

现代语音编码技术

吴家安 主编



科学出版社
www.sciencep.com

现代语音编码技术

吴家安 主编

科学出版社

北京

内 容 简 介

本书全面系统地阐述了现代语音编码的原理、技术及应用,讨论了目前语音编码的若干热点研究课题,介绍了语音编码的发展趋势和方向。全书由六个相对独立而又相互紧密联系的单元组成,共计十五章。主要内容包括现代语音编码技术导论、矢量量化编码、时域波形编码、子带编码、变换域编码、参数编码、混合激励线性预测编码、混合编码、多带激励编码、低速率和极低速率语音编码、变速率语音编码、宽频带高音质声频编码、感知音频编码、MPEG 音频压缩编码以及语音及音频编码的国际标准和地区标准等。本书取材新颖、结构合理、深入浅出、阐述清晰、内容丰富、实用性强,包含近二十几年来语音编码技术的许多新成果和新进展。

本书可作为通信、电子、计算机、信号处理、雷达、导航、无线电遥测、遥控、自动控制、广播、电视及其他信息技术专业的高等院校本科生及大专生的教学用书,也可供从事相关领域研究的科技工作者参考。

图书在版编目(CIP)数据

现代语音编码技术/吴家安主编. —北京:科学出版社,2007

ISBN 978-7-03-020507-0

I. 现… II. 吴… III. 语音数据处理—编码 IV. TN912.3

中国版本图书馆 CIP 数据核字(2007)第 179171 号

责任编辑:刘宝莉 / 责任校对:陈玉凤

责任印制:刘士平 / 封面设计:王 浩

科 学 出 版 社 出 版

北京东黄城根北街 16 号

邮政编码: 100717

<http://www.sciencep.com>

双 青 印 刷 厂 印 刷

科学出版社发行 各地新华书店经销

*

2008 年 1 月第 一 版 开本: B5 (720×1000)

2008 年 1 月第一次印刷 印张: 19

印数: 1—4 000 字数: 363 000

定价: 31.00 元

(如有印装质量问题, 我社负责调换〈双青〉)

前 言

20 世纪 80 年代至今的二十几年间,是语音编码技术发展极为迅速的时期。现代语音编码技术主要是介绍在这二十几年间发展起来的新的语音编码技术,它现在仍然处于迅速发展之中,是通信和信息技术中备受关注的重要领域。现代语音编码技术的发展、进步和广泛应用,是现代通信技术领域中的一个闪光点,对于现代通信及信息技术的进步和发展起到了重要的推动作用。

众所周知,应用最早和最广泛的语音编码是 PCM(脉冲编码调制)。国际电报电话咨询委员会(CCITT)于 1972 年制定的 G. 711 建议规定的 A 律和 μ 律这两种 PCM 编码标准,都是非线性量化 PCM 编码方法。北美、日本使用 μ 律压扩方法,其他国家和地区使用 A 律压扩方法。编码器输出的速率通常为 64 kb/s。PCM 是人们最熟悉的语音编码技术,而且 PCM 对于促进通信的数字化、推动通信和信息技术的迅速发展曾经起过极为重要的作用。但是,由于 PCM 的编码速率过高,占用信道带宽过大,不适应通信和信息技术发展的需要。为了压缩编码速率,减少传输占用的带宽,人们一直致力于研究开发新的语音编码技术。这种强烈的客观需求是推动语音编码技术发展的巨大动力。最近二十几年来,随着计算机、微电子、信号处理等相关技术的迅速发展和广泛应用,尤其是数字信号处理算法和器件(DSP 芯片)的飞速发展和应用,为中、低速率语音编码器的发展和应用准备了必要条件。正是在这种情况下,从 20 世纪 80 年代以来的二十几年间,语音编码技术进入一个飞速发展的时期。在这一时期,研究和开发出了 32 kb/s、24 kb/s、16 kb/s、8 kb/s、4.8 kb/s、4 kb/s、2.4 kb/s、2 kb/s 等速率的一系列语音质量良好的编码算法和编码技术,其中,32 kb/s、16 kb/s、8 kb/s、4.8 kb/s 等速率的语音编码技术已经由 ITU-T(CCITT)制定了国际标准,并进入了实用阶段;其余速率的语音编码技术也已经有了一些地区性标准,并正在成为制定新的国际标准的竞争热点,必将进一步推动语音编码技术沿着低速率和极低速率方向发展。

进入 21 世纪以来,语音编码技术开始进入一个新的全面、快速发展的时期。近年来在数字声广播、消费电子技术以及电话电视会议迅速发展的影响下,语音编码也在朝着宽频带高音质声频编码的方向发展,受到人们的广泛关注。其应用领域主要是数字声广播、电话电视会议系统、家电及娱乐电子技术。由于受到社会需求的强大推动,这是一个相当重要的发展动向。此外,第三代移动通信(3G)的语音编码技术是语音编码另一个备受关注的研究课题。3G 的三大主流技术 TD-SCDMA、WCDMA 和 CDMA2000 都采用了变速率语音编码技术。TD-SCDMA 拟采用自适应

应多速率(AMR)语音编码;WCDMA 拟采用 AMR 语音编码,同时将可选模式声码器(SMV)作为备用技术;CDMA2000 选用变速率 CELP(QCELP)声码器或增强型变速率编码器(EVRC)。在数字卫星通信等系统中,也有的采用了变速率语音编码技术。变速率语音编码由于能够在保证重建信号质量的条件下进一步压缩编码速率,其应用前景十分广阔,是语音编码另一个重要的发展趋势。

近年来随着 DSP 芯片技术、语音学、语义学、神经生理学和心理学的迅速发展和广泛应用,发展起来一种利用人类听觉系统特性的感知编码技术,在音频压缩编码中得到广泛应用。感知编码器首先分析输入信号的频率和振幅,然后将其与人的听觉感知模型进行比较。感知编码器利用听觉感知模型去除那些与听觉感知无关的冗余部分,使得比特率大幅度降低,而编码信号质量并未下降。感知编码已经在 MPEG 音频压缩编码中应用,取得了良好效果,是语音编码又一个重要的发展动向。

国际标准化组织(ISO)和国际电工技术委员会(IEC)组织运动图像专家组(MPEG)于 1992 年 11 月制定了关于视频和音频信号压缩的国际标准 ISO/IEC 11172,即 MPEG-1 标准。这个标准有系统、视频和音频三大主要部分,最大音频比特率为 1.856 Mb/s。MPEG-1 中有关音频压缩的标准,即 11172-3,已经成功应用在 VCD、CD-ROM、ISDN、数字音频广播以及视频游戏等领域中,它支持每声道为 224~32 kb/s 的 32 kHz、44.1 kHz 和 48 kHz 的 PCM 数据。现在 MPEG-2、MPEG-4 等的音频层编码标准也已经制定出来,目前非常流行的 MP3、MP4 等都属于 MPEG 音频编码。因此,MPEG 音频编码也是一个十分值得注意的发展方向。

为了适应语音编码技术快速发展和应用的需要,从 20 世纪 80 年代至今的二十几年间,国际电报电话咨询委员会(CCITT)、国际电信联盟标准化部(ITU-T)和其他国际及地区标准化组织,制定了一系列的语音及音频编码的国际和地区标准。这些标准不但是学习、研究和应用现代语音及音频编码技术的重要依据,而且从另一个侧面反映出了这一时期语音及音频编码发展的历程和发展的趋势。

学习和掌握现代语音编码技术的基础知识、分析方法、关键技术和算法十分重要,只有了解和掌握这些已经成熟或基本成熟的方法和技术,才能很好地适应现代通信和信息技术发展的需要,做好我们的本职工作,也才能研究开发新的方案和算法,提出具有自主知识产权的技术方案。

本书编者根据多年从事语音编码教学和科研工作的经验、体会及成果,并参考了大量的文献和资料,试图总结现代语音编码技术的发展成就,介绍各种主要的实用编码技术及编码标准,全面、系统地阐述现代语音编码的原理、技术及应用,讨论目前的热点研究课题及最新成果,指出语音编码今后的发展趋势及方向,以期给广大读者以启示。

本书内容由相对独立而又相互紧密联系的六个单元共计十五章组成。第一单元是现代语音编码技术的入门知识,包括第一章“现代语音编码技术导论”和第二章“矢量量化编码”,分别介绍现代语音编码的基本知识、基本概念和基本技术。第二单元集中讨论经典的语音编码技术,即波形编码技术,包括第三章“时域波形编码”、第四章“子带编码”和第五章“变换域编码”,分别介绍时域波形编码、频域波形编码和变换域波形编码的原理、技术及应用。第三单元重点介绍现代参数编码技术,包括第六章“参数编码”和第七章“混合激励线性预测编码”,分别讨论 LPC 声码器基本原理、技术、应用及新的进展。第四单元较为详细地阐述语音混合编码的技术,包括第八章“混合编码”和第九章“多带激励编码”,分别介绍语音混合编码的基本原理、技术、应用及其新进展。第五单元讨论语音编码技术的几个热门研究课题,包括第十章“低速率和极低速率语音编码”、第十一章“变速率语音编码”、第十二章“宽频带高音质声频编码”、第十三章“感知音频编码”和第十四章“MPEG 音频压缩编码”,分别探讨了语音以及音频编码技术的几个主要发展趋势和方向。第六单元包括第十五章,介绍语音及音频编码的国际标准和地区标准,供读者在学习、研究和应用现代语音编码技术时参考。

本书可作为通信、电子、计算机、信号处理、雷达、导航、无线电遥测、遥控、自动控制、广播、电视及其他信息技术专业的高等院校本科生及大专生的教学用书,建议教学时数为 60 学时。书中打※号的章节内容,教师在教学过程中可以根据实际情况适当精简。如果使用本书对大专生教学时,可以省略这些章节的内容不讲。

本书由吴家安主编,吴家安、张会生、吴海锋、郭鸿基参编。吴家安、张会生制定了编写提纲和计划,吴家安编写了第一章、第三~十章和第十二~十五章的全部内容以及第二章、第十一章的部分内容,吴海锋编写了第二章的大部分内容,郭鸿基编写了第十一章的部分内容,吴海锋和郭鸿基绘制了书中的部分图表。全书由吴家安、张会生整理定稿。

本书在编写过程中,得到了西北工业大学的领导和老师的大力支持,电子信息工程专业的甄钊博同学绘制了书中的大部分图表,在此一并表示感谢。

由于编者水平有限,书中难免存在不妥之处,敬请读者不吝指正。

目 录

前言

第一章 现代语音编码技术导论	1
1.1 现代语音编码概述	1
1.1.1 语音编码和现代语音编码的概念及发展	1
1.1.2 对语音编码的要求	2
1.1.3 现代语音编码技术的作用和意义	2
1.1.4 语音编码系统的构成	3
1.2 语音编码的分类	4
1.2.1 传统的分类方法	4
1.2.2 按照编码速率的分类方法	5
1.2.3 按照被编码信号所在的域的分类方法	5
1.2.4 按照编码所使用的主要技术的分类方法	5
1.2.5 按照编码算法是否依赖于某种模型假定的分类方法	6
1.2.6 按照被编码信号的属性的分类方法	6
1.2.7 按照编码速率是否固定的分类方法	6
1.2.8 按照编码出现和应用时间以及技术发展情况的分类方法	7
1.3 衡量语音编码性能的主要指标	7
1.3.1 编码质量	7
1.3.2 编码速率	8
1.3.3 编译码复杂程度	9
1.3.4 编译码时延	9
1.3.5 坚韧性	10
1.4 提高语音编码质量的基本途径	10
1.5 语音编码发展现状及今后的发展方向	13
1.5.1 语音编码的发展现状	13
1.5.2 现代语音编码的发展趋势及方向	16
1.6 本书的内容结构安排	17
复习思考题	17
练习题	18
第二章 矢量量化编码	19

2.1 概述	19
2.1.1 量化的分类	19
2.1.2 矢量量化的发展及应用	19
※2.2 标量量化	21
2.2.1 无记忆标量量化	21
2.2.2 有记忆标量量化	34
2.3 矢量量化原理	37
2.3.1 矢量量化的定义	37
2.3.2 失真测度	39
2.3.3 矢量量化器的结构	41
2.3.4 矢量量化器的速率	41
2.3.5 最佳矢量量化器	42
2.3.6 矢量量化器的设计算法	44
2.4 无记忆矢量量化器	49
2.4.1 基本矢量量化器	50
2.4.2 树搜索矢量量化器	52
2.4.3 多级矢量量化器	59
2.4.4 乘积码矢量量化器	64
※2.5 有记忆矢量量化器	68
2.5.1 反馈矢量量化器	68
2.5.2 自适应矢量量化器	70
复习思考题	71
练习题	71
第三章 时域波形编码	73
3.1 PCM	73
3.1.1 PCM 的基本原理	73
3.1.2 PCM 时分复用原理	78
3.1.3 PCM 的 A 律 13 折线编码	82
3.2 DPCM	88
3.3 ADPCM	90
3.4 DM 和 ADM	91
3.4.1 DM	91
3.4.2 ADM	92
3.5 CCITT G. 721 建议的高质量 32 kb/s ADPCM	92
3.6 CCITT G. 727 建议的镶嵌式 ADPCM	99

复习思考题.....	101
练习题.....	102
第四章 子带编码	103
4.1 概述	103
4.2 子带编码的工作原理	104
4.3 子带编码的比特分配	106
4.4 整数带滤波器组	109
※4.5 正交镜像滤波器组	111
4.6 7kHz 带宽 SB-ADPCM 高音质声频编码系统	112
4.6.1 CCITT G. 722 关于 7 kHz 带宽高音质声频编码方案的主要内容	112
4.6.2 编、译码器的组成	113
4.6.3 CCITT G. 722 编码方案的特点	116
4.6.4 G. 722 编码标准的应用场合	116
复习思考题.....	117
练习题.....	117
第五章 变换域编码	118
5.1 变换编码概述	118
5.1.1 什么是变换编码? 为什么进行变换?	118
5.1.2 正交变换与正交矩阵	118
5.2 几种常用的正交变换	120
5.2.1 K-L 变换	120
5.2.2 W-H 变换	122
5.2.3 DCT(离散余弦变换)	124
5.2.4 Haar 变换	127
5.3 变换域系数的量化和编码	129
5.3.1 变换矩阵的选择	129
5.3.2 量化系数的选择	130
复习思考题.....	130
练习题.....	131
第六章 参数编码	132
6.1 参数编码概述	132
6.1.1 语音信号的产生模型	132
6.1.2 参数编码的类型	134
6.2 线性预测原理	136
6.2.1 预测编码的基本原理	136

6.2.2 线性预测	138
6.2.3 线性预测方程组及预测系数的确定	139
※6.3 线性预测的几种推演参数	140
6.3.1 反射系数	141
6.3.2 对数面积比系数	141
6.3.3 倒谱系数	141
6.3.4 预测器多项式的根	142
6.3.5 全极点系统的冲激响应	142
6.3.6 全极点系统冲激响应的自相关系数	143
6.3.7 预测误差滤波器冲激响应的自相关系数	143
※6.4 线谱对	143
6.4.1 线谱对分析的基本原理	143
6.4.2 线谱对参数的求解	145
6.5 线性预测声码器	145
6.5.1 考虑语音短时和长时相关性的语音生成模型	146
6.5.2 LPC 声码器	147
6.5.3 特征参数的提取	148
6.6 LPC-10 声码器	150
6.6.1 LPC-10 的编码器	150
6.6.2 LPC-10 的译码器	151
6.6.3 LPC-10 参数的编码与译码	152
6.7 LPC-10e 声码器	153
6.7.1 LPC-10e 的编码器	153
6.7.2 LPC-10e 的译码器	155
6.8 LPC 声码器存在问题及其改进措施	156
6.8.1 LPC 声码器存在的主要问题	156
6.8.2 LPC 声码器的改进措施	157
复习思考题	158
练习题	159
第七章 混合激励线性预测编码	160
7.1 概述	160
7.2 MELP 声码器编码原理	160
7.3 MELP 声码器译码原理	166
7.3.1 基音周期的译码	166
7.3.2 增益的译码和抑制	166

7.3.3 参数的插值	167
7.3.4 混合激励的生成	167
7.3.5 自适应谱增强	167
7.3.6 线性预测合成	168
7.3.7 增益校正	168
7.3.8 脉冲整形滤波	168
复习思考题	168
练习题	169
第八章 混合编码	170
8.1 语音混合编码的一般原理	170
8.1.1 感觉加权滤波器	170
8.1.2 合成分析法	172
8.1.3 激励源的改进	173
8.2 多脉冲激励线性预测编码	173
8.2.1 多脉冲激励线性预测编码的原理框图	173
8.2.2 最佳激励脉冲参数的估值	174
8.2.3 准最佳激励脉冲参数的估值	175
8.2.4 多脉冲线性预测编码的应用	176
8.3 规则脉冲激励线性预测编码	176
8.3.1 RPELPC 的基本原理	176
8.3.2 RPE-LTP 的编码方案	178
8.3.3 GSM 系统的 RPE-LTP 编码器	179
8.3.4 GSM 系统的 RPE-LTP 译码器	180
8.4 码激励线性预测编码	180
8.4.1 CELPC 的基本原理	181
8.4.2 美国政府标准(FED-STD-1016)4.8 kb/s 的 CELPC 声码器	182
8.4.3 美国电子工业协会电信协会(EIA/TIA)标准 8 kb/s 的 VSELPC 声码器	184
8.4.4 CCITT G. 728 建议的 16 kb/s 低延时 CELPC(LD-CELPC)声码器	186
8.4.5 ITU-T G. 729 建议的 8 kb/s 共轭结构代数码激励线性预测(CS- ACELP)声码器	189
复习思考题	191
练习题	191
第九章 多带激励编码	192
9.1 概述	192

9.2 MBE 语音模型.....	193
※9.3 MBE 语音参数的估计.....	197
9.3.1 某个子频带内的 MBE 语音产生模型	197
9.3.2 MBE 模型参数的估计	198
9.3.3 MBE 语音合成	199
9.4 MBE 声码器的应用.....	200
复习思考题.....	200
练习题.....	201
※第十章 低速率和极低速率语音编码.....	202
10.1 低速率语音编码概述.....	202
10.2 波形内插语音编码.....	203
10.2.1 狭义波形内插法和广义波形内插法	203
10.2.2 浊音语音的波形内插	203
10.2.3 广义内插(GWI)模型	207
10.2.4 实用 GWI 语音编码系统	209
10.3 正弦变换编码.....	213
10.3.1 语音信号的正弦分析	213
10.3.2 语音信号的正弦合成	214
10.3.3 低速率正弦编码器	215
10.3.4 与原型波形内插相结合	216
10.4 极低速率语音编码概述.....	217
10.5 1200~400 b/s 的语音编码技术	218
10.5.1 帧填充技术	218
10.5.2 利用矢量量化技术	218
10.6 400 b/s 以下的语音编码技术——语音识别与合成技术	219
10.6.1 识别合成型声码器的基本原理	219
10.6.2 关于识别合成型声码器编码速率的估计	220
复习思考题.....	220
※第十一章 变速率语音编码.....	222
11.1 概述.....	222
11.1.1 变速率语音编码的必要性和可能性	222
11.1.2 变速率语音编码中使用的相关新技术	223
11.1.3 变速率语音编码的速率控制方式	223
11.2 实现变速率语音编码的关键技术.....	224
11.2.1 语音激活检测(VAD)技术	224

11.2.2 速率判决(RDA)技术	225
11.2.3 差错隐藏(ECU)技术	225
11.2.4 舒适背景噪声(CNA)生成技术	225
11.3 增强型变速率语音编码	226
11.4 变速率 CELPC(QCELPC)语音编码	227
11.4.1 QCELPC 概述	227
11.4.2 QCELPC 编码器	228
11.4.3 QCELPC 译码器	232
11.5 自适应多速率(AMR)语音编码	233
11.6 可选模式声码器(SMV)	235
复习思考题	236
练习题	236
※第十二章 宽频带高音质声频编码	237
12.1 宽频带声频的发展状况	237
12.2 宽带声频编码技术	238
12.2.1 宽带声频波形编码的分类	238
12.2.2 CCIR 建议的宽频带声频编码方式	239
12.3 宽频高音质声频编码器实例	239
12.3.1 编码系统组成	239
12.3.2 设计特点	240
12.3.3 编码系统的工作原理	240
复习思考题	241
※第十三章 感知音频编码	242
13.1 心理声学基础	242
13.1.1 人类听觉系统的听觉阈值	242
13.1.2 频域掩蔽	242
13.2 感知音频编码/译码系统的组成	245
13.3 信宿(听觉)模型	245
13.3.1 单音的听觉模型	245
13.3.2 考虑掩蔽效应的听觉模型	247
13.4 感知音频编码的时-频分析信源模型	248
13.5 根据听觉(信宿)模型进行量化编码	248
13.6 感知音频编码器	248
13.6.1 音频编码-1(AC-1)	248
13.6.2 音频编码-2(AC-2)	249

13.6.3 音频编码-3(AC-3)——数字杜比	249
13.6.4 Apt-X100 音频编码器	254
复习思考题	254
第十四章 MPEG 音频压缩编码	255
14.1 MPEG 音频编码概述	255
14.1.1 声音	255
14.1.2 模拟音频	256
14.1.3 数字音频	256
14.2 MPEG-1 音频压缩编码	256
14.2.1 MPEG-1 音频层概述	256
14.2.2 MPEG-1 音频层的数据结构	257
14.2.3 心理声学模型	257
14.2.4 MPEG-1 的音频层第 I 层	260
14.2.5 MPEG-1 的音频层第 II 层	261
14.2.6 MPEG-1 的音频层第 III 层	263
14.2.7 MP3	265
14.3 MPEG-2 音频压缩编码	267
14.3.1 MPEG-2 音频压缩编码概述	267
14.3.2 MPEG-2 音频编码与译码	267
14.3.3 MPEG-2 AAC(高级音频编码)	269
14.4 MPEG-4 音频压缩编码	271
14.4.1 MPEG-4 音频压缩概述	271
14.4.2 交互操作	271
14.4.3 音频编码	272
复习思考题	274
练习题	274
第十五章 语音及音频编码的国际标准和地区标准	275
15.1 概述	275
15.1.1 主要的国际标准化组织	275
15.1.2 主要的地区标准化组织	277
15.2 各标准化实体制定语音编码标准的一般工作程序	278
15.2.1 建立参考条款	278
15.2.2 语音编码器的性能测量	279
15.2.3 工作日程	280
15.3 CCITT(ITU)已经公布的电话带宽语音编码标准	280

15.4	ITU-T 近十几年来公布的电话带宽语音编码标准	281
15.5	CCITT 已经公布及 ITU-T 计划的宽带语音编码标准	281
15.6	国际海事卫星组织 4.15 kb/s 的 IMBE 语音编码标准	282
15.7	各主要地区标准化组织的语音编码标准	282
15.7.1	北美数字蜂窝移动通信语音编码标准	282
15.7.2	欧洲数字蜂窝移动通信语音编码标准	283
15.7.3	美国的保密通信语音编码标准	283
15.7.4	日本的数字蜂窝移动通信语音编码标准	284
15.8	MPEG 音频压缩编码标准	284
15.8.1	MPEG-1 音频压缩编码标准	284
15.8.2	MPEG-2 音频压缩编码标准	285
15.8.3	MPEG-4 音频压缩编码	286
	复习思考题	287
	参考文献	288

第一章 现代语音编码技术导论

最近二十几年来,随着数字通信、计算机、信号处理、微电子等相关技术的发展和广泛应用,语音编码技术发展非常迅速,取得了一系列突破性的成果,极大地促进了数字通信的发展和普及,是现代通信以及信息技术的一个亮点。所谓现代语音编码技术,主要就是指最近二十几年发展起来并得到广泛应用的语音编码技术。本章主要介绍现代语音编码技术的基本概念和基本知识,如什么是语音编码、对语音编码的要求、现代语音编码技术的作用和意义、现代语音编码系统的构成、分类以及主要性能指标等,以使读者对现代语音编码有一个初步的概要的了解。这些基本概念和基本知识,是学习和研究现代语音编码技术的入门知识和重要基础。读者应深刻理解,牢固掌握。

1.1 现代语音编码概述

1.1.1 语音编码和现代语音编码的概念及发展

在现代通信中,随着科学技术的迅速发展,图像、数据等非语音信息在通信信息总量中所占的比例大大提高,而且这种提高的趋势仍然会继续下去。但是,到目前为止,在大多数通信系统中,传输最多的信息仍然是语音信号。在可以预见的未来的通信中,尽管语音信号在通信信息总量中所占的比例会有所下降,但仍然是传输最多的信息。

语音信号是模拟信号,不能直接在数字通信系统中传输,必须首先进行模/数转换和数/模转换。这种模/数转换和数/模转换就称为语音编译码,其作用是在发送端将语音模拟信号转换为数字信号,到了接收端,再将收到的语音数字信号还原为语音模拟信号。语音编译码又简称为语音编码。可见,语音编码技术在数字通信中具有十分重要的关键的作用。

语音编码属于信源编码。自从 20 世纪 70 年代以来的三十多年中,尤其是 20 世纪 80 年代至今的二十几年来,随着科学技术的迅速发展,特别是随着计算机技术、微电子技术、信号处理技术、语音学、语义学、神经生理学和心理学以及编码理论的发展和进步,语音编码技术取得许多突破性进展,研究出许多实用的编码方案,这些方案在不断研究、改进和应用中日趋成熟,形成了各种实用的语音编码技术。现代语音编码技术已成为通信技术中一个相当重要的学科,在各种通信网络中

都得到广泛应用。

现代语音编码技术主要是指从 20 世纪 80 年代至今二十几年间发展起来的新的语音编码技术,它现在仍然处于迅速发展之中,是通信和信息技术中备受关注的重要领域。

1.1.2 对语音编码的要求

各种通信网络工作的环境各不相同,传输信息也不完全一样,用户情况千差万别,因而对语音编码提出的要求也就不尽相同。综合各种通信网络对语音编码的共同要求,大致有以下几点:

(1) 编码速率要适合在常用话音信道内传输,一般要求编码速率在 16~2 kb/s 范围内的情况较多。

(2) 在一定编码速率下,语音质量应尽可能高,即译码后恢复语音的保真度要尽量高,一般要求达到长话质量(网络质量),即 MOS 评分 4 分以上,或要求达到通信质量,即 MOS 评分 3.5 分左右。

(3) 编译码时延要小。总时延一般要求不大于 65 ms。

(4) 编译码算法复杂度不能太大,以适于用大规模集成电路实现。

(5) 坚韧性好,有较好的抗误码性能。

上述这些要求之间往往是互相矛盾的。例如,为了使语音质量好,编码速率就应高一些,但这又会使其占用的信道带宽增大。信道带宽是有限的,编码速率过高就不能在信道内传输。因此,在实际应用中要根据具体情况综合分析和比较,选择最佳的编码方案。

1.1.3 现代语音编码技术的作用和意义

应用最早和最广泛的语音编码是 PCM(脉冲编码调制)。CCITT 于 1972 年制定的 G. 711 建议规定的 A 律和 μ 律这两种 PCM 编码标准,都是非线性量化 PCM 编码方法。北美、日本使用 μ 律压扩方法,其他国家和地区使用 A 律压扩方法。其编码器输出的速率通常为 64 kb/s。由于 PCM 的编码速率过高,不适应通信和信息技术发展的需要,为了压缩编码速率,减少传输占用的带宽,人们一直在致力于研究开发新的语音编码技术。这种强烈的客观需求是推动语音编码技术发展的巨大动力。另一方面,最近二十几年来,随着计算机、微电子、信号处理等相关技术的迅速发展和广泛应用,尤其是随着数字信号处理算法和器件(DSP 芯片)的飞速发展和应用,为中、低速率语音编码器的发展和应用准备了必要条件。正是在这种情况下,从 20 世纪 80 年代以来的二十几年间,语音编码技术进入一个飞速发展的时期。在这一时期,研究和开发出了 32 kb/s、24 kb/s、16 kb/s、8 kb/s、4.8 kb/s、4 kb/s、2.4 kb/s、2 kb/s 等速率的一系列语音质量良好的编码算法和编码技术,