

21世纪高等学校应用型规划教材

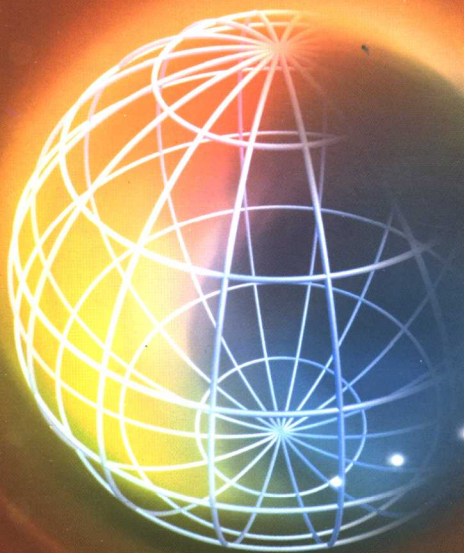


计算机系列

多媒体技术

原理与实践

郑成增 主编
陈志锋 副主编
朱巧明 主审
郑 淼



中国电力出版社

www.infopower.com.cn



多媒体技术

原理与实践

郑成增 主编
郑 淼 副主编
陈志锋 主审
朱巧明 主审



中国电力出版社
www.infopower.com.cn

内容提要

本书是 21 世纪高等学校应用型规划教材计算机系列之一,是作者在长期从事多媒体技术开发,积累了
一定实践经验的基础上编写的。为了适应人们的学习认知规律,本书在讲述多媒体技术原理的基础上,
按照以理论知识点为线索,以实践应用为目的的技术定位模式,精心编排全书的内容,并配有多媒体
网络课件的开发过程演示及部分课件内容(附书光盘)。本书充分吸收了面向 21 世纪多媒体技术实践
课程改革的成果,采用以基础带动技术应用,以技术促进应用实践,以实践加强知识能力的培养,使
学生学有所得、学有所用。全书分三部分,前两部分是基础知识,包括多媒体技术原理和系统应用;
第三部分以实验的形式,通过实例详细指导学生掌握常用多媒体软件的操作与使用。

本书可作为面向应用型人才培养的高等学校计算机相关专业的教材使用,也可作为相关技术人员的参
考书。

图书在版编目(CIP)数据

多媒体技术原理与实践 / 郑成增主编. —北京: 中国电力出版社, 2005.9

21 世纪高等学校应用型规划教材·计算机系列

ISBN 7-5083-1531-6

I. 多... II. 郑... III. 多媒体技术—高等学校—教材 IV. TP37

中国版本图书馆 CIP 数据核字(2005)第 093313 号

丛 书 名: 21 世纪高等学校应用型规划教材·计算机系列

书 名: 多媒体技术原理与实践

出版发行: 中国电力出版社

地 址: 北京市三里河路 6 号

邮政编码: 100044

电 话: (010) 68358031 (总机)

传 真: (010) 68316497, 88383619

本书如有印装质量问题, 我社负责退换

服务电话: (010) 88515918 (总机)

传 真: (010) 88518169

E-mail: infopower@cepp.com.cn

印 刷: 北京同江印刷厂印刷

开本尺寸: 185×260

印 张: 18.5

字 数: 404 千字

书 号: ISBN 7-5083-1531-6

版 次: 2005 年 9 月北京第 1 版

印 次: 2005 年 9 月第 1 次印刷

印 数: 0001—4000 册

定 价: 29.80 元(附光盘)

版权所有, 翻印必究

21 世纪高等学校应用型规划教材·计算机系列

编 委 会

主任委员:

宗 健 常明华

副主任委员:

顾元刚 陈 雁 杨翠南 林全新 华容茂 曹泰斌

魏国英 邵晓根 庄燕滨 邓 凯 吴国经 常晋义

许秀林 谢志荣 张家超 陶 洪 龚兰芳 刘广峰

丁 雁 方 岩 王一曙

委 员:(以姓氏笔画为序)

丁志云 及秀琴 石振国 李 翊 吕 勇 朱宇光

任中林 刘红玲 刘 江 刘胤杰 许卫林 杨劲松

杨家树 杨伟国 郑成增 张春龙 闵 敏 易顺明

周维武 周 巍 胡顺增 袁太生 高佳琴 唐学忠

徐煜明 曹中心 曾 海 颜友钧

序 言

进入 21 世纪,世界高等教育已从精英教育走向了大众教育。我国也适应这一潮流,将高等教育逐步推向大众化。培养应用型人才已成为国家培养国际人才的重要组成部分,且得到了社会各界的广泛支持。于是一大批有规模、有实力、规范化、以培养应用型人才为己任的高等学校得到了长足发展。这类高校办学的一个显著的特点是按照新时代需求和当地的需要来培养学生,他们重视产学研相结合,并紧密地结合当地经济状况,把为当地培养应用型人才作为学校办学的主攻方向。

这类学校的教学特点是:在教授“理论与技术”时,更注重技术方法的教学。在教授“理论与实践”时,更注重理论指导下的可操作性,更注意实际问题的解决。因此,这些学生善于解决生产中的实际问题,受到地方企事业单位的普遍欢迎。

为满足这类高校的教学要求,达到培养应用型人才的目的,根据教育部有关重点建设项目的要求和相关教学大纲,我们组织了多年在这类高校中从教,并具有丰富工程经验的资深教授、高级工程师、教师来编写这套教材。

在这套教材的编写中,我们提倡“实用、适用、先进”的编写原则和“通俗、精练、可操作”的编写风格,以解决多年来在教材中存在的过深、过高且偏离实际的问题。

实用——本套教材重点讲述本行业中最广泛应用的知识、方法和技能。使学生学习后能胜任岗位工作,切实符合当地经济建设的需要和社会需要。

适用——本套教材是以工程技术为主的教材,所以它适用于培养应用型人才的所有高校(包括本科、专科、技术学院、高职等),既符合此类学生的培养目标,又便于教师因材施教。

先进——本套教材所选的内容是当今的新技术、新方法。使学生在掌握经典的技术和方法之后,可用教材中的新技术、新方法去解决工程中的技术难题,为学生毕业后直接进入生产第一线打下坚实的基础。

通俗——本套教材语言流畅、深入浅出、容易读懂。尽量避开艰深的理论和长篇的数学推导,尽量以实例来说明问题,在应用实例中掌握理论,使学生轻松掌握所学知识技能,达到事半功倍的效果。

精练——本套教材选材精练、详细而不冗长、简略得当,对泛泛而谈的内容将一带而过,对学生必须掌握的新技术、新方法详细讲,讲透、讲到位,为教师创造良好的教学空间和结合当地情况调整教学内容的余地。

可操作——本套教材所有的实例均是容易操作的,且是有实际意义的案例。把这些案例连接起来,就是一个应用工程的实例。通过举一反三的应用,使学生能够在更高层次上创造性地应用教材中的新思想、新技术、新方法去解决问题。

本套教材面向培养应用型人才的高等学校,同时也可作为社会培训高级技术人才的教材和需要加深某些方面知识技能的人员的自学教材。

编委会

代 序

21 世纪的人类社会步入了信息化社会，技术、科学和社会的发展正在迎接这个时代的全面到来，IT 业的进程在不断加快，数字化、信息化带动工业化、改造传统产业在不断提速，这些带有革命性的变化都为社会主体共同分享技术进步和信息资源，提高劳动生产率和生活质量提供了一个前所未有的发展空间。

尤其网络多媒体技术经过十几年的不断摸索和研究，自 20 世纪 80 年代中后期开始，就成为人们非常关注的热点之一。多媒体技术的出现从根本上改变了昔日基于字符的各种计算机处理，不但产生了丰富多彩的信息表现能力，而且还能形成可视听媒体的人机界面，在一定程度上改变了人们的交互方式、生活方式、工作方式和整个经济社会的面貌，给人类的工作、生活带来一场革命。正如专家们预测的计算机发展方向是 MODN：M 为多媒体技术（Multimedia），O 为开放系统（Open System），D 为小型化（Downsizing），N 为计算机网络技术（Network Computing），其中多媒体技术位于第一，“交互性多媒体技术已经成为一股不可阻挡的潮流”。

顺应潮流的发展，培养大批能适应这种发展要求的高水平专业人才是一项重要的战略任务。要完成这项任务，培养人才的“一剧之本”——教材是教学过程中确保教学质量的关键之一。我翻阅了该书的全稿，可以看出，本书是作者在自己多年的教学实践并翻阅了大量的文献、资料基础上，根据培养应用型高级人才的需要，遵照理论上够用、原理上讲通、实践上系统强化的人才培养模式，编写了此教材。我相信本书的出版必将为应用型高等院校多媒体技术与实践教学发挥重要作用。

苏州大学校长
钱培德 教授、博导
于苏州

前 言

为了贯彻教育部“面向 21 世纪教学内容与课程体系改革”的精神，遵照面向 21 世纪高等工程教育教材规划研讨会成果，我们组织编写了《多媒体技术原理与实践》。

多媒体技术是以计算机技术为核心，综合处理形、声（文字、图像、图形、声音和视频等）信息的数字化处理技术。随着计算机技术的发展，多媒体技术已成为当今信息社会中的主流技术之一，它正在推动着许多产业的发展，改变着人们的生产和生活方式，人们有理由相信，多媒体技术的广泛应用必将对 21 世纪的社会发展产生巨大的影响。

本书是作者在长期从事多媒体技术开发，积累了一定实践经验的基础上编写的。为了适应人们的学习认知规律，在讲述多媒体技术原理的基础上，以其理论知识点为线索，突出实践应用之目的的认知定位模式，精心编排全书的内容，并配有作者已开发完成的“教育部远程教育资源库项目”——《Visual Basic 入门教程》多媒体网络课件的主要章节内容（附书光盘）。

全书在编写过程中，参阅了大量的国内外资料，在此，谨向其中的诸位作者表示感谢。

本教材由常州工学院郑成增教授主编，郑淼、陈志锋副主编，眭莉、许晓宇、常睿等参编。

本书的第 1、2、3 章由郑成增编写，第 4、5、6、7、11、14 章由郑淼编写，第 8、12 章由陈志锋编写，第 9 章由许晓宇编写，第 10 章由常睿编写，第 13 章由眭莉编写，全书由郑成增教授主编并统稿。

苏州大学校长、博士生导师钱培德教授在通览本书后，欣然为本书作序；苏州大学朱巧明教授在百忙之中审阅了全书，并提出若干中肯的修改意见；另外，本书在编写出版过程中得到了中国电力出版社的关心和支持。谨此一并表示深切的谢意。

信息社会在日新月异的发展、进步，多媒体技术新的思想、方法和工具会不断推陈出新，作者衷心希望得到各位读者的支持，共同探讨多媒体技术的发展动向和相关教学及实践中的体会，以便在本书修订中得到及时的体现。由于编者水平有限，书中难免有疏漏之处，敬请读者批评指正。

作 者

E-mail: zhengcz@czu.cn

2005 年 5 月

目 录

序 言

代 序

前 言

第 1 章 基本概念和基础知识.....	1
1.1 多媒体的基本概念	1
1.2 多媒体	3
1.3 多媒体技术	12
1.4 声音、颜色和图像	15
1.5 多媒体技术的应用	25
1.6 多媒体技术的由来和发展	30
思考与练习	32
第 2 章 超文本与超媒体技术.....	33
2.1 超文本与超媒体	33
2.2 超媒体系统的组成和原理	37
2.3 多媒体信息交换及其标准	43
思考与练习	43
第 3 章 多媒体计算机系统	44
3.1 多媒体硬件子系统	44
3.2 多媒体软件子系统	48
3.3 多媒体操作系统	50
3.4 Windows XP 的新特性	51
3.5 Windows 多媒体工具和实用程序.....	54
思考与练习	63
第 4 章 多媒体音频技术	64
4.1 音频技术基础	64
4.2 音频编码算法和标准	67
4.3 音乐合成与 MIDI 规范	73
4.4 语音识别技术	77
4.5 文字-语音转换技术	79
4.6 声音处理硬件及软件	80
思考与练习	85

第 5 章 多媒体图像视频技术	87
5.1 图像视频技术基础	87
5.2 图像文件格式	91
5.3 图形、图像处理常用硬件	100
5.4 图形、图像处理常用软件	106
思考与练习	108
第 6 章 多媒体数据压缩技术	109
6.1 多媒体数据压缩技术概述	109
6.2 量化	112
6.3 统计编码	115
6.4 变换编码	119
6.5 数据压缩编码国际标准	122
第 7 章 多媒体数据存储技术	129
7.1 CD-ROM 概述	129
7.2 CD-ROM	133
7.3 光盘驱动器	134
7.4 DVD 简介	139
思考与练习	143
第 8 章 多媒体网络技术	144
8.1 多媒体通信网络的系统组成及特点	144
8.2 音频和视频信息处理的网络需求	146
8.3 多媒体通信网	150
8.4 Internet 工作原理	155
8.5 多媒体网络应用	160
思考与练习	170
第 9 章 音频的处理与制作	171
9.1 实验 1 声音的录制及文件格式转换	171
9.2 实验 2 声音的编辑	175
9.3 实验 3 MIDI 音乐的制作	182
第 10 章 视频的编辑和制作	189
10.1 实验 1 视频的截取与生成	189
10.2 实验 2 效果处理	194
第 11 章 二维动画的编辑与处理	200
11.1 实验 1 字母变形动画	200
11.2 实验 2 探照灯效果	204
11.3 实验 3 路径动画制作	210

第 12 章 图像的编辑与处理	216
12.1 实验 1 立体几何体 1	216
12.2 实验 2 立体几何体 2	231
12.3 实验 3 兽皮中国画	236
12.4 实验 4 巧克力按钮的制作	243
第 13 章 多媒体编辑与创作工具	253
13.1 实验 1 入门小实例	253
13.2 实验 2 建立用户交互	258
13.3 实验 3 使用判断和计算图标	262
13.4 实验 4 使用导航和框架图标	264
13.5 实验 5 使用变量、函数和表达式	268
13.6 实验 6 综合实验	278
第 14 章 多媒体网络课件的制作流程与规范	280
14.1 多媒体网络课件开发流程与人员组织	280
14.2 多媒体网络课件脚本设计与规范	282

第1章 基本概念和基础知识

本章要点:

- 数据、信息、媒体及多媒体的概念
- 媒体分类、特点及媒体元素的概念
- 多媒体技术及其他相关知识
- 多媒体技术的应用
- 多媒体技术的由来和发展

计算机诞生的初衷主要是为了科学计算,处理的主要是数字,后来也用来处理文字,从而实现了数字和文本处理的计算机化,给人们提供了不少方便,大大促进了计算机技术的发展,拓展了计算机的应用范围。但是,仅数字和文本方式的交互与人的自然交互形式相距甚远。因为人体本身就是一个绝妙的多种媒体信息集约系统,人类通过视觉、听觉、触觉、嗅觉和味觉来感知外部世界,获取有关图像、图形、文本、数据、声音等,再通过大脑神经网络处理后,返传到身体的相应部位,从而获取不同形式的信息。在人的感知系统中,视觉所获取的信息占60%以上,听觉获取的信息占20%左右,其余的大都通过触觉、嗅觉、味觉、手势甚至情感等方式获取。虽然在数字和文本处理计算机化的环境下,人类靠自身的感官系统利用数字、文本形式也能表达信息内容,但直观性差,不能听其声、见其物,而且有的难以准确、清楚地表达。

经过十几年的不断摸索和研究,自20世纪80年代中后期开始,多媒体计算机技术成为人们关注的热点之一。多媒体技术的出现从根本上改变了昔日基于字符的各种计算机处理。首先,是语音和图像的实时获取、传输及存储,使人们获取和交互信息流的渠道豁然开朗,可听其声,观其色,千里之外,近在咫尺,在一定程度上改变了人们的交互方式、生活方式和工作方式。给人类的工作、生活带来了一场革命。正如专家们预测的计算机发展方向是MODN:M为多媒体技术(Multimedia),O为开放系统(Open System),D为小型化(Downsizing),N为计算机网络技术(Network Computing)。其中多媒体技术位于第一,“交互性多媒体技术已经成为一股不可阻挡的潮流”。

本章主要介绍关于多媒体技术的一些基本概念和基础知识。

1.1 多媒体的基本概念

近年来,随着计算机技术的飞速发展,多媒体技术得到了广泛应用,人们对媒体概念

的认识进一步加深,已初步形成共识。

1.1.1 数据、信息、媒体与多媒体

要弄清什么是多媒体 (Multimedia), 首先要知道什么是媒体 (Media) 以及媒体与数据、媒体与信息的关系。

1. 数据

一般意义的**数据**是指客观实体或系统的原始记录,用以记录和负载信息的,由于客观世界的多样化,使得数据的种类也呈现出多样性。这种数据只有经过加工,并赋予一定意义后,才形成**信息**。

2. 信息

信息是数据经过加工后形成的,其实质是客观实体或系统运动特征及内涵的抽象,以数据的形式存在和运作,是人类管理活动的基础,也是管理过程的归宿和表征。

3. 媒体

所谓**媒体**,在通常意义下是指媒介、传媒、媒质或介质,是中间物质。在计算机领域赋有两种含义:一是指用以存储信息的实体,如磁带、磁盘、光盘和半导体存储器等;另一种是指信息的载体,客观世界中存在着各种各样的信息形式,不同的形式称作不同的**信息媒体**,如数字、文字、声音、图像和图形,多媒体技术中的**媒体**是指后者,是一种形象化了的信息表现,即**信息载体**。

1.1.2 媒体的分类

媒体是信息的载体,是信息的存在形式和表征,由于信息在被人们感知、抽象,作出表示,使之呈现,实现存储或传输过程中采用了多种载体,根据国际电信联盟 (ITU) 电信标准部 (TSS) 的 ITU-TI.347 建议,媒体可分为以下 6 类。

(1) **感觉媒体 (Perception Medium)**: 这是一种使人产生感觉的媒体,它直接作用于人的感官,如眼、耳、鼻、舌、手等。例如,语音、音乐、图形、图像、文字、动画、气味等等,它是人们的感觉器官所能感觉到的信息的自然种类。

(2) **表示媒体 (Representation Medium)**: 说明交换信息的类型,定义信息的特征。一般以编码的形式描述,如声音编码、图像编码和文本编码等。

(3) **呈现媒体 (Presentation Medium)**: 是人们用以获取信息或再现信息的物理手段,输入或输出信息的设备。如显示器、打印机和音箱等输出设备,键盘、鼠标、扫描仪和摄像机等输入设备,都可称之为**呈现媒体**。

(4) **存储媒体 (Storage Medium)**: 存储数据的物理设备,如磁盘、磁带、光盘、内存等。

(5) **传输媒体 (Transmission Medium)**: 传输数据的物理设备,如电缆、光纤、无线电波等。

(6) 交换媒体 (Communication Medium): 在系统之间交换信息的手段和类型, 可以是存储媒体或传输媒体, 也可以是两种媒体的组合。

1.2 多媒体

1.2.1 多媒体的概念

多媒体就是用多种方法、多种形态传输 (传播) 信息的介质、载体。例如, 电影是一个很好的多媒体例子: 在有声电影发明以前, 电影采用文字显示的方法表示人物对话的内容, 后来, 人们把第二种媒质——声音加到电影中, 使我们不但能听到演员的对话而且能听到动听的声音。现代个人计算机程序用到了电影的所有组成部分, 包括活动图像、音响效果、语言、文字、音乐、动画以及静物摄影等等, 譬如, 将活动图像与音响效果混合起来就可做成视频游戏。

多媒体译自英文 Multimedia, 该词是由 Multiple 和 Media 构成的复合词, 与之对应的词是单媒体 Monomedia。对于多媒体并没有一个明确的、统一的定义, 从字面上来看可以这样说: 多媒体是有两种或两种以上媒体的有机集成体, 或者说是由单媒体复合而成的有机集成体。但是, 这种字面解释与科学技术领域“多媒体”术语的含义有着很大的区别, ITU 对多媒体含义的表述是: 使用计算机交互式综合技术和数字通信网技术处理多媒体文本、图形、图像和声音, 使多种信息建立逻辑连接, 集成为一个交互系统。在此, 可以看出, 多媒体本身是计算机技术与视频、音频和通信等技术的集成产物, 是把文字、音频、视频、图形、图像和动画等多媒体信息通过计算机进行数字化采集、获取、压缩/解压缩、编辑、存储等加工处理后, 再以单独或合成形式表现出来的。从使用者的角度看, 多媒体是一个丰富多彩的感官世界, 它能使人的眼睛、耳朵、手指, 特别是大脑兴奋起来。随着科学技术的不断进步, 多媒体的含义和范围还将进一步扩展。

1.2.2 信息媒体的基本特性

多媒体的基本特性主要有信息媒体的多样性、交互性和集成性等几个方面。

1. 信息媒体的多样性

信息媒体的多样性是多媒体的主要特性之一。多媒体扩展和放大了计算机处理的信息空间和种类, 不再局限于数值和文本, 而是广泛采用图像、图形、视频和音频等信息形式来表征内涵和反映思想。使计算机表达人类的思维不再局限于线性的、单调的、狭小的范围内, 而有了更充分、更自由的余地, 即计算机变得更加人性化。在人类的日常生活中, 接触最频繁的信息就是眼睛看到的图像和耳朵听到的声音。但对于应用而言, 声像信号的输入 (获取) 与输出 (表现) 并不一定相同, 如果二者完全一样, 则只能称为记录和重放, 效果显然不是最理想的。如果能对声像信号进行加工、变换, 即通常所说的创作, 就会大

大丰富信息的表现力并增加表现效果。多媒体可使计算机处理的信息多样化或称多维化,使之在信息交互过程中有更加广阔和更加自由的空间,满足人类感官全方位的信息需求。

2. 信息媒体的交互性

多媒体的第二个重要特性是交互性。交互性是指向用户提供更加有效的控制和使用信息的手段,交互可以增加对信息的注意和理解,延长信息保留的时间。打开电视机,会显示图像、声音和文字,由于观众只能被动地收看,因此,人与电视节目之间的关系是非交互式的。交互式工作是计算机固有的特点(从存储单元调出一个文件修改后再存入存储单元,随意访问,这便是交互式工作),但是,在引入多媒体概念之前,人机对话只在单一的文本空间中进行,这种交互的效果和作用十分有限,只能“使用”信息,很难做到自由控制和干预信息的处理。而多媒体的交互性是指人们可以使用键盘、鼠标、触摸屏、声音或数据手套等设备,通过计算机程序来控制各种媒体的播放。人与计算机之间,人通过操纵计算机而驾驭多媒体,人是主动者,而多媒体是被驾驭的对象。

当多媒体的交互性引入后,人处于参与、开发的活动地位,活动(Activity)本身就介入了信息转变为知识的过程。人们借助于活动可以获得更多的信息,可以改变信息的组织过程,获得许多奇特的理想效果。

交互性被赋予多媒体信息空间后,就带来了巨大作用。从数据库中检索出某人的照片、声音及其文字材料,只是多媒体交互性的初级应用;通过交互特征使用户介入到信息过程中(不仅仅是提取信息),则为应用的中级阶段;当我们完全进入到一个与信息环境一体化的虚拟信息空间遨游时,才能达到交互应用的高级阶段。这就是虚拟现实(Virtual Reality, VR),也是当今多媒体研究的重要内容之一。

3. 信息媒体的集成性

多媒体中的集成性是信息系统构架层次的一次飞跃,这是媒体信息的又一重要特性。这种集成性主要表现在两个方面,即多种信息媒体的集成和处理这些媒体设备的集成。对前者而言,各种信息媒体应该成为一个有机体,而不应是分散游离的,尽管可能是多通道的输入或输出,这种集成也应包括信息的多通道统一获取,多媒体信息的统一存储与组织,多媒体信息的合成等各方面。总之,不应再像早期那样,只使用唯一的形态进行获取和理解信息,而应更加看重媒体之间的关联及其所蕴涵的结构信息。另外,多媒体的各种设备应该成为一个统一体,亦即从硬件来说,应该具有能够处理多媒体信息的高速及并行的CPU系统,大容量的存储、适合多媒体多通道的输入输出外设,以及宽带的通信网络接口;对于软件来说,应该有集成一体化的多媒体操作系统,适合于多媒体信息管理和使用的软件系统及创作工具,高效的各类应用软件等。这些还要在网络的支持下,集成构造出支持各类信息的应用系统。

4. 信息媒体的协同性

每一种媒体都有其自身的规律,各种媒体之间必须有机配合、协调一致。多媒体之间的协调,以及在时间和空间上的一致性,称为信息媒体的协同性。

5. 信息媒体的实时性

信息媒体的实时性就是在人的感官系统允许的情况下，进行即时媒体交互，就好像面对面（Face to Face）一样，图像和声音等各种交互媒体信息都表现得很连续，也很逼真。

6. 信息媒体的海量数据

信息媒体的海量数据是媒体的又一个重要特性，现以声音和视频图像为例加以说明。对声音数据进行采样、量化时，通常采用 44.1kHz 的采样频率，为了达到较大的动态范围和信噪比，每一样本需用 16 位二进制数表示，这样，对一路双声道立体声而言，信息量为每秒 178KB 或每分钟 10.6MB。对于视频图像同样需要采样量化即数字化处理，所形成的数据量也是非常庞大的。信息媒体的海量数据特点导致一系列的存储、传输等方面的技术难点需要解决。

1.2.3 媒体元素

1. 文字

文字是组成计算机文本（Text）文件的基本元素。纯文字的文本文件常有 .txt 扩展名，而 .doc 则是 Word 所采用的加入了排版命令的特殊文本文件。

文字用编码的方式在计算机内存储和交换。英文字符常用的编码是 ASCII 码，占 7 位，扩展后占 1 字节（8 位）。例如，英文字母 A 的 ASCII 编码为 01000001。汉字是一种象形文字，由于字的种类较多，常用的也有近万个（全部汉字有 6 万~8 万个），所以常用 2 字节的编码来表示一个汉字。例如，汉字“啊”的中国国标 GB—2312 编码为 101100001 0100001。文字的存储和传输就是使用它的编码，所以存储和传输的数据量还相对较小。但为了在屏幕上显示或使用打印机打印汉字字符，还需要建立字模库。字模库中所存放的是字符的形状信息。它可以用平面二进制位图（BitMap，BMP）即点阵方式表示，也可以用“矢量”方式表示。位图最典型的是用 1 来表示有笔划经过，0 表示空白，如图 1-1 所示。

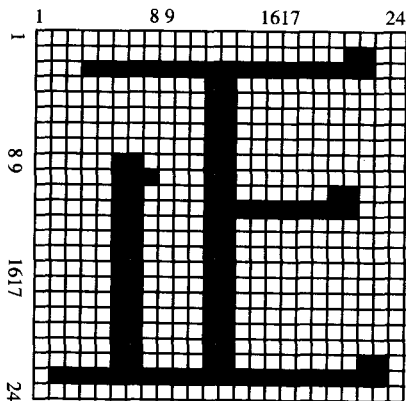


图 1-1 24×24 点阵汉字示例

位图方式占存储量相当大,例如,采用 64×64 点阵来表示一个汉字(其精度基本上可以提供给激光打印机输出),一个汉字占 $64 \times 64 \times 8 = 512 \text{ (B)} = 0.5 \text{ (KB)}$;一种字体(例如,宋体)的一、二级国标汉字(6 763 个)所占的存储量为 $0.5 \text{ KB} \times 6763 = 3382 \text{ KB}$,接近 3.3 MB 。汉字最常用的字体有宋体、仿宋体、楷体和黑体(又称等线体)4种,此外,隶书、魏碑、综艺等字体也比较常用。由于字体多,字模库所占的硬盘空间也是相当大的。

矢量表示法是通过字的点、线等特征的描述来进行表示,存储量较小,且字形可以随意放大而不产生“锯齿”形失真。

目前,将文字输入计算机的方法主要有4种:键盘输入、OCR输入、手写输入和语音输入。

键盘输入时,通常使用普通的英文键盘,而输入法则主要分为拼形和拼音两大类。拼音输入法相对简单,会汉语拼音就能输入,但输入速度较慢,最高只能达到 $20 \sim 30$ 个汉字每分钟;拼形输入法(例如“五笔字型输入法”)要会拆字,并熟背与汉字部件的对应键,这需要一段时间的训练,但由于重复字少,可以实现盲打,输入速度可达 $50 \sim 100$ 个汉字每分钟。

OCR (Optical Character Recognition/Reader, 光学字符识别或阅读器)汉字识别输入,主要用于对大量印刷体文字的输入。印刷体文字,经图文扫描器按图形模式进入计算机,经软件识别出文字,再用相应汉字的编码进行表示。目前 OCR 已可同时识别各种字体字号的汉字,且识别率可达 98% 以上。

联机手写识别输入近年发展极快,在一块巴掌大的手写板上用专用笔或用手指写字,计算机能感受出手写的笔划和笔顺,从而识别出文字,甚至对连笔字也能准确识别。

语音输入也已从单字、单词输入发展到连续语音的输入,但作为文字输入的技术,其准确率还不够理想。语音输入往往与讲话人的发音、音调有关,且需要专人对计算机进行反复的“训练”。

2. 声音

声音(Audio)是人们用来传递信息最方便、最熟悉的方式之一,它是携带信息的极其重要的媒体。声音是通过一定介质(如:空气、水等)传播的一种连续的机械波,在物理学上称为声波。声音的强弱体现在声波压力的大小上(和振幅相关),声调的高低体现在声波的频率上。频率在 $20 \sim 20 \text{ kHz}$ 的波,称为音频波;频率小于 20 Hz 的波,称为亚音波;频率大于 20 kHz 的波,称为超音波。我们说话时产生的声音波的频率范围约为 $300 \sim 3 \text{ 000 Hz}$,英文通常用 Speech、Voice 等词来表示。音乐波的频率范围可达到 $10 \sim 20 \text{ 000 Hz}$,英文用 High-Fidelity Audio 来表示,一般就用 Audio 表示。

我们经常谈论的声音质量通常用声音信号的范围来衡量。现在公认的声音质量分为以下4级。

- (1) 数字激光唱盘 (Compact Disc-Digital Audio, CD-DA) 质量。
- (2) 调频无线电广播 (Frequency Modulation, FM) 质量。
- (3) 调幅无线电广播 (Amplitude Modulation, AM) 质量。

(4) 电话 (Telephone) 质量。
它们的频率范围如图 1-2 所示。

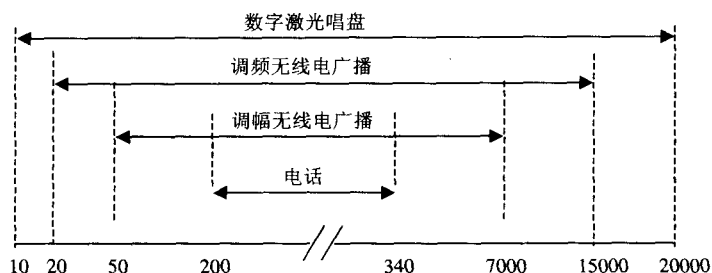


图 1-2 常见声音频率范围示意

从图中可以看到，数字激光唱盘的声音质量最高，电话的声音质量最低。低于 200Hz 的低频信号可以用来增强话音的自然度和说话风度。高于 7 000Hz 的高频信号用来提高话音的易懂度和区分能力。

声音的频率范围为 20~20kHz，如 1.2.2 节所述，在对它进行数字化转换时，采样频率不应低于 40kHz，在多媒体技术中常用的标准采样频率为 44.1kHz。量化精度现在常用 16 位，质量更高的也有用 24 位的。

为了取得立体声音响效果，有时需要进行“多声道”录音，最起码有左右两个声道，较好的则采用 5.1 或 7.1 声道的环绕立体声。所谓 5.1 声道是指：含左、中、右、左环绕、右环绕 5 个有方向性的声道，以及 1 个无方向性的低频加强声道，如图 1-3 所示。MPEG-2 (运动图像专家组) 数字影视标准和杜比 AC-3 (声音编码 3 型) 标准都采用 5.1 声道的环绕立体声。

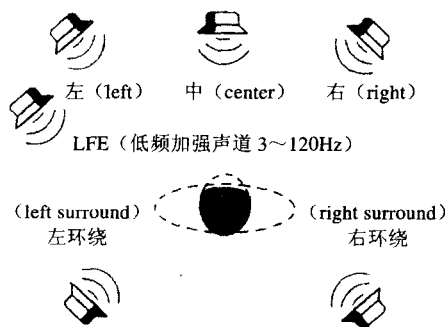


图 1-3 5.1 声道示意

采样频率越高，量化精度越高，声道数越多，则声音质量就越好，而数字化后的数据量也就越大。例如，在采用 44.1kHz 采样，精度为 16 位 (即 2 字节)，左右两个声道的情况下，每秒声音所占数据量为：

$$44.1\text{k} \times 2 \times 2 = 176.4 \text{ (Kb/s)}$$

1s 的声音就为 176KB 容量，一张软盘只够存储 8s 的声音，这对存储和传输的负担都