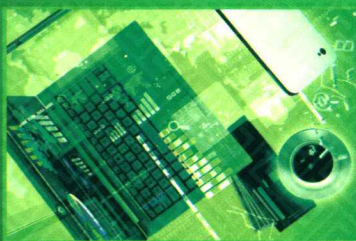
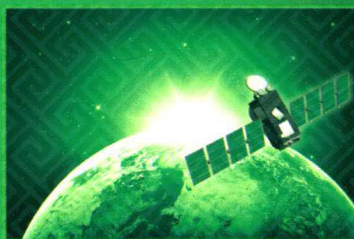


❖ 普通高等教育电子信息类规划教材 ❖

数字音频 技术及应用

THE TECHNOLOGY AND APPLICATIONS
OF DIGITAL AUDIO



谢明 编著



机械工业出版社
CHINA MACHINE PRESS



本书赠送电子教案

普通高等教育电子信息类规划教材

数字音频技术及应用

谢 明 编著



机械工业出版社

本书从基础理论到专业知识,从基本原理到实际系统和仿真设计,从简单到复杂,系统地介绍了音频信息处理的基础理论、基本方法和简要算法。本书共10章,包括绪论、音频信息处理与识别系统、音频信息采集与数字化、音频信息变换、音频信息编码、音频信息滤波、音频信息增强、音频信息的信噪分离、音频信息的分割与合成、音频信息的编辑。

本书可以供从事通信与信息工程、信号与信息处理、信息技术与电子工程、计算机网络与多媒体技术、自动化与智能化、生物医学工程等方面的教学、科研、工程、技术人员学习参考,也可作为大专院校相关专业的本科和研究生教材。

图书在版编目(CIP)数据

数字音频技术及应用/谢明编著. —北京:机械工业出版社,2017.7

普通高等教育电子信息类规划教材

ISBN 978-7-111-57411-8

I. ①数… II. ①谢… III. ①数字技术-应用-音频设备-高等学校-教材 IV. ①TN912.271

中国版本图书馆CIP数据核字(2017)第165558号

机械工业出版社(北京市百万庄大街22号 邮政编码100037)

策划编辑:尚晨 责任编辑:尚晨

责任校对:张艳霞 责任印制:常天培

唐山三艺印务有限公司印刷

2017年8月第1版·第1次印刷

184mm×260mm·13.75印张·326千字

0001-3000册

标准书号:ISBN 978-7-111-57411-8

定价:39.00元

凡购本书,如有缺页、倒页、脱页,由本社发行部调换

电话服务

网络服务

服务咨询热线:(010)88379833

机工官网:www.cmpbook.com

读者购书热线:(010)88379649

机工官博:weibo.com/cmp1952

教育服务网:www.cmpedu.com

封面无防伪标均为盗版

金书网:www.golden-book.com

前 言

声音的感知、传输、处理、识别、存储，自从动物形成以来，就产生并存在于动物界。最早最原始的声音感知，是动物的声音传感细胞、组织或器官。最早最原始的声音传输，是动物的发声细胞、组织或器官。最早最原始的声音处理，是动物的神经细胞、组织、器官或大脑。最早最原始的声音识别，也是动物的神经细胞、组织、器官或大脑。声音感知、传输、处理、识别、存储的能力，随着动物的进化、进步、升级而增强。

音频信息处理理论与技术广泛应用于人类社会的各个方面。可以说，只要有声音的时间、地点、场合、领域，都需要音频信息处理与识别。例如人，只要不是听力障碍的人，只要没睡着，耳和大脑都在接收、处理、识别音频信息。在空间科学领域，包括航天、航空，需要进行音频信息采集、传输、处理、识别、通信等，特别是传输、通信、加密等理论与技术的研究。在军事领域，包括空间、地面、海上海下军事对抗，需要音频信息技术，特别是噪声与抗噪、声呐定位、探测、跟踪、制导等理论与技术研究。在工业生产、制造、建筑领域，需要音频信息技术，特别是噪声与抗噪、超声探伤、测量、定位、切削、加工等理论与技术研究。在农业生产领域，需要音频信息技术，特别是超声灭害、育种、催生、音乐助长、虫害探测、定位、跟踪、诱导等理论与技术研究。在信息工程领域，包括信息系统、信息处理、信息安全、有线通信、无线通信、移动通信、有线网络、无线网络、移动网络等，更是离不开音频信息技术。在生物医学、生命科学领域，需要音频信息技术，特别是超声诊断、检测、定位、治疗、聋哑病理、治理、康复、人工听觉等理论与技术研究。在科学研究、高等教育领域，大部分科研院所、高等院校都有音频信息理论与技术的研究单位和研究课题和项目。这些课题和项目涉及各个领域的应用。在人类的生活领域，包括社会活动、物质生活、精神生活、文化娱乐、智慧城市、智慧社区、智慧家居、智慧服务等，特别是多媒体信息与技术、语言语音技术、音像音响技术等，也离不开音频信息的理论与技术研究。总之，音频信息技术无时不在，无处不有。

本书作者根据自己多年的教学探索、研究、实践和经验，经过总结、提炼、升华和创新，编写了本书。本书从基础理论到专业知识，从基本原理到实际系统和仿真设计，从简单到复杂，深入浅出、图文并茂、有案有例、系统地介绍了音频信息处理的基础理论、基本方法和简要算法。本书共有 10 章，包括绪论、音频信息处理与识别系统、音频信息采集与数字化、音频信息变换、音频信息编码、音频信息滤波、音频信息增强、音频信息的信噪分离、音频信息的分割与合成、音频信息的编辑。本书可以供从事通信与信息工程、信号与信息处理、信息技术与电子工程、计算机网络与多媒体技术、自动化与智能化、生物医学工程等方面的教学、科研、工程、技术人员学习参考，也可作为大专院校相关专业的本科和研究生教材。

本书在编写过程中，得到学校和学院各级部门的热情鼓励和大力支持，得到同事们的热心关怀和友好建议，也得到机械工业出版社的大力支持。在此，对他们一并表示最真诚的感谢。由于本人水平有限，书中会有一些不足之处，敬请读者批评指正。

谢 明

目 录

前言

第 1 章 绪论	1
1.1 序言	1
1.2 声学基础	2
1.3 人类听觉感知基础	6
1.3.1 人类听觉感知系统	6
1.3.2 人类听觉感知的特性	7
1.3.3 人类听觉感知效应	7
1.3.4 人类听觉感知力与评价	8
1.4 音频信息处理理论与技术	9
1.5 本章小结	11
第 2 章 音频信息处理与识别系统	12
2.1 音频信息处理与识别系统结构	12
2.2 音频信息处理与识别硬件系统	13
2.3 音频信息处理与识别软件系统	15
2.3.1 音频信息处理软件系统	16
2.3.2 音频信息识别软件系统	17
2.4 本章小结	19
第 3 章 音频信息采集与数字化	20
3.1 概述	20
3.2 声音传感器	20
3.3 前置放大器	22
3.4 A-D 模数转换器	25
3.5 音频信息采样	28
3.5.1 等间隔周期的线性采样	29
3.5.2 非等间隔非周期的非线性采样	31
3.6 音频信息量化	33
3.6.1 等间隔线性量化函数	33
3.6.2 非等间隔非线性量化函数	34
3.6.3 非等间隔自适应量化函数	35
3.7 音频信息的描述	36
3.7.1 时间域描述	36
3.7.2 频率域描述	36

3.7.3 软件域描述	36
3.8 音频信息文件格式	37
3.9 本章小结	38
第4章 音频信息变换	39
4.1 正交变换	39
4.2 傅里叶变换	40
4.2.1 一维连续傅里叶变换	40
4.2.2 二维连续傅里叶变换	40
4.2.3 一维离散傅里叶变换	41
4.2.4 二维离散傅里叶变换	41
4.2.5 矩阵与快速傅里叶变换	42
4.2.6 快速傅里叶变换	43
4.2.7 傅里叶变换的性质	44
4.3 余弦变换	45
4.3.1 一维连续余弦变换	46
4.3.2 二维连续余弦变换	46
4.3.3 一维离散余弦变换	46
4.3.4 二维离散余弦变换	47
4.3.5 矩阵与快速余弦变换	47
4.4 沃尔什变换	49
4.4.1 一维沃尔什变换	49
4.4.2 二维沃尔什变换	53
4.4.3 快速沃尔什变换	54
4.4.4 沃尔什变换的性质	54
4.5 哈尔变换	56
4.5.1 一维哈尔变换	57
4.5.2 二维哈尔变换	58
4.5.3 哈尔函数的性质	59
4.5.4 快速哈尔变换	60
4.6 Gabor 变换	61
4.6.1 一维连续 Gabor 变换	61
4.6.2 一维连续 Gabor 变换的另一种形式	63
4.6.3 一维离散 Gabor 变换	63
4.6.4 二维连续 Gabor 变换	64
4.6.5 二维离散 Gabor 变换	64
4.6.6 Gabor 变换的性质	64
4.7 小波变换	65
4.7.1 一维连续小波变换	66
4.7.2 一维离散小波变换	68

4.7.3	二维连续小波变换	70
4.7.4	二维离散小波变换	70
4.7.5	快速小波变换	70
4.7.6	小波变换的性质	71
4.8	KL 变换	72
4.8.1	KL 变换及其逆变换	73
4.8.2	KL 变换的性质	74
4.9	Hibert 变换	74
4.9.1	连续信号的 Hilbert 变换	75
4.9.2	离散信号的 Hilbert 变换	76
4.10	本章小结	77
第 5 章	音频信息编码	78
5.1	概述	78
5.2	霍夫曼编码	78
5.3	仙农 - 范诺编码	81
5.4	算数编码	84
5.5	行程编码	87
5.6	LZW 编码	87
5.7	余弦变换编码	90
5.8	小波变换编码	92
5.9	矢量量化编码	94
5.10	预测编码	96
5.11	PCM 编码	98
5.12	子带编码	99
5.13	国际编码标准	101
5.14	本章小结	103
第 6 章	音频信息滤波	104
6.1	概述	104
6.2	低通滤波	105
6.2.1	理想低通滤波	105
6.2.2	指数低通滤波	106
6.2.3	梯形低通滤波	106
6.2.4	高斯低通滤波	107
6.2.5	巴特沃尔斯低通滤波	107
6.3	高通滤波	108
6.3.1	理想高通滤波	108
6.3.2	指数高通滤波	109
6.3.3	梯形高通滤波	110
6.3.4	高斯高通滤波	110

6.3.5	巴特沃尔斯高通滤波	110
6.4	带通滤波	111
6.4.1	理想带通滤波	111
6.4.2	指数带通滤波	112
6.4.3	梯形带通滤波	113
6.4.4	高斯带通滤波	113
6.4.5	巴特沃尔斯带通滤波	114
6.5	带阻滤波	114
6.5.1	理想带阻滤波	114
6.5.2	指数带阻滤波	115
6.5.3	梯形带阻滤波	116
6.5.4	高斯带阻滤波	117
6.5.5	巴特沃尔斯带阻滤波	117
6.6	梳状滤波	118
6.6.1	理想梳状滤波	118
6.6.2	指数梳状滤波	118
6.6.3	梯形梳状滤波	119
6.6.4	高斯梳状滤波	119
6.6.5	巴特沃尔斯梳状滤波	120
6.7	频域滤波器的参数	120
6.8	复原滤波	122
6.9	时域滤波	125
6.9.1	均值滤波	125
6.9.2	中值滤波	126
6.9.3	微分滤波	127
6.9.4	积分滤波	128
6.9.5	微分积分滤波	128
6.9.6	线性组合滤波	129
6.9.7	高斯-拉普拉斯滤波	129
6.9.8	Gabor 滤波	129
6.10	卡尔曼滤波	130
6.10.1	基本卡尔曼滤波	130
6.10.2	扩展的卡尔曼滤波	131
6.11	本章小结	133
第7章	音频信息增强	134
7.1	概述	134
7.2	时间域增强	134
7.2.1	加减增强	134
7.2.2	乘除增强	135

7.2.3	线性增强	135
7.2.4	指数增强	135
7.2.5	对数增强	136
7.2.6	幂函数增强	136
7.2.7	高斯增强	136
7.2.8	巴特沃斯增强	136
7.2.9	平滑增强	137
7.2.10	锐化增强	137
7.3	频率域增强	138
7.3.1	加减增强	138
7.3.2	乘除增强	138
7.3.3	线性增强	139
7.3.4	指数增强	139
7.3.5	对数增强	139
7.3.6	幂函数增强	139
7.3.7	高斯增强	140
7.3.8	巴特沃斯增强	140
7.3.9	平滑增强	140
7.4	直方图增强	141
7.4.1	概率统计直方图	142
7.4.2	时域幅度直方图增强	142
7.5	模式增强	148
7.5.1	加减增强	149
7.5.2	乘除增强	149
7.5.3	线性增强	149
7.5.4	指数增强	150
7.5.5	对数增强	150
7.5.6	幂函数增强	150
7.5.7	高斯增强	150
7.5.8	巴特沃斯增强	151
7.5.9	平滑增强	151
7.5.10	锐化增强	152
7.5.11	微分锐化增强	152
7.5.12	微分积分锐化增强	152
7.6	特殊效果增强	152
7.6.1	延时增强	153
7.6.2	回声增强	153
7.6.3	混响增强	155
7.6.4	调制增强	156

7.7	本章小结	158
第8章	音频信息的信噪分离	160
8.1	概述	160
8.2	时间域信噪分离	160
8.2.1	微分信噪分离	160
8.2.2	积分信噪分离	161
8.3	频率域信噪分离	161
8.3.1	高通滤波谱减信噪分离	161
8.3.2	低通滤波谱减信噪分离	162
8.3.3	带阻滤波谱减信噪分离	162
8.3.4	带通滤波谱减信噪分离	162
8.4	变换域信噪分离	163
8.4.1	直方图变换信噪分离	163
8.4.2	Gabor 变换信噪分离	164
8.4.3	小波变换信噪分离	165
8.5	噪声对消	166
8.5.1	噪声模型	166
8.5.2	噪声对消	170
8.5.3	RLSE 算法	173
8.6	本章小结	174
第9章	音频信息的分割与合成	175
9.1	概述	175
9.2	端点检测的分割	175
9.2.1	功率谱单阈值法	175
9.2.2	局部最小平均功率法	177
9.2.3	功率谱双峰谷点法	178
9.2.4	功率谱多峰谷点法	180
9.3	包络检测的分割	181
9.3.1	检波法包络检测	181
9.3.2	低通滤波法包络检测	183
9.3.3	极值定理法包络检测	184
9.3.4	包络检测目标分割	186
9.4	Gabor 滤波和变换的分割	187
9.4.1	Gabor 滤波的目标信息分割	187
9.4.2	Gabor 变换的目标信息分割	188
9.5	小波变换的分割	189
9.6	幅度合成	190
9.6.1	加性合成	190
9.6.2	乘性合成	191

9.6.3	综合合成	191
9.6.4	调制合成	192
9.6.5	卷积合成	192
9.7	频率合成	192
9.7.1	加性合成	192
9.7.2	乘性合成	193
9.7.3	综合合成	193
9.7.4	卷积合成	194
9.8	变换合成	194
9.8.1	对数变换合成	194
9.8.2	Gabor 变换合成	195
9.8.3	小波变换合成	195
9.9	本章小结	195
第 10 章	音频信息的编辑	196
10.1	概述	196
10.2	线性编辑	196
10.3	非线性编辑	198
10.4	算术编辑	199
10.5	本章小结	207
参考文献		209

第1章 绪 论

1.1 序言

声音，自从宇宙形成以来，就形成并存在于宇宙。声音的感知、传输、处理、识别、存储，自从动物形成以来，就产生并存在于动物界。最早的，也是最原始的声音感知，是动物的声音传感细胞、组织或器官。最早的，也是最原始的声音传输，是动物的发声细胞、组织或器官。最早的，也是最原始的声音处理，是动物的神经细胞、组织、器官或大脑。最早的，也是最原始的声音识别，也是动物的神经细胞、组织、器官或大脑。声音感知、传输、处理、识别、存储的能力，随着动物的进化、进步、升级而增强。

随着人类世界和人类社会的快速进步和高速发展，与时俱进的现实化和现代化，最早的最原始的声音处理已经不能适应人类世界和人类社会的进步和发展。人类需要更好的理论、先进的方法、高级的手段、崭新的设施进行声音处理，去适应世界的变革，去满足社会的需求，去解决社会的问题，去维系人类的生成。因此，声音处理的研究具有重大的理论意义和社会价值。

人类从自然界接收的信息主要有五大信息：视觉信息即图像和视频信息、听觉信息即声音和音频信息、触觉信息即冷暖软硬等信息、嗅觉信息即气味信息、味觉信息即味道信息。这些信息中，视觉信息占60%，听觉信息占20%，其他信息占20%。可见，音频信息是人类和自然界进行信息互通的第二大信息。因此，音频信息感知、传输、处理、识别、存储的研究具有重大的理论意义和社会价值。

音频信息处理，最早的恐怕要算是1876年3月10日美籍英国人亚历山大·格雷厄姆·贝尔（Alexander Graham Bell）发明的贝尔电话（Bell Telephone）。贝尔电话是把声音转换成音频电信号，音频电信号通过金属线从电话发送端传输到电话接收端，在电话接收端，再把音频电信号转换成声音。贝尔电话的声电转换传感器是磁铁弹簧片话筒，电声转换器也是电磁弹簧片听筒。那时，贝尔电话还没有音频信息处理和识别功能。贝尔电话也是最早的有线通信系统（Cable Communication）。后来，贝尔电话经过不断地改进、创新，逐步发展成了当代的电信网络电话和电信网络通信（Telecommunication）。如今的电信网络电话和通信具有高级先进的声电转换和电声转换以及强大的智能化的音频信息传输、处理、识别、存储功能。贝尔有线电话也发展成了当代的无线电话（Wireless Phone）和移动电话（Mobile Phone）。如今，移动电话比电信电话功能更强大，智能化程度更高。

音频信息存储，最早的恐怕要算是1877年8月15日美国人托马斯·阿尔瓦·爱迪生（Thomas Alva Edison）发明的留声机（Gramophone）和唱片（Microgroove）。爱迪生留声机是把声音转换成波形轨道存储在介质唱片上，被称为留声，即录音。回放时再从介质唱片上读取轨道波形转换成声音，被称为放声，或放唱。声音转换成波形轨道的传感器是一片振动片、一个杠杆和一颗唱针组成的机械装置，称为录音头。录音时，声音振动振动片，振动片

通过杠杆推动唱针运动，唱针在唱片上刻划出声音强弱的波形轨道。唱片材质为赛璐珞 (Celluloid)。轨道波形转换成声音的传感器和录音头一样，被称为唱头或拾音头。放音时，唱针在轨道中运动，通过杠杆推动振动片振动发声。那时，爱迪生留声机只有声音的录放功能，还没有声音的传输、处理和识别功能。爱迪生留声机当时也被称为说话机。后来，爱迪生留声机经过不断地研究、改进、创新，发展成了当代的有线无线电声音视网络系统。如今的有线无线电声音视网络系统具有高级先进的声电转换和电声转换，强大的智能化的音频信息传输、处理、识别、存储功能。

当前，海量音频信息感知、传输、处理、识别的速度、精度、灵敏度、分辨率、质量、自动化程度、智能化程度等还不够高，海量音频信息的存储空间和存储容量还不够大，存取的速度、自动化程度、智能化程度等还不够高。因此，音频信息处理的新理论、新方法、新技术、新工艺、新设备的探索、研究、创新、开发不能停滞不前，需要持续不断的努力。

音频信息长远的主要研究方向是：

1) 海量音频信息的高速度、高精度、高灵敏度、高分辨率、高质量、高自动化、高智能化的感知。

2) 海量音频信息的高速度、高精度、高效率、高自动化、高智能化的传输、处理、识别。

3) 海量音频信息的大空间、大容量、高速度、高自动化、高智能化的存储。

音频信息的感知、传输、处理、识别、存储中，音频信息的感知和存储是音频信息的源头和终点，音频信息的识别是终极目标，音频信息处理是实现终极目标的关键，音频信息传输是连接音频信息感知、处理、识别、存储的桥梁。因此音频信息处理的研究必不可少，且非常关键。

音频信息处理以声学为基础，以人类听觉为参照，以数学物理电子学理论、方法、技术为主导，进行研究和发展的。

1.2 声学基础

宇宙万事万物，都在永恒地运动。运动是绝对的，静止是相对的。万事万物的运动是多姿多彩的、变化莫测的，有平动、转动、自转、摆动、振动、膨胀、收缩、分解、合成、复合运动等。其中，振动有简单的振动、复杂的振动。复杂的振动可以是简单振动的合成。一般简单的振动是简谐振动，一般复杂的振动可以是简谐振动的合成。一般简谐振动释放的能量是一种波，这种波是正弦波。一般合成简谐振动释放的能量也是一种波，这种波是正弦波的合成。

正弦波可以表示为：

$$s_s = A \sin(\omega t + \varphi) \quad (1.1)$$

其中，下标 s 表示简谐波； A 是正弦波振动的幅度，量纲是米 (m)； ω 是振动的角频率，量纲是弧度/秒 (rad/s)； t 是振动的的时间，量纲是秒 (s)， φ 是振动的初始相位，量纲是弧度 (rad)。

正弦波的合成可以表示为：

$$s_c = \sum_{i=-\infty}^{\infty} A_i \sin(\omega_i + \varphi_i) \quad (1.2)$$

其中,下标 c 表示合成波;下标 i 表示单个简谐波。

简谐波的角频率 ω 与频率 f 及周期 T 之间的关系为:

$$\omega = 2\pi f = 2\pi/Tf = 1/TT = 1/f \quad (1.3)$$

其中,频率 f 的量纲是赫兹 (Hz),即周/秒 (c/s);周期 T 的量纲是秒/周 (s/c)。

波有能量,简谐波的能量密度可以表示为:

$$E = A^2 \quad (1.4)$$

波可以在真空中传播。波传播的速度 v 可以表示为:

$$v = \lambda f = \lambda/T = \lambda\omega/2\pi \quad (1.5)$$

这里, λ 是波的波长,量纲是米/周 (m/c)。

浩瀚的宇宙,存在无穷无尽的波,有光波、电磁波、声波等。其中,声波是能够被人类和其他动物的听觉细胞、组织、器官、系统感知到的波。声波的频率分布从 20 Hz ~ 20 kHz,又被称为音频声。低于 20 Hz 的波叫次声波,高于 20 kHz 的波叫超声波。声波可以是单频率的波,一般为简谐波,即正弦波,又被称为纯音。声波也可以是多频率的波,一般可以认为是一些单频率正弦波的合成,又被称为复合音,如图 1-1 所示。图中,第一幅是一段声音波形,其余七幅是组成声音波形的正弦波。这段声音波形可以表示为:

$$s(t) = 10\sin(20\pi t) + 8\sin(40\pi t) + 5\sin(60\pi t) + 6\sin(80\pi t) + \sin(100\pi t) \\ + 3\sin(120\pi t) + 2\sin(140\pi t)$$

神秘的自然界,存在无数的声波。这些声波可以分类成自然声音、动物声音、人类声音、人造声音、突发声音等。自然声音如风声、雨声、雷电声、海浪声、流水声等;动物声音如鸣叫、会话、歌唱、运动响声等;人类声音如语音、歌声、呼叫声、笑声、哭声、读书声等;人造声音如乐音、钟声、机器声等;突发声音如爆炸声、垮塌声等。这些不同的声音有不同的频率范围,不同的中心频率,不同的幅度,不同的特性。

声音的频率范围一般用频带宽度 B 来表示:

$$B = f_h - f_l \quad (1.6)$$

其中, f_h 是上截止频率,频率 $f > f_h$ 的幅度都非常小,可以忽略不计。 f_l 是下截止频率,频率 $f < f_l$ 的幅度都非常小,可以忽略不计。

声音的中心频率一般用 f_0 来表示:

$$f_0 = (f_h + f_l)/2 \quad (1.7)$$

或者

$$f_0 = f | Af = \max \{ [Af_h, Af_l] \} \quad (1.8)$$

这里,符号 $x | y$ 表示条件 y 下的 x , Af 是频率为 f 的幅度 A , $[x, y]$ 是从 x 到 y 的闭区间。

声音传播的能量或强度(声强),一般用平均能流密度来描述。平均能流密度就是单位时间内通过单位面积的声音的平均能量:

$$P = \mu v A^2 \omega^2 / 2 \quad (1.9)$$

其中, μ 是传播介质的密度, v 是声波的传播速度, A 是声波的幅度, ω 是声波的频率。一般情况下,空气的密度大约为 $\mu = 1.29 \text{ kg/m}^3$,声音在空气中的传播速度大约为 $v = 340 \text{ m/s}$ 。

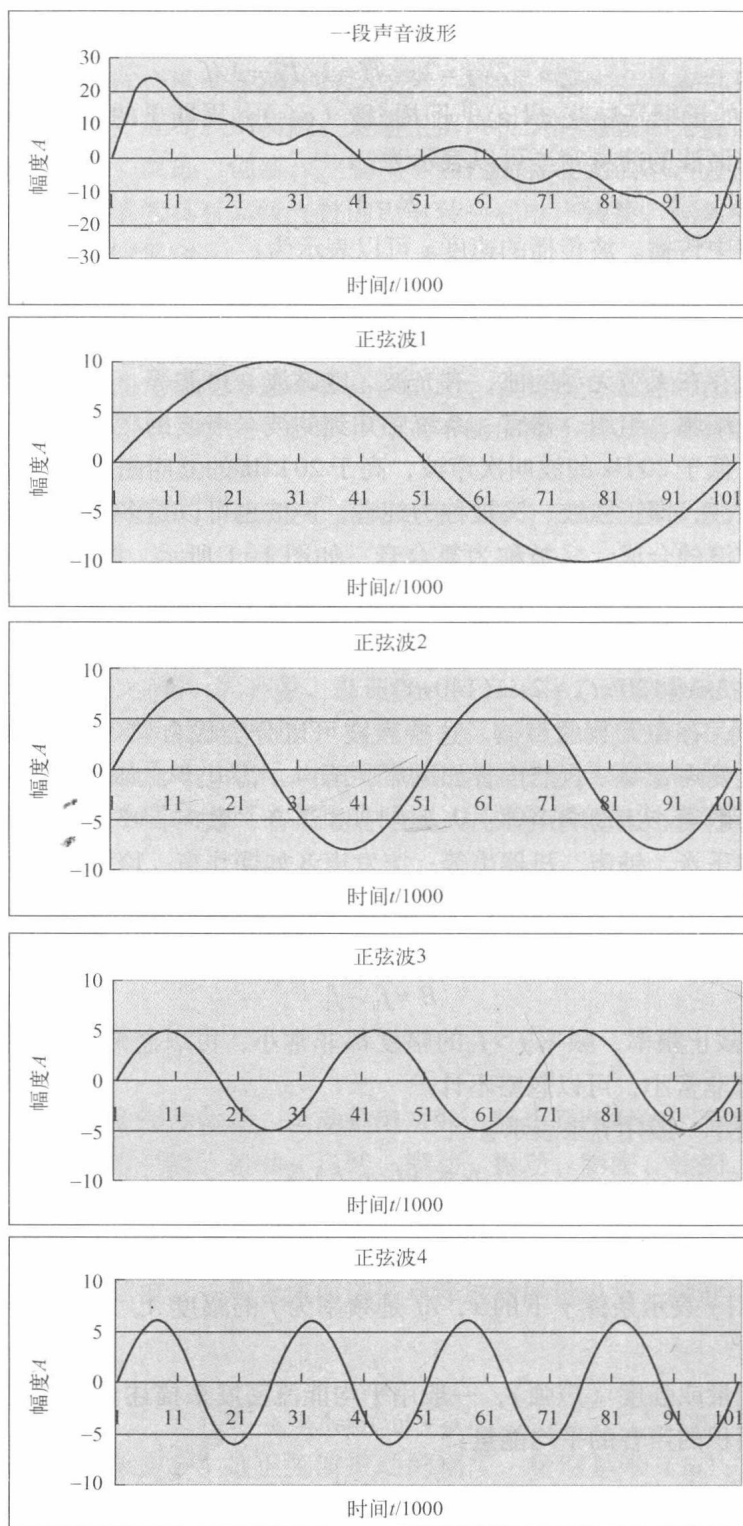


图 1-1 一段声音波形及其组成的正弦波

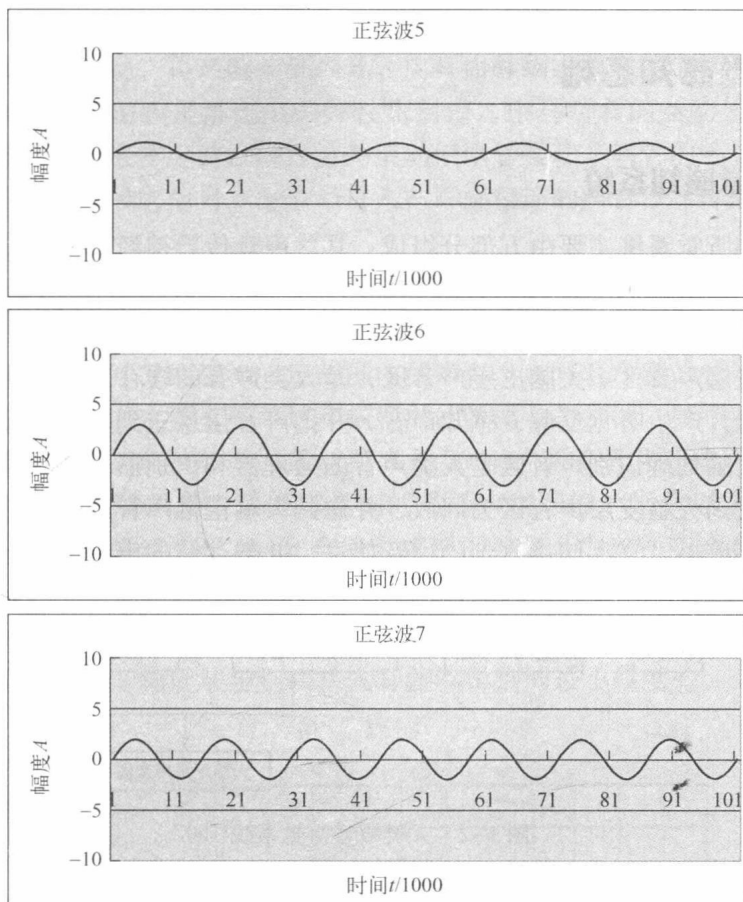


图 1-1 一段声音波形及其组成的正弦波 (续)

声音传播的压强（声压），一般用有声音传播时介质中心的压强与无声音传播时的压强之差来表示，采用声压级 L_p 来计量：

$$L_p = 10 \log(P^2/P_0^2) = 20 \log(P/P_0) \quad (1.10)$$

其中， $P_0 = 2 \times 10^{-5}$ ，单位 Pa（帕），是参考声压，接近人类的正常听觉阈值。

声音传播过程中存在反射、折射、绕射、衍射、干涉、衰减、吸收等现象。声音的反射就是声音在一种介质中传播时遇到另一种介质在这两种介质界面处反射回到原来介质中继续传播。声音的反射有全反射、漫反射和多径反射。漫反射又叫散射，就是反射后声音分多个方向传播。反射回来听到的声音叫回声。多径反射是在多种界面之间多次反射。多径反射听到的声音叫混响声音。声音的折射就是声音在一种介质中传播时遇到另一种介质在这两种介质界面处改变方向在另一种介质中继续传播。声音的绕射就是声音在传播过程中遇到较大障碍物时绕过障碍物弯曲传播路径继续向前传播。声音的衍射就是声音在传播过程中绕过圆形或球形障碍物后或者穿过较窄的通道后形成幅度加强和削弱的条纹状分布。这种条纹叫衍射条纹。声音的干涉就是几种声音在传播过程中相遇叠加形成幅度加强和削弱的条纹状分布。这种条纹叫干涉条纹。声音的衰减就是声音在传播过程中能量逐步减弱甚至消失。声音的吸收就是声音在一种介质中传播时遇到另一种介质，进入另一种介质后能量迅速衰减消失。

1.3 人类听觉感知基础

1.3.1 人类听觉感知系统

人类对声音的听觉系统主要由五部分组成：耳、声音传导神经、大脑皮层声音区、大脑处理识别声音区和大脑记忆声音区，如图 1-2 所示。耳，是人类的声音感知器，它把感知到的声音信息传给声音传导神经。声音传导神经，是人类声音的传输通道，它把接收到的声音信息传给大脑皮层声音区。大脑皮层声音区，是人类声音的缓存器和反应器，它临时存储接收到的声音信息，产生听觉反应，感知声音，并把声音信息送到大脑处理识别声音区和大脑记忆声音区。大脑处理识别声音区是人类声音的处理器和识别器，它处理识别接收到的声音信息，把结果送到大脑皮层声音区去回放，并送到大脑记忆声音区去存储。大脑记忆声音区，是人类声音存储器，它存储接收到的声音信息，并把存储的声音信息送入大脑皮层声音区去回放，送到大脑处理识别区去处理识别。

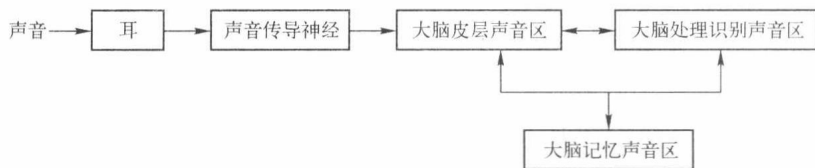


图 1-2 人类声音听觉系统

人耳的结构，主要由三部分组成：外耳、中耳和内耳。外耳由耳廓、耳孔和外耳道构成。耳廓形状非常奇特，具有集音作用，可以接收来自各个方位的声音，并把这些声音都反射或多径反射进入耳孔。外耳道，是一条较长的声音管道，长约 2.5 cm，端部被中耳的鼓膜封闭，形成谐振腔，谐振频率约 3500 Hz。它把由耳孔进来的声音聚焦共振放大，投射到中耳的鼓膜上，把空气振动声波转换成机械振动声波。耳道隔离中耳鼓膜与外耳，保护中耳鼓膜不受外部伤害，并收集排放污垢，杀菌消毒，保护耳道畅通清洁健康。中耳由鼓膜、鼓室、听骨、耳肌、韧带和咽鼓管等构成，如图 1-3 所示。耳肌连接鼓膜和听骨，韧带支持

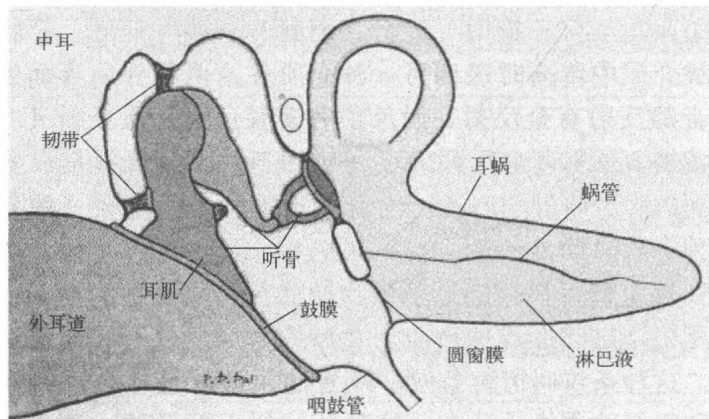


图 1-3 人耳结构原理图

(摘自 http://amuseum.cdsm.cn/AMuseum/perceptive/page_2_ear/page_2_3/page_2_3_2.htm，版权属该作者所有)