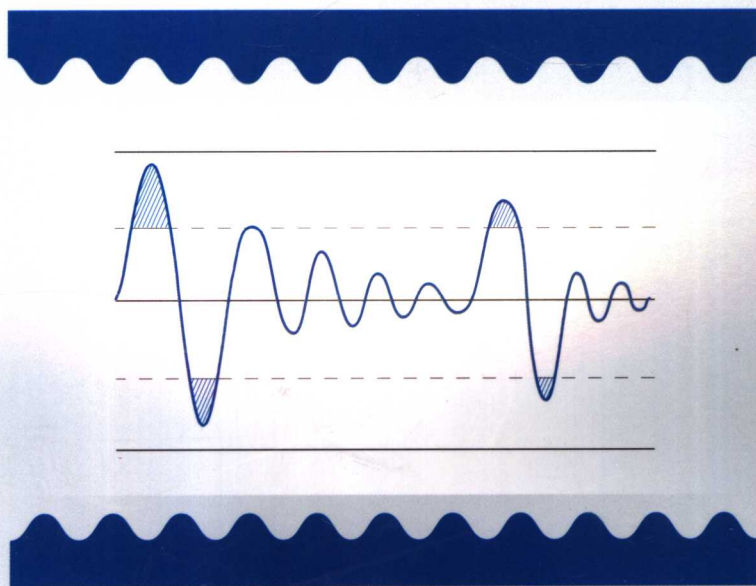


低速率语音编码

Low Rate Speech Coding

■ 王洪 唐凯 编著 ■



国防工业出版社

National Defense Industry Press

低速率语音编码

Low Rate Speech Coding

王洪 唐凯 编著

国防工业出版社

· 北京 ·

内 容 简 介

本书依据低速率语音编码的算法体制作作为研究线索,以编码速率作为指引,以编码国际标准作为具体研究对象而深入研究与讨论,介绍了目前最新的低速率语音编码国际标准和低速率语音编码新技术,同时介绍了低速率语音编码的分析方法。

本书的特点是讲解了目前最前沿的低速率语音编码技术,反映了语音编码技术的最新水平与发展趋势,具有与本学科学术水平相适应的先进性;理清了研究、分析低速率语音编码技术的工作头绪,吸纳了众家之成果;提供了大量的实际参数、图表,与实际工作联系紧密,具有很强的可操作性与适用性,同时,通过我们在书中反映出来的研究过程,读者能够在以后的研究工作中得到很大的启发。

本书可作为大学本科生高年级、研究生的选修课本或参考书,科研人员的参考书。尤其适合从事语音编码研究与分析工作的工程技术人员和科研人员作为参考书。

图书在版编目(CIP)数据

低速率语音编码 / 王洪, 唐凯编著. —北京: 国防工业出版社, 2006. 2

ISBN 7-118-04345-1

I. 低... II. ①王... ②唐... III. 语音数据处理—编码 IV. TN912. 3

中国版本图书馆 CIP 数据核字(2006)第 006613 号

※

国防工业出版社出版发行

(北京市海淀区紫竹院南路 23 号 邮政编码 100044)

腾飞胶印厂印刷

新华书店经售

*

开本 787×1092 1/16 印张 15 $\frac{3}{4}$ 字数 352 千字

2006 年 2 月第 1 版第 1 次印刷 印数 1—4000 册 定价 30.00 元

(本书如有印装错误, 我社负责调换)

国防书店: (010) 68428422

发行邮购: (010) 68414474

发行传真: (010) 68411535

发行业务: (010) 68472764

前 言

语音编码技术从 20 世纪 30 年代末开始,经历了不到 70 年的探索与研究,已经取得了无数令人惊叹的成果。尤其是到了 20 世纪 80 年代后期和 90 年代,计算机技术、大规模集成电路技术以及现代通信技术的迅猛发展极大地带动了语音编码技术的进步。

随着军事通信、移动通信和因特网的飞速发展,语音通信技术也在不断地进行更新并与之相融合。语音编码作为数字语音通信过程中的重要组成部分,它在总体上有 3 个大的发展方向:低速率语音编码、变速率语音编码和不压缩语音编码。

低速率语音编码是未来语音编码的主要发展方向,这是因为在现代通信中,信道利用的有效性和经济性仍是研究的重要目标,低速率语音编码技术是语音通信中不可缺少的一个重要研究方向。在实际通信中,有些信道难以扩宽且质量很差,如短波信道;有的信道正被广泛使用,短期内难以更新;有些特殊条件下的通信,如强干扰下的军事通信等;还有些昂贵的信道,节省的每一个比特都意味着经济上的好处。在这些情况下,低速率语音编码极具吸引力。随着数字通信业务的发展,高质量的低速率语音编码技术将成为研究的热点,它必将在保密通信、语音邮件、网络通信、IP 电话等领域有广泛的应用前景,特别是在信息化战场上广泛应用于单兵作战通信设备上。

低速率语音编码的新技术和新算法却是层出不穷。经常会有不同的国际组织推出不同算法标准的低速率语音编码新技术,这就给研究和分析低速率语音编码技术的工程人员带来了困难。如何从纷繁复杂的低速率语音编码新技术中理出头绪,从总体上把握和跟踪低速率语音编码技术的发展呢?作者经过阅读大量的参考文献和自己对低速率语音编码技术的研究与分析工作,总结出了一套研究和分析低速率语音编码技术的方法。那就是依据低速率语音编码的算法体制作为研究线索,以编码速率作为指引,以编码国际标准作为具体研究对象而深入研究与讨论。

本书正是在这种方法与思路的指导下,介绍了目前最新的低速率语音编码国际标准和低速率语音编码新技术,同时介绍了低速率语音编码的分析方法。

全书共分 10 章。第一章从总体上简要介绍了语音编码技术的现状与发展方向。第二章介绍了语音编码的基本理论和常用技术,以及语音编码的性能评价方法与常用的码流特征分析方法,为以后各章的具体分析与讨论打下坚实的理论基础。第三章阐述了现代低速率语音编码的研究思路进展情况和低速率语音编码的最新成果,着重介绍了低速率语音编码的算法体制。第四章~第八章以低速率语音编码的算法体制作为研究线索,以编码速率作为指引,介绍了 5 种常用的低速率语音编码国际标准算法的编译码原理,同时对这些算法的性能进行了测试以及编码后的码流特征进行了分析。为了完善低速率语音编码的算法体制研究,第九章和第十章分别介绍了谐波激励线性预测语音编码和波形

内插语音编码算法的编译码原理。

书后附有“英汉名词术语对照”、“常见语音编码标准码流的零—变化率直方图”和“常见语音编码标准码流的帧结构比较表”，具有很高的实用价值。最后列出了全书的“参考文献”。

本书的特点是讲解了目前最前沿的低速率语音编码技术，反映了语音编码技术的最新水平与发展趋势，具有与本学科学术水平相适应的先进性；理清了研究、分析低速率语音编码技术的工作头绪，吸纳了众家之成果；提供了大量的实际参数、图表，与实际工作联系紧密，具有很强的可操作性与适用性。同时，通过我们在书中反映出来的研究过程，读者能够从中得到很大的启发。

由于时间比较仓促，而且低速率语音编码技术还处于不断地发展之中，因此，书中难免有不妥之处，恳请读者批评指正。同时，书中引用了大量的文献资料，在此向原作者表示深深的谢意。

作者

2005年10月

目 录

第一章 语音编码概览	1
1.1 引言	1
1.2 语音编码总体发展方向	1
1.3 语音编码分类	2
1.3.1 按语音信号带宽分类	2
1.3.2 按语音编码方法分类	3
1.3.3 按语音编码处理域分类	4
1.3.4 按语音编码速率分类	4
1.4 语音编码国际标准	5
1.4.1 高速率语音编码国际标准	5
1.4.2 中速率语音编码国际标准	7
1.4.3 低速率语音编码国际标准	7
1.4.4 变速率语音编码国际标准	8
1.5 两类语音编码应用系统	9
1.6 现代通信系统中的语音编码技术	9
1.6.1 移动通信系统中的语音编码技术	9
1.6.2 因特网中的语音编码技术	10
1.7 小结	11
第二章 语音编码研究与分析基础	12
2.1 语音编码理论基础	12
2.1.1 基于语音的信息传递	12
2.1.2 人类的语音生成和语音感知	13
2.1.3 语音信号数学模型	14
2.1.4 语音编码通用模型	16
2.1.5 语音可压缩编码的依据	17
2.1.6 语音编码的极限速率	18
2.2 语音编码常用的基本技术	19
2.2.1 语音信号数字化	19
2.2.2 预处理技术	20
2.2.3 线性预测技术	21
2.2.4 矢量量化技术	25
2.2.5 基音周期估计	28

2.2.6	合成分析法	30
2.2.7	感知加权滤波器	31
2.3	语音编码的性能评价方法	33
2.3.1	编码速率	33
2.3.2	合成语音质量	33
2.3.3	编解码延时	38
2.3.4	算法复杂度	38
2.3.5	其它性能	38
2.4	语音编码的码流特征分析方法	39
2.4.1	码流的空间分布特征	39
2.4.2	帧长的提取	40
2.4.3	静音帧的特征	41
2.4.4	零一概率的统计	41
2.4.5	零一变化率的统计	41
2.4.6	游程长度分布的统计	42
2.4.7	参数分布特征	42
2.5	小结	42
第三章	低速率语音编码概述	43
3.1	低速率语音编码的发展与应用	43
3.1.1	低速率语音编码的发展	43
3.1.2	低速率语音编码的应用	44
3.2	低速率语音编码新进展	45
3.2.1	低速率语音编码的研究思路进展情况	45
3.2.2	低速率语音编码的最新成果	46
3.3	低速率语音编码的算法体制	49
3.3.1	码激励线性预测模型	49
3.3.2	混合激励线性预测模型	50
3.3.3	正弦激励模型	51
3.3.4	插值编码模型	52
3.4	小结	53
第四章	2.4Kb/s 二元激励线性预测语音编码	54
4.1	概述	54
4.2	LPC-10 声码器编码原理	54
4.2.1	预处理	55
4.2.2	反射系数的计算与编码	55
4.2.3	增益均方根的计算与编码	57
4.2.4	基音周期的计算与编码	58
4.2.5	清/浊音的判决与编码	59
4.2.6	编码参数的比特分配与发送	59

4.3	LPC-10 声码器译码原理	60
4.3.1	参数译码与插值	60
4.3.2	激励源的获取	61
4.3.3	语音合成	62
4.4	LPC-10e 声码器的改进	62
4.4.1	激励源的改进	62
4.4.2	基音周期计算方法的改进	64
4.4.3	声道滤波器参数量化的改进	64
4.5	合成语音质量测试与模拟结果	66
4.5.1	合成语音质量测试	67
4.5.2	模拟结果	67
4.6	码流特征分析	68
4.6.1	帧长与编码比特结构分析	69
4.6.2	静音帧的特征分析	70
4.6.3	零一概率和游程长度分布的统计	70
4.6.4	零一变化率的统计和参数分布特征分析	70
4.7	小结	71
第五章	4.8Kb/s 码激励线性预测语音编码	72
5.1	概述	72
5.2	FS-1016 标准编码原理	72
5.2.1	编码框图与编码过程	73
5.2.2	线性预测分析与量化	73
5.2.3	自适应码本搜索与编码	75
5.2.4	固定码本搜索与编码	79
5.2.5	编码参数的比特分配	80
5.2.6	FS-1016 标准编码特点	80
5.3	FS-1016 标准译码原理	81
5.3.1	译码框图与译码过程	81
5.3.2	后置滤波与语音合成	82
5.4	FS-1016 标准算法的进一步研究	82
5.4.1	感知加权滤波器的应用	82
5.4.2	利用递归的快速算法	83
5.5	合成语音质量测试与模拟结果	85
5.5.1	合成语音质量测试	86
5.5.2	模拟结果	86
5.6	码流特征分析	87
5.6.1	帧长与编码比特结构分析	88
5.6.2	静音帧的特征分析	89
5.6.3	零一概率和游程长度分布的统计	89

5.6.4	零一变化率的统计和参数分布特征分析	89
5.7	小结	90
第六章	4.75Kb/s 代数码激励线性预测语音编码	91
6.1	概述	91
6.2	AMR 标准 4.75Kb/s 编码原理	91
6.2.1	预处理	92
6.2.2	线性预测分析与量化	92
6.2.3	开环基音分析	104
6.2.4	脉冲响应的计算	106
6.2.5	目标信号的计算	106
6.2.6	自适应码本搜索和增益控制	106
6.2.7	代数码本的结构和搜索	108
6.2.8	自适应码本增益和固定码本增益的量化	111
6.2.9	修改存储器	112
6.2.10	编码参数的比特分配	112
6.3	AMR 标准 4.75Kb/s 译码原理	113
6.3.1	译码与语音合成	114
6.3.2	后置处理	115
6.4	合成语音质量测试与模拟结果	116
6.4.1	合成语音质量测试	116
6.4.2	模拟结果	117
6.5	码流特征分析	118
6.5.1	帧长与编码比特结构分析	119
6.5.2	静音帧的特征分析	119
6.5.3	零一概率和游程长度分布的统计	119
6.5.4	零一变化率的统计和参数分布特征分析	120
6.6	小结	121
第七章	2.4Kb/s 混合激励线性预测语音编码	122
7.1	概述	122
7.2	MELP 声码器编码原理	123
7.2.1	预处理	123
7.2.2	基音周期的计算与量化	123
7.2.3	子带声音强度的计算与量化	127
7.2.4	增益的计算与量化	129
7.2.5	线性预测分析与量化	130
7.2.6	傅里叶级数幅值的计算与量化	132
7.2.7	纠错编码	133
7.2.8	编码参数的比特分配与发送	134
7.3	MELP 声码器译码原理	135

7.3.1	基音周期译码与纠错译码	135
7.3.2	增益浮码与抑制	135
7.3.3	参数插值	137
7.3.4	混合激励生成与滤波	137
7.3.5	语音合成	139
7.4	MELP 声码器的关键技术	140
7.4.1	多带混合激励	141
7.4.2	使用非周期脉冲	141
7.4.3	残差谐波谱处理技术	141
7.4.4	自适应谱增强技术	142
7.4.5	脉冲整形滤波	142
7.5	合成语音质量测试与模拟结果	143
7.5.1	合成语音质量测试	143
7.5.2	模拟结果	143
7.6	码流特征分析	144
7.6.1	帧长与编码比特结构分析	145
7.6.2	静音帧的特征分析	145
7.6.3	零一概率和游程长度分布的统计	146
7.6.4	零一变化率的统计和参数分布特征分析	146
7.7	小结	147
第八章	4. 15Kb/s 改进型多带激励语音编码	148
8.1	概述	148
8.2	基于 MBE 模型的语音分析与合成	148
8.2.1	MBE 语音产生模型	149
8.2.2	基于 MBE 模型的语音分析原理	149
8.2.3	基音周期的估计	151
8.2.4	清/浊音的判决	160
8.2.5	谱包络参数确定	160
8.2.6	基于 MBE 模型的语音合成	161
8.3	IMBE 声码器编码原理	162
8.3.1	编码基本过程	162
8.3.2	改进型基音周期的估计	162
8.3.3	改进型清/浊音的判决	163
8.3.4	改进型谱包络参数确定	164
8.3.5	编码参数的比特分配	165
8.4	IMBE 声码器译码原理	167
8.4.1	清音部分的合成	167
8.4.2	浊音部分的合成	168
8.4.3	重建语音产生	169

8.5	合成语音质量测试与模拟结果	169
8.5.1	合成语音质量测试	169
8.5.2	模拟结果	170
8.6	码流特征分析	171
8.6.1	帧长与编码比特结构分析	171
8.6.2	静音帧的特征分析	171
8.6.3	零一概率和游程长度分布的统计	172
8.6.4	零一变化率的统计和参数分布特征分析	172
8.7	小结	173
第九章	2.4Kb/s 谐波激励线性预测语音编码	174
9.1	概述	174
9.2	HELP 声码器模型与基本原理	175
9.2.1	正弦语音产生模型的原理	175
9.2.2	HELP 语音模型的原理	178
9.2.3	HELP 语音模型的参数提取	178
9.3	HELP 声码器编码原理	181
9.3.1	线性预测参数的编码	182
9.3.2	基音周期的提取与编码	183
9.3.3	清/浊音的判决与编码	187
9.3.4	谱包络参数的提取与编码	189
9.3.5	编码参数的比特分配	191
9.4	HELP 声码器译码原理	191
9.4.1	残差重建模型	191
9.4.2	相位合成	193
9.5	小结	195
第十章	3.75Kb/s 波形内插语音编码	196
10.1	概述	196
10.2	WI 声码器基本原理	196
10.2.1	波形内插原理	196
10.2.2	WI 声码器基本结构	197
10.2.3	特征波形的傅里叶级数表示	197
10.3	WI 声码器编码原理	199
10.3.1	线性预测分析与量化	200
10.3.2	基音周期的提取与编码	204
10.3.3	特征波形的提取	207
10.3.4	特征波形的对齐	208
10.3.5	特征波形的编码	210
10.3.6	特征波形功率的归一化	215
10.3.7	特征波形功率的编码	216

10.3.8	编码参数的比特分配	218
10.4	WI 声码器译码原理	218
10.4.1	特征波形功率的去归一化	218
10.4.2	瞬时基音与瞬时特征波形的内插	219
10.4.3	二维特征波形表面到一维残差信号的转换	220
10.4.4	线性预测合成	221
10.5	小结	222
附录A	英汉名词术语对照	223
附录B	常见语音编码标准码流的零一变化率直方图	229
附录C	常见语音编码标准码流的帧结构比较表	234
参考文献		236

第一章 语音编码概览

1.1 引言

语音是语言的声学表现,是人类交流信息最自然、最有效、最方便的基本手段,也是人类进行思维的一种依托。随着信息科学技术的飞速发展,语音处理技术在最近 20 多年中取得了重大进展。语音信号处理技术的 4 大分支——语音增强、语音编码、语音合成和语音识别(包括说话人识别、语种识别、关键词识别和连续语音识别)都在蓬勃发展,学术思想丰富多彩,研究开发成果累累,并且陆续地被推向实际应用。

语音编码技术是语音信号处理技术的一个重要组成部分,这是因为长期以来,在通信网的发展过程中,解决语音信号传输效率是一个十分关键的问题,极其重要。目前科研人员已通过两个途径研究这一课题:一是研究新的调制方法与技术,来提高信道传输信息的比特率,指标是每赫兹带宽所传送的比特数;二是压缩语音编码的比特率,例如标准脉冲编码调制(PCM)编码,对 3.4kHz 频带语音信号需用 64Kb/s 编码比特率进行传送,而压缩这一比特率显然可以提高信道传送的话路数。这对任何频率资源有限的传输环境来说,无疑是极为重要的,尤其是在无线通信技术决定今后通信发展命运的今天更显得重要。

1.2 语音编码总体发展方向

语音通信是人类通信最基本、最重要的方式之一。语音信号的数字化传输和存储,在可靠性、抗干扰能力、快速交换等方面远胜于模拟信号,且灵活方便、易于保密、价格低廉,所以数字化语音在通信系统中所占比重越来越大。语音编码是数字语音通信中的一项重要技术,为了使同样的信道容量能传输更多路的语音信号和节省存储空间,语音编码技术随着通信技术的发展也取得了很大的进展,并广泛应用于短波、超短波、地面微波和卫星通信系统中。

随着军事通信、移动通信和因特网的飞速发展,语音通信技术也在不断地进行更新并与之相融合。语音编码作为数字语音通信过程中的重要组成部分,它在总体上有 3 个大的发展方向:低速率语音编码、变速率语音编码和不压缩语音编码。

低速率语音编码。在现代通信中,信道利用的有效性和经济性仍是研究的重要目标,低速率语音编码技术是语音通信中不可缺少的一个重要研究方向。因为在实际通信中,有些信道难以扩宽且质量很差,如短波信道;有的信道正被广泛使用,短期内难以更新;有些是在特殊条件下的通信,如强干扰下的军事通信等;还有些信道价格昂贵,节省每一个比特都意味着经济上的好处。在这些情况下,低速率语音编码极为重要。随着数字通信业务的发展,高质量的低速率语音编码技术将成为研究的热点,它必将在保密通信、语音

邮件、网络通信、IP 电话等领域有广泛的应用前景,特别是在信息化战场上,它可广泛应用于单兵作战的通信设备上。

变速率语音编码。它是指在语音通信过程中,通信系统根据需要动态地调整语音编码速率,在合成语音质量和系统容量中取得灵活折衷,最大限度地发挥系统的效能。在当前应用广泛、前景广阔的码分多址(CDMA)移动通信系统中,采用的变速率语音编码算法对于系统的容量和通话质量有非常重要的影响。随着技术的成熟,它的应用领域也越来越广阔,不仅限于移动通信系统,在 IP 电话、因特网和卫星通信中都有很好的应用前景。

不压缩语音编码。低速率语音编码和变速率语音编码是未来语音编码的主要发展方向,但也有一些通信系统仍然采用不压缩语音编码。这是因为目前传输带宽增长很快,特别是光纤通信的发展,使得传输成本大幅度下降,为语音编码付出的费用与所节省的传输费用相比变得不合算了,语音不压缩传输所需费用反而更低。因此这些系统就采用了不压缩的语音编码算法,大多数直接利用脉冲编码调制(PCM)进行传输,这样可以保证很高的语音通信质量和比较简单的通信设备,极大地降低了通信费用。不压缩语音编码主要应用于地面微波通信和光纤通信系统中。

1.3 语音编码分类

语音编码既是一个人们熟知的领域,又是一个新的领域。自从 1937 年 A. H. Reeves 提出脉冲编码调制(PCM)编码以来,语音编码技术就开创了语音数字化通信的历程。近年来,随着通信技术和人类社会信息化进程的加快,语音编码技术也正在迅速地发展。各种语音编码的新技术和新算法更是层出不穷,让人有眼花缭乱和应接不暇的感觉。因此,对语音编码算法进行合理地分类,可以让读者从整体上把握语音编码技术的各个发展领域。

根据对语音信号不同的处理角度,语音编码可有 4 种不同的分类方法:按语音信号带宽分类、按语音编码方法分类、按语音编码处理域分类和按语音编码速率分类。下面分别进行介绍。

1.3.1 按语音信号带宽分类

按语音信号带宽分类是根据语音信号本身在频域内所占的频谱宽度进行的一种语音编码分类方法。据此方法可以将语音编码分为 3 种:普通话带语音编码、宽带语音编码和高宽带语音编码。

目前应用得最多的是普通话带语音编码。标准的话路频带(话带)语音信号是 300Hz~3400Hz,再加上少量的保护带宽,则一般意义上所说的一路标准话路频带宽度就为 4kHz。由于人类发声的频率范围主要集中在 1kHz~3kHz 左右,因此这种设定对于传递语音的主要信息是可以接受的。根据奈奎斯特准则,带宽为 4kHz 的模拟信号变成数字信号时,采样频率应为 8kHz 才能在还原时不会出现频谱重叠。因此,通常所说的语音编码(本书所讲的语音编码也是这样)一般都指的是这种对话带范围内的语音信号经 8kHz 采样后的数字语音信号进行的有损压缩过程。

语音编码的另一研究领域也日益受到人们的重视,并且得到十分迅速的发展,这就是宽带语音编码。它可以用于电视会议系统等对语音要求较高的应用领域中。宽带语音信

号带宽在 7kHz 左右,采样率一般在 20kHz 左右。国际电信联盟标准部(ITU-T)第 16 工作组于 2002 年 1 月公布的自适应多速率宽带(AMR-WB)语音编码器的标准为 G.722.2 协议,其语音信号带宽为 50Hz~7000Hz,采样率为 16kHz,属于宽带语音编码的一种。AMR-WB 语音编码器已被第三代伙伴计划(3GPP)选定为全球移动通信系统(GSM)和第三代宽带码分多址(WCDMA)通信系统的语音编码器,并将被应用于因特网协议(IP)电话、第三代移动通信、综合业务数字网(ISDN)宽带电话、ISDN 可视电话和电视会议等领域。这标志着无线和有线业务第一次采用同样的语音编码器。

另外,还有对更高带宽的语音信号的编码,这就是高宽带语音编码。它不仅是对话音信号的处理,而且还将人耳听力范围内的音频信号(带宽在 20kHz 左右)进行压缩,此时的采样率一般都在 50kHz 左右。这一研究领域的应用范围也十分广泛,比如在电影、高保真音响、电脑音乐及艺术创作等方面的应用。该技术的应用将为人类进入多媒体信息社会提供更为丰富的声音资源。

1.3.2 按语音编码方法分类

语音信号是一种时变的准周期信号,它可以近似地被看做由许多振幅和相位都随时间变化的正弦波构成的信号,因此可以用语音的抽样波形来描述语音信号,同时也可以语音的参数特征来描述语音信号。人们根据描述语音信号的不同方法将语音编码分为 3 类:波形编码、参数编码和混合编码。这种按语音编码方法进行的分类也是最传统和最常用的一种分类方法,目前被大多数相关学者所采用。

波形编码的指导思想是将语音信号当做一般波形信号来处理,使重构的语音波形与原始的语音波形尽可能保持一致,即在编码端以波形逼近为原则对语音信号进行压缩编码,而在译码端根据这些编码数据恢复出语音信号的波形。波形编码具有适应能力强、语音质量好和抗噪性能强等优点,但所需要的编码速率较高,一般在 16Kb/s~64Kb/s 速率能得到较高的语音编码质量。当速率进一步降低时,其性能会下降得很快。常用的波形编码方式包括脉冲编码调制(PCM)、自适应增量调制(ADM)、自适应差分脉码调制(ADPCM)、自适应预测编码(APC)和自适应变换编码(ATC)等。

参数编码一般也称做声码器(Voice Coder,简称为 Vocoder),它的出发点与波形编码不一样,其基本思想是通过对语音信号特征参数的提取来进行编码,力图使重构的语音信号具有尽可能高的可懂度,即保持原始语音信号的语义,而合成的语音信号波形可能与原始的语音信号波形有较大的差别。这类编码的优点是编码速率比较低,通常都在 16Kb/s 以下,甚至可以低到 2.4Kb/s 以下。它的主要问题是合成语音质量较差,特别是自然度较低;另外它对说话环境的噪声比较敏感,需要在较安静的环境下才能给出较高的可懂度。共振峰声码器和线性预测声码器都是典型的参数编码器。

自从 20 世纪 30 年代末提出脉冲编码调制原理和声码器(Vocoder)的概念后,语音信号编码就一直沿着上述两个方向发展。然而,由于这两种编码方法各自存在一定的不足,因此人们就想到了混合编码的方法。混合编码是上述两类方法优点的有机结合,由于突破了波形编码和参数编码的界限,因此混合编码得到了最广泛的应用。与参数编码相同的是,混合编码也是基于语音产生模型的假定并采用了合成分析技术,同时它又利用了语音的时间波形信息,增强了重构语音的自然度,使得合成语音质量有了明显的提高,其代

价是编码速率相应上升,一般都在 $2.4\text{Kb/s} \sim 16\text{Kb/s}$ 之间。这类编码器的典型代表有:多脉冲激励线性预测编码(MPLPC)、规则脉冲激励线性预测编码(RPE-LPC)和码激励线性预测(CELP)编码等。

1.3.3 按语音编码处理域分类

语音信号是一种典型的非平稳信号,但由于语音的形成过程是与发音器官的运动密切相关的,这种物理运动比起声音振动速度来讲要缓慢得多,因此语音信号常常可假定为短时平稳信号,即在 $10\text{ms} \sim 30\text{ms}$ 这样的时间段内,语音信号的某些物理特征参量可近似地看做是不变的。这就决定了语音信号可以进行时域处理。同时语音信号的短时平稳性决定了语音频谱在 $10\text{ms} \sim 30\text{ms}$ 这样短的时间段中也是平稳的,因此也可以对语音信号进行频谱分析,即频域处理。

人们根据语音信号处理域的不同将语音编码分为两类:时域编码和频域编码(变换域编码)。顾名思义,时域编码就是对语音信号在时间域内进行相关处理并进行压缩编码;频域编码就是将语音信号由时间域变换到频域内进行相关处理并进行压缩编码。

利用语音信号本身的性质和人类的听觉特性,可以确定时域编码和频域编码的分类关系,如图 1-1 所示。

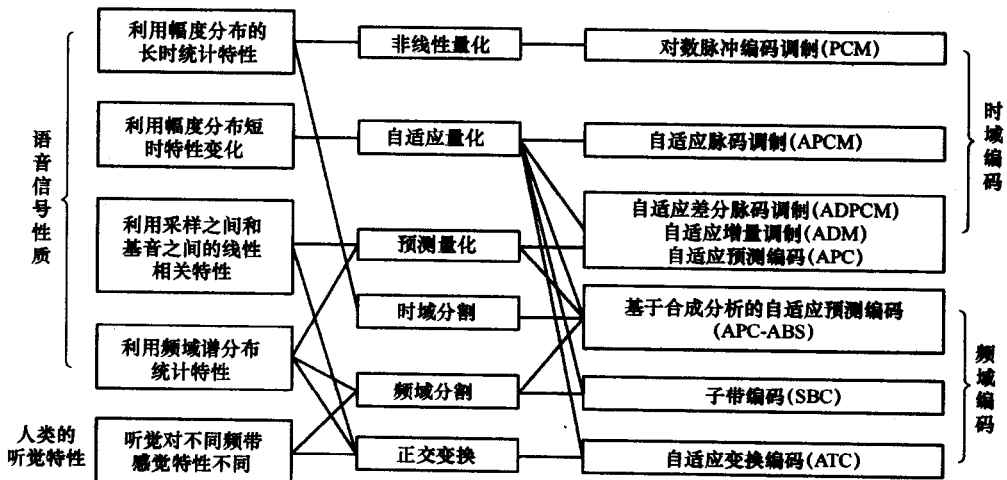


图 1-1 时域编码和频域编码的分类关系图

1.3.4 按语音编码速率分类

按语音编码速率分类是根据语音信号进行压缩编码后进行传输或存储所需要的数据速率(比特率)进行的一种语音编码分类方法。据此方法可以大致将语音编码分为 3 种:高速率(32Kb/s 以上)语音编码、中速率($4.8\text{Kb/s} \sim 32\text{Kb/s}$)语音编码和低速率(4.8Kb/s 以下)语音编码,如图 1-2 所示。

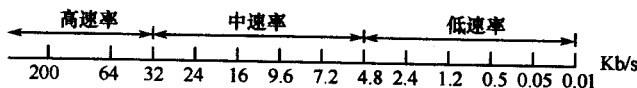


图 1-2 语音编码速率划分图

从图 1-2 可以看出,编码速率分类和编码方法分类基本上是对应的。波形编码方法通常称为高速率编码,其比特率一般都在 32Kb/s 以上;参数编码方法通常称为低速率编码,其比特率一般都在 4.8Kb/s 以下;介于中间的编码方法属于中速率语音编码,也就是混合编码方法。

当然,编码速率分类和编码方法分类的对应并不是绝对的,这种对应是一个相对的概念,有时它们相互之间存在着交叉关系。另外,也有学者按编码速率将语音编码分为 5 种:高速率(32Kb/s 以上)语音编码、中高速率(16Kb/s~32Kb/s)语音编码、中速率(4.8Kb/s~16Kb/s)语音编码、低速率(1.2Kb/s~4.8Kb/s)语音编码和极低速率(1.2Kb/s 以下)语音编码。它和图 1-2 的分类方法没有本质的区别,只是此方法分得更详细而已。本书在图 1-2 的基础上对语音编码进行速率分类,因此本书所讲的低速率语音编码指的是 0.01Kb/s~4.8Kb/s 的语音编码,它包含了极低速率语音编码。

1.4 语音编码国际标准

当前,语音编码技术不仅受到研究部门、应用部门的重视,而且推动了标准的制订,因为标准是工业生产的一个重要前提,对通信体制的确定有很大影响。语音编码技术在国际标准化工作中可称为最活跃的领域之一。

目前,以国际电信联盟标准部(ITU-T)为代表的各种国际机构和组织提出了大量的国际语音编码标准建议,了解这些语音编码标准可以帮助我们系统地学习、掌握和应用语音编码知识。但大量的国际语音编码标准又让读者很难从整体上了解和把握这些标准的全貌,绝大多数读者只是了解这些标准的一部分,他们十分需要一个完整的语音编码国际标准资料。

基于这个原因,本书以编码速率为主线将目前已公布的国际语音编码标准进行了详尽地分类总结和相关资料的介绍。这些国际标准被分为:高速率语音编码国际标准、中速率语音编码国际标准、低速率语音编码国际标准和变速率语音编码国际标准。

1.4.1 高速率语音编码国际标准

高速率语音编码的编码速率在 32Kb/s(含)以上,它具有适应能力强、话音质量好和抗噪性能强等优点,一般用于具有较高传输带宽的通信系统中。表 1-1 给出了高速率语音编码国际标准的相关资料,其中“×”符号表示暂时没有找到相关资料,而表中各个英文缩写的中文含义请参见表 1-2。

表 1-1 高速率语音编码国际标准的相关资料表

编码速率/(Kb/s)	标准名称	制定机构	公布时间	编码算法	编码方法	语音质量			时延/ms	复杂度/MIPS	帧长		主要用途
						MOS	DRT	DAM			/ms	比特	
64	G. 711	ITU-T	1972	PCM	波形	4.3	95	73	×	0.01	1/8	8	公用网
64	G. 722	ITU-T	1988	SB-ADPCM	波形	4.1	×	×	×	5	1/8	8	公用网
56	G. 722	ITU-T	1988	SB-ADPCM	波形	4.0	×	×	×	5	1/8	7	公用网