

MODERN
COMMUNICATIONS

现代通信理论与技术丛书

Variable Rate Speech Coding

变速率语音编码

● 王炳锡 王洪 编著

西安电子科技大学出版社
[http:// www.xduph. com](http://www.xduph.com)

现代通信理论与技术丛书

变速率语音编码

王炳锡 王洪 编著

西安电子科技大学出版社

2004

内 容 简 介

本书主要讲解了目前最前沿的变速率语音编码及其相关技术。首先简单介绍了变速率语音编码技术的历史与发展现状,再简明地讲解了语音编码的基本原理及其性能评价方法,对在变速率语音编码中最常用的码激励线性预测原理进行了单独介绍,然后对变速率语音编码的相关技术及其背景进行了比较性的讲解,最后对 QCELP、EVRC、AMR 和 SMV 四种变速率语音编码技术的原理进行了详细的、深入浅出的讲解,并根据我们的研究与实验结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。该书是国内第一本讲解变速率语音编码技术的书籍,具有很高的实用价值和参考价值。

全书既简明扼要地介绍了语音压缩编码的基本原理,又详细地讲解了国际上最新、最流行的变速率语音编码标准,章节之间紧密配合、前后呼应,具有很强的系统性。通过使用本书,读者可以了解目前最前沿的变速率语音编码技术,掌握语音编码的基本知识和理论,获得语音编码技术的基本技能。同时,通过书中反映出来的研究过程和研究方法,读者能够在以后的研究工作中得到很大的启发。

本书可作为高等学校理工科通信和信息处理及相关专业的高年级本科生和研究生(硕士、博士)的教材和参考书,也可供相关专业的工程技术人员和科研人员参考。

图书在版编目(CIP)数据

变速率语音编码=Variable Rate Speech Coding/王炳锡,王洪编著.

—西安:西安电子科技大学出版社,2004.6.

ISBN 7-5606-1382-9

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

中国版本图书馆 CIP 数据核字(2004)第 028690 号

策 划 臧延新

责任编辑 龙 晖 臧延新

出版发行 西安电子科技大学出版社(西安市太白南路2号)

电 话 (029)88242885 88201467 邮 编 710071

http://www.xduph.com E-mail: xdupfxb@pub.xaonline.com

经 销 新华书店

印刷单位 西安文化彩印厂

版 次 2004年6月第1版 2004年6月第1次印刷

开 本 787毫米×1092毫米 1/16 印张 22.5

字 数 528千字

印 数 1~4 000册

定 价 34.00元

ISBN 7-5606-1382-9/TN·0265

XDUP 1653001-1

如有印装问题可调换

本社图书封面为激光防伪覆膜,谨防盗版。

序

通信网是国家的重要基础设施之一,通信对国民经济和社会发展的影响非常深远。目前通信理论和技术发展很快,通信网在我国也得到了迅猛的发展。改革开放 20 多年来,我国的电话普及率已增长超过 100 倍。目前我国的第二代蜂窝网(GSM)规模居世界第一位;互联网的上网人数居世界第二位。与此相应,在通信领域从事科技工作的人员数量也大为增加。

在这种形势下,新生力量的培养和现有人员的继续教育就成为一项不可忽视的重要任务。加强通信领域新理论与新技术的传播,更是培养通信人才所必须。为了将通信理论和技术最新进展和发展前景及时地介绍给广大科技工作人员和学生,西安电子科技大学出版社依托本校的学科优势,邀请国内通信领域具有丰富理论和实践经验的专家、教授,成立了“现代通信理论与技术丛书”编委会,负责规划这套丛书的出版工作,力争能够出版一批具有较高学术水平和实用价值的著作。

这套丛书主要面向从事通信领域研究和开发的科技工作者、工程师、高年级大学生和在读研究生,以及希望了解该领域发展的各类相关人员。丛书突出内容新颖、全面系统、理论联系实际的特色,在简明扼要讲述物理概念和基本理论的基础上,尽可能地引导读者运用理论解决应用中的实际问题。为了满足目前研究生教学的迫切需要,一些优秀的研究生教材也将纳入到这套丛书中。

这套丛书的编写和出版将是一项艰巨的工作,感谢为出版这套丛书提供了支持和帮助的专家们,他们严谨的态度和深厚的学识是这套丛书顺利出版的保障。也衷心地希望广大读者和有关人员提出宝贵的意见和建议,使这套丛书日臻完善。

樊昌信

现代通信理论与技术丛书编委会

主 任：樊昌信

副主任：王育民 裴昌幸 朱世华

委 员：（以姓氏笔画为序）

王永生 王新梅 卢朝阳 刘乃安

刘增基 张 辉 吴成柯 李兵兵

李建东 易克初 苟彦新 杨家玮

常义林 曾兴雯 葛建华 鲍长春

廖桂生

前 言

《变速率语音编码》是《语音编码》一书的姊妹篇。变速率语音编码是继定速率语音编码之后语音压缩编码技术的新发展。两者的理论基础是相同的,但变速率语音编码技术有了新的发展和进步。

变速率语音编码的特点是“变”:一是为什么要“变”,二是要有可能“变”,三是怎么“变”,四是“变”的效果如何。广大的科技工作者在回答这四个问题时做了大量的工作,历经近十年的努力,这些问题几乎全面地被解决了。各种标准协议的公布以及变速率语音编码技术在移动通信、互联网、IP 电话和语音存储等领域得到实际应用,为回答这些问题给出了最好的答案。

人类在进行语音通信时,大约有 70% 左右的空闲时间没有讲话,始终用一个速率进行语音编解码对信道资源是一个浪费。因此,能否在无话时编码速率低一点,讲话时编码速率高一点,使平均速率降低下来,这个朴素的想法是有可能实现的。因为无话段的信道噪声和有话段的声信号特征是截然不同的,语音激活技术可以得到有效的应用。大家知道,现代通信技术越来越复杂,影响通信质量的因素很多,综合考虑各种因素可得到三种速率控制方式:源控制、信道控制和网络控制。这些速率控制方式可以使收发双方形成默契,并使速率变化有规律可循,以达到降低速率的目的。传输速率降下来后,接收方要得到满意的合成语音,这是变速率压缩编码的关键。在这里除了利用一般的语音合成技术、参数内插技术和后置滤波技术之外,还利用了速率判决技术、语音激活检测技术、舒适背景噪声技术、差错隐藏技术等。这些技术的应用使变速率之后的语音合成效果几乎没有降低。

20 世纪 90 年代,在无线通信的多址系统中,变速率语音编码得到了广泛的应用,如 E-TDMA、PRMA,尤其是 CDMA 移动通信的飞速发展,把变速率语音编码推上了前台。本书旨在系统研究 QCELP、EVRC、AMR 和 SMV 四种算法标准,从原理上给予深刻分析,从实践上给予模拟,比较各种标准的优劣,从编解码算法的各个环节上探讨其算法的改进,为广大读者提供研究资料、思路和方法,以利于变速率语音编码技术得到更好的发展。

变速率语音编码的每一种标准都有多种速率可选,这样不仅仅是从语音压缩编码考虑,更重要的是要配合 CDMA 系统的需要,既要编码速率低,又要合成语音质量好,同时最大限度地发挥系统的效能,这就要求速率判决算法灵活。简单地说,平均速率的降低是用算法复杂度换来的。

当然,更低速率的语音编码技术已有报导,如 800 b/s、600 b/s 语音压缩编码算法。作者现场测听过编解码效果,认为其音质、可懂度都很好。不要很久,相信其中的一些创新技术在未来极有可能融合在变速率语音编码中。

本书的结构安排如下:全书共分八章,第一章简要介绍了定速率语音编码向变速率语音编码的发展历程和变速率语音编码技术的现状与发展方向;第二章简明地讲解了语音编

码的基本原理及其性能评价方法；第三章对在变速率语音编码中最常用的码激励线性预测原理进行了单独介绍；第四章比较性地讲解了各种变速率语音编码的相关技术；第五、六、七、八章分别对 QCELP、EVRC、AMR 和 SMV 四种变速率语音编码技术的原理进行了详细的、深入浅出的讲解和研究。书末是“英汉名词对照”，其中收录了语音编码技术涉及的部分专业名词。第一章由王炳锡编写，第二、六章由丁琦编写，第三、五章由李莉编写，第四、七章由王洪编写，第八章由曹建军编写。全书内容由王炳锡统筹指导，由王洪整理修改，最后由王炳锡校订统稿。

本书的特点是讲解了目前最前沿的变速率语音编码技术，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；由浅入深地安排章节，先理论后应用，内容编排科学，知识结构合理，章节之间紧密配合、前后呼应，具有很强的科学性和系统性；紧密结合作者自己的研究与实验成果，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。通过学习本书，读者可以了解目前最前沿的变速率语音编码技术，掌握语音编码的基本知识和理论，获得语音编码技术的基本技能。另外，通过我们在书中反映出来的启发式研究过程，读者既可以提高自学能力，又可以在创新思维方面得到很大的提高。本书可作为大学高年级本科生、硕士生的选修课本或参考书，也可作为科研人员的参考书。

本书的编著得到了解放军信息工程大学各级领导的关心和支持，得到了国内广大学者的支持和帮助。尤其是课题组同志和研究生们的研究成果充实、完善了本书的内容，增强了本书的可读性和可操作性。西安电子科技大学出版社为本书的出版付出了大量的劳动，在此表示衷心的感谢。书中引用了大量的文献资料，在此也向原作者表示深深的谢意。

因本人的研究水平和阅历有限，书中错漏之处在所难免，恳切希望读者不吝赐教，与作者联系，共同研讨问题。联系地址为河南省郑州市 1001 信箱 837 号(邮编 450002)。

王炳锡

2004 年 3 月

于解放军信息工程大学

目 录

第一章 绪论	1
1.1 概述	1
1.2 语音编码算法综述	2
1.3 定速率语音编码向变速率语音编码的过渡	3
1.4 变速率语音编码的发展现状	4
1.5 变速率语音编码的速率控制方式	6
1.6 变速率语音编码的发展方向	6
参考文献	7
第二章 语音编码基础	8
2.1 概述	8
2.2 语音信号及其模型	8
2.2.1 发音的生理机构与过程	8
2.2.2 语音听觉器官的生理	9
2.2.3 语音听觉的心理和两种听觉实验	9
2.2.4 语音信号模型	11
2.3 语音信号的线性预测分析	14
2.3.1 线性预测的基本原理	16
2.3.2 线性预测误差滤波	16
2.3.3 语音信号的线性预测分析	18
2.3.4 线性预测分析的解法	20
2.3.5 线性预测分析的应用	22
2.4 语音信号的矢量量化	24
2.4.1 矢量量化的基本原理	31
2.4.2 失真测度	31
2.4.3 快速搜索问题	32
2.4.4 最佳矢量量化器和码本的设计	34
2.4.5 降低复杂度的矢量量化系统	36
2.4.6 有记忆的矢量量化	38
2.5 语音信号编码	40
2.5.1 基本原理	41
2.5.2 语音信号的剩余度	41
2.5.3 两种编码方法	42
2.6 语音编码的性能评价方法	43
2.6.1 编码速率	44
2.6.2 合成语音质量	44

2.6.3 编解码延时	46
2.6.4 算法复杂度	46
2.6.5 其它性能	46
2.7 小结	47
参考文献	47
第三章 码激励线性预测声码器	49
3.1 概述	49
3.2 CELP 编码原理	49
3.3 知觉加权	51
3.4 CELP 码本搜索	52
3.4.1 CELP 码本搜索算法	53
3.4.2 自适应码本搜索	54
3.4.3 固定码本搜索	57
3.5 自适应后置滤波	58
3.5.1 短时后置滤波器	59
3.5.2 频谱补偿滤波器	60
3.5.3 长时后置滤波器	61
3.5.4 自动增益控制(AGC)	61
3.6 小结	62
参考文献	62
第四章 变速率语音编码的相关技术	63
4.1 概述	63
4.2 语音激活检测(VAD)技术	63
4.2.1 VAD 技术的基本原理	63
4.2.2 VAD 参数的计算	64
4.2.3 VAD 判决方法	71
4.2.4 变速率语音编码中 VAD 技术的比较	72
4.3 速率判决(RDA)技术	73
4.3.1 RDA 技术的基本原理	73
4.3.2 信源控制速率(SCR)的原理	74
4.3.3 信道控制速率(CCR)的原理	81
4.3.4 变速率语音编码中 RDA 技术的比较	84
4.4 差错隐藏(ECU)技术	85
4.4.1 ECU 技术的基本原理	85
4.4.2 差错检测的方法	85
4.4.3 差错隐藏的方法	86
4.4.4 变速率语音编码中 ECU 技术的比较	93
4.5 舒适背景噪声(CNA)生成技术	93
4.5.1 CNA 生成技术的基本原理	93
4.5.2 AMR 中的 CNA 生成技术	93
4.6 小结	96

参考文献	97
第五章 QCELP 变速率语音编码	98
5.1 概述	98
5.2 TIA/EIA/IS-96C 8k QCELP 概述	98
5.2.1 编码	100
5.2.2 解码	101
5.2.3 延时要求	101
5.3 8k QCELP 编码原理	102
5.3.1 预处理	102
5.3.2 线性预测分析和量化	102
5.3.3 知觉加权	108
5.3.4 速率判决	108
5.3.5 自适应码本搜索	109
5.3.6 固定码本搜索	111
5.3.7 编码器的存储器更新	119
5.3.8 空帧	120
5.3.9 参数格式化输出	120
5.4 8k QCELP 解码器功能说明	124
5.4.1 解码器的工作流程	124
5.4.2 后置处理	125
5.5 13k QCELP 编解码原理	126
5.5.1 TIA/EIA/IS-733 13k QCELP 概述	126
5.5.2 LSP 参数的量化	128
5.5.3 速率判决	135
5.5.4 基音延迟的量化	137
5.5.5 固定码本搜索	137
5.5.6 参数格式化输出	143
5.5.7 13k QCELP 解码器	147
5.6 QCELP 编解码器工作流程总结	149
5.6.1 8k QCELP 编解码器工作流程	149
5.6.2 13k QCELP 编解码器工作流程	153
5.7 QCELP 性能测试及模拟结果	156
5.7.1 平均编码速率测试	156
5.7.2 合成语音质量测试	157
5.7.3 模拟结果	157
5.7.4 QCELP 与其它语音编码算法的比较	159
5.8 小结	160
参考文献	160
第六章 增强型变速率语音编解码器(EVRC)	161
6.1 概述	161
6.2 编码器原理	162

6.2.1	预处理	163
6.2.2	模型参数估计	167
6.2.3	速率判决算法	170
6.2.4	LSP 参数的量化	172
6.2.5	各种速率下的编码量化过程	173
6.2.6	数据包格式	185
6.3	解码器原理	186
6.3.1	差错隐藏技术	186
6.3.2	全速率和半速率解码	187
6.3.3	1/8 速率解码	189
6.3.4	后置滤波器	190
6.4	EVRC 性能测试及模拟结果	192
6.4.1	平均编码速率测试	192
6.4.2	合成语音质量测试	193
6.4.3	模拟结果	193
6.4.4	EVRC 与其它语音编码算法的比较	194
6.5	EVRC 算法特色及改进	194
6.5.1	噪声抑制算法	194
6.5.2	速率判决算法	195
6.5.3	RCELP 算法	195
6.5.4	量化和码本搜索	196
6.5.5	对 EVRC 算法的改进	197
6.6	小结	200
	参考文献	201

第七章	自适应多速率 (AMR) 语音编码	202
7.1	概述	202
7.2	编码器原理	205
7.2.1	预处理	205
7.2.2	线性预测分析和量化	206
7.2.3	开环基音分析	229
7.2.4	脉冲响应的计算	233
7.2.5	目标信号的计算	233
7.2.6	自适应码本搜索和增益控制	233
7.2.7	代数码本的结构和搜索	237
7.2.8	自适应码本增益和固定码本增益的量化	243
7.2.9	修改存储器	245
7.3	解码器原理	245
7.3.1	解码和语音合成	246
7.3.2	后置处理	248
7.4	AMR 语音编码器帧结构	249
7.4.1	AMR IF1 帧结构	249
7.4.2	AMR IF2 帧结构	264

7.5	AMR 性能测试及模拟结果	265
7.5.1	平均编码速率测试	266
7.5.2	合成语音质量测试	266
7.5.3	模拟结果	267
7.6	采用 AMR 的实际系统模型	270
7.7	小结	271
	参考文献	271
第八章	可选模式声码器(SMV)	273
8.1	概述	273
8.2	编码算法原理	275
8.2.1	预处理	276
8.2.2	线性预测(LPC)分析	283
8.2.3	参数内插	285
8.2.4	感觉加权滤波	288
8.2.5	开环基音分析	289
8.2.6	语音帧分类	290
8.2.7	信号修正	295
8.2.8	速率和编码类型选择	302
8.2.9	自适应码本搜索	304
8.2.10	固定码本搜索	307
8.2.11	量化机制	316
8.3	解码算法原理	323
8.3.1	解码器算法概述	323
8.3.2	LSP 的解码	323
8.3.3	激励解码	324
8.3.4	差错隐藏算法	325
8.3.5	后处理	327
8.4	SMV 中的关键技术	329
8.4.1	概述	329
8.4.2	信号修正机制	330
8.4.3	增益判定机制	331
8.4.4	固定码本结构	331
8.5	SMV 性能测试及模拟结果	332
8.5.1	平均编码速率测试	332
8.5.2	合成语音质量测试	332
8.5.3	模拟结果	333
8.5.4	SMV 与其它语音编码算法的比较	337
8.6	小结	337
	参考文献	338
附录	英汉名词对照	339

第一章 绪 论

1.1 概 述

语言是人类进行交流的重要手段,通信系统中最常见的数据形式就是语音。语音通信是人类通信最基本、最重要的方式之一。随着移动通信与互联网的飞速发展,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化,且灵活方便,易于保密,价格低廉,所以从 20 世纪 50 年代以来,数字化语音在通信系统中所占的比重越来越大。语音编码是数字语音通信中的一项重要技术。为了压缩数字语音传输的比特率,以使同样的信道容量能传输更多路的语音信号,节省存储空间,语音压缩编码也有了很大的发展,并在有线/无线电话的话带语音信号、会议电视的宽带语音信号、HDTV 和高保真音乐等的音频信号等领域有广泛的应用。

通信系统都是围绕着通信传输的数量与质量两个类型的三种指标(有效性、可靠性和安全性)进行不断优化的。有效性是指占用尽可能少的信道资源(如频段、时隙和功率)传送尽可能多的信源信息,它是通信的数量指标;可靠性主要是指在传输中抵抗各类客观自然干扰的能力,但是在军事通信中它也包含电子对抗;安全性则指的是传输中的安全保密性能,即收端防窃听、发端防伪造和篡改的能力等。移动通信中的各类新技术,都是以解决移动通信中的有效性、可靠性和安全性为目标而设计的。

移动通信属于无线通信。在无线通信中有效性的要求显得非常突出,这是由于无线信道的频率资源是有限的。提高移动通信的有效性主要通过改进信源(语音)编码技术来完成。近年来,通信系统发展迅速,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,不但要求编码码率较低以增加系统容量,而且要求合成音质较高以保证通话质量。用传统的编码方式,很难同时满足这两个要求。在这种形势下,人们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。由于移动通信市场竞争异常激烈,因此对变速率语音编码的研究成为一个热点。近几年来,变速率语音编码技术发展得很迅速,并不断有新的国家标准和国际标准公布。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、互联网等方面也有很好的应用前景。

1.2 语音编码算法综述

语音编码大致分为四种方式：时域波形编码、变换域编码、参数编码和混合编码。

1. 时域波形编码

时域波形编码不基于声学模型，只针对语音波形进行编码。这种方法在降低量化每个语音样本比特数的同时，又保持了相对良好的语音质量。波形编码主要有脉冲编码调制(PCM)、增量调制(DM)、自适应增量调制(ADM)、自适应差分脉码调制(ADPCM)和自适应预测编码(APC)等。线性 PCM 是用同等的量化级进行量化的，没有利用声音的性质，所以信息没有得到压缩；对数 PCM 利用了语音信号幅度的统计特性，对幅度按对数变换压缩，将压缩的结果做线性编码，在接收端解码时，按指数进行扩展，这种方法在数字电话通信中得到了广泛的应用，现有的 PCM 采用编码速率为 64 kb/s 的 A 律或 μ 律对数压缩方法。由于对数 PCM 广泛应用于通信系统中，而线性 PCM 可以直接进行二进制运算，因此一般速率低于 64 kb/s 的语音编码系统多是先进行对数 PCM 到线性 PCM 变换后，再进行语音信号数字处理的。PCM 的最大缺点是数码率高，在传输时所占频带较宽。差分脉码调制(DPCM)是根据相邻采样值的差值信号进行编码的，ADPCM 是在 DPCM 的基础上发展起来的，其量化器与预测器的参数能根据输入信号的统计特性自适应于最佳或接近于最佳参数状态。ADPCM 是语音编码中复杂程度较低的一种方法。增量调制是指根据信号的增量进行编码，用一比特量化前后样点的增量。这种方法较简单，实现容易，但由于量阶固定，因此当信号变化幅度大或者是小信号时，信噪比(SNR)会下降。为了改进这种方法的动态范围，引进了自适应技术，让量阶的大小随输入信号的统计特性变化而变，这种方法称为自适应增量调制(ADM)。连续可变斜率增量调制(CVSD)是让量阶的大小随音节时间间隔(5~20 ms)中信号的平均斜率变化，信号的斜率是通过输出连“0”或连“1”来表示的。ADM 编码器简易、同步简单、成本低，是数字移动通信中较好的一种语音编码方法。APC 是根据语音的统计特性，由信号过去的采样值预测出当前样值的一种编码方法，它是通过自适应预测器来提高预测精度的，预测得越精确，编码速率越低，这种方法可以做到低速率(10 kb/s 以下)，并且音质与电话音质相似。

2. 变换域编码

变换域编码方式也是不基于声学模型的编码方法，主要有子带编码(SBC)和自适应变换编码(ATC)。SBC 利用带通滤波器将语音频带分成若干子带，并且分别进行采样、编码，编码方式可以用 ADPCM 或 ADM，SBC 速率可以达到 9.6 kb/s。可变 SBC 可使子带的设计不固定，而是随共振峰变化，使编码效率进一步提高，这种方式在码率为 4.8 kb/s 时可具有相当于 7.2 kb/s 的固定 SBC 的语音质量。ATC 先将语音信号在时间上分段，每一段信号一般有 64~512 个采样，再将每段时域语音数据经正交变换到频域，得到相应的各组频域系数，然后分别对每一组系数的每个分量单独量化、编码和传输，在接收端解码得到的每组系数再进行频域至时域的反变换，恢复时段信号，最后将各时段连接成语音信号，ATC 编码在速率为 12~16 kb/s 时可得到优质语音。

3. 参数编码

参数编码是基于模型的编码方法,它分析并提取语音信号的特征参数,且只传送能够合成语音信息的参数,通过参数还原语音。典型的声码器有谱带式、共振峰式和按线性预测(LP)分析所组成的声码器等。谱带式声码器发送语音信号的三种信息,其中一种信息是使语音信号通过 10~20 个并联带通滤波器,通过检波得到信号的包络值,再用 50 Hz 或 30 Hz 的帧频传送。另一种信息是声带音调,通过音调控制器从语音中分析出基音频率,并送出相应的电压信号。第三种是清/浊音判决信息,将上述信息通过采样、量化、编码、合成再发送出去,在接收端进行清/浊音检测,浊音由一个周期脉冲序列发生器来产生周期脉冲,其频率与基音频率相等,清音由一个伪随机噪声发生器来产生随机脉冲,发生器的输出由浊音、清音检测控制开关的交替通断,再被发送端送来的相应信息调制,就得到合成的语音,其速率可压缩到 2.4 kb/s。共振峰式声码器是利用语音频带中的共振峰信息进行编码的,它的速率可压缩到 1.2 kb/s,这种方法存在的问题是要准确地提取共振峰的频率比较困难。LP 声码器是一种比较有实用价值的声码器,典型的方法如美国国家安全局于 1975 年及 1986 年选定的 LPC-10 及改进型 LP-10e,码率为 2.4 kb/s,用 10 阶线性预测的方法提取声道参数,采用区分浊音和清音的二元激励,清音用白噪声而浊音用周期为基音周期的脉冲序列激励 LPC 合成滤波器合成语音,用这种方法还原出来的语音的清晰度、可懂度仍很高。

4. 混合编码

混合编码结合上述几种编码方式的优点,在保留参数模型技术精华的基础上,应用波形编码准则去优化激励信号,从而在 4.8~9.6 kb/s 的码率上获得了较高质量的合成语音。其代表是合成分析(A-B-S, Analysis-By-Synthesis)线性预测编码。它采用感觉加权技术,在闭环的基础上寻找主观意义上失真最小的激励矢量。由于采用的激励信号模型不同,这类方法派生出的多种新的编码方法都能在 9.6 kb/s 码率上获得较高的语音质量。典型的方法有规则脉冲激励线性预测(RPE_LP)、多脉冲激励线性预测(MPE_LP)和码激励线性预测(CELP)等。MPE_LP 采用数目有限、非均匀间隔的冲激序列作为激励,在 RPE_LP 中信号同样为冲激序列,但其间隔为常数,所以 RPE_LP 除需指明序列的起始位置外,无需对每一个脉冲位置进行描述。CELP 采用矢量量化技术,传送给解码器的是码本索引而不是样值矢量,CELP 能高效地压缩数码率,但是建立码本搜索码字的运算量比较大。

1.3 定速率语音编码向变速率语音编码的过渡

传统的定速率语音编码从总体来讲,较高速率的编码算法对话音质量较易保证,但占用网络资源较大;较低速率的编码算法占用网络资源小,但对话音质量较难保证。

语音激活检测(VAD)技术的出现和发展,使对有无语音进行判断成为可能,从而可以对背景噪声和激活的语音部分以不同的速率进行编码,降低平均速率,也就是采用变速率语音编码的方法。人类在进行语音通信时,大约有 70%左右的空闲时间没有讲话,

始终用一个速率进行语音编解码对信道资源是一个浪费。因此能否在无话时编码速率低一点,讲话时编码速率高一点,使平均速率降低下来?这个朴素的想法是有可能实现的,那就是采用变速率语音编码技术。自从变速率语音编解码算法诞生以来,因为其不但可以根据需要动态调整编码速率,在合成语音质量和系统容量之间取得灵活的折衷,最大限度地发挥系统的效能,而且非常适合分组交换网络,故得到了很大的发展,逐渐成为新的研究热点。

变速率语音压缩编码理论上仍属于 CELP,但在“变”上有了新的研究,由此引入了相关的先进技术。这些相关技术主要包括:用来检测语音通信时是否有话音存在的话音激活检测(VAD, Voice Activity Detector)技术、为突出“变”字而进行速率判决(RDA, Rate Decision Algorithm)的自适应技术、为避免语音帧丢失后带来负面效应的差错隐藏(ECU, Error Concealment Units)技术、为克服背景噪声不连续的舒适背景噪声(CNA, Comfort Noise Aspects)生成技术等等。这些相关技术的应用使变速率语音编码之后的语音合成效果几乎没有降低。

1.4 变速率语音编码的发展现状

移动通信系统从第一代发展到第二代,包含了从模拟语音信道到数字语音信道的转变,用数字化方法研究语音处理技术,使人们能更加有效地产生、传输、存储和获取语言信息,这对于促进社会的发展具有十分重要的意义。作为语音数字化处理的一个重要分支,数字语音编码技术得到了长足发展。随着通信技术的高速发展,频率资源变得更加宝贵,语音编码技术可以压缩语音信号的传输带宽,增加通信系统的容量。实际上,语音信号存在着大量冗余,如信息冗余、时间冗余、谱间冗余、听觉冗余和知识冗余等,采用各种语音编码技术的目的就是为去除语音信号的冗余度,压缩原始语音数据,合成出可懂度和自然度较好的语音。

第三代移动通信系统的特点之一就是采用变速率语音编码技术。变速率编码算法的进一步研究,其算法符合声音的特性,适配分组交换网络,在第三代移动通信系统中具有很好的发展前景。当然,变速率语音编码的传统应用还包括语音存储、分组应用和用于数字电路倍增设备(DCME)的数字语音插空(DSI)。近年来,它在无线通信的多址系统中应用得越来越广泛。例如,在 E-TDMA(Extended-Time Division Multiple Access)、PRMA(Packet Reservation Multiple Access)和 CDMA(Code Division Multiple Access)中都得到了变速率语音编码。

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。20 世纪 90 年代,CDMA 移动通信系统的飞速发展,及其对语音编码技术的进一步要求,把变速率语音编码技术推上了前台。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、AMR-WB(本书不作介绍,有兴趣的读者可参见参考文献[11])和 SMV 五种算法标准,如表 1.1 所示。

表 1.1 当前变速率语音编码所采用的标准协议对照表

算法名称	公布单位或协议	公布时间	编码所用速率/(kb/s)	算法特点
8 kb/s QCELP	美国通信工业协会 (TIA) 公布的 IS-96 协议	1995 年 5 月	包括四种速率: 全速率 8.55、半速率 4.0、1/4 速率 2.0 和 1/8 速率 0.8	这是 CDMA 系统中最早采用的变速率语音编码算法, 它采用单级编码数据速率判决技术, 对线谱频率(LSF)使用标量量化
13 kb/s QCELP	美国通信工业协会 (TIA) 公布的 IS-733 协议	1998 年 3 月	包括四种速率: 全速率 13.3、半速率 6.2、1/4 速率 2.7 和 1/8 速率 1.0	它使用了两级编码数据速率判决技术, 对线谱频率(LSF)使用矢量量化
EVRC	TIA 公布的 IS-127 协议, 规定了增强型变速率编解码 (Enhanced Variable Rate Codec) 算法	1997 年 1 月	使用三种编码速率: 全速率 8.55、半速率 4.0 和 1/8 速率 0.8, 没有使用 1/4 速率编码	这个协议可以很好地解决 IS-96 语音质量不高和 IS-733 码率较高等问题, 它既保持了与 IS-733 相近的合成语音音质, 又具有较低的平均码率
AMR	第三代伙伴计划 (3rd Generation Partnership Project, 简称 3GPP) 公布的自适应多速率 (Adaptive Multi Rate) 语音编解码算法	1999 年 8 月	支持八种速率模式: 12.2、10.2、7.95、7.40、6.70、5.90、5.15 和 4.75	它的基本概念是以更加智能的方式解决信源和信道编码的速率分配问题, 使得无线资源的配置与利用更加灵活和高效
AMR-WB	ITU(国际电信联盟)公布的自适应多速率宽带 (Adaptive Multi Rate - WideBand) 编码器的标准 G.722.2, 其输入原始信号为 50 Hz~7 kHz, 采样率为 16 kHz, 相对于 200 Hz~3.4 kHz 为宽带	2000 年 12 月	支持九种速率模式: 6.6、8.85、12.65、14.25、15.85、18.25、19.85、23.05 和 23.85	它已被 3GPP 选定为 GSM 和 3G 无线 W-CDMA 的宽带编码器, 并将应用于 IP 电话、第三代移动通信、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编码器
SMV	3GPP2 将可选模式声码器 (Selectable Mode Vocoder) 算法公布为宽带扩频通信系统的一种候选变速率语音编码标准	2001 年 12 月	采用四种速率: 全速率 8.55、半速率 4.0、1/4 速率 2.0 和 1/8 速率 0.8	这个标准的特点是采用变速率编码机制, 可以根据工作模式调整平均码率, 可以在合成语音质量和系统容量之间取得灵活的折衷, 可以在保证语音质量的同时最大限度地发挥系统的效能