

21 世纪高等学校规划教材

多媒体技术及应用

林筑英 林建勤 欧晓鸥 等 编著



中国电力出版社

<http://jc.cepp.com.cn>

中国美术学院美术考级教材

多媒体技术及应用

中国美术学院美术考级教材



中国美术学院美术考级教材

7P37/266

21 世纪高等学校规划教材

多媒体技术及应用

林筑英 林建勤 欧晓鸥 等 编著



中国电力出版社

<http://jc.cepp.com.cn>

内容提要

本书为 21 世纪高等学校规划教材之一。本书较系统地介绍了多媒体技术及应用的基本概念、基本原理和基本方法,详细介绍了多媒体信息获取与处理、网络多媒体技术、多媒体应用系统设计与开发等知识。主要内容分为 5 章,包括多媒体技术基础、多媒体计算机系统环境及使用、多媒体信息的获取与处理、网络多媒体技术及应用、多媒体应用系统设计。本书在传授基本理论知识的同时,强调综合应用技能和实际动手能力的培养。

本书可作为高等院校本科各专业或研究生相关课程的教材,也适合作为教育培训的多媒体应用教材,以及广大喜爱开发多媒体应用软件、制作多媒体网页等读者的自学或参考用书。

图书在版编目(CIP)数据

多媒体技术及应用 / 林筑英等编著. —北京: 中国电力出版社, 2008

21 世纪高等学校规划教材

ISBN 978-7-5083-7173-3

I. 多… II. 林… 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 (2008) 第 134426 号

丛 书 名: 21 世纪高等学校规划教材

书 名: 多媒体技术及应用

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) 68362602

传 真: (010) 68316497, 88383619

服务电话: (010) 58383411

传 真: (010) 58383267

E-mail: infopower@cepp.com.cn

印 刷: 北京丰源印刷厂

开本尺寸: 184mm×260mm

印 张: 18.25

字 数: 412 千字

书 号: ISBN 978-7-5083-7173-3

版 次: 2008 年 9 月北京第 1 版

印 次: 2008 年 9 月第 1 次印刷

印 数: 0001—3000 册

定 价: 28.00 元

敬告读者

本书封面贴有防伪标签, 加热后中心图案消失
本书如有印装质量问题, 我社发行部负责退换

版 权 专 有 翻 印 必 究

前言

进入 21 世纪以来,以多媒体技术和网络技术为核心的信息技术飞速发展,推动着各行各业发生深刻变革。随着多媒体技术和网络技术的广泛应用,信息技术的数字化、多媒体化、智能化和信息网络的全球化成为信息社会的主要特征,人们的生产、生活方式也发生了巨大变化。网络多媒体技术在互联网络、教育培训、信息管理、视频会议、影视娱乐、视频点播、电子出版、休闲旅游、商业广告、过程模拟等各个领域发挥着重要的作用。

围绕高速发展的多媒体和网络技术,本书较系统地介绍了多媒体技术及应用的基本概念、基本原理和方法,详细介绍了多媒体信息获取与处理、网络多媒体技术、多媒体应用系统设计与开发等知识。目的是使读者了解多媒体信息表示、获取、存储、处理的基本原理及主要技术,了解多媒体输入输出设备的基本组成及工作原理,理解网络多媒体技术的基本原理及流媒体技术,理解多媒体应用系统的设计方法与开发技术。本书主要内容包括多媒体技术基础、多媒体计算机系统环境与使用、多媒体信息的获取与处理、网络多媒体技术及应用、多媒体应用系统设计。

本书有针对性地以任务驱动、由浅入深、由易到难、循序渐进、图文并茂的风格编写,在注重理论体系的同时,强调突出综合应用技能、实践能力和创新能力的培养。书中以当前最新的音频、视频、图像、动画等多媒体创作软件的实践应用及丰富的制作实例,将重要的知识点融于实例当中,融合了学科间的交叉与渗透,充分反映了多媒体应用的新技术、新方法。

本书主要由林筑英教授编写,林建勤、欧晓鸥、张琦、林宏伟、唐翠芳、杨源参加了本书部分章节的编写工作。其中第 1 章由林筑英编写,第 2 章由林宏伟和林筑英编写,第 3 章由唐翠芳、张琦和林筑英编写,第 4 章由林宏伟、杨源和欧晓鸥编写,第 5 章由林建勤、张琦编写,全书由林筑英审阅并最后做了必要的修改,硕士研究生金贻同学参与了全书有关表格和图形的绘制工作。

另外书中还参考和引用了一些国内外的论著、论文和研究成果,谨在此一并表示衷心的感谢。感谢中国电力出版社在本书的撰写和出版过程中给予的大力支持。

由于作者水平有限,书中不当之处在所难免,敬请同行专家和广大读者批评指正。

作者

2008 年 6 月

目 录

前 言	
第 1 章 多媒体技术基础	1
1.1 多媒体概述	1
1.2 多媒体技术的基本概念	5
1.3 多媒体的关键技术	8
1.4 多媒体技术的发展与应用	27
1.5 多媒体技术的发展趋势及主要研究方向	41
练习与思考题	52
第 2 章 多媒体计算机系统环境与使用	54
2.1 多媒体个人计算机	54
2.2 多媒体系统组成	59
2.3 多媒体输入输出设备	69
练习与思考题	92
第 3 章 多媒体信息的获取与处理	94
3.1 图像信息的获取与处理	94
3.2 图像处理软件 Photoshop	102
3.3 音频信息的获取与处理	121
3.4 视频信息的获取与处理	132
3.5 动画制作原理及其制作过程	145
3.6 动画制作软件 Flash	151
练习与思考题	172
第 4 章 网络多媒体技术及应用	174
4.1 多媒体网络通信	174
4.2 流媒体技术	204
4.3 多媒体网页制作	216
练习与思考题	238
第 5 章 多媒体应用系统设计	239
5.1 多媒体应用系统设计概述	239
5.2 多媒体应用系统开发过程	247
5.3 多媒体应用系统开发工具	257
5.4 多媒体应用系统开发实例	266
练习与思考题	282
参考文献	284

第 1 章 多媒体技术基础

通过本章学习，可以了解或掌握：

- 多媒体与多媒体技术的基本概念
- 多媒体元素组成与分类
- 多媒体的关键技术
- 多媒体技术的应用和发展方向

自 20 世纪 90 年代以来，多媒体技术的迅速兴起与蓬勃发展，带来了计算机世界的一场革命。有人说它是继印刷术、无线电、电视技术、计算机技术等之后的又一个新的技术革命，是信息处理技术的第四次革命性的飞跃。多媒体计算机的出现，提高了计算机的应用水平，扩大了计算机技术的应用领域，使计算机除了能够处理文字信息外，还能处理图像、声音、视频和动画等信息；同时也促使多媒体技术渗透到人们生活的各个领域，在信息社会中处于十分重要的地位。多媒体技术的应用将对人类的生产方式、工作方式及生活方式带来巨大的变革，将使人类社会进入到一个前所未有的新时代。

1.1 多媒体概述

1.1.1 多媒体定义

1. 多媒体

多媒体中的“媒体”一词，包括了两方面的含义，一是存储和传输信息的实体，如磁带、磁盘、光盘、半导体存储器等；二是信息表示和传播的载体，如数字、文字、声音、图形、图像、动画等，多媒体技术中的媒体通常是指后者。随着计算机技术与通信技术的发展，使人们有能力把这些媒体信息在计算机内均以数字形式来表示，并综合起来形成一种全新的媒体概念——多媒体（Multimedia）。

多媒体是由两种以上单一媒体融合而成的信息综合表现形式，是多种媒体的综合、处理和利用的结果。通过不同形式的“媒体”，反映了不同的信息表示与信息交流方式；而多媒体的“多”，在强调信息媒体多样性的同时，更强调各媒体间的有机结合及人与信息媒体之间的交互作用，具体表现为多种媒体表现、多种感官作用、多种设备支持、多个学科交叉、多个领域应用等。因此，多媒体是建立在一定信息处理技术之上的、融合两种以上媒体的一种人机交互式信息媒体或系统。

多媒体的实质是将不同表现形式的媒体信息数字化并集成，通过逻辑链接形成有机整体，同时实现交互控制，所以数字化与交互集成是多媒体的精髓。虽然“多媒体”这个词

语在新闻传媒、文化娱乐等领域中由来已久，并发挥着重要的作用，但是在计算机领域，多媒体与人们经常接触到的传媒主要有三点不同：

- (1) 传统的传媒基本是模拟信号，而多媒体信息是数字化的。
- (2) 传统的传媒只能让人们被动接受信息，而多媒体可让人们主动与信息媒体进行交互。
- (3) 传统的传媒一般是单一形式，而多媒体是两种以上不同媒体信息的有机集成。

例如传统的电视节目为什么不是多媒体，就因为它是基于模拟信号处理并线性播放的，人们只能被动收看，不能操作和控制播放内容及播放过程，缺乏交互能力。

进入 21 世纪，以多媒体技术和网络技术为核心的信息技术飞速发展，并以惊人的速度进入社会的各个领域，推动着各行各业发生深刻变革。在信息化程度不断提高和科学技术日新月异的今天，“数字化”、“网络化”与“多媒体”相结合的现代信息处理技术，给“多媒体”概念更注入了新的含义和活力，使得现代人对“多媒体”津津乐道，并且习惯用新的观念去谈论它。

2. 多媒体的信息表达元素

多媒体的信息表达元素主要有文本、图形和图像、音频、视频、动画。部分多媒体信息如图 1-1 所示。

(1) 文本 (Text)。文本是指以文字和各种专用符号表达的信息形式，包括各种文字和各种字体、大小尺寸、格式及色彩的文本。人类使用文字来传情达意、记录文明，文本数据可在计算机的任意一个文本编辑软件里输入和制作，也可以使用扫描仪扫描获得相应的文字。在多媒体中，文字一直被认为是最基本、最重要的成分，在现实生活中，文本也是使用最多的一种信息存储传递方式。

(2) 图形 (Graphic) 和图像 (Image)。图形和图像都是多媒体中的可视元素。图形是指从点、线、面到三维空间的黑白或彩色几何图形，也称为矢量图形 (Vector Graphic)。图像是由称为像素 (Pixel) 的点构成的矩阵图，也称为位图 (Bitmap)。

矢量图形是通过一组指令集来描述的。这些指令描述构成一幅图形的所有直线、圆、圆弧、矩形、曲线等几何元素的位置、维数、色彩、大小和形状。矢量图形主要用于线型的图画、美术字、统计图和工程制图等。它占据存储空间较小，但不适于表现复杂的、色彩逼真的图画。用来生成矢量图形的软件通常称为绘图 (Draw) 软件。

位图是由描述图像 (Image) 中各个像素点的强度与颜色的数位集合组成的，即用二进制位 (bit) 来定义图中每个像素的颜色、亮度和属性。位图适合表现比较细致、层次和色彩比较丰富、包含大量细节的图像，如照片和图画等。位图的特点是显示速度快，色彩较逼真但占用的存储空间较大。

(3) 音频 (Audio)。音频是指在 20Hz~20kHz 频率范围的连续变化的声波信号，可分为语音、音乐与合成音效三种形式，多媒体中的音频包括高保真的音乐声、语音声、各种

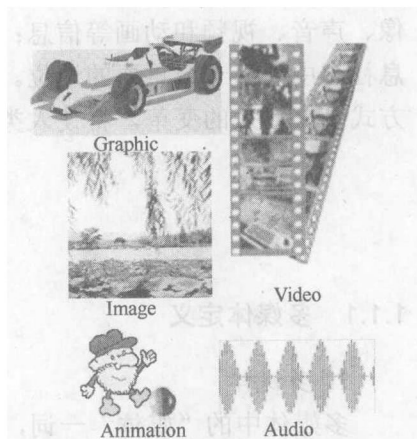


图 1-1 图像、视频和声音媒体

效果声和背景声等。多媒体计算机中只有经过数字化后的声音才能播放和处理,数字化的音频文件有多种格式,常见的有波形(WAV)音频、乐器数字接口(MIDI)音频、光盘数字(CD-DA)音频等。

(4) 视频(Video)。视频是指从摄像机、录像机、影碟机等影像设备得到的连续活动图像信号,由若干有联系的真实图像数据连续播放形成,其类似于电影和电视,有声有色,有动作有情感。多媒体计算机上的视频是数字化视频,可更方便地进行存储、重放和各种特殊效果的处理。在多媒体信息表达元素中,视频是最新和最具魅力的一种,可更有效地表达内容及表现主题。

(5) 动画(Animation)。动画是借助计算机生成的一系列连续运动的图画。用多媒体计算机实现的动画有两种,一种叫造型动画(Cast-Based Animation),另一种叫帧动画(Frame Animation)。帧动画是由一幅幅连续的画面组成的图形或图像序列。与视频影像不同的是,动画是由计算机绘制生成的非真实图形或图像。造型动画则是对每一个活动的对象(称为动元或角色)分别进行设计,赋予每一个对象一些特征(如形状、大小、颜色等),然后用这些对象组成完整的帧画面。动画可增强视觉效果,起到强调主题、增加生动情趣的作用。

1.1.2 媒体种类和特点

1. 人类感知媒体的分类

在人类社会,信息的表现形式是多种多样的,一般把这些表现形式称为媒体(medium)。媒体是承载信息的载体,是信息的表现形式。

人类利用视觉、听觉、触觉、嗅觉和味觉来感知各种信息或感受各种媒体。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,其余部分有触觉、嗅觉、味觉、脸部表情、手势等。人类感知媒体可分为视觉类媒体、听觉类媒体、触觉类媒体、嗅觉类媒体和味觉类媒体。

视觉类媒体是通过视觉来感知和传递信息的,包括文本、图形、图像、动画、视频等;听觉类媒体是通过听觉来感知和传递信息的,包括波形声音、语音和音乐等;触觉类媒体是与环境进行交互的媒体,例如人的皮肤可感觉环境的温度、湿度、振动、压力等。多媒体技术已逐步将触觉类媒体引入到实际系统中,使人与环境的信息交流更充分。在触觉类媒体中,触摸、指点是最常见的形式,通过指点可以确定媒体对象的位置、大小、方向和方位,执行特定的过程和相应的操作。目前在计算机中还难以实现触觉类媒体、嗅觉类媒体和味觉类媒体,有待于虚拟现实技术的进一步研究和发展。直至今日,计算机以及与之相类似的设备还远远没有达到人类感官的水平,在信息交互方面与人的感官空间还相差甚远。多媒体技术就是要把计算机处理的信息多维化,通过信息的捕获、处理与展现,使之在交互过程中具有更加广阔和更加自由的空间,满足人类感官空间全方位的多媒体信息需求。

2. 媒体的五种类型

这里所指的媒体是表示和传输信息的载体。可分为感觉媒体、表示媒体、显示媒体、存储媒体和传输媒体五种类型,这些媒体类型与计算机系统之间存在如图1-2所示的概念关系。

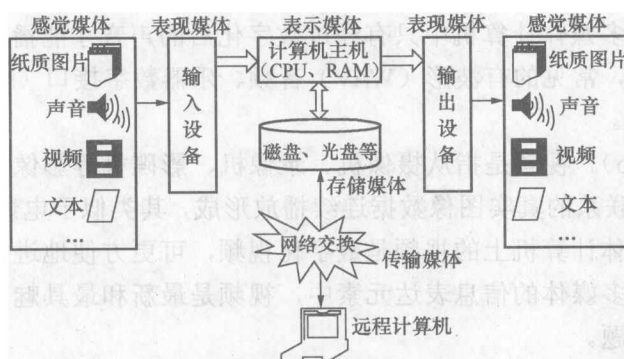


图 1-2 媒体类型与计算机系统

感觉媒体 (Perception): 是直接作用于人的感觉器官, 使人能直接产生感觉的媒体。包括人类的语言、音乐、自然界的各种声音、静止图像、图形、活动图像、动画和文本等。

表示媒体 (Presentation): 是为传输感觉媒体而研究出来的中间媒体, 借助此媒体以便能更有效地将感觉媒体从一处传向另一处。它包括上述感觉媒体的各种编码, 例如语言编码、图像编码、电报码、条形码等。

显示媒体 (Display): 是用于在通信中使电信号和感觉媒体之间产生转换所用的媒体。有两种显示媒体: 输入显示媒体和输出显示媒体。前者包括键盘、鼠标、摄像机、扫描仪、光笔和话筒等, 后者包括显示器、扬声器和打印机等。

存储媒体 (Storage): 是用于存储表示媒体的介质。它提供机器随时调用和终端远距离调用的可能性。存储介质有硬盘、软盘、光盘、磁带和半导体存储器等。

传输媒体 (Transmission): 是将表示媒体从一点传输到另一点的物理实体。包括各种导线 (如双绞线)、各种电缆 (如同轴电缆)、光纤、无线电波、红外线等。

归纳起来, 五种媒体的作用、表现形式等特点如表 1-1 所示。

表 1-1 媒体种类及特点

媒体类型	作用	表现形式	内容
感觉媒体	人的感知客观环境	视觉、听觉……	文本、图形、图像、视频、动画、声音等
表示媒体	定义信息的表达特征	计算机数据格式	图像编码、声音编码、ASCII 码、条形码等
显示媒体	表达信息	输入、输出信息	键盘、鼠标、摄像机、扫描仪、光笔、话筒、显示器、扬声器、打印机等
存储媒体	存储信息	保存与读取信息	硬盘、软盘、光盘、磁带和半导体存储器等
传输媒体	数据信息的传输	信息传输的介质	电缆、光缆、无线电波、红外线等

3. 媒体的特性

从信息表达的角度来看, 媒体具有以下特性:

(1) 有格式的数据才能表达信息的含义。由于媒体的种类不同,它们所具有的格式也有所不同,只有对该格式能够理解和解释,才能对其内容即承载的信息进行表达,这种格式主要是指媒体种类的识别和解释。

(2) 不同的媒体所表达信息的程度不同。由于每种媒体都有自己本身的承载信息的形式特征,人们对不同信息的接受程度有差异,这种差异有时表现为信息量的不同。一般说来,越是接近人类原始表达的信息越丰富,越是抽象化(即远离原始信息表达)的信息,其信息量越少,但越精确。

(3) 媒体之间的关系也代表着信息。媒体的多样化关键不在于能否接收多种媒体的信息,而在于媒体之间的信息表示的合成效果。由于多种媒体来源于多个感觉通道,以不同的形式表达,具有一种“感觉相乘”的效应,所以将远远超出各个媒体单独表达时的效果。

(4) 媒体可以进行相互转换。媒体转换是指媒体从一种形式转换为另一种形式。一般说来,媒体的转换总是要损失一定的信息,但损失的信息对接收者来说是否重要,将取决于具体的应用及应用要求。也有一些媒体目前尚不能转换,尤其是不能直接转换。

可以说,多媒体是计算机技术与各种媒体信息处理相结合的产物,是融合两种或两种以上媒体的一种人机交互式信息交流和传播媒体。“多媒体”概念的提出,是基于人类已经有了要将多种媒体信息进行综合处理的要求,并随着科学技术的发展,人类也逐步具有处理多种媒体信息的能力,从而使多媒体成为一种现实。事实上,现在所说的“多媒体”是指能够同时获取、处理、编辑、存储和展示多种媒体信息的技术,即主要是指怎样综合处理和应用多种媒体信息的所有一整套的技术和方法,而不再是指多种媒体信息本身。因此,“多媒体”也被视为“多媒体技术”的同义语。

1.2 多媒体技术的基本概念

1.2.1 多媒体技术定义

多媒体技术起源于计算机数据处理、通信、大众传媒等技术的发展与融合,为了实现多种媒体信息的综合处理,计算机厂家试图将视听节目的处理能力扩展到电脑产品;而家用电器制造商则希望利用计算机、激光等新技术更新家电产品(如电视机)的功能和性能;通信产品制造商更是为此不断研发能支持多种媒体信息传输的通信网络。可见,多媒体技术是以计算机技术为主体,结合通信技术、微电子技术、激光技术、广播电视技术等多种技术而形成的用来综合处理多种媒体信息的、交互性的、数字化的信息处理技术。

由于多媒体技术是一些新技术融合的产物,还在不断地发展和完善之中,因此目前还没有一个严格且完善的定义。G. R. Wichman (1991) 曾经说:“多媒体是一个没有定义的定义”,“你问 10 个不同的人有关多媒体的定义,你将得到至少 10 种不同的回答。”因此,关于多媒体技术的定义,大体上有如下一些描述:

多媒体技术是基于计算机技术的综合技术,它包括数字信号处理技术、音频和视频技术、计算机硬件和软件技术、人工智能和模式识别技术、通信和图像技术。它是正处于发

展过程中的一门综合性的高新技术。

多媒体技术是传统的计算媒体（如文字、图形、图像、动画等）与视频、音频相结合，且为了知识创造和表达交互式应用的结合体。多媒体技术是把文字、图像、图形、动画、音频、视频等多种媒体信息，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，再以单独或合成形式表现出来的一体化技术。

多媒体技术是以计算机（或微处理芯片）为中心，把数字、文字、图形、图像、声音、动画、视频等不同媒体形式的信息集成在一起，进行加工处理的交互性综合技术。是能综合处理多种媒体信息，使多种信息建立逻辑联结并集成为一个具有交互性的计算机系统。

多媒体技术提供了三种新的能力，即在通信中结合电视的音像能力、出版发行能力、计算机交互式处理能力。因此多媒体技术是指能够同时获取、处理、编辑、存储和展示两个以上不同类型信息媒体的技术。

多媒体技术的实质就是将各种媒体数字化，然后利用计算机对这些数字化信息进行加工处理，它不仅是各种媒体信息的集成，也是处理各种媒体信息的设备和软件的集成，通过逻辑和物理连接而形成的有机整体，能实现交互控制，并进而集成为一种交互式的多媒体计算机系统。

多媒体技术是一种全数字技术，是一种将数字声音、数字图像、数字电视、计算机图形和通用计算机集成的，一个具有人机交互功能的可编程环境的技术。

总之，多媒体技术使计算机成为能综合处理多种媒体信息，集文字、数字、图像、图形、声音和视频于一体，并进而集成为综合的多媒体系统。多媒体技术赋予传统的计算机技术更高层次的新含义，从狭义上讲，是指人类用计算机等设备交互处理多媒体信息的方法 and 手段（如 I/O、传输、存储、处理等）；从广义上讲，是指包括了对信息处理的所有技术和方法的一个技术领域。多媒体技术把计算机技术、声像技术、出版技术及网络通信技术结合起来，使计算机进入家庭、艺术及社会生活的各个方面。多媒体技术将极大地影响人们的生活及生产方式，成为对人类有重大影响的技术。

1.2.2 多媒体技术基本特性

多媒体技术的基本特性主要包括信息载体的多样性、集成性、交互性及数字化等。

1. 信息载体的多样性

信息载体的多样性是相对计算机而言的，是指信息媒体的多样性或多维化。最早的计算机主要只能处理数值运算，后来逐渐地转向能处理文字信息和辅助进行绘图并发展了三维图形动画技术，一直发展到今天已经可以处理数字视频、音频等多种数字媒体信息。具有多媒体功能的计算机从根本上改变了过去基于字符的计算机信息处理的单一模式，可以综合处理文本、图形、图像、声音、动画和视频等多种信息，呈现出信息载体的多样性，使人们交流或获取信息时既能听其声，又能见其人。同时多媒体信息的多维化也使得计算机更加人性化。

2. 集成性

多媒体技术的集成性是指将多种媒体有机地组织在一起共同表达事物，做到“图、文、

声、像”一体化。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多媒体信息的软、硬件技术的集成。

多种信息媒体的集成主要指多媒体信息的多通道统一获取、统一存储、组织以及表现合成等各方面。其中，多媒体信息的组织和表现合成是采用超文本思想通过超媒体的方式实现的，为人们构造了一种非线性的信息组织结构。

处理多媒体信息的软、硬件技术的集成主要指计算机系统的硬件和软件集成，硬件应将视频、音频、存储等外部设备集成为能处理多媒体信息的高性能计算机，使之具备相应高效的输入/输出处理能力，软件应有集成一体的多媒体操作系统、多媒体信息处理系统、多媒体应用开发与创作工具等。

多媒体的集成性可以说是计算机系统结构的一次飞跃，当以往单一的图像、声音、交互等项技术统一在多媒体计算机系统下时，一方面意味着各项单独的技术已经发展到成熟阶段，另一方面意味着以往各自独立的发展受到一定局限且不能满足应用需要，必须通过集成才能达到更好的应用目标。多媒体技术的集成性将各种媒体信息有机地进行同步，给人们提供了用计算机来表现、传播和处理信息的数字化方法，能用“眼睛看”和“耳朵听”来接收处理信息，集成为一个完整的多媒体信息系统。

3. 交互性

交互性是多媒体技术的关键特征，也是多媒体技术与电视、电影等单向信息提供手段的主要区别。交互性是指人与机器、人与人、机器与机器之间的交互性，或称人机会话。例如在播放多媒体节目时，人工能够干预播放节目的内容和次序，通过人机交互方式来交换信息并完成交互的任务，交互性提供了更加有效的控制和使用信息的手段，实现了人们对信息的主动选择和控制，人们可以方便地按照思维习惯和意愿，主动地选择和接收信息。除了操作上可通过键盘、鼠标、触摸屏等控制外，在媒体综合处理上也希望做到随心所欲，如屏幕上影视图像可以任意定格、缩放，可根据需要配上解说词和文字说明等。

交互性不仅可增加对信息的注意力和理解，延长信息保留时间，使人们获取信息和使用信息的方式由被动变为主动。而且交互活动本身也作为一种媒体加入了信息传递和转换的过程，使人们获得更多的信息。借助于交互性，人们不是被动地接收文字、图形、声音和图像，而是可以主动地随时参与信息的组织过程，进行编辑、检索、提问和回答，甚至可以控制信息的传播过程，研究学习感兴趣的知识。一般的家电产品是不具备这种功能的。随着多媒体技术的发展，触摸屏及手写输入技术使人机交互更加直观、方便，更接近自然，增强了计算机的友好性，缩短了人与计算机之间的距离。虚拟现实的三维交互传感设备（如数据手套、头盔显示器和立体眼镜）的应用，更使人机交互方式发生了革命性的变化。交互性向人们提供了更加有效的控制和使用信息的手段，也为多媒体技术的应用开辟了更加广阔的空间。

4. 数字化

数字化是指各种媒体信息都是以数字化“0”或“1”的二进制形式存储到计算机中进行处理。多媒体信息处理的关键设备是计算机，而能用计算机记录和传播的信息媒体，无论是文字、图像、声音或视频，都要进行数字化。只有把各种媒体数字化，使其在计算机

中存储和传播时都为一系列比特（bit）的排列组合，即数字化的排列组合，才能使这些信息按照一定的结构存储起来，解决多媒体数据类型繁多、数据类型之间差别大的问题，使各种信息融合在统一的计算机平台上，实现人机的交互作用。因此多媒体信息的全数字化是多媒体技术发展的核心所在，也是多媒体技术实现的唯一可行的方法。以全数字化方式加工处理的多媒体信息，具有精度高、定位准确和质量效果好等特点。

正是由于计算机中数字化技术和交互式的处理能力，才使多媒体技术成为可能，才能对多种信息媒体进行统一处理。这就是为什么一般具有声音、图像的电视机、录像机等还谈不上是“多媒体”的原因。多媒体技术中的“多媒体”并不仅指多媒体信息本身，更主要的是强调处理和应用它的整套软、硬件技术。因此“多媒体”也是多媒体技术的同义语。

数字化、集成性、交互性是多媒体技术的主体特性，此外，还具有其他一些特性，例如：

实时性：是指多媒体技术必须要支持实时处理。多媒体技术是多种媒体集成的技术，由于媒体信息中的声音和视频图像都是与时间密切相关的连续媒体，很多场合要求实时处理。例如播放时，声音和视频图像都不能出现停顿现象，在处理信息时要有严格的时序要求和很高的速度要求。另外，声音和视频图像信息的实时压缩、解压缩、传输与同步处理等，在交互操作、编辑、检索、显示等方面也都要求有实时性。在网络上，这一问题会更加突出，对系统结构、媒体同步、多媒体操作及应用服务等都提出相应的实时化要求。

分布性：由于多媒体信息数据的多样性，多媒体应用的开发工作需要各种专业人员的介入，包括计算机专业、文字写作、影视制作、宣传广告以及计算机动画等方面的人才协同工作。由于原始素材往往分布在不同的空间和时间里，因此多媒体系统还应具有分布性特征。

综合性：多媒体系统并不是各种设备的简单组合，而是以计算机为控制中心，处理来自它所连接的各种设备的多媒体数据。以往各项单媒体技术都是单一的应用，例如有的仅有声音而无图像，有的仅有静态图像而无动态视频等。如今的多媒体系统是将各种媒体设备综合集成起来，经过多媒体技术处理，能够发挥多种媒体的综合作用，产生综合的客观效应。随着多媒体技术的发展，这种综合系统效应越来越明显。“ $1+1>2$ ”的系统特性将在多媒体系统中得到充分体现。

1.3 多媒体的关键技术

由于多媒体系统需要将不同的媒体数据表示成统一的结构码流，然后对其进行变换、重组和分析处理，以进行进一步的存储、传送、输出和交互控制。因此，多媒体技术是一门多学科的综合技术，所涉及的技术领域相当广泛，汇集了计算机体系结构，计算机系统软件，视频音频信号的获取、处理、特技及显示输出等技术。从对各种媒体信息的处理、数据的压缩与解压缩，到多媒体硬件体系结构、多媒体软件系统；从开发多媒体应用的制作工具、动画技术、数据库技术，到与网络技术相结合，都体现出多媒体技术涉及许多领域，其关键技术主要集中在音频和视频处理技术、数据压缩技术、专用芯片技术、光盘存储技术、多媒体网络通信技术和多媒体数据库技术等方面。由于这些技术取得了突破性的进展，多媒体技术才得以迅速地发展，成为具有强大的处理声音、文字、图像等媒体信息

能力的高科技技术。多媒体的关键技术可以归纳为：各种媒体信息的处理技术、数据压缩技术、多媒体计算机系统技术、多媒体数据库技术、多媒体网络通信技术、多媒体同步技术、超文本与超媒体技术。

1.3.1 各种媒体信息的处理技术

各种媒体的处理技术除了可以借鉴和继承各个分支学科的原有技术，如超文本与超媒体技术、计算机图形学、动画制作等技术的原理和方法之外，还有音频和视频处理技术。

1. 音频处理技术

多媒体应用中的一种重要媒体是音频，包括语音、音响和音乐。其作用是直接通过说话来表达信息，或制造某种音响效果和气氛，或演奏优美的音乐。多媒体系统的音频技术主要包括音频数字化、音频压缩技术、语音处理、语音合成及语音识别等技术。

(1) 模拟音频信号的数字化。声音是连续变化的，而计算机内的音频必须是以数字形式存储和处理。音频的数字化就是将模拟（连续的）声音波形转化为数字化（离散化）信号，以便利用数字计算机进行处理的过程。在这一处理技术中，主要包括采样和量化。

采样就是把连续的音频信号转化为一系列有限的离散数值（抽样点）。采样频率，即取样频率是指每秒钟取得声音样本的次数（对声音进行采样的次数 / s）。采样频率越高，音频的真实感质量就越好，但需要存储的音频数据量也越大。

量化是将采样的取值空间划分为若干个子区间，在同一子区间内不同的值用此区间内某一确定的数值代替，即量化是用有限的离散数值代替无限的连续模拟量的多对一的映射操作，量化位数（采样精度）即子区间的个数（用二进制来表示有 8 位、16 位）。

音频数字化的质量相应地由采样频率和量化数据位数来决定。由于人耳听觉的频率上限约为 20kHz，为了保证声音不失真，采样频率应在 40kHz 左右。目前经常使用的采样频率有 44.1kHz、22.5kHz 和 11.025kHz 三种。22.05kHz 只能达到 FM 广播的声音品质，44.1kHz 则是理论上的 CD 音质界限。量化数据位数是每个采样点能表示的数据范围，目前经常采用的有 32 位、16 位、8 位三种。量化位数越多，音质越好，但需要的存储数据量也越大。

(2) 语音处理。语音处理主要包括音频压缩、语音合成和语音识别。其中难度最大、最吸引人的技术是语音识别技术，是指让机器通过识别和理解过程把语音信号转变为相应的文本或命令的高技术。较有代表性的语音识别系统如 IBM 推出的 Via Voice，国内如 Dutty++ 语音识别系统。最新的 IBM Via Voice 9.0 识别技术，识别率比传统的语音识别大大提高。有较强的抗噪声能力，对不同口音的人具有超强的口音适应能力，可直接听写到 WPS、Word、Excel、Outlook 等应用软件中，速度达 150~350 字 / min。

① 语音识别基本原理。语音识别过程实际上是一种认识过程。就像人们听语音时，并不把语音和语言的语法结构、语义结构分开一样。计算机识别系统也要利用这方面的知识，并通过有效地描述语法和语义来实现。计算机的语音识别过程与人对语音的识别处理过程基本上是一致的。但让计算机识别人说话的语音是非常困难的（在某种程度上就像一个外语不好的人听外国人讲话一样），它与不同的说话人、不同的说话速度、不同的说话内容以及不同的环境条件有关。即语音信号具有多变性、动态性、瞬时性和连续性等特点，这些

特点给语音识别造成了较大的难度。目前主流的语音识别技术是建立在统计模式识别基本理论上的,其语音识别系统大致分为三部分:

语音特征提取:其目的是从语音波形中提取出随时间变化的语音特征序列。

声学模型与模式匹配(识别算法):声学模型通常通过学习算法来获取语音特征。在识别时将输入的语音特征同声学模型进行匹配与比较,从而得到最佳的识别结果。

语言模型与语言处理:语言模型包括由识别语音命令构成的语法网络或由统计方法构成的语言模型,语言处理可以进行语法、语义分析。

语言模型对大词汇量的语音识别系统特别重要。当分类发生错误时可以根据语言学模型、语法结构、语义学进行判断纠正,对一些同音字则必须通过上下文结构才能确定词义。语言学理论包括语义结构、语法规则、语言的数学描述模型等方面。

②语音识别技术的应用。语音识别技术发展至今,中小词汇量非特定人语音识别系统的识别精度已达98%,而特定人语音识别系统的识别精度则更高,已能满足通常应用的要求。复杂的语音识别系统已完全可以制成专用芯片。语音识别产品已经进入市场和服务领域,例如有些交换机、电话机、手机包含了语音识别拨号功能;语音记事本、语音智能玩具等也包括了语音识与语音合成功能。人们可通过电话网络用语音识别对话系统查询有关的机票、旅游、银行信息,取得较好的效果。语音处理对新闻发布、语音广播等语音业务能提供更加人性化的人机交互方式。随着语音合成技术的不断发展,未来的计算机一定能够像人一样,用自然而流畅的语言表达自己,与人类自由地进行交谈。

③语音合成技术的发展方向。语音合成技术正在走向市场,随着信息社会需求的发展,人机口语对话系统,即语音合成研究已开始从文字到语音的转换阶段向概念到语音的转换阶段发展,这不仅对语音合成技术提出了更高的要求,而且涉及计算机语言的生成及人类大脑的高级神经活动。为了适应社会的需求,语音合成技术将朝着提高合成语音的自然度、丰富合成语音的表现力(进一步实现语气、语调的变化)、降低语音合成技术的复杂度、多语种文语合成等方向发展。

(3) MIDI 音频。MIDI 音频是一种数字音乐接口的国际标准,它定义了电子音乐设备与计算机的通信接口,规定了使用数字编码来描述音乐乐谱的规范。MPC 就是根据 MIDI 文件中存放的对 MIDI 设备的命令,即每个音符的频率、音量、通道号等指示信息进行音乐合成,通过合成器实现的。MID 文件的优点是短小, MIDI 文件本身不包含表征声音特点的波形数据,因此 MIDI 文件比波形采样文件小得多,可以高质量地描述各种声乐品质。MIDI 文件本身记录的是数字化的演奏信息。例如,某个时刻某种乐器的某个键被按下,直到什么时候这个键才松开,以及按下的力度是多少等, MIDI 演奏信息的还原取决于 MIDI 合成器。

2. 视频技术

在多媒体系统中,视频(Video)是由一幅幅单独的画面,也称为帧(Frame)的序列组成,这些画面以一定的速率连续地投射在屏幕上,使观察者具有连续运动的感觉。典型的帧率从 24fps 到 30fps,这样的视频图像看起来是光顺和连续的。通常,伴随着视频图像还有一个或多个音频轨,以提供声音。

视频处理技术包括视频图像信号的获取和视频信号的压缩与存储等主要技术。视频图

像信号主要来自外界的视频设备的电视信号,通过专用的视频卡采集视频信号并把模拟信号进行数字化处理后送入计算机。在多媒体技术中,视频技术包括视频信号数字化和视频编码技术两个方面。

(1) 视频信号数字化。要在多媒体计算机系统中处理视频信息,就必须对不同信号类型、不同标准格式的模拟视频信号进行数字化处理,形成数字视频。

数字视频是以离散的数字信号方式表示、存储、处理和传输的视频信息,所用的存储介质、处理设备以及传输网络都是数字化的。数字视频具有以离散的数字信号形式记录视频信息;用逐行扫描的方式在输出设备(如显示器)上还原图像;用数字化设备编辑处理;通过数字化宽带网络传播等特点。

数字视频信息采用数字化设备进行存储、传输及处理。例如采用数字摄像设备直接拍摄视频画面并通过数字宽带网络传输,使用数字设备(数字电视接收机或模拟电视机顶盒、计算机)接收播放或用数字化设备将视频信息存储在数字存储介质(VCD、DVD 光盘等)上。

视频信号数字化是指在一段时间内,以一定的速度将模拟视频信号经模数(A/D)转换和彩色空间变换转为计算机可显示和处理的数字信号的处理过程。视频数字化使得计算机可以显示和处理视频信号,且视频数字化后的色彩、清晰度及稳定性也有了明显的改善。在模拟视频信号的数字化处理过程中,主要包括采样、量化和编码。

采样:是将连续的视频波形信号变为离散的数字信号。

量化:是将视频图像的幅度信号变为离散值。

编码:将数字化的视频信号进行编码、压缩存储,或将压缩的数字视频解码还原成为电视信号再输出。

通常的视频数字化方法有两种,即复合数字化和分量数字化。其中,复合数字化是指先用一个高速模数(A/D)转换器对全彩色电视信号进行数字化,然后在数字域中分离亮度和色度以获得 YCbCr 分量、YUV 分量或 YIQ 分量,最后再转换成 RGB 色彩模式。分量数字化是指先把复合视频信号中的亮度和色度进行分离,得到 YUV 分量或 YIQ 分量,然后用三个模数(A/D)转换器对三个分量分别进行数字化,最后再转换成 RGB 分量。在这两种视频数字化方法中,分量数字化是采用较多的一种模拟视频数字化方法。

(2) 视频数字化过程。由于视频信号既是空间函数,又是时间函数,而且又采用隔行扫描的显示方式,所以视频信号的数字化过程远比静态图像的数字化过程复杂。首先多媒体计算机系统必须具备连接不同类型的模拟视频信号的能力,可将录像机摄像机、电视机、VCD 机、DVD 机等提供不同视频源接入多媒体计算机系统,然后再进行具体的数字化处理。

如果采用分量采样的数字化方法,则基本的视频数字化过程如图 1-3 所示。

视频数字化过程包括如下步骤:

- ①按分量采样方法采样,得到隔行样本点。
- ②将隔行样本点组合、转换成逐行样本点。
- ③进行样本点的量化。
- ④彩色空间的转换,即将采样得到 YUV 或 YCbCr 信号转换为 RGB 信号。
- ⑤对得到的数字化视频信号进行编码、压缩。