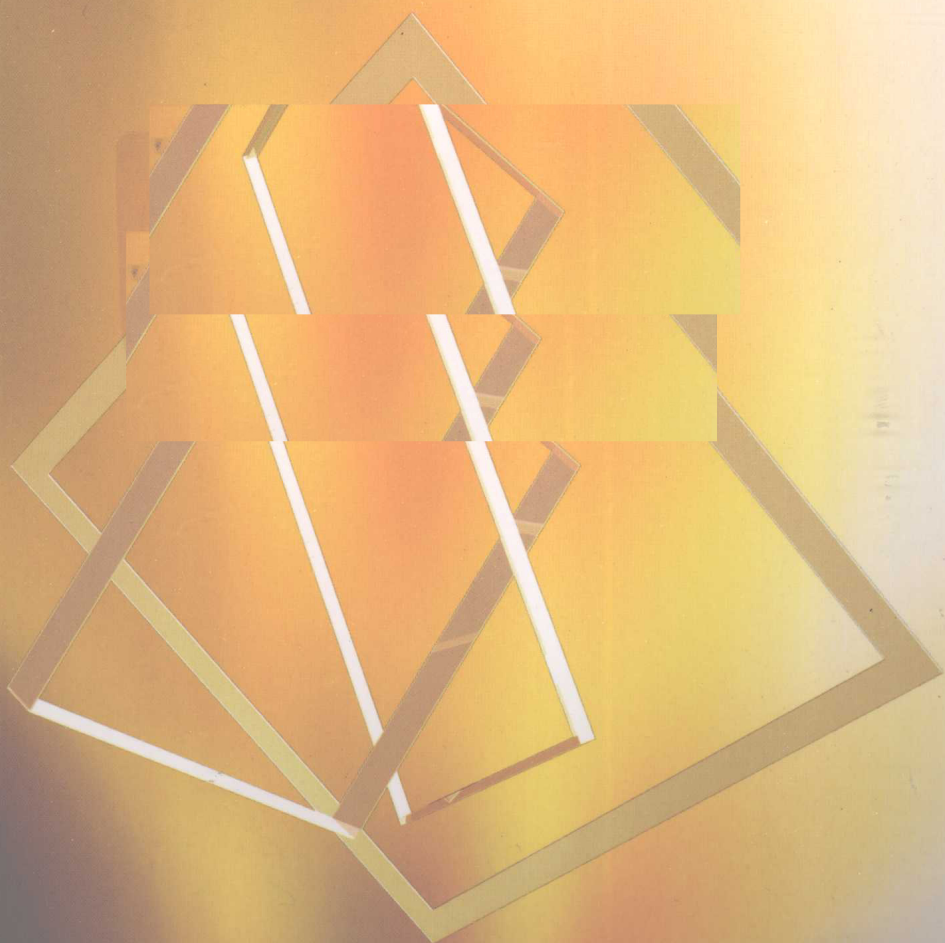


SHUZI DIANSHI
YUANLI JI YINGYONG

数字电视 原理及应用

杨 峰 白新跃 何 建◎编 著
唐继勇◎主 审



电子科技大学出版社

数字电视 原理及应用

杨 峰 白新跃 何 建◎编 著
唐继勇◎主 审

常州大学图书馆
藏书章

出 版 社：电子工业出版社
编 著：何建、白新跃、杨峰
主 审：唐继勇
主 页：www.hesicp.com
电 子 邮 箱：hesicp@hesicp.com
发 行 所：北京人民邮电出版社
印 刷 厂：北京人民邮电出版社
开 本：182mm×260mm
定 价：2010年4月第一版
印 次：2010年4月第一次印刷
书 号：ISBN 978-7-204-0476-8
定 价：28.80元

◆ 邮购本请与本社发行部联系。电话：(058) 83505353, 83526053
◆ 邮购地址：北京市丰台区方庄东里508号 邮编：100079



电子工业出版社

图书在版编目(CIP)数据

数字电视原理及应用 / 杨峰, 白新跃, 何建编著. —成都: 电子科技大学出版社, 2010.4

ISBN 978-7-5647-0476-6

I. 数… II. ①杨… ②白… ③何… III. 数字电视 IV. TN949.197

中国版本图书馆 CIP 数据核字 (2010) 第 061069 号

内 容 简 介

本书系统地介绍了数字电视的基础理论、系统组成、相关标准和关键应用。全书共 6 章, 主要内容包括数字音频信号及其压缩编码、数字视频信号与压缩编码、节目复用与 PSI/SI 信息、信道编码与传输、数字电视接收终端, 以及数字电视增值业务的重点——视频点播。每章都附有习题, 以指导读者加深对本书内容的理解。

本书层次分明, 突出了理论联系实际、原理与应用相结合的特点。可以作为高等院校通信工程、电子信息工程、广播电视工程、计算机应用等相关专业高年级本科生和研究生的数字电视技术教材, 也可以作为广播电视领域内的研究人员和工程技术人员的学习参考书。

数字电视原理及应用

出 版: 电子科技大学出版社(成都市一环路东一段 159 号电子信息产业大厦 邮编: 610051)

策划编辑: 李小锐

责任编辑: 李小锐

主 页: www.uestcp.com.cn

电子邮件: uestcp@uestcp.com.cn

发 行: 新华书店经销

印 刷: 郫县犀浦印刷厂

成品尺寸: 185mm×260mm 印张 12.75 字数 310 千字

版 次: 2010 年 4 月第一版

印 次: 2010 年 4 月第一次印刷

书 号: ISBN 978-7-5647-0476-6

定 价: 26.80 元

■ 版权所有 侵权必究 ■

◆ 邮购本书请与本社发行部联系。电话: (028) 83202323, 83256027

◆ 本书如有缺页、破损、装订错误, 请寄回印刷厂调换。

前 言

数字电视自 20 世纪七八十年代逐步发展而来,已经成为全球电视行业的发展趋势。数字电视是指从电视节目采编、制作、播出到传输、发射、接收各个环节全部采用数字编码和数字传输技术的新一代电视,具有抗干扰能力强、频率资源利用率高、图像质量好等优点,可实现更多的业务服务,如交互电视、远程教育、会议电视、电视商务、视频点播等,主要包括地面数字电视、卫星数字电视、有线数字电视、网络电视等形态。

中国数字电视产业已经进入了全面发展阶段,有线数字电视继续向高峰推进,地面数字电视正式实施,北京奥运会已经采用地面高清信号免费向全世界转播,卫星直播开始进入产业培育期。数字电视的投融资政策进一步放开,机卡分离体制逐步开始实行,数字电视产业化进程加快。另外,广播电视和电信将会互相进入,三网融合的发展进程加快。总之,数字电视将带来一场深刻的革命,这不仅仅是技术革命,而且将带来广播电视运营体制管理方式以及用户收听、收看方式的根本性变革,甚至对整个信息产业的发展产生深远影响。

针对数字电视的当前状况,为了适应中国的广播电视事业发展,作者编写了本书。全书共分 6 章。

第 1 章介绍数字音频信号及其压缩编码,主要内容包括人耳听觉机构、人耳听觉系统、声音信号数字化、数字音频压缩编码、MPEG 音频压缩编码标准和 AC-3 音频压缩编码标准。

第 2 章介绍数字视频信号与压缩编码,主要内容包括人眼视觉特性分析、摄像机成像原理、逐行扫描及隔行扫描、彩色视频信号的组成、数字视频压缩编码基础、预测编码、正交变换编码、统计编码、MPEG-2 视频、视频基本码流结构等。

第3章介绍节目复用与 PSI/SI 信息, 主要包括 ES 到 PES 封装、节目流 PS、传输流 TS、节目专用信息 PSI、业务信息 SI 等。

第4章介绍信道编码与传输, 主要包括信道编码、数字调制技术和信道传输标准。

第5章介绍数字电视接收终端, 主要包括机顶盒基本原理和结构、DVB-S 接收机、DVB-C 机顶盒、DMB-T/H 终端等。

第6章介绍数字电视增值业务的重点——视频点播, 主要包括有线 NVOD 系统、有线 TVOD 系统和 PVOD 系统。

本书由唐继勇总体规划, 第2章、第3章由杨峰编著, 第1章、第4章由白新跃编著, 第5章、第6章由何建编著, 最后由杨峰进行统稿, 唐继勇主审定稿。由于数字电视技术更新发展很快, 加之作者水平有限, 本书难免有疏漏和不足之处, 敬请读者批评指正。

编 者

2010年2月

于电子科技大学

目 录

第 1 章 数字音频信号与压缩编码	1
1.1 人耳听觉机构	1
1.2 人耳听觉系统	2
1.2.1 声音三要素	2
1.2.2 人耳的掩蔽效应	5
1.2.3 临界频带 (Critical Band)	8
1.3 声音信号数字化	9
1.3.1 连续时间信号的取样	10
1.3.2 离散时间信号的量化	11
1.3.3 编码	12
1.4 数字音频压缩编码	12
1.4.1 数字声音信源参数	12
1.4.2 数字音频压缩编码基础	13
1.4.3 声音信号的感知编码 (Perceptual Coding)	13
1.5 MPEG 音频压缩编码标准	14
1.5.1 MPEG-1 音频标准	14
1.5.2 MPEG 数字音频 ES 序列结构	18
1.6 AC-3 音频压缩编码标准	22
1.6.1 AC-3 编码原理概述	23
1.6.2 AC-3 编码系统	24
1.6.3 AC-3 数字音频 ES 流结构	24
思考与练习	29
第 2 章 数字视频信号与压缩编码	30
2.1 人眼视觉特性分析	30
2.1.1 人眼分辨明暗与色彩的原理	30
2.1.2 人眼明暗视觉的特性	32
2.1.3 人眼彩色视觉的特性	34
2.2 摄像机成像原理	35
2.2.1 成像原理	36
2.2.2 感光元件	37
2.3 逐行扫描及隔行扫描	39

2.3.1	水平扫描和垂直扫描的参数	39
2.3.2	逐行扫描	40
2.3.3	隔行扫描	41
2.4	彩色视频信号的组成	42
2.4.1	三基色原理	42
2.4.2	亮度和色差信号编码	43
2.4.3	色差信号频谱搬移和色度信号形成	45
2.4.4	彩色全电视信号和 U、V 信号	46
2.4.5	取样结构	47
2.4.6	亮度信号取样频率的选择	48
2.4.7	色度取样格式	49
2.5	数字视频压缩编码基础	51
2.5.1	压缩的必要性	51
2.5.2	压缩的可能性	51
2.6	预测编码	52
2.6.1	预测编码基本原理	52
2.6.2	预测方法	53
2.6.3	预测量化器	55
2.6.4	图像帧间编码中的运动处理	56
2.7	正交变换编码	59
2.7.1	正交变换的性质	59
2.7.2	二维 DCT 变换	60
2.7.3	量化器	62
2.8	统计编码	63
2.8.1	游程长度编码	63
2.8.2	哈夫曼编码	65
2.8.3	算术编码	67
2.9	MPEG-2 视频	69
2.9.1	MPEG-2 的型和级	70
2.9.2	视频结构	71
2.9.3	I.B.P 图像压缩编码原理	73
2.10	视频基本码流结构	76
2.10.1	序列头	78
2.10.2	序列扩展	80
2.10.3	GOP 头	82
2.10.4	PICTURE 头	83
	思考与练习	84

第 3 章 节目复用与 PSI/SI 信息	85
3.1 ES 到 PES 封装	85
3.2 节目流 PS	92
3.2.1 节目流 (PS)	93
3.2.2 PS 包头语法结构	93
3.2.3 PS 系统头语法结构	95
3.2.4 PS 流形成过程	96
3.3 传输流 (TS)	96
3.3.1 MPEG-2 系统层 TS 规范	96
3.3.2 MPEG-2 传送流语法结构	98
3.4 节目专用信息 (PSI)	105
3.4.1 节目关联表 (PAT)	106
3.4.2 PMT 表	108
3.4.3 CAT 表	111
3.4.4 系统复用	112
3.4.5 系统解复用	114
3.5 业务信息 (SI)	115
3.5.1 数字电视业务信息的构成	115
3.5.2 网络信息表 (NIT)	117
3.5.3 业务描述表 (SDT)	120
3.5.4 事件信息表 (EIT)	123
3.5.5 时间和日期表 (TDT)	125
思考与练习	126
第 4 章 信道编码与传输	127
4.1 信道编码	127
4.1.1 基本概念	127
4.1.2 能量扩散	128
4.1.3 BCH 码	129
4.1.4 LDPC 码 (低密度奇偶校验码)	131
4.1.5 RS 编码	133
4.1.6 交织	135
4.1.7 卷积编码	137
4.2 数字调制技术	140
4.2.1 四相相移键控 QPSK (Quaternary Phase Shift Keying)	140
4.2.2 正交幅度调制 QAM (Quadrature Amplitude Modulation)	141
4.2.3 TDS-OFDM	143

4.3 信道传输标准	145
4.3.1 DVB-S 的信道编码与调制	145
4.3.2 DVB-C 的信道编码与调制	147
4.3.3 DMB-T/H 的信道编码与调制	149
思考与练习	153
第 5 章 数字电视接收终端	154
5.1 机顶盒基本原理和结构	154
5.1.1 机顶盒工作原理	154
5.1.2 机顶盒结构	156
5.1.3 机顶盒软件	156
5.1.4 加解扰处理	157
5.2 DVB-S 接收机	158
5.2.1 DVB-S 接收机系统结构	158
5.2.2 DVB-S 接收机方案	159
5.3 DVB-C 机顶盒	163
5.3.1 QAMi5516 芯片	163
5.3.2 QAMi5516 硬件系统	165
5.4 DMB-T/H 终端	165
5.4.1 EM8511 平台结构	165
5.4.2 基于 EM8511 的硬件方案	167
思考与练习	168
第 6 章 视频点播	169
6.1 有线 NVOD 系统	169
6.1.1 NVOD 系统实现原理	170
6.1.2 NVOD 系统组成及工作流程	180
6.2 有线 TVOD 系统	181
6.2.1 TVOD 系统结构	181
6.2.2 TVOD 系统组成	183
6.3 PVOD 系统	192
6.3.1 PVR 机顶盒	193
6.3.2 PVOD 推播系统	193
思考与练习	195
参考文献	196

第1章 数字音频信号与压缩编码

数字音频是一种利用数字化手段对声音进行录制、存放、编辑、压缩或播放的技术，它是随着数字信号处理技术、计算机技术、多媒体技术的发展而形成的一种全新的声音处理手段。

1.1 人耳听觉机构

人耳的听觉过程是一个复杂的生理过程，它直接和人耳的机械结构和听觉神经系统特点有关。人耳能经得起自然界最强的声音，同时也能感觉到极微小的声音。例如，对于一个20岁左右健康的年轻人，他可听到的最小声压约为 $2 \times 10^{-5} \text{ Pa}$ ，这只有一个大气压的五十亿分之一。这种极微小的声压，在频率接近1000Hz时，使人耳鼓膜产生数量级仅为 $1 \times 10^{-9} \text{ cm}$ 的位移，这个距离小于氢分子直径的十分之一。

人耳听觉系统不仅是一个极灵敏的“传声器”，而且还具有“声分析器”的作用，对声音有相当大的选择能力。例如，在噪声场中，人们能够听出某些特殊频率的声音，能够选择出所需要的声音信息，就好像有一组十分可靠“声滤波器”一样。与此同时，人耳听觉系统还具有判别声音响度、音调和音色的本领。

人耳听觉机构由具有不同作用的3部分组成，即外耳、中耳和内耳，如图1-1所示。

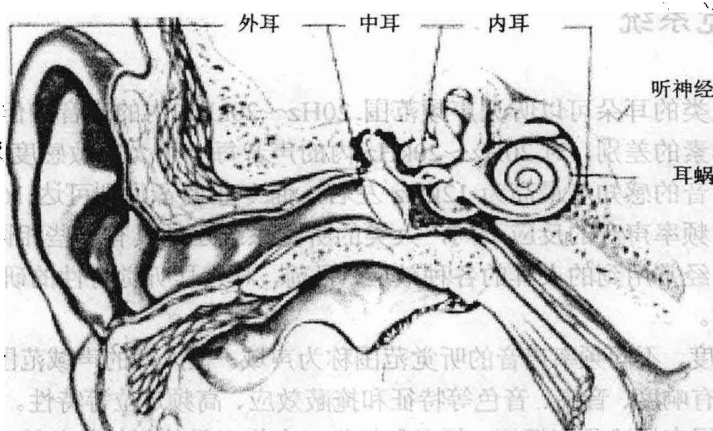


图1-1 人耳听觉机构

外耳包括耳翼和外耳道。声音通过外耳道进入耳内。外耳腔体在听觉的中频段(3000Hz)左右产生共鸣。在外耳道的末端，有一薄膜，称作鼓膜。

鼓膜及鼓膜以内称为中耳。中耳由鼓膜和3块听小骨，即锤骨、砧骨和镫骨组成，它们是相互衔接的。声波由外耳道进入后推动鼓膜振动，由于声振动的作用使连接于鼓膜的3

个听小骨也随之振动。卵形窗内是一个卵形锥体，它是声音进入内耳的正常通道。听小骨的振动就是通过镗骨与卵形窗上的弹性膜传入内耳的。整个中耳起“阻抗变换器”的作用。它使低阻抗的空气和从鼓膜开始直至耳蜗中的淋巴液高阻抗进行匹配。

内耳是人耳听觉系统和听觉器官中最复杂和最重要的部分。内耳中有一个专司听觉的部分，它是一个螺旋形的圆通道，称作耳蜗。耳蜗是一个由骨质所组成，具有蜗牛形状的中空器官，内部充满一种无色的淋巴液体，如图 1-2 所示。另外，耳蜗处还有一个也是被弹性膜遮住的称为圆形窗的小孔。圆形窗的作用相当于一个施压装置。

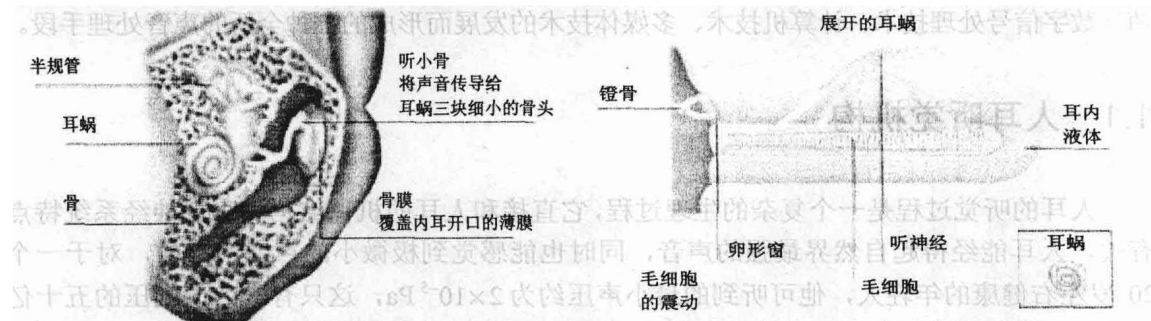


图 1-2 展开的耳蜗及毛细细胞的振动

在内耳中，接受声音振动后，起“感觉”部分的是一个螺旋线似的胶质薄膜，称为基底膜。这是一个十分重要的振动膜，它从卵形窗直到耳蜗顶端沿耳蜗的整个通道分布着。耳蜗中的淋巴液被基底膜分隔成两部分，只是在耳蜗基底膜的底端蜗孔处被分隔的两部分淋巴液才混合在一起。沿基底膜表面分布着专司听觉的毛状神经末梢约 25 000 条，其中最重要的听觉神经主干为前庭神经和蜗神经。

1.2 人耳听觉系统

研究表明，人类的耳朵可以听见声频范围 20Hz~20kHz 内的声音，但由于各个人的年龄及听音习惯等因素的差别，对 20Hz~20kHz 内的声音每一个人的敏感度是不同的。例如，一般人对高频的声音的感知度最高为 12kHz 左右，而专业调音师则可达 16kHz 左右。但是尽管各个人对不同频率声音的反应不同，人类的听觉系统还是具有一些相同的共性的东西。也就是声学学科中经常用到的人耳的各种效应，目前，对人耳听觉特性的研究主要集中在心理声学 and 语言声学。

人耳对不同强度、不同频率声音的听觉范围称为声域。在人耳的声域范围内，声音听觉心理的主观感受主要有响度、音高、音色等特征和掩蔽效应、高频定位等特性。其中响度、音高、音色可以在主观上用来描述具有振幅、频率和相位三个物理量的任何复杂的声音，故又称为声音“三要素”。而在多种音源场合，人耳掩蔽效应等特性更重要，它是心理声学的基础。

1.2.1 声音三要素

1. 响度

响度，又称声强或音量，它表示声音能量的强弱程度，主要取决于声波振幅的大小。对

微小的声音，只要响度稍有增加人耳即可感觉到，但是当声音响度增大到某一值后，即使再有较大的增加，人耳的感觉却无明显变化。我们把人耳对声音响度的这种听觉特性称为“对数式”特性。

另外人耳对不同频率的声音，听觉响度也不相同。例如，我们播放一个从 20Hz 逐步递增到 20kHz 增益相同的正弦交流信号，就会发现虽然各频段增益一样，但我们听觉所感受到的声音响度却不相同。

在 20Hz~20kHz 整个可听声频率范围内，上下限频率共 10 个倍频程，如表 1-1 所示。

表 1-1 倍频程的频率范围

频程	频率范围 (Hz)
1	20~40
2	40~80
3	80~160
4	160~320
5	320~640
6	640~1280
7	1280~2500
8	2500~5000
9	5000~10000
10	10 000~20 000

我们把可听声按倍频关系分为三份，确定低、中、高频段。即

- 低频频段 20Hz~160Hz (3 倍频)
- 中音频段 160Hz~2500Hz (4 倍频)
- 高频频段 2500Hz~20 000Hz (3 倍频)

人耳对中音频段感受到的声音响度较大，且较平坦。高频频段感受到的声音响度随频率的升高逐渐减弱，为一斜线。低频频段在 80Hz 以下急剧减弱，斜线陡率较大。我们把低频频段的急剧减弱称为低频“迟钝”现象。

声音的响度一般用声压或声强来计量，声压的单位为帕 (Pa)，它与基准声压比值的对数值称为声压级，单位是分贝 (dB)。对于响度的心理感受，一般用单位宋 (Sone) 来度量，并定义 1kHz、40dB 的纯音的响度为 1 宋。响度的相对量称为响度级，它表示的是某响度与基准响度比值的对数值，单位为方 (phon)，即当人耳感到某声音与 1kHz 单一频率的纯音同样响时，该声音声压级的分贝数即为其响度级。可见，无论在客观和主观上，这两个单位的概念是完全不同的，除 1kHz 纯音外，声压级的值一般不等于响度级的值，使用时要注意。

响度是听觉的基础。正常人听觉的强度范围为 0~140dB。固然，超出人耳的可听频率范围（即频域）的声音，即使响度再大，人耳也听不出来（即响度为零）。但在人耳的可听频域内，若声音弱到或强到一定程度，人耳同样是听不到的。当声音减弱到人耳刚刚可以听见时，此时的声音强度称为“听阈”。一般以 1kHz 纯音为准进行测量，人耳刚能听到的声压为 0dB（通常大于 0.3dB 即有感受）、声强为 10~16W/cm² 时的响度级定为 0 方。而当声

音增强到使人耳感到疼痛时,这个阈值称为“痛阈”。仍以 1kHz 纯音为准来进行测量,使人耳感到疼痛时的声压级约达到 140dB 左右。

通常认为,对于 1kHz 纯音,0~20dB 为宁静声,30~40dB 为微弱声,50~70dB 为正常声,80~100dB 为响音声,110~130dB 为极响声。而对于 1kHz 以外的可听声,则由响度曲线来描述。所谓的响度曲线,就是描述响度、声压级以及声源频率之间的关系曲线,如图 1-3 所示。图 1-3 是将听起来与 1kHz 纯音响度相同的各频率的声音的声压求得后用曲线连接起来的结果,每条线上的点均具有同一响度,称作等响曲线,又称为响度的灵敏度曲线。

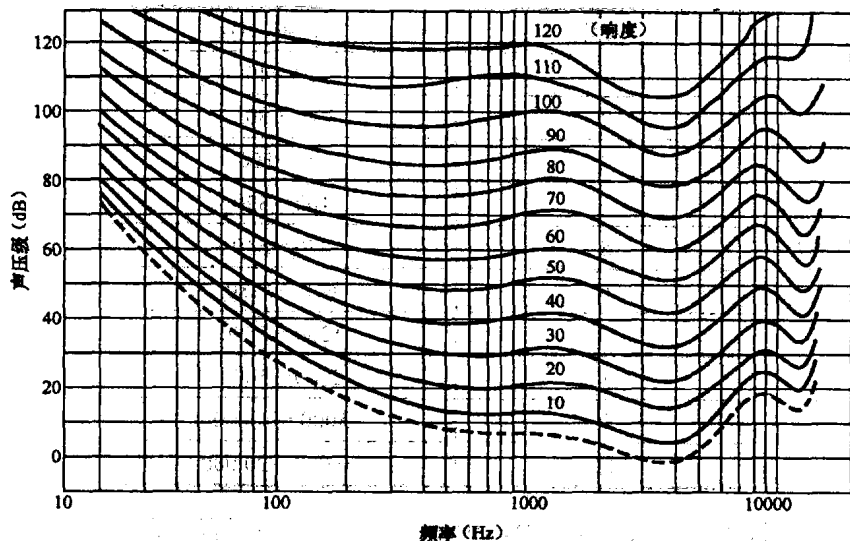


图 1-3 等响度曲线

由图 1-3 可见:

- (1) 响度与声压有关,声压级越高,相应的响度随之增大。
- (2) 人耳对 3~5kHz 声音最敏感,幅度很小的声音信号都能被人耳听到,而在低频区(如小于 800Hz)和高频区(如大于 5kHz)人耳对声音的灵敏度要低得多。
- (3) 人耳的灵敏度是随响度而变化的,声音越响,灵敏度响应越平坦。
- (4) 响度级较小时,高、低频声音灵敏度降低较明显,而低频段比高频段灵敏度降低更加剧烈,一般应特别重视加强低频音量。通常 200Hz~3kHz 语音声压级以 60~70dB 为宜,频率范围较宽的音乐声压以 80~90dB 最佳。
- (5) 图中最低的一条等响曲线描述的是最小可闻阈,它表示在整个可闻声频频段内,正常听力的人耳刚好能察觉的最小声压级。在这条最小的可闻阈线以下的区域,为不可闻区,它表示虽然存在一定的声压,但人耳却听不到。

2. 音高

音高也称音调,表示人耳对声音调子高低的主观感受。客观上音高大小主要取决于声波基频的高低,频率高则音调高,反之则低,单位用赫兹(Hz)表示。主观感觉的音高单位是“美”,通常定义响度为 40 方的 1kHz 纯音的音高为 1000 美。赫兹与“美”同样是表示音高的两个不同概念而又有联系的单位。

人耳对响度的感觉有一个从闻阈到痛阈的范围。人耳对频率的感觉同样有一个从最低可

听频率 20Hz 到最高可听频率 20kHz 的范围。响度的测量是以 1kHz 纯音为基准, 同样, 音高的测量是以 40dB 声强的纯音为基准。实验表明, 音高与频率之间的变化并非线性关系, 除了频率之外, 音高还与声音的响度及波形有关。

音高的变化与两个频率相对变化的对数成正比。不管原来频率多少, 只要两个 40dB 的纯音频率都增加一个倍频程 (即一倍), 人耳感受到的音高变化则相同。在音乐声学中, 音高的连续变化称为滑音, 一个倍频程相当于乐音提高了一个八度音阶。根据人耳对音高的实际感受, 人的语音频率范围可放宽到 80Hz~12kHz, 乐音较宽, 效果音则更宽。

3. 音色

音色又称音品, 由声音波形的谐波频谱和包络决定。声音波形的基频所产生的听得最清楚的音称为基音, 各次谐波的微小振动所产生的声音称为泛音。单一频率的音称为纯音, 具有谐波的音称为复音。每个基音都有固有的频率和不同响度的泛音, 借此可以区别其他具有相同响度和音调的声音。声音波形各次谐波的比例和随时间的衰减大小决定了各种声源的音色特征, 其包络是每个周期波峰间的连线, 包络的陡缓影响声音强度的瞬态特性。声音的音色纷呈, 变化万千, 高保真 (Hi-Fi) 音响的目标就是要尽可能准确地传输、还原重建原始声场的一切特征, 使人们真实地感受到诸如声源定位感、空间包围感、层次厚度感等各种临场听感的立体环绕声效果。

从以上主观描述声音的三个主要特征看, 人耳的听觉特性并非完全线性。声音传到人的耳内经处理后, 除了基音外, 还会产生各种谐音及它们的和音和差音, 并不是所有这些成分都能被感觉。人耳对声音具有接收、选择、分析、判断响度、音高和音品的功能, 例如, 人耳对高频声音信号只能感受到对声音定位有决定性影响的时域波形的包络 (特别是变化快的包络在内耳的延时), 而感觉不出单个周期的波形和判断不出频率非常接近的高频信号的方向, 以及对声音幅度分辨率低, 对相位失真不敏感等。这些涉及心理声学和生理声学方面的复杂问题。

1.2.2 人耳的掩蔽效应

一个较弱的声音 (被掩蔽音) 的听觉感受被另一个较强的声音 (掩蔽音) 影响的现象称为人耳的“掩蔽效应”。被掩蔽音单独存在时的听阈分贝值, 或者说在安静环境中能被人耳听到的纯音的最小值称为绝对闻阈。实验表明, 3~5kHz 绝对闻阈值最小, 即人耳对它的微弱声音最敏感; 而在低频和高频区绝对闻阈值要大得多。在 800~1500Hz 范围内闻阈随频率变化最不显著, 即在这个范围内语言可懂度最高。在掩蔽情况下, 提高被掩蔽弱音的强度, 使人耳能够听见时的闻阈称为掩蔽闻阈 (或称掩蔽门限), 被掩蔽弱音必须提高的分贝值称为掩蔽量 (或称阈移)。

已有实验表明, 纯音对纯音、噪音对纯音的掩蔽效应结论如下:

(1) 纯音间的掩蔽

- 对处于中等强度时的纯音最有效的掩蔽是出现在它的频率附近。
- 低频的纯音可以有效地掩蔽高频的纯音, 而反过来则作用很小。

(2) 噪音对纯音的掩蔽

噪音是由多种纯音组成, 具有无限宽的频谱。若掩蔽声为宽带噪声, 被掩蔽声为纯音, 则它产生的掩蔽门限在低频段一般高于噪声功

率谱密度 17dB, 且较平坦, 超过 500Hz 时大约每十倍频程增大 10dB。

若掩蔽声为窄带噪声, 被掩蔽声为纯音, 则情况较复杂。其中位于被掩蔽音附近的由纯音分量组成的窄带噪声即临界频带的掩蔽作用最明显。所谓临界频带是指当某个纯音被以它为中心频率, 且具有一定带宽的连续噪声所掩蔽时, 如果该纯音刚好能被听到时的功率等于这一频带内噪声的功率, 那么这一带宽称为临界频带宽度。临界频带的单位叫巴克 (Bark), 1Bark 等于一个临界频带宽度。频率小于 500Hz 时, 1Bark 约等于 $\text{freq}/100$; 频率大于 500Hz 时, 1Bark 约等于 $9+4\log(\text{freq}/1000)$, 即约为某个纯音中心频率的 20%。通常认为, 20Hz~16kHz 范围内有 24 个子临界频带。而当某个纯音位于掩蔽声的临界频带之外时, 掩蔽效应仍然存在。

1. 频域掩蔽

所谓频域掩蔽是指掩蔽声与被掩蔽声同时作用时发生掩蔽效应, 又称同时掩蔽。这时, 掩蔽声在掩蔽效应发生期间一直起作用, 是一种较强的掩蔽效应。通常, 频域中的一个强音会掩蔽与之同时发声的附近的弱音, 弱音离强音越近, 一般越容易被掩蔽; 反之, 离强音较远的弱音不容易被掩蔽。例如, 一个 1000Hz 的音比另一个 900Hz 的音高 18dB, 则 900Hz 的音将被 1000Hz 的音掩蔽。而若 1000Hz 的音比离它较远的另一个 1800Hz 的音高 18dB, 则这两个音将同时被人耳听到。若要让 1800Hz 的音听不到, 则 1000Hz 的音要比 1800Hz 的音高 45dB。

图 1-4 描述的是, 由于一个高强度的正弦 f_1 的存在, 使得最小可闻阈提升, 而对另一个正弦音 f_2 (不同幅度和频率) 掩蔽的情况。当第二个音在实线下面时, 人耳听不到 f_2 这个声音。

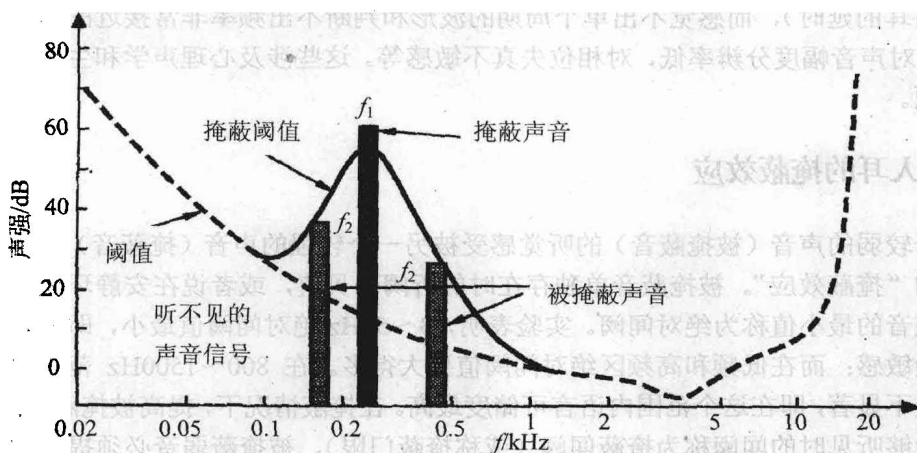


图 1-4 同时掩蔽作用示意图

在实验中发现, 掩蔽作用既和频率有关也和掩蔽信号的强度有关。也就是说, 掩蔽信号频率不同, 其掩蔽程度也是不同的; 掩蔽信号的强度不同, 其掩蔽作用也不同。

掩蔽阈是指音调音在有掩蔽声存在时刚刚听到时的阈值。如图 1-5 所示的是三条掩蔽曲线, 频率是对数刻度。由图可见, 音调音掩蔽阈的宽度随频率而变化。掩蔽曲线是非对称的, 其高频段一侧曲线的斜率要缓一些, 低频音容易对高频音产生掩蔽。

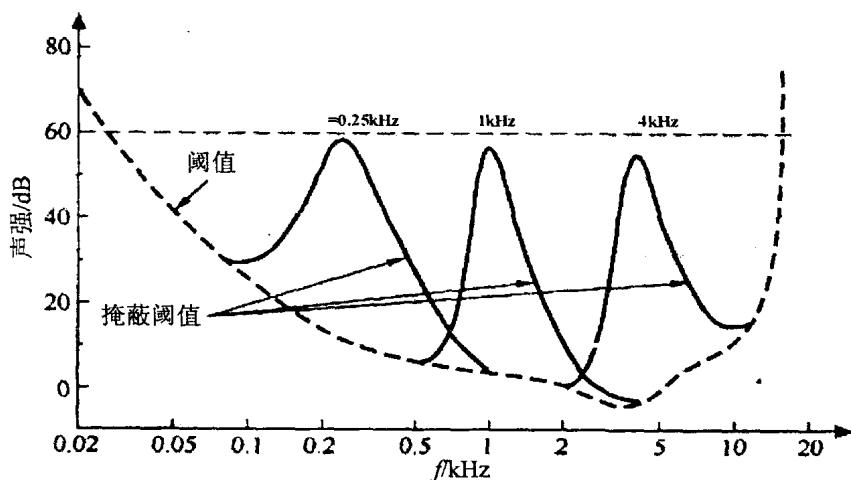


图 1-5 掩蔽阈随频率变化的曲线

实验还表明,掩蔽阈随声压级的变化而变化。如图 1-6 所示的是中心频率为 1kHz 的窄带噪声产生的一系列掩蔽曲线。低声压级掩蔽声影响的掩蔽频段相对窄一些,随着掩蔽声的声压级提高,掩蔽阈曲线就加宽了,同时其高频一侧的曲线斜率也下降了,而其低频一侧的斜率基本保持不变。

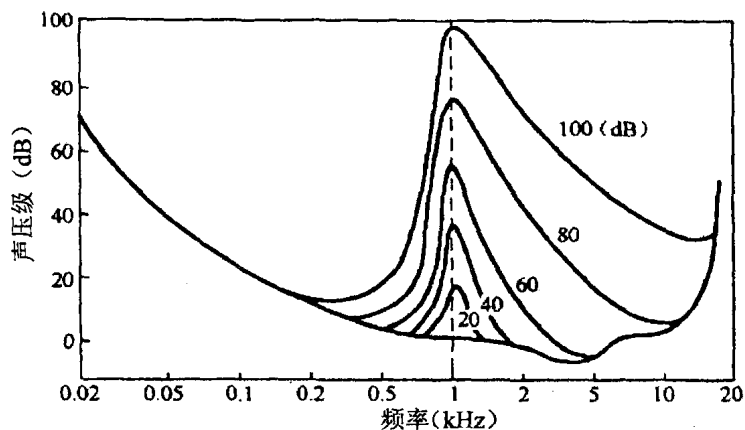


图 1-6 掩蔽阈随声压级的变化曲线

2. 时域掩蔽

所谓时域掩蔽是指掩蔽效应发生在掩蔽声与被掩蔽声不同时出现时,又称异时掩蔽。异时掩蔽又分为导前掩蔽和滞后掩蔽。若掩蔽声音出现之前的一段时间内发生掩蔽效应,则称为导前掩蔽,否则称为滞后掩蔽。

产生时域掩蔽的主要原因是人的大脑处理信息需要花费一定的时间,异时掩蔽也随着时间的推移很快会衰减,是一种弱掩蔽效应。一般情况下,导前掩蔽只有 3~20ms,而滞后掩蔽却可以持续 50~100ms。

(1) 导前掩蔽 (Pre-masking): 一个信号被在此之后发生的另一个信号所掩蔽,称为前掩蔽。也就是说,第二个声音影响了在时间上先于它的声音的听觉能力。

(2) 滞后掩蔽 (Post-masking): 在一个信号开始之前结束的另一个信号也可以掩蔽这

个信号,这称为后掩蔽。也就是说一个声音虽然已经结束了,但它对另一声音的听觉能力仍然存在影响。

总之,较强的音调音不论是发生在较弱的音调音之前还是之后,都将掩蔽掉较弱的音调音。时域掩蔽曲线如图 1-7 所示。

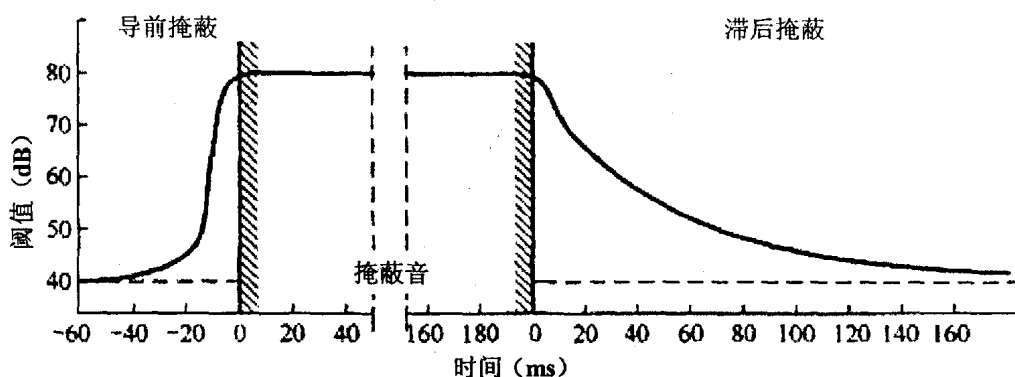


图 1-7 时域掩蔽曲线

声掩蔽现象可以用来压缩数字声频码率,一是可以将会被掩蔽的信号分量在传输之前就去除,因为这部分信号即使传输了也不会被听见;二是可以不理睬可能被掩蔽的量化噪声。

1.2.3 临界频带 (Critical Band)

所谓临界频带是指,如果掩蔽信号覆盖一定的频率范围,它的带宽逐渐增大时,掩蔽效应并不随着带宽的增大而改变,直到带宽增加到超过某个值,掩蔽效应就不再保持不变,这个带宽就是临界频带。

信号临界频带的概念表明人的耳朵好似一组多通道的实时分析器,各分析器具有不同的灵敏度和带宽。临界频带也代表掩蔽者的最小分辨率。例如,窄带噪声由于一个大的正弦波的存在而受到掩蔽,该正弦波的频率在一定范围内改变,掩蔽效应保持不变,直到两个正弦波之间的频率超过临界频带,两者的掩蔽效应才不同。

为了将临界频率带宽的概念量化,又提出了临界频带比值的概念,其定义为:当一个在噪声中的正弦波刚好被检测出来时,在扫过它的频率时会有 3dB 的增量,这个带宽就叫临界频带。

临界频率的带宽大小是频率的函数,随着频率的增加,临界频率的带宽也随之改变。研究表明,低频段的临界频带要比高频段的窄得多,故相比之下人耳能从低频段获得更多信息。在声音强度不高时,各个临界频带是相互独立的,有关临界频带的划分列于表 1-2 中。

表 1-2 人耳听觉阈范围内临界频带

临界频带编号 (Bark)	中心频率 (Hz)	临界频带 (Hz)	下端截止频率 (Hz)	上端截止频率 (Hz)
1	50	80	20	100
2	150	100	100	200
3	250	100	200	300