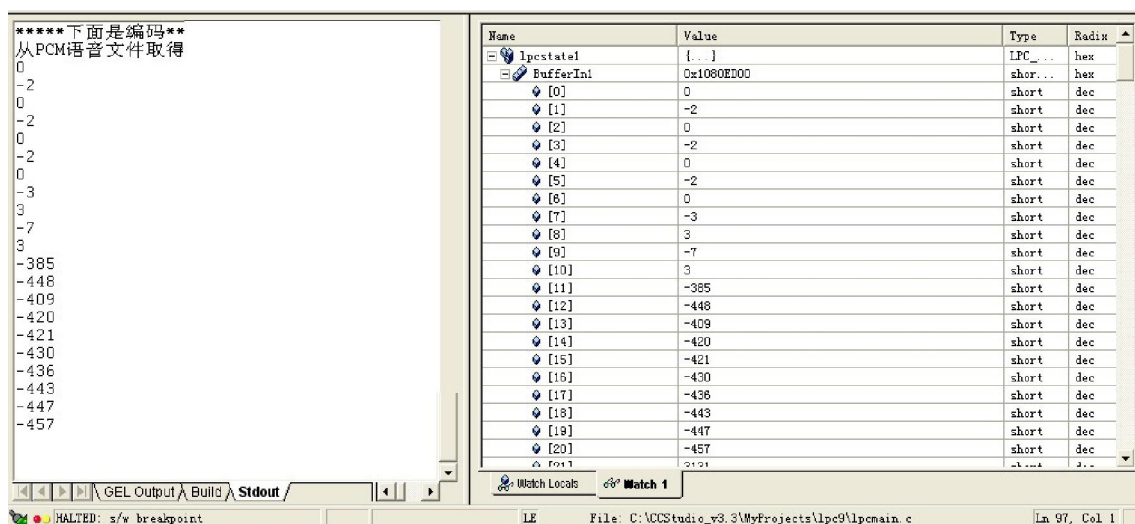


图 5-4 CCS 中取得的 PCM 数据

5.1.2 LPC 编码得到的线性数据流

LPC 编码时是按帧提取，每个帧有 240 个样本，每一帧可以压缩成 12 个值。其文本文件如一个时。

原始 PCM 语音文件总共有 22887 个样本数。分析其中数据，包括 12 个无符号字节数，从 encode_txt 文件中提取两个数组来说明一下：

{103 0 134 0 26 41 54 14 232 249 248 252} 和 {107 0 138 215 4 40 53 36 229 8 254 248}

每个数组的第一数存储的信息是基音周期；第二个数表示增益的均方根值；后面十个数是 10 个预测系数。

下表表示的 LPC 编码得到的数据流情况

表 LPC 编码数据流

编码比特流	浊音	清音
基音周期 per	8bits	0 (无基音)
增益 G	8bits	8bits
k_1	8bits	8bits

k_2	8bits	8bits
k_3	8bits	8bits
k_4	8bits	8bits
k_5	8bits	8bits
k_6	8bits	8bits
k_7	8bits	8bits
k_8	8bits	8bits
k_9	8bits	8bits
k_{10}	8bits	8bits
总计	196bits	196bits

如下图所示

Name	Value	Type	Radix
lpcstate1	{...}	LPC_ST...	hex
BufferIn1	0x1080ECA0	short[...	hex
BufferOut1	276885120.0 ""	unsign...	float
[0]	0.0	unsign...	float
[1]	0.0	unsign...	float
[2]	133.0	unsign...	float
[3]	192.0	unsign...	float
[4]	202.0	unsign...	float
[5]	55.0	unsign...	float
[6]	239.0	unsign...	float
[7]	254.0	unsign...	float
[8]	16.0	unsign...	float
[9]	249.0	unsign...	float
[10]	10.0	unsign...	float
[11]	2.0	unsign...	float
lpcstate2	{...}	LPC_ST...	hex
i	12	int	dec

Watch Locals Watch 1

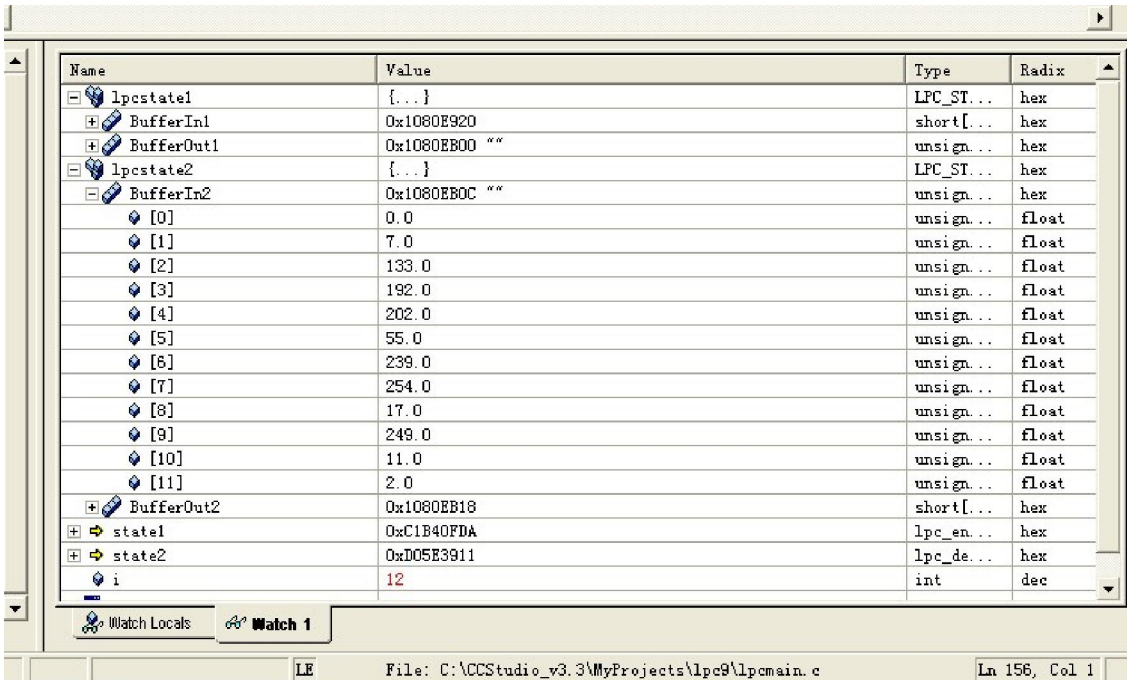
LE File: C:\CCStudio_v3.3\MyProjects\lpc9\lpcmain.c Ln 95, Col 1

图 BufferOut1 数组数据

5.2 LPC 解码调试

LPC 解码是使用编码成的二进制比特流，进行解码，得到一系列参数，然后解码得到一系列语音数据。

可以使用观察查看编码后得到的数组



Name	Value	Type	Radix
lpcstate1	{...}	LPC_ST...	hex
BufferIn1	0x1080E920	short[...]	hex
BufferOut1	0x1080EB00 ""	unsign...	hex
lpcstate2	{...}	LPC_ST...	hex
BufferIn2	0x1080EBOC ""	unsign...	hex
[0]	0.0	unsign...	float
[1]	7.0	unsign...	float
[2]	133.0	unsign...	float
[3]	192.0	unsign...	float
[4]	202.0	unsign...	float
[5]	55.0	unsign...	float
[6]	239.0	unsign...	float
[7]	254.0	unsign...	float
[8]	17.0	unsign...	float
[9]	249.0	unsign...	float
[10]	11.0	unsign...	float
[11]	2.0	unsign...	float
BufferOut2	0x1080EB18	short[...]	hex
state1	0xC1B40FDA	lpc_en...	hex
state2	0xD05E3911	lpc_de...	hex
i	12	int	dec

Watch Locals Watch 1

File: C:\CCStudio_v3.3\MyProjects\lpc9\lpcmain.c Ln 156, Col 1

图 BufferIn2 数组

Name	Value	Type	Radix
lpcstate1	{...}	LPC_STATE1	hex
+ BufferIn1	0x1080E920	short[240]	hex
+ BufferOut1	0x1080EB00 ""	unsigned...	hex
lpcstate2	{...}	LPC_STATE2	hex
+ BufferIn2	0x1080EB0C ""	unsigned...	hex
- BufferOut2	0x1080EB18	short[160]	hex
[0]	133	short	dec
[1]	81	short	dec
[2]	17	short	dec
[3]	-23	short	dec
[4]	-163	short	dec
[5]	-251	short	dec
[6]	0	short	dec
[7]	0	short	dec
[8]	0	short	dec
[9]	0	short	dec
[10]	0	short	dec
[11]	0	short	dec
[12]	0	short	dec
[13]	0	short	dec
[14]	0	short	dec
[15]	0	short	dec
[16]	0	short	dec
[17]	0	short	dec
[18]	0	short	dec
[19]	0	short	dec
- Don't	0	short	dec

Watch Locals Watch 1

LE File: C:\CCStudio_v3.3\MyProjects\lpc9\lpcdec.c Ln 166, Col 1

图 BufferOut2 数组中的数据

其解码得到的时域波形如下图所示

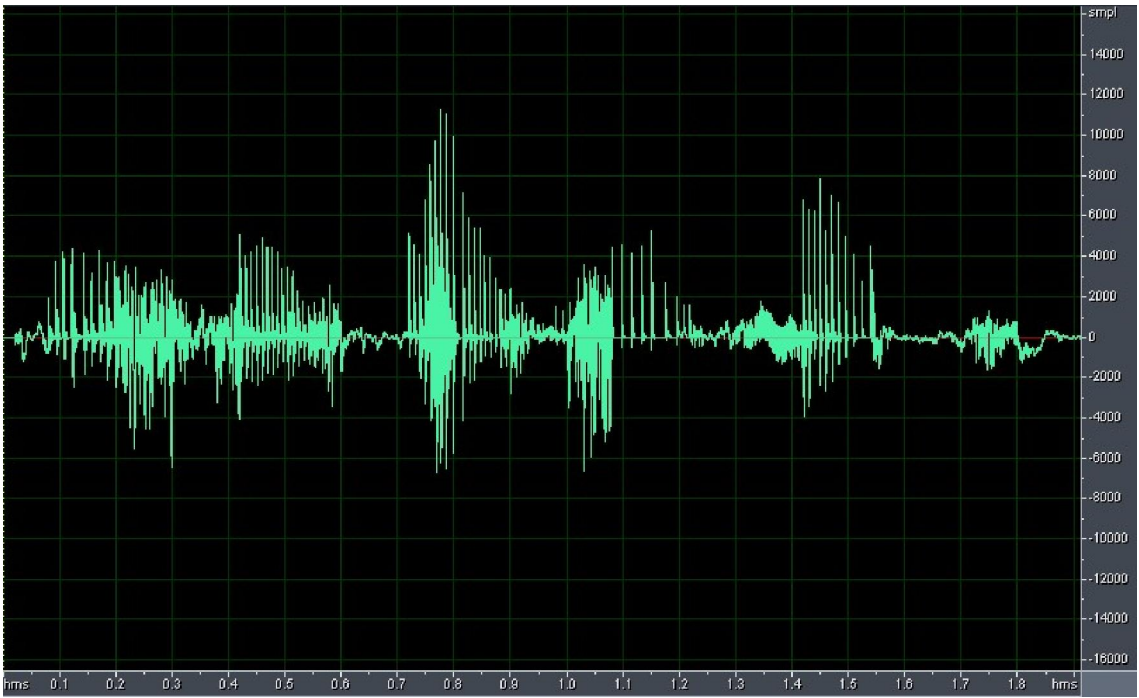


图 解码后的波形

5.3 失真分析

比较原始语音与解码之后的合成语音，发现合成语音音质比较差，与原始波形较大的差别，失真较严重。仔细分析其原因可能如下。

1) 实际情况下，人说话的情况很复杂，在语音信号产生模型上，将语音信号简单地分为受周期脉冲激励和受白噪声激励两种情况与实际情况比较有很大差别，因此简单区分会造成非常多的失真。在实践中，这个二元激励的模型是有局限性的。首先，此模型建立了一个参数随时间缓慢变化的结构，即所谓的“短时”平稳概念，“准稳”的原理对于塞音这样的暂音此模型欠佳。其次，模型将语音信号截然分为受周期脉冲激励和噪声激励两种情况，这与实际情况并不完全符合，将清浊音开关换为叠加号更为合理，这样激励信号可以是上述两种激励相叠加，这更加符合实际情况，而且，此模型的传输函数不包含有限传输点，而像鼻音、擦音这样一些音的声道传输函数中是包含有限零点的，也受到限制。最后，浊擦音不是简单的浊音和清音的叠加，它的发音是很复杂的过程，此模型不能给出模拟。尽管如此，此模型仍能合成出较满意的语音，它一直是分析语音的最重要的基础#。

2) LPC 解码时，实际中，激励信号既含有周期性成分又含有类噪声成分，很难以帧为单位严格将他们分开。

3) 基音周期比较难准确提取，因此无法正确判断清音与浊音。

5.4 LPC 改进措施

从编码分析可知，编码得到的线性数据流没有携带同步及误差校正的信息，因此在传输时产生误码时无法校正，影响效果。

针对上一节提出的 LPC 编解码出现的问题，提出的改进措施如下。

1. 可以进一步减小编码得到的数据流的大小

使用了三角波激励之后的解码波形如下

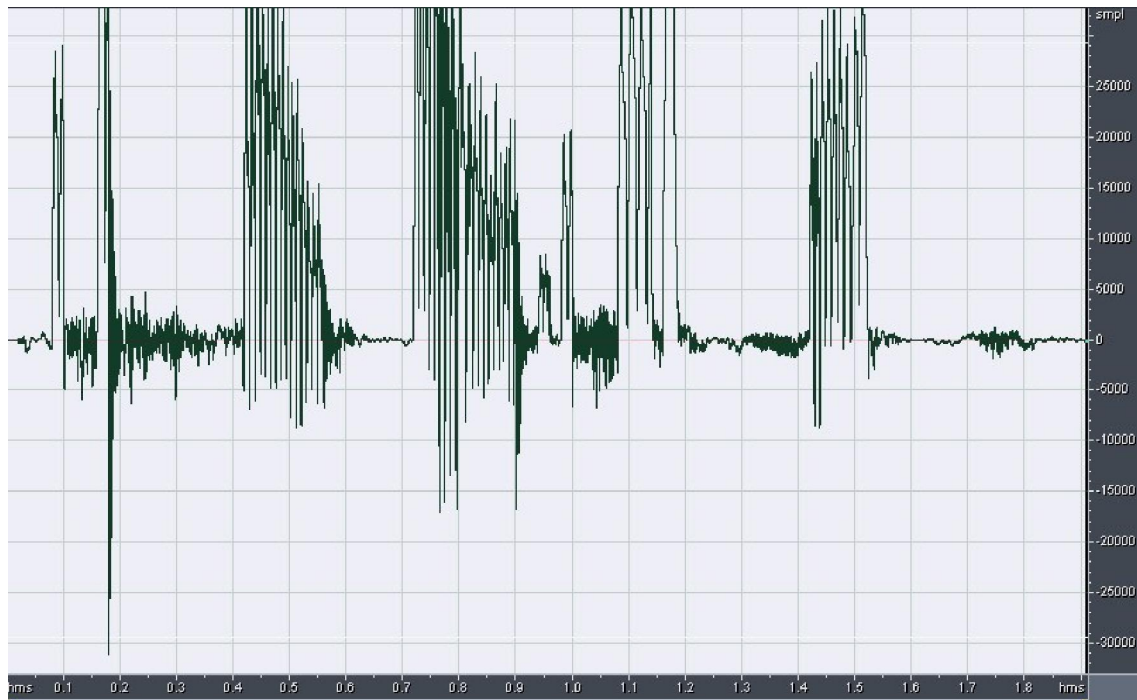


图 三角波激励后解码波形

第6章 工作总结与展望

本章主要是描述论文期间的工作以及一些不足之处、需要改进的地方。

6.1 论文期间完成的工作

总结全部的工作，主要包括以下几个方面：

1. 查阅语音编码相关资料，了解了语音编码技术的产生发展及原理、语音压缩编码的不同分类及性能衡量标准。熟悉常见的诸如 mp3, wav, pcm 等音频编码形式。

2. 选定了 LPC 作为研究课题，对其编码过程进行了深入的研究。研究了语音信号产生模型，语音压缩编码中最常用的线性预测原理，并对此进行了较为详细的分析

3. 对 LPC 解码过程进行了深入研究，研究如何由参数合成语音。

4. 深入学习了 CCStudio 的使用，DSP/BIOS 的使用，*.lib, *.cmd 文件的编写及调用，如何在调试中解决问题。

5. 研究了 TI 公司的 C64x 性能及 DSP 系统的各种开发技术，并对 AIC33 语音编解码芯片的原理及各种工作方式等进行了研究。

6.2 工作展望

本课题期间由于自己的能力和时间限制，还存在很多不足之处。存在问题及需要进一步研究的方向如下：

1. LPC 得到的编码数据流，可以利用诸如 Hamming 码进行参数编码，以便提高抗干扰能力。

2. DM6437EVM 开发板有 AIC33 语音芯片，因此可以直接做到使用麦克风从外界采集语音数据然后由 AIC33 进行模数转换，使用 McBSP 协议将数据传递给 DM6437，然后调用语音编解码算法进行处理，再将数据传递给 AIC33，数模转换

之后通过扬声器或者耳机回放出来。

3. 得到的

参考文献

- [1]吴家安主编. 现代语音编码技术.北京:科学出版社,2007.78~80
- [2]鲍长春编著.数字语音编码原理.西安:西安电子科技大学出版社,2007.128~132
- [3] Richard B.Chase, Nicholas J. Aquilano, F. Robert Jacobs. 运营管理(原书第 9 版).北京:机械工业出版社,2003.145~163
- [4]Madeleine E Pullman, Gary Thompson. Strategies for integrating capacity with demand in service networks, Journal of Service Research, 2003, Vol.5,No.6:16~20
- [5]Mark M. Davis 等. 服务管理(第 1 版).北京:人民邮电出版社, 2006.132~148
- [6]Madeleine E Pullman, Gary M Thompson, Evaluating capacity- and demand-management decisions at a ski resort, Cornell Hotel and Restaurant Administration Quarterly, 2002, Vol.43,No.6:25~37
- [7]陈敏. 基于需求管理的成都地区供电企业服务营销策略研究:[工商管理硕士(MBA)学位论文]. 成都:四川大学,2004
- [8]Kenneth J Klassen, Thomas R Rohleder, Combining operations and marketing to manage capacity and demand in services, The Service Industries Journal, 2001, Vol.21,No.2:25~30
- [9]Rhyne, David M, the Impact of Demand Management on Service System Performa, the Service Industries Journal, 1988, Vol.8,No.4:26~31
- [10]Douglas R Moodie, Demand management: The evaluation of price and due date negotiation strategies using simulation, Production and Operations Management, 1999, Vol.8, No.2:15~23
- [11]Eitan Altman, Uri Yechiali, Analysis of customers' impatience in queues with server vacations, Queuing Systems, 2006, Vol.52,No.4:61~63
- [12]盛骤, 谢式千等. 概率论与数理统计(第 3 版). 浙江大学出版社, 2001.63~65
- [13]James A Fitzimmons 等. 服务管理(第 5 版). 北京:机械工业出版社, 2007.185~207
- [14]Ward Whitt, Sensitivity of Performance in the Erlang-A Queuing Model to Changes in the Model Parameters, Operations Research, 2006, Vol.54,No.2:14~18
- [15]James A Fitzimmons 等. 服务管理(第 5 版). 北京:机械工业出版社, 2007.298~304
- [16]蔡滢. 浅谈消费者的需求管理. 商业时代学术评论, 2006, Vol.7,No.5:44~48
- [17]罗鹏. 服务营销中的排队管理. 营销策略, 2006, Vol.6,No.7:50~51
- [18]李平英. 排队现象及其管理研究. 山东农业大学学报, 2003, Vol.1, No.9:68~72

- [19]姜兵. 银行营业网点客户排队问题研究:[高级工商管理硕士(EMBA)学位论文]. 南京: 南京大学, 2005
- [20]乐友正. 信号与系统(第 1 版). 北京:清华大学出版社, 2004.37~42
- [21]Charles L. Phillips 等.Signals, Systems, and Transfor(第 3 版).北京:机械工业出版社, 2006.58~63
- [22]James A Fitzimmons 等. 服务管理(第 5 版). 北京:机械工业出版社, 2007.129~137
- [23]Kenneth J Klassen, Thomas R Rohleder, Combining operations and marketing to manage capacity and demand in services, The Service Industries Journal, 2001, Vol.21,No.2:27~29
- [24]张准. 论服务业中的顾客需求管理. 内蒙古财经学院学报, 2004, Vol.4,No.6:44~46
- [25]刘洋. 中国移动自有营业厅服务管理研究:[硕士学位论文].北京:北京邮电大学, 2006

致谢

我要首先感谢我的毕业设计指导老师胡全，本论文是在胡全老师的悉心指导下完成的。胡老师严谨治学态度、敏锐洞察力和严于律己、宽以待人的崇高品质为我树立了良好的榜样。在各个方面都锻炼我、培养我，使我各方面能力都有了很大的提高，在此，谨向胡老师表示最诚挚的感谢。

感谢清水河畔程序员版的一些同学，谢谢你们在编程方面的帮助！

非常感谢我的同学李坤，在论文排版方面的一些帮助。

衷心感谢我的父母和姐姐！感谢你们毫无怨言地为我付出，在挫折时给我鼓励。正是由于你们，我才能够不断的成长。

感谢一切帮助我、支持我的人们，谢谢你们！

最后，谨向百忙之中阅读、评审论文的各位老师表示感谢！

附录

附录一：LPC 编解码的主程序源代码

```
/*  
*****  
*****/  

```

附录二：顾客调查问卷

附录三：仿真分号停机策略的 MATLAB 程序命令

外文资料原文

DESIGN OF A DIGITAL SPEECH PROCESSOR

Bin Li, Junbi Chen

COMASIC DESIGN CO., LTD, Shijiazhuang Hebei 050081

E-mail:lb@zeus.comerc.ac.cn

Abstract: This paper describes the design of a flexible hardware structure called Digital Speech Processor(DSP), which is the core module of implementing several code excited linear prediction(CELP) algorithms such as IS-96,IS-54,RCR-27,G.728 etc. In China, there are a lot of communication systems being used, therefore, it is very difficult to determine a common speech coding algorithm, instead, many kinds of speech coding algorithms are coexisting. For this reason, finding a effective way to implement these speech coding algorithms can reduce the product developed time and quickly go to market. DSP described in this paper is more flexible for implementing many speech coding algorithms mentioned above. The structure of DSP is presented in this paper, also included VHDL descriptions of some modules and demand.

The character of several algorithms based on CELP

Code-excited linear prediction coding is a class of speech coding algorithm used widely in the low rate speech communication system. This kind of speech coding algorithm is the most promising schemes especially at 4 to 8 Kbps. For the moment, there are so many international standards of this kind of algorithm such as IS-96, IS-54, RCR-27 and G.728. IS-96 is a variable rate speech compression algorithm used in CDMA digital cellular system, its output rate can be varied from 1.2Kbps to 9.6Kbps depends on speech activity. IS-54 is a vector-sum excited linear prediction coding algorithm that output rate is 7.95Kbps used in DAMPS digital cellular system. RCR-27 is the same as IS-54, except for output rate is 6.7Kbps. This kind of algorithm also

includes G.728 developed by CCITT in 1992, which is a low-delay CELP algorithm and its output rate is 16Kbps.

Although the output rate of these algorithms is not same. The coding process is very similar. The first three algorithms need to determine formant, pitch and codebook parameters. The formant parameters are computed directly from the input speech samples, including the autocorrelation function of the input speech samples, Levinson-Durbin's recursion, covariance lattice algorithm, quantizing and coding the formant parameters. The computation of this part is very intensive. The pitch and codebook parameters are determined through an analysis by synthesis approach whereby an exhaustive search is performed over possible parameter values. This part also needs to compute repeatedly a large amount of mean square errors between the synthesized speech and the original speech. The computation of this part is much more intensive than that in determination of formant parameters. The output parameter of G.728 is only one, but the process to determine this parameter also has contained the computation above.

Because of so many arithmetic computations, if the implementation of these speech coding algorithms in terms of full hardware according to the algorithm process, the scale of the circuit will be too large and don't fit the ASIC manufacture. On the other hand, if we will implement another algorithm, the circuits must be redesigned because the hardware structure is firm. This will be lengthened the design period and, at the same time. The design can not be reused for the reasons above, we presented a effective and flexible hardware structure for implementing these algorithms. The core of this structure is a ware called digital speech processor (DSP). The effectiveness is that the hardware scale is not very large when it is used to implement the speech algorithms while the flexibility is that with this hardware structure many speech compression algorithms can be implemented.

The structure of DSP

Fig 1 shows a block diagram of digital speech processor DSP contains three main parts: 1) Arithmetic & Control unit, including a arithmetic logic unit(ALU), a multiplier and a controller; 2) Memory unit, including a data RAM and a instruction ROM; 3) I/O port. Three parts are connected through a internal bus. The major functions of each part

are as follows. The instruction ROM contains various instruction sets and some constants that will be used while the data RAM contains intermediate results. The controllers reads the instructions in the ROM and send off controlling signals, then, the ALU and multiplier finish the operation according to thus signals. The I/O ports are used to communicate with the external parts. The controller is the core of the DSP and capable of interpreting various instructions. By sending off correspondent controlling signals, the controllers can control the operation of the other parts.

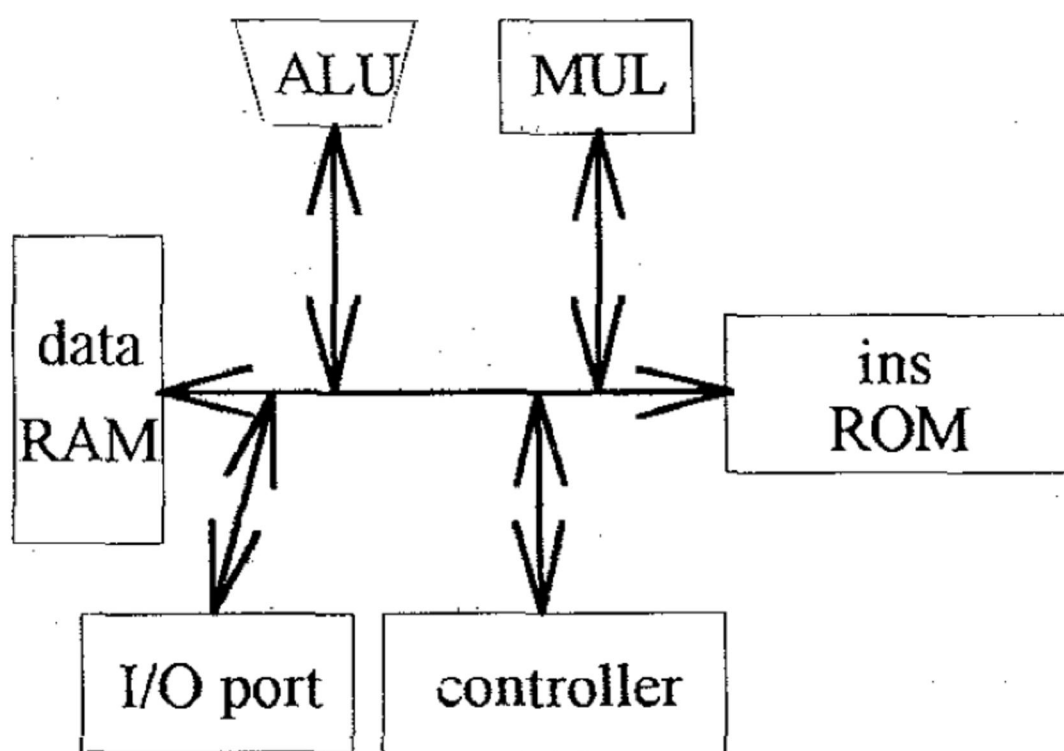


Fig.1 DSP structure diagram

In fact, we can program the instruction sets according to the algorithm flow and put them into the ROM. Each algorithm can be corresponded to instruction set. The computation of each algorithm can be achieved by the DSP. It's easier for implementing the algorithm in software than in hardware. This can reduce the design periods and shorten the time to market.

DSP is a specific circuit, which features are to implement some arithmetic modules with direct logic other than with software through the ALU. The clock frequency is the

major question that should be thought in the DSP design. Each algorithm has specified a certain delay time when all the operations of algorithms should be achieved. This delay is usually very short, eg. for IS-96, the maximum delay is 20ms. According to the computations of IS-96 that should be achieved, the DSP clock frequency is about 40MHz.

The addition, multiplies, square, digital filtering, FFT and soon is the major arithmetic function achieved by the DSP. Among these arithmetic functions, the square, digital filtering and FFT can be simply achieved through basic addition and multiply. In order to keep the scale of DSP not too large, we can only design the adder and a multiplier, then, the square, digital filtering and FFT can be achieved through programming mode. These program is putted into ROM as a firmware and, when required, can be used as subprogram call. The high speed adder and multiplier is the necessary when implementing method above. Therefore the key modules of DSP are a high speed adder and a high speed multiplier.

翻译文稿

数字语音处理器的设计

Bin Li, Junbi Chen

COMASIC DESIGN CO., LTD, Shijiazhuang Hebei 050081

E-mail:lb@zeus.comerc.ac.cn

摘要: 本文描述了一种灵活的对数字语音处理器 (DSP) 硬件结构的设计方法。DSP, 是实现几个码激励线性预测编码算法 (CELP) 的核心模块, 例如 IS-96, IS-54, RCR-27, G.728 等算法。目前在中国有许多不同的通信系统投入使用, 存在着许多不同种类的语音编码算法, 而确定一个通用的语音编码算法是比较困难的。出于这个原因, 找到一种有效的方法来实现这些语音编码算法可以大大减少产品开发时间, 并迅速投入市场。本文提速的 DSP 可以更灵活地实现上述的几个语音编码算法, 还给出了 DSP 的架构及用 VHDL 语言描述的一些模块和需求。

几个基于 CELP 的算法的特征

码本激励线性预测编码 (CELP) 是在中低速率语音通信系统广泛使用的一类语音编码算法, 是传输速率在 4 到 8 Kbps 领域中最有前途的。目前有许多使用这种算法的国际标准, 如 IS-96, IS-54, RCR-27 和 G.728 等。IS-96 是应用在在 CDMA 数字蜂窝系统的一种可变速率语音压缩算法, 其传输速率根据不同语音可从 1.2Kbps 到 9.6Kbps。IS-54 是一种矢量和激励线性预测编码算法, 应用于 DAMPS 数字蜂窝系统, 传输速率是 7.95Kbps。IS-54 也是应用于 DAMPS 数字蜂窝系统, 除了传输速率是 6.7Kbps。CELP 算法还包括由国际电话电报咨询委员会 (CCITT) 于 1992 年制定的 G.728 算法。G.728 算法是一种低时延编码算法, 输出速率是 16Kbps。

尽管不同算法的传输速率是不一样的, 但是不同算法的编码处理过程是相似的。以上 3 个算法都需要确定共振峰, 基音和码本参数。求共振峰是直接对输入语音样本计算的, 过程包括对输入语音样本求自相关函数, Levinson-Durbin 递归

算法，协方差格型算法，量化和最后的共振参数编码。这部分计算非常复杂。而基音和码本参数的求解，是通过搜索所有可能的参数值后分析合成的。这一部分需要反复计算大量的合成语音与原始语音误差的均方根值，比求共振峰参数更复杂。相比而言，G 728 算法包含一个参数，但求解此参数的过程也包括上述计算。

上一段描述了语音的算法处理过程，可以得知语音编码算法拥有庞大的算术计算。如果要设计出能够实现这些语音编码算法的全部硬件电路，电路的规模将变得过于庞大，而不适合 ASIC 制造。另一方面，如果电路要实现另一个算法，因为硬件结构无法更改，那么电路又需要重新设计，设计周期将大大延长。基于上述原因，本文提出了一种高效而灵活的硬件架构来实现这些算法。架构的核心部分是数字语音处理器（DSP）。高效是表现在用来实现的语音编码算法时硬件规模不是很大；灵活表现在硬件结构能够实现很多不同的语音编码算法。

DSP 的架构

图 1 展示的是数字语音处理器（DSP）的框图，由 3 个主要部分组成：1）算法及控制单元，包括一个算术逻辑单元（ALU），一个乘法器和一个控制器；2）存储器，包括存储数据的 RAM 和存储指令的 ROM；3）I/O 端口。三个部分通过内部总线连接。每个部分的主要功能如下：ROM 存储各种指令集和需要使用的一些常量，RAM 存储中间结果；控制器读取 ROM 中的指令和发送控制信号，然后，ALU 及乘法器根据控制信号完成操作；I/O 端口是用来与外部交换数据。控制器是 DSP 的核心，解释各种指示及发出控制信号，以控制其他部分的操作。

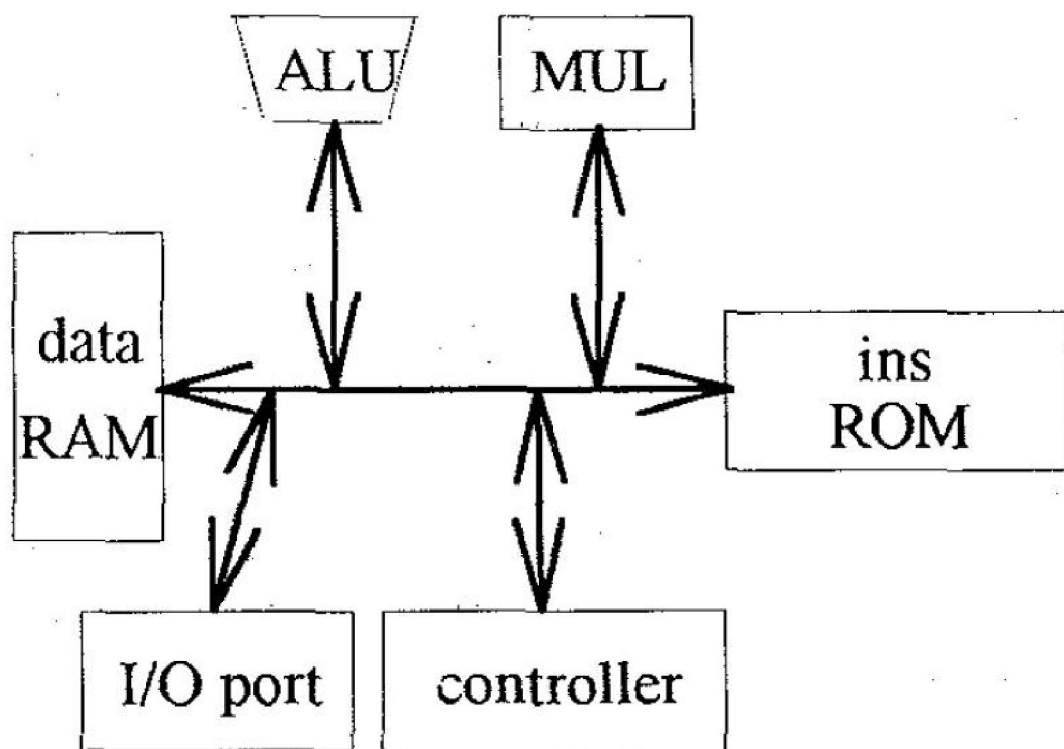


图 1 DSP 的架构图

事实上，我们可以根据算法流程编程系列指令，把指令写入 ROM 中。每个算法都可以写出对应的一系列指令，通过 DSP 实现。用软件实现算法要比硬件实现容易的多。这可以减少设计周期，缩短产品上市时间

DSP 是一种特殊设计的电路，特点是拥有实现一些算法的算术逻辑模块而不是通过 ALU 计算的软件流程。时钟频率是 DSP 设计需要思考的主要问题。在所有不同业务的算法都需要实现时，而每个算法具有一定的延迟。延迟通常是很短的，例如 IS-96 的最大延迟是 20ms。根据 IS-96 算法的要求，DSP 的时钟频率大概是 40MHz。

加法，乘法，平方，数字滤波，FFT 等主要数学运算功能都可以由 DSP 来实现。在以上的数学运算中，平方、数字滤波和 FFT 可以通过基本的加乘运算来实现。为了保证 DSP 的规模不会太大，我们只能设计一个加法器和一个乘法器。平方、数字滤波与 FFT 可以通过编写程序来实现。这些程序将作为固件存入 ROM 中，并在必要时，可作为子程序被调用。实现上述运算，高速加法器和乘法器是必要的。因此，DSP 的核心模块是一个高速加法器和一个高速乘法器。