

中国通信学会普及与教育工作委员会推荐教材

21世纪高职高专电子信息类规划教材

21 Shiji Gaozhi Gaozhuān Dianzi Xinxilei Guihua Jiaocai

数字音视频 技术

解相吾 解文博 编著

- 技术先进，反映数字音视频技术最新动态
- 结构系统，体现宽基础、薄理论、强技能
- 突出实用，符合高职高专“双证制”需要



 人民邮电出版社
POSTS & TELECOM PRESS

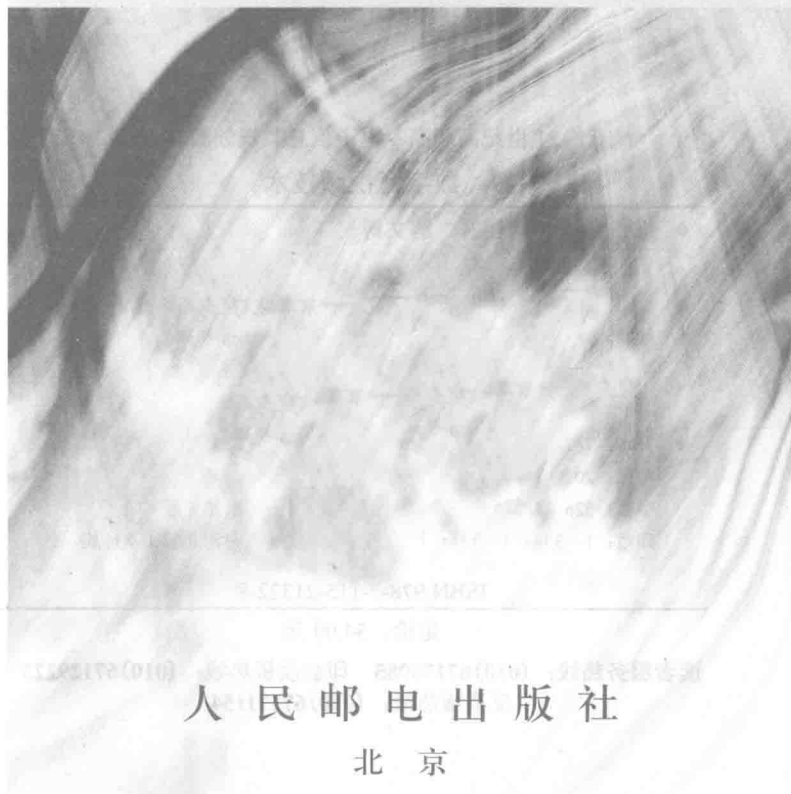
中国通信学会普及与教育工作委员会推荐教材

21世纪高职高专电子信息类规划教材

21 Shiji Gaozhi Gaozhuān Dianzi Xinxilei Guihua Jiaocai

数字音视频 技术

解相吾 解文博 编著



人民邮电出版社

北京

图书在版编目(CIP)数据

数字音视频技术 / 解相吾, 解文博编著. — 北京: 人民邮电出版社, 2009. 11

21世纪高职高专电子信息类规划教材
ISBN 978-7-115-21322-8

I. ①数… II. ①解… ②解… III. ①数字技术—应用—音频设备—高等学校: 技术学校—教材②数字技术—应用—视频信号—高等学校: 技术学校—教材 IV. ①TN912.27②TN941.3

中国版本图书馆CIP数据核字(2009)第187090号

内 容 提 要

本书以数字音视频技术为主线,从实际应用出发,结合典型设备,深入浅出地介绍了数字音视频技术的基本原理和相关技术。全书共分9章,对数字音视频信源编码技术、信道编码技术、调制技术、存储技术及网络传输技术进行了系统阐述;对数字音视频技术在消费电子领域、广播电视领域、多媒体计算机领域及多媒体通信领域的实际应用进行了全面的介绍;精心设计了相关技能训练项目,是一部体系新颖、内容全面的通用基础教材。

本书可作为高职院校和其他高等院校电子信息、应用电子、通信技术、广播电视、多媒体技术等专业的教材,也可以供相关领域的技术人员参考。

21世纪高职高专电子信息类规划教材

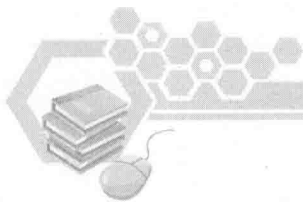
数字音视频技术

- ◆ 编 著 解相吾 解文博
责任编辑 蒋 亮
- ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街14号
邮编 100061 电子函件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
三河市海波印务有限公司印刷
- ◆ 开本: 787×1092 1/16
印张: 20.5
字数: 526千字
印数: 1—3 000册
- 2009年11月第1版
2009年11月河北第1次印刷

ISBN 978-7-115-21322-8

定价: 34.00元

读者服务热线: (010)67170985 印装质量热线: (010)67129223
反盗版热线: (010)67171154



数字音视频技术作为核心的新兴产业,已经成为当前及今后一段时期提升区域经济发展规模和档次的战略性产业。

本书从技术和应用两大方面对信源编码技术、信道编码技术、调制技术、存储技术及网络传输技术进行了系统的阐述,对数字音视频技术在消费电子领域、广播电视领域、多媒体计算机领域及多媒体通信领域的实际应用进行了全面的介绍,让学生掌握数字音视频技术的基本原理,了解数字音视频技术的发展和应用,为学习后续课程打好基础。

在编写过程中,我们遵照循序渐进的原则,结合高职高专实行“双证制”的需要,采用富有弹性的模块化结构,对知识与能力进行有机的综合。内容编排力求由浅入深,通俗易懂,对数字音视频领域中所涉及的各种技术和应用进行了详尽的阐述,并配以丰富的图表和图片,以利于学生形象地掌握最基本的原理和相关技术。本书主要有以下特点。

(1) 技术上的先进性。本书在取材上力求反映数字音视频技术的最新动态,注意吸收新知识、新理论和新技术的成果,重点讲述数字音视频领域中的核心技术和主流技术。

(2) 知识上的完整性。数字音视频技术课程涉及内容广泛,单独讲解某种设备和技术难免有偏概全,不利于完整了解整个系统应有的功能。为便于学生理解,本书对数字音视频技术的相关设备、基本概念、体系结构和相关性进行了完整描述。

(3) 结构上的系统性。现实中的数字音视频领域规模庞大,技术复杂,设备种类繁多,必须科学地、有机地组合在一起才能发挥作用。全书各章由典型设备、基本知识与主要技术、专业技能实训三大模块组成,充分体现“宽基础、薄理论、强技能”的专业教学特色。

(4) 突出了实用性。本书结合高职高专的教学特点,强调理论与实际相结合。由于现代产品的集成度很高,在介绍工作原理时,不必一览无余地对其内部的每一个元器件进行分析。我们只须讲解主要芯片、关键部件和电路的主要功能就可以了。

(5) 注重能力的培养。本书内容较多,涉及面宽,从实用的角度出发,在注意提高学生兴趣的同时,结合职业资格考证的需要,精心安排技能实训项目,让学生在实践中学,提高他们解决实际问题和实践操作的能力。

由于各校的实际情况不同,建议教师采用有所教有所不教和有所学有所不学的策略灵活处理各章内容。

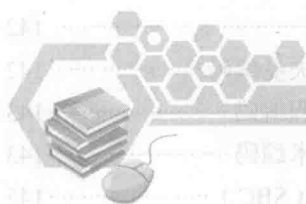
本书由解相吾、解文博编写。第8章、第9章解文博编写,其余各章由解相吾编写,全书由解相吾统稿。徐小英、陈炯尧、吴嘉明、关天军、陈杰辉、钟科科、杨远辉等为本书的资料收集和整理提供了大量的帮助。

本书在编写过程中,参考了大量的文献和资料,书后的参考文献仅列出其中的一部分,其他出处实难一一指出,在此特向本书所引用资料的作者表示衷心的感谢。

由于编写时间仓促,编者水平有限,书中疏漏之处在所难免,欢迎广大读者批评指正。

编者

2009年9月



第1章 概述	1	3.2.2 差分脉冲编码调制	52
1.1 数字音视频技术的基本概念与主要特点	1	3.2.3 自适应差分脉冲编码调制	54
1.2 数字音视频系统的组成	2	3.2.4 增量调制和自适应增量调制	56
1.3 数字音视频系统的业务与应用	5	3.2.5 子带编码	58
1.4 数字音视频技术的发展趋势	10	3.2.6 矢量编码	61
技能训练一：数码视听设备的连接操作	12	3.3 参数编码	62
思考与练习	13	3.3.1 语音生成模型	62
第2章 数字音频技术基础	14	3.3.2 线性预测编码	64
典型设备：CD机	14	3.4 混合编码	65
基本知识与主要技术	20	3.4.1 多脉冲线性预测编码	65
2.1 声学原理	20	3.4.2 规则脉冲激励/长项预测编码	66
2.1.1 声音物理特性	20	3.4.3 码激励线性预测编码	67
2.1.2 人的听觉特性	22	3.4.4 矢量和激励线性预测编码	69
2.2 声音数字化	25	3.4.5 多带激励语音编码	70
2.2.1 数字化概述	26	3.4.6 混合激励线性预测编码	71
2.2.2 模/数变换	30	3.5 MPEG音频编码	72
2.2.3 数/模变换	33	3.5.1 MPEG-1音频	73
2.3 数字音频的主要性能参数	35	3.5.2 MPEG-2音频	74
2.4 数字音频文件的常见格式	37	3.5.3 MPEG-4音频	78
技能训练二：CD/VCD机芯结构观察与拆装	40	3.5.4 MPEG-7音频	79
技能训练三：CD机的电路观测	41	3.5.5 杜比AC-3	80
思考与练习	42	3.6 G.7xx系列编码标准	81
第3章 音频压缩编解码	43	技能训练四：MP3播放器的整机认识与拆装	83
典型设备：MP3播放器	43	技能训练五：MP3播放器的升级扩容	83
基本知识与主要技术	46	思考与练习	84
3.1 音频压缩概述	46	第4章 数字视频编码基础	85
3.1.1 音频压缩的可行性	46	典型设备：DV（数字摄录机）	85
3.1.2 音频压缩编码技术分类	48		
3.2 波形编码	50		
3.2.1 脉冲编码调制	50		



基本知识及主要技术	88	5.3 统计编码	142
4.1 光与色	88	5.3.1 霍夫曼编码	142
4.1.1 光的特性与度量	88	5.3.2 游程编码	143
4.1.2 色彩	90	5.3.3 算术编码	143
4.2 视觉与图像	92	5.4 子带编码(SBC)	145
4.2.1 视觉现象	92	5.4.1 子带编码原理	145
4.2.2 图像	93	5.4.2 子带滤波	146
4.3 视频	96	5.4.3 二维子带编码	147
4.3.1 视频信号的获取和显示	97	5.5 小波变换编码	148
4.3.2 视频信息和信号特点	98	5.6 分形编码	149
4.3.3 模拟视频	100	5.6.1 分形的概念	149
4.3.4 数字视频	103	5.6.2 分形编码原理	150
4.4 视频信号的数字化	106	5.7 知识(模型)基编码	151
4.4.1 A/D 变换器中的舍入法 量化	106	5.7.1 基于模型的编码	151
4.4.2 并串型 A/D 变换器	107	5.7.2 语义基图像编码	152
4.4.3 PCM 编码器	109	5.7.3 物体基图像编码	153
4.4.4 D/A 变换器	110	技能训练七: VCD 机整机电路结构 观测	154
4.5 视频编码	111	技能训练八: VCD 机信号检测与 调整	155
4.5.1 视频压缩原理	112	思考与练习	156
4.5.2 视频编码的方案体系	113	第 6 章 数字视频编码标准	158
4.5.3 视频压缩编码的分类	116	典型设备: DVD 影碟机	158
技能训练六: DV 机的视频采集	120	基本知识及主要技术	160
思考与练习	121	6.1 概述	160
第 5 章 视频压缩编码	122	6.2 ITU-R BT.601 数字视频标准	161
典型设备: VCD 影碟机	122	6.3 H.26X 标准	162
基本知识及主要技术	126	6.3.1 H.261 标准	162
5.1 预测编码	126	6.3.2 H.263 标准	167
5.1.1 预测编码的基本内容	126	6.3.3 H.264 标准	168
5.1.2 预测编码的类型	126	6.4 H.323 可视通信系统标准	173
5.1.3 预测器	128	6.5 JPEG 标准	174
5.1.4 后向预测和双向预测	128	6.5.1 JPEG	174
5.1.5 像素块预测	129	6.5.2 JPEG 2000	177
5.1.6 量化与编码	131	6.6 MPEG 标准	181
5.1.7 具有运动补偿的帧内插	132	6.6.1 MPEG-1 标准	182
5.2 变换编码	135	6.6.2 MPEG-2 标准	186
5.2.1 变换编码的基本原理	135	6.6.3 MPEG-4 标准	191
5.2.2 正交变换矩阵	138	6.6.4 MPEG-7 标准	196
5.2.3 离散余弦变换(DCT)	138		



6.6.5 MPEG-21 标准	198	7.10 残留边带调制	238
技能训练九: DVD 机解码电路的 检测	199	7.11 TCM 网格编码调制	240
技能训练十: DVD 机主要信号波形 测量	200	7.11.1 TCM 编码的基本原理	240
思考与练习	201	7.11.2 TCM 编码的子集划分 与译码	241
第 7 章 数字音视频信号的传输技术	202	7.12 网格编码交织与去交织	242
典型设备: 数字电视	202	7.12.1 网格编码交织原理	242
基本知识 with 主要技术	207	7.12.2 TCM 接收端的去交织	244
7.1 信道编码的基本内容	207	7.13 正交频分复用和编码正交频 分复用	244
7.2 差错控制编码	210	7.13.1 OFDM 数字调制技术	245
7.2.1 差错控制方式	210	7.13.2 克服多径干扰和常规 电视干扰的措施	246
7.2.2 差错编码的基本原理	211	7.13.3 COFDM 与其他地面传输 系统的比较	249
7.2.3 简单差错编码	214	技能训练十一: 数字卫星电视接收 系统的安装与调试	250
7.3 里德-所罗门码 (RS 码)	215	技能训练十二: 机顶盒的安装与 调试	251
7.3.1 RS 码的结构特点	215	思考与练习	252
7.3.2 RS 码的编码	217	第 8 章 数字音视频存储技术	253
7.3.3 RS 码的截短选择	218	典型设备: DC (数码相机)	253
7.3.4 RS 码的译码	218	基本知识 with 主要技术	257
7.4 卷积码	218	8.1 光盘存储技术	258
7.4.1 卷积码的产生	218	8.1.1 光存储技术	258
7.4.2 卷积码的译码	220	8.1.2 光盘的分类	258
7.5 信道编码中的交织技术	221	8.1.3 光盘的结构与格式	260
7.5.1 交织过程	221	8.2 光盘驱动器	265
7.5.2 卷积交织与解交织	226	8.2.1 CD-ROM 驱动器的工作 原理	265
7.6 Turbo 码	227	8.2.2 光盘驱动器的工作模式	267
7.6.1 软输出译码及并行级 联卷积码	228	8.2.3 光盘驱动器的种类	268
7.6.2 Turbo 码的译码	229	8.2.4 光盘驱动器的性能指标	272
7.6.3 Turbo 码性能分析	229	8.3 其他存储技术及其设备	273
7.7 数字电视信号的调制	230	8.3.1 半导体存储设备	274
7.8 数字调相与解调	231	8.3.2 磁存储技术	277
7.8.1 二相移相键控信号的 形成与解调	231	8.3.3 网络存储技术	279
7.8.2 多进制相位调制与解调	232	技能训练十三: 数码相机的维护 保养	281
7.8.3 多进制相位调制特点与 应用	235		
7.9 正交振幅调制 (QAM) 与 解调	235		



技能训练十四：刻录 CD/DVD 光盘	282	9.5 交互电视与视频点播	303
思考与练习	282	9.5.1 概述	303
第9章 数字音视频技术的业务应用	284	9.5.2 VOD 系统的体系结构	304
9.1 数字音频 (DAB) 广播	284	9.6 远程教育	307
9.1.1 DAB 发展概况	284	9.6.1 远程教育分类	308
9.1.2 DAB 系统	286	9.6.2 远程教育的系统功能模块	308
9.2 可视电话	290	9.7 远程医疗	309
9.2.1 可视电话的发展	290	9.7.1 远程医疗系统的构成	309
9.2.2 可视电话的分类	291	9.7.2 远程医疗系统的类型	309
9.3 视频会议	292	9.8 多媒体计算机技术	310
9.3.1 视频会议概述	293	9.8.1 多媒体计算机系统	310
9.3.2 视频会议系统的基本构成	296	9.8.2 MPC 系统的数字视频	313
9.3.3 视频会议系统的新发展	297	9.8.3 MPC 系统的数字音频	314
9.4 视频监控	298	9.9 新业务应用	316
9.4.1 视频监控系统的发 展特点	299	技能训练十五：视频会议系统终端的 安装与调试	318
9.4.2 视频监控系统的结构	299	思考与练习	318
		参考文献	320

第1章

概述

学习目标

知识目标：了解数字音视频技术的基本概念，掌握数字音视频系统的结构组成，熟悉数字音视频系统的业务与应用，把握数字音视频技术的未来发展趋势。

技能目标：掌握 DVD 与 TV、音频功率放大器的连接方法；正确操作使用 DVD、TV、音频功率放大器。

随着现代科技的不断发展，以信息技术产业为代表的高新技术产业，得到了迅猛发展，推动了全球产业结构转型和优化升级，带来了人类生产和生活的深刻变化。数字音视频技术已成为当前最流行、使用最频繁、应用范围最广泛的新技术，它与人们的生活、工作和娱乐紧密联系在一起，日益深刻地改变着人们的生活方式，掀起了一场声势浩大的信息革命。激光视盘机（VCD 和 DVD）、数字媒体播放器（MP3 和 MP4）、数码相机（VC）、数字摄录机（DV）、数字电视和高清晰度电视（HDTV）以及移动数字视频（CMMB）等数字音视频设备层出不穷，把人们带进了一个五彩缤纷的奇妙世界。

1.1 数字音视频技术的基本概念与主要特点

数字音视频技术是信息领域的基础技术之一，随着大规模集成电路、计算机数字技术的发展，传统的影视传媒、消费类电子以及通信行业几乎全部实现了数字化。数字化促进了这些行业的迅速发展，同时也将原来不同的行业——计算机、通信、影视传媒、消费类电子等汇聚在一起。所谓数字化，是指信息的采集、传输、交换和处理过程全面采用数字化技术。

数字音视频技术是对音视频信息（文本、图形、图像、声音、动画、视频等）进行采集、获取、压缩/解压缩、编辑、存储、传输及再现等环节全部采用数字化的技术。数字音视频技术的发展推动了音视频产品的发展，音视频产品的数字化进一步提高了产品的技术含量。与传统模拟技术相比，数字音视频技术有如下特点。



(1) 传输效率较高。音视频数字信号经过压缩后,可以在 6~8MHz 的传输信道内传输 2~4 套标准清晰度电视 (SDTV) 节目或一套高清晰度电视 (HDTV) 节目。

(2) 信息传输存储灵活方便。数字信号便于存储、控制、修改,存储时间与信号特点无关,存储媒体的存储容量大,存储媒体可以是 CMOS 型半导体存储器,也可以是计算机的硬盘、高密度激光盘等。

(3) 信息传输存储的可靠性高。数字信号的检错、容错、纠错能力很强,在数字信号传输放大过程中如出现误码,很容易实现检错与纠错。

(4) 抗干扰能力强。不会产生噪波和失真的累积。

(5) 有效保护信息和进行版权管理。便于实现加密/解密技术和加扰/解扰技术,便于专业应用(军用、商用、民用)或条件接收、视频点播、双向互动传送等。

(6) 具有可扩展性、可分级性和可操作性。易于与其他系统配合使用,在各类通信信道和网络中传输,构成一个灵活、通用、多功能的综合业务信息传输网。

(7) 便于与其他数字设备融合。因为它们的信号语言是相同的,只要有一套数字信号传输、编码、调制协议,就可以做到互联、互通。以音视频数字化为代表的消费类电子,正逐渐与电子计算机、通信技术相融合。

(8) 易于集成化和大规模生产,其性能一致性好,且成本低。

1.2 数字音视频系统的组成

数字音视频信息系统的整体模型如图 1-1 所示。其中,信源编码和信源解码统称为信源编码,信源编码主要解决有效性问题,旨在通过对信源的压缩、扰乱、加密等一系列处理,用最少的数码传递最大的信息量,使信号更适宜传输存储;信道编码和信道解码统称为信道编码,主要解决可靠性问题,旨在尽量使处理的信号在传输/存储过程中不出错或少出错,即使出了错也要能自动检错和尽量纠错,通常包括调制和纠错编码,前者主要用于解决码间干扰产生的错误,后者主要用于解决“噪声”引起的突发性错误(如光盘刮伤、污迹等);格式编/解码统称为格式编码,主要解决高效性问题,旨在通过对所存储/传输数据的组织达到提高数据存取速度的目的。传输通道或介质统称为信道,实际信道可以是由光缆或电缆构成的有线信道,也可以是高频无线线路、微波线路或卫星中继等构成的无线信道。存储介质可以是磁带、磁盘、光盘等。无论是何种介质,都将受到不同性质的噪声干扰。信源和信宿指的是音视频的采集和重放等终端设备。

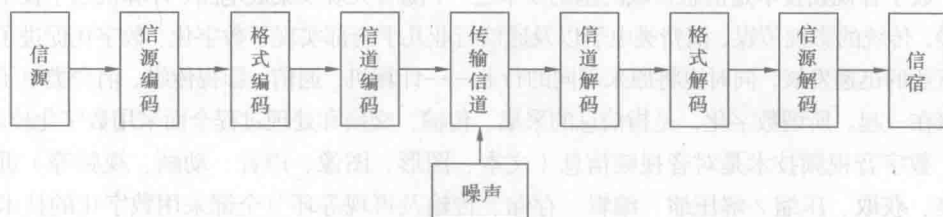


图 1-1 数字音视频信息系统模型

数字音视频系统涉及的方面很多,应重点了解以下内容。



1. 数字音视频信息处理

(1) 信息处理。信息处理主要包括信息获取, 信息变换, 信息特征的抽取与选择、分类与识别, 信息安全与授权等方面的内容。

信息获取是信息处理的基础, 包括传感技术和提取技术两个主要方面。传感技术是自然界与信息处理系统的界面和接口, 灵敏、精确、可靠的传感器件是传感技术的核心。目前传感技术发展的总趋势是应用微电子技术、超高精密加工以及超导、光导与粉末等新材料, 使新型传感器具有集成化、多功能化和智能化的特点。提取技术是从已经获取的信号中提取感兴趣的信息, 这是信号处理技术的一种应用。

各种声音信号、图像信号是音视频信源信息的载体, 从这些信号中提取音视频信息的本质特征, 进行计算机自动分类和识别是信息利用的基本方式, 也是信息处理的主要内容。

音视频信源的主要特征有几何特征、数学特征和结构特征等。特征抽取与选择的主要任务是针对给定识别任务, 按照预定判别准则, 选择那些数量少, 对提高正确识别贡献最大的参数作为特征, 因而它主要是一个统计优化问题。特征选择的优劣最直接影响到识别性能。由于自然界待识别的音视频信息的复杂多样性, 特征选择也就成为最重要和最复杂的问题之一。特征选择还应考虑到不变性特征的选取, 如图像识别中的尺寸、位置, 如旋转不变性的选择等。不变性的数学基础是群的性质, 群论是不变性选择的主要理论依据。

分类与识别也称模式识别, 有统计模式识别和结构(句法)模式识别之分。近年来, 有将两者结合起来的发展趋势。模式识别的重要应用是语音识别、图像识别和文字识别, 即视觉和听觉信息的识别。

信息安全与授权也是信息处理的重要内容。信息安全主要是指其保密性, 即信息系统的信息不被泄露给非授权用户、实体或过程, 或供其利用。信息授权是指信息系统的信息可被授权实体访问并按需使用, 即信息系统的服务在需要时, 允许授权用户或实体使用的特性。

(2) 信息的数字化处理。数字信号处理(Digital signal Processing, DSP)技术通常是指利用计算机或/和专用处理设备(或器件), 对数字化的信号进行采集、滤波、检测、均衡、变换、调制、压缩、降噪、估计等处理, 以得到更能满足人们需要的信号形式。例如, 对音视频信号的降噪处理的目的是减少原音视频信号中的噪声或干扰信息, 使信号变换成更便于人们识别和理解的形式。

数字信号处理是相对于模拟信号处理而言的, 显然这里的“处理”就是对数字化的信息信号施加某种数学运算。自从1965年快速傅立叶算法发明以来, 数字信号处理技术得到长足发展, 随着超大规模集成电路技术和计算机技术的发展, 各种快速数字信号处理器件大量问世, 并得到广泛应用, 数字信号处理比模拟信号处理有更显著的优点, 例如, 功能强大、精度高、灵活性强、稳定性好、抗干扰能力强、设备体积小等。因此, 更易于大规模集成化和实现多维处理, 可以用于更广泛的领域。

2. 数字视频处理系统

图1-2是数字化视频处理系统的模型, 由5个功能模块构成: 输入模块、输出模块、存取/通信模块、控制与存储模块, 以及作为核心的视频处理模块。

(1) 视频输入设备。视频输入设备的主要功能是将待处理视频/图像信号输入系统, 根据应用需求不同, 往往有不同的设备形式。如摄像头、数字摄录像机、红外/X光摄像机、数字照相机、



激光视盘机、触摸屏、视频采集卡、扫描仪等静态或动态图像形成/存储设备,以及诸如电视调谐器等视频传输设备。



图 1-2 视频处理系统模型

(2) 视频输出设备。视频输出设备的主要功能是将经过系统处理后的视频/图像信号以用户能感知的形式显示出来。目前,最常用的视频/图像输出设备有:各类电子显示设备,如阴极射线荧光屏(CRT)、液晶显示屏(LCD)、等离子体显示屏(PDP)、电致发光显示器(ELD)和荧光管显示器(VFD)等;各类硬拷贝设备,如打印机、彩色绘图仪和硬拷贝机等。

(3) 存储与控制设备。控制设备主要用于在视频处理过程中对视频处理设备进行控制,如键盘、鼠标、控制杆、各种开关等;存储设备主要用于在视频/图像处理过程中对视频/图像信息本身和其他相关信息进行暂时或永久保存,如各种RAM、ROM、闪存、硬盘、光盘、磁带机等。

(4) 用户存取/通信设备。用户存取/通信设备的主要功能是使用户按需将已处理好的,或还需进一步处理的视频/图像信号取出或送入视频处理模块。存取一般是指本地的存取操作,如光盘、磁盘、磁带或各种存取器件;通信则是相当于远端的存取操作,如基于LAN、数字通信网等的通信设备等。

(5) 视频处理设备。实际的视频处理设备大到分布式计算机组、大型计算机(工作站),小至一台PC、甚至一块DSP芯片,但无论什么形式它都是一个复杂的软、硬件系统。目前,视频处理设备一般分为两类,一类是软件型视频处理系统,计算机插入一个视频处理卡,视频处理卡中有刷新视频存储器等专用硬件,而处理工作由计算机软件完成;另一类是专用型计算机视频/图像处理系统,用专用硬件进行视频/图像处理,而主机仅发布命令和进行部分处理,视频/图像硬件处理可以脱机进行。

3. 数字语音信息处理系统

数字语音信息处理系统的构成与数字视频/图像信息处理系统相似,是指把语音信号数字化并在数字状态下进行记录、传输、重放以及其他加工处理等一整套技术。研究对象是语音信息的获取、表示、传输与处理(编码、存储、变换、识别、综合、理解)的方法、规律及其利用。它以物理声学、生理声学、心理声学、语言学、语音学为基础,涉及包括电声技术、电子技术、微电子技术、数字技术、计算机控制技术和精密加工技术在内的多个学科领域。

自1874年电话发明以来,语音处理取得迅猛发展,建立了语音产生的声学模型,发展了语音信号处理的理论和方法,开发出包括语图仪在内的研究语音信号处理算法的仪器设备,研制出多种语音处理专用大规模集成电路,实现了多种应用范围较广、实用性较强的自动语音机器。目前,语音处理研究的重点是完善语言产生模型,加强语言感知模型研究,建立一个完整的语言产生模型;利用听觉的心理特性,如声音的掩蔽效应,实现大压缩比高效编码,并将其应用于语音通信、广播电视和电声产品。



4. 数字化音视频信息系统的性能指标

数字音视频信息系统的根本任务是传播音视频信息。人们总希望从音视频信息系统中所获取的音视频信息愈多愈好,失真愈小愈好,使用愈安全愈好,操作愈方便愈好。因此,衡量一个音视频信息系统优劣程度的最重要的性能指标,是它传输信息的有效性、可靠性、安全性和高效性。这些指标是相互矛盾而又互为联系的,并在一定的条件下是可转换的。例如,人们不可能同时获得有效性和可靠性的高指标,只能根据具体要求获得一个折中指标,即只能在满足一定可靠性指标下,尽量提高信息传输的有效性,或是在维持一定有效性的指标下,使传输信息的可靠性尽量提高。这里所说的可靠性,是指信息传输/存储的可靠性,而不是指设备和元件器工作的可靠性。影响信息传输/存储可靠性的基本因素是信道内外的噪声(干扰)。无疑,系统对噪声的抵抗能力越强,其传输信息的可靠程度就越高,因此,信息传输/存储的可靠性可用系统的抗噪声性能来体现。

(1) 有效性。数字音视频信息系统的有效性是指在给定信道内能够传输的信息的多少,通常用码元传输速率或信息传输速率来进行度量。

码元传输速率又称为波特率或传码率。码元传输速率是指单位时间(通常为 s)内传输的码元数目,记为 R_B ,其单位为波特(baud 或 B),它反映了数据的调制信号波形变换的频繁程度。

信息传输速率是指单位时间内传输的信息量,记为 R_b ,其单位为比特/秒(bit/s)。信息传输速率又称为传信率、码率或比特率。根据信息量的定义,一个二进制码元含有一比特(bit)的信息量。因此,在二进制情况下,码元传输速率与信息传输速率在数值上是相等的,但含义不同,单位不同。一个多进制码元所含的信息量为 $I = \log_2 m = k(\text{bit})$ 。式中, $m = 2^k (k = 1, 2, 3, \dots)$ 表示进制数,如 $m = 4, I = 2\text{bit}$,即一个四进制码元含有 2bit 信息量。在多进制的情况下,信息传递速率与码元传输速率存之间的关系为: $R_b = R_B \log_2 m$ 。

(2) 可靠性。数字音视频信息系统的可靠性是指信宿所接收到的信息的准确程度,通常用码元差错概率或信息差错概率来进行度量。码元差错概率是指所传输的码元总数目中发生差错的码元数目所占的比值(取统计平均值,下同),记为 P_e ,码元差错概率简称误码率。信息差错概率是指所传输的信息比特总数目中发生差错的信息比特数目所占的比值,记为 P_b 。信息差错概率又称为比特差错概率或误比特率。在二进制情况下,码元差错概率和信息差错概论在数值上相等,即 $P_e = P_b$,但含义不同。

导致数字信息在传输/存储过程产生差错的主要原因是不同数字信息系统的性能及其在传输/存取过程中受到的不同干扰。因此,要从多种途径来研究提高系统可靠性的方法。目前,主要有三类方法:选择合理的调制技术、改善信道及存储介质以及采用抗干扰编码(即信道编码或纠错编码)。

(3) 保密性。数字音视频信息系统的保密性是指数字音视频信息系统通过授权技术,以防范在系统中传输/存储信息被非授权信宿终端所接收、正确解码和利用的有效程度。

(4) 便利性。数字音视频信息系统的便利性是指信宿终端用户按需从系统中检索出目标信息的简捷程度。

1.3 数字音视频系统的业务与应用

当前,信息技术的发展日新月异,信息技术的普及和应用对于经济、政治、社会、文化,包括军事等方面,都产生了深刻的影响。以数字音视频为代表的信息技术已催生出许多热点应用,



主要包括广播电视、视频点播、可视电话、视频会议、远程教育、远程医疗、远程监控、远程交易、互动游戏等。归纳起来,可分为如下3个方面。

1. 在广播电视与消费类电子产品中的应用

早期的音响设备和广播电视均采用模拟技术,近20年来随着数字技术的发展逐步被数字技术取代。最早出现的数字音响设备CD机成功地取代了磁带录音机,随后数字化革命进入视频领域,VCD机将数字视频引入家庭。美国休斯公司卫星直播、数字电视Direct TV的成功,标志着数字电视时代的到来。数字音频广播DAB、数字电视广播DVB将从卫星直播到有线电视再到地面广播逐步普及。为保证现有的大量电视机能继续使用,数字电视将以在普通电视机上加装机顶盒的形式来实现。近几年问世的MP3、MP4、CMMB等信息家电产品,更进一步加快了数字化的进程。让消费者开始明白“任何内容、任何时间”的真正含义,并了解了多功能、海量存储设备在家庭及便携式产品市场中实现的可能性。

数字电视(DTV)技术一直是数字音视频领域研究的一个热点和难点。美国大联盟方案和ATSC标准确定之后,DTV已进入技术实施阶段。数字电视技术将把消费者和电视行业带入一个新的时代。与传统模拟电视相比,DTV除提高图像和声音的接收质量外,更主要的变革是使传统电视从被动式的、单纯的娱乐媒体,变成集视听娱乐、通信业务和信息服务于一体的交互式家用电器,形成一项新的基于TV的交互式业务。新的数字电视标准有足够的带宽允许电视台同时播放多套电视节目,或者同时传送数据和电视节目,开创了基于TV的交互式数据业务。根据DTV的发展趋势,可以预测,在未来10年内,消费者完全可以在家里通过数字电视,将以电影级别的真体验在欣赏电视节目的同时,享受电视购物、电视银行、电视商务、电视通信、电视游戏、电视办公自动化、电视实时点播和电视网上游览的乐趣。

图1-3从存储和传输两个角度简略地勾勒出了各种广播电视和视听消费电子类设备和产品数字化的发展情况。

2. 在多媒体计算机中的应用

在计算机领域,由于多媒体技术的发展而出现了多媒体PC。多媒体PC进入家庭,成为一种可看电影、听音乐、玩电子游戏的消费电器。多媒体计算技术(multimedia computing technology)是指计算机综合处理文本、图形、图像、视频和音频等多种媒体信息,使多种信息建立逻辑链接,集成为一个系统并具有交互性。多媒体技术的基本特征是信息载体多样性、集成性和交互性。

从开发和生产厂商以及应用的角度看,多媒体计算机可以分为两大类:一类是家用电器制造商研制的电视计算机(Teleputer),它把CPU放到家用电器中,通过编程控制管理电视机、音响等家用电器,因此又称为“灵巧”电视(Smart TV);另一类是计算机制造商研制的计算机电视(Compuvision),它采用微处理器(80x86、68xxx)作为CPU,其他设备还有VGA卡、CD-ROM、音响设备以及扩展的多窗口系统,通常被称为TV-Killer。

要把一台普通的计算机变成多媒体计算机要解决的关键技术有:①视频音频信号获取技术;②多媒体数据压缩编码和解码技术;③视频音频数据的实时处理和特技;④视频音频数据的输出技术。图1-4是多媒体计算机系统的基本构成。

多媒体计算机需要计算机交互式地综合处理图像、声音、文字信息,尤其是图像和声音信息数据量大,处理速度要求高,用过去的通用计算机很难完成上述任务。例如,多媒体计算机中用



的最多的静态图像压缩和解压缩算法，如果选用 Motorola 公司生产的 25MHz 的 68 030 处理器芯片，完成一幅分辨率为 $512 \times 512 \times 24\text{bit}$ 的彩色静态图像的 JPEG 算法大约需要 15min，用 IBM PC/386 (33M) 完成上述算法大约只需 1min。为了较好地解决多媒体计算机综合处理图像、声音和文字的信息问题，可以采用下述 3 种方式。

(1) 选用专用芯片设计专用接口卡单独解决，例如使用声霸卡解决声音的输入、输出和实时编码、解码的处理问题；使用视霸卡解决视频信号的输入、输出及多窗口的彩色键连问题；使用视频信号压缩编码和解码卡解决视频信号的压缩和解压缩问题；使用局域网的 ISDN 网接口卡解决局域网和远程网的多媒体通信问题。这种方式就是多媒体个人计算机 (MPC) 的标准和结构。

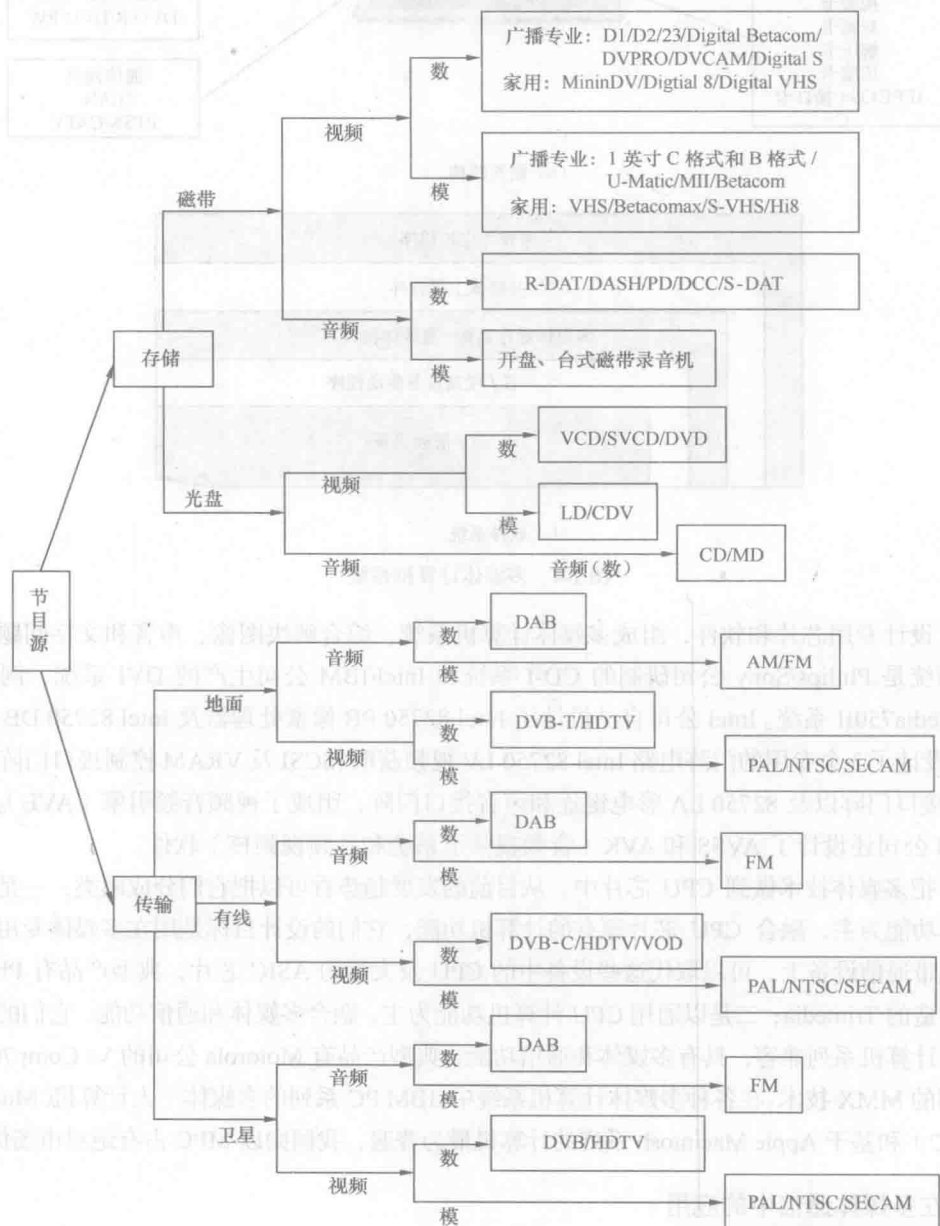
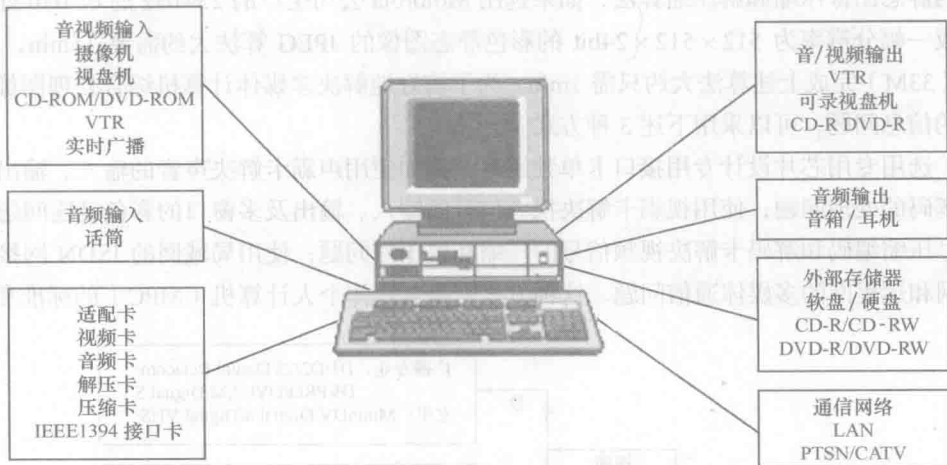
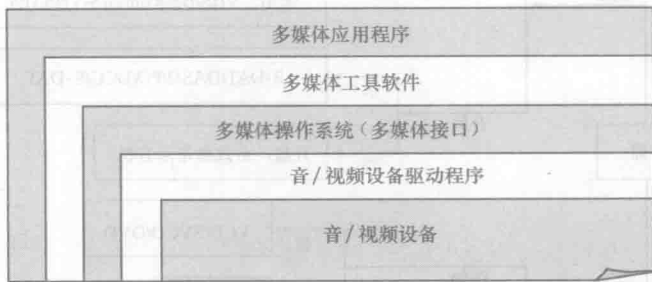


图 1-3 广播电视和视听消费电子类设备和产品的数字化发展



(a) 硬件结构



(b) 软件系统

图 1-4 多媒体计算机系统

(2) 设计专用芯片和软件，组成多媒体计算机系统，综合解决图像、声音和文字问题。比较成功的系统是 Philips/Sony 公司研制的 CD-I 系统和 Intel/IBM 公司生产的 DVI 系统，例如 DVI Action Media750II 系统。Intel 公司自己设计了 Intel 82750 PB 像素处理器及 Intel 82750 DB 显示处理器，还设计了 3 个专用的门阵电路 Intel 82750 LV 视频获取、SCSI 及 VRAM 控制接口门阵，82750 LH 主机接口门阵以及 82750 LA 彩电键连和声音接口门阵，组成了视频音频引擎 (AVE)。同时，Intel/IBM 公司还设计了 AVSS 和 AVK (音频视频子系统和音频视频核) 软件。

(3) 把多媒体技术做到 CPU 芯片中。从目前的发展趋势看可以把它们分成两类：一是以多媒体和通信功能为主，融合 CPU 芯片原有的计算机功能，它们的设计目标是用在多媒体专用设备、家电及宽带通信设备上，可以取代这些设备中的 CPU 及大量的 ASIC 芯片，典型产品有 Philips 公司设计制造的 Trimedia；二是以通用 CPU 计算机功能为主，融合多媒体和通信功能，它们的设计目标与现有计算机系列兼容，具有多媒体和通信功能，典型产品有 Motorola 公司的 Ve Comp701 以及 Intel 公司的 MMX 技术。在各种多媒体计算机系统中，IBM PC 系列的多媒体个人计算机 (Multimedia PC, MPC) 和基于 Apple Macintosh 系列的计算机最为普遍，我国则以 MPC 占有绝对市场优势。

3. 在多媒体通信中的应用

数字技术的进步推动了信号的数字化，使文本、声音、图形、图像等各种单媒体信号能够以



数字的形式,按照一定的逻辑关系融合成多媒体信号。信号的数字化为多媒体的产生提供了条件。与传统的电话、电报、传真、计算机通信等单一媒体的通信方式相比较,利用多媒体通信,用户不仅能声像图文并茂地交流信息,而且对通信的全过程具有完备的交互控制能力。

多媒体通信系统是指完成多媒体通信业务的系统,包括通信终端、通信设备、传输通路、多媒体应用服务设备等。多媒体通信系统具有分布协同多媒体环境,能够通过网络完成多媒体信息的处理和传送,并支持交互式以及广播和多播方式。

多媒体通信是信息高速公路建设中的一项关键技术,是多媒体、通信、计算机和网络等技术相互渗透和发展的产物,是计算机、通信和广播电视三网合一的必然趋势。计算机联网通信的要求推动了数据通信的发展,从传输到交换全部实现了数字化。而数字通信网络技术的发展反过来又推动计算机技术进入到以网络为中心的计算机时代。自美国政府1993年提出国家信息基础设施建设倡议以来,全世界掀起了建设信息高速公路的热潮。国际互联网尤其是环球网近两年来获得巨大成功,被人们看成是未来信息高速公路的雏型。互联网的成功对传统通信业产生了重大冲击。接入互联网收发电子邮件、浏览环球网络将成为人们生活中娱乐、教育、社交活动的重要手段。人们在家庭中更习惯于使用电视机而不是计算机,电视机是人们日常获得信息的主要手段,未来的电视机将是一种可以用来浏览互联网的设备。

多媒体通信的业务大致可以分为6类。

(1) 会议业务:在人与人之间多点通信且双向信息交互的业务。典型的业务应用包括点对多点的远程多媒体教学、多点对多点的多媒体互动会议等。

(2) 谈话业务:在人与人之间点对点通信且双向信息交互的业务。典型的业务应用包括可视电话等。

(3) 分配业务:在机器与人之间点对多点通信且单向信息交换的业务。典型的业务应用包括广播型影视放送、互动型视频点播等。

(4) 检索业务:在人与服务器之间点对点通信且单向信息交换的业务。典型的业务应用包括实时交通信息查询、网上图书馆、网上书店、各种电信业务费用查询及支付、人才中介等。

(5) 采集业务:在机器与机器或机器与人之间多点对一点通信且单向信息交换的业务。典型的业务应用包括远程故障诊断、远程医疗等。

(6) 消息业务:在人与服务器之间点对点或点对多点通信且单向信息交换的业务。典型的业务应用包括多媒体短信、语音或视频电子邮件等。

在实际应用中,上述业务并不是孤立存在的。如在远程教育中用到多种多媒体业务;实时授课为谈话业务;作业提交为消息业务;资料查询为检索业务。

多媒体应用主要包括以下方面。

(1) 办公自动化。

(2) 远程会议。

(3) 交互式电视与视频点播。

(4) 交互式影院与数字化电影。

(5) 出版。

(6) 信息查询。

(7) 数字化图书馆。

(8) 家庭信息中心。