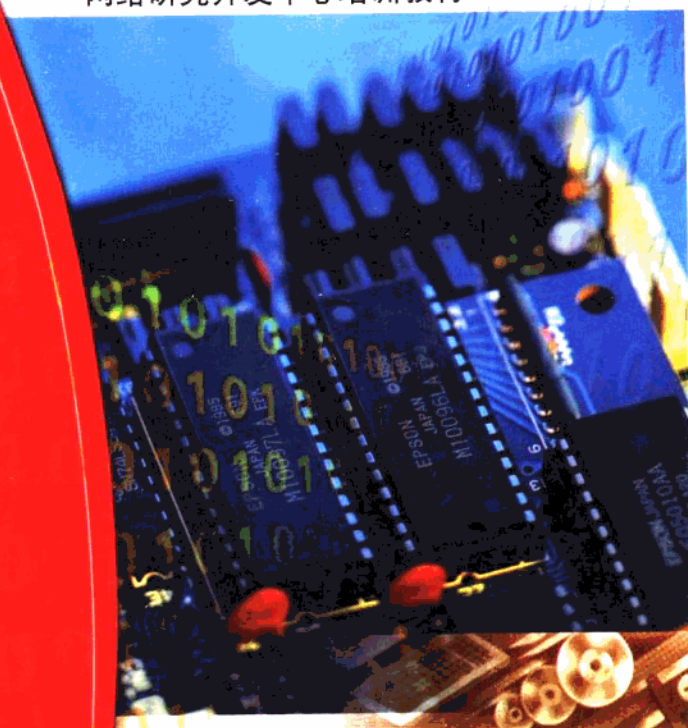


大恒图书

刘占全 张维 王强 黎连业 编著

多媒体网络 应用技术指南

中国科学院计算所(二部)
网络研究开发中心培训教材



大恒电子出版社

计算机

丛书

前 言

多媒体计算机技术是当前计算机应用的热点之一,它涉及电子、娱乐、电视传播、教育培训、印刷出版、远程通信等诸多的领域,使文字、图形、图像、动画、语音和音乐在现实的世界中变得更纷繁复杂、形形色色。作为多媒体网络技术就是研究如何表示、再现、存储、传递和处理这些信息的技术。它涉及了计算机、图形、网络、通信等不同学科的多项技术,形成了多媒体计算机原理、多媒体数据、多媒体通信、多媒体网络、多媒体文档以及表现技术等。加上新技术。新方法层出不穷,作为某个专业的研究应用人员不可能会全面地了解多媒体网络技术和应用的具体情况,因此也需要一本覆盖面大、组织结构合理、深入浅出、简明易懂、技术全面的书籍。《多媒体网络应用技术指南》就是这样一本好书。它回答了读者以下几个问题:

- 1、什么是多媒体技术?
- 2、多媒体存储介质 CD-ROM 的原理构成?
- 3、多媒体技术的应用范围?
- 4、多媒体计算机网络的要求?
- 5、多媒体技术在实现过程中使用了哪些技术?
- 6、多媒体光盘如何刻录、需要哪些硬件、软件环境?
- 7、通过本书的学习,能否就能马上制作 CD 光盘?

本书自始至终围绕着“多媒体”三字,介绍其基本知识、新技术、对网络环境的要求、光盘制作等内容,满足读者的上述七点需求。本书由九章组成,它们是:

- 第一章 多媒体计算机网络技术概述
- 第二章 多媒体通信
- 第三章 多媒体存取介质 CD-ROM
- 第四章 多媒体网络所需的性能和网络环境
- 第五章 多媒体系统的数据压缩技术
- 第六章 文件的种类与 Internet 应用
- 第七章 多媒体网络系统中的有关标准
- 第八章 国内外较为成功的多媒体应用系统介绍
- 第九章 多媒体一体化网络方案和多媒体光盘刻录

本书的特点是介绍的知识新,既讲多媒体和多媒体的应用,又讨论多媒体需要的网络环境,以及当前的热门视频点播、视频广播、视频会议,使读者通过本书的阅读对多媒体有一个全面的、系统了解。最后介绍多媒体载体—CD-ROM 光盘的刻录、制作,使读者读了这本书后,既能应用多媒体,又能制作多媒体光盘。

本书可作为计算机科研教学,计算机网络、通信等领域的科技人员、教学人员参考用书,也可作为大学计算机通信等学科的大学生、研究生的教科书。

本书的第一章、第三章的部分内容是在作者的听课笔记、读书笔记的基础上整理而成。

本书在写作过程中,参考了大量的国内外书刊、杂志、专业报纸以及国外公司惠赠的技术资料。在此向这些技术资料、文章书籍的作者、编著、译者和公司表示衷心的感谢!

参加本书写作的人员有:刘占全、刘淑芝、黎萍、王强、李淑春、刘令、王瑞、王毅、张维。李淑春同志完成了第九章的写作。黎连业同志审校了本书的全部稿件,并提出了许多宝贵意见。作者对他们表示衷心的感谢。

由于作者水平有限,对书中的错误和不当之处,欢迎读者指正。

作者

2000 年 5 月

目 录

第一章 多媒体计算机网络技术概述.....	1
1.1 多媒体是计算机网络技术发展的必然趋势.....	2
1.2 多媒体计算机技术的发展现状.....	2
1.3 多媒体数据的特点与处理技术.....	8
1.4 多媒体系统中的媒体组成.....	10
1.5 多媒体计算机系统结构和 PC 机的有关标准.....	16
1.6 多媒体网络	18
第二章 多媒体通信.....	23
2.1 多媒体通信简述.....	23
2.2 时延的有关问题.....	27
2.3 带宽的有关问题.....	29
2.4 文件传输.....	30
2.5 音频通信.....	31
2.6 视频通信.....	32
第三章 多媒体存取介质 CD-ROM.....	34
3.1 CD-ROM 的发展过程.....	34
3.2 CD-ROM 盘的构成原理.....	37
3.3 CD-ROM 驱动器简述.....	49
3.4 CD-ROM 驱动器的接口标准.....	61
3.5 CD-R 可写小型光盘.....	67
3.6 CD-ROM 驱动器和可擦写光盘驱动器的应用.....	70
第四章 多媒体网络所需的性能和网络环境.....	72
4.1 多媒体网络的性能需求.....	72
4.2 多媒体网络的环境.....	72
4.3 多媒体网络建设的建议.....	93
第五章 多媒体系统的数据压缩技术.....	94
5.1 概述.....	94
5.2 静态图像压缩标准 JPEG.....	99
5.3 运动视频和音频压缩标准 MPEG.....	104
5.4 视频信号编码简述.....	116
5.5 用于压缩的 H.320 标准总览.....	119
第六章 多媒体文件的种类与 Internet 应用.....	122

6.1 文件的种类.....	122
6.2 声音文件.....	125
6.3 MIME 文件应用.....	129
6.4 WWW 服务.....	136
第七章 多媒体网络系统中的有关标准.....	140
7.1 多媒体通信标准.....	140
7.2 多媒体技术标准.....	141
7.3 Internet 标准.....	143
7.4 其他专用标准.....	144
第八章 国外较为成功的多媒体应用系统介绍.....	147
8.1 电视会议的四个浪潮.....	147
8.2 网络视频应用要解决的问题.....	147
8.3 我国黎明公司视频解决方案的介绍.....	151
8.4 国外国内多媒体应用系统介绍.....	153
8.5 思科公司的视频介绍.....	157
第九章 多媒体一体化网络方案和多媒体光盘的刻录.....	165
9.1 “三网合一”的“一体化”网络简述.....	165
9.2 一体化网络解决方案.....	166
9.3 多媒体光盘与光盘刻录机.....	174
9.4 多媒体光盘刻录操作方法.....	178

第一章 多媒体计算机网络技术概述

1.1 多媒体是计算机网络技术发展的必然趋势

随着电子、计算机、声像和通信技术的发展,多媒体技术在计算机应用中掀起了一个不小的高潮,特别是进入 20 世纪 90 年代以来,把高清晰度电视(HDTN)、高保真音响、高性能的录像机及智能化的计算机融为一体,给多媒体系统赋予了新的内容。目前多媒体技术正在迅猛推广,全世界已形成一股研究和开发多媒体技术及产品的热潮。可以说 20 世纪 90 年代是属于多媒体技术的年代,到 21 世纪多媒体技术和应用将进入千家万户。

多媒体一词来源于视听工业。它首先是用来描述计算机控制的、多投影仪的幻灯片演示,并且配有声音通道。现在,在计算机应用领域中,多媒体是指文本(text)、图(image)、声(audio)、象(video)等这些单媒体进行组合,变为今天的数字、文字、图形、图像、视频、动画产生的数字声音、数字图象、数字视频、计算机网络集成在一起的具有人机交互功能的系统。

既然应用多媒体技术是 90 年代计算机的时代特征,是 90 年代计算机又一次革命。何谓多媒体呢?媒体(Medium)在计算机领域中有两种含义,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;一是指信息的载体,如数字、文字、声音、图形和图象。多媒体技术中的媒体是指后者。关于多媒体计算机技术(Multimedia Computing),Lippincott 和 Robinson 在 1990 年 2 月份 Byte 杂志上的两篇文章中,分别给出了不太严格的定义,概括起来是:计算机交互式综合处理多种媒体信息—文本、图形、图象和声音,使多种信息建立逻辑连接,集成为一个系统并具有交互性。

多媒体系统是信息技术发展的必然结果。促进多媒体技术趋于成熟的技术很多,其中关键的是:①CD-ROM 解决了多媒体信息存储问题;②高速计算机网使多媒体信息的实时传送成为可能;③高速微处理器技术、专用集成电路技术和亚微米半导体集成电路工业的发展,为多媒体技术提供了高速处理的硬件环境;④各种多媒体数据压缩算法、人机交互和分布式处理系统等技术使多媒体系统的产生成为可能。

我们对多媒体技术作了一番赞颂,其实它是计算机技术和网络应用发展的必然趋势。为什么我们要这样讲呢?在计算机发展的初期,人们只能用数值媒体承载信息。当时只能通过 0 和 1 两种符号表示信息,即用纸带和卡片有孔和无孔表示信息,纸带机和卡片机是主要的输入输出设备。0 和 1 既不直观,也不方便,输入输出的内容很难理解,而且容易出错,出了错也不容易发现。这一时代是使用机器语言的时代,因此计算机应用只能限于极少数计算机专业人员。扩大计算机应用的范围是一件困难的事。

50 年代到 70 年代,出现了高级程序设计语言,开始用文字作为信息的载体,人们可以用文字(如英文)编写源程序,输入计算机,计算机处理的结果也可以用文字表示输出。这样,人与计算机交往就直观、容易得多,计算机的应用也就扩大到具有一般文化程度的科技人员。这时的输入输出设备主要是打字机,键盘和显示终端。使用英文文字同计算机交往,对于文化水平较低,特别是非英语国家,普及应用仍然是件困难的事情。

80 年代开始,人们致力于研究将声音、图形和图象作为新的信息媒体输入输出计算机,这将使计算机的应用更为直观、容易。在 1984 年起“领头羊”的 Apple 公司的 Macintosh 个人计算机,首先引进了“位映射”的图形机理,用户接口开始使用 Mouse 驱动窗口技术和图符(Window and Icon),受到广大用户的欢迎。这使得文化水平较低的公众,包括儿童

在内都能使用计算机，由于 Apple 采取发展多媒体技术、扩大用户层的方针，使得它在个人计算机市场上成为当时一家重要的大公司。经过短短几年的努力，国际上在下述几项技术又有了突出的进展：

- ①超大规模集成电路的密度增加了 16 倍；
- ②超大规模集成电路的速度增加了 8 倍；
- ③CD-ROM 可作为低成本、大容量 PC 机的只读存储器；
- ④双通道 VRAM 的引进；
- ⑤网络技术的广泛使用。

这五项计算机基本技术的进展，有效地带动了数字视频压缩算法和视频处理结构的改进，促使八十年代中期的单色文本/图形系统转变成今天的彩色丰富、高清晰度显示系统，同时能够做到全屏幕、全运动的视频图象，高清晰度的动态图象，视频特技，三维实时的全电视信号以及高速真彩色图形。同时还有高保真度的音响信息。尤其是 Internet 的迅速发展，把多媒体的应用推上了历史舞台。人们可以在网上购物，网上看时事新闻、体育比赛、电视会议等，拓宽和加深了计算机应用的广度和深度。我们有理由说：无论从半导体还是计算机技术进步的角度，或是从普及计算机应用、拓宽计算机处理信息类型看，利用多媒体是计算机技术发展的必然趋势。许多计算机应用的行家都认为多媒体技术可以被看成是实现图象和语音识别之前的过渡技术。它充分发挥了计算机运算快、综合处理能力强等优点，用交互式技术来弥补目前计算机对于图象和语音理解和识别的不足。就目前来看，多媒体计算机技术最大贡献是改善了人机接口界面，拓宽了计算机的应用领域；从长远来看，它有可能对计算机机理和体系结构产生深远的影响。

1.2 多媒体计算机技术的发展现状

多媒体计算机技术从 80 年代中期开始，至今已有十多年的历史。在这十几年间，世界上很多国际性的大公司都在研制开发多媒体计算机技术，其中卓有成效影响较大，成绩斐然的公司现简述如下：

1、Commodore 公司的 Amiga 系统

Commodore 公司在 1985 年率先在世界上推出了第一个多媒体计算机系统 Amiga。在 1989 年秋美国的 Comdex 博览会上，Commodore 公司展示了 Amiga 系统一个完整的系列。到目前为止，该公司已推出 Amiga500，1000，2000，2500 以及 3000 等多种型号的产品。提高视频和音响信息的处理速度，Commodore 公司在 Amiga 系统中采用了三个专用芯片：即 Agnus (8370)，Paula (8364) 以及 Denise (8362)。现对它们作一简要介绍。

(1)Agnus(8370)是专用的动画制作芯片，芯片中有 5 个 DMA 控制逻辑：视频游标 DMA，音频 DMA，位平面 DMA，软盘和刷新电路 DMA 以及位映射控制部件的 DMA 控制逻辑线路及其需要的控制寄存器，它们通过内部总线与专用芯片内部的图形协处理器连在一起。作为 Agnus 的功能主要表现在：

- ①用硬件显示移动数据，允许高速的动画制作；
- ②显示同步协处理器；
- ③控制 25 个通道的 DMA，使 CPU 以最小的开销处理磁盘、声音和视频信息；
- ④从 28MHz 振荡器产生系统时钟；
- ⑤为视频 RAM (VRAM) 和扩展 RAM 卡提供所有的控制信号；
- ⑥为 VRAM 和扩展 RAM 提供地址。

(2)Paula (8364) 是专用音响处理及外设接口芯片，芯片中音响处理器、磁盘控制器、异步通信接口以及电位计通道接口都连接到内部总线的设备码译码器上。音响处理器是由 2 路

数据寄存器、两个音响控制计数器及 4 路 D/A 变换器组成。它可以通过 DMA 的方式和 Amiga 系统的存储器以及其他设备交换音响信息，在 Paula 的音响处理器中处理音响信息，最后经过 D/A 变换器，可把 4 路两对立体声信号输出到音响设备中。盘控制逻辑也通过 DMA 的方式将 Amiga 系统中存储器数据，通过盘控制器输出到盘上；反之可将盘上数据通过盘控制器读入到 Amiga 系统的存储器中。此外，还有异步通信接口和电位计通道控制逻辑，都以 I/O 方式进行数据传输。该芯片的主要功能是输出 4 路两个立体声道、9 个八音阶，使用音频放大和频率调制，还有异步通信接口、盘控制器以及电位计通道接口。

(3)Denise (8362) 是专用的图形芯片，它有：位平面数据寄存器，位平面控制以及位平面串行输出器；硬件游标数据寄存器，硬件游标串行连续化器以及位置比较逻辑；碰撞控制逻辑、碰撞检测逻辑以及碰撞存储逻辑；优先排队控制逻辑以及位平面排队和控制寄存器；彩色选择译码器以及 32 位彩色输出寄存器；Mouse 计数器。由上述可见，它就是多功能的彩色图形控制器，它可以控制不同分辨率的输出，从 320×200 到 640×400 ；在电视机和 RGB 彩色监视器的屏幕上可同时显示 4096 种颜色；有 8 个可重复使用的“硬件游标”控制器。

为了适应不同用户对多媒体技术的需要，Commodore 公司提供一个多任务 Amiga 操作系统，它有上下拉的菜单，多窗口，图符 (Icon) 以及 PM (Presentaion Manager) 等功能。同时，配备了大量应用软件，如能绘制动画，制作电视片头及作曲等专用软件。尤其是 Amiga Vision 多媒体的著作系统，为用户提供一个完备的图符编程语言 (A Complete Iconic Programming Language)。

2、Philips/Sony 公司的 CD-I 系统

Philips/Sony 公司于 1986 年 4 月公布了基本的 CD-I 系统，同时还公布了 CD-ROM 的文件格式，这就是以后的 ISO 标准。该系统把高质量的声音、文字、计算机程序、图形、动画以及静止图象等都以数字的形式存放在容量为 650MB 的 5 英寸只读光盘上。用户可通过与该系统相连的家用电视机、计算机显示器和 CD-I 系统进行通信、使用鼠标器、操纵杆或遥控器等定位装置选择人们感兴趣的视听材料进行播放，可完成培训或教育任务。对于 Philips/Sony 公司的 CD-I 系统中期的概况简述如下：

CD-I 基本系统的结构如图 1-1 所示，整个系统也称 CD-I 译码器。该系统可分成两部分：一部分是 CD-ROM 驱动装置，它有 CD 驱动器，可以使用 CD-I 光盘或 CD-DA 光盘。CD-ROM 驱动器中还有 CD-ROM 译码器，通过它把 CD-ROM 驱动器连接到系统总线。另一部分是多媒体控制器 MMC，它由音频信号处理器，视频信号处理器，68000 微处理器，RAM，ROM，不挥发的 RAM 以及定位装置组成。

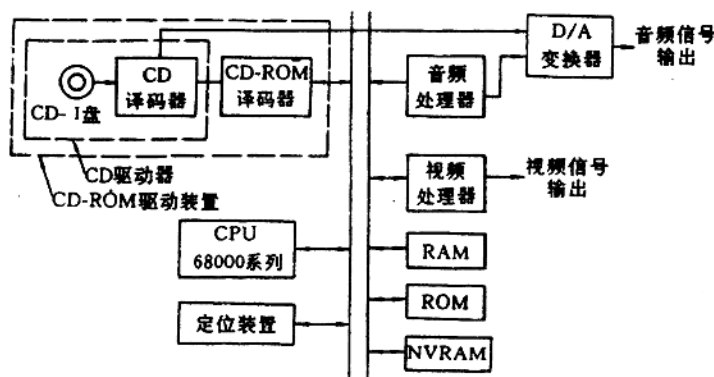


图 1-1 CD-I 基本系统结构

从 CD-I 系统结构图很容易看出，它有两种工作方式。一种是不需要其他计算机，CD-I 系统与家用电视机、录象机及音响设备连接在一起，在紧凑光盘实时操作系统的管理控制

下, 编译来自光盘的音频、视频和程序数据, 并把声音图象数据分别通过音频处理器和视频处理器送给音响设备和电视机或录像机, 用户可通过鼠标或操纵杆定位装置移动显示屏幕的游标, 向 CD-I 系统发出指令, 运行交互式的培训或教育节目。另一种方式是 CD-I 基本系统可以作为多媒体控制器连接到其他微型计算机、工作站以及小型计算机上。

CD-I 基本系统提供 4 种不同音质的运行方式, 如表 1-1 所示, 一种选用 CD-DA 光盘,

表 1-1 CD-I 音频方式

级别	采样率 (kHz)	位数/ 样本	频率响应 (kHz)	数据速率 (字节/ 秒)	通道数	数据流 百分数	播放时 (小时)
CD-DA PCM 超级 HiFi	44.1	16	20	171,100	1 立体声	100%	1
C D I	A (LV)	8	17	85,100	2 立体声	50%	2
				42,500	4 单声道	25%	4
A D P	B (FM)	4	17	42,500	4 立体声	25%	4
				21,300	8 单声道	12.5%	8
C M	C (AM)	4	8.5	21,300	8 立体声	12.5%	8
				10,600	16 单声道	6.25%	16

它采用脉冲编码调制 (PCM) 系统, 直接通过 CD 译码器, 输出到 D/A 变换器, 再送给音响设备输出, 会得到超级 HiFi 音响效果; 其它 3 种是选用 CD-I 光盘, 它采用自适应差分脉冲编码调制 (ADPCM), 选用不同的采样频率、量化精度, 使得 A 级相当于 Laser Vision 的音质, B 级相当于 FM 调频广播的音质, C 级相当于 AM 调幅广播的音质。基本系统还定义了 3 种不同的图象分辨率 (如表 1-2 所示) 以及几种视频工作方式 (如表 1-3 所示)。

表 1-2 CD-I 图象分辨率

	NTSC	PAL/SECAM
正常分辨率	360×240	384×280
双倍分辨率	720×240	768×280
高分辨率	720×480	768×480

表 1-3 CD-I 视频方式

方式	位数/像素	显示颜色 数目	数据量/幅画面 (千字节)		
			360×240	720×240	720×480
DYUV	16/2	16,772,216	86.4	—	345.6 *
RGB 5:5:5	16	32,768	172.8	—	—
CLUT8	8	256	86.4	—	345.6 *
CLUT7	7	128	87.4	—	345.6 *
CLUT4	4	16	—	86.4	172.8
RL7		128		—	345.6 *
RL3		8		86.4 以下	172.8

*表示 CD-I 基本系统不提供

<NTSC 方式>

为了改进 CD-I 基本系统的特性, Motorola 公司为 CD-I 系统开发了一套新的专用电路,

sony 公司为 CD-I 扩充硬件, 增强功能的 CD-I 系统如图 1-2 所示。该系统微处理器采用了 Motorola 公司高性能的嵌埋式微处理器 MC68340, 同时 Motorola 公司为 CD-I 所开发的大规模集成电路专用芯片有: 视频系统控制器、视频合成器、全运动视频信号控制器以及视频信号 D/A 转换器。其中全运动视频信号控制器是一个复杂的专用芯片, 它处理视频信号的压缩编码和解压缩问题, 为 TV 提供全屏运动的图像; 视频系统控制器主要用于内存管理; 视频合成器主要处理位映射图像; 视频信号 D/A 转换是对数字式视频信号进行模拟量的转换, 送给 RGB 驱动电路, 供给彩色监视器。此外, 连接 CD 驱动器选用了 56001 数字信号处理器, 主要用它处理语音信号, 再通过静态存储器以及串行音频 D/A 变换, 分左右两个通道输出到音响设备中去。其他还有: EPROM, 不发挥的 RAM 以及视频帧存储器 DRAM。从上述结构我们可以清楚地看到, 增强型的 CD-I 系统在全屏幕运动视频及音响处理方面比起 CD-I 基本系统有了较大的改进。

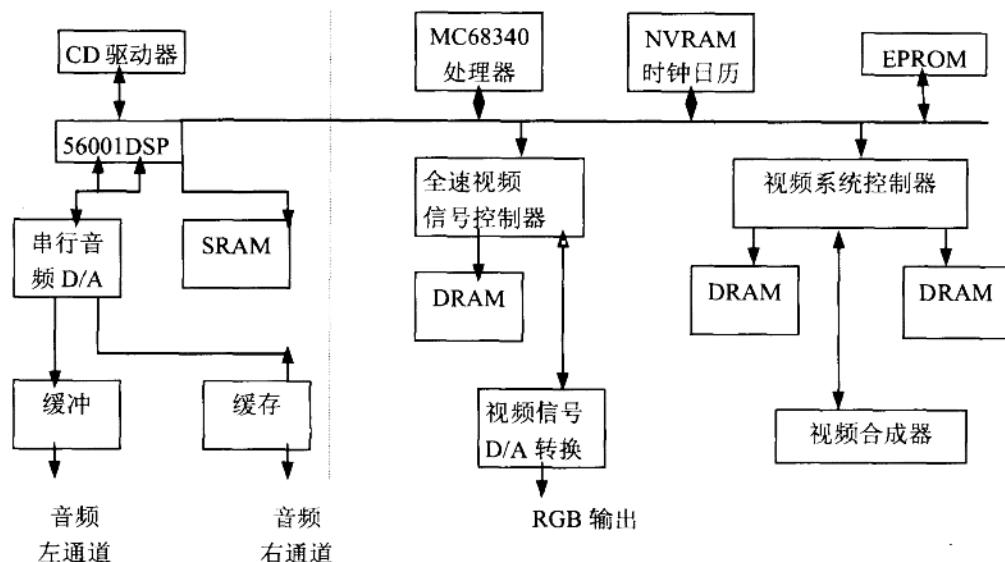


图 1-2 增强功能的 CD-I 系统

3、Apple 公司的 Hypercard

Apple 公司的 Macintosh 系统具有公认的良好图形特性, 它是桌上出版和桌上展示系统的先驱。Apple 公司的多媒体系统也有人称之为桌上媒体, 它实质上是把高质量的音响及活动的视频图像加到原来的 Macintosh 系统中, 能够把上述特性连在一起的是 HyperCard 及其兼容软件。HyperCard 是以卡片 (Card) 为节点的超级文本 (Hypertext) 系统, 基本的信息单元是卡片或称节点, 一个卡片可充满整个屏幕。一组卡片称为卡堆 (Stack), 可以认为卡堆是 HyperCard 中的文件, 同类和相关的卡片可在一个卡堆内。每个卡片不仅是字符, 还包括图形、图像和声音。HyperCard 系统提供了许多命令或工具, 通过鼠标器或键盘实现控制完成卡片的浏览、编辑、制作, 信息的输入、修改、检索。它能把简单的数据库、复杂的文本程序、编程语言及著作系统组成一个快速灵活的软件包。HyperCard 的数据库和所有的 MAC 的数据格式兼容, 并开发有直接的连接电路、光扫描器以及 CD-ROM 驱动器连接。为了使 HyperCard 和这些外部设备相连接, Apple 公司已经公布了一个多媒体协议和驱动程序标准集, 叫做 AMCA (Apple Media Control Architecture)。AMCA 是系统级的结构, 用来访问视频光盘、音频光盘以及录像带的信息, 软件工作人员不用为多媒体外部设备写专门的驱动程序。

Apple 公司原来选用 Mac SE 和 Mac II 作为多媒体计算机的平台, 现在选用了 68030 微处

理器作为 CPU，直接寻址最多可安装 8MB 内存，视频适配器板可在 16M 种不同颜色中同时显示其中 256 种颜色。音响媒体接口板和 HyperCard 软件兼容，能够提供良好的语音、音响效果，通过语音分析和识别能够代替键盘、鼠标以及操纵杆的功能。

为了快速、实时地处理视频和语音信号，Apple 公司正在和 MIT 的媒体实验室合作，组成新一代技术研究小组开发视频和音频信号压缩编码和解码技术。为了传输视频信号，他们提出了高速的宽带网以及对称的压缩编码和解码技术，并已研制出了这种样机。

4、Intel 和 IBM 公司的 DVI 系统

DVI 技术 (Digital Video Interactive) 于 1983 年在 RCA 公司的戴维·沙诺夫研究中心 (David Sanaoff Research Center in Princeton, New Jersey) 开始了开发工作，在 1987 年 3 月第二次 MicroSoft CD-ROM 会议上，首次公布了 DVI 技术的研究成果，1988 年 10 月 Intel 公司从 GE 公司买来了 DVI 技术，1989 年 Intel 和 IBM 公司在国际市场上推出了 DVI 技术第一代产品 Action Media 750，1991 年又在美国 comdex 展示会上推出了第二代的 DVI 技术的产品 ActionMedia 750 II。

DVI 技术第一代的产品 ActionMedia 750 I 的结构图如图 1-3 所示，其核心是两块专用的 DVI 接口板：DVI 视频板、DVI 音频板以及 DVI 多功能板。IBM PC/AT，386，486 或其兼容型计算机作为工作平台，同时配有 CD-ROM 驱动器，带有放大器和音响效果的 RGB 彩色监视器，组成了 DVI 用户系统。在此基础上再配置与多媒体有关的外设：视频信息数字化器（连接到 DVI 视频板上），音响信号数字化器（连接到 DVI 音响板上），扩展的视频 RAM（连接到视频板上），大容量的光盘或硬盘、磁带机、录像机、音响设备、监视器以及摄像机或扫描仪等组成 DVI 开发系统。

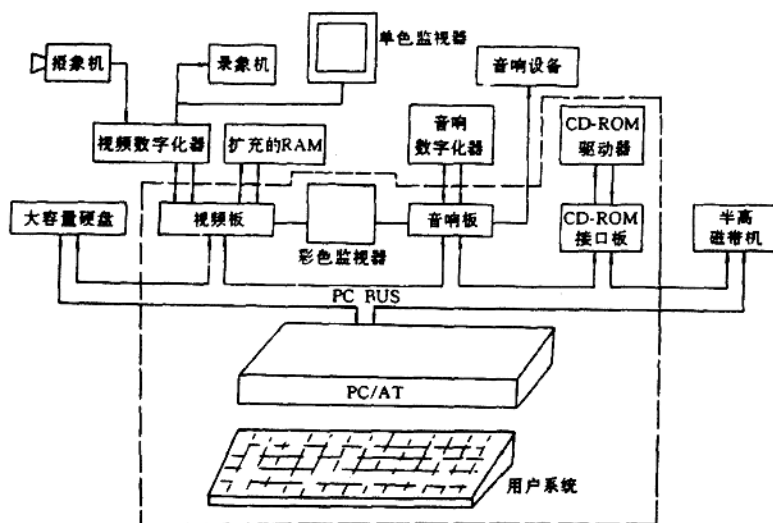


图 1-3 DVI 系统结构图

DVI 技术硬件的核心部件是 Intel 公司生产的专用芯片：VDPI (82750PA, 82750PB) 和 VDP2 (82750DA, 82750DB)，称之为视频象素处理器和视频显示处理器。A 型提供 12.5M/s 操作速度，B 型提供 25M/s 操作速度。82750PA/PB 是象素处理器，采用微码编程，可以高速执行象素处理的各种算法。82750DA/DB 是显示处理器，它可与 82750PA/PB 并行处理、显示处理好的帧存储器中的位映射图。它具有较强的图形功能，同时通过编程，适应不同分

辨率、不同像素格式及不同同步格式的多种型号的显示器。

DVI 技术软件的核心部件上 AVSS (Audio/Video Sub System) 和 AVK (Audio Video Kernel)。AVSS 是在 DOS 环境下, 加上 RTX (实时执行部件)、视频驱动器、音响驱动器, 多功能驱动器以及驱动器接口模块, 运行音响视频的子系统。AVK 在 Windows 环境下运行, 因此它就不局限在 DOS 操作系统环境, 可以在其他种类的操作系统环境下运行。AVSS 和 AVK 最主要的任务是: 为音频和视频数据流相关同步提供需要的实时任务调度, 实时的数据压缩和解压缩, 实时地拷贝和改变比例尺, 建立位映射, 管理控制它们并将其送至显示缓冲区等。

Intel 公司 DVI 产品发展规划如表 1-4 所示, 1989 年到 1991 年推出其第一代产品 AM750 I, 1991 年到 1993 年将推销第二代产品 AM750 II, 1992 年到 1995 年将推销第三代产品 S3 (性能指标详见表 1-4)。第二代的产品性能指标比第一代产品提高了一倍, 价格仅是原来的三分之一。Intel 公司副总裁 David House 预测, 1995 年 Intel 公司要把 DVI 技术放到母板上, 2000 年时放在芯片里。House 说“从 71 年 Intel 4004 问世以来, 每 18 个月 CPU 中晶体管的数量增加一倍, 到本世纪末 CPU 中晶体管的数目可达到 0.5-1 亿。”他还说: “在本世纪末, Intel 公司生产的 Intel786, 运算速度可达 2000MIPS, 包括 DVI 运动视频, 并和 386 结构兼容。”

表 1-4 Intel 公司 DVI 产品发展规划

	第一代产品 1989-1991	第二代产品 1991-1993	第三代产品 1992-1995
专用芯片	i750A	i750B	V3
硬件板	AM750 Series I	AM750 Series II	S3
软件	AVSS-DOS (2.13)	AVSS-DOS (2.2)	AVK-Window AVK
压缩 算法	PLV1.0/1.5 RTV1.5 PIC1.0	PLV2.0 PIC2.0 (JPEG)	PLV2.0 RTV2.0 PIC2.0 (JPEG) MPEG PX64 PIC3.0 (JPEG) PLV3.0
工具	MDT-Limited MPT.Lumena Authology	MDTLimited MDTLimited Third Party Tools	V3 Tools

目前 DVI 技术有丰富的软件支持, 例如有工具软件、库函数。到 91 年底, 全世界已有 80 多家厂商为 DVI 技术编制开发工具和各种应用软件。其中, 美国有 50 多家, 欧洲有 10 家, 此外还有加拿大和日本。例如美国 Time Arts 公司编制的 LUMENA 图象编辑工具, 以及很多著作语言都为用户提供了极大的方便。

全世界还有很多厂商和公司在从事多媒体计算机技术的研制、开发以及设计制造工作, 最值得一提的是 Next 计算机公司, 被称为“计算机奇才”的斯蒂夫·乔伯 (Steve Jobs) 是该公司的总裁, 他设计的 Next 计算机从一开始就考虑了多媒体技术的需要, 它的设计思想是卓越的。该机利用 PostScript 实现高级绘图功能, 采用 DSP 进行高保真的声音信号处理, 提供图象和声音的 Nextmail 邮件系统等。

1.3 多媒体数据的特点与处理技术

1.3.1 多媒体数据的特点

多媒体数据大体可分为 5 种状况, 现简述如下:

1、数据类型复杂

多媒体数据实际上是由多种不同类型数据组成的。通常包括正文、图形、图像、声音、视频图像、动画等不同数据类型, 而且同一类型数据可以有不同表示。例如, 可以用编码形式表示, 也可用二进制非编码形式表示; 可以用内部数据结构表示(如图形数据常用图段、层次、边界、结构几何等多种不同数据结构表示); 也可采用无结构的位图表示形式, 特别是, 这些内部数据结构都随具体应用而变。多媒体数据这一复杂性不仅使多媒体数据的存储、检索以及建立多媒体数据的处理技术各不相同, 而且使多媒体计算系统功能较普通微机、工作站的功能要复杂得多。

2、数据信息量大

现以声音和视频图像数据为例加以说明。对声音数据进行采样并量化时, 根据人耳能感知的最低频率为 20KHz 这一特点, 通常采用 44.1KHz 的采样频率, 而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 176K 字节或每分钟 10M 字节。图形(图像)和视频图像的信息量与屏幕分辨率(x 方向像素数乘以 y 方向像素数)、表示每一像素的信息量(通常用二进制位数表示)、帧刷新频率以及是否压缩等因素有关。多媒体数据的大信息量的特点导致一系列技术难点需要解决, 主要包括: 高速处理器技术, 大容量廉价存储技术(包括内存、帧存和外存), 具有高压缩比的实时图像数据压缩和解压缩技术, 以及高速通信网络技术等。

3、数据的实时性

多媒体数据中的声音和视频图像数据都是与时间有关的信息, 很多场合要求实时处理, 如声音和视频图像信息的实时压缩与解压缩、传输与同步等。另外, 在交互操作(如编辑、检索、显示等)方面要求的实时操作系统支持。因此, 多媒体计算机系统要有很高的运算速度, 除通用的高速处理器芯片技术外, 很多算法均需要专用硬件支持。因而, 高速专用集成电路是多媒体计算机的重要组成部分, 例如 Intel 公司的 DVI 采用 82750 专用处理器芯片, 不仅增强了压缩和解压缩处理的实时性, 还使系统软件的实时性设计得到简化。

4、数据的分布性

由于多媒体数据的多样性, 多媒体应用的开发工作要求各种专业人员介入, 包括计算机开发、文字写作、影视制作、广告宣传以及动画设计等人员的协同工作, 因而原始素材往往分布在不同空间和时间里, 使得分布式多媒体数据库的建立和管理以及多媒体通信成为多媒体计算系统的关键技术。

5、数据的交互性

多媒体技术的特点之一就是有很强的人机交互性, 这也是它有别于传统声像技术之处。在多媒体技术的实际应用中, 主要方法是“选择和观察”。例如, 应用鼠标器单击屏幕上的文字, 调用实物图片或解释性的视频图像片断, 或调用其他背景材料供用户观看和决策, 也可以用鼠标单击图像或视频图像上的某一区域, 调用有关文字说明材料或其他图像、声音等材料供用户观察, 按用户所希望的顺序重新组织有关材料。这种应用方法的基础就是人机交互技术, 这种交互操作是一种实时操作, 要求整个系统的软/硬件系统都能实时响应。

1.3.2 多媒体数据处理技术

多媒体数据在处理技术上大致可分为 4 种情况, 现分别简述如下:

1、多媒体信息的表示及压缩编码技术

随着多媒体技术的发展,计算机中信息的类型不仅有正文、数字,而且还有图形、声音、图像、动画、视频等。它们也被称为基本信息类型。信息的表示主要分为两种方式,即模拟方式和数字方式。在多媒体计算技术中都采用数字方式。鉴于数字化多媒体信息量大,因而必须对多媒体数据进行压缩。目前存在多种图像信息(静止、活动)和声音信息的压缩标准,在多媒体技术领域通常采用 JPEG 和 MPEG 两种标准。前者用于静止图像,有失真的压缩比可达 50:1;后者用于视频图像及其伴音,在允许噪声存在的前提下,其压缩比亦可达 50:1。若将这一技术用于每秒 27M 字节的 $640 \times 480 \times 24$ 的动态图像,则可压缩到每秒 550K 字节。

2、多媒体信息存储技术

微机数据需要量以成倍的速度增长,虽然硬盘的容量越来越大,但依然满足不了用户的要求。尤其是,随着多媒体技术的日益普及,几百兆甚至上千兆容量的硬盘已难以容纳下多媒体程序运行时所需要的图形、图像、声音和音乐等庞大的数据文件。数字化的多媒体对存储技术提出两方面的要求:其一是大容量存储技术,其二是足够的数据传送带宽的支持和多媒体的实时处理功能。

3、多媒体数据库技术

多媒体数据类型不同,表示方式也各不相同。当使用数据库技术来支持多媒体应用时,需要将多媒体数据对象各种表示的固有特性(如是否采用编码形式或结构形式等)映射到相应的表示形式,如正文文件、图像参数文件、图像数据文件、图形结构表等等。多媒体数据库应能处理数据对象的上述各种表示方式,包括很多复杂数据对象是由异构的子对象组成的情况,例如在图形上叠加图像等。

不同对象表示形式、存取方式、绘制方法等各不相同,因此,多媒体数据库还应包括处理不同对象的相关方法库。多媒体数据库与方法库紧密相关联,以便进行数据对象的组合、分解和变换等操作。另外,为了管理数据对象方便起见,应建立数据对象的说明,以便于定义数据对象的二级属性。因此,数据对象、数据对象的说明以及与对象相关联的方法是多媒体数据库的三个组成部分。

除了管理数据类型复杂外,多媒体数据库的另一特点是存在着时间上的限制,这主要是指实时性和同步要求都很严格。

上面仅仅提出了建立多媒体数据库的一些具体要求,完全是概念性的,并未与哪一种数据库理论联系在一起。事实上,上述概念性要求完全可以采用现有的关系数据库或面向对象数据库来实现。

4、多媒体数据通信技术

多媒体数据的分布性以及计算机支持的协同工作(CSCW)等应用领域均要求在计算机网络上传送声像数据。这里,传输速率不是问题的本质,因为无论是每秒 10M 位的以太网,还是每秒 100M 位的光纤网(FDDI)都能满足传输压缩的视频图像数据的要求(如 $640 \times 480 \times 24$ 的动态图像可压缩为 550Kbps)。现有局域网传输声像数据所遇到的问题是多媒体数据在时间上是连续的,因此要求不间断地传输。在 CSCW 应用中,对同时在网络上传输多路双向声音和图像的要求更高,因为在一个会议室里可以有多个摄像机、监视器和话筒同时发送和接收声像数据。现有局域网是基于各结点可共享网络带宽的思想设计的,它假设各结点间传输的数据在时间上是相互独立的,从而可以把数据打成包,分别传送。因此从这个观点来看,现有局域网技术不符合多媒体通信要求。新的国际通信标准 B-ISDN(宽带综合服务数字网)的异步传输功能能满足多媒体通信的要求。除了带宽问题外,多媒体通信技术中仍有许多特殊问题需要解决,例如:相关数据类型的同步,多媒体设备的控制,不同终端和网络服务器的动态适应,超媒体信息的实时性要求,可变视频数据流的处理,网络频谱及信道分配;高性能和高可靠性以及网络和工作站的连接结构等等。

1.4 多媒体系统中的媒体组成

1.4.1 媒体的分类

在前面我们已经叙述过了多媒体系统中的文字、图形、图像、音频、数据的特点和数据处理方法等，但是媒体又可以根据不同的标准进行分类，目前把媒体分为感知媒体、描述媒体、表示媒体、存储媒体、传输媒体和信息交换媒体。

1、感知媒体

人们借助感知媒体来接收环境的信息。在计算机环境中，人们主要通过视觉和听觉器官来感知外部世界。视觉媒体表现为文字、图像和视频，听觉媒体表现为音乐、噪声和语音等。

2、描述媒体

描述媒体的特征是信息的计算机内部表示方法。在计算机中使用不同的格式来描述媒体信息或对媒体内容进行编码。其编码方法主要有：

①正文字符用 ASCII 或 EBCDIC 编码。

②图形可根据 CEPT 或 CAPTAIN 视频正文标准编码，图形标准 GKS、PHGS 和 CGM 也可用来对图形信息进行编码。

③音频流可用简单的 PCM (Pulse Coding Method) 脉冲编码表示。

④图像可用传真标准或 JPEG 格式进行编码。

⑤组合的音频/视频序列可用不同的 TV 标准格式（如 PAL、NTSC 或 SECAM）进行编码，在计算机中存储时则可以用 MPEG 格式进行编码。

3、表示媒体

表示媒体指的是通信中电信号和感知媒体之间转换用的媒体。表示媒体有两种：一种是面向输入的表示媒体，如键盘、摄像机、光笔、话筒；另一种是面向输出的表示媒体，如显示器、打印机等。

4、存储媒体

存储媒体是指保存媒体数据的介质，主要例子有缩微胶片、磁盘、CD-ROM 等。

5、传输媒体

传输媒体是指传送连续媒体数据的信息载体，它与存储媒体含义不同，表示媒体可以在网络（同轴电缆、光纤）上进行传输。

6、信息交换媒体

信息交换媒体是指用于信息存储和传输的媒体，它包括信息传输和信息存储两方面的含义，例子有电子邮件系统。

1.4.2 多媒体的文本格式

文本也称文字、正文。是人们与计算机之间进行信息交流的主要方式。在计算机发展的早期，使用终端为一般的字符终端，屏幕上显示的是文字信息。由于人们在现实生活中经常用语言进行交流，所以开始时也觉得比较方便。后来出现了图形、图像、声音等，这样就相继有图形终端、图文终端、多媒体终端等。在现实世界中，文字是人们之间进行异步通信的主要形式，例如书、报纸、信、电子邮件等。作为文本格式有两种情况。即正文格式和文档格式。它们分别又有各自的两种状况，现简述如下：

1、正文格式

正文格式分为未格式化正文和格式化正文

(1)未格式化正文 (unformatted text, 也称 plain text)

组成正文的字符来自于特定的字符集，对于西文而言，字符集中的字符数目相对较少，而汉字字符集中的字符数目就很多，未格式化文字中的字符有固定的大小，并且风格单一。

这种文字形式可读性相对较差, 由于不包括定义格式的命令, 故交换信息比较容易。

未格式化正文用于书写计算机程序或向计算机提交命令。

(2) 格式化正文 (formatted text, 也称 rich text)

格式化正文对正文中的字符、字号和格式都有一定的规定, 例如章节标题用大号黑体, 而参考文献中的期刊名用斜体表示, 从而增加可读性。格式化文档在显示时, 所得到的结果与打印输出所得到的结果类似, 也就是所谓的所见即所得 (WYSIWYG: What You See Is What You Get)。格式化正文中的格式定义多种多样, 如多列正文、居中、对齐等。

2、文档格式

格式化正文为书写、保存、显示、打印具有一定格式要求的文档提供了可能性。文档可以是文章、书籍等。格式化正文具有内部结构, 如它们通常由标头、题目、章、节、表、脚注、参考文献和索引组成。为了在计算机上处理格式化文档, 必须具有描述文档结构的能力。描述这种结构的规则集就称为文档格式。文档格式主要分为两种: 结构描述格式和页面描述格式。

(1) 结构描述格式

文档除了正文体外, 还包括控制信息, 控制信息用来标识不同的结构部件 (如章、节)。这样, 在文档真正输出时, 就会自动产生特定的文档元素 (如脚注号、索引号等)。一般来说, 结构描述格式定义的文档作者如何指定结构的元素 (如一新章或索引的开始) 以及结构信息编码的方法, 这种格式既可以只规定文档的逻辑结构, 也可以规定文档的逻辑结构和布局结构。结构描述格式的例子有 Tex、Letex、SGML 和 HTML 等

(2) 页面描述格式

这些格式也称为页面描述语言, 它们使用一种编程语言来描述可打印文档的页面。尽管我们可以手工编程, 但绝大多数情况下是用一些特定的应用系统自动生成。它们必须由打印中的解释器进行解释执行, 从而产生输出。当然, 解释器也可以用软件仿真, 软件在一般的微机系统上执行, 从而可驱动显示屏在显示器上产生仿真输出, 或在多种不同的打印机上产生输出。页面描述格式通常也可以描述图形和图像等其他元素。使用最为普遍的页面描述格式 (页面描述语言) 是 Adobe System 公司的 PostScript。

1.4.3 图形

对于图形我们从图形的特性、格式、图形与图像关系来叙述。

1、图形特性

图形是可修改型“文档”, 原因是它们的格式保留在结构信息中。图形可以由用户借助于基于计算机的图形编辑器来创建, 也可以通过程序自动生成。图形画面由线、曲线或圆之类的对象组成。这些对象可以被检测到, 可以进行有关编辑操作 (删除、增加、移动、修改、旋转和延伸等)。图形对象的一些属性 (如线宽、颜色、填充式样和线型等) 也是可修改的。图形有时也称为向量图形或矢量图形。

2、图形格式

在表示图形时, 通常保留了图形的语义内容。PHIGS 和 GKS 是两个常用的图形标准。如果在一幅图的画面中也包含有正文信息, 那么这些正文信息通常属 rich text, 即格式化正文, 它们也是可修改的, 它们的正文属性也可适当改变。正文由图形编辑器 (或图文编辑器) 中的文字编辑功能来生成。

3、图形与图像的关系

图形和图像都是非正文信息, 它们可以被显示或打印。我们可以把图形和图像显示到计算机的屏幕上, 也可以把它们输出到打印机上。但是, 如果相应的设备只能处理字符 (如字符终端), 那么就不能输出图形或图像。图形类似于书或杂志中的线画图, 而图像则类似

于书中的照片。正常情况下,对不涉及到计算机的线画图和连续色调的照片之间的差别很明显;如果涉及到计算机,那么图形和图像的差别就没那么明显。表 1-5 对图形和图像进行了对比。

表 1-5 计算机图形/图像对比

图形	图像
可以修改	可以修改
文档格式包含结构信息	文档格式不包含结构信息
在表示中保留了语义内容	语义内容并未保留
描述成对象	描述成位,用一个个像素表示

1.4.4 图像

1、图像特性

图像是一种可修改的文档,但对应的文档格式不包含结构信息,图像可能起源于真实世界(如用照相机拍摄的照片),也可能由计算机生成。从真实世界产生数字化图像的方法有以下三种:

- (1)图像的扫描输入
- (2)帧捕捉设备从视频序列捕捉到的图像帧
- (3)来源于 CD-ROM 或电影胶片

上述这些图像在生成过程中都要进行数字化处理,所以称为扫描静止图像。另一方面,图像也可以借助计算机来生成。例如用图形编辑器编辑生成一幅图形,把它按图像格式保存,即把图形转换成图像;另外一个例子是用计算机生成三维真实感图形,有照片的效果,这样的高质量图形也可保存为图像,我们把这类图像称为合成静止图像。产生数字化静止图像的方法如图 1-4 所示。

图像可以用图像编辑器进行处理。图像编辑器和图形编辑器类似,但图形编辑器中操作的对象可以是一些基本图元(如圆、曲线),而图像编辑器中操作的对象只是最基本的像素,这样由图像编辑器产生编辑结果(文档)就不包含语义信息。图像编辑操作的种类很多,有些非常简单,有些也比较复杂。复杂的例子有:从几个画面创建复合图像,图像叠加,颜色修改,变形,对图像的某些部分进行拉伸等。

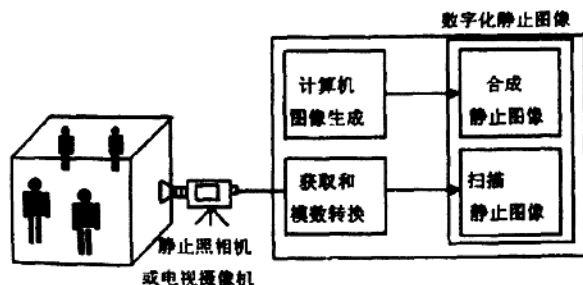


图 1-4 生成数字化静止图像的方法

2、图像格式

计算机图像以位图的形式表示。一个简单单位图可看成是一个二维矩阵,每个矩阵元素是像素。像素是表示图像的最基本元素。表示每个像素的数值称为幅值。对一个像素进行编码的可用位数也称为幅值深度或像素深度。图像的典型像素深度为 1、2、4、8、12、16 或 24 位。当像素深度为 1 位时,表示黑白图像,只有两种颜色;如果像素深度为 2,那么对应的

图像可有 $2+2=4$ 种颜色。像素的数值可表示二值图像中的黑白点，也可以表示连续色调单色图像中的灰度，当然也能表示彩色图像中的彩色属性。

虽然图形或正文是可修改的文档，但是它们也可转换为位图格式，并以图像的方式来保存。把计算机图形转换为位图也是产生合成静止图像的一种手段。在把图形转换为图像时，图形画面的一些结构信息就会完全丢失。下面举一个简单的例子来说明。设图形画面只由一根黑线组成，这个画面使用这根线的空间坐标来描述，在转换为图像后，“画面中有一条黑线”的信息就不复存在，图像只由一系列的像素点组成，在保存图像时，必须分别保存这些像素点。典型的图像格式（也称位图格式）标准有 TIFF 和 GIF 等。表 1-6 总结图像和图形的类型及来源。

表 1-6 图像和图形的类型及来源

类型		来源		
名字	存储形式	名字	特征	产生技术
静止图像	位图	扫描静止图像	从真实世界捕捉	• 打印页面的扫描 • 由模拟静止视频摄像机捕捉，然后数字化 • 由数字静止视频摄像机捕捉 • 由模拟视频摄像机捕捉，再由计算机视频板上的 A/D 转换器生成
		合成静止图像	由计算机辅助生成	• 用画面程序手工创建 • 屏幕画面捕捉 • 把图形转换为位图 • 由程序生成
图形	对象	计算机图形		• 通过图形编辑器手工创建 • 由计算机程序生成

3、存储需求

图像所需的存储量远多于图形或正文的存储量，原因是位图图像忽略了访问信息。在不使用压缩算法的情况下，两幅相同大小的图像所占的存储空间完全一样，而不管这两幅图像的复杂度相差有多大（可假设一个图像中保存照片，另一个图像中保存的是简单的几何线画图）。因此，可以得出这样的结论：图形存储方法远比图像存储方法所需的空间小，只要有可能，我们就应该用图形保存画面。

然而，在某些情况下我们无法把图像按图形格式来保存。第一个原因是，图像的语义有时很复杂，不能为计算机识别，而模式识别只能解决部分问题。光学字符识别（OCR）是把位图转换为文字信息的典型例子，它远远未普及使用，目前绝大多数计算机扫描仪或传真机等设备所产生的仍然是位图格式。有些程序检测位图图像中的形状边界，把图像转变为对象，这种技术也称“自动跟踪”。

第二个原因是有照片效果的图像只用计算机图形技术很难绘制出来。

第三个原因在于计算机图形生成的时间比显示图像的时间要长很多。这是因为图形格式中保存的是抽象描述信息，在生成前要对它们进行解释。如果要生成的是三维真实感图形，那么所需的时间取决于场景的复杂程度，有时可能要几十分钟。这样，在某些情况下，人们还是要使用位图格式来保存相应的图形画面。

4、图像压缩

图像数据的一大特点是数据量大。数据量大不仅影响到存储，而且也影响到传输，在传输时可能会产生较大的延迟。为了解决这个问题，需要对图像进行压缩处理。压缩方法很