

3.1 多媒体技术的基本概念

“多媒体”一词译自英文“Multimedia”，即“Multiple”和“Media”的合成，其核心词是媒体。媒体(medium)在计算机领域有两种含义：即媒质和媒介。(媒质：存储信息的实体，如磁盘、光盘、磁带、半导体存储器等。媒介：传递信息的载体，如数字、文字、声音、图形和图像等。)多媒体就是运用计算机综合处理多种媒体信息(如文本、声音、图形、图像、动画等)，使多种信息建立逻辑链接(Logical Link)，并集成一个具有交互性(Interactive)的系统。

多媒体是超媒体系统中的一个子集，超媒体系统是使用超链接构成的全球信息系统，全球信息系统是因特网上使用 TCP/IP 和 UDP/IP 的应用系统。二维的多媒体网页使用 HTML 来编写，而三维的多媒体网页使用 VRML 来编写。在目前许多多媒体作品使用光盘存储器发行，在将来多媒体作品更多地使用网络来发行。

通常概念的“媒体”，可分为以下五种类型：感觉媒体、表示媒体、显示媒体、存储媒体和传输媒体。

(1) 感觉媒体：能直接作用于人们的感觉器官，从而能使人产生直接感觉的媒体。如语音、音乐、各种图像、动画、文本等。

(2) 表示媒体：为了传送感觉媒体而人为研究出来的媒体。借助于此种媒体，便能更有效地存储或传送感觉媒体。如语言编码、电报码等。

(3) 显示媒体：用于通信中使电信号和感觉媒体之间产生转换用的媒体。如输入/输出设施、键盘鼠标器、显示器、打印机等。

(4) 存储媒体：用于存放某种媒体的媒体，如纸张、磁带、磁盘、光盘等。

(5) 传输媒体：用于传输某些媒体的媒体，如电话线、电缆光纤等。

3.2 多媒体系统的组成

多媒体计算机是一组硬件和软件设备，它结合了各种视觉和听觉媒体，能够产生令人印象深刻的视听效果。多媒体计算机无非就是具有了多媒体处理功能的计算机(如早期的 586 机型)，它的硬件结构与一般所用的计算机并无太大的差别，只不过是多了一些软硬件配置而已。一般用户如果要拥有多媒体计算机大概有两种途径：一是直接购买具有多媒体功能的 PC；二是在基本的 PC 上增加多媒体套件而构成多媒体计算机。今天，对计算机厂商和开发人员来说，多媒体计算机已经成为一种必须具有的技术规范。

多媒体计算机的基本配置：一般来说，多媒体个人计算机(MPC)的基本硬件结构可以归纳为七部分：(1)至少有一个功能强大、速度快的中央处理器(CPU)；(2)可管理、控制各种接口与设备的配置；(3)具有一定容量(尽可能大)的存储空间；(4)高分辨率显示接口与设备；(5)可处理音响的接口与设备；(6)可处理图像的接口设备；(7)可存放大量数据的配置等。

1. 多媒体系统的层次结构与组成

一般的多媒体系统应该包括如下 5 个层次的结构：多媒体硬件系统、多媒体软件系统、多媒体应用程序接口、创作工具和多媒体应用系统。

(1) 多媒体硬件系统：包括计算机硬件、声音/视频处理器、多种媒体输入、输出设备及信号转换装置、通信传输设备及接口装置等。其中，最重要的是根据多媒体技术标准而研制生成的多媒体信息处理芯片、光盘驱动器等。

(2) 多媒体软件系统：包括多媒体文件系统、多媒体操作系统和多媒体通信系统。其中，最重要的是多媒体操作系统，也称为多媒体核心系统(Multimedia kernel system)，具有实时任务调度、多媒体数据转换和同步控制对多媒体设备的驱动和控制，以及图形用户界面管理等。

(3) 多媒体应用程序接口(API)。

(4) 创作工具：或称为媒体处理系统工具、多媒体系统开发工具软件，是多媒体系统重要组成部分。

(5) 多媒体应用系统：根据多媒体系统终端用户要求而定制的应用软件或面向某一领域的用户应用软件系统，它是面向大规模用户的系统产品。

2. 常用多媒体设备简介

包括输入设备中的扫描仪、数码相机和语音输入系统、手写输入系统、IC 卡输入系统，输出设备中的各种打印机和绘图仪，光盘驱动器、声卡、音箱、视频卡、电视接收卡、SCSI 卡及摄像头等多媒体适配器，网络设备等。

1) 触摸屏

触摸屏是一种常见的多媒体界面，是随着多媒体技术发展而兴起的一种新型输入设备，它提供了一种人与计算机非常简单、直观的输入方式。

从原理上来看，触摸屏主要分为：红外线式、电阻式、电容式、表面声波式及压力式。目前常用的是电阻式和电容式。电阻式触摸屏由二层膜组成，膜之间有网格触点阵列，对膜的压力会造成电阻的变化，从而定位压点的位置。与电阻式触摸屏略有不同的是，电容式触摸屏上镀有一层金属膜，通过触摸金属膜而产生的电流变化来定位压点的位置。

2) DVD

DVD 原本称为数字视盘(Digital Video Disk)，现在一般称 DVD 为数字通用光盘(Digital Versatile Disk)。关于 DVD 的技术实际上有很多：DVD-ROM 用作存储电脑数据；DVD-Video 用作存储图像；DVD-Audio 用作存储音乐；DVD-R 只可写入一次刻录碟片；DVD-RAM 可重复写入刻录碟片。

DVD 是按照国际标准组织(ISO)和国际电工委员会(IEC)制定的 MPEG-2 标准的基本级进行制作的，是一种体积小、容量大的存储设备。激光头采用红色半导体激光器，比 CD 用的激光波长短 15% 以上，信号读取效率比 CD 高 20% 以上。采用波长更短的蓝色半导体

激光器的 DVD 机还会进一步提高容量。

3) 顶置型摄像机

顶置型摄像机通过一个布满成千上万电荷耦合设备——CCD 的微型芯片将光转换成电脉冲,光线越强,电荷量越大。CCD 可以把亮度分级,但并不认识颜色,彩色摄像机通常用三个 CCD 芯片来建立真彩色合成。摄像机用三个彩色滤色镜来为 CCD 提供合适的光线:红色、绿色、蓝色。CCD 的精度决定了最高分辨率,这是选购摄像机时就考虑的一个重要参数,当然镜头的质量和图像处理技术也是一个重要的性能指标。

和单独使用的数字相机及数字摄像机不一样,这类摄像机自身没有存储器来存放图像,而是直接把数据实时送往系统。

3.3 多媒体文件格式及标准

1. 常见音频文件格式

WAV(Waveform audio format)是微软与 IBM 公司所开发的一种声音编码格式,在 Windows 平台受到广泛的支持。由于此音频格式未经过压缩,所以在音质方面不会出现失真的情况,但也因此在众多音频格式中体积较大。

MP3 全称是动态影像专家压缩标准音频层面 3(Moving Picture Experts Group Audio Layer III)。是当今较流行的一种数字音频编码和有损压缩格式,它设计用来大幅度地降低音频数据量,而对于大多数用户的听觉感受来说,重放的音质与最初的不压缩音频相比没有明显的下降。

WMA(Windows Media Audio)是微软公司开发的一种数字音频压缩格式。一般情况下相同音质的 WMA 和 MP3 音频,前者文件体积较小;而“Windows Media Audio Professional”达到比 Dolby Digital(杜比数字)更优秀的音质。

2. 常见静态图像文件格式

BMP 取自位图 Bit Map 的缩写,也称为 DIB(与设备无关的位图)是微软视窗图形子系统(Graphics Device Interface)内部使用的一种位图图形格式。BMP 图像文件的扩展名为 .bmp。BMP 文件存储数据时,图像的扫描方式是按从左到右、从下到上的顺序。BMP 文件通常是不压缩的,所以它们通常比同一幅图像的压缩图像文件格式要大很多。

JPEG 图像文件是目前使用的最广泛、最热门的静态图像文件,这是由于 JPEG 格式的图像文件具有高压缩率、高质量、便于网络传输的原因,它的扩展名为 .jpg。JPEG 是 Joint Photographic Experts Group(联合摄影专家小组)的缩写,该小组是 ISO 下属的一个组织。JPEG 采用的是有损压缩,由于它采用了高效的 DCT 变换、哈夫曼编码等技术,造成在高压缩比的情况下,仍然有着很高的图像质量。

标签图像文件格式(Tagged Image File Format, TIFF)是一种主要用来存储包括照片和艺术图在内的图像的文件格式。它最初由 Aldus 公司与微软公司一起为 PostScript 打印开发。它存储的图像细微层次的信息非常多,图像的质量也得以提高,故而非常有利于原稿的复制。TIFF 格式在业界得到了广泛的支持,图像处理应用,桌面印刷和页面排版应用,扫描、传真、文字处理、光学字符识别和其他一些应用等都支持这种格式。

3. 常见动态影像文件格式

MPEG 是 Motion Picture Experts Group(运动图画专家小组)的缩写,是 ISO 下属的一个组织。该小组于 1988 年组成,至今已经制定了 MPEG-1、MPEG-2、MPEG-3、MPEG-4、MPEG-7 等多个标准。用 MPEG 格式来存储动态影像文件,能节省大量的存储空间(压缩倍数为几百倍),同时影像的质量也很好(全屏幕、全运动、真彩色),MPEG 文件的扩展名一般为 .dat 或 .mpg。

AVI 是英语 Audio Video Interleave(“音频视频交织”或译为“音频视频交错”)的首字母缩写,由微软在 1992 年 11 月推出的一种多媒体文件格式,它的扩展名为 .avi。现在所说的 AVI 多是指一种封装格式。即 AVI 本身只是提供了一个框架,内部的图像数据和声音数据格式可以是任意的编码形式。

3.4 信息媒体数字化技术

各种模拟信息,例如模拟的音频、图像和视频,其数字化时的实现技术有所不同。但其数字化的基本过程是一致的,即采样、量化和编码三个步骤。

1. 采样

采样(Sample)是对模拟信号进行周期性抽取样值的过程,即将信号从连续时间域上的模拟信号按照一定时间间隔采样,然后转换到离散时间域上的离散信号的过程。这个过程由模数转换器(ADC)(又称采样器)实现。经过对模拟信号采样而得到的信号称为离散信号(是连续信号的离散形式)。为了保证在采样之后数字信号能完整地保留原始信号中的信息,能不失真地恢复成原模拟信号,采样频率应不小于输入模拟信号频谱中最高频率的两倍。

2. 量化

采样把模拟信号变成了时间上离散的脉冲信号,但脉冲的幅度仍然是模拟的,还必须进行离散化处理,才能最终用数码来表示。量化(Quantization)指将信号的连续取值近似为有限多个(或较少的)离散值的过程。量化在有损数据压缩中起着相当重要的作用。

3. 编码

经采样和量化后得到的数据量非常大,所以使用编码对数据进行压缩与传输。在计算机科学和信息论中,数据压缩或者编码是按照特定的编码机制用比未经编码少的数据位(或者其他信息相关的单位)表示信息的过程。

3.5 音频处理技术

3.5.1 声音的基本特征

模拟音频信号有两个重要物理参数:频率和幅度。频率是声波每秒钟振动的次数,表示声音的音调;幅度是从信号的基线到当前波峰的距离,表示声音的强弱;周期是指信号在两个波峰或谷底之间的相对时间。周期和频率之间的关系是互为倒数。

3.5.2 声音数字化的过程

如果要用计算机对音频信息进行处理,则首先要将模拟音频信号转变成数字信号。音频信号的数字化,是将模拟音频信号每隔一定时间间隔截取一段,并将所截取的信号振幅值用一组二进制脉冲序列表示,使连续的模拟音频信号等价地转换成离散的数字音频信号。即模拟音频数字化过程由采样、量化和编码三个步骤组成。

数字音频的技术指标主要是指采样频率和量化位数(或量化深度)。

一秒钟内采样的次数称为采样频率。采样的频率越高,丢失的信息就越少,数字化的声音就越接近源音质,存储量越大。音调越高的声音需要的采样频率也越高。

量化位数是指每个样本量化后一共可取多少个离散的数值,或用多少个二进制数位来表示。采样的位数的越高,则量化的精度就越高,数字化的声音也就越接近源音质。

3.5.3 音频信息编码

音频信息编码一般可分为波形编码、参数编码和混合编码三种类型。

波形编码方式:以数字序列编码的方式尽可能重新构建信源的波形。在时间轴上对模拟信源按一定的速率进行采样,然后将幅度样本分段量化,并用数字序列表示。解码是其反过程,将收到的数字序列恢复成模拟信号。

参数编码方式:是分析并提取信源信息模型中必要的、关键的但不是全部的特征参数,将上述参数信息通过采样、量化、编码,然后合成发送出去;在接收端通过接收到的参数取值的编码,还原出信源信息。

混合编码方式:是结合波形编码和参数编码的优点,总体上使用参数编码的保留低带宽需求优点,在重点的部分信息应用波形编码获得较高质量的合成语音,增强了语音的自然度。

3.5.4 MIDI 技术

MIDI 是 Music Instrument Digital Interface(乐器数字接口)的缩写。MIDI 是用来将电子乐器相互连接,或将 MIDI 设备与计算机连接成系统的一种通信协议。通过它各种 MIDI 设备都可以准确传送 MIDI 信息。

MIDI 的特点在于它处理音乐的方式,不是将声音编码而是将 MIDI 音乐设备上产生的每一个活动编码记录下来。在 MIDI 文件中,只包含产生某种声音的指令,这些指令包括使用什么 MIDI 乐器、乐器的音色、声音的强弱、声音持续时间的长短等。计算机将这些指令发送给声卡,声卡按照指令将声音合成出来。例如,在 MIDI 音乐设备的键盘上演奏时,MIDI 文件记录下按了哪一个键,力有多大,时间有多长。

3.6 数字图像处理技术

3.6.1 图形与图像的概念

图形图像作为一种视觉媒体,早已成为人类信息传输、思想表达的重要方式之一。数字图像处理与计算机图形学(Computer Graphics),无论在概念上还是在实用方面都是各自独

立发展起来而又难以分清的技术领域。

数字图像处理是指将图像信号转换成数字信号并利用计算机对其进行处理的过程。数字图像处理主要研究的内容包括图像变换、图像编码压缩、图像增强和复原、图像分割、图像描述、图像分类(识别)等。

计算机图形学的主要研究内容就是研究如何在计算机中表示图形,以及利用计算机进行图形的计算、处理和显示的相关原理与算法。可以说,计算机图形学的一个重要研究内容就是要利用计算机产生令人赏心悦目的真实感图形。图形通常由点、线、面、体等几何元素和灰度、色彩、线型、线宽等非几何属性组成。

图形与图像两个概念间的区别越来越模糊,但还是有区别的:图像在计算机中以具有颜色信息的点阵来表示,它强调图形由哪些点组成,记录点及它的灰度或色彩。而图形(graphics)在计算机中由场景的几何模型和景物的物理属性表示,它更强调场景的几何表示,记录图形的形状参数与属性参数。它的显示形式是基于线条信息的矢量图和基于明暗(Shading)处理后的图像图。

3.6.2 图像的颜色模型

颜色模型是使用一组值(通常使用三个、四个值或者颜色成分)表示颜色方法的抽象数学模型。建立颜色模型可看作建立一个 3D 的坐标系统,其中每个空间点都代表某一特定的彩色。

颜色模型可分为面向硬设备的颜色模型和面向视觉感知的颜色模型。面向硬设备的颜色模型非常适合在输出显示场合使用,例如 RGB(Red, Green, Blue)颜色模型、CMY(Cyan, Magenta, Yellow)颜色模型。面向视觉感知的颜色模型与人类颜色视觉感知比较接近,其独立于显示设备,包含 HSI(色调 Hue、饱和度 Saturation、亮度 Intensity)颜色模型、HSV(Hue, Saturation, Value)颜色模型、LAB 颜色模型。

3.6.3 图像的数字化过程

光学图像、照片以及人的眼睛看到的一切景物,都是模拟图像,图像无法直接用计算机处理。为了使图像能在电子计算机中作处理运算,必须将模拟图像转化为离散数字所表示的图像,即所谓的数字图像。转化为数字图像的过程称为图像数字化。

图像的数字化过程同样分为采样、量化与编码三个步骤。

图像采样是指把图像分割成为 $M \times N$ 个小区域,用特定的数值来表示每一个小区域的亮度、色彩等特征。 $M \times N$ 表示图像的分辨率。

采样后得到的亮度值(或色彩值)在取值空间上仍然是连续值。把采样后所得到的这些连续量表示的像素值离散化为整数值的操作叫作量化。

从信息论的观点来看,描述图像的数据是信息量(信息源)和信息冗余之和。图像数据压缩编码的本质就是减少这些冗余量。目前数字图像编码的国际标准有 JPEG 和 JPEG 2000。

3.6.4 数字图像的技术指标

数字图像的技术指标主要有图像分辨率(采样频率)和图像深度(量化位数),它们是影响数字图像质量的重要因素。

其中图像分辨率有以下几个方面的含义。

图像分辨率：数字化图像水平与垂直方向像素的总和。例如，800 万像素的数码相机，图像最高分辨率为 3264×2448 等。

屏幕分辨率：一般用显示器屏幕水平像素 $\times$ 垂直像素表示，如 1024×768 等。

印刷分辨率：图像在打印时，每英寸像素的个数，一般用 dpi(像素/英寸)表示。例如，普通书籍的印刷分辨率为 300dpi，精致画册印刷分辨率为 1200dpi。

3.7 视频处理技术

广义地说多媒体的视频技术包括：图像的数字化、压缩编码、数字图像处理及传输、图像编辑和变换、图像的存储、检索和组织管理技术等。

3.7.1 视频的概念

视频与图像是两个既有联系又有区别的概念。就数字媒体的语境而言，数字视频中的每帧画面均形成一幅数字图像，对视频按时间逐帧进行数字化得到的图像序列即为数字视频。因此，可以说图像是离散的视频，而视频是连续的图像。

需要指出的一点是视频数字化的概念是建立在模拟视频占主角的时代，现在通过数字摄像机摄录的信号本身已是数字信号，并且这种趋势越来越明显。

3.7.2 视频数字化的过程

视频数字化是将模拟视频信号经模数转换和彩色空间变换转为计算机可处理的数字信号。与其他媒体的数字化过程类似，视频数字化过程首先必须把连续的图像函数 $f(x, y)$ 进行空间和幅值的离散化处理，空间连续坐标 (x, y) 的离散化，叫作采样； $f(x, y)$ 颜色的离散化，称之为量化。两种离散化结合在一起，叫作数字化，离散化的结果称为数字视频。

编码技术主要分成帧内编码和帧间编码，前者用于去掉图像的空间冗余信息，后者用于去除图像的时间冗余信息。

3.7.3 数字视频压缩标准

20 世纪 90 年代以来 ITU-T(国际电信联盟)和 ISO(国际标准化组织)制定了一系列音视频编码技术标准(信源编码技术标准)和建议，主要有两大系列 ISO 制定的 MPEG 系列标准，数字电视采用的是 MPEG 系列标准 ITU 针对多媒体通信制定的 H. 26x 系列视频编码标准。这些标准和建议的制定极大地推动了多媒体技术的实用化和产业化。

视频编码标准并非一个单一的算法，而是一整套的编码技术与方案，这些技术综合起来就达到了完整的压缩效果。

3.8 多媒体技术的应用

多媒体技术是当今信息技术领域发展最快、最活跃的技术，本文通过对多媒体技术的应用现状和发展趋势的分析，使我们展望到，随着日益普及的高速信息网，它正被广泛应用在

咨询服务、图书、教育、通信、军事、金融、医疗等诸多行业。多媒体技术的应用领域涉及多媒体出版。国家新闻出版署对电子出版物定义为“电子出版物”，是指以数字代码方式将图、文、声、像等信息存储在磁、光、电介质上，通过计算机或类似设备阅读使用，并可复制发行的大众传播媒体。该定义明确了电子出版物的重要特点。电子出版物的内容可分为电子图书、辞书手册、文档资料、报刊杂志、教育培训、娱乐游戏、宣传广告、信息咨询、简报等，许多作品是多种类型的混合。

多媒体办公自动化系统。多媒体技术为办公室增加了控制信息的能力和充分表达思想的机会，许多应用程序都是为提高工作人员的工作效率而设计的，从而产生了许多新型的办公自动化系统。由于采用了先进的数字影像和多媒体计算机技术，把文件扫描仪、图文传真机、文件资料微缩系统和通信网络等现代化办公设备综合管理起来，将构成全新的办公自动化系统，成为新的发展方向。

计算机会议。计算机会议系统是基于多媒体计算机技术的一类视频会议系统，也称为多媒体计算机会议系统，它为 CSCW 系统提供了一种重要的协同工作环境和工具。是多媒体网络的重要应用，不同地点的人员可以通过它来传送文件、讨论问题、协调工作、共享信息等。人们无须关心地理位置上的差异，只需把自己的方案、档案资料准备好，就可以随时给予“与会”各方，面对面地讨论问题。多媒体会议系统可以提高办公室自动化的质量和效率。身处不同地理位置的双方可以共享、修改、存储、显示数据和文件。文件会议和白板应用可以集语音、传真、文件、图像和视频于一体，可以直接将传真文字识别成文件保存，实现无纸办公。

多媒体信息查询系统。近年来随着计算机网络的全面普及，多媒体信息查询发展很快。IBM 公司数字图书馆方案将物理信息转化为数字多媒体形式，通过网络安全地发送给世界各地的用户。自然语言查询和概念查询对返回给用户的信息进行筛选，使相关数据的定位更为简单和精确。聚集功能将查询结果组织在一起，使用户能够简单地识别并选出相关的信息。摘要功能能够对查询结果进行主要观点的概括，这样用户不必查看全部文本就可以确定所要查找的信息。IBM Almaden 研究中心推出了 QBIC 系统。该系统开创了图像信息查询的全新领域。图像可以按照颜色、灰度、纹理和位置进行查询。查询要求将以图形方式表达，如从颜色表中选取颜色，或从例图中选择图像的纹理。查询结果可以按照相关的序列指导子序列查询的进行。这种方法能够使用户更为快速和简便地对可视化信息进行筛选和确定。

其他的应用领域包括交互式电视与视频点播、交互式影院与数字化电影、数字化图书馆、家庭信息中心、远程教育、远程医疗、计算机支持下的协同工作、虚拟现实、媒体空间等。

习 题

一、填空题

1. 通常概念的“媒体”，可分为以下五种类型：_____、_____、_____、_____和_____。
2. 一般多媒体系统包括如下五个层次的结构：_____、_____、_____、_____和_____。

3. _____或称为多媒体系统开发工具软件,是多媒体系统重要组成部分。
4. 多媒体操作系统。也称为_____,具有实时任务调度、多媒体数据转换和同步控制对多媒体设备的驱动和控制,以及图形用户界面管理等。
5. _____是一组硬件和软件设备,它结合了各种视觉和听觉媒体,能够产生令人印象深刻的视听效果。

二、简答题

1. 什么是多媒体?
2. 什么是多媒体计算机?
3. 多媒体技术的应用领域有哪些?