

保密类别 _____

编号 200512613038

毕 业 论 文

声码器研究

学 院 信息工程学院

专 业 电子信息工程

班 级 2005 级宽带网络方向

姓 名 王非非

指导教师 杜伟韬

中 国 传 媒 大 学

2009 年 5 月 10 日

声码器研究

王非非

摘 要

近二十年来，随着通信技术的高速发展，频带资源愈发紧张，以更小的传输码率获得更优的通话质量逐渐成为了人们追求的目标。语音编码作为实现这一目标的主导力量，在近十几年来也得到了突飞猛进的发展。

语音编码主要分为波形编码，参数编码和混合编码三类。参数编码（又称声码器），以及以其为主导的混合编码方式，具有比特率低，保密性好的优点，在现代移动通信技术中得到了广泛应用。从 2G 时代的 GSM，到是 3G 时代的 CDMA 以及网络电话 VoIP，声码器都发挥了无可替代的作用。

本课题先对当前的语音编码技术和标准进行综述，进而对 GSM 中所采用的 RPE-LTP（规则脉冲激励，长时预测编码）算法标准进行介绍，并对其算法原理进行了深入的研究，并用 MATLAB 实现了该算法。

关键字：语音编解码 声码器 规则脉冲激励，长时预测编码

The Study of Vocoder

Wang Feifei

ABSTRACT

In the past two decades, communication technology is developed very fast. As a result, the frequency band resource is more and more limited. People are in pursuit of using lower bit-rate to get better voice quality. As a powerful tool to realize this goal, Speech compress coding gets rapid development in the past ten years.

Speech compress coding can be classified into three types: Waveform Coding, Parameter Coding and Mixed Coding. Parameter Coding (also called Vocoder) has the advantages of low bit-rates and high security. It's widely applied in the area of modern communication technology. Vocoder plays a non-substitutive role in the speech coding of the GSM of 2G, CDMA of 3G, and the network telephone VoIP.

This subject first describes the present speech compress coding technologies and standards. And then the speech coding standards RPE-LTP (Regular Pulse Excitation – Long term prediction) is introduced. Secondly, the algorithm principles of RPE-LTP are studied deeply. Finally the RPE-LTP codec is implemented in MATLAB language.

Keywords: Speech compress coding Vocoder Regular Pulse Excitation – Long term prediction

目 录

摘 要	I
ABSTRACT	II
绪 论	1
(一) 引言	
(二) 声码器简介	
(三) 课题选题依据及背景	
(四) 算法实现软件 MATLAB	
一、数字语音压缩编码理论概要	3
(一) 语音编码的基本目标	
(二) 语音编码的分类	
(三) 语音编码标准介绍	
(四) 语音编码算法的评价方法	
二、声码器中语音信号处理的关键技术	8
(一) 人体发声机理及声学特征	
(二) 语音信号产生的数字模型	
三、RPE-LTP 算法研究	10
(一) 编码原理	
(二) 解码原理	
(三) 比特分配	
(四) 结果评价	
结 论	26
参考文献	27
附录	28
后 记	30

绪 论

（一）引言

语言作为人类所特有的工具，是人们相互之间情感交流和思想疏通的最主要途径。其载体语音信号是人与人之间相互交流最直接，最快捷，最有效的信息形式。作为最基本的通信手段，语音通信在通信领域中一直以来占据着相当大的比重。随着通信技术的高速发展，频带资源的逐渐匮乏，人们对语音编码技术的要求越来越高^[1]，传输速率从最初固定电话 A/μ 律 PCM 的 64kbps 到移动通信 GSM 的 13kbps，再到现在的军用电话，卫星通信的 0.6kbps，语音编码技术不断地追求着更少的比特率和更高的语音质量，并逐渐向极限逼近。

在移动通信技术从 2G 向 3G 过渡，以及以后向 4G 发展的过程中，单一的语音数据传输服务也在向丰富多彩的多媒体业务发展，对语音编码的要求也越来越高，要求语音的传输速率更低，并且能够灵活变化，以为其他业务省出频带资源。国际电信联盟（ITU）最近的报告指出，目前全球的手机用户达到 41 亿，普及率达到了 61.1%^[2]，其潜在的经济价值是巨大的。

（二）声码器简介

声码器（VOCODER，又称参数编码）作为语音编码的一个重要分类，是一种能对语言信号自动进行分析-合成的装置。它利用参数编码实现语音通信，是压缩通信频带和进行保密通信的有力工具。根据人发声的特点，声码器在发送端对语言信号进行分析，提取出语言信号的特征参量加以编码和加密，以取得和信道的匹配，经信道传递到接收端，再根据收到的特征参量恢复原始语言波形。分析可在频域中进行，对语言信号作频谱分析，鉴别清浊音，测定浊音基频，进而选取清-浊判断、浊音基频和频谱包络作为特征参量加以传送。分析也可在时域中进行，利用其周期性提取一些参数进行线性预测，或对语言信号作相关分析。

声码器的明显优点是码率低，因而适合于窄带、昂贵和劣质信道条件下的数字电话通信，能满足节约频带、节省功率和抗干扰编码的要求。低速码率对语音存储和语音加密处理也都很有利。声码器的缺点是音质不如普通数字电话好，而且工作过程较复杂，造价较高。目前声码器广泛应用于军事工业和卫星通信，并且在移动通信标准中，也随处可见声码器的身影。

（三）课题选题依据及背景

本课题在广泛了解现有的各式各样的声码器基础上，对声码器的普通原理进行了深入的学习。但由于其种类繁多，复杂度较高，研究难度较大，通过调研，最终选定了技术上已经非常成熟的 GSM（全球移动通讯系统）^[3]中使用的混合语音编码算法 RPE-LTP(Regular Pulse Excitation – Long term prediction)。

RPE-LTP 算法音质较佳，速率居中，编码复杂度也相对较小，适合作为对声码器的入门性研究。其广泛的应用和可靠的质量保证也是选择该算法的一个重要原因。

RPE-LTP 算法^[4]在参数编码方面使用了 LP（Linear Prediction，线性预测）技术和 LTP（Long Term Prediction，长时预测分析）技术，其中 LP 技术是现代声码器的一个重要基础，对它进行深入学习对以后研究其他声码器有着十分重要的作用。RPE-LTP 算法使用 RPE（规则脉冲激励）来作为激励，这是区别于纯粹的参数编码的一个重要特征，参数编码中通常使用码本搜索或预测基音周期作为激励。规则脉冲激励是一组在位置上和幅度上均优化的脉冲序列，用它们来代替残差信号（信号与预测信号之差值）进行编解码，既可以简化编码又能使合成波形失真较小。

（四）算法实现软件 MATLAB

MATLAB 是美国 MathWorks 公司出品的商业数学软件^[5]，用于算法开发、数据可视化、数据分析以及数值计算的高级技术计算语言和交互式环境。它具有友好的用户界面和强大的数字信号处理能力，包含了丰富的数字信号处理相关函数，比 C 语言具有更高的算法开发效率。使用 MATLAB 实现算法并进行仿真，对在此基础上进行 C 代码的编写以及后续硬件的实现有着重要的辅助作用。

一、数字语音压缩编码理论概要

语音编码始终与电话的发展有着不解之缘。在1876年贝尔(Alexander Graham Bell)发明电话^①以后的几十年里,人们在享受电话方便的同时却对电话信号及电话系统所知甚少。从上世纪二十年代开始,随着美国贝尔实验室的建立,电话系统也渐渐开始得到了理论化系统化的研究。在长达半个多世纪的研究中,通过各国学者的大量努力,对人类的发音和听觉机理进行了深入的研究,建立了逼近实际的发声和听觉模型,电话通信技术也得到了长足的发展。

无论是固定电话还是移动电话,都是按着从模拟向数字的方向发展的。随着数字电话逐渐取代模拟电话,数字语音编码技术也开始登上历史舞台。从七十年代开始,各种语音编码标准相继出台,语音编码技术逐渐丰富和成熟,特别是在近十几年里,得到了突飞猛进的发展,成为国际标准化工作中最活跃的领域,已经具备了完善的理论和技术体系,并进入实用阶段。

从目前的通信技术的发展不难预测,语音编码以其本身的必要性,在以后的很长一段时间,依然会是专家学者们努力探索的一个重要领域,保持较高的科研热度。

(一) 语音编码的基本目标

语音编码最初是源于人们对压缩频带的需要,在其发展过程中一直追求着更低的码率,并且在相同的码率上,追求更高的合成语音质量。语音编码一直在码率和音质的矛盾中发展,更低的码率是以音质降低为代价的,而较高的音质则意味着更高的传输带宽。

语音编码技术在向着低速率和音质无失真两个极端方向发展,但更为一般的是两者的折中。当前许多成熟的语音编码技术大多是在保证音质能满足通信要求的前提下,选择适中的传输速率。

(二) 语音编码的分类

语音编码按照编码方法目前主要分为三类^[6]: 波形编码, 参数编码和混合编码。

1. 波形编码

波形编码直接对语音时域或频域波形样值进行编码, 力图使重建语音波形保持原

^①另有一种说法电话发明者是德国的科学教师菲利普·雷斯

语音信号的波形形状。它通常将语音信号作为一般的波形信号来处理，因而能够保留语音信号原始样值的细节变化，具有适应能力强，话音质量高等优点。但波形编码的最大缺点是传输码率高，通常在 16~64 Kbps 的速率上具有较好的编码质量，当码率进一步降低时，性能会迅速下降，无法满足低速率传输的要求，适用于固定线路，长途通话等。

常用的波形编码方式有：脉冲编码调制（PCM），自适应增量调制（ADM），自适应差分脉码调制（AD-PCM）和自适应预测编码（APC）等。

2. 参数编码

参数编码，即本课题的主要研究对象，又叫做声码器，是通过对语音信号特征参数的提取和编码，力图使重建语音信号具有尽可能高的可懂性，即保持原语音的语意，而重建语音信号的波形同原始语音信号的波形可能会有相当大的差别，因而无法准确还原原始信号的细节特征，从而自然度较差，无法准确分辨说话人。此外参数编码对讲话环境噪声比较敏感，只有在安静的环境下才能给出较高的可懂度。参数编码的优势在于编码速率很低，一般可低至 1.2~2.4 Kbps 以下，而目前在军事工业中甚至到了 600bps^[7]。

常见的声码器有共振峰声码器和线性预测声码器。

3. 混合编码

混合编码是在参数编码模型的基础上，先求得模型的特征参数，然后改变其激励信号，使这些参量能够更好地还原出原始信号，获得更好的语音自然度。从其原理上可以看出，混合编码器更依赖于参数编码，它们具有相同的编码原理，只是激励的不同选择，因而可以将参数编码和混合编码归为一类，在下文中为了叙述方便，统称为声码器。

这类声码器的典型代表有：规则脉冲激励预测声码器（RPE-LTP），码激励线性预测声码器（CELP）以及混合激励线性预测声码器（MELP）。

（三）语音编码标准介绍

随着语音编码技术的逐渐成熟和越来越广泛的应用，人们对语音编码公共标准的需求也变得越来越迫切。制定语音编码标准的机构应运而生，并且制定了大量的标准。广泛地了解这些标准有利于对语音编码系统有个全面的认知，了解声码器在这个体系中所占有的重要地位。

目前从事语音编码标准制定的主要组织有国际电信联盟电信标准部（ITU-T）、欧洲电信标准协会（ETSI）、北美电信工业协会（TIA）等。

下面按照编码速率进行分类^[8]，对目前比较常用的重要的编码标准进行介绍。

1. 高速率语音编码国际标准

高速率语音编码的编码速率一般指在 32 Kbps（含）以上，一般采用波形编码方法进行编码。表 1-1 给出了高速率语音编码标准的相关资料。

表 1-1 高速率语音编码国际标准

编码速率 (Kbps)	标准名称	制定机构	公布时间	编码算法	编码方法	MOS 评分	主要用途
64	G.711	ITU-T	1972	PCM	波形	4.3	公用网
56	G.722	ITU-T	1988	SB-ADPCM	波形	4.0	公用网
40	G.726	ITU-T	1990	ADPCM	波形	4.2	公用网

2. 中速率语音编码国际标准

中速率语音编码的编码速率一般指在 4.8~32 Kbps 之间，一般采用混合编码方法进行编码。表 1-2 给出了中速率语音编码标准的相关资料。

表 1-2 中速率语音编码国际标准

编码速率 (Kbps)	标准名称	制定机构	公布时间	编码算法	编码方法	MOS 评分	主要用途
16	G.727	ITU-T	1990	ADPCM	波形	2.2	PCME
13	GSM06.10	ETSI	1988	RPE-LTP	混合	3.7	移动通信
12.2	GSM06.60	ETSI	1991	ACELP	混合	3.6	移动通信
8	G.729	ITU-T	1996	CS-ACELP	混合	4.0	公用网
6.5	GSM06.20	ETSI	1994	VSELP	混合	3.5	移动通信
5.3	G.723	ITU-T	1995	ACELP	混合	3.6	因特网

3. 低速率语音编码国际标准

低速率语音编码的编码速率一般指在 4.8 Kbps（含）以下，采用声码器技术进行编码，在保密电话和军事工业中有着广泛的应用。表 1-3 给出了低速率语音编码标准的相关资料。

表 1-3 低速率语音编码国际标准

编码速率 (Kbps)	标准名称	制定机构	公布时间	编码算法	编码方法	MOS 评分	主要用途
4.8	FS-1016	NSA	1991	CELP	混合	3.2	保密电话
4.8	STC-1	ETSI	1989	STC	混合	3.5	保密电话
2.4	MELP	DDVPC	1997	MELP	混合	3.2	保密电话
2.4	FS-1015	NSA	1981	LPC-10	参数	2.5	军事通信

4. 变速率语音编码国际标准

除了上述的语音编码标准外，随着码分多址（CDMA）技术的飞速发展，及其对主音编码技术的进一步要求，变速率语音编码技术被推到了前台。这种编码技术能够根据信道的占用情况灵活改变传输速率，能够有效地提高频带利用率。表 1-4 给出了变速率语音编码标准的相关资料。

表 1-4 变速率语音编码国际标准

编码速率 (Kbps)	速率个数	标准名称	制定机构	公布时间	编码算法	编码方法	主要用途
6.6-23.85	9	G.722.2	ITU-T	2000	ACELP AMR-WB	混合	WCDMA 通信系统
1.0-13.3	4	IS-733	TIA	1998	QCELP	混合	3G CDMA
4.75-12.2	8	AMR26.090	3GPP	1999	AELP AMR	混合	WCDMA 通信系统
0.8-8.55	4	IS-893	3GPP2	2001	EX-CELP SMV	混合	3G CDMA 系统

（四）语音编码算法的评价方法

语音压缩虽然会降低传输码率，但合成语音质量也会下降，并且会增加复杂度和编解码时延，因此评价一个语音编码算法的优劣要从实际出发，对其传输速率、语音质量，算法复杂度和编解码时延^[9]等方面进行综合考虑。

1. 语音质量

语音质量包括声音的可懂度、清晰度和自然度。若所统计的语言单位为有意义的语言单位，测试结果称为可懂度，如单词可懂度、单句可懂度。当所统计的语言单位为意义不连贯的语言单位，则称清晰度，如语音清晰度、音节清晰度。自然度则反映语音的韵调特性，是对讲话人的辨识水平的反映。

由于语音信号本身的特点，语音质量的测试一直没有适合的客观评价方法。目前比较公认的方法是进行广泛的人为主观评价。目前的主观评价方式有平均意见评分测试（Mean Opinion Score, MOS），用于评测可懂度或清晰度的判断韵字测试（Diagnostic Rhymer Test, DRT）和评估综合质量的判断满意度测试（Diagnostic Acceptability Measure, DAM）等。

平均意见评分测试通常采用 5 级评分标准，即优、良、中、差、坏，用数字 5~1 对应表示这 5 个等级。参加测试的人员在听完测试语音后，选择相应的等级评分，全体测试人员的平均分就是被测语音质量的 MOS 分。在数字语音通信中，通常认为 MOS 分在 4.0~4.5 时为高质量，达到长途电话网的质量要求。MOS 分在 3.5 左右时为通信

质量，能感觉到重建话音质量有所下降，但不妨碍正常通话。MOS 分在 3.0 以下时，称为合成语音质量，只具有足够的可懂度。

2. 编码速率

如果数字音频信号的抽样频率为 f_s ，对每个抽样的幅度值用 R 位二进制编码表示，就得到了数字音频信号的传输速率 I （也可称为码率或比特率），即：

$$I = f_s \times R \quad (1-1)$$

为该数字音频信号在通信线路上每秒钟应传送的位数，单从数据压缩的角度看，传输速率常常是体现数字音频编码器压缩算法技术水平的最主要的指标。

3. 编码复杂度

复杂度主要是考虑在编码算法硬件实现时所耗费的成本和功耗，这也是实际应用中必须要注意的一个问题。复杂度主要取决于算法所采用技术的运算量。

4. 编解码延时

语音的编解码过程会增加延时，延时过大会使人感觉到通话对方反应“迟钝”，并且由于固定电话中固有的回声现象，如果往返总延时超过 100ms 时，发话者就能从手机中听到自己的回声，导致正常通信困难。

二、声码器中语音信号处理的关键技术

（一）人体发声机理及声学特征

声音是由于空气振动而产生的自然界的一种物理现象，引起空气振动的部分称作声源，也叫做声音激励。人的发音器官包括：肺、气管、喉（含声带）、咽、鼻和口。这些器官共同形成一条形状复杂的管道，其发声机理类似于管乐器，首先由气源产生气流并穿过一条狭长的有孔管道，气流引起管乐器的笛膜之类的发声部分（相当于人体的声带）的振动产生声波，声波经过管乐器的孔引起谐振。改变谐振孔开启的位置就可以调整音调。

语音激励根据发音的机制不同可以分为两类。当气体通过紧绷的声带时，声带会周期性地振动，产生准周期性脉冲序列空气流，这时发出的声音称为浊音（Voiced Sound）。浊音通常是一个准周期信号，其频率称为基音频率，通常在 60~450 Hz 范围内。当气体通过完全舒展的声带时，产生空气湍流，这时发出的声音称为清音（Unvoiced Sound）。在发清音时声带不振动，声道被噪声状的随机波所激励。其幅度一般较小，波形与噪声相似。

此外，由于口腔、鼻腔及咽腔的不同位置和形状变化，会使声道具有不同的固有频率。通过改变声道形状来产生不同的共振峰，从而产生出不同的元音和辅音。在声波离开声道时，由于嘴唇的辐射作用，也会对声波产生一定的影响。

（二）语音信号产生的数字模型

对应人体发声机理，语音信号的数字模型分为三个部分：激励模型，声道模型，辐射模型^[10]。

激励模型：声码器中最基本的激励模型为二元激励，其结构如图 2-1 所示。对于浊音，由于产生原因是声门周期性地开启和关闭，因此可以用一个斜三角脉冲来代替，其周期等于基音周期。对于清音，由于是气体湍流，可以用一个白噪声发声来模拟。增益因子 A_v 和 A_u 分别用于调节浊音和清音的幅度或能量。

比较常用的还有码激励模型。该模型将预先计算好的各种可能的激励进行编号，称为码书，在实际应用中通过判决选择最佳激励。码书通常由固定码本和自适应码本两部分组成。

本文主要研究的 RPE-LTP 算法采用的激励是规则脉冲激励，在以后的章节中会再详细阐述。

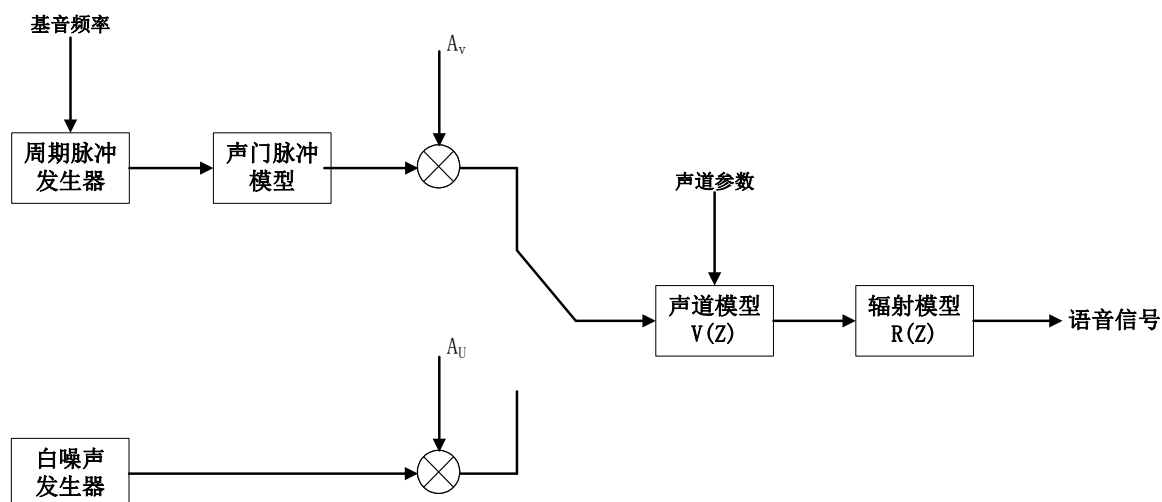


图2-1 语音信号产生的数字模型

声道模型：典型的声道模型主要有无损声管模型和共振峰模型。本文主要论述无损声管模型。

无损声管模型假定声道是由多个等长的具有不同截面积的管子串联而成的系统，并且截面积近似不变，因此在短时间内声道可以表示为形状稳定的管道。通过理论分析，可推导出无损声管模型的传递函数为：

$$V(z) = \frac{1}{\sum_{i=0}^p \alpha_i z^i} \quad (2-1)$$

它是一个全极点函数，其中， $\alpha_0 = 1$ ， α_i 为实数， p 为全极点滤波器的阶，即近似为 p 段声管的串联。 p 值越大，模型的传输函数与声道传输函数的吻合程度越高。实际应用中， p 一般取 8~12 即能满足要求。

辐射模型：主要考虑人嘴在发声时对声音的影响。

三、RPE-LTP 算法研究

RPE-LTP 算法全称规则脉冲激励，长时预测编码，是 ETSI（欧洲电信标准协会）在 1988 年制定的用于泛欧 900MHz 数字蜂窝系统的全速率语音编码标准。

RPE-LTP 算法是一种中速率混合编码方式，在激励上采用规则脉冲序列代替简单的二元激励。所谓规则脉冲激励是指一组在位置上和幅度上均优化的脉冲序列，该序列保有原信号较完整的基音周期信息和声音细节特征，因而能还原出可懂度高且较为自然的语音。

RPE-LTP的编码净比特率为13Kbps，加上信道抗误码编码后为22.8kbps，当再加上其它信息码及保护间隔后，信道传送速率每话路为24.7kbps^[11]。其主观平均评分(MOS)达3.8。该编译码方案抗误码性能较好，编译码延时约为30ms，同时这种算法杂度较小，硬件易于实现，适于实际应用。

（一）编码原理

RPE-LTP语音编码器主要包括预处理、线性预测（LP）分析、短时分析滤波、长时预测及规则脉冲激励码编码五大部分。图3-1给出了RPE-LTP语音编解码器的结构框图。

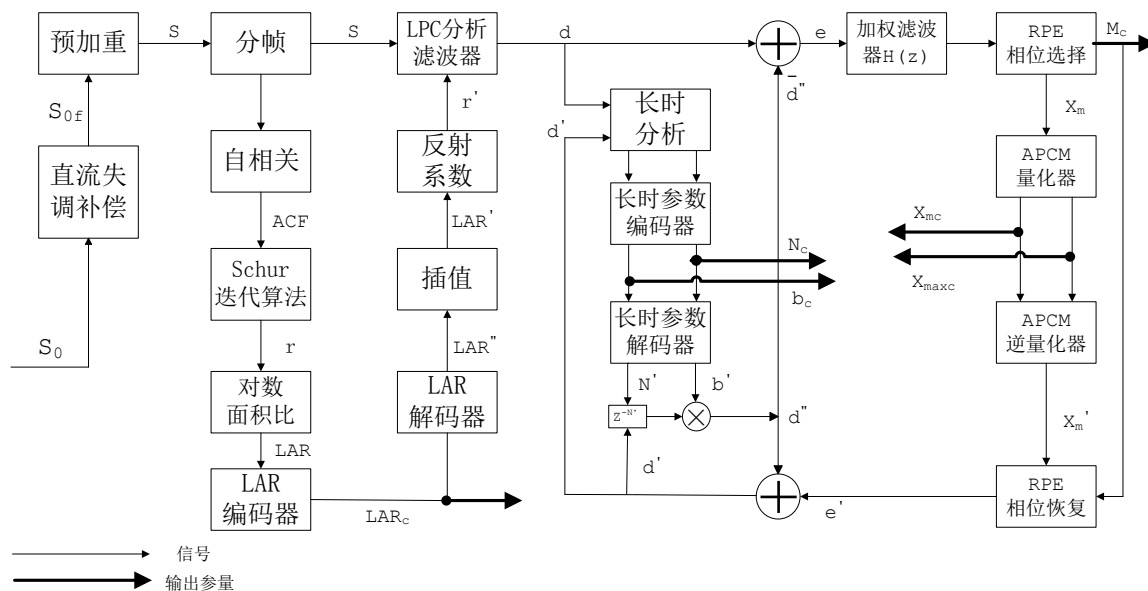


图3-1 RPE-LTP全速率语音编码原理图
Fig.3.1 Block diagram of the RPE-LTP encoder

1. 预处理

预处理包括偏移补偿和预加重两个步骤。

偏移补偿的目的是减少语音信号中直流分量的影响。输入信号为 $S_0(n)$ ，偏移补偿算法如下：

$$S_{of}(n) = S_0(n) - S_0(n-1) + \alpha \times S_{of}(n-1) \quad (3-1)$$

$$\alpha = 32735 \times 2^{-15} = 0.9989929$$

预加重主要考虑人嘴对发声的影响。由于语音信号的平均工功率谱受声门激励和口鼻辐射的影响，高频端大约在 800Hz 以上按 6dB/倍频或 20dB/10 倍频。所以其高频部分的频谱要比低频部分难求，为此要在预加重（Pre-emphasis）处理。预加重的目的是提升高频部分，使信号的频谱变得平坦，保持在低频到高频的整个频带中，能用同样的信噪比求频谱，以便于声道参数分析，在本算法中主要是为了提高 LPC 的分析性能。通常使用一个具有 6dB/倍频的提升高频特性的预加重数字滤波器来实现，它一般是一个一阶的数字滤波器：

$$H(Z) = 1 - \beta z^{-1} \quad (3-2)$$

在本算法中采用的是 6dB/倍频的提升方案，取 $\beta = 28180 \times 2^{-15} = 0.8599853$ ，预处理算法如下：

$$S(n) = S_{of}(n) - \beta \times S_{of}(n-1) \quad (3-3)$$

图 3-2 为 MATLAB 仿真的该预加重滤波器的幅频响应曲线。

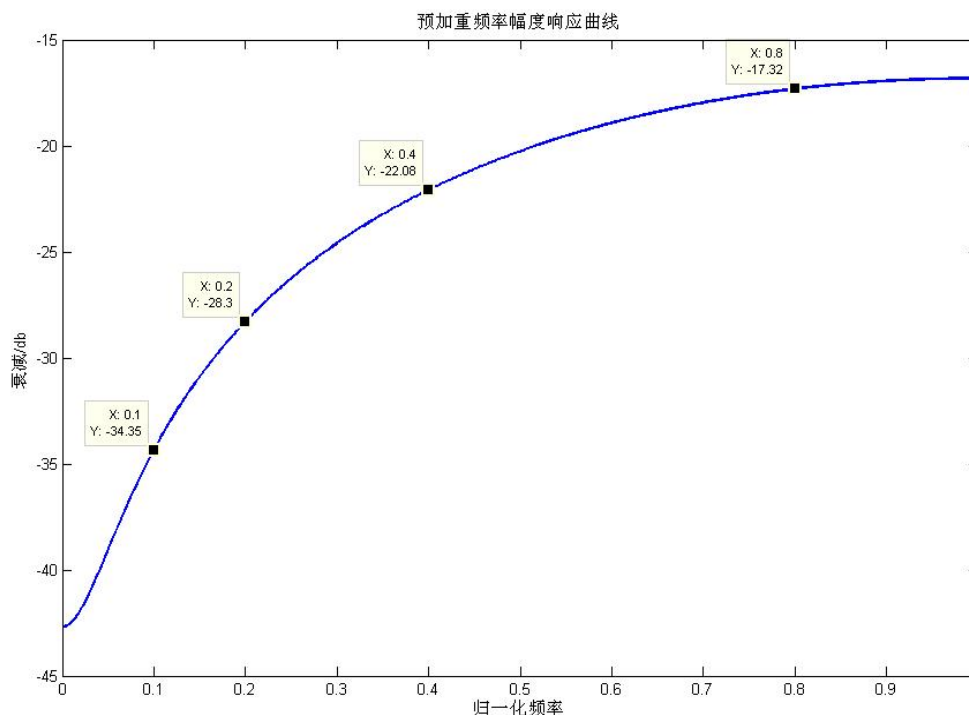


图3-2 预加重频率-幅度响应曲线
Fig.3.2 The amplitude-frequency response curve of Pre-emphasis

2. 线性预测分析

(1) 语音信号的分帧处理

语音信号从整体上来看是一个长时非平稳过程，不能用处理平稳信号的数字信号处理技术来对其进行处理。但语音信号在 10ms~30ms 的短时范围内具有短时平稳性，可以用处理平稳信号的数字处理技术来对其进行近似处理。

所以所有语音编码算法都必须建立在“短时”的基础上，即进行短时分析，将语音信号分为一段一段来分析其特征参数，其中每一段称为一“帧”，因此帧长一般也取 10ms~30ms。这样语音编码就是在帧结构上来进行数据处理的过程。

在该算法中取帧长为 20ms,由于输入数据为 8000Hz 采样，所以一帧应取 160 个样点。如果输入语音数据长度不是 160 的整数倍，则在最后一帧后补 0。

(2) 线性预测的理论推导

线性预测 (Linear Prediction) [12] 在上世纪 60 年代首先被应用到语音分析和合成中后，便成为一种很重要的技术，几乎普遍地应用于语音信号处理的各个方面，特别是在声码器中，得到了广泛的应用，是声码器最关键的一个技术要素。

线性预测的基本思想是：由于语音样点之间存在相关性，所以可以用过去的样点值来预测现在或未来的样点值，即一个语音的采样能够用过去若干个语音采样或它们的线性组合来逼近。通过使实际语音采样和线性预测采样之间的误差在某个准则下达到最小值来决定唯一的一组预测系数。这里的预测系数是指过去样点线性组合的加权系数，它能够反映语音信号的特性。

线性预测为语音信号的数字模型提供了一个非常合适的声道参数估计方法，能够有效地降低编码码率，并且比较准确地表达频谱幅度，可以很好地还原语音。线性预测的模型如下图所示：

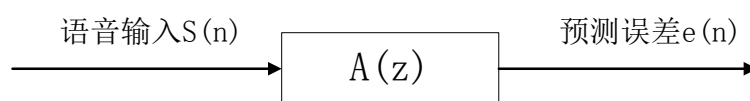


图3-3 语音信号的线性预测模型
Fig.3.3 The Linear Prediction Model of Voice signal

设输入信号为 $S(n)$ ，则其预测值为

$$\hat{S}(n) = \sum_{i=1}^p \alpha_i S(n-i) \quad (3-4)$$

其中 $\{\alpha_i\}_{i=1,2,\dots,p}$ 称为线性预测器系数， p 称为预测阶数。 $e(n)$ 称为预测误差，其表

达式为:

$$e(n) = S(n) - \hat{S}(n) = S(n) - \sum_{i=1}^p \alpha_i S(n-i) \quad (3-5)$$

线性预测的实质就是设计一个误差滤波器, 使预测误差 $e(n)$ 在某个预定的准则下最小, 在该算法中采用的是最小均方差准则, 即 $E[e^2(n)]$ 最小。为了得到 $E[e^2(n)]$ 最小的 α_i 可将 $E[e^2(n)]$ 对各个系数求偏导, 并令其结果为零, 即:

$$\frac{\partial E[e^2(n)]}{\partial \alpha_j} = -2E[e(n)s(n-j)] = 0, \quad 1 \leq j \leq p \quad (3-6)$$

经代入计算可得:

$$r(j) - \sum_{i=1}^p \alpha_i r(j-i) = 0, \quad 1 \leq j \leq p \quad (3-7)$$

式中 $r(j) = E[e(n)s(n-j)]$ 是待测信号 $s(n)$ 的自相关序列。由此式求得的 α_i 将使得预测误差的输出均方值, 即输出功率最小。令这一最小均方误差为正向预测误差功率 E_p , 即

$$E_p = E[e^2(n)]_{\min} = E\left[e(n)\left\{s(n) - \sum_{i=1}^p \alpha_i s(n-i)\right\}\right] \quad (3-8)$$

又由式 (3-6) 可得 $E[e(n)s(n-j)] = 0$, 所以有:

$$E_p = E[e(n)s(n)] = r(0) - \sum_{i=1}^p \alpha_i r(i) \quad (3-9)$$

由式 (3-7) 和式 (3-9) 组成的方程组是完整的, 因而求最小 $e(n)$ 的关键就是求解下面的方程组:

$$\left\{ \begin{array}{l} r(j) - \sum_{i=1}^p \alpha_i r(j-i) = 0 \end{array} \right. \quad (3-10a)$$

$$\left\{ \begin{array}{l} r(0) - \sum_{i=1}^p \alpha_i r(i) = E_p, \quad 1 \leq j \leq p \end{array} \right. \quad (3-10b)$$

(3) 短时分析滤波参数的计算

在短时分析滤波中需要称为“反射系数”的参数, 其含义在下面的论述中会详细

介绍。反射系数的计算需要在线性预测中计算得出，其步骤如下：

1) 按下式计算出信号 $s(n)$ 的前9个自相关系数 $ACF(j)$

$$ACF(j) = \sum_{n=j}^{159} S(n)S(n-j) \quad , \quad j = 0,1,\dots,8 \quad (3-11)$$

2) 推算方程组 (3-10) 的解。采用的方法是采用Schur递推算法，求出前8阶的反射系数 $r(i)$ ，其流程图见附录A-1。

3) 根据级联声管模型可知，反射系数是反映声管模型截面积变化的一组参数，且这组参数与短期格形滤波器的滤波参数相等，即反射系数就是格形滤波器的滤波参数。由于反射系数具有分布不均匀的特点，为了合理的配置量化比特，需要将反射系数转换为对数域表示的形式后再对其进行量化和编码，即转化为LAR (Log-Area Ratio) 参数：

$$LAR(i) = \log_{10}\left(\frac{1+r(i)}{1-r(i)}\right) \quad (3-12)$$

在实际运算中，为了简化对数运算，以折线函数来代替对数函数，有：

$$\begin{cases} LAR(i) = r(i) & |r(i)| < 0.675 & (3-13a) \\ LAR(i) = \text{sign}[r(i)] \times 2 \times |r(i)| - 0.675 & 0.675 \leq |r(i)| < 0.95 & (3-13b) \\ LAR(i) = \text{sign}[r(i)] \times 8 \times |r(i)| - 6.375 & 0.95 \leq |r(i)| < 1 & (3-13c) \end{cases}$$

4) 反射系数对数面积比的量化与编码^[13]

$$LAR_c(i) = N_{\text{int}}[A(i) \times LAR(i) + B(i)] \quad (3-14)$$

其中 $N_{\text{int}}[z] = \text{int}[z + \text{sign}(z) \times 0.5]$ ，即取最靠近变量的整数。量化系数A(i)和B(i)以及LAR(i)的值如表3-1所示。

表3-1 反射系数对数面积比LAR_c(i)量化表

Table 3.1: Quantization of the Log.-Area Ratios LAR(i)

LAR序号i	A(i)	B(i)	LAR _c (i)取值范围
1	20.000	0.000	-32~+31
2	20.000	0.000	-32~+31
3	20.000	4.000	-16~+15
4	20.000	-5.000	-16~+15
5	13.637	0.184	-8~+7

6	15.000	-3.500	-8~+7
7	8.334	-0.666	-4~+3
8	8.824	-2.235	-4~+3

3. 短时分析滤波

短时分析滤波对经过预处理和分帧后的语音信号进行短时分析滤波。去除掉语音之间的冗余成分。从语音发音原理上讲，这一部分就是利用语音信号的实际采样值，通过分析计算出语音的激励源数据的过程。经过短时分析滤波所得到的信号称为短时滤波余量信号。

由于在LPC分析过程中利用的是由帧内自相关数据计算出的反射系数，为了防止在语音帧交替过程中因滤波参数发生突变而出现的短暂寄生现象，需要对反射系数进行处理。处理的过程包括反射系数量化编码的解码、对数域反射系数插值^[14]及对数域反变换。短期分析滤波器采用格形滤波器^[15]，格型结构的滤波器是由FIR的基本结构演变而来的，广泛应用在信号建模、谱估计和自适应滤波中，它是受自回归信号建模理论的启发而建立的。采用这种形式的滤波器，是由于格型滤波器的模块化结构便于实现高速并行处理，并且它对有限长的舍入误差不灵敏。其结构图如图3-4所示。

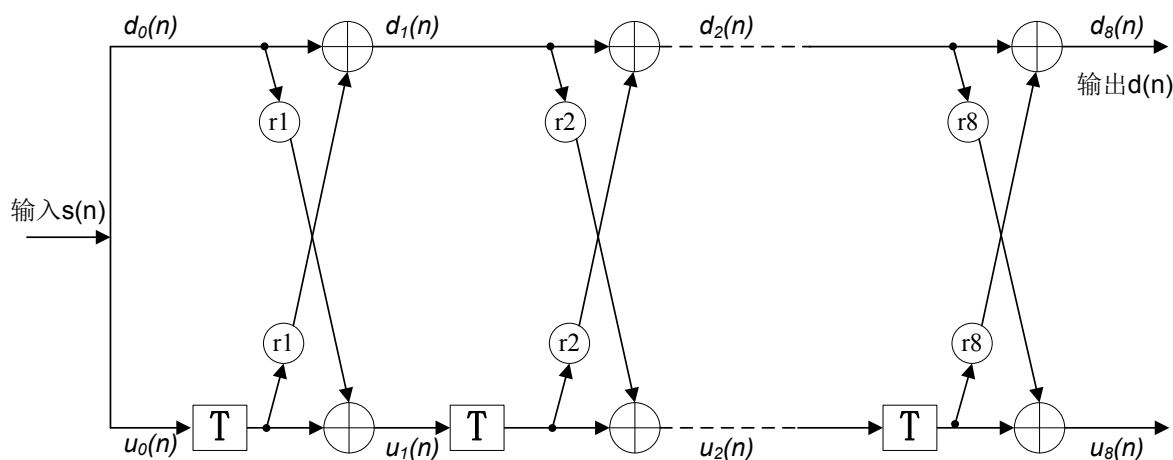


图3-4 短时分析滤波器
Figure 3.4 Short term analysis filter

其数学表达示为：

$$\left\{ \begin{array}{l} d_0(k) = s(k) \quad (3-15a) \\ u_0(k) = s(k) \quad (3-15b) \\ d_i(k) = d_{i-1}(k) + r'_i \times u_{i-1}(k-1) \quad , \quad i = 1 \cdots 8 \quad (3-15c) \\ u_i(k) = u_{i-1}(k-1) + r'_i \times d_{i-1}(k) \quad , \quad i = 1 \cdots 8 \quad (3-15d) \\ d(k) = d_8(k) \quad (3-15e) \end{array} \right.$$

其中 r' 为插值后重构的反射系数。计算 r' 的过程如下：

(1) 对LARc反量化解码：

$$LAR''(i) = \frac{LAR_c(i) - B(i)}{A(i)} \quad , \quad i = 0, \cdots, 7 \quad (3-16)$$

(2) 按照表3-2进行反射系数间的插值

表3-2：当前帧LAR系数的插值

Table 2.2: Interpolation of LAR parameters (J=actual segment)

k	$LAR'_j(i)$
0...12	$0.75 \times LAR_{j-1}''(i) + 0.25 \times LAR_j''(i)$
13...26	$0.50 \times LAR_{j-1}''(i) + 0.50 \times LAR_j''(i)$
27...39	$0.25 \times LAR_{j-1}''(i) + 0.75 \times LAR_j''(i)$
40...159	$LAR_j''(i)$

(3) 求对数域的反变换得到 r' ：

$$r'(i) = \begin{cases} LAR'(i) & |LAR'(i)| < 0.675 \quad (3-16a,b,c) \\ sign[LAR'(i)] \times [0.500 \times |LAR'(i)| + 0.337500] & 0.675 \leq |r(i)| < 1.225 \\ sign[LAR'(i)] \times [0.125 \times |LAR'(i)| + 0.796875] & 1.225 \leq |r(i)| < 1.625 \end{cases}$$

4. 长时预测分析

语音信号 $s(n)$ 经短时预测滤波之后，其残差信号 $d(n)$ 进入长时预测分析，以进一

步去除信号的冗余度。在 5ms 内，我们可以认为图 2-1 中信号的基音周期和激励增益等特征参数没有改变，因而在长时预测中我们将短期滤波余量信号按每 40 个为一组划分子帧，在子帧间利用语音信号在短时间内具有相关性的特点进行长期预测分析。预测是根据先前子帧的重构短期余量信号与当前子帧短期滤波余量信号的互相关，寻找出相关性最好的重构短期余量信号的位置及相应的预测增益并计算出预测值，最后将当前子帧的短期余量信号与预测值相减，得出长期预测余量。

长时预测分析分为两部分：第一部分称为长时分析，用于估计增益参数 b 和最佳延时样点数 N ；第二部分称为长时预测，通过恢复出的本子帧量化预测系数 N' ， b' 和以前的 120 个短时重构残差信号了对当前子帧的短时残差信号进行预测，然后用本子帧的 40 个短时残差减去 40 个短时预测残差得到 40 个长时残差信号。

(1) 子帧划分

每个帧划分为 4 个子帧，每个子帧 40 个样点。为叙述方便，设子帧号 $j=0,1,2,3$ ，设每一帧的第一个样点序号为 k_0 ，则该帧的样点序号表示为 $n=k_0+k_j$ ，其中 $k_j=40*j$ 。

(2) LTP 参数计算

一帧中的四个子帧分别对应一个长时相关延时 $N_j(j=0, \dots, 3)$ 和一个相应的增益因子 $b_j(j=0, \dots, 3)$ 。计算方法如下：

1) 计算当前帧短时残差信号 $d(k_j+i)$ 与重构的前三个子帧 $d'(k_j+i')$ 的互相关函数 $R_j(\lambda)$ ，其中 $i'=(-1, \dots, -120)$ ，用 i 表示为 $i'=i-\lambda$ ，延迟因子 $40 \leq \lambda \leq 120$ ，即保证了在本帧外而又不超过前三个子帧。

$$R_j(\lambda) = \sum_{i=0}^{39} d(k_j+i) \times d'(k_j+i-\lambda) \quad (3-17)$$

其中 $j=0, \dots, 3$ ； $k_j= k_0+40*j$ ， $\lambda =40, \dots, 120$ 。 $d'(k_j+i')$ 的计算方法将下文长时合成滤波器中说明。

2) 找出重构样点与该子帧相关性最大的位置：

$$R_j(N_j) = \max \{R_j(\lambda)\} \quad (3-18)$$

$j=0, \dots, 3$ ； $\lambda =40, \dots, 120$

3) 计算增益因子：

$$\left\{ \begin{array}{l} b_j = \frac{R_j(N_j)}{S_j(N_j)} \quad j=0, \dots, 3 \quad (3-19a) \\ S_j(N_j) = \sum_{i=0}^{39} d'^2(k_j + i - N_j) \quad j=0, \dots, 3 \quad (3-19b) \end{array} \right.$$

4) LTP参数的编解码

➤ 长时相关延迟 $N_j(j=0, \dots, 3)$ 可以在(40, ..., 120)的范围内取值, 用7bit对它进行编码:

$$N_c = N_j \quad (3-20)$$

假设传输无差错则解码端:

$$N'_j = N_c \quad (3-20)$$

➤ 长时预测增益 $b_j(j=0, \dots, 3)$ 按下面的算法进行2bit编码:

$$\left\{ \begin{array}{l} b_j \leq DLB(i), i = 0, \text{ 则 } b_c = 0 \\ DLB(i-1) < b_j \leq DLB(i), i = 1,2, \text{ 则 } b_c = i \\ DLB(i) < b_j, i = 3, \text{ 则 } b_c = 3 \end{array} \right. \quad (3-21)$$

在解码端:

$$b'_j = QLB(b_c), j = 0, \dots, 3 \quad (3-22)$$

精度和量化电平如表2.3所示。

表3-3: LTP增益量化表

Table3.3:Quantization table for the LTP gain

i	判决电平DLB(i)	量化电平QLB(i)
0	0.2	0.10
1	0.5	0.35
2	0.8	0.65
3		1.00

(3) 长时分析滤波器

长时分析滤波器的作用是计算长时残差信号, 算法如下:

$$e(k_j + i) = d(k_j + i) - d''(k_j + i) \quad (3-23)$$

式中 $d''(k_j + i)$ 表示残差信号的预测值, 在减法前应先算出, 它是通过对重构残差

信号 $d'(k_j + i - N'_j)$ 调整和加权得到的。

$$d'(k_j + i) = b'_j \times d''(k_j + i - N'_j) \quad (3-24)$$

其中 N'_j 和 b'_j 是经过编码再解码后的 N_j 和 b_j 值。

(4) 长时合成滤波器

在每个子帧内，计算重构残差信号 $d'(k_j + i)$ ：

$$d'(k_j + i) = e'(k_j + i) + d''(k_j + i) \quad (3-25)$$

式中 $e'(k_j + i)$ 是 $e(k_j + i)$ 经过量化编码再经译码重构的长时残差信号，其计算方法将在下节中详细说明。

5. 规则脉冲编码

规则脉冲编码的作用是选择最合适的激励脉冲序列。长期预测所得信号经过加权滤波后按照三选一进行下采样分组，一共可得相位不同的4个子抽样序列，如图3-5所示。

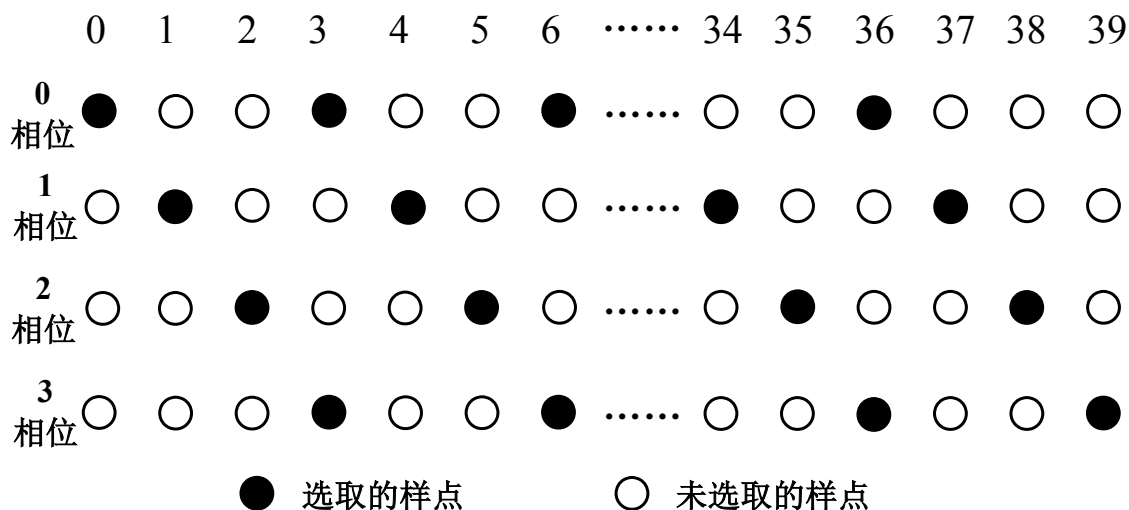


图3-5 规则脉冲序列相位示意图

Figure 3.5 The Posture fig of regular impulsion pulse sequence

规则脉冲编码就是选择出其中能量最大的一组并对其进行编码，用作重构语音时的激励脉冲序列。经过短时、长时预测之后得到的LP残差信号，在这里进行加权滤波、规则脉冲序列抽取和量化编码。

(1) 加权滤波器

运用FIR“块滤波”^[16]算法计算每子帧(40个样点)的 $e(k)$ 与脉冲冲击相应 $h(i)$ 的卷积,其中 $i=0, \dots, 10$,具体响应见表3-4。

表3-4 加重滤波器的脉冲冲击响应

Table 3.4: Impulse response of block filter (weighting filter)

i	5	4(6)	3(7)	2(8)	1(9)	0(10)
$H(i)*2^{13}$	8192	5741	2054	0	-374	-134

$$|H(\Omega)|=2.779;$$

“块滤波”算法只产生常规卷积运算50个样点的40个中心样点,其算法如下:

$$x(k) = \sum_{i=0}^{10} h(i)e(k+5-i), k=0, \dots, 39 \quad (3-26)$$

当 $k+5-i < 0$ 或 $k+5-i > 39$ 时使 $e(k+5-i)=0$ 。

(2) 基于RPE坐标选取的自适应抽样

对滤波信号 x 进行3:1的等间隔抽样,以 $m=0, \dots, 3$ 为抽样起点,共出现四种抽样序列,每个序列含有13个非零样值:

$$x_m(i) = x(k_j + m + 3 \times i), i=0, \dots, 12 \quad (3-27)$$

根据均方差准则,被选取的序列应具有最大能量,即

$$E_m = \max \sum_{i=0}^{10} x_m^2(i), m=0, \dots, 3 \quad (3-28)$$

这样每个子帧的40个长时余量信号就被一个13脉冲长的序列代替,并将位置信息 M 用2bit编码成 M_c 以在解码端恢复长时余量信号。

(3) RPE序列的APCM量化

对选定的脉冲子序列进行APCM方式进行量化,由于每个RPE序列都由13个样点 $x_m(i)$ 组成,先找出 $x_m(i)$ 的绝对值 $|x_m(i)|$ 中的最大值 x_{\max} ,对 x_{\max} 用6bit进行量化,得到量化码 $x_{\max c}$,对 $x_{\max c}$ 解码后得到 x'_{\max} , x_{\max} 的量化表见附录A-2。

(4) 归一化

将RPE序列的13个样点分别除以 x'_{\max} 取归一化样点 $x'(i)$:

$$x'(i) = \frac{x_m(i)}{x'_{\max}}, i=0, \dots, 10 \quad (3-29)$$

并按照表3-5进行量化得到 $x_{mc}(i)$ 。

表3-5 归一化RPE样点的量化

Table 3.5: Quantization of the normalized RPE-samples

$x' \times 2^{15}$ 间隔	$x'_m \times 2^{15}$	x_{mc}
-32768...-24577	-28762	0
-24576...-16385	-20480	1
-16384...-8193	-12288	2
-8192...-1	-4096	3
0...8192	4096	4
8192...16383	12288	5
16383...24575	20480	6
24575...32767	28762	7

(5) 长时残差信号的重构

经上述步骤产生的量化参数 M_c ， $x_{\max c}$ 以及 x_{mc} 将分为两路，一路送到接收端用以恢复语音信号；另一路经解码产生重构长时余量信号。先解码去归一化得到重构RPE序列 x'' ：

$$x''(i) = x'_m(i) \times x'_{\max c}, i = 0, \dots, 12 \quad (3-30)$$

然后根据位置参数 M_c 对 x'' 进行插零值，以得到重构长时余量信号。

$$\begin{cases} e'(k_j + k) = x''(i) & k = M_c + 3 \cdot i & (3-31a) \\ e'(k_j + k) = 0 & k \text{ 为其它值} & (3-31b) \end{cases}$$

(二) 解码原理

图 3-6 是 RPE-LTP 解码器的原理图，图中 M_c ， $x_{\max c}$ ， x_{mc} ， b_c ， N_c ， LAR_c 等表示解码器收到的信息码。解码器工作过程与编码器相反，先得到量化的规则脉冲序列 e' ，再得到重建短时残差信号 d' ，然后让 d' 通过综合滤波器恢复语音信号，最后去加重得到解码声音信号。

解码器主要由下面四部分组成，大部分的子块算法与编码器的相似，只有短时合成滤波器和去加重滤波器作为新的子块加到解码器中。合成滤波器也采用格型结构，只是分析滤波器是一个全零点系统，而合成滤波器是一个全极点的滤波器。

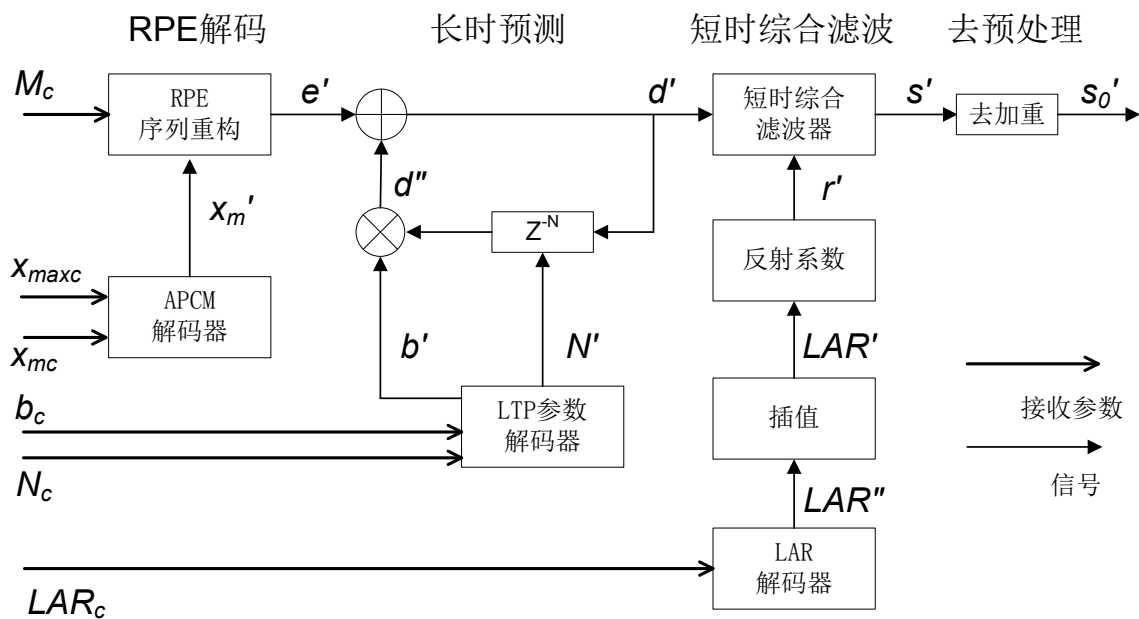


图3-6 RPE-LTP解码器原理框图
Figure 3.6: Block diagram of the RPE-LTP decoder

1. RPE 编码

长时合成滤波器的输入信号可通过 RPE 解码、去归一化和定位来完成。在此阶段，先根据 x_{maxc} 和 x_{mc} 去归一化，再根据 M_c 在一定的位位置插 0，恢复信号的取样频率。

2. 长时预测

重建的长时残差信号 e' 被输入到长时合成滤波器，根据输入的增益参数 b_c 和位置参数 N_c 产生重建的短时残差信号 d' ，并把四个子帧合并为一帧，作为短时合成滤波器的输入。

3. 短时合成滤波

$$\begin{cases} s_0(k) = d'(k) & (3-32a) \\ s_i(k) = s_{i-1}(k) - r'(9-i) \times v_{8-i}(k-1) & i=1, \dots, 8 \quad (3-32b) \\ v_{9-i}(k) = v_{8-i}(k-1) + r'(9-i) \times s_i(k) & i=1, \dots, 8 \quad (3-32c) \\ s'_i(k) = s_8(k) & (3-32d) \\ v_0(k) = s_8(k) & (3-32e) \end{cases}$$

图3-7为短时合成滤波器的格型结构图，其解法如式3-32。

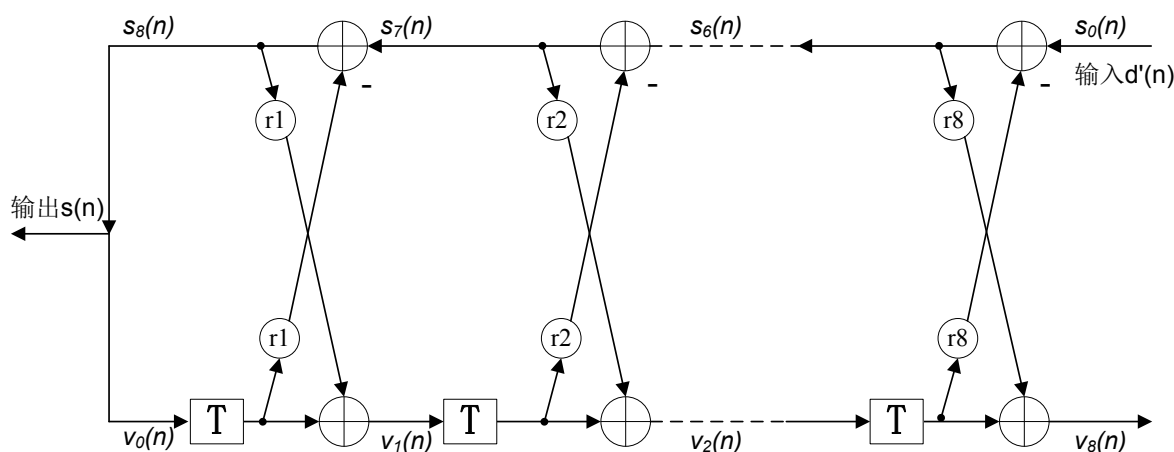


图3-7 短时合成滤波器
Figure 3.7 Short term synthesis filter

4. 后处理

将 $s'_i(k)$ 去加重得到输出信号 $s'_0(k)$ 。

$$S_0(k) = S'(k) + \beta \times S_0(k-1) \quad (3-33)$$

$$\beta = 28180 \times 2^{-15} = 0.8599853$$

(三) 比特分配

GSM 全速率 RPE-LTP 语音编码方案中各参数所用的比特分配如表 2.7 所示，语音帧长为 20ms，每帧比特数为 260bits，故总码率为 13kbps。

表 3-6 GSM 13kbit/s RPE-LTP 编码方案比特分配情况

Table 3.6 The bit assignment of GSM 13kbit/s RPE-LTP coder

参数	Bit/每帧
8 个 LP 参数 $LARc$	6+6+5+5+4+4+3+3=36
4 个 LTP 增益参数 bc	2*4=8
4 个 LTP 时延参数 Nc	7*4=28
4 个 RPE 序列相位参数 Mc	2*4=8
4 个子帧最大非零样值 $Xmaxc$	6*4=24
52 个 RPE-LTP 样值 Xm	3*13*4=156
总计	260

(四) 结果评价

主要是针对声音质量的评价。先用一段语音信号作为输入进行测试，主观评价其还原的语音质量；再用一个 320 点（2 帧）的样本进行测试，对其编解码前后的时、频域波形进行比较，分析算法还原输入信号的质量以及存在的问题。图 3-8 是一个 320 点的样本输入和输出的时域波形比较，上面为输入信号，下面为输出信号。

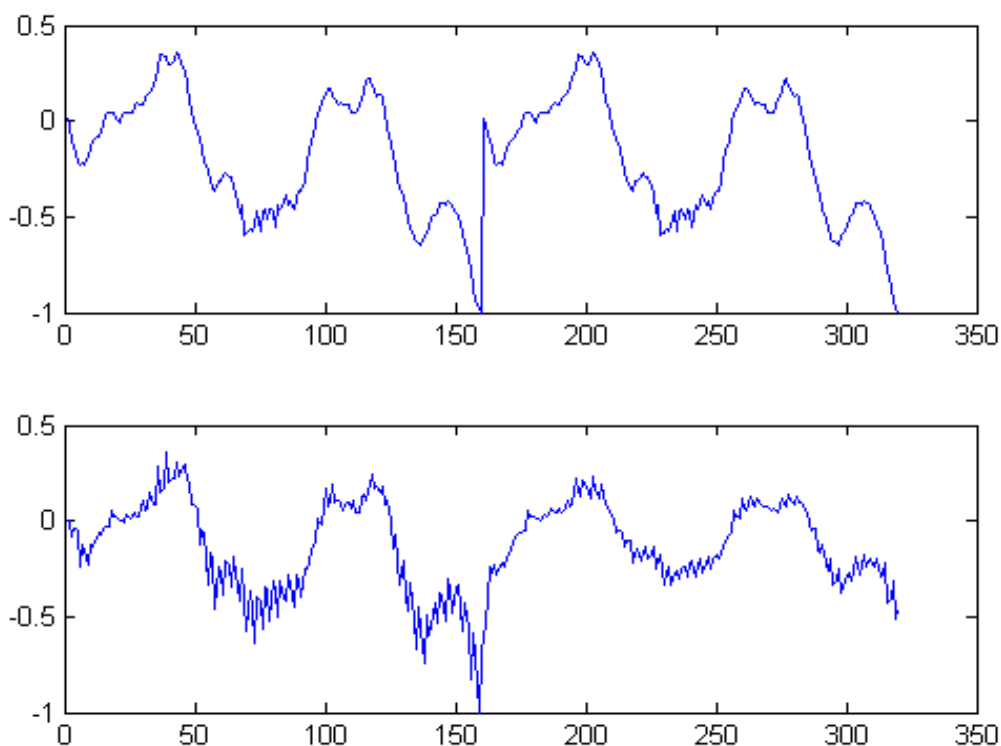


图3-8 输入信号与编解码输出信号时域对比

Fig.3.8 The compare of input signal and output signal in time domain

由上图时域分析可以看出，该算法能够较为准确地还原信号，但是出现了高频失真，是因为在编解码过程中使用了变采样，由于滤波不彻底，产生了高频镜像，造成高频部分被抬高。并且由于选取样本的前后两帧是一样的，可以明显地从图中看出，前一帧的还原效果优于后一帧，这是由于第一帧并未采用长时帧间预测，后一帧在长时预测后的还原效果不理想，长时预测部分的算法还需要进一步进行改进。

在分析完时域的基础上，再对频域进行对比。为了进行较为准确的频域分析，采用音频处理软件 Cool Edit Pro 2.1 进行处理。时域到频域的变换采用 1024 点 FFT，加的窗为汉明窗。其频域对比图如图 3-9 所示。

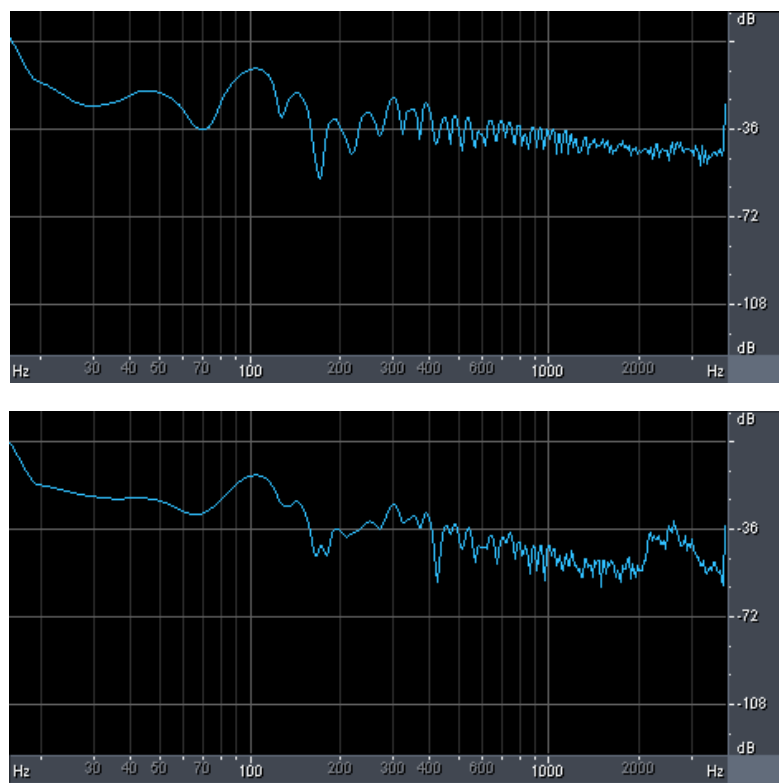


图3-9 输入信号与编解码输出信号频域对比

Fig.3.9 The compare of input signal and output signal in frequency domain

从图中可以看出，输出信号在 2000Hz 以上有一个明显的失真，这也验证了在时域中看到的高频失真是由于变采样造成的。

结论

由测试结果可知，RPE-LTP 算法能够得到懂性很强，自然度较高的还原声音质量，完全可以满足正常移动通信的要求。并且由于自身用 MATLAB 所实现的 RPE-LTP 算法虽然基本上可以得到与标准较为接近的音质，但是在一些细节处理上，如变采样的处理，以及长时预测的参数处理，效果并不是很理想，因而与标准还存在一些差距，音质还有提高的空间。

在本设计中，主要做了以下几个方面的工作：

- ◇ 由于研究的方向是声码器，因而对声码器做了广泛的调研，并且对几种主流算法进行了细致的分析和比较，结合自身情况，选定较为合适的 RPE-LTP 算法进行研究。
- ◇ 对 RPE-LTP 算法进行系统的、深入的研究，结合数字信号处理和语音信号处理的相关文献书籍，掌握标准中的算法思路和原理。在此过程中提高了对语音和滤波器方面的认识，并熟悉了声码器的几个通用原理。
- ◇ 用 MATLAB 实现算法。在此过程中遇到了很多问题，有软写编程上的，也有原理理解上的，经过不断的改进对算法有了更深的认识，语音也在逐渐改善。
- ◇ 语音质量的评测。

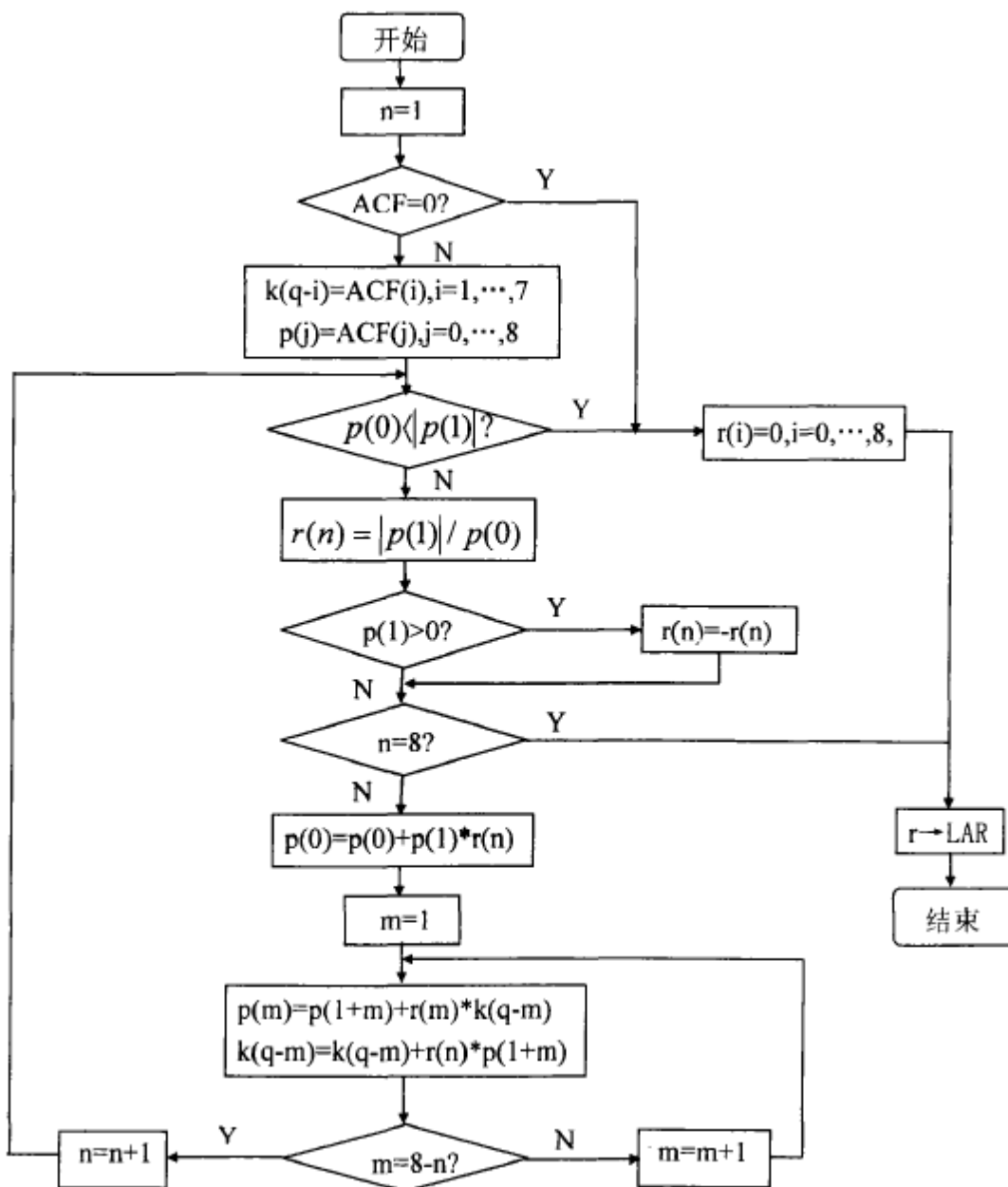
由于对声码器的研究处于刚刚起步的阶段，对其原理的理解还不够深入和系统，目前实现的算法还存在不足，并且由于没有进行硬件实验，无法得知在实际应用中可能会出现的问题，因而还需要进一步进行研究和改进。

参考文献

- [1] 李昌立, 吴善培. 数字语音—语音编码实用教程[M]. 北京: 人民邮电出版社, 2004.2.
- [2] 赛迪网 . 国际电信联盟 : 目前全球手机用户达到 41 亿 [EB/OL] . http://news.ccidnet.com/art/1032/20090303/1694841_1.html
- [3] 许霭云, 吴坚. GSM 和 CDMA 中的语音编码技术[J]. 江苏通信技术, 1999, (5): 11-17.
- [4] ETSI. ETSI GSM 06.10: Digital cellular telecommunications system (Phase2+); Full rate speech; Transcoding [EB/OL]. <http://www.etsi.org>, Nov 2000.
- [5] 刘波. MATLAB 信号处理[M]. 北京: 电子工业出版社, 2006.
- [6] 鲍长春. 数字语音编码原理[M]. 西安: 西安电子科技大海军出版社, 2007.1.
- [7] 丛键, 张知易. 一种 600bps 极低速率语音编码算法[J]. 电子与信息学报, 2007, (2): 429-432.
- [8] 王洪, 唐凯. 低速率语音编码[M]. 北京: 国防工业出版社, 2006.5.8.
- [9] Barrett, P. A. and Rix, A. W. Applications of speech quality measurement for 3G[C]. Third International Conference on 3G Mobile Communication Technologies. May,2002,8-10.
- [10] 赵力. 语音信号处理[M]. 北京: 机械工业出版社, 2003.21.26.
- [11] 郑敏, 温启荣. 泛欧 GSM 13kb/s RPE-LTP 编译码算法探讨[J]. 重庆邮电学院学报, 1997, (1): 18-22.
- [12] McCart, R.Connell, J.A . comparison of LSF quantization techniques for 3G mobile communications[C],Irish Signals and System Conference 2004,2004,633-638.
- [13] Miguel A.Ferrer-Ballester, Anibal R.Figueiras.Vidal. Efficient Adaptive Vector Quantization of LPC Parameters[J]. IEEE Transactions on Speech and Audio Processing, 1995, (3): 3-14.
- [14] 鲍长春. 一种有效的典型波形块内插方法[J]. 电子科学学刊, 1999, (21) : 734-751.
- [15] Oppenheim, A.V, and R. W.Schafer . Discrete-Time Signal Processing[M]. Prentice-Hall, 1989.263.271.
- [16] 程佩青. 数字信号处理教程[M]. 北京: 清华大学出版社, 2001.204.205.

附录

附图 A-1 利用 Schur 递归算法进行 LPC 分析流程图



附表A-2 x_{\max} 量化表

x_{\max}	x'_{\max}	$x_{\max c}$	x_{\max}	x'_{\max}	$x_{\max c}$
0...31	31	0	2048...2303	2303	32
32...63	63	1	2304...2559	2559	33
64...95	95	2	2560...2815	2815	34
96...127	127	3	2816...3071	3071	35
128...159	159	4	3072...3327	3327	36
160...191	191	5	3328...3583	583	37
192...223	223	6	3584...3839	3839	38
224...255	255	7	3840...4095	4095	39
256...287	287	8	4096...4608	4608	40
288...319	319	9	4609...5119	5119	41
320...351	351	10	5120...5631	5631	42
352...383	383	11	5632...6143	6143	43
384...415	415	12	6144...6655	6655	44
416...447	447	13	6656...7167	7167	45
448...479	479	14	7168...7679	7679	46
480...511	511	15	7680...8191	8191	47
512...575	575	16	8192...9215	9215	48
576...639	639	17	9216...10239	10239	49
640...703	703	18	12240...11263	11263	50
704...767	767	19	11264...12287	12287	51
768...831	831	20	12288...13311	13311	52
832...895	895	21	13312...14335	14335	53
896...959	959	22	14336...15359	15359	54
960...1023	1023	23	15360...16383	16383	55
1024...1151	1151	24	16384...18431	18431	56
1152...1279	1279	25	18432...20479	20479	57
1280...1407	1407	26	20480...22527	22527	58
1408...1535	1535	27	22528...24575	24575	59
1536...1663	1663	28	24576...26623	26623	60
1664...1791	1791	29	26624...28671	28671	61
1792...1919	1919	30	28672...30719	30719	62
1920...2047	2047	31	30720...32767	32767	63

后记

在本课题的整个完成过程中，我收获了很多。新颖的课题让我对其充满好奇并积极探索，大量的专业知识需求也让我对通信原理，数字信号处理以及语音信号处理等课程有了更加深入的学习和理解，是对我四年学习的一个综合检验。在整个过程中，我对课题的理解从无到有，从有到优，算法一遍又一遍的修改和优化，结果也越来越好。

设计最终取得的结果比较让人满意，这与学校提供的优越的实验室环境是密不可分的。特别要感谢我的辅导老师杜伟韬老师，他循循善诱，对我悉心指导，使我对语音信号、数字信号等基本概念有了深入的理解，并教会了我算法开发的整个流程，以及很多经验心得，让我少走了许多弯路。在我遇到难题时，杜老师总是稍加点拨，而不是合盘托出，让我养成了独立钻研的好习惯，对问题的理解也更加深刻。杜老师在工作上严谨的作风和钻研的精神会让我受益终身。

感谢所有曾经给我帮助的老师，你们的辛勤教导我会铭记在心。

感谢跟我一起并肩作战，走过四年光阴的同学朋友们，你们的鼓励和帮助是我莫大的财富。

感谢我的大学，希望我能送给你一个完美的送别礼物。