



国外电子与通信教材系列

# 离散时间语音信号处理

## ——原理与应用

Discrete-Time Speech Signal Processing:  
Principles and Practice

[美] Thomas F. Quatieri 著

赵胜辉 刘家康 谢 湘 等译

匡镜明 审校

电子工业出版社  
Publishing House of Electronics Industry  
北京·BEIJING

## 内 容 简 介

本书可称为语音信号处理领域最新最好的指导书和参考资料。作者在所开设的麻省理工学院研究生课程的基础上,介绍了语音信号处理的主要原理、重要应用以及最新研究动态,并且指出了新的研究方向的进展和局限性。全书在理论和应用之间达到了极好的平衡。首先讲解了用于理解离散时间语音信号处理的完整理论基础,然后介绍了语音信号处理方面的重要研究进展,其中包括正弦语音处理、语音时频分析以及非线性声学语音产生模型,而这些进展情况在以往任何一本语音信号处理教科书中都不曾提及。本书在应用部分深入介绍了以下内容:语音编码、语音增强、语音综合、说话人识别、语音降噪、语音信号恢复、动态范围压缩等。值得注意的是,在本书的原理和应用部分恰当地加入了非常完整的实例和 MATLAB 习题。

本书适合作为数字语音处理及相关方向的研究生教材,也可供有关领域的研究人员参考阅读。

Authorized translation from the English language edition, entitled *Discrete-Time Speech Signal Processing: Principles and Practice*, ISBN: 013242942X by Thomas F. Quatieri, published by Pearson Education, Inc, publishing as Prentice Hall PTR, Copyright © 2002.

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage retrieval system, without permission from Pearson Education, Inc.

Simplified Chinese language edition published by Publishing House of Electronics Industry, Copyright © 2004.

This edition is authorized for sale only in the People's Republic of China excluding Hong Kong, Macau and Taiwan.

本书中文简体专有翻译出版版权由 Pearson 教育集团所属的 Prentice Hall PTR 授予电子工业出版社。其原文版权及中文翻译出版版权受法律保护。未经许可,不得以任何形式或手段复制或抄袭本书内容。

此版本仅限在中华人民共和国境内(不包括香港、澳门特别行政区以及台湾地区)发行与销售。

版权贸易合同登记号 图字: 01-2003-0599

### 图书在版编目(CIP)数据

离散时间语音信号处理——原理与应用 / (美) 夸特瑞 (Quatieri, T. F.) 著; 赵胜辉等译.

-北京: 电子工业出版社, 2004.8

(国外电子与通信教材系列)

书名原文: *Discrete-Time Speech Signal Processing: Principles and Practice*

ISBN 7-5053-9947-0

I. 离… II. ①夸… ②赵… III. 离散系统-语言信号处理-教材 IV. TN912.3

中国版本图书馆 CIP 数据核字 (2004) 第 082786 号

责任编辑: 陶淑毅

印刷: 北京智力达印刷有限公司

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

经销: 各地新华书店

开本: 787 × 1092 1/16 印张: 39.75 字数: 1018 千字

印次: 2004 年 8 月第 1 次印刷

定价: 59.00 元

凡购买电子工业出版社的图书,如有缺损问题,请向购买书店调换;若书店售缺,请与本社发行部联系。联系电话:(010) 68279077。质量投诉请发邮件至 zlt@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

## 序

2001年7月间,电子工业出版社的领导同志邀请各高校十几位通信领域方面的老师,商量引进国外教材问题。与会同志对出版社提出的计划十分赞同,大家认为,这对我国通信事业、特别是对高等院校通信学科的教学工作会很有好处。

教材建设是高校教学建设的主要内容之一。编写、出版一本好的教材,意味着开设了一门好的课程,甚至可能预示着一个崭新学科的诞生。20世纪40年代MIT林肯实验室出版的一套28本雷达丛书,对近代电子学科、特别是对雷达技术的推动作用,就是一个很好的例子。

我国领导部门对教材建设一直非常重视。20世纪80年代,在原教委教材编审委员会的领导下,汇集了高等院校几百位富有教学经验的专家,编写、出版了一大批教材;很多院校还根据学校的特点和需要,陆续编写了大量的讲义和参考书。这些教材对高校的教学工作发挥了极好的作用。近年来,随着教学改革不断深入和科学技术的飞速进步,有的教材内容已比较陈旧、落后,难以适应教学的要求,特别是在电子学和通信技术发展神速、可以讲是日新月异的今天,如何适应这种情况,更是一个必须认真考虑的问题。解决这个问题,除了依靠高校的老师 and 专家撰写新的符合要求的教科书外,引进和出版一些国外优秀电子与通信教材,尤其是有选择地引进一批英文原版教材,是会有好处的。

一年多来,电子工业出版社为此做了很多工作。他们成立了一个“国外电子与通信教材系列”项目组,选派了富有经验的业务骨干负责有关工作,收集了230余种通信教材和参考书的详细资料,调来了100余种原版教材样书,依靠由20余位专家组成的出版委员会,从中精选了40多种,内容丰富,覆盖了电路理论与应用、信号与系统、数字信号处理、微电子、通信系统、电磁场与微波等方面,既可作为通信专业本科生和研究生的教学用书,也可作为有关专业人员的参考材料。此外,这批教材,有的翻译为中文,还有部分教材直接影印出版,以供教师用英语直接授课。希望这些教材的引进和出版对高校通信教学和教材改革能起一定作用。

在这里,我还要感谢参加工作的各位教授、专家、老师与参加翻译、编辑和出版的同志们。各位专家认真负责、严谨细致、不辞辛劳、不怕琐碎和精益求精的态度,充分体现了中国教育工作者和出版工作者的良好美德。

随着我国经济建设的发展和科学技术的不断进步,对高校教学工作会不断提出新的要求和希望。我想,无论如何,要做好引进国外教材的工作,一定要联系我国的实际。教材和学术专著不同,既要注意科学性、学术性,也要重视可读性,要深入浅出,便于读者自学;引进的教材要适应高校教学改革的需要,针对目前一些教材内容较为陈旧的问题,有目的地引进一些先进的和正在发展中的交叉学科的参考书;要与国内出版的教材相配套,安排好出版英文原版教材和翻译教材的比例。我们努力使这套教材能尽量满足上述要求,希望它们能放在学生们的课桌上,发挥一定的作用。

最后,预祝“国外电子与通信教材系列”项目取得成功,为我国电子与通信教学和通信产业的发展培土施肥。也恳切希望读者能对这些书籍的不足之处、特别是翻译中存在的问题,提出意见和建议,以便再版时更正。

吴佑寿

中国工程院院士、清华大学教授  
“国外电子与通信教材系列”出版委员会主任

## 出版说明

进入 21 世纪以来,我国信息产业在生产和科研方面都大大加快了发展速度,并已成为国民经济发展的支柱产业之一。但是,与世界上其他信息产业发达的国家相比,我国在技术开发、教育培训等方面都还存在着较大的差距。特别是在加入 WTO 后的今天,我国信息产业面临着国外竞争对手的严峻挑战。

作为我国信息产业的专业科技出版社,我们始终关注着全球电子信息技术的发展方向,始终把引进国外优秀电子与通信信息技术教材和专业书籍放在我们工作的重要位置上。在 2000 年至 2001 年间,我社先后从世界著名出版公司引进出版了 40 余种教材,形成了一套“国外计算机科学教材系列”,在全国高校以及科研部门中受到了欢迎和好评,得到了计算机领域的广大教师与科研工作者的充分肯定。

引进和出版一些国外优秀电子与通信教材,尤其是有选择地引进一批英文原版教材,将有助于我国信息产业培养具有国际竞争能力的技术人才,也将有助于我国国内在电子与通信教学工作中掌握和跟踪国际发展水平。根据国内信息产业的现状、教育部《关于“十五”期间普通高等教育教材建设与改革的意见》的指示精神以及高等院校老师们反映的各种意见,我们决定引进“国外电子与通信教材系列”,并随后开展了大量准备工作。此次引进的国外电子与通信教材均来自国际著名出版商,其中影印教材约占一半。教材内容涉及的学科方向包括电路理论与应用、信号与系统、数字信号处理、微电子、通信系统、电磁场与微波等,其中既有本科专业课程教材,也有研究生课程教材,以适应不同院系、不同专业、不同层次的师生对教材的需求,广大师生可自由选择 and 自由组合使用。我们还将与国外出版商一起,陆续推出一些教材的教学支持资料,为授课教师提供帮助。

此外,“国外电子与通信教材系列”的引进和出版工作得到了教育部高等教育司的大力支持和帮助,其中的部分引进教材已通过“教育部高等学校电子信息科学与工程类专业教学指导委员会”的审核,并得到教育部高等教育司的批准,纳入了“教育部高等教育司推荐——国外优秀信息科学与技术系列教学用书”。

为作好该系列教材的翻译工作,我们聘请了清华大学、北京大学、北京邮电大学、东南大学、西安交通大学、天津大学、西安电子科技大学、电子科技大学等著名高校的教授和骨干教师参与教材的翻译和审校工作。许多教授在国内电子与通信专业领域享有较高的声望,具有丰富的教学经验,他们的渊博学识从根本上保证了教材的翻译质量和专业学术方面的严格与准确。我们在此对他们的辛勤工作与贡献表示衷心的感谢。此外,对于编辑的选择,我们达到了专业对口;对于从英文原书中发现的错误,我们通过作者联络、从网上下载勘误表等方式,逐一进行了修订;同时,我们对审校、排版、印制质量进行了严格把关。

今后,我们将进一步加强同各高校教师的密切关系,努力引进更多的国外优秀教材和教学参考书,为我国电子与通信教材达到世界先进水平而努力。由于我们对国内外电子与通信教育的发展仍存在一些认识上的不足,在选题、翻译、出版等方面的工作中还有许多需要改进的地方,恳请广大师生和读者提出批评及建议。

电子工业出版社

## 教材出版委员会

- 主任** 吴佑寿 中国工程院院士、清华大学教授
- 副主任** 林金桐 北京邮电大学校长、教授、博士生导师  
杨千里 总参通信部副部长、中国电子学会会士、副理事长  
中国通信学会常务理事
- 委员** 林孝康 清华大学教授、博士生导师、电子工程系副主任、通信与微波研究所所长  
教育部电子信息科学与工程类专业教学指导委员会委员
- 徐安士 北京大学教授、博士生导师、电子学系副主任  
教育部电子信息与电气学科教学指导委员会委员
- 樊昌信 西安电子科技大学教授、博士生导师  
中国通信学会理事、IEEE 会士
- 程时昕 东南大学教授、博士生导师  
移动通信国家重点实验室主任
- 郁道银 天津大学副校长、教授、博士生导师  
教育部电子信息科学与工程类专业教学指导委员会委员
- 阮秋琦 北方交通大学教授、博士生导师  
计算机与信息技术学院院长、信息科学研究所所长
- 张晓林 北京航空航天大学教授、博士生导师、电子工程系主任  
教育部电子信息科学与电气信息类基础课程教学指导委员会委员
- 郑宝玉 南京邮电学院副院长、教授、博士生导师  
教育部电子信息与电气学科教学指导委员会委员
- 朱世华 西安交通大学教授、博士生导师、电子与信息工程学院院长  
教育部电子信息科学与工程类专业教学指导委员会委员
- 彭启琮 电子科技大学教授、博士生导师、通信与信息工程学院院长  
教育部电子信息科学与电气信息类基础课程教学指导委员会委员
- 徐重阳** 华中科技大学教授、博士生导师、电子科学与技术系主任  
教育部电子信息科学与工程类专业教学指导委员会委员
- 毛军发 上海交通大学教授、博士生导师、电子信息学院副院长  
教育部电子信息与电气学科教学指导委员会委员
- 赵尔沅 北京邮电大学教授、教材建设委员会主任
- 钟允若 原邮电科学研究院副院长、总工程师
- 刘 彩 中国通信学会副理事长、秘书长
- 杜振民 电子工业出版社副社长

# 序



我国正处在一个研究生教育迅速发展的时期，教育质量是其能否真正起到促进社会经济发展作用的核心问题。选择和出版高质量的优秀研究生教材是解决上述核心问题的关键之一。《离散时间语音信号处理——原理与应用》一书的翻译出版正符合了这种需求。

语音是人们交流思想和进行社会活动最基本的手段，深入理解语音信号数字处理的基础理论、算法原理、研究方法和难点，对从事语音研究的科技人员是创新研究的囊中之宝，优秀的教材就是此宝的载体。美国 Prentice Hall 出版公司在国际数字信号处理权威、麻省理工学院的 Alen V. Oppenheim 教授的组织主持下，出版了 27 册关于信号处理的丛书，本书正是其中之一。作者 Thomas F. Quatieri 教授自 20 世纪 80 年代起就开始从事语音信号处理的研究，发表过许多语音、声频的增强、编码和识别方面的论文，主创了麻省理工学院的“数字语音处理”的研究生课程，在多年教学的基础上完成了此书的写作，是一部经过长期凝练的语音信号处理领域的优秀研究生教材。它的翻译出版对国内语音信号处理领域研究生的培养质量的提高将有重要的意义。

全书共分为 14 章，内容涉及离散时间信号处理基础、语音的产生及分类、语音产生的声学理论、零-极点语音模型的分析与综合、同态信号处理、短时傅里叶变换分析与综合、滤波器组分析与综合、正弦分析与综合、频域基音估计、非线性测量和建模技术、语音编码、语音增强和说话人识别等。全书详细讲述了涉及语音信号处理的基础理论、建模机理、算法推导、实用范例和 MATLAB 习题求解。内容系统深入、易于阅读、图文并茂，作为教学用书是非常出色的。更为宝贵的是，全部内容除标注文献来源之外，多数是作者自己的研究教学心得，首次在教学用书中引入。

尽管国内目前已有许多本语音信号处理方面的书籍，也都成为了研究生的教学参考用书，但相比之下，此书的独到之处是：基础理论更为完整，分析过程更为深入，解释更为深入浅出，习题也更为全面丰富。书中最有特色的几个章节，如语音产生的声学理论，全极点模型的分析合成，短时傅里叶变换、滤波器组以及正弦分析与合成等正是国内书籍所薄弱的环节，因此具有重要的意义。该书的引入对语音信号处理领域教学质量的提高是大有裨益的。

语音信号处理涉及的内容很宽，包括识别与合成，论证与检索；应用面也很广，涉及多媒体与网络等。此书涵盖的是基础理论和计算原理，旨在提高分析解决问题的实践能力，作为教材这一点是非常重要的。

我相信此书的出版将对改善研究生培养质量，提高语音信号处理的科技水平，增强语音研究的创新能力，适应 21 世纪语音通信和语音智能处理的广泛应用和发展的要求，起到重要的推动作用。

上海交通大学信息科学研究所教授  
中国电子学会常务理事、会士  
中国通信学会学术委员会副主任、学报编委、会士  
中国电子学会信号处理分会主席  
IEEE 北京计算机分会、信号处理分会主席  
国家模式识别实验室等 5 个国家级重点实验室学术委员会委员、副主任  
国务院学位委员会学科评议组成员  
国家自然科学基金委员会评审组成员

袁保宗

# 译 者 序

作为语音信号处理领域的教学和科研工作者，我们深感国内缺少内容系统深入而又体现该领域最新成果的学术专著和研究生教材。一个偶然的机，我们见到了这本 Thomas F. Quatieri 博士的语音信号处理专著。欣喜之余，有了翻译此书的想法，恰好电子工业出版社也正计划出版此书的中文版，于是贸然承担了重任。一方面是想与国内的众多同行共飨，另一方面也打算用做所开设的研究生课程“语音信号数字处理”的教材。这样，我们在教学科研工作之余，开始了此书的翻译工作。然而，翻译此书的工作量和难度是当初始料未及的。原著中长句较多，我们便在翻译时尽量按中文习惯组织译文，很多的夜晚和节假日都是伴着此书度过的。

原著是作者 Thomas F. Quatieri 博士长期科研和教学工作的结晶，我们认为它具有以下特色：（1）在内容上主要涉及语音信号处理范畴，而不包含语音识别、语音合成等与语言密切相关的应用领域；（2）正文中插入了许多实例，这些实例对于读者学习和理解有很大的帮助；（3）内容丰富的习题成为本书的一个有机组成部分，而其中的 MATLAB 习题更能使读者亲身体会语音信号处理的乐趣。此外，在作者与原著出版社（Prentice Hall PTR）的合作网站（<http://www.phptr.com/quatieri/>）上提供了全部 MATLAB 习题所需要的语音数据文件、MATLAB 脚本和工作空间。

这里，我们首先要感谢本书作者 Thomas F. Quatieri 先生，他一直关心着翻译工作的进展情况，提供了本书已发现的勘误，并对我们在翻译过程中发现的一些问题给予了及时的答复；特别感谢北京交通大学的袁保宗教授，在百忙之中欣然为本书作序；非常感谢北京理工大学的李鸿纪教授，他为本书的翻译提出了许多宝贵的意见和建议；感谢中国科学院自动化所的陶建华博士，他为本书一些术语的译法提出了很好的建议；另外，在翻译过程中，我们得到了北京理工大学各级领导，特别是匡镜明教授（北京理工大学校长）的大力支持，在此一并表示衷心的感谢。

在本书的翻译工作中，赵胜辉负责第1章、第5章、第7章、第8章、第9章、第12章以及原书序、前言、缩略语表等的翻译，刘家康负责第2章、第6章、第10章和第11章的翻译，谢湘负责第3章、第4章、第13章和第14章的翻译。赵胜辉对全书进行了统稿，最后由匡镜明教授对全书进行了审校。此外，北京理工大学现代通信实验室的一些博士研究生和硕士研究生也参加了翻译工作，付出了辛勤的劳动，他们是赵军辉、晋艳伟、朗玥、王海涛、黄石磊、肖国军、钟颖、尹安容等。

由于译者水平有限，书中难免存在错误和不妥之处，热诚欢迎读者批评指正。



# 原 书 序

作为人类最常用的交流方式，“听”与“说”是人们孜孜不倦进行探索的课题——从 Von Kempelen 的“会说话的机器”的时代算起，至今已有 150 多年的历史。随着电话的出现及其极为迅速广泛的应用，设计制造出带宽更加有效和质量更高的通信系统已成为工程师和科技人员在 70 多年时间里所追求的目标。该领域的研究和探索在极大程度上一直为现实世界的应用所驱动，而这些应用领域目前已扩展成不仅仅包括语音合成器，而且还包括自动语音识别系统、说话人确认系统、语音增强系统、高效语音编码系统和语音修正系统。工程师们的目标是设计制造那些真正可用而且经济可行的系统，这些系统可广泛用于已有的和新开发的通信信道中。

自从 20 世纪 60 年代集成电路出现以后，通信信道和语音信号处理终端系统已从模拟方式演变为全数字系统。最早从事此项重要技术变革的实验室包括贝尔电话实验室、麻省理工学院林肯实验室、IBM Thomas Watson 研究实验室、BB&N 语音研究小组和得克萨斯仪器公司，以及众多著名的大学研究小组。得克萨斯仪器公司在 20 世纪 70 年代推出的听写机 (Speak-and-Spell product) 采用了大量数字集成电路技术，导致整个技术、制造行业和销售行业的蓬勃发展，而由于快速发展的集成电路技术的应用，无数的数字系统和产品也应运而生。

由于功能更为强大的集成电路的出现，工程师们开始努力改进现有系统。这就意味着在反思并研究目前已有的语音产生和分析模型，以求对所涉及的物理过程有一个更为完整的理解，同时也意味着必须设计和使用更强有力的数学工具和算法，以便应付更深入的分析所带来的额外复杂性。某些方法已开始广泛应用，这部分归因于其早期的成功、可行性以及易于分析和实现。但这样一来，改变系统的某一部分而不影响其他部分就变得愈加困难。这种逻辑设计的过程是很复杂的，并且要在数字集成电路成本不断降低和功能不断增强之间进行折中。

在所有这些进程中，林肯实验室在语音领域开展了大量和广泛的研究。本书作者一直在非常积极地参与其中许多项目的工程开发和研究工作，为项目的成功做出了重要贡献。此外，他还在麻省理工学院常年开设语音分析和处理方面的研究生课程，该课程构成了本书的基础。

通过本书，读者可对语音产生和感知的基本科学原理，以及语音信号的表示、分析和处理所需的基本数学工具有一个深入完整的认识。而且，作者通过大量的应用实例介绍了设计方法、系统性能，并对结果进行了详细分析和点评。作者亲历了许多系统的研究工作，而且以其敏锐的目光和超常的洞察力指出了系统的工作原理及其局限所在。读者通过阅读本书就可以跟随作者了解和掌握这些系统。

请认真阅读，不断思索，始终质疑，努力实践，倾听所得的语音结果，查阅所附的大量参考文献，并通过本领域一位活跃的研究人员的丰富经验，充分领略数字技术用于语音通信这一广阔领域的神奇魅力。相信你会受益匪浅的！

James F. Kaiser

美国 Duke 大学电气与计算机工程系客座教授

# 前 言

本书内容来自于我在麻省理工学院讲授的研究生课程“数字语音处理”(我自1990年秋开始讲授该课程),也是我在麻省理工学院林肯实验室长期研究工作的结晶。这样,无论过去还是现在,原理从未远离实践,而理论与应用相伴。本书也可追溯到我童年时期对诸如信号和符号处理、声音以及技术之类事物的好奇之心。在用细绳连接两个易拉罐进行通信时,在摆弄摩尔斯式电码玩具时,在我家的地下室里怀着激动的心情操作老式业余无线电设备时,我便开始感受到那种神奇的魔力。本书的目的是向读者详细讲解离散时间语音信号处理的原理,介绍语音信号处理研究及其应用方面的最新动态,并以我本人对这个迅速发展的领域的不懈思索与兴趣来感染读者。

全书共分14章,第1章对各章内容进行了综述,第2章至第11章是本书的“理论”部分,第12章至第14章则涵盖了语音编码、语音增强和说话人识别等应用领域。而其他应用领域,如语音修正、降噪、信号恢复和动态范围压缩等,则在第2章至第11章中穿插介绍。本书未涉及包括语音识别和语言识别在内的更广范围的领域,若涉及这些就会成为一本语音技术总览,而即使如此也不能满足该领域目前的需求。本书的特色是不仅介绍那些语音建模和处理方法的成功之处,还指出了其局限性。这样可使读者质疑传统思想,并揭示所需的突破之处。本书一项重要的原则就是:与众不同的视角于突破而言尤为重要。这一点可从已故哲学家 Thomas Kuhn 那里得到印证:“发现源自对异常之关注,亦即从承认自然以其特有的方式推翻了正统科学理论所遵守的范例演绎之结果开始。”<sup>①</sup>

本书正文部分增加了大量的实例和习题。每章的习题部分都包括一些MATLAB习题,提供了语音信号及其处理方法的实践经验。这些MATLAB习题所需要的脚本、工作空间以及信号位于我在 Prentice Hall 的关联网站 (<http://www.phptr.com/quatieri/>) 上。在这个网站上还有声音演示,具体说明了各章中的原理和应用。本书中未单独作为一章介绍的应用则以实例或习题的形式出现,如采用正弦模型的说话人分离,以及应用同态处理实现的旧录音唱片恢复。我在麻省理工学院讲授语音处理课时,发现这种方法非常有效,特别是这些实例和习题都是关于理论的引人入胜的展示,并可由此一窥应用方面的最新动态。

本书的安排也使得这些内容可以不同的深度和广度进行讲授。例如,一学期的离散时间语音信号处理课程可将重点放在第2章至第9章的基础部分,若要看重讲解语音编码,则可包括第12章,而以其他应用作为实例和习题。如果是两学期的课程,则第一学期讲授第2章至第9章,将基础部分深入介绍;第二学期主要介绍第10章至第14章的理论与应用,并增加语音识别方面的辅助内容。

在这里我要衷心感谢我的同事、朋友和学生,是他们审阅了原稿的一些章节,并就章节的选题和风格进行了讨论。他们是:Walt Andrews, Carlos Avendano, Joe Campbell, Mark Clements, Jody 和 Michael Crocetta, Ron Danisewicz, Bob Dunn, Carol Epsy-Wilson, Allen Gersho, Terry

<sup>①</sup> T. Kuhn, *The Structure of Scientific Revolution*, University of Chicago Press, 1970.

Gleason, Ben Gold, Mike Goodwin, Siddhantan Govindasamy, Charles Jankowski, Mark Kahrs, Jim Kemerling, Gernot Kubin, Petros Maragos, Rich McGowen, Michael Padilla, Jim Pitton, Mike Plumpe, Larry Rabiner, Doug Reynolds, Dan Sinder, Elliot Singer, Doug Sturim, Charlie Therrien, Lisa Yanguas。此外,我还要感谢在我研究生课上的学生,他们对我的语音处理课程讲义提出了许多建设性意见;还有我的助教 Babak Azifar, Ibrahim Hajjahmad, Tim Hazen, Hanfeng Yuan 和 Xiaochun Yang,他们在习题解答和课程反馈等方面给予了很大帮助。不能忘记 Gary Kopec 和 Tom Hanna,他们既为我的同事亦是朋友,我感激他们对我的鼓励和影响,这些已凝聚于本书中。

特别感谢 Jim Kaiser,他以其特有的一丝不苟和无可挑剔的认真细致几乎审阅了全部内容,并且因其出色的工作和非凡的创造力成为我效仿的对象,激励我完成本书的写作;感谢 Bob McAulay,我们一起并肩工作多年,充满战斗力且硕果累累,而合作的研究成果为本书第9章、第10章以及第12章中正弦分析与合成及其应用中的一部分奠定了基础;同样要感谢 Hamid Nawab 和 Rob Baxter,前者与我在20世纪80年代早期富于成果的工作促成了第7章的内容,后者与我富于启发性的讨论则对第11章中的时频分布部分有极大影响。另外,我对下列麻省理工学院林肯实验室的管理人员深表谢意:Cliff Weinstein, Marc Zissman, Jerry O'Leary, Al McLaughlin 和 Peter Blankenship。是他们给予了我在这所学院执教、在林肯实验室开展研究工作的机会,并提供了一个令人兴奋和开放的研究环境。我也非常幸运地得到了 Al Oppenheim 的大力支持,他为我在麻省理工学院电气工程系与计算机科学系的执教铺平了道路,提出了写作本书的建议,并一直鼓励我在数字信号处理方面的工作。还要感谢执行编辑 Faye Gemmellaro、出版商 Bernard Goodwin 以及 Prentice Hall 出版社的其他人员,他们的关心和辛勤工作是本书质量的重要保证。

最后,我要向我的妻子 Linda 表示我最深的谢意,她的爱、支持与鼓励对于这样一项工作来说是不可或缺的,她使这一切具有意义。

Thomas F. Quatieri  
麻省理工学院林肯实验室<sup>①</sup>

---

① 这项工作得到美国国防部的项目资助,空军项目合同号是 F19628-00-C-0002。本书中的观点、说明、结论和建议为作者本人提出,无需美国空军认可。

# 目 录

第 1 章 引言	1
1.1 离散时间语音信号处理	1
1.2 语音传输通路	2
1.3 基于语音产生和感知的分析与综合	2
1.4 应用	3
1.5 本书概要	5
1.6 小结	6
参考文献	7
第 2 章 离散时间信号处理基础	8
2.1 引言	8
2.2 离散时间信号	8
2.3 离散时间系统	10
2.4 离散时间傅里叶变换	11
2.5 测不准原理	15
2.6 $z$ 变换	17
2.7 频域 LTI 系统	20
2.8 LTI 系统的特性	24
2.8.1 差分方程实现	24
2.8.2 幅度相位关系	25
2.8.3 FIR 滤波器	27
2.8.4 IIR 滤波器	28
2.9 时变系统	29
2.10 离散傅里叶变换	31
2.11 连续信号和系统转换为离散时间形式	32
2.11.1 采样定理	32
2.11.2 系统响应的采样	34
2.11.3 微分方程的数值模拟	35
2.12 小结	35
习题	35
参考文献	41
第 3 章 语音的产生与分类	42
3.1 引言	42

3.2	解剖学和生理学对发声的解释	43
3.2.1	肺	43
3.2.2	喉	44
3.2.3	声道	50
3.2.4	根据声源的语音分类	54
3.3	语音的语谱图分析 (Spectrographic Analysis)	55
3.4	语音的分类	60
3.4.1	语言的元素	60
3.4.2	元音	62
3.4.3	鼻音	64
3.4.4	摩擦音	65
3.4.5	爆破音	68
3.4.6	过渡语音	70
3.5	韵律学: 语音的音调	73
3.6	语音感知	75
3.6.1	声学音征	76
3.6.2	语音感知模型	76
3.7	小结	77
	习题	78
	参考文献	82
<b>第4章</b>	<b>语音产生的声学理论</b>	<b>85</b>
4.1	引言	85
4.2	声音物理学	85
4.2.1	基础知识	85
4.2.2	波动方程	88
4.3	均匀声管模型 (Uniform Tube Model)	91
4.3.1	无损情况	91
4.3.2	能量损耗的影响	97
4.3.3	边界效应	100
4.3.4	完整的模型	102
4.4	基于声管级联的离散时间模型	104
4.4.1	声音在级联声管模型中的传播	105
4.4.2	离散时间模型	109
4.4.3	完整的离散时间模型	114
4.5	声带/声道相互作用	118
4.5.1	一个描述声源/声道相互作用的模型	118
4.5.2	共振峰频率和带宽的调制	121
4.6	小结	124
	习题	124
	参考文献	133

第 5 章 零-极点语音模型的分析与综合	136
5.1 引言	136
5.2 时间依赖处理	136
5.3 确定性信号的全极点建模	137
5.3.1 公式化表示	137
5.3.2 误差最小化	140
5.3.3 自相关法	144
5.3.4 Levinson 递归及其相关特性	150
5.3.5 逆滤波器的格形表示	155
5.3.6 频域解释	159
5.4 随机语音信号的线性预测分析	160
5.4.1 数学表示	161
5.4.2 误差最小化	162
5.4.3 自相关法	163
5.5 “良好度”准则 ( criterion of “goodness” )	163
5.5.1 时域评价	163
5.5.2 频域评价	165
5.6 基于全极点模型的语音综合	168
5.7 零-极点估计	170
5.7.1 线性化	171
5.7.2 语音方面的应用	172
5.7.3 高基频说话人: 采用双分析窗	176
5.8 声门波导数的分解	176
5.8.1 模型	176
5.8.2 估计	178
5.9 小结	180
附录 5.A 随机过程的性质	180
附录 5.B 线性预测分析中格形滤波器的推导	184
习题	186
参考文献	195
第 6 章 同态信号处理	197
6.1 引言	197
6.2 概念	198
6.3 卷积同态系统	200
6.4 类语音序列的复倒谱	203
6.4.1 有理 $z$ 变换序列	203
6.4.2 与有理 $z$ 变换序列相卷积的脉冲序列	206
6.4.3 同态滤波	208

6.4.4	离散复倒谱	209
6.5	频谱根同态滤波	212
6.6	周期序列的短时同态分析	215
6.6.1	倒频域观点	215
6.6.2	频域观点	218
6.7	短时语音分析	219
6.7.1	浊音语音的复倒谱	219
6.7.2	清音语音的复倒谱	223
6.8	分析/综合结构	224
6.8.1	零相位和最小相位合成	224
6.8.2	混合相位合成	226
6.8.3	频谱根解卷积	227
6.9	线性预测与同态滤波的对比	228
6.9.1	特性	228
6.9.2	同态预测	228
6.10	小结	230
	习题	231
	参考文献	238
<b>第7章</b>	<b>短时傅里叶变换分析与综合</b>	<b>240</b>
7.1	简介	240
7.2	短时分析	240
7.2.1	傅里叶变换观点	240
7.2.2	滤波观点	243
7.2.3	时-频分辨率折中	247
7.3	短时综合	247
7.3.1	数学表示	248
7.3.2	滤波器组求和法 (FBS)	249
7.3.3	叠接相加 (OLA) 法	252
7.3.4	时-频抽样	254
7.4	短时傅里叶变换幅度	256
7.4.1	信号表示	257
7.4.2	由时-频样点重建信号	259
7.5	由修正的 STFT 或 STFTM 进行信号估计	260
7.5.1	STFT 综合方法的启发式应用	262
7.5.2	由修正的 STFT 进行最小方差信号估计	264
7.5.3	由修正的 STFTM 进行 LSE 信号估计	265
7.6	时域修正和语音增强	266
7.6.1	时域修正	266
7.6.2	降噪	271

7.7 小结 .....	272
附录 7.A 乘性修正的 FBS 法 .....	272
习题 .....	273
参考文献 .....	281
<b>第 8 章 滤波器组分析与综合 .....</b>	<b>283</b>
8.1 引言 .....	283
8.2 再谈 FBS 法 .....	283
8.3 相位声码器 .....	286
8.3.1 准周期性信号的分析与综合 .....	286
8.3.2 应用 .....	292
8.3.3 采用正弦分析 / 综合的原因 .....	296
8.4 相位声码器中的相位相干 .....	297
8.4.1 时间包络的保持 .....	297
8.4.2 准周期性信号的相位相干 .....	300
8.5 恒定 Q 分析 / 综合 .....	301
8.5.1 推动 .....	301
8.5.2 小波变换 .....	302
8.5.3 离散小波变换 .....	305
8.5.4 应用 .....	309
8.6 听觉模型 .....	312
8.6.1 听觉处理的 AM-FM 模型 .....	314
8.6.2 听觉谱模型 .....	316
8.6.3 听觉神经处理的阶段 / 紧张论 (Phasic / Tonic View) .....	318
8.7 小结 .....	321
习题 .....	321
参考文献 .....	329
<b>第 9 章 正弦分析与综合 .....</b>	<b>333</b>
9.1 引言 .....	333
9.2 正弦语音模型 .....	334
9.3 正弦参数的估计 .....	337
9.3.1 浊音 .....	340
9.3.2 清音 .....	342
9.3.3 分析系统 .....	343
9.3.4 帧-帧之间的峰值匹配 .....	344
9.4 正弦综合 .....	347
9.4.1 立方相位内插 .....	347
9.4.2 叠接相加内插 .....	351
9.4.3 实例 .....	352



9.4.4	应用	354
9.4.5	时-频分辨率	356
9.5	激励源 / 滤波器相位模型	358
9.5.1	信号模型	358
9.5.2	应用	359
9.6	加性确定-随机模型	368
9.6.1	信号模型	369
9.6.2	分析 / 综合	370
9.6.3	在信号修正中的应用	371
9.7	小结	372
附录 9.A	正弦模型的推导	373
附录 9.B	最优立方相位参数的推导	375
	习题	377
	参考文献	389
<b>第 10 章</b>	<b>频域基音估计</b>	<b>393</b>
10.1	引言	393
10.2	基于相关的基音估计器	393
10.3	基于“梳状滤波器”的基音估计	394
10.4	基于谐波正弦波模型的基音估计	397
10.4.1	谐波正弦波模型的参数估计	397
10.4.2	利用先验幅度的谐波正弦波模型的参数估计	399
10.4.3	清浊音检测	403
10.4.4	时-频分辨率的观点	406
10.4.5	通过谐波正弦波重建的评价	408
10.5	声门脉冲起始估计	408
10.5.1	基于起始时间的相位模型	409
10.5.2	起始时间估计	410
10.5.3	正弦波幅度包络估计	413
10.5.4	最小相位正弦波重建	414
10.6	多频带基音和清浊音估计	415
10.6.1	谐波正弦波模型	415
10.6.2	多带清浊音判决	417
10.7	小结	417
	习题	418
	参考文献	422
<b>第 11 章</b>	<b>非线性测量与建模技术</b>	<b>424</b>
11.1	引言	424
11.2	回顾 STFT 和小波变换	425