

信息科学与工程系列专著

数字语音编码技术

Digital Speech Coding Technologies

李 晔 崔慧娟 唐 昆 等编著



电子工业出版社

PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

<http://www.phei.com.cn>

信息科学与工程系列专著

数字语音编码技术

Digital Speech Coding Technologies

李 晔 崔慧娟 唐 昆 杨行峻 编著

电子工业出版社

Publishing House of Electronics Industry

北京 • BEIJING

内 容 简 介

本书全面、系统地阐述了数字语音编码的相关技术。全书共 15 章, 内容包括: 语音编码概述和声学基础, 语音编码中的数字分析和矢量量化, 语音编码的评价指标和关联技术, 语音信号的波形编码、参数编码和混合编码, 语音编码领域的热点问题——低速率语音编码、宽带语音编码、变速率语音编码和嵌入式语音编码, 语音编码系统的硬件实现, 以及语音编码的国内外标准。

读者对象: 从事语音编码和语音通信的科技研发人员和工程人员, 以及电子通信类及相关专业的高校教师、高年级本科生和研究生。

未经许可, 不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有, 侵权必究。

图书在版编目 (CIP) 数据

数字语音编码技术 / 李晔等编著. —北京: 电子工业出版社, 2013.1

(信息科学与工程系列专著)

ISBN 978-7-121-19028-5

I. ①数… II. ①李… III. ①语音数据处理—编码 IV. ①TN912.3

中国版本图书馆 CIP 数据核字 (2012) 第 278208 号

策划编辑: 张来盛 (zhangls@phei.com.cn)

责任编辑: 杨 博

印 刷: 涿州市京南印刷厂

装 订: 涿州市京南印刷厂

出版发行: 电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

开 本: 787×1 092 1/16 印张: 22.5 字数: 571 千字

印 次: 2013 年 1 月第 1 次印刷

定 价: 69.00 元

凡所购买电子工业出版社图书有缺损问题, 请向购买书店调换。若书店售缺, 请与本社发行部联系, 联系及邮购电话: (010) 88254888。

质量投诉请发邮件至 zltts@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线: (010) 88258888。

前言

语音是人类最直接、最方便的沟通手段,在通信技术发展的过程中,语音通信一直是科研人员和系统开发人员最为关注的问题之一。

语音通信最早可以追溯到 1876 年, Bell 发明的电话机首次利用声/电、电/声转换技术实现了远距离的语音信号传输。早期的语音通信装置采用的是模拟系统,在可靠性、抗干扰能力、快速交换等方面存在很多问题。自从 20 世纪 60 年代集成电路出现以后,通信信道和通信终端开始从模拟系统演变为全数字系统。与模拟通信相比,数字通信在可靠性和抗干扰能力等方面具有很大优势,不足之处是占用的频带较宽。为了节省数字语音信号传输所需的带宽,人们开始深入研究语音编码技术。

几十年来,数字语音通信技术几乎覆盖了通信的每一个领域,语音编码技术领域的研究成果也层出不穷,其中包括很多突破性的研究成果和实用的语音编码方案。借助于先进的数字信号处理芯片,这些研究成果已经得到了广泛应用。随着应用的普及,语音编码算法标准化的工作也取得了很大的进展。例如,主要的标准制定组织 ITU,已经制定了一系列语音编码标准,其编码速率越来越低,编码带宽从窄带到宽带,甚至到超宽带和全频带,可支持的速率也越来越多,而且其中部分标准具有互通性很强的嵌入式结构。

本书的编写,旨在帮助读者全面、系统地了解数字语音编码的相关技术。本书在讲述语音编码基础理论及相关语音编码算法的同时,力图反映数字语音编码技术的国内外最新进展和当前语音编码技术中的热点问题,并将编著者的相关研究成果囊括其中,以便读者参考和借鉴。此外,还讲述了实用语音通信系统中将涉及的部分关联技术,如语音增强、回声抵消、同步、抗信道误码和抗丢包等技术,以及硬件实现的相关技术,这部分内容对于科技研发人员和工程人员具有较高的实践参考价值。

本书共 15 章。其中第 1 章从整体上介绍语音编码的应用、分类以及发展趋势;第 2 章介绍语音信号的产生机理及听觉系统感知;第 3 章介绍语音编码中的数字分析技术;第 4 章介绍语音编码中的矢量量化技术;第 5 章介绍语音编码的评价指标;第 6~8 章分别介绍语音信号的波形编码、参数编码及混合编码;第 9~12 章分别介绍目前语音编码领域的 4 个热点:低速率语音编码、宽带语音编码、变速率语音编码和嵌入式语音编码;第 13 章介绍语音编码过程中经常会遇到的关联技术;第 14 章介绍语音编码系统硬件实现的相关原理及技术;第 15 章介绍当前语音编码的国际、地区及行业应用标准。

本书由李晔主持编写并统稿。其中,第 6~8 章由崔慧娟、唐昆编写,第 2、3 章由杨行峻、李晔编写,其余各章由李晔编写。

在本书的编写过程中,参阅了大量的国内外相关文献,其中包括 Rainer Martin 等编著的《Advances in Digital Speech Transmission》(Wiley Blackwell, 2008),杨行峻等编著的《语音信号数字处理》(电子工业出版社,1995),W. B. Kleijn 等编著的《Speech Coding and

Synthesis》(Elsevier Science, 1995), 以及清华大学电子工程系赵铭、李悦、张鑫、李军林、陆希玉、田秋玲、周高洪、詹杰、张仁志、李燕成和雷鸣等人的博士论文或硕士论文; 山东省计算中心、山东省计算机网络重点实验室的马晓凤、郝秋赟、樊燕红和姜竞赛等做了大量的图文编辑及校对工作。另外, 本书的出版得到电子工业出版社及相关编辑的大力支持。在此, 谨对有关原作者、同事和编辑表示诚挚的感谢。

囿于编著者的水平, 书中难免出现一些疏漏和错误, 恳请读者批评指正, 并将意见发送至 liye@keylab.net。同时, 由于语音编码技术的发展日新月异, 书中的部分表述具有较强的时效性, 也恳请读者谅解, 并参阅最新资料。

编著者

2012年9月



目 录

第 1 章 绪论	1
1.1 概述	1
1.2 语音编码的应用	1
1.3 语音编码的分类	2
1.3.1 按编码方法分类	2
1.3.2 按编码速率分类	3
1.3.3 按编码语音带宽分类	3
1.3.4 按编码信号所在域分类	3
1.4 语音编码的硬件与系统	4
1.5 语音编码的发展趋势	4
1.6 全书的组织	5
参考文献	5
第 2 章 语音编码的声学基础	6
2.1 语音信号的时域波形	6
2.2 语音的产生机理与模型	7
2.2.1 肺与喉	8
2.2.2 声道	9
2.2.3 语音信号的产生模型	9
2.3 语音信号的声学特性	10
2.3.1 语音信号的分类	10
2.3.2 基音与四声	11
2.4 人耳听觉特性	12
2.4.1 正常人的听域与听阈	12
2.4.2 响度级 (Loudness Level) 和响度 (Loudness)	12
2.4.3 音调	14
2.4.4 掩蔽效应 (Masking Effect)	14
2.4.5 言语听觉	15
参考文献	17
第 3 章 语音编码中的数字分析技术	18
3.1 短时分析技术基础	18
3.1.1 预滤波、采样、A/D 变换	18
3.1.2 存储和加窗	18

3.2	短时能量、短时平均幅度和短时过零率	19
3.3	短时自相关函数和短时频谱	19
3.4	短时基音周期分析	21
3.5	语谱图分析	24
3.6	线性预测分析	25
3.6.1	线性预测分析的基本概念	26
3.6.2	线性预测分析的求解	29
3.6.3	德宾 (Durbin) 递推算法	31
3.6.4	LPC 模型阶数 P 和激励增益 G 的确定	40
3.7	各种 LPC 参数以及它们之间的关系	41
	参考文献	48
第 4 章	语音编码中的矢量量化技术	49
4.1	概述	49
4.2	矢量量化的基本原理	49
4.2.1	矢量量化的定义	49
4.2.2	矢量量化的评价指标	51
4.2.3	矢量量化的关键技术	52
4.3	矢量量化器	53
4.3.1	穷尽搜索矢量量化器	53
4.3.2	多级矢量量化器	53
4.3.3	分裂矢量量化器	55
4.3.4	分类矢量量化器	56
4.3.5	形状-增益矢量量化器	57
4.3.6	预测矢量量化器	58
4.3.7	有限状态矢量量化器	59
4.4	矢量量化的码本设计	60
4.4.1	LBG 算法	60
4.4.2	模拟退火码本设计算法	62
4.4.3	其他码本设计算法	63
4.5	矢量量化的码字搜索	63
4.5.1	部分失真搜索算法	64
4.5.2	基于不等式的快速搜索算法	64
4.5.3	带有时变加权的快速搜索算法	67
4.6	矢量量化的码字索引分配	70
4.6.1	基于 BSA 的码字分配算法	70
4.6.2	基于模拟退火的码字分配算法	71
4.7	LSF 参数的矢量量化	71
4.7.1	P-MSMQ 算法	72
4.7.2	带有级间预测的多级矩阵量化算法	74

4.7.3 带有级间预测的多级码本联合优化	76
4.8 多参数多模式码本尺寸联合调整	78
4.8.1 矢量参数多模式量化	78
4.8.2 参与码本尺寸联合调整的参数	79
参考文献	84
第 5 章 语音编码的评价指标	86
5.1 评价语音编码的考虑要素	86
5.1.1 说话人和语种	86
5.1.2 信号能量	87
5.1.3 背景噪声	87
5.1.4 多次转接	88
5.1.5 信道误码	88
5.1.6 非语音信号	88
5.2 语音编码的主观测量	89
5.2.1 音质评价方法	89
5.2.2 可懂度评价方法	90
5.3 语音编码的客观测量	90
5.3.1 信噪比与谱失真	90
5.3.2 PSQM 与 PESQ	91
5.4 语音编码的复杂度和时延	91
参考文献	92
第 6 章 语音信号的波形编码	94
6.1 语音信号的标量量化	94
6.1.1 均匀量化	96
6.1.2 非均匀量化	98
6.1.3 自适应量化	105
6.2 自适应差分脉冲编码	110
6.2.1 DPCM 的工作原理	110
6.2.2 自适应预测	114
6.2.3 CCITT G.721 高质量 32 kb/s ADPCM	116
6.2.4 镶嵌式 ADPCM	122
6.3 变换域编码	124
6.3.1 概述	124
6.3.2 最佳正交变换——KLT	130
6.3.3 离散余弦变换 (DCT)	132
6.3.4 自适应比特分配	134
参考文献	136

第 7 章 语音信号的参数编码	138
7.1 线性预测编码的基本原理	138
7.2 LPC-10 声码器	139
7.3 LPC-10e 声码器	144
7.3.1 激励源的改善	144
7.3.2 基音提取方法的改进	147
7.3.3 声道滤波器参数量化的改进	147
7.3.4 LSF 参数的矢量量化	148
7.4 本章小结	149
参考文献	149
第 8 章 语音信号的混合编码	151
8.1 混合编码的基本原理	151
8.1.1 合成分析法 ABS (Analysis By Synthesis)	151
8.1.2 感觉加权滤波器	151
8.1.3 中、低速率混合编码	153
8.2 多脉冲及规则脉冲激励线性预测声码器 (MPLPC 及 RPE-LPC)	154
8.2.1 多脉冲线性预测声码器	154
8.2.2 规则脉冲激励线性预测声码器 (RPE-LPC)	158
8.3 码激励线性预测声码器 (CELP)	168
8.3.1 CELP 编码算法	168
8.3.2 CELP 码本搜索算法	169
8.3.3 美国政府标准 4.8 kb/s CELP 声码器 FED-STD-1016	170
8.3.4 美国 EIA/TIA 8 kb/s VSELP 声码器	173
8.3.5 16 kb/s 短延时码激励线性预测 (LD-CELP) 声码器	178
参考文献	183
第 9 章 低速率语音编码	185
9.1 波形插值编码	185
9.1.1 波形插值编码原理	186
9.1.2 特征波形的提取	186
9.1.3 特征波形的描述	186
9.1.4 特征波形的对齐	188
9.1.5 特征波形的量化	189
9.1.6 特征波形的重建和线性预测合成	190
9.2 正弦变换编码	192
9.2.1 STC 编码参数的提取	192
9.2.2 STC 编码合成算法	194
9.2.3 正弦变换的低速率编码模型	195

9.3 混合激励线性预测编码	196
9.3.1 MELP 编码器原理	196
9.3.2 MELP 编码参数量化	199
9.3.3 MELP 解码器原理	201
9.3.4 MELPe 算法	204
9.4 正弦激励编码	204
9.4.1 SELP 编码器原理	204
9.4.2 SELP 模型参数分析算法	205
9.4.3 SELP 解码器原理	207
9.5 多带激励编码	209
9.5.1 多带激励语音模型	209
9.5.2 多带激励语音模型参数的提取	210
9.5.3 多带激励语音合成	217
9.6 极低速率语音编码	219
9.6.1 速率在 400~1 200 b/s 左右的编码技术	220
9.6.2 速率在 150~300 b/s 左右的编码技术	223
9.7 低速率语音编码展望	224
参考文献	225
第 10 章 宽带语音编码	228
10.1 概述	228
10.2 子带编码	229
10.2.1 子带编码工作原理	229
10.2.2 子带编码的比特分配	231
10.2.3 整数带滤波器组	233
10.2.4 正交镜像滤波器组	235
10.3 G.722 宽带语音编码	238
10.3.1 G.722 编码器原理	239
10.3.2 G.722 解码器原理	240
10.4 频带扩展 (BWE) 技术	242
10.5 联合语音音频编码 (USAC)	244
10.5.1 USAC 的组合型算法	244
10.5.2 USAC 的统一型算法	245
10.6 宽带语音编码展望	246
参考文献	246
第 11 章 变速率语音编码	247
11.1 变速率语音编码的发展现状	247
11.2 变速率语音编码的相关技术	248
11.2.1 语音端点检测技术 (VAD)	248

11.2.2	速率判决技术	252
11.2.3	差错隐藏技术	253
11.2.4	舒适噪声生成技术	256
11.3	AMR 变速率语音编码	259
11.3.1	AMR 编码器原理	259
11.3.2	AMR 参数表示及量化	263
11.3.3	AMR 解码器原理	269
11.3.4	AMR 帧结构	272
11.4	变速率语音编码展望	272
	参考文献	273
第 12 章 嵌入式语音编码		274
12.1	概述	274
12.2	嵌入式语音编码的发展现状	275
12.3	嵌入式语音编码的有关理论	275
12.3.1	设计理念与编码器的最优性	276
12.3.2	分层矢量量化	278
12.4	嵌入式语音编码的实现	279
12.4.1	嵌入式 DPCM 编码	279
12.4.2	嵌入式 CELP 编码	280
12.4.3	现有的 CELP 编码的嵌入式扩展	281
12.4.4	嵌入式参数量化	282
12.5	嵌入式语音编码标准 G.729.1	282
12.5.1	G.729.1 编码器原理	283
12.5.2	G.729.1 解码器原理	287
12.5.3	G.729.1 性能分析	287
12.5.4	G.729.1 应用: 宽带电话系统	288
12.6	嵌入式语音编码在网络中的应用要素	289
12.6.1	伸缩性的实现和利用	289
12.6.2	非等重错误保护及加密	291
12.7	嵌入式语音编码展望	291
	参考文献	292
第 13 章 语音编码的关联技术		293
13.1	语音增强	293
13.1.1	语音增强算法的基本原理与分类	293
13.1.2	基于短时谱估计的语音增强算法	294
13.2	回声抵消	298
13.2.1	电学回声抵消	299
13.2.2	声学回声抵消	300

13.2.3	回声抵消中的语音检测	302
13.3	语音通信的同步	303
13.3.1	帧同步系统	303
13.3.2	帧同步系统的进一步改进	306
13.4	语音通信的抗信道误码方案	306
13.4.1	信道编码	307
13.4.2	差错检测与恢复	313
13.5	网络语音通信的抗丢包方案	315
13.5.1	语音算法的丢包鲁棒性	316
13.5.2	多描述语音编码	316
13.5.3	滑动窗算法	317
13.5.4	交织及前向纠错技术	318
13.5.5	丢包隐藏技术	318
13.6	本章小结	318
	参考文献	318
第 14 章	语音编码系统的硬件实现	320
14.1	概述	320
14.2	数字信号处理器	320
14.3	基于 DSP 的语音编码系统设计	321
14.3.1	TMS320C54X/55X 系列数字处理芯片	322
14.3.2	语音编码系统硬件平台的组成	322
14.3.3	硬件平台的音频接口电路	323
14.3.4	硬件平台的控制电路	324
14.3.5	语音编码系统低功耗设计	327
14.4	DSP 开发工具与开发环境	328
14.4.1	DSP 开发工具	328
14.4.2	DSP 开发环境	329
14.5	DSP 软件编程技巧	329
14.5.1	DSP 软件定点化	329
14.5.2	DSP 软件优化	332
	参考文献	334
第 15 章	语音编码的标准	335
15.1	语音编码的国际标准	335
15.2	语音编码的区域或行业标准	337
15.2.1	北美标准化组织	337
15.2.2	欧洲标准化组织	338
15.2.3	日本标准化组织	339
15.2.4	中国标准化组织	340

15.2.5	第三代合作伙伴计划 (3GPP, 3GPP2)	340
15.2.6	国际海事卫星组织	341
15.2.7	保密通信标准	342
15.3	VOIP 中应用的语音编码算法	342
15.3.1	iLBC 算法	343
15.3.2	iSAC 算法	343
15.3.3	Speex 算法	344
15.3.4	SILK 算法	344
	参考文献	345

第1章 绪 论

1.1 概述

语音是人类交流信息最自然、最有效、最方便的手段,因此大多数通信系统都要包含语音传输的功能。随着科学技术的发展,虽然图像、数据等非语音信息在信息传递中所占的比例越来越大,但是有效地传递语音信息仍然是众多通信系统必备的功能之一。

语音通信最早可以追溯到 1876 年贝尔发明的电话机,该装置首次利用声电、电声转换技术实现了远距离的语音信号传输。早期的语音通信装置采用的是模拟语音通信,但它在可靠性、抗干扰能力、快速交换等方面存在很多问题。自从 20 世纪 60 年代集成电路出现以后,通信信道和通信终端开始从模拟系统演变为全数字系统。与模拟通信相比,数字通信在可靠性和抗干扰能力等方面具有很大优势,但不足之处是占用的频带较宽。为了节省数字语音信号传输所需要的带宽,语音编码技术开始得到深入研究。

语音编码属于信源编码范畴,自从 20 世纪 30 年代末提出脉冲编码调制(PCM)原理以及声码器(Vocoder)概念后,语音编码一直沿着这两个方向发展,它们也可以称为语音信号的波形编码与声码化编码,或称为非参数编码与参数编码,参数编码有时也称为模型编码。

随着计算机技术的发展以及大规模、超大规模集成电路的出现,语音编码技术取得了很多突破性的进展,研究出了很多实用的语音编码方案,应用在不同的语音通信系统中,并在不断地研究、改进中得以完善,逐渐形成国际性或者地区性的语音编码标准。

语音编码领域发展非常迅速,相关文献浩如烟海。本书在讲述语音编码基础理论及相关语音编码算法的同时,力争将近十几年来语音编码领域的研究热点展现给读者,同时用前瞻性的眼光提出语音编码领域的未来发展趋势。

1.2 语音编码的应用

几乎语音编码领域的所有研究课题都是受应用驱动的,由于数字化语音的传输和存储,在可靠性、抗干扰、速交换、易保密等方面都远远胜过模拟语音,从 20 世纪 50 年代以来,在通信系统中数字语音所占的百分比不断增加。目前,在 ISDN(综合业务数字通信网)、卫星通信、移动通信、微波接力通信等系统中无一例外采用的都是数字化语音传输和存储。最近几年,被戏称为“最后一个数字化电子行业”的对讲机领域也开始研究应用数字化语音编码方案了,我国数字对讲机的语音编码标准也在制定当中。

移动通信系统和因特网(Internet)是语音编码技术的两个重要应用领域。移动通信的发展日新月异,除第一代移动通信系统采用模拟式语音传输外,从第二代移动通信系统开始,全部采用数字语音编码方式。目前广泛使用的“全球通”、“神州行”等第二代数字蜂窝移动通信系统采用了 13 kb/s 的 RPE-LTP(Regular Pulse Excitation-Long Term Prediction)语音编码技术。北美第二代数字移动电话(CDMA)采用的是 Qualcomm(高通)公司的 QCELP

(Qualcomm Code Excited Linear Prediction) 语音编码算法。QCELP 语音编码算法不仅可以工作在 4/4.8/8/9.6 kb/s 等固定速率上,还可以变速率地工作于 0.8~9.6 kb/s 之间。在第三代移动通信系统中,变速率语音编码技术得到了广泛应用。TD-SCDMA 标准采用了 AMR (Adaptive Multi-rate) 语音编码技术, WCDMA 标准主要采用了 AMR 语音编码技术, cdma2000 则主要采用了 QCELP 或 EVRC (Enhanced Variable Rate Codec) 语音编码技术。近些年,由于成本低廉,网络电话得到深入研究,目前世界多个标准组织和工业实体提出了很多语音编码方案。其中包括国际电信联盟的 G.711 (速率为 64 kb/s), G.723.1 (速率为 5.3 kb/s 或 6.3 kb/s), G.729A (速率为 8 kb/s) 编码方案。GIPS (Global IP Sound) 公司、Skype 公司等业界企业也有自己的编码方案,如 iLBC、SILK 等编码算法等。考虑到语音数据包在网络环境中传输的特殊性,以及不同通信网络的异构性等因素,嵌入式多速率语音编码算法也在近几年成为一个重要的研究领域。随着网络带宽的不断增长,终端处理能力的不断增强,用户对话音质量要求的提高,宽带、超宽带、全频带语音编码技术得到了广泛的研究,很多传统的窄带语音编码技术得到深入研究并扩充到宽带编码版本上。

除了移动通信系统和因特网外,语音压缩编码技术还广泛应用于保密通信、卫星通信以及水下通信等领域。在这些应用领域中,传输带宽通常较窄,超低速率语音压缩编码算法的研究成为一个热点。

1.3 语音编码的分类

1.3.1 按编码方法分类

按编码方法分类,语音编码可以分为波形编码、参数编码以及混合编码。

语音信号的波形编码力图使重建语音信号的波形维持原语音信号的波形形状。该类编码算法通常将语音信号作为一般的波形信号来处理,具有适应能力强、话音质量好的优点,但需要的编码速率高。脉冲编码调制 (PCM)、自适应增量调制 (ADM)、自适应差分编码 (ADPCM)、自适应预测编码 (APC)、自适应子带编码 (ASBC)、自适应变换编码 (ATC) 等都属于该类编码器。它们分别在 16~64 kb/s 的编码速率上给出高的编码质量,当速率进一步降低时,其性能下降较快。

语音信号参数编码通过对语音信号特征参数的提取及编码,力图使重建语音信号具有尽可能高的可懂度,即保持原语音的语义,而重建信号的波形可能同原始语音信号有较大的差别。此类编码器的优点是编码速率低,可以达到 2.4 kb/s 甚至更低,主要问题是合成语音质量较差,自然度较低,连熟人都一定不能听出讲话人是谁。另外,该类编码器对环境噪声比较敏感,需要在安静的环境才能给出较高的可懂度,对信道误码也比较敏感。通道声码器、共振峰声码器以及目前广泛使用的线性预测声码器都是典型的参数编码。

混合编码采用参数编码和波形编码的混合编码形式。该类算法的共同特点是:先进行线性预测分析,去掉语音信号的短时相关性,然后利用合成分析法及感觉加权均方误差最小原则分析出合适的替代余量信号的最佳激励源,最后对 LP 参数和激励信号源进行编码和传送。由于混合编码方案的激励模型和误差计算与时域波形相联系,使合成语音具有较强的跟踪输入语音变化的性能,从而能够改善合成语音的质量以及抗噪声的能力。

1.3.2 按编码速率分类

按语音编码速率分类是根据语音信号压缩编码后进行存储或传输所需要的数据速率进行分类的一种方法。例如，按照编码后的数据速率，可以将窄带语音编码划分为 5 类：

- (1) 高速率语音编码，其编码速率为 32 kb/s 以上；
- (2) 中高速率语音编码，其编码速率为 16~32 kb/s；
- (3) 中速率语音编码，其编码速率为 4.8~16 kb/s；
- (4) 低速率语音编码，其编码速率为 1.2~4.8 kb/s；
- (5) 超低速率或极低速率语音编码，其编码速率低于 1.2 kb/s。

目前，语音编码的国际标准主要集中于中速率以上，低速率以及超低速率语音编码算法的研究仍然是一个热点。

1.3.3 按编码语音带宽分类

通常将人耳可以听到的频率在 20 Hz~20 kHz 的声波称为音频信号。标准的话路频带语音信号是 300~3400 Hz，若加上少量保护带宽，一般意义上说的标准话路频带宽度应该是 4 kHz，根据奈奎斯特采样定律，带宽为 4 kHz 的语音信号采样频率应为 8 kHz，量化精度一般在 16 位左右。通常所说的窄带语音编码指的就是这种频带范围内的语音信号编码。

随着网络传输带宽的不断增加，语音编码的另一研究领域也开始引起广泛重视，这就是宽带语音编码。宽带语音编码的频率范围通常在 50~7 000 Hz，加上保护带宽，通常为 8 kHz，采样频率一般为 16 kHz。国际电信联盟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，其频率带宽就在 50~7 000 Hz，采样频率为 16 kHz，属于宽带语音编码的范畴。

在多媒体领域中，处理的主要是音频信号，它包括音乐、语音、风声、雨声、鸟鸣声、机器声等。音频编码主要是针对频率范围较宽的音频信号进行的编码，主要应用于数字广播和数字电视广播、消费电子产品等。典型的编码有 MPEG1/MPEG2 的 layer1、2、3 和 MPEG 4 AAC 的音频编码等。按带宽对语音编码划分如表 1-1 所示。

表 1-1 按带宽对语音编码划分

信号类型	频率范围/Hz	采样率/kHz	量化精度/采样位数
电话语音（窄带）	200~3 400	8	13~16
宽带语音（宽带）	50~7 000	16	16
调频广播（超宽带）	20~15 k	32	16
高质量音频（全带）	20~20 k	44.1	16

1.3.4 按编码信号所在域分类

语音信号具有短时平稳性，通常假定语音信号在 5~30 ms 之间的时间段内，某些物理特征参数是可以看做不变的，所以语音信号的参数编码及混合编码通常以语音帧为单位进行。按照被编码的语音信号所在域可以将编码算法分为时域编码和频域编码。时域编码是指对语音的时域信号进行编码，如 PCM 编码等。而频域编码是指对语音的频域信号进行编码，子带编码（SBC）就是一种频域编码方法。

1.4 语音编码的硬件与系统

绝大多数语音编码系统需要实时工作，这对于系统的硬件环境提出了较高的要求。随着语音编码算法的日益复杂，许多算法的运算量达到几十个 Million Instructions Per Second (MIPS)。但相应的，近 20 年来，随着微电子行业的迅猛发展，DSP（数字信号处理器）技术也产生了巨大的飞跃，这就为语音编码算法的实用化铺平了道路。在 DSP 领域最成功的企业是美国的德州仪器公司（Texas Instruments, TI 公司），自从 1982 年推出 TMS320 系列的第一款定点 DSP TMS3210 以来，其 DSP 已经发展了若干系列，处理能力最高可达 GHz。除 TI 公司外，美国 AD 公司、Motorola、NEC 等公司的 DSP 也各有特点。具体选择哪一种硬件平台来实现语音编码，要根据应用的实际需求综合考虑，灵活决定。

根据 DSP 所能处理的数据格式，可以将其分为定点 DSP 和浮点 DSP。定点 DSP 速度较快且价格便宜，但编程困难，精度较低；浮点 DSP 精度高，但其价格通常较贵。在为某种语音编码算法选择合适的 DSP 时，需要综合考虑运算量、存储量以及性价比等众多因素。在语音编码领域，TI 的 TMS320C54X 及 TMS320C55X 系列 DSP 得到了广泛应用，这两款 DSP 均为定点 DSP。

1.5 语音编码的发展趋势

经过近几十年的发展，语音编码领域涌现了很多的成果。目前在语音编码领域有几个研究热点：低速率语音编码、宽带语音编码、变速率语音编码以及嵌入式语音编码。

ITU 已经制定的语音编码标准主要集中在中速率以上。如何在低速率甚至超低速率上获得较高的语音质量，仍然是语音编码领域的热点。在低速率语音编码领域，目前研究较多的模型包括混合激励线性预测模型、正弦激励线性预测模型、正弦变换编码、多带激励编码以及波形插值编码等。

除了致力于制定更低速率的语音编码标准外，ITU 正在对已经制定的语音编码标准进行全频带扩展，使其能够适应音频的应用，例如 AMR 从窄带 AMR-NB 到宽带 AMR-WB 再到超宽带 AMR-WB+，G.729 全频带扩展到 G.729.1 等。ISO/MPEG 标准化组织目前正在制定下一代音频编/解码标准 USAC（Unified Speech and Audio Coding），即语音/乐音联合编解码器。USAC 可以对以任意比例混合的语音/乐音信号进行编/解码，同时，无论是语音成分还是乐音成分，其编码性能至少不亚于当前最好的专业语音编码器或乐音编码器的编码性能。USAC 标准具有广泛的应用前景，可以广泛地应用于全球数字广播，第三、四代移动通信等领域。

近年来，随着移动通信的发展，尤其是第三代移动通信的发展，对语音压缩编码算法提出了更高的要求，不但要求编码速率较低以增加系统容量，而且要求合成语音音质较高以保证通话质量。用传统的压缩编码方式，很难同时满足这两个要求，在这种形势下，提出了变速率语音压缩编码的方法。在移动通信系统中采用变速率语音压缩编码，可以根据需要动态地调整编码速率，在合成语音质量和系统容量中取得灵活的折中，最大限度地发挥系统的效能。20 世纪 90 年代，随着 CDMA 移动通信系统的飞速发展，各种通信组织相继推出了不同的变速率语音编码技术标准，主要包括 QCELP、EVRC、AMR 和 SMV 4 种算法标准。

随着网络技术的不断发展，以网络为载体的语音信号传输得到了广泛研究。在现实的网络条件下，由于网络拥塞、过延时和缓冲器溢出等原因，决定了网络传输中必然经常出现包丢失等问题，这对语音信号的实时传输具有很大影响，传统的编码器因为本身的特点，包丢