

# 多媒体技术与应用

- ◆ 多媒体课件概述
- ◆ 文本技术与应用
- ◆ 数字音频技术与应用
- ◆ 图形图像技术与应用
- ◆ 动画技术与应用
- ◆ 数字视频技术与应用
- ◆ 演示型多媒体课件设计与制作



主编 于萍



清华大学出版社

非外借

高等学校计算机应用规划教材

# 多媒体技术与应用

主 编 于 萍

副主编 孙启隆 齐长利 何保锋

清华大学出版社

## 内 容 简 介

本书主要介绍多媒体课件制作的实用技术,内容包括文本、图形、图像、音频、视频和动画等多媒体的处理与制作技术。全书共分为7章:第1章是多媒体课件概述;第2章介绍文本技术与应用;第3章介绍数字音频技术与应用;第4章以Photoshop软件为工具,介绍图形图像技术与应用;第5章以Flash软件为工具,介绍动画技术与应用;第6章介绍数字视频技术与应用;第7章将以上媒体素材整合,介绍演示型多媒体课件的设计与制作。

本书涵盖知识面广,基本包括了典型的多媒体处理软件,具有较强的实用性。本书可作为高等学校、师范院校的教材,也可作为多媒体制作技术人员和爱好者的自学教材或参考书。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

### 图书在版编目(CIP)数据

多媒体技术与应用 / 于萍 主编. —北京:清华大学出版社, 2019

(高等学校计算机应用规划教材)

ISBN 978-7-302-52169-3

I. ①多… II. ①于… III. ①多媒体技术—高等学校—教材 IV. ①TP37

中国版本图书馆 CIP 数据核字(2019)第 011505 号

责任编辑:王 定

封面设计:孔祥峰

版式设计:思创景点

责任校对:牛艳敏

责任印制:沈 露

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座

邮 编:100084

社总机:010-62770175

邮 购:010-62786544

投稿与读者服务:010-62776969, [c-service@tup.tsinghua.edu.cn](mailto:c-service@tup.tsinghua.edu.cn)

质 量 反 馈:010-62772015, [zhiliang@tup.tsinghua.edu.cn](mailto:zhiliang@tup.tsinghua.edu.cn)

印 刷 者:北京富博印刷有限公司

装 订 者:北京市密云县京文制本装订厂

经 销:全国新华书店

开 本:185mm×260mm

印 张:15.5

字 数:348千字

版 次:2019年5月第1版

印 次:2019年5月第1次印刷

定 价:48.00元

产品编号:077758-01

# 前言

## P R E F A C E

计算机的教育应用和信息技术的发展迅猛异常，新的思想、新的方法和新的技术对多媒体课件的制作带来很大冲击。本书介绍了多媒体课件的设计原理与制作技术，更全面、系统地研究多媒体课件，不仅讨论它的设计与制作，而且重视它的运行环境和实用方法；与时俱进，努力将现代信息技术运用到多媒体课件设计中；力求与目前主流软件接轨，争取给读者以更实际和具体的帮助。本书旨在帮助多媒体课件设计人员掌握多种媒体信息的获取与编辑方法，并可根据实际课程内容制作出精良的多媒体课件。

全书共有7章，主要内容如下：第1章介绍了多媒体技术和多媒体计算机系统的基本知识、多媒体技术的教育应用及多媒体课件基础；第2章介绍文本技术与应用，包括文本素材的获取与编辑、文本设计、OCR识别技术、PDF文件处理及电子书制作；第3章介绍数字音频技术与应用，包括数字音频基础、常用音频文件格式及格式转换、音频素材的获取与编辑、音频处理软件 Adobe Audition；第4章介绍图形图像技术与应用，包括图形图像基础、Photoshop 概述及其处理图像的方法；第5章介绍动画技术与应用，包括动画基础、动画素材的获取与编辑、使用 Flash 制作动画的方法；第6章介绍数字视频技术与应用，包括数字视频基础、数字视频文件格式及格式转换、视频素材的获取与编辑、使用 Camtasia Studio 和 Premiere 处理视频；第7章介绍演示型多媒体课件设计与制作方法，包括 PowerPoint 2010 使用技巧、幻灯片切换、超链接、母版设计和打包等知识。

由于本书涉及的知识面较广，知识点多，构成一个完整体系难度较大，不足之处在所难免，恳请广大读者批评指正。

本书提供课件、素材及习题参考答案，下载地址如下：



课件



素材



习题参考答案

编者

2019年2月

# 目录

## C O N T E N T S

第1章 多媒体课件概述	1
1.1 多媒体技术概述	2
1.1.1 媒体的常见形式	2
1.1.2 媒体的分类	4
1.1.3 多媒体及其特点	6
1.1.4 多媒体技术的典型应用	8
1.2 多媒体计算机系统	12
1.2.1 多媒体硬件系统	12
1.2.2 多媒体软件系统	13
1.3 多媒体技术的教育应用	16
1.4 多媒体课件基础	18
1.4.1 多媒体课件的相关概念	18
1.4.2 多媒体课件的分类	18
1.4.3 多媒体课件的开发过程	19
1.5 习题	21
第2章 文本技术与应用	23
2.1 文本素材的获取与编辑	24
2.1.1 文本素材的获取	24
2.1.2 文本素材的编辑	25
2.1.3 文本素材的获取与编辑实例	28
2.2 文本设计	33
2.3 OCR 识别技术	34
2.4 PDF 文件处理	37
2.5 电子书制作	42
2.6 习题	46
第3章 数字音频技术与应用	50
3.1 数字音频基础	51
3.1.1 音频的基本概念	51
3.1.2 音频音质与数据量	51
3.1.3 音频压缩编码的国际标准	52
3.2 常用音频文件格式及格式转换	53
3.2.1 音频文件格式	53
3.2.2 音频文件格式转换	54
3.3 音频素材的获取与编辑	55
3.3.1 音频素材的获取	55
3.3.2 音频素材的编辑	56
3.4 音频处理软件 Adobe Audition	57
3.4.1 Adobe Audition 简介	57
3.4.2 声音录制	58
3.4.3 音频的编辑	60
3.4.4 制作音频效果	62
3.4.5 使用音频插件	63
3.4.6 使用 Adobe Audition 制作音频实例	64
3.5 习题	66
第4章 图形图像技术与应用	67
4.1 图形图像基础	68
4.1.1 图形图像的基本概念	68
4.1.2 常用的图像文件格式	70
4.2 Photoshop 概述	70
4.2.1 Photoshop 的工作界面	70
4.2.2 Photoshop 的基本操作	73
4.3 图像的选取	77
4.3.1 选区的创建	77
4.3.2 选区的编辑	82
4.3.3 图像选择的应用实例	85

4.4	图像的处理	87	5.4.2	高级动画制作	158
4.4.1	绘制图像	87	5.4.3	Flash 制作动画实例	166
4.4.2	修复图像	91	5.5	习题	170
4.4.3	裁剪图像	94	<b>第 6 章</b>	<b>数字视频技术与应用</b>	<b>174</b>
4.4.4	修饰图像	96	6.1	数字视频基础	175
4.4.5	合成图像	97	6.1.1	数字视频的基本概念	175
4.4.6	为图像配文字	101	6.1.2	动态图像压缩编码技术及 国际标准	176
4.4.7	图像特效	104	6.2	数字视频文件格式及格式 转换	177
4.5	使用 Photoshop 制作图形图像 实例	108	6.2.1	数字视频文件格式	177
4.5.1	实例介绍	108	6.2.2	数字视频格式转换	179
4.5.2	实例操作步骤	108	6.3	视频素材的获取与编辑	180
4.6	习题	112	6.3.1	视频素材的获取	180
<b>第 5 章</b>	<b>动画技术与应用</b>	<b>114</b>	6.3.2	视频素材的编辑	181
5.1	动画基础	115	6.3.3	视频素材的获取与编辑 实例	182
5.1.1	动画的基本概念	115	6.4	使用 Camtasia Studio 处理 视频	183
5.1.2	动画的分类	115	6.4.1	视频剪辑	184
5.2	Flash 概述	116	6.4.2	为视频配音	184
5.2.1	Flash 相关概念	116	6.4.3	为视频添加字幕	186
5.2.2	Flash 操作简介	118	6.4.4	视频转场特效	187
5.2.3	Flash 动画文件格式	120	6.5	Adobe Premiere Pro CS4 简介	188
5.2.4	绘制矢量图	120	6.5.1	创建项目并配置项目设置	188
5.2.5	选取工具	123	6.5.2	视频采集与导入素材	190
5.2.6	对象操作	124	6.5.3	装配序列	191
5.2.7	绘图工具	127	6.5.4	在序列中编辑素材	194
5.2.8	关于颜色	135	6.5.5	输出	197
5.2.9	文本	139	6.6	习题	199
5.2.10	元件	143	<b>第 7 章</b>	<b>演示型多媒体课件设计与 制作</b>	<b>202</b>
5.2.11	帧	146	7.1	PowerPoint 工作环境	203
5.3	动画素材的获取与编辑	148	7.2	保存演示文稿	204
5.3.1	动画素材的获取	148			
5.3.2	动画素材的编辑	150			
5.3.3	动画素材的获取与编辑 实例	153			
5.4	使用 Flash 制作动画	155			
5.4.1	基本动画制作	155			

7.3 设置幻灯片背景.....	207	7.7 自定义动画.....	220
7.3.1 设置渐变色填充背景.....	208	7.7.1 动画效果.....	220
7.3.2 设置纹理填充背景.....	209	7.7.2 为对象设置动画效果.....	221
7.3.3 设置图片填充背景.....	210	7.7.3 设置效果选项.....	223
7.3.4 设置图案填充背景.....	211	7.8 音频与视频对象.....	225
7.3.5 制作水印.....	211	7.8.1 插入音频.....	225
7.3.6 配色方案.....	213	7.8.2 插入视频.....	228
7.4 幻灯片切换.....	214	7.8.3 全屏播放视频.....	229
7.5 超链接.....	215	7.8.4 格式化视频.....	229
7.5.1 利用“超链接”按钮创建 超链接.....	215	7.9 幻灯片放映.....	230
7.5.2 利用“动作”按钮创建 超链接.....	216	7.9.1 创建自定义放映.....	230
7.6 设计母版.....	217	7.9.2 设置放映方式.....	231
7.6.1 幻灯片母版.....	217	7.9.3 控制演讲者放映.....	233
7.6.2 标题母版.....	218	7.10 打包成 CD.....	235
7.6.3 讲义母版.....	219	7.11 将演示文稿保存为视频.....	236
7.6.4 备注母版.....	219	7.12 习题.....	238
		参考文献.....	240



# 多媒体课件概述

多媒体技术是当今信息技术领域发展最快、最活跃的技术，它为人们展现了一个多姿多彩的视听世界，令人耳目一新。多媒体技术的出现使得我们的计算机世界丰富多彩起来，也使得计算机世界充满了人性的气息。多媒体技术自问世起即引起人们的广泛关注，并迅速由科学研究走向应用。目前，多媒体技术广泛地应用于教育教学、工业控制、信息管理、办公自动化系统及游戏、娱乐等领域，逐步深入到人们生活的各个方面。

在教育领域，多媒体计算机技术与计算机辅助教学相结合产生了多媒体计算机辅助教学(Multimedia Computer-Assisted Instruction, MCAI)。多媒体计算机辅助教学已成为当今计算机教学应用最普及的方式。它代表了教育教学领域中计算机应用技术的最新发展方向，是教育信息化的重要手段，不仅能促进教学方法的更新和发展，而且有助于改变传统的教育思维模式。多媒体计算机辅助教学受到了广大教育工作者的青睐。

多媒体课件是多媒体计算机辅助教学普遍使用的工具，其通过生动的画面和形象的演示改变了传统教学中一支粉笔和一块黑板的教学手段，运用多媒体技术为学习者提供视觉与听觉的多重感官刺激，有利于调动学习积极性，便于因材施教，有助于取得更好的学习效果。

本章主要介绍多媒体技术的基础知识、多媒体计算机系统、多媒体技术的教育应用和多媒体课件的相关知识。



## 1.1 多媒体技术概述

多媒体技术是一门涉及文本、数值、声音、图形、图像、动画和视频等媒体信息的综合技术,涉及计算机、通信、电视和心理学等多个学科。多媒体技术被认为是继造纸术、印刷术、电报、电话、广播电视和计算机之后,人类处理信息技术的又一大飞跃,是计算机发展史上的一次革命。目前,多媒体技术已广泛应用于军事、教育、音乐、美术、游戏、娱乐、医疗等多个领域,为这些领域的发展和研究带来了勃勃生机,并改变着人们生活、工作和娱乐的方式。

### 1.1.1 媒体的常见形式

通常,在计算机领域中,媒体(Media)有两种含义:一是指存储信息的实体,如磁盘、光盘、软盘等存储设备,一般称为媒质;二是指传播信息的载体,如文本、数值、声音、图形、图像、视频等,一般称为媒介。多媒体计算机技术中的媒体指的是后者。

#### 1. 文本

文本(Text)包括汉字、英文字母、数字、英文标点符号和中文标点符号等,通常由文字编辑软件(如 Microsoft Word、记事本、写字板或 WPS 文字处理软件等)生成。需要注意的是:中文标点符号和英文标点符号是不同的两类文本。这是因为,中文和英文使用的编码形式不同,中文使用汉字标准信息交换码,每个汉字占用 2B,而英文使用美国标准信息交换代码(American Standard Code for Information Interchange, ASCII),每个英文字符占用 1B。

#### 2. 数值

数值(Number)包括整数和实数。整数由正负号和数字组成,如 12, 4, 0, -9, 在计算机中,整数通常是用补码形式表示的。实数由正负号、数字和小数点组成,如 3.14, 14.58, 100.0, 对于实数,计算机中可使用定点或浮点形式表示。

计算机中对于数值的处理有数值运算和非数值运算。数值运算是数值进行常规的数学运算,可应用于求解方程的根、矩阵的秩、数值积分和数值微分等数学问题,是计算机最为传统的功能。非数值运算涉及的对象是文本、图形、图像、声音、视频和动画等,随着计算机的普及,非数值处理技术将计算机的应用拓宽到模式识别、情报检索、人工智能和计算机辅助教学等领域。

#### 3. 声音

声音(Audio)是由物体振动产生的声波。发出振动的物体叫做声源。声音是通过介质(空气、固体或液体)传播并能被人或动物听觉器官所感知的波动现象。声音具有一定的频率范围。人耳可以听到的声音的频率范围为 20Hz~20kHz。高于这个范围的声音称为超声波,

低于这一范围的声音称为次声波。声音的波形图如图 1.1 所示。

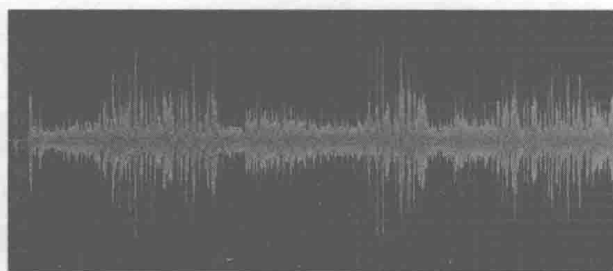
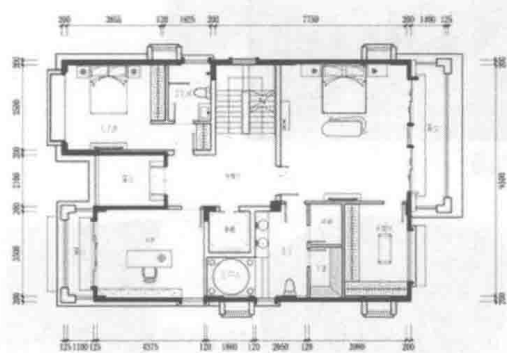


图 1.1 声音的波形图

#### 4. 图形和图像

图形(Graph)和图像(Image)都是多媒体系统中的可视元素。图形和图像的示例如图 1.2 所示。图形是矢量图,是人们根据客观事物制作生成的,不是客观存在的。图形的元素包括点、直线、弧线、圆和矩形等。通常,图形在屏幕上显示要使用专用软件(如 AutoCAD 和 Microsoft Visio 等)将描述图形的指令转换成屏幕上的形状和颜色,由于图形在本质上是由数学的坐标和公式来描述的,所以一般只适用于描述轮廓不是很复杂、色彩不是很丰富的对象,如几何图形、工程图纸和 3D 造型等,并且图形在进行缩放时不会失真,可以适应不同的分辨率。

图像是由扫描仪、摄像机等输入设备捕捉实际的画面产生的数字图像,是由像素构成的位图。就像细胞是组成人体的基本单位一样,像素是组成一幅图像的基本单位。对图像的描述与分辨率和色彩的颜色位数有关,分辨率与颜色位数越高,占用存储空间越大,图像越清晰,但图像在缩放过程中会损失细节或产生锯齿。基于以上特点,图像适用于显示含有大量细节,如明暗变化、场景复杂、轮廓色彩丰富的对象,并且可通过图像处理软件(如 Paint、Brush、Photoshop 等)对图像进行处理以得到更清晰的图像或产生特殊效果。



(a) 图形



(b) 图像

图 1.2 图形和图像

#### 5. 视频

视频(Video)指的是将一系列静态影像以电信号方式加以捕捉、记录、处理、存储、传送与重现的各种技术,是多幅静止图像与连续的音频信息在时间轴上同步运动的混合媒体。

当连续的图像变化每秒超过 24 帧画面以上时, 根据视觉暂留原理, 人眼无法辨别单幅的静态画面, 多帧图像随时间变化而产生运动感, 继而产生平滑连续的视觉效果, 因此视频也被称为运动图像。图 1.3 是多幅静态图像构成的白云运动画面。

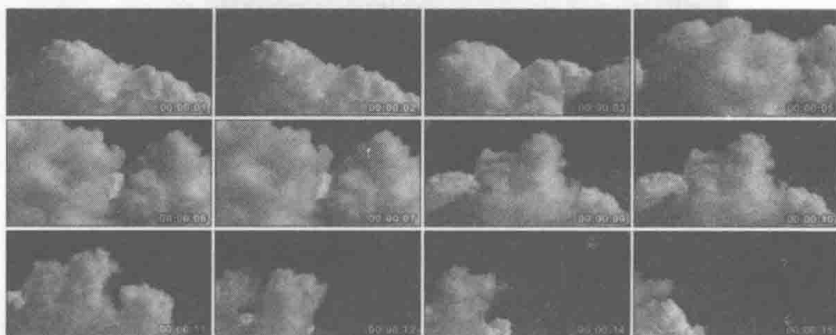


图 1.3 视频

## 6. 动画

动画(Animation)也是一种视频, 指的是采用动画制作软件(如 Adobe Flash CS3、3ds Max 等)生成的一系列可供实际播放的连续动态画面。动画是一门幻想艺术, 更容易直观表现和抒发人们的感情, 扩展人类的想象力和创造力。目前, 动画已成功应用到多个领域, 如娱乐行业的动漫游戏、建筑行业的建筑结构展示、军事行业的飞行模拟训练和机械行业的加工过程模拟。图 1.4 为机械手模拟动画演示界面。

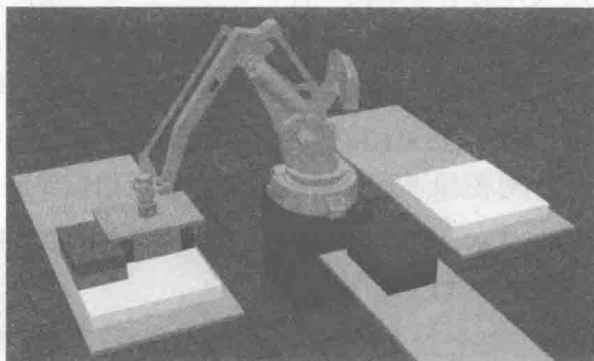


图 1.4 动画

### 1.1.2 媒体的分类

1993 年, 国际电信联盟(International Telecommunication Union, ITU)将媒体划分为以下 5 种类型。

#### 1. 感觉媒体

感觉媒体(Perception Media)指的是能够直接作用于人的感觉器官, 并能够使人直接产生感觉的一类媒体。人的感觉包括视觉、听觉、触觉和味觉等。视觉是人类感知信息最重要的途径。据统计, 依靠视觉获取的外部信息量约占人类获取的总信息量的 70%, 人类通

过视觉器官可感知到符号、图形、图像、视频和动画等媒体信息。除视觉外，人类可通过听觉器官感知声音信息，如语音、音乐和音响等；可通过嗅觉器官感知气味信息；可通过触觉器官(如神经末梢)感知对象的位置、大小、方向、方位、温度和质地等性质。

## 2. 表示媒体

表示媒体(Representation Media)是为了对感觉媒体进行有效加工、处理和传输而人为构造出的一种媒体，表示媒体通常表现为对各种感觉媒体的编码。例如，英文字符的美国标准信息交换代码(ASCII)表示形式，汉字的标准信息交换码和机内码表示形式，整数的补码表示形式，实数的定点数表示形式和浮点数表示形式，语音的脉冲编码调制形式，图像的JPEG编码和视频的MPEG编码形式等。

## 3. 显示媒体

显示媒体(Presentation Media)是指显示感觉媒体的物理设备，即将感觉媒体和计算机中电信号相互转换的一类媒体。显示媒体又分为两类：输入显示媒体和输出显示媒体。其中，输入显示媒体是完成将感觉媒体转换为计算机中的电信号，包括键盘、鼠标、话筒、摄像机、扫描仪、手写笔等，如图 1.5 所示。输出显示媒体的功能则是将计算机中的电信号转换为感觉媒体，包括显示器、打印机、投影仪和扬声器等，如图 1.6 所示。

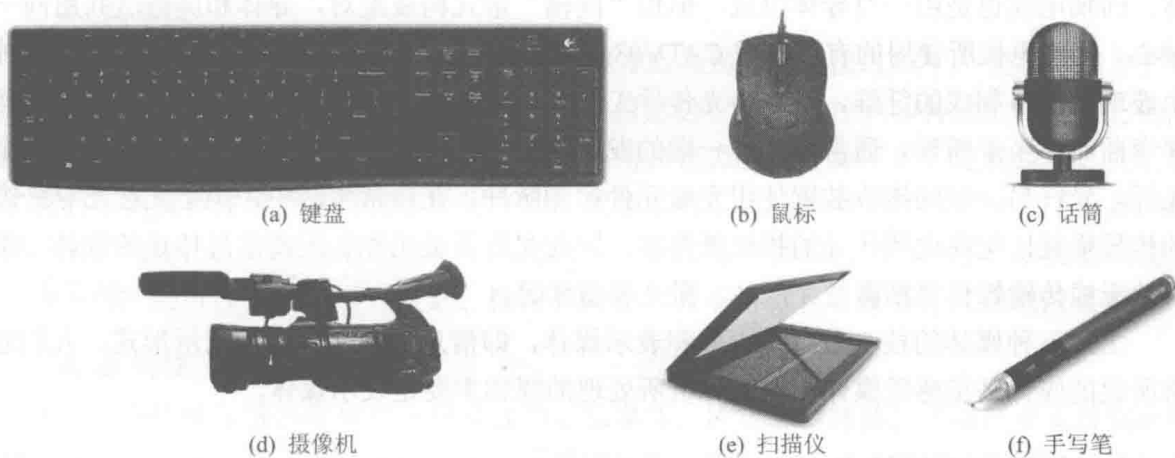


图 1.5 输入显示媒体

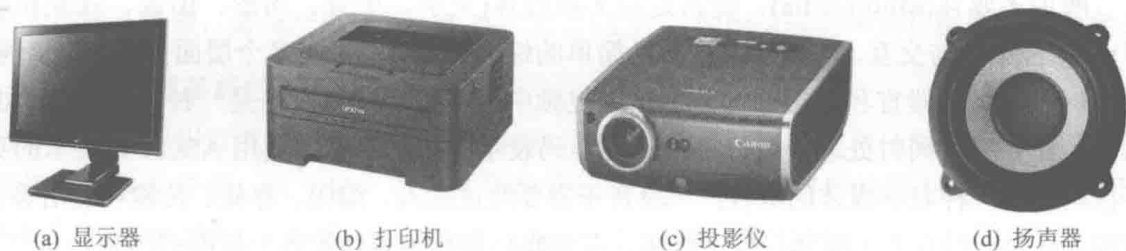


图 1.6 输出显示媒体

## 4. 存储媒体

存储媒体(Storage Media)又称存储介质，是指将感觉媒体转换为表示媒体后，用于存储

表示媒体的物理介质。常见的存储媒体包括硬盘、软盘和光盘等。目前，软盘已经很少使用，而硬盘技术越来越成熟，已成为主要的存储媒体。

### 5. 传输媒体

传输媒体(Transmission Media)是将表示媒体从一处传送到另一处的物理载体，是通信网络中发送方和接收方之间的物理通路。传输媒体可分为有线和无线传输媒体两大类。常见的有线传输媒体包括双绞线、同轴电缆和光纤，如图 1.7 所示。



图 1.7 有线传输媒体

双绞线是将两根绝缘导线螺旋对扭在一起形成的，这种结构可以有效地减少两根导线之间的辐射电磁干扰，其最早被应用于电话通信中模拟信号的传输，是最为常用的传输媒体。同轴电缆也是由一对导体组成，但按“同轴”形式构成配对，导体和屏蔽层共用同一轴心。闭路电视所使用的有线电视 CATV(Cable Television)电缆就是同轴电缆。光纤是一种由玻璃或塑料制成的纤维，可作为光传导工具。微细的光纤封装在塑料护套中，使得它能够弯曲而不至于断裂。通常，光纤一端的发射装置使用发光二极管或激光将光脉冲传送到光纤，光纤另一端的接收装置使用光敏元件检测脉冲。在日常生活中，由于光在光导纤维的传导损耗比电在电线传导的损耗低得多，因此光纤常被用作长距离信息传递的载体。常见的无线传输媒体包括微波、红外、激光和蓝牙等。

上述五种媒体的核心是感觉媒体和表示媒体，即信息的存在形式和表示形式。人们通常所说的媒体是指感觉媒体，但计算机所处理的媒体主要是表示媒体。

### 1.1.3 多媒体及其特点

所谓多媒体(Multimedia)，通常是指多种媒体(文字、声音、图形、图像、视频和动画等)的综合集成与交互。这种综合绝不是简单的综合，而是发生在多个层面上的综合。例如，人们可通过听觉器官和视觉器官分别感受视频中的声音和图像，这是一种感觉媒体层面的综合；计算机可同时处理使用标准信息交换码表示的中文字符和使用 ASCII 码表示的英文字符，这是一种表示媒体的综合；一段音乐需要经过输入、编码、存储、传输和输出等多个过程，这是一种在表示媒体、显示媒体、存储媒体和传输媒体等各个层面的综合。

多媒体具有以下特点。

#### 1. 大数据量

多媒体技术将计算机所能处理的信息空间扩展和放大，使之不再仅仅局限于数字和文

本。计算机由原来的无声世界进入到有声世界；由原来的静止画面进入到动态画面乃至活动影像。处理的媒体类型包括文本、声音、图形、图像、视频和动画等。其中，声音、图形、图像、视频和动画都需要占用较大的存储空间。例如，一幅分辨率为  $4032 \times 3024$  的 JPEG 格式的图像，大约占用 2.16MB 的存储空间；一首播放时间为 5 min，采样频率为 44.1 kHz 的 MP3 音乐，大约占用 5.3 MB 的存储空间。

## 2. 集成性

多媒体的集成性可以理解为两种情况：一种是多媒体信息的集成，即多媒体系统总是同时处理多种媒体信息。在多媒体系统中，各种媒体信息不再采用单一方式进行采集与处理，而是多通道同时统一采集、存储与加工处理，更加强调各种媒体之间的协同关系。例如，文字编辑系统同时处理包括文字和图像的集成媒体信息；视频处理软件同时处理包括声音和图像的集成媒体信息。另一种是多媒体处理设备的集成，在硬件方面，多媒体系统包括能处理多媒体信息的高速并行的中央处理器、多通道的输入输出接口及外设和大容量的存储器，并将这些硬件设备集成为统一的系统；在软件方面，多媒体系统包括多媒体操作系统、满足多媒体信息管理的软件系统、高效的多媒体应用软件和创作软件等。这些多媒体系统的硬件和软件被集成为处理各种复合信息媒体的信息系统。

## 3. 交互性

所谓交互性，就是把人的活动作为一种媒体加入到信息传播过程中，使参与信息交互的各方，不论是发送方还是接收方，都可以对信息进行编辑、控制和传递的特性。它向用户提供了更加有效地控制和使用信息的手段，同时也为应用开辟了更广泛的领域。交互性使我们在获取和使用信息时变被动为主动，增加了对信息的注意和理解，延长了信息的保留时间。例如，在计算机辅助教学中，可以人为地改变节目的内容和顺序，研究感兴趣的某些方面；还可以主动地进行检索、提问和回答，而不是像看电视那样被动地接收信息。

## 4. 实时性

由于多媒体系统需要处理各种复合的信息媒体，决定了多媒体技术必然要支持实时处理。接收到的各种信息媒体在时间上必须是同步的。例如，语音和活动的视频图像必须严格同步，因此要求实时性，甚至是强实时(Hard Real Time)。例如，电视会议系统的声音和图像不允许存在停顿，必须严格同步，包括“唇音同步”，否则传输的声音和图像就失去意义。

## 5. 编码方式多样性

多媒体的编码方式也呈现多样性。例如，文本中的英文字符使用 ASCII 码；中文字符使用汉字信息交换码；汉字输入时可使用拼音编码方案或五笔字型编码方案；语音使用脉冲编码调制 PCM 形式；图像使用 JPEG 编码；视频使用 MPEG 编码。

由于现在的多媒体信息都是由计算机进行处理，因此，多媒体除了具有多种媒体的含义以外，还包括处理和应用多媒体信息的一整套技术。即多媒体技术是将文本、声音、图形、图像、视频和动画等多种媒体信息通过计算机进行采集、处理、存储和传输的各种技术的统称。在处理这些多媒体信息时，首先，将自然界存储的各种媒体编码成计算机可以

处理的二进制编码形式；然后，在计算机上完成对多媒体信息的各种处理，如字体的大小设置，图像的旋转、滤镜、模糊，声音的淡入淡出，视频的剪辑等，这些处理都是基于计算机中存储的多媒体的二进制信息进行的，都是数字处理，当需要保存处理后的多媒体信息时，涉及多媒体的存储；最后，将处理完成的多媒体信息从一台计算机传输到另一台计算机时，涉及多媒体的网络通信。在以上多媒体信息进行采集、处理、传输和存储的过程中产生了多种多媒体技术，这些技术包括多媒体信息编码技术、多媒体信息数字化处理技术、大容量存储技术、多媒体信息压缩技术和多媒体网络通信技术等，它们相辅相成，共同促进多媒体技术的发展。

### 1.1.4 多媒体技术的典型应用

多媒体技术是一种实用性很强的技术，当使用者通过人机接口访问任何种类的电子信息时，多媒体都可以作为一种适当的手段。多媒体大大改善了计算机的人机界面，集文本、声音、影像于一体，更接近于人类自然的信息交流方式，提高了计算机的易用性和可用性，同时增强了信息的记忆能力与效率。多媒体技术不仅使计算机产业日新月异，而且极大地改变了人们传统的思维、学习、工作和生活的方式。

#### 1. 基于内容的图像检索

基于内容的图像检索(Content-Based Image Retrieval, CBIR)，这一概念是 1992 年由 T.Kato 首次提出的。他在论文中首次构建了一个基于色彩与形状进行查询的图像数据库，并根据一定的检索功能进行实验。此后，基于图像特征提取以实现图像检索的过程以及 CBIR 这一概念，被广泛应用于统计学、模式识别、信号处理和计算机视觉等各种研究领域。目前，CBIR 是计算机视觉领域中关注大规模数字图像内容检索的研究分支。典型的 CBIR 系统允许用户输入一张图片，以查找具有相似甚至相同内容的其他图片。而传统的图像检索都是基于关键词的文本检索，即通过图片的名称和文字信息来实现检索功能。

基于 CBIR 技术的图像检索系统，在建立图像数据库时，系统对输入的图像进行分析并分类统一建模，然后根据各种图像模型提取图像特征存入特征库，同时对特征库建立索引以提高查找效率。用户在输入查询条件时，可以采用一种特征或几种综合特征来检索，然后系统根据相似性匹配算法计算关键图像特征与特征库中图像特征的相似度，最后根据相似度从大到小的顺序将匹配图像反馈给用户。用户可根据自己的满意程度，选择是否改变查询条件继续查询，以达到满意的查询结果。

目前，国外已经出现一些可以真正实现基于内容的图像检索系统。例如，IBM 公司的 QBIC 系统、哥伦比亚大学的 VisualSEEK 图像检索系统和 WebSEEK 图像及视频搜索引擎等。国内基于内容的图像检索技术也取得了一定的研究成果，仍有较大进步空间。

#### 2. 语音识别

语音识别技术，也称为自动语音识别(Automatic Speech Recognition, ASR)，是将人类的语音中的词汇内容转换为计算机可以识别的文本或命令。语音识别技术的处理对象为语音，是模式识别的一个重要发展分支。语音识别是一门综合人工智能、数字信号处理、模

式识别、语言学、声学 and 认知科学等多学科领域的综合技术，包括多方面的研究领域：根据对说话人说话方式的要求，可分为孤立字(词)、连接词和连续语音识别系统；根据对说话人的依赖程度，可分为特定人和非特定人语音识别系统；根据词汇量的大小，可分为小词汇量、中等词汇量、大词汇量和无限词汇量语音识别系统。

语音识别技术有着非常广泛的应用领域和市场前景。它的出现使人们甩掉键盘，通过语音方式与计算机进行沟通，使用语音要求、请示、命令或询问以得到正确的响应。目前已应用于声控语音拨号系统、声控智能玩具、工业控制领域、军事领域等。微软公司在 Windows 操作系统中已经开发了语音识别系统，图 1.8 为 Windows 中的语音识别界面。



图 1.8 Windows 中的语音识别界面

用户可通过学习语音识别教程掌握语音识别的常用命令。Windows 中的语音识别常用命令如表 1.1 所示。

表 1.1 Windows 中的语音识别常用命令

功 能	说出的内容
按项目名称单击任何项目	单击文件、开始、查看
单击任何项目	单击回收站、单击计算机、单击文件名
双击任何项目	双击回收站、双击计算机、双击文件名
切换到某个打开的程序	切换到画图、切换到写字板、切换到程序名
沿一个方向滚动	向上滚动、向下滚动、向左滚动、向右滚动
在页面中滚动确切的距离	向下滚动 2 页、向上滚动 10 页
在文档中插入新段落或换行	新段落、换行
在文档中选择字词	选择字词
选择某个字词并开始对其更正	更正字词
选择并删除特定字词	删除字词
显示适用命令的列表	我可以说什么
更新当前可用的语音命令列表	刷新语音命令
让计算机听您说话	开始聆听



(续表)

功 能	说出的内容
让计算机停止聆听	停止聆听
移动语音识别麦克风栏	移动语音识别
最小化语音识别	最小化语音识别
将光标放到特定字词之前	转到字词
将光标放到特定字词之后	转到字词后面
请勿在下一个字词前插入空格	无空格
转到光标所在句子开头	转到句子开头
转到光标所在段落开头	转到段落开头
选择当前文档中的字词	字词到字词
选择当前文档中的所有文本	选择全部文本
选择光标位置之前的多个字词	选择前 20 个字词
选择最后听写的文本	选择它
在屏幕上清除选定内容	清除选定内容
删除前一个句子	删除前一个句子
删除选定的文本或最后听写的文本	删除这个
在键盘上按任意键	按键盘键、按 a、按大写字母 B、按 Shift+a
在不首先说“按”的情况下直接按某些键盘键	Delete、Backspace、Enter、Page Up、Page Down、Home、End、Tab
,	逗号
单击某个带编号的项目	19 确定、5 确定
双击某个带编号的项目	双击 19、双击 5
右键单击某个带编号的项目	右键单击 19、右键单击 5
关闭程序	关闭这个、关闭画图、关闭文档
最小化	最小化这个、最小化画图、最小化文档
最大化	最大化这个、最大化画图、最大化文档
还原	还原这个、还原画图、还原文档
剪切	剪切这个、剪切
复制	复制这个、复制
粘贴	粘贴
删除	删除这个、删除
撤消	撤消这个、撤消
为下一个命令插入由字母组成的字词	由字母组成的字词
插入数字形式的数	由数字组成的数

### 3. 虚拟现实

虚拟现实(Virtual Reality, VR)技术是一种可以搭建和体验虚拟世界的计算机技术,它通过计算机生成一种逼真的模拟环境,是一种多源信息融合交互式的三维动态视景和实体行为的系统仿真。人们可以借助传感头盔、眼镜和数据手套等专业设备,通过视觉、触觉和听觉进入虚拟空间,实时感知和控制虚拟世界中的各种对象,获得身临其境的感受。虚拟现实技术是仿真技术与计算机图形学、传感技术和网络技术等多种技术的融合,是仿真技术的一个