

高等院校通信与信息专业规划教材

# 数字音频原理及应用

卢官明 宗 昉 编著

摇 摇

摇

摇 摇 摇 摇

摇

摇



机械工业出版社

摇摇本书使用通俗易懂的语言，系统全面地介绍了数字音频技术的基础理论，数字音频设备的工作原理、性能指标及操作方法，数字声音广播的系统组成及关键技术。全书共分 7 章，主要介绍了音频信号的数字化、数字音频压缩编码的基本原理及相关标准、信道编码与调制技术、数字音频记录重放系统、电子乐器数字接口、数字音频文件格式与接口标准、音频处理与控制设备、数字音频工作站、数字声音广播、音频测量与分析等内容。每章都附有小结与习题，以指导读者加深对本书主要内容的理解。

本书注重选材，内容丰富，层次分明。在加强基本概念、基本原理的同时，注重理论与实际应用相结合，有很强的实用性。

本书可作为高等院校广播电视工程、现代教育技术、电子信息和通信类专业的本科生教材或教学参考书，也可作为数字音响工程、影视节目制作、多媒体应用与开发等领域的技术人员的岗位培训和自学用书。

## 图书在版编目 (CIP) 数据

数字音频原理及应用 / 官明, 宗昉编著. —北京: 机械工业出版社, 2005.12

高等院校通信与信息专业规划教材

ISBN 7-111-17600-0

Ⅰ. ①数… Ⅱ. ①官…②宗… Ⅲ. ①音频设备—数字技术—高等学校—教材 Ⅳ. ①TN912.2

中国版本图书馆 CIP 数据核字 (2005) 第 176000 号

机械工业出版社 (北京市百万庄大街 22 号 邮政编码 100037)

责任编辑: 陈振虹 版式设计: 霍永明 责任校对: 陈延翔

责任印制: 杨 莹

印刷厂印刷 · 新华书店北京发行所发行

2005 年 12 月第 1 版第 1 次印刷

787mm×1092mm 1/16 印张 · 200 千字

00001—00001

定价: 20.00 元

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换

本社购书热线电话 (010) 68995199 网络订购: [www.cmpbook.com](http://www.cmpbook.com)

[www.ltd.com.cn](http://www.ltd.com.cn) [www.cmpbook.com](http://www.cmpbook.com)

封面无防伪标均为盗版

高等院校通信与信息专业规划教材  
编委会名单  
(按姓氏笔画排序)

编委会主任 乐光新  
编委会副主任 张文军 张思东 杨海平 徐澄圻  
编委会委员 王金龙 冯正和 刘增基 李少洪  
邹家禄 吴镇扬 赵尔沅 南利平  
徐惠民 彭启琮 解月珍  
秘书长 胡毓坚  
副秘书长 许晔峰

# 出版说明

为了培养 21 世纪国家和社会急需的通信与信息领域的高级科技人才，配合高等院校通信与信息专业的教学改革和教材建设，机械工业出版社会同全国在通信与信息领域具有雄厚师资和技术力量的高等院校，组成阵容强大的编委会，组织长期从事通信与信息专业教学工作的骨干教师编写了这套面向普通高等院校的通信与信息专业系列教材，并将陆续出版。

这套教材力求做到：专业基础课教材概念清晰、理论准确、深度合理，并注意与专业课教学的衔接；专业课教材覆盖面广、深度适中，不仅体现相关领域的最新进展，而且注重理论联系实际。

这套教材的选题是开放式的。随着现代通信与信息技术的发展，我们将不断更新和补充选题，使这套教材及时反映通信与信息领域的新发展和新技术。我们也欢迎在教学第一线有丰富教学经验的教师及通信与信息领域的科技人员积极参与这项工作。

由于通信与信息技术发展迅速，而且涉及领域非常宽，这套教材的选题和编审中如有缺点和不足之处，诚恳希望各位老师和同学提出宝贵意见，以利于今后不断改进。

机械工业出版社  
高等院校通信与信息专业规划教材编委会

# 前摇摇言

广播电视是当今最具影响力和竞争力的现代化大众传播媒体，因其丰富的信息资源和庞大的受众群体而成为具有深远社会影响和巨大经济价值的重要信息行业。20世纪末，一场席卷全球的数字风暴引发了广播电视行业的革命。这场以数字技术、网络技术为主的变革最先冲击了广播电视技术平台，涉及到广播电视节目的采集、制作、存储、播出、传输及接收等各个环节。数字化、网络化不仅使整个广播电视节目制作和传输质量都有了显著改善，信道资源利用率大大提高，更重要的是使传统的广播电视媒体从形态、内容到服务方式发生了革命性的改变。

为了适应我国广播影视业飞速发展的需要，迫切需要一大批具备广播电视工程的基本理论、基本知识和基本技能的高级专门人才。为此，多所高校最近设立了广播电视工程专业，与此同时相应地要进行课程建设。为了满足本科教学和相关技术人员自学的需求，笔者根据近年来的教学和科研经验，结合当今最新的技术发展，编写了本书。本书可作为高等院校广播电视工程、现代教育技术、电子信息和通信类专业的本科生教材或教学参考书，也可作为数字音响工程、影视节目制作、多媒体应用与开发等领域的技术人员岗位培训和自学用书。

本书共分15章，其中第1章概述了数字音频技术的发展、所涉及的关键技术及应用领域；第2章讲述了音频信号的数字化，介绍了模数转换器的原理及性能指标；第3章介绍了数字音频压缩编码的基本原理以及国际通用标准和国家标准音频压缩编码标准；第4章讲述了信道编码与调制技术；第5章介绍了数字音频记录重放系统；第6章介绍了电子乐器数字接口；第7章介绍了数字音频文件格式与接口标准；第8章讲述了音频处理与控制设备的工作原理、性能指标及操作方法；第9章介绍了数字音频工作站的组成及主要功能；第10章介绍了数字音频广播（DAB）、数字中短波调幅广播（DRM）及网络广播的系统组成、关键技术及特点；第11章介绍了音频系统的性能指标测量与分析。从第2章起，每章都附有小结与习题，以指导读者加深对本书主要内容的理解。

在本书的编著过程中，参考和引用了前人的研究成果、著作和论文，具体出处见参考文献。在此，谨向这些文献的著作者表示敬意和感谢！

本书除第1章由宗昉编写外，其余各章由卢官明编写，宗昉也参加了资料的收集、整理工作。

由于时间仓促，作者的理论水平和教学经验有限，错误、疏漏之处在所难免，恳请广大读者批评指正，提出宝贵意见和建议。

作摇者





摇摇愿愿愿愿均衡器的种类 .....	员苑	摇摇愿愿愿愿模拟设备之间的同步 .....	员缘
摇摇愿愿愿愿摇图示均衡器的基本原理 .....	员愿	摇摇愿愿愿愿模拟设备与数字设备之间的 同步 .....	员缘
摇摇愿愿愿愿愿均衡器的技术指标 .....	员怨	摇摇愿愿愿愿愿数字音频设备之间的同步 .....	员苑
摇摇愿愿愿愿愿声反馈抑制器 .....	员园	摇摇愿愿愿愿愿同步方法综合运用实例 .....	员苑
摇摇愿愿愿愿愿声反馈的产生原因及预防措施 .....	员园	摇摇愿愿愿愿愿音频编辑软件 悦霖巷巷愿愿愿 .....	员苑
摇摇愿愿愿愿愿声反馈抑制器的工作原理 .....	员员	摇摇愿愿愿愿愿录音 .....	员愿
摇摇愿愿愿愿愿云非数位型声反馈抑制器 .....	员圆	摇摇愿愿愿愿愿编辑制作 .....	员愿
摇摇愿愿愿愿愿效果处理器 .....	员猿	摇摇愿愿愿愿愿保存与发送 .....	员员
摇摇愿愿愿愿愿室内声的组成以及对调音的 影响 .....	员猿	摇摇愿愿愿愿愿小结 .....	员员
摇摇愿愿愿愿愿摇延时器 .....	员源	摇摇愿愿愿愿愿习题 .....	员圆
摇摇愿愿愿愿愿摇混响器 .....	员愿	第 愿章摇数字声音广播 .....	员猿
摇摇愿愿愿愿愿摇听觉激励器 .....	员员	摇摇愿愿愿愿愿概述 .....	员猿
摇摇愿愿愿愿愿摇听觉激励器的作用 .....	员员	摇摇愿愿愿愿愿数字音频广播 .....	员源
摇摇愿愿愿愿愿摇听觉激励器的工作原理 .....	员员	摇摇愿愿愿愿愿阅粤粤粤粤粤系统的构成 .....	员源
摇摇愿愿愿愿愿摇听觉激励器上的功能键及调试 方法 .....	员圆	摇摇愿愿愿愿愿阅粤粤粤粤粤系统技术参数 .....	员苑
摇摇愿愿愿愿愿摇听觉激励器的应用 .....	员圆	摇摇愿愿愿愿愿阅粤粤粤粤粤月的覆盖方式 .....	员苑
摇摇愿愿愿愿愿愿调音台 .....	员猿	摇摇愿愿愿愿愿阅粤粤粤粤粤月数据广播 .....	员愿
摇摇愿愿愿愿愿愿调音台的基本功能 .....	员猿	摇摇愿愿愿愿愿阅粤粤粤粤粤月接收机原理简述 .....	员怨
摇摇愿愿愿愿愿愿调音台的分类 .....	员缘	摇摇愿愿愿愿愿阅粤粤粤粤粤月系统的特点 .....	员圆
摇摇愿愿愿愿愿愿调音台的基本构成 .....	员苑	摇摇愿愿愿愿愿愿数字调幅 (粤粤) 广播系统 .....	员猿
摇摇愿愿愿愿愿愿调音台的信号流程 .....	员怨	摇摇愿愿愿愿愿愿调幅广播的发展历史和现状 .....	员猿
摇摇愿愿愿愿愿愿调音台的技术指标 .....	员圆	摇摇愿愿愿愿愿愿粤粤粤粤粤的诞生 .....	员源
摇摇愿愿愿愿愿愿调音台与效果处理器的连接 方式 .....	员员	摇摇愿愿愿愿愿愿基于粤粤粤粤粤标准的数字粤粤 广播系统 .....	员缘
摇摇愿愿愿愿愿愿数字调音台 .....	员圆	摇摇愿愿愿愿愿愿粤粤粤粤粤系统的技术特点 .....	员愿
摇摇愿愿愿愿愿愿小结 .....	员猿	摇摇愿愿愿愿愿愿网络广播 .....	员怨
摇摇愿愿愿愿愿愿习题 .....	员猿	摇摇愿愿愿愿愿愿网络广播的特点 .....	员怨
第 愿章摇数字音频工作站 .....	员源	摇摇愿愿愿愿愿愿网络广播技术 .....	员圆
摇摇愿愿愿愿愿愿概述 .....	员源	摇摇愿愿愿愿愿愿流媒体技术原理 .....	员员
摇摇愿愿愿愿愿愿数字音频工作站的主要功能 .....	员源	摇摇愿愿愿愿愿愿网络广播软件系统 .....	员苑
摇摇愿愿愿愿愿愿数字音频工作站的组成 .....	员缘	摇摇愿愿愿愿愿愿小结 .....	员圆
摇摇愿愿愿愿愿愿愿主机 .....	员缘	摇摇愿愿愿愿愿愿习题 .....	员员
摇摇愿愿愿愿愿愿愿音频处理软件 .....	员苑	第 愿章摇音频测量与分析 .....	员圆
摇摇愿愿愿愿愿愿愿音频接口 .....	员怨	摇摇愿愿愿愿愿愿电平测量 .....	员圆
摇摇愿愿愿愿愿愿愿数字音频工作站的附件 .....	员圆	摇摇愿愿愿愿愿愿愿测量方法 .....	员圆
摇摇愿愿愿愿愿愿愿信号转换器 .....	员圆	摇摇愿愿愿愿愿愿愿电平测量单位 .....	员源
摇摇愿愿愿愿愿愿愿遥控台 .....	员圆	摇摇愿愿愿愿愿愿愿声音电平的监测仪表 .....	员苑
摇摇愿愿愿愿愿愿愿同步器 .....	员猿	摇摇愿愿愿愿愿愿愿串音和隔离度测量 .....	员愿
摇摇愿愿愿愿愿愿愿音频设备间的同步实现 .....	员缘	摇摇愿愿愿愿愿愿愿噪声的测量 .....	员怨
		摇摇愿愿愿愿愿愿愿相位和频率测量 .....	员圆

摇摇电测电引言 .....	圆圆	摇摇电测电交互调制法 .....	圆圆
摇摇电测电相位测量 .....	圆圆	摇摇电测电白噪声法 .....	圆圆
摇摇电测电频率测量 .....	圆圆	摇摇电测电眼图及抖动测量 .....	圆圆
摇摇电测电信号频谱分析 .....	圆圆	摇摇电测电眼图 .....	圆圆
摇摇电测电时域和频域的关系 .....	圆圆	摇摇电测电抖动 .....	圆圆
摇摇电测电周期性矩形脉冲的频谱 .....	圆圆	摇摇电测电小结 .....	圆圆
摇摇电测电谐波分析仪(选频电平表) .....	圆圆	摇摇电测电习题 .....	圆圆
摇摇电测电频谱分析仪 .....	圆圆	附录摇摇缩略语英汉对照 .....	圆圆
摇摇电测电非线性失真的测量 .....	圆圆	参考文献摇摇 .....	圆圆
摇摇电测电基波抑制法 .....	圆圆		

# 第 1 章 绪 论

因特网、多媒体通信、数字电视、数字音频广播、信息高速公路等这些话题，在各种媒体上每天都能听到或见到，它们已经广泛而深入地渗透到社会与生活的各个方面，对人类社会的发展产生了深刻的影响。其中扮演着重要角色的是以计算机为核心的信息处理与通信技术。这些都是伴随着数字技术的进步而实现的。

传统的模拟录音技术是通过各种话筒把各种声音、音乐转换成模拟电压信号，通过录音机等录音设备，把模拟电压信号转换为磁信号记录在磁性媒介上，比如各种录音带上。重放时，可以通过录音机等设备把磁信号重新变为模拟电压信号，通过功率放大器推动扬声器来重现声音。但是，模拟磁性录音性能受电磁性的影响较大，模拟电压信号在放大和传输过程中，会受到各种噪声和干扰的影响，磁带的频率特性在每次重放过程中都会有一些变化，这些都会对音质产生影响。对于把声波的大小以原有形态直接加以记录和传输的模拟方式来说，它的音质是依赖于所用的材料和器件。虽然可以通过改善材料和器件来提高音质，但所需费用随音质的提高成几何级数增长，而且用模拟方法来提高音质是有限度的。

运用数字技术来改善音响系统特性的尝试起始于 20 世纪 60 年代后半期。当时（1965 年）日本的调频立体声广播正值由实验阶段向正式开播过渡，为了改善音质，对于从节目制作到发射、接收的整个调频立体声系统进行了重新评估。结果表明，发射机与传输信道等环节固然会使音质恶化，但对所制作的节目最终加以收录的模拟磁带录音机却是影响音质的关键环节。同时，也得出了这样的结论：无论在模拟技术领域里如何加以改进，也不能实现音质的大幅度改善。因此，日本广播协会技术研究所决定在用磁带录音机录音的环节上尝试引入数字技术，于 1967 年完成了一台试制机。其后加以各种改进，直到 10 年之后的 1977 年，终于推出了 PCM-1 型数字磁带录音机。这种数字录音机的走带机构与盒式磁带等都直接利用了当时业已商品化的录像机，另由脉冲编码调制（Pulse Code Modulation, PCM）处理器进行 A/D 转换。由于当时的数字化技术水平还不够高，因而其最初上市的 PCM-1 型机作为数字录音机的品质比现在的品质水平要低，但与模拟磁带录音机相比已有显著的改善。

为这一成果所鼓舞，其后作为存储媒介的磁带与唱片以及使用传输媒介的声音广播都加快了数字化的速度。唱片在 1982 年推出了激光唱盘（Compact Disc, CD）数字音响系统，磁带在 1987 年使用了小型盒带的数字录音磁带（Digital Audio Tap, DAT）系统，广播也同在 1987 年推出了卫星电视直播的 B 模式立体声伴音，从而使数字音频技术走向了实用化。

所谓数字音频技术是指把模拟声音信号通过采样、量化和编码过程转换成数字信号，然后再进行记录、传输以及其他加工处理；在重放时再将这些记录的数字音频信号还原为模拟信号，获得连续的声音。

模拟信号在时间和幅度上都是连续的，幅度的微小变化都会引起声音质量的变化。而数字音频技术是通过把模拟信号进行时间上的离散化和幅度上的量化处理以后，变为一连串数字信号加以存储或传输。理论上除了把模拟信号转变为数字信号的数字化过程和把数字信号重新还原为模拟信号的过程会引入一些误差以外，在对数字信号的存储和传输过程中不会引

起音质的变化，这是越来越多采用数字音频技术的主要原因之一。

除此之外，采用数字音频技术的另一个重要原因是计算机技术的飞速发展，使数字音频处理变得越来越容易。在模拟音频技术为主的年代，声音的处理主要是依靠各种专业设备进行处理，声音的混合、延迟、改变都是通过各种设备来完成的。由于各种设备的放大、滤波、延迟等电路都有可能引入噪声和畸变，因此，为了获得尽可能高的音质，这些设备的造价是非常昂贵的，这就在某种程度上影响了这项技术的发展。数字音频技术则不同，当音频信号变为数字形式以后，所有的声音处理实际上都是一种数字的处理。基于数字信号处理的理论和各种算法，数字音频处理既可以通过软件在计算机上实现，也可以通过硬件来实现。以软件为主的实现方法的优点是成本低、灵活，一台计算机配上一块声卡和相应的软件就可以作各种处理，而且可以反复修改、多次加工；主要缺点是非实时，但随着计算机处理能力的不断提高，这个缺点正在逐步得以克服。基于数字信号处理器（DSP）的以硬件为主的实现方法能够满足实时操作的要求，也可以通过软件实现一些功能，具有较强的灵活性和较高的性能/价格比。

数字音频技术所涉及的关键技术有：

### (1) A/D（模拟/数字）和 D/A（数字/模拟）转换

一旦把模拟信号变换成数字信号之后，数字信号本身的处理就纯粹属于数字处理。然而，作为数字音频系统的入口与出口的 A/D 和 D/A 转换器却是模拟信号与数字信号的分界点。一个数字音频系统的性能究竟能在多大程度上逼近原信号所确定的理论界限，可以说就取决于 A/D、D/A 转换器及其外围电路的性能。对音频的质量来说，音频信号通过 A/D 和 D/A 变换后，越接近原始的模拟音质就越好。

采样频率和量化比特数是衡量 A/D 转换器性能的两个重要参数。为了提高数字音频的质量，需要提高采样频率和量化精度。目前，采样频率已从 44.1kHz 发展到 192kHz；量化比特数由 16bit 提高到 24bit。但如果只是简单地通过增加量化比特数和提高采样频率的措施来改善 PCM 数字音频的质量，则效果有限。为此，飞利浦和索尼公司共同推出一种称为直接流数字（Direct Stream Digital, DSD）技术，用于超级音频 CD（Super Audio CD, SACD），支持立体声和 5.1 环绕声。DSD 音频格式简化了信号流程，将模拟音频直接以 2.8224MHz 的频率进行过采样，通过  $\Sigma-\Delta$  调制和噪声整形技术以 1bit 脉冲密度方式进行编码。虽然 DSD 格式表示的声音信号是数字化数据，但是它又与真正的声波非常接近，可很好地记录模拟系统的信息。DSD 音频格式的发展将更有利于与模拟音频系统的配合。

### (2) 压缩编码技术

随着采样频率的提高，量化比特数的不断增加，数字化后的音频信号数据量非常巨大，这不利于存储和传输。压缩编码的目的是在尽量不降低音质的条件下减少数据量，以使数字系统实现小型化、多媒体化与廉价化。常用的音频压缩编码方法有子带编码（Subband Coding）、改进的离散余弦变换（Modified Discrete Cosine Transform, MDCT）编码、自适应变换听觉编码（Adaptive Transform Acoustic Coding, ATRAC）、自适应谱感知熵编码（Adaptive Spectral Perceptual Entropy Coding, ASPEC）、自适应掩蔽模型的通用子带综合编码和复用（Masking pattern adapted Universal Subband Integrated Coding And Multiplexing, MUSI-CAM）、杜比公司的 AC-3（Audio Code Number 3）、MPEG-2/4 音频编码等。

到目前为止，音频压缩编码技术的水平是：立体声编码数码率已降到 64kbit/s 甚至低到

48kbit/s，语音编码已降到 2kbit/s。随着因特网传输信息量的日益增大，要求下载速度快，故数码率的进一步降低是人们关心的热点。对音频来说，降低数码率的关键在于充分利用人耳的听觉特性，创建新型的心理声学模型以及新型的量化编码方式。

### (3) 数字信号处理技术

数字信号处理就是对数字信号进行变换、压缩、滤波、均衡、估值等各种处理。数字信号处理器 (DSP) 是数字音频设备的核心组成部分。数字调音台、数字音频工作站以及对声音进行处理的数字化设备是利用 DSP 进行幅度、频率及延时处理，实现模拟设备中的幅度压缩/限制、音量控制、频率均衡、混响、回声、效果音处理等功能，通过数字滤波器滤除信号中的各种噪声和干扰。

### (4) 信道编码和调制技术

数字信号在记录或传输过程中会产生误码，为了提高传输的可靠性或者能从光盘、磁盘等存储媒介中读出正确的数据，必须进行纠错编码，并将其码型变换成适合于信道传输的形式。

在数字音频记录重放系统中，通常采用交叉交织理德-索罗门码 (Cross Interleaved Reed-Solomon Code, CIRC)、理德-索罗门乘积码 (Reed-Solomon Product Code, RSPC) 进行纠错编码；采用 8-10 调制 (由 8bit 变换成 10bit)、EFM (Eight to Fourteen Modulation) 调制进行码型变换。而在数字声音广播系统中，通常采用截短删余卷积码及编码正交频分复用 (Coded Orthogonal Frequency Division Multiplexing, COFDM) 技术。

数字音频技术的应用领域包括：

#### (1) 消费电子类数字音响设备

CD 唱机、数字磁带录音机 (DAT)、MP3 播放机以及 MD (Mini Disc) 唱机已广泛应用了数字音频技术。

#### (2) 广播节目制作系统

在声音节目制作系统，如录音、声音处理加工、记录存储、非线性编辑等环节使用了数字调音台、数字音频工作站等数字音频设备。但在这里值得强调的是：数字化时代的音频技术，并不能丢弃模拟技术，而要将两者有机的结合，取长补短，用数字化技术去追求模拟的音质，用数字化手段来弥补模拟音频设备的弱点。

#### (3) 多媒体应用

在多媒体上的应用体现在 VCD、DVD、多媒体计算机以及因特网。VCD 采用 MPEG-1 编码格式记录声音和图像；DVD-Audio 格式支持多种不同的编码方式和记录参数，可选的编码方式包括 MLP (Meridian Lossless Packing)、Dolby AC-3、数字影院系统 (Digital Theater System, DTS)、DSD 等等，而且是可扩充的、开放的，并可以应用未来的编码技术；因特网上采用 MP3 的音频格式传输声音，以提高下载能力。

#### (4) 广播电视数字化

在数字视频广播 (DVB) 和数字音频广播 (DAB) 系统中，声音编码采用 MUSICAM 编码方法，符合 MPEG-1 Layer II 标准。在中、短波数字调幅广播 (DRM) 系统中，声音压缩编码采用 MPEG-4 高级音频编码 (Advanced Audio Coding, AAC)、码激励线性预测 (Code Excited Linear Prediction, CELP) 以及谐波矢量激励编码 (Harmonic Vector Excitation Coding, HVXC) 技术。

## 第 2 章 音频信号的数字化

虽然自然界中的音频信号是模拟信号，经过数字化处理后的音频信号必须还原为模拟信号，才能最终转换为声音。但是，由于音频信号数字化后可以避免模拟信号容易受噪声和干扰的影响，可以扩大音频的动态范围，可以利用计算机进行数据处理，可以不失真地远距离传输，可以与图像、视频等其他媒体信息进行多路复用，以实现多媒体化与网络化。所以，音频信号的数字化是一种必不可少的技术手段。

### 2.1 声音的频率范围

声音是通过空气传播的一种连续的波，叫声波。声音的传播携带了大量的信息，是人类传播信息的一种主要媒体。声音的强弱体现在声压的大小上，音调的高低体现在声音的频率上。声音用电信号表示时，声音信号在时间和幅度上都是连续的模拟信号。

频率是单位时间内信号振动的次数，一般用  $f$  表示，单位是赫兹 (Hz)。

在声学 and 电学领域里，频率一般是指正弦波信号的频率。任何信号都可以认为是各种频率的正弦波叠加，或者说任何信号都含有正弦波的各种频率成分。人们通过对各种频率成分含量的分析，可以了解该信号的许多特性。例如，人的声音信号可以分解为各种频率正弦信号的叠加，通过频谱分析可以知道，男声的高频成分要比女声的高频成分少且幅度小，男声的低频成分要比女声的低频成分多且幅度大，故男声声音较低沉浑厚，女声声音较尖细。由此可见，对信号频率的分析是非常重要的。

把频率低于 20Hz 的声波称为次声波；频率高于 20kHz 的声波称为超声波，这两类声音是人耳听不到的。人耳可以听到的声音是频率在 20Hz~20kHz 之间的声波，称之为音频 (audio) 信号。而人的发音器官发出的声音频率约在 80~3400Hz 之间，但人说话的信号频率通常在 300~3000Hz 之间，人们把这种频率范围的信号称为语 (话) 音信号 (speech, voice)。在多媒体技术中，处理的信号主要是音频信号，它包括音乐、语音、风声、雨声、鸟叫声、机器声等。对于音频信号，人们是否都能听到，主要取决于各个人的年龄和耳朵的特性。一般来讲，青年人要比老年人听到的频率范围要宽，因为随着年龄的增长，人耳对高频声的听力会逐渐降低。

自然界中发声体发出的声音从频率角度看分为两类，即纯音和复合音 (复音)。纯音是单一频率成分的音，如音叉发出的声音；复合音是由两种以上频率成分构成的音，在听觉上是多于一个音调的声音。自然界中绝大多数发声体发出的声音是复合音，如语音、音乐或噪声大多是复合音。复合音都可以分解为许多纯音之和。如果复合音的大多数纯音都集中在高频部分，就称为高频声；集中在低频部分，就称为低频声。当然，所谓高频声和低频声都是相对而言的，习惯上把频率低于 60Hz 的声音称为超低音，把频率为 60~200Hz 的声音称为低音，把频率为 200Hz~1kHz 的声音称为中音，把频率为 1~5kHz 的声音称为中高音，而把频率高于 5kHz 的声音统称为高音。在复合音分解的信号中，频率最低的一个纯音成分称

为基音；比基音频率高整数倍的纯音成分称为泛音。

## 2.2 音频信号的数字化

信号的数字化就是将连续变化的模拟信号转换成离散的数字信号，一般需要完成采样、量化和编码三个步骤，如图 2-1 所示。采样是指用每隔一定时间间隔的信号样本值序列代替原来在时间上连续的信号，也就是在时间上将模拟信号离散化。量化是用有限个幅度值近似原来连续变化的幅度值，把模拟信号的连续幅度变为有限数量、有一定间隔的离散值。编码则是按照一定的规律，把量化后的离散值用二进制数码表示。上述数字化的过程又称为脉冲编码调制（Pulse Code Modulation, PCM），通常由模拟/数字（A/D）转换器来实现。

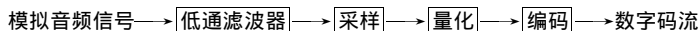


图 2-1 音频信号的数字化

数字音频信号经过处理、记录或传输后，当需要重现声音时，还必须还原为连续变化的模拟信号。将数字信号转换成模拟信号的转换称为数字/模拟（D/A）转换。

### 2.2.1 采样

模拟信号不仅在幅度取值上是连续（连续的含义是指在某一取值范围内可以取无穷多个数值）的，而且在时间上也是连续的，即每个时刻都存在一个信号幅度值与之对应。要使模拟信号数字化，首先要在时间上进行离散化处理，即在时间上用有限个采样点来代替连续无限的坐标位置。采样（sampling，又称取样或抽样）就是从时间上连续变化的模拟信号取出若干个有代表性的样本值，来代表这个连续变化的模拟信号。通过后面的分析可以知道，只要满足一定的采样条件，由这些时间上离散的样本值序列可以恢复出原来的模拟信号，即采样过程不会造成模拟信号信息的损失。

#### 1. 采样定理

采样是每隔一定的时间间隔，抽取信号的一个瞬时幅度值（样本值）。采样的时间间隔称为采样周期；每秒内采样的次数称为采样频率。采样后所得出的一系列在时间上离散的样本值称为样值序列。根据奈奎斯特（Nyquist）采样定理，一个带宽限制的模拟信号可以用一个样值序列信号来表示而不会丢失任何信息，只要采样频率  $f_s$  大于或等于被采样信号的最高频率  $f_m$  的 2 倍，就可以通过理想低通滤波器，从样值序列信号中无失真地恢复出原始模拟信号。也就是说，在满足奈奎斯特采样定理的条件下，在时间上离散的信号包含有离散前模拟信号的全部信息。

下面用傅里叶函数变换原理来证明上述定理的正确性，变换过程如图 2-2 所示。

假设原模拟信号的时间函数为  $f(t)$ ，通过傅里叶变换得到频率函数  $F(\omega)$ ，分别如图 2-2a、图 2-2b 所示。

设采样脉冲序列是周期为  $T_s$  的冲激脉冲序列，其时间函数为

$$\delta(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT_s) \quad (2-1)$$

通过傅里叶变换可以得到采样脉冲序列  $\delta(t)$  的频率函数为

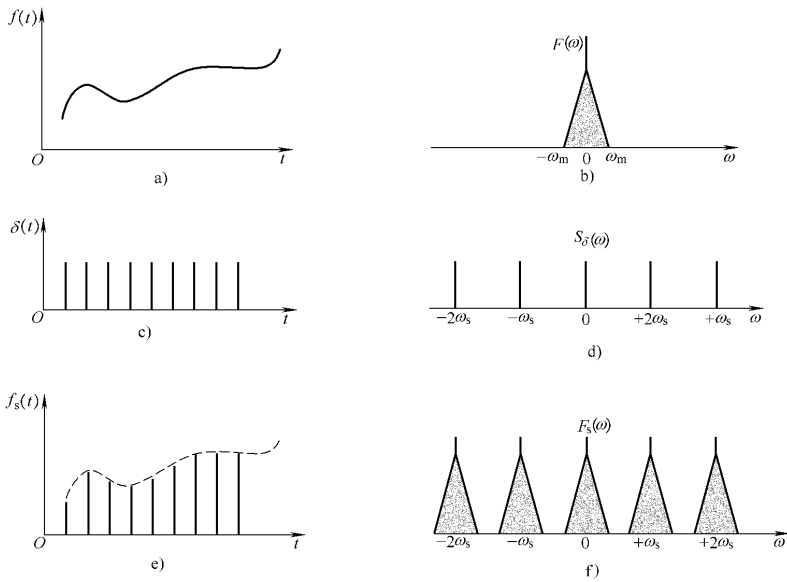


图 2-2 傅里叶变换

$$S_{\delta}(\omega) = \omega_s \sum_{n=-\infty}^{\infty} \delta(\omega - n\omega_s) \quad (2-2)$$

采样脉冲序列  $\delta(t)$  及其相应的频率函数  $S_{\delta}(\omega)$  分别如图 2-2c、图 2-2d 所示。

采样过程可以看成是原模拟信号  $f(t)$  对采样脉冲序列进行幅度调制，即二者相乘，所以，采样后得到的样值序列信号的时间函数为

$$f_s(t) = f(t) \sum_{n=-\infty}^{\infty} \delta(t - nT_s) \quad (2-3)$$

通过傅里叶变换可以得到样值序列信号  $f_s(t)$  的频率函数为

$$\begin{aligned} F_s(\omega) &= \frac{1}{2\pi} [F(\omega) * S_{\delta}(\omega)] = \frac{1}{2\pi} F(\omega) * \omega_s \sum_{n=-\infty}^{\infty} \delta(\omega - n\omega_s) \\ &= \frac{1}{T_s} \sum_{n=-\infty}^{\infty} \int_{-\infty}^{\infty} F(\tau) \delta(\omega - n\omega_s - \tau) d\tau = \frac{1}{T_s} \sum_{n=-\infty}^{\infty} F(\omega - n\omega_s) \end{aligned} \quad (2-4)$$

样值序列信号  $f_s(t)$  及其频率函数  $F_s(\omega)$  分别如图 2-2e、图 2-2f 所示。图中， $\omega_m$  为原信号的最高角频率， $\omega_s$  为采样信号的角频率， $\omega_s = 2\pi f_s = 2\pi/T_s$ ， $T_s$  为采样间隔， $f_s$  为采样频率。

上述分析结果表明，原信号被采样之后的频谱  $F_s(\omega)$  为原信号频谱  $F(\omega)$  按周期  $\omega_s$  作重复延拓。不难看出，只要满足  $f_s \geq 2f_m$ ，则在样值序列信号的频谱中完整地包含有原信号的频谱成分，即包含原模拟信号的全部信息，通过一个低通滤波器（该低通滤波器称为内插滤波器），就能从样值序列信号中恢复出原信号。现分析如下。

已知样值序列信号的频率函数为

$$F_s(\omega) = \frac{1}{T} \sum_{s_n=-\infty}^{\infty} F(\omega - n\omega_s) \quad (2-5)$$

设计一个理想的低通滤波器，其频率函数为

$$H(\omega) = \begin{cases} 1, & |\omega| \leq \frac{\omega_s}{2} \\ 0, & |\omega| > \frac{\omega_s}{2} \end{cases} \quad (2-6)$$

这是一个频域上的门函数，相应的时域函数为

$$h(t) = \frac{1}{2\pi} \int_{-\frac{\omega_s}{2}}^{\frac{\omega_s}{2}} H(\omega) e^{j\omega t} d\omega = \frac{1}{2\pi} \int_{-\frac{\omega_s}{2}}^{\frac{\omega_s}{2}} e^{j\omega t} d\omega = \frac{\omega_s}{2\pi} \frac{\sin \frac{\omega_s}{2} t}{\frac{\omega_s}{2} t} \quad (2-7)$$

样值序列信号通过理想的低通滤波器后，其输出频率函数为

$$F(\omega) = F_s(\omega)H(\omega)$$

通过傅里叶反变换，可以得到相应的时间函数为

$$f(t) = f_s(t) * h(t) = \sum_{n=-\infty}^{\infty} f(nT_s) \frac{\sin \frac{\omega_s}{2}(t - nT_s)}{\frac{\omega_s}{2}(t - nT_s)} \quad (2-8)$$

上式说明，利用理想的低通滤波器可以由时间上离散的各样值序列不失真地恢复出时间上连续的原模拟信号。

## 2. 混叠失真与限带滤波

前面已证明，当采样频率满足  $f_s \geq 2f_m$  时，可以不失真地恢复出原模拟信号。如果不满足上述采样定理的条件，即采样频率  $f_s$  小于信号最高频率  $f_m$  的 2 倍，或信号的实际最高频率超过了  $f_s/2$ ，则采样后的信号频谱会发生频谱混叠现象，如图 2-3 所示。这时，即使用理想的矩形低通滤波器也无法不失真地恢复出原模拟信号。因此而产生的失真称为频谱混叠失真。

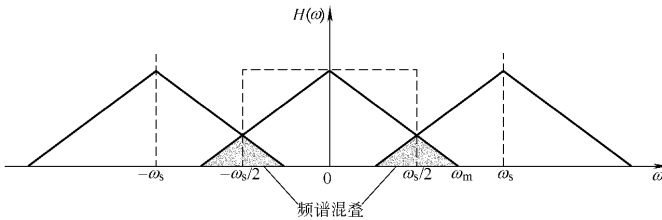


图 2-3 频谱混叠失真

为了防止产生频谱混叠失真，当采样频率确定后，就必须限制原模拟信号的上限频率。音频信号的频率通常在 20kHz 以下，因此，在 A/D 转换器之前可设置一个上限频率合适的低通滤波器（通常称为前置滤波器，也称防混叠滤波器），先对原模拟信号进行限带滤波，以滤除频率高于  $f_s/2$  的频谱分量。相应地，在 D/A 转换器之后要设置内插低通滤波器（也称防镜像滤波器），以滤除多余的高频分量，只把原信号频谱取出来。这种多余的高频分量

本来是人耳听不出的，但是如果后级的放大器或扬声器有非线性之处，就会由于产生了交扰调制而落在人耳的可听域之内，造成音质失真。由于实际的前置低通滤波器及内插低通滤波器的频率特性不可能是理想的门函数，即其幅频特性不是矩形，有一定的截止边沿宽度。为了较好地防止频谱混叠失真，实际中应使采样频率稍大于信号最高频率  $f_m$  的 2 倍，通常选为

$$f_s = (2.1 \sim 2.5) f_m \quad (2-9)$$

### 3. 采样保持电路

在对采样点的样本值进行量化的操作是需要一段时间的。虽然这段时间很短，但在这段时间里，输入的模拟信号可能会发生变化，这意味着在量化所需的这段操作时间内必须使该采样值保持不变。采样保持（简称为 S/H）电路就是用来完成采样和保持功能的电路，其电路原理如图 2-4 所示。它是用控制信号使模拟开关的接通与断开来实现采样保持功能的。

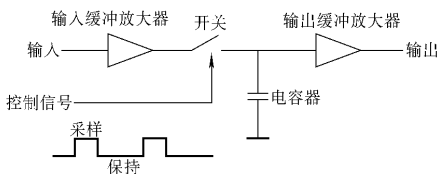


图 2-4 采样保持电路原理图

开关电路可以用场效应晶体管（FET）、二极管电桥等构成。当开关闭合时（采样阶段），输入电压对电容器充电，同时直接作为输出电压；当开关断开后（保持阶段），用电容器上所充的电荷使输出电压保持开关断开前的电压（采样值）。采样保持电路的输入信号和输出信号的示意图分别如图 2-5a 和图 2-5b 所示。

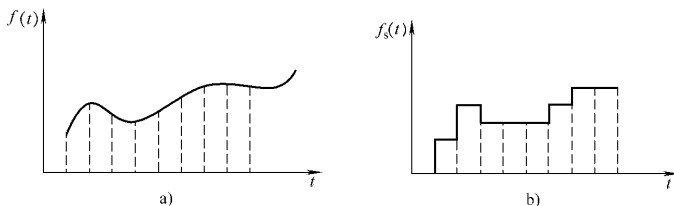


图 2-5 采样保持示意图

a) 输入信号 b) 输出信号

采样保持电路的主要性能参数有以下三项：

#### (1) 失调误差

失调误差也称偏移误差，主要发生在输入/输出缓冲放大器处。不过，一个音响系统是没有必要记录、重放纯直流信号的，因而也就没有必要把失调误差看得有多么严重。

#### (2) 捕捉时间

捕捉时间是指从采样保持电路接到采样指令起、到输出电压已能对输入电压信号加以跟随为止的一段时间。理想的采样保持电路的捕捉时间应该接近于零，否则，A/D 转换器输出的变换值就不是采样时刻信号的样值，而是捕捉过程中输入信号的平均值。捕捉时间取决于缓冲放大器的响应速度、瞬态响应特性、模拟开关的切换时间、电容的充电电流与容量等。

#### (3) 平顶降落

平顶降落（drop）是指采样保持电路在对采样的电压值加以保持的过程中所产生的电压降落。在采样保持电路接到保持指令、模拟开关已断开的状态下，本来希望输出的电压等于