

多媒体

技术实用教程

刘淑芬 胡亮 李春阳 编著



★ 吉林大学出版社

前 言

在现代科技高度发展的今天,信息的处理、信息的传输和信息的共享成了推动现代社会进步的关键,多媒体技术的产生和发展是为适应这一需要的产物。

多媒体技术的出现被认为是计算机的下一场革命。多媒体的广泛应用构成了九十年代的时代特征。多媒体技术的意义在于:大大增强了计算机处理和存储信息的能力,由于多媒体技术充分利用了文字、图像、声音等各种信息,并模拟了人的形象思维和逻辑思维的能力来综合处理和运用这些信息,因而使计算机获取并处理各种信息的能力大大增加;多媒体技术的应用使计算机系统的人机界面更加友好,操作更加方便,信息的表达方式更加符合人的习惯。有人说未来的多媒体系统将像“傻瓜”照像机一样“聪明”,使用方便,功能齐全,多媒体技术使音像技术、计算机技术和通信技术这三大信息处理技术紧密结合在一起,为信息处理技术的发展开拓了新的途径。虽然多媒体技术目前还处于发展初期,但强大的社会需求,促使多媒体技术加速走向标准化和实用化道路。

面对世界性多媒体技术研究和开发热潮,许多单位和各方面专家正研究和应用多媒体技术,人们迫切需要多媒体技术资料 and 参考书。为了满足广大读者学习和开发多媒体的需要,我们编写了这本《多媒体技术实用教程》,对多媒体技术开发者、学习者和使用者在阅读时有较强的实用性和操作性。

本书内容共分七章:1. 多媒体计算机概论;2. 多媒体计算机系统;3. 多媒体数据压缩编码技术;4. CD-ROM 与 CD-I 系统;5. 多媒体信息处理系统;6. 多媒体计算机的安装与使用;7. 多媒体计算机网络。对于这些内容,读者可根据自己的实际需要阅读,不一定要顺序地阅读。

本书由吉林大学李春阳副教授主写第一章,刘淑芬副教授主写了第二、三、四章,胡亮博士主写了第五、六、七章,还有李伟军、韩志安、刘淑清、董凤霞、胡明、冯毅同志参加了部分编写、材料翻译及整理工作。由刘淑芬同志审阅了全书各章。

在编写工程中,我们参阅了国内外有关多媒体技术的书刊及论文,受到同行专家、教授的热情帮助和支持,我们谨向他们表示诚挚的谢意。

作者

1996年7月

7

目 录

第一章 多媒体计算机技术概论

1.1 多媒体和多媒体计算机技术	(1)
1.2 利用多媒体是计算机技术发展的必然趋势	(2)
1.3 多媒体计算机技术的发展现状	(3)
1.3.1 Commodore 公司的 Amiga 系统	(3)
1.3.2 Philips/Sony 公司的 CD-I 系统	(4)
1.3.3 Apple 公司的 Hypercard	(6)
1.3.4 Intel 和 IBM 公司的 DVI 系统	(7)
1.4 多媒体计算机关键技术	(8)
1.4.1 视频和音频数据压缩和解压技术	(8)
1.4.2 多媒体计算机硬件体系结构的关键	(11)
1.4.3 多媒体计算机系统核心软件	(12)
1.5 多媒体系统的应用	(14)
1.5.1 教育和训练	(15)
1.5.2 演示系统	(15)
1.5.3 多媒体电子出版物	(15)
1.5.4 多媒体通信	(16)
1.5.5 视频会议	(16)
1.5.6 其它应用	(16)

第二章 多媒体计算机系统

2.1 多媒体系统	(17)
2.1.1 DVI 的硬件系统	(17)
2.1.2 DVI 的软件系统	(18)
2.2 多媒体硬件系统	(18)
2.2.1 视频卡	(18)
2.2.2 音频卡	(21)
2.2.3 光存储器	(22)
2.3 多媒体软件系统	(22)
2.3.1 多媒体软件支持环境	(22)
2.3.2 多媒体程序设计	(23)
2.3.3 多媒体系统开发平台	(23)
2.4 Windows 的多媒体功能	(24)
2.4.1 Windows 的多媒体扩展	(24)
2.4.2 Windows 3.1 对多媒体的支持	(26)
2.4.3 检查多媒体 Windows 驱动程序	(27)

2.4.4 检查 Windows 的波形音频驱动程序	(30)
2.5 CD-I 的光盘实时操作系统	(31)
2.5.1 CD-RTOS 的特性	(31)
2.5.2 CD-RTOS 的结构	(32)
2.5.3 文件管理程序	(33)
2.5.4 同步和控制	(34)
2.5.5 CD-I 基本系统的主要技术指标	(36)
2.6 超级文本与超级媒体	(37)
2.6.1 简述	(37)
2.6.2 超级文本的定义及其功能	(38)
2.6.3 超级文本的数据模型	(39)
第三章 多媒体系统的数据压缩编码技术	
3.1 概述	(41)
3.2 视频信号编号	(41)
3.2.1 彩色电视信号	(41)
3.2.2 彩色图像信号的编码与解码	(42)
3.3 图像的数字编码与压缩方法	(43)
3.3.1 图像数据的抽样与量化	(43)
3.3.2 图像编码与压缩方法	(43)
3.4 预测编码技术	(46)
3.4.1 预测编码	(46)
3.4.2 自适应预测	(47)
3.4.3 自适应量化	(48)
3.5 变换编码	(50)
3.5.1 二维快速余弦变换	(50)
3.5.2 余弦变换编码的自适应量化	(51)
3.6 熵编码	(52)
3.6.1 矢量量化	(52)
3.7 多媒体数据压缩编码标准	(53)
3.7.1 P×64kb/s 电视图像压缩编码	(53)
3.7.2 MPEG 运动图像压缩编码	(57)
3.7.3 JPEG 静态图像压缩编码	(61)
第四章 CD-ROM 与 CD-I 系统	
4.1 光盘的诞生与发展	(70)
4.1.1 LV 激光视盘	(70)
4.1.2 CD-DA 激光唱盘	(71)
4.1.3 CD-ROM	(72)
4.2 CD-ROM 盘	(72)
4.2.1 盘片	(72)
4.2.2 EFM 调制编码	(73)

4.2.3	帧的数据结构	(74)
4.2.4	扇区的数据结构	(74)
4.2.5	盘的数据结构	(77)
4.3	CD-ROM 文件结构	(79)
4.3.1	结构设计概要	(79)
4.3.2	文件	(80)
4.3.3	逻辑块及逻辑扇区	(80)
4.3.4	路径表	(81)
4.3.5	卷	(82)
4.3.6	目录	(82)
4.4	CD-ROM 驱动器简介	(83)
4.4.1	CD-ROM 驱动器的结构	(83)
4.4.2	光头	(84)
4.4.3	伺服机构	(84)
4.4.4	EFM 解调	(85)
4.4.5	错误检测和校正处理	(86)
4.5	CD-I 系统	(86)
4.5.1	CD-I 光盘的数据格式	(87)
4.5.2	CD-I 扇区格式	(89)
4.5.3	CD-I 基本系统	(90)
4.5.4	CD-I 视频子系统	(92)
4.5.5	CD-I 音频子系统	(98)
第五章 多媒体信息处理系统		
5.1	多媒体系统中的音频信号	(103)
5.1.1	信号	(103)
5.1.2	复合信号	(103)
5.1.3	信号的采样和量化	(104)
5.1.4	音频信号的存贮和处理	(107)
5.2	声音质量	(108)
5.3	多媒体系统中的图像	(110)
5.3.1	图像的组成	(111)
5.3.2	分辨率和象素深度	(113)
5.3.3	真彩色、伪彩色和调配色	(114)
5.3.4	矢量图和点位图	(115)
5.4	图像处理系统	(117)
5.4.1	图像输入设备	(118)
5.4.2	运算装置	(119)
5.4.3	存贮器	(120)
5.4.4	图像输出设备	(120)
5.5	图像处理软件	(122)

5.5.1 图像处理专用语言	(123)
5.5.2 交互式图像分析和命令语言	(123)
5.5.3 图像处理算法库	(124)
5.5.4 图像处理系统的注意事项	(124)
5.6 图像输入方法	(124)
5.6.1 图像输入的一般问题	(125)
5.6.2 图像输入系统	(127)
第六章 多媒体系统的安装与使用	
6.1 多媒体计算机的标准与组成	(128)
6.1.1 MPC 标准	(128)
6.1.2 组成	(128)
6.2 多媒体一般的安装过程	(130)
6.2.1 IRQ 地址	(131)
6.2.2 通道	(132)
6.2.3 基本地址	(132)
6.2.4 安装程序	(132)
6.2.5 跳接和 DIP	(133)
6.3 硬件安装	(134)
6.3.1 注意事项	(134)
6.3.2 多媒体硬件连接	(134)
6.3.3 内部连接	(134)
6.3.4 菊花链连接	(135)
6.4 软件安装	(136)
6.4.1 软件安装内容及安装顺序	(136)
6.4.2 软件的有效期	(137)
6.5 声音卡的安装与使用	(139)
6.6 CD-ROM 驱动器安装和使用	(141)
6.7 视卡安装与使用	(144)
6.8 多媒体的开发	(147)
6.8.1 多媒体信息的获取	(148)
6.8.2 媒体控制接口 (MCI)	(152)
6.8.3 音频服务	(152)
第七章 多媒体计算机网络	
7.1 计算机网络概述	(155)
7.1.1 网络体系结构	(155)
7.1.2 IEEE 标准	(156)
7.1.3 使用网络的优点	(158)
7.2 分布式多媒体网络技术	(158)
7.2.1 分布多媒体技术的产生	(158)
7.2.2 分布式多媒体技术的通信平台	(159)

7.2.3 多媒体信息的传输机制	(160)
7.2.4 分布式多媒体数据库	(161)
7.3 综合业务数字网 ISDN	(162)
7.3.1 ISDN 系统结构	(162)
7.3.2 宽带 ISDN (B-ISDN) 及其信息传送方式	(164)
7.4 ATM 和 SONET	(165)
7.4.1 ATM	(166)
7.4.2 SONET	(166)
7.4.3 高速数据业务	(167)
7.5 分布式多媒体系统	(168)
7.5.1 分布式多媒体系统结构	(168)
7.5.2 宽带网环境	(170)
7.5.3 多媒体数据库	(171)
7.6 MTP 电影传输协议	(172)
7.6.1 更新彩色查找表 CLUT	(172)
7.6.2 优化 CLUT 更新方案	(173)
7.6.3 可靠的传输	(173)
7.6.4 结论	(174)
附表1 亮度 AC 系数表	(175)
附表2 色度 AC 系数表	(178)
参考书及参考资料	(181)

第一章 多媒体计算机技术概论

多媒体计算机技术是当前计算机工业热点课题之一。各种多媒体计算机技术的产品及其开发在发达国家方兴未艾；在我国，无论在多媒体计算机硬件和软件，还是在多媒体计算机技术的应用方面都在竞相争上，生产者、研究者、使用者和学习者等不同层次的群体都在迅速增长。

多媒体计算机技术是指计算机综合处理多种媒体信息：文本、图形、图像和声音，使多种信息建立逻辑连接，集成为一个系统，并且具有交互性。简单地说：计算机综合处理声、文、图信息；具有集成性和交互性。多媒体计算机技术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、特技以及显示输出等技术。可以说多媒体计算机是趋于人性化的多维信息处理系统。

其实，目前多媒体技术大多只利用了人的视觉和听觉，虚拟现实（Virtual Reality）中用到触觉，而味觉和嗅觉尚未集成进来。对于视觉也主要在可见光部分，其它光谱段尚未有效地利用。随着技术的进步，多媒体的含义和范围必将进一步扩展。或者说，这个领域已从量的积累发展到质的飞跃。所以，在认识多媒体技术时，立足点也要移过来。

1.1 多媒体和多媒体计算机技术

“多媒体”一词译自英文“Multimedia”，而该词又是由 multiple 和 media 复合而成的。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等，中文常译为媒质；二是指传递信息的载体，如数字、文字、声音、图形和图像等，中文常译作媒介。与多媒体对应的是单媒体（Monomedia），从字面看，多媒体是由单媒体复合而成的。

从语释的角度出发，对多媒体可定义为：“多媒体是两种以上以下媒体组成的组合体：文本、图形、动画、静态视频、动态视频、声音。据此可以说，多媒体是信息表示的多样性。随着技术的进步，多媒体的含义和范围已经扩展，如仍基于语释来定义，我们不妨说多媒体的“多”是说其多种媒体表现，多种感官作用，多种仪器设备，多学科交汇，多领域应用；多媒体的“媒”是指人与客观世界之中介；多媒体的“体”是说其综合集成一体化，包括各种信息及其关系的码流一体化，设备控制一体化，并且具有实时交互控制环境。据此，可以定义多媒体技术是能对多种载体（媒介）上的信息和多种存储体（媒质）上的信息进行处理的技术。多媒体计算机技术使计算机交互式综合处理多种媒体信息——文本、图形、图像和声音，使多种信息建立逻辑连接，集成为一个系统并具有交互性的技术。

事实上，多媒体的提出，不仅仅是人们有了把多种媒体信息做统一处理的需要和愿望，更重要的是其发展的技术基础条件已经成熟，人类已经拥有了其科学、技术和产业发展的能力。多媒体的目标是要尽可能实现像人类在其临境的自然景象下那种信息交流的高保真效果、通信带宽和交互控制能力。如果从多媒体本征的角度，给出较为抽象的定义，可以定义为多媒体计算机系统是趋于人性化的多维信息处理系统。

1.2 利用多媒体是计算机技术发展的必然趋势

自计算机诞生以来,人们一直在探讨、开拓人和计算机之间的信息沟通方式。由于科学和技术发展的影响,直到今天才有了长足的前进。计算机发展的初期,人们只能用数值这种媒体承载信息。即只能通过0和1两个符号表示信息,而且用纸带和卡片的有孔和无孔记录数值“1”和“0”,纸带机和卡片机是主要的输入输出设备。用“0”和“1”综合表示信息很不直观,也不方便,对其表示的内容难理解,难记忆,并且容易出错,一旦错了也不容易发现、定位。这一时代称为机器语言时代。因此计算机应用面很小,只能限于极少数计算机专业人员。

50年代到70年代,出现了高级程序设计语言,开始使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,由计算机翻译成机器可识别的代码,并且可以将计算机处理的结果用文字表示输出出来。这样,人与计算机之间的信息交流就容易、直观很多,计算机的应用范围也扩大了,具有一般文化程度的科技人员也可以使用。这时期的输入输出设备主要是打字机、键盘和字符显示终端。使用文字承载信息,与计算机交互是一个很大的进步。而使用英文文字同计算机交互,对文化水平低,特别是非英语国家,仍然是困难的事。

80年代开始,人们致力于将图形、图像和声音作为信息的载体与计算机进行交互,这将使计算机的使用更为直观、容易。1984年Apple公司的Macintosh个人计算机,首先引进“位映射”的图形机理,用户接口开始使用Mouse驱动的窗口技术和图符(Windows and Icon),受到广大用户的欢迎。这就使得文化水平较低的公众,包括儿童在内都能使用计算机。

现代科学为延伸和扩展人的信息获取和信息传播能力,进行了不懈的探索和卓越的创造。在80年代之前,广播电影工业、印刷出版工业、邮电通信工业和计算机工业基本上是各自独立发展,它们按照各自的应用目标,研究发展其支持技术,在各自的应用领域内逐步优化提高其效能,形成了各种不同的媒体形式、技术规范、工业产品。

今天,国际上在下述几项技术又有了突出的进展:

- 超大规模集成电路的密度增加了16倍;
- 超大规模集成电路的速度增加了8倍;
- CD-ROM可作为低成本、大容量个人计算机的只读存储器(5英寸盘片的容量达600多兆);
- 双通道VRAM的引入;
- 网络技术的广泛应用。

这五项计算机基本技术的进展,有力的带动了数字视频压缩算法、技术和视频处理器结构的发展,使得几年前单色文体/图形子系统转变成今天的彩色丰富、高清晰度的显示子系统,同时能够作到全屏幕、全运动的视频图像,高清晰度的静态图像,视频特效及三维实时全电视信号及高速真彩色图形。同时还有高保真的声音信息。

现代文明和信息社会的发展,呼唤着科技界提供集成化的多媒体信息系统,它植根于现代多种高技术沃土,并推动信息化社会在一个新的发展阶段上前进。

多媒体对各种信息工具和方法手段的集成,已产生许多它们原来没有的新功能、新概念,创造了一套属于多媒体的技术规范和设计范型,并正在不断深化。虽然目前时日尚短,但已现端倪。

●多媒体系统能将不同媒体数据都表示成统一的结构码流，能够对其进行变换、重组和分析处理，能够进行存储、传送和交互控制。

●多媒体系统实行新的技术标准体制，以适应系统级集成和规范相关产品的性能指标，如：H. 261, JPEG, MPEG I、II、N, MHPE, Hy Time, PREMO, AVIs 等。有一些原来较成熟的标准（如语音、电视等）在不断修订或重新颁布，还有一些全新的标准正在研制、竞争和筛选之中。

●建设支持多媒体通信和资源共享的高速宽带网络环境，在“全球信息高速公路”建设中，每一个国家和地区都以支持多媒体通信为技术指标，这将使得多媒体信息系统能以置身于前所未有的共享空间的条件来研究设计，从而能够实现过去难以想象的应用目标。

●多媒体信息系统强调应用的双向性设计，即各终端用户既是信息的接受者，又是信息的提供者，同时还是在授权范围内的信息控制者。例如视频会议的与会者，异地设计的会商，多媒体教学（CAI），以及其它交互控制的共享空间方式。

●多媒体信息系统设计的目标是赋予其对客观世界的信息级的自然模拟和处理能力。对多媒体数据对象之间的语义、时态和空间关系的分析及表示模式，基于内容理解的数据压缩和信息抽象化，使我们能够克服多媒体的数据“爆炸”，使建设全景化虚拟世界数据库（Virtual World Database）能够成为现实。

总之，多媒体计算机的特点已逐渐明确，其中所涉及的许多技术已取得了巨大进展，有些支持技术正成为研究热点。计算机利用多媒体技术是必然的发展趋势。

人和计算机的信息交互最方便最自然的途径是使计算机具有看、听和说的能力。但是图像和语言的识别和处理都属于约束不充分（Under constrained）的问题。也就是说，只根据图像和声音本身不能提供充分的约束求得唯一解，还必须要有知识的引导，这就涉及到人工智能的研究。

1.3 多媒体计算机技术的发展现状

目前世界上许多国际性大公司都在研制开发多媒体计算机技术。首先值得提出的是 Apple 公司，他在 84 年将位映射（bitmap）概念引入计算机来对图进行处理，并使用了窗口和图形符号（icon）作为用户接口。进一步在 87 年将“超级卡（Hypercard）”引入计算机，成为一种用户可方便使用的、能处理多种信息媒体的计算机。

在目前卓有成效的公司和他们的系统简介如下。

1.3.1 Commodore 公司的 Amiga 系统

Commodore 公司在 1985 首先在世界上推出第一个多媒体计算机系统 Amiga。到 1989 年秋形成了 Amiga 系统的一个完整系统。到目前为止，该公司已推出 Amiga 500, 1000, 1500, 2000, 2500 及 3000 等型号的产品，它们分别配置 Motorola 公司生产的 68000, 68020 以及 68030 不同型号的 CPU 以及不同容量的 RAM。并且为了提高处理视频信息和音频信息的速度，Comodor 公司在 Amiga 系统中使用了三个专用芯片：Agnus（8370），Paula（8364），以及 Denise（8362）。

Amiga 系统结构与 68000 微机系统以及 CD-I 系统非常相似，只是在系统总线上连接了上述三个很有特色的专用芯片。下面概要地介绍一下三个芯片的结构及功能。

(1) Agnus (8370) 是专门用于动画制作的芯片, 芯片中有五个 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。因为 Agnus 中有较多的控制寄存器, 所以设有寄存器地址译码器以及寄存器地址存储器译码器等, 此外芯片还有与系统总线接口电路, 包括缓冲器、多路开关和时钟发生器等。

Agnus 的功能概括起来可归纳为:

- 用硬件显示移动数据, 允许高速的动画制作;
- 控制 DMA 的 25 个通道, 使 CPU 以最小的开销处理盘、声音和视频信息;
- 产生系统时钟;
- 为视频 RAM (VRAM) 和扩展 RAM 卡提供所有的控制信号;
- 为 VRAM 和扩展 RAM 提供地址。

(2) Paula (8364) 是音响处理和外设接口专用芯片, 片中包括音响处理器、盘控制器、异步通信接口以及电位计通道接口, 都连到内部总的设备译码器上。音响处理器由两路数据寄存器、两个音响控制计数器及四路 D/A 变换器组成。它以 DMA 方式和 Amiga 系统的存储器及其它设备交换音响信息, 在 Paula 的音响处理器中处理音响信息, 经 D/A 变换后, 把四路两对立体声信号输出到音响设备中。盘控制逻辑也是以 DMA 方式将 Amiga 系统存储器中的数据输出到盘上, 或将盘上的数据取到系统存储器中。异步通信接口和电位计通道接口, 都以 I/O 方式进行数据传输。该芯片的主要功能是输出四路两个立体声道及音质控制。

(3) Denise (8362) 是专用图形芯片, 包括位平面数据寄存器、位平面控制以及位平面串行输出器; 游标数据寄存器、游标串行连续化器以及位置比较逻辑; 碰撞控制逻辑、碰撞检测逻辑及碰撞存储逻辑; 优先排队控制逻辑、位平面排队和控制寄存器; 彩色选择译码器及 32 位彩色输出寄存器; 鼠标 (Mouse) 计数器。它是一个多功能的彩色图形控制器, 可以控制不同的分辨率和 4096 种颜色。

该公司为了适应不同用户对多媒体技术的需要, 提供一个多任务 Amiga 操作系统, 提供上下拉菜单、多窗口、图符和 PM (presentation Manager) 等功能。提供动画绘制、电视片头制作和乐曲制作等大量应用软件。最近又推出 Amiga Vision 多媒体著作系统, 为用户提供一个完备的图符编辑语言 (A Complete logic programming language)。

1.3.2 Philips/Sony 公司的 CD-I 系统

1986 年 4 月 Philips/Sony 公司公布了基本的 CD-I 系统及 CD-ROM 文件格式, 这就是后来的 ISO 标准。这个系统将高质量的声音、文字、计算机程序、图形、动画及静止图像都以数字的形式存在只读光盘 (5 寸, 650MB) 上。该系统可与家用电视机、计算机显示器进行通信, 并可使用鼠标器、操纵杆或遥控器等定位装置选择视听材料进行播放。

CD-I 基本系统的结构如图 1-1 所示。它可分为两个大部分: (1) CD-ROM 驱动器, 包括 CD 译码器和 CD-ROM 译码器及其它控制电路; (2) 多媒体控制器 MMC, 包括音频信号处理器的 D/A 变换器、视频信号处理器、CPU (68000 系列), RAM、ROM 及 NVRAM、定位装置等。

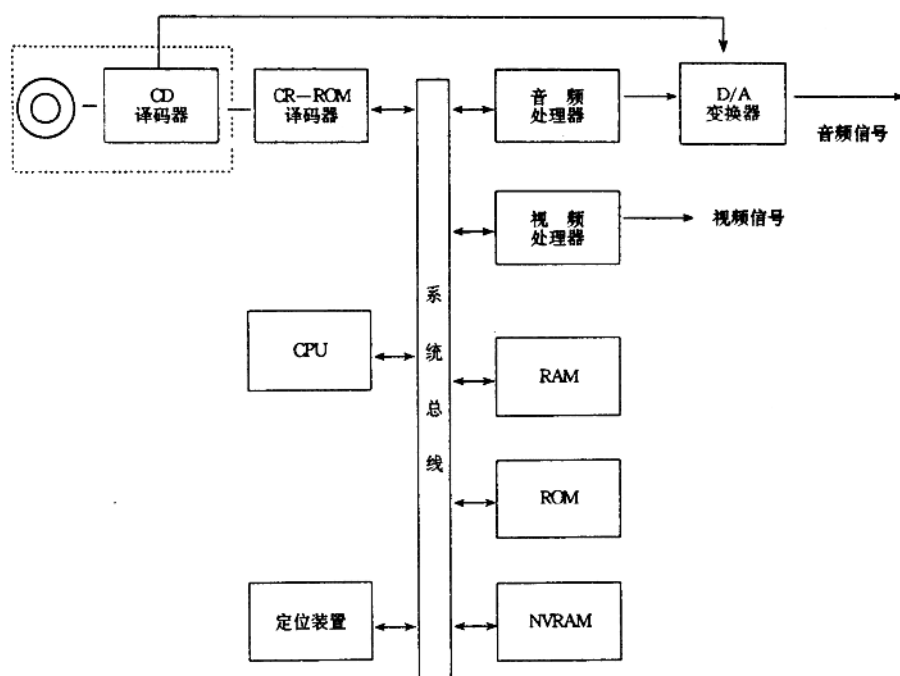


图1-1 CD-I基本系统的结构

该系统有两种工作方式：

1. 与家用电视机、录像机及音响设备连接，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，把声音和图像数据分别送到音频处理器和视频处理器，经过处理后再送往音响设备和电视机或录像机。用户可以通过操纵杆或鼠标等定位装置移动屏幕上的游标，向 CD-I 系统发出指令控制系统的运行。

2. 与微型计算机、工作站以及小型机系统相连，作为多媒体控制器使用。

CD-I 基本系统提供四种不同音质的运行方式，一种是选用 CD-DA 光盘，它采用脉冲编码调制 (PCM) 系统，可得到超级 HiFi 的音响效果；其它三种是选用 CD-I 光盘，它采用自适应差分脉冲码调制 (ADPCM)，选用不同的采样频率、量化精度，可以得到 Laser Vision 的音质 (A 级)，FM 调频广播的音质 (B 级)，和 AM 调幅广播的音质 (C 级)。该系统还定义了三种图像分辨率及若干种视频工作方式。

增强型 CD-I 系统结构框图如图 1-2 所示，它采用了一套全新的专用电路芯片。

该系统采用 Motorola 公司高性能的嵌入式微处理器 MC68340 芯片做主控；视频系统控制器用于内存管理；视频合成器处理位映射图像；全运动视频信号控制器处理视频信号的压缩编码和解压缩问题，为 TV 提供全屏的运动图像；视频 D/A 转换器对数字式视频信号进行模拟量的转换，送到 RGB 驱动电路，输出到彩色监视器；56001 数字信号处理器 (DSP) 处理语言信号，通过静态存储器及串行音频 D/A 变换，分左右两个声道输出到音响设备。

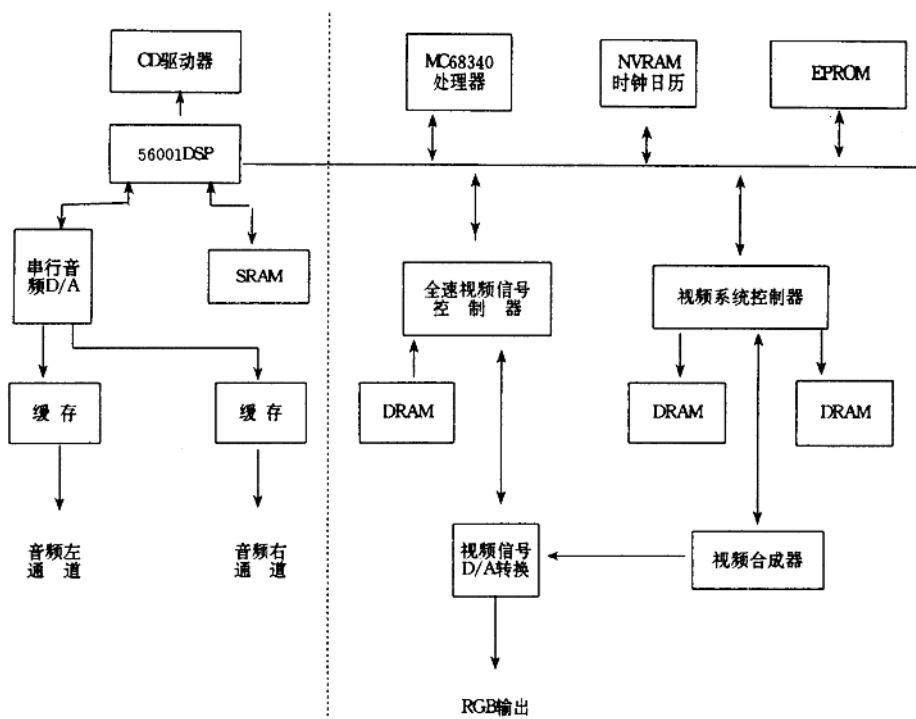


图1-2 增强型CD-I系统结构框图

1.3.3 Apple 公司的 Hypercard

Apple 公司的 Macintosh 系统具有良好的图形特性是众所周知的,它是桌上出版和桌上展示系统的先驱。它用 Hyper Card 及相应软件把高质量的音响及活动的视频图像加到 Macintosh 系统形成自己的多媒体计算机系统。Hyper Card 是以卡片 (Card) 为节点的超级文本 (Hypertext) 系统,基本的信息单元或称为节点,一个卡片可以充满整个屏幕。一组卡片称为卡堆 (stack),可以认为卡堆是 Hyper Card 中的文件,同类的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。Hyper Card 系统还提供许多命令和工具,利用这些命令和工具可完成卡片的浏览、编辑、制作,信息的输入、修改和检索。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。Hyper Card 的数据库和所有的 MAC 的数据库格式兼容,并开发出直接的连接电路、光扫描器以及 CD-ROM 驱动器连接软件。为了使得 HyperCard 和这些外部设备相连接,Apple 公司公布了一个多媒体协议和驱动程序标准集,称为 AMCA (Apple Media Control Architecture)。AMCA 是系统级的结构,用来访问视频光盘、音频光盘及录像带的信息。

Apple 公司原来的多媒体计算机平台是 Mac SE 和 Mac II,现选用 68030 CPU,直接寻址可达到 8 兆字节内存,视频适配器板 16M 种不同颜色中显示其中的 256 种颜色。音响媒体接口板及 HyperCard 软件,能够提供良好的语音、音响效果,通过语言分析和识别能够替代键盘、鼠标以及操纵杆的功能。

为了快速、实时地处理视频和语言信号 Apple 公司正在加强合作, 研制开发高速宽带网以及对称的视频和音频信号的压缩编码和解码技术。

1.3.4 Intel 和 IBM 公司的 DVI 系统

DVI (Digital Video Interactive) 技术 1983 年在 RCA 公司开始开发, 1987 年首次公布 DVI 技术的研究成果, 1989 年 Intel 和 IBM 公司在国际市场上推出 DVI 技术的第一代产品 Action Media 750, 1991 年推出了 DVI 技术的第二代产品 Action Media 750 I。

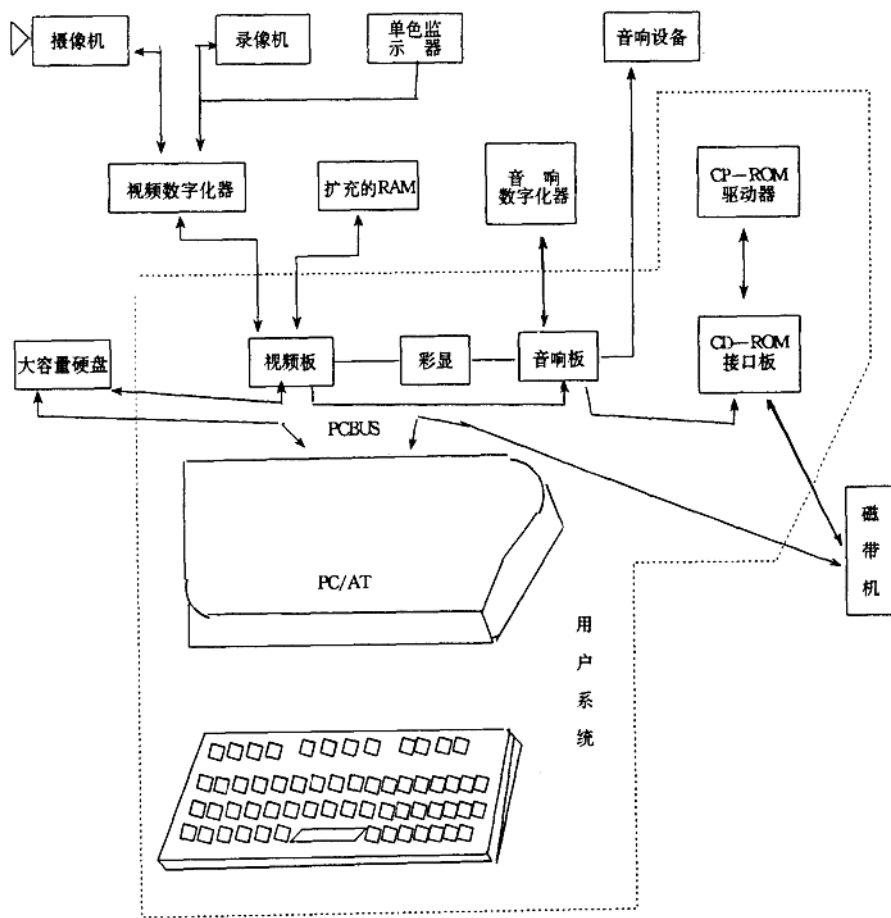


图1-3 Action Media 750 I 的结构框图

DVI 技术的第一代产品 Action Media 750 I 的结构框图如图 1-3 所示。

其核心是两块专用的 DVI 接口板: DVI 视频板、DVI 音频板和 DVI 多功能板。工作平台包含各种机型, 如 IBM PC/AT, 386, 486 或其它兼容型计算机。当配上 CD-ROM 驱动器, 带放大器和音响效果的 RGB 彩色显示器, 就组成了用户系统。在此基础上再配上视频信息数字化仪 (连接到 DVI 视频板上), 音响信号数字化仪 (连接到 DVI 音频板上), 扩展视频 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就组成了 DVI 开发系统。

DVI 技术硬件的核心部件是 Intel 公司生产的专用芯片: VDP1 (82750PA, 82750PB) 和 VPP2 (82750DA, 82750DB), 称为视频像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。像素处理器 (88750PA/PB), 采用微码编程, 可以高速执行像素处理的各种算法。显示处理器 (82750DA/DB), 它与像素处理器 (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, 通过编程可适应不同分辨率、不同像素格式及不同同步格式的多种型号显示器。

DVI 技术软件的核心部件是 AVSS (Audio/Vidio Sub System) 和 AVK (Audio Vedio Kernel)。AVSS 是在 DOS 环境下, 配之 RTX (实时执行) 部件、视频驱动器、音响驱动器、多功能驱动器及驱动器接口模块, 处理音响视频的子系统。AVK 在 Windows 环境下运行, 因此它就不局限于 DOS 操作系统, 也可在其它操作系统环境下运行。AVSS 和 AVK 的主要功能是, 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, 管理它们、控制它们并将它们送至显示缓冲区等。

DVI 技术发展很快, 到 1995 年 Intel 公司将把 VDI 技术放到母板上, 到 2000 年将放在芯片中。本世纪末, Intel 公司生产的 Intel786, 运算速度可达到 2000 MIPS, 包括 DVI 运动视频, 并和 386 结构兼容。目前 DVI 技术有丰富的支持软件, 例如工具软件、库函数、著作语言等。还有很多厂商和公司继续从事多媒体计算机技术的研制、开发以及设计制造工作。值得一提的是 Next 计算机公司的 Next 计算机的卓越设计思想。它利用 Postscript 实现高级绘图功能, 采用 DSP 进行高保真度的声音信号处理, 提供声音和图像的 Nextmail 邮件系统等。

1.4 多媒体计算机关键技术

多媒体技术可以认为是将文本、图形、声音和静态图像、活动图像与计算集成在一起的技术。目前, 图形图像的通信技术已经成熟, 三维图形硬件在技术上也已成熟, 多媒体技术的软硬件产品琳琅满目。多媒体信息压缩的标准已推出 JPEG、MPEG、P×64 (H261)、MHEG 等多种。

多媒体技术主要趋向是进一步解决信息压缩、集成、交互和同步等问题。集成指不同媒体信息、不同视听设备及软硬件的有机结合; 交互是指提供人们多种交互控制方式; 同步是指媒体在时基方面的配合和制约。

多媒体计算机关键技术包括下述几个方面: (1) 系统平台技术: 多媒体计算机硬件结构中的关键芯片 (专用), 音频/视频信号的压缩与解压缩, 音频/视频信号的获取与播放, 音频/视频信号的混合及同步, 数字信号处理等; (2) 多媒体系统软件: 如多媒体操作系统, 多媒体编辑、动画处理, 多媒体信息的混合重叠以及多媒体数据库管理系统等; (3) 人工智能输入/输出技术, 如语言识别、图像识别及处理, 语言合成, 语言和文字间的转换以及图/文/表分离技术等; (4) 多媒体网络技术: 包括网络管理技术, 高速网络协议, 开放式文件结构等, (5) 其它开发、制作工具及应用软件……。

1.4.1 视频和音频数据压缩和解压缩技术

如果没有数据压缩技术的进步, 多媒体计算机就难以得到实际的应用。多媒体计算机的

关键问题是实时地综合处理大量的声音、文本、图像等信息。数字化的图像和声音信号的数据是非常大的。例如一幅 640×480 的中等分辨率的彩色图像 (24bit/像素) 的数据量约为 7.37Mbit/帧, 如果以每秒 30 帧或 25 帧的速度播放时, 则视频信号的传输速率为 220Mbit/s 或 184Mbit/s。如果存放在 600Mb 的光盘中, 只能播放 20 秒钟或 26 秒钟。对于音频信号, 以激光唱片 CD-DA 声音数据为例, 如果采样频率为 44.1KHz, 量化为 16 bit 两通过立体声, 600 Mb 光盘只能存放 1 小时的数据, 其传输速度为 1.4 Mbit/s。综上所述, 视频和音频数字信号数据量很大, 同时要求传输速度很快。如以 IBM PC/AT 或其它兼容机为例, 其总线传输速度约为 150kb/s。如果处理上述视频和音频数据, 所需要的数据压缩比将接近 200 倍, 否则 IBM PC/AT 机将无法完成此任务。

自从 1948 年 Oliver 提出脉冲编码调制 (PCM) 编码理论开始, 至今已有四十多年的历史, 编码技术日趋成熟。目前常用的编码方法有: PCM; 预测编码; 变换编码 (K-L, DCT, FFT, FWT 变换); 插值和外推 (空域子采集, 时域子采集, 自适应); 统计编码 (哈夫曼, 算术编码, 香农—弗诺)。此外, 还有矢量量化, 子带编码, 行程编码, 比特平面编码等。混合编码也是一种常用的方法, 它可以把上述几种方法混合使用。在研究、选用编码时要特别注意以下两个问题:

- 该编码方法一定能用计算机或者用 VLSI 电路快速实现;
- 一定要符合当前的国际标准。

目前公认的关于压缩编码的国际标准是:

(1) JPEG (Joint Photographic Expert Group) 标准

“联合国图片专家组” (JPEG) 经过多年细致的工作, 于 1991 年 3 月提出了 ISO CD 10918 号建议草案: “多灰度静止图像的数字压缩编码”。其主要内容包括: ①基本系统 (Baseline System) 提供顺序扫描重建图像, 实现信息有丢失的图像压缩, 而重建图像的质量要达到难以觉察出图像损伤的要求。它采用 8×8 像素自适应 DCT 算法、量化以及哈夫曼型的熵编码器。②扩展系统 (Extended System) 采用累进工作方式, 编码过程采用具有自适应能力的算术编码。③无失真的预测编码, 采用帧内预测编码及哈夫曼编码 (或算术编码), 可以保证重建图像数据与原始数据完全相同 (即均方根误差等于零)。

(2) CCITT H. 261 (电视电话/会议电视 $P \times 64K$ bit/s) 标准

CCITT 第十五研究组进行视频编码和解码器的标准化工作, 于 1984 年提出“数字基群传输会议电视”的 H. 120 建议。其图像压缩采用“帧间条件修补法”的预测编码、变字长编码以及梅花型 (Quincunx) 亚抽样/内插复原技术。该组又于 1988 年提出电视电话/会议电视 H. 261 建议 $P \times 64K$ bit/s, 其中 P 是一个可变参数, 可取值 1 到 30, 当 $P=1$ 或 2 时, 支持四分之一中间格式 (QCIF——Quarter Common Intermediate Format) 每秒帧数较低的视频电话; 当 $P \geq 6$ 时可支持通用中间格式 (CIF——Common Intermediate Format) 每秒帧数较高的电视会议。

$P \times 64K$ bit/s 视频编码压缩算法, 采用混合编码方法, 即基于 DCT 的变换编码及带有运动预测的差分脉冲编码调制 (DPCM) 预测编码的混合。在低速率 64K bit/s 或 128K bit/s (即 P 取 1 或 2) 时, 除采用 QCIF 外, 还可采用亚帧 (Sub-frame) 技术, 即隔 1 (或 2, 3) 帧处理一帧, 这种方法的压缩比可以达到 48:1。

(3) 运动图像专家的 MPEG-1 标准

MPEG-1 标准, 作为 ISO CD11172 号建议, 于 1992 年通过。它是运动图像专家组的

“用于数字存储媒体运动图像及其伴音速率为 1.5M bit/s 的压缩编码”的简称。该标准包括三个部分：MPEG 视频，MPEG 音频和 MPEG 系统。因为视频和音频需要同步，故 MPEG 压缩算法，应将视频和音频统一考虑，产生一个电视质量的视频和音频压缩形式的速率约为 1.5M bit/s 的 MPEG 单一的位流。

MPEG 视频压缩算法采用了两个基本技术：运动补偿即预测编码和插补编码，变换域 (DCT) 压缩技术。在 MPEG 中 DCT 不仅用于帧内压缩，对帧间预测误差再作 DCT 变换，可减少空域冗余，从而达到进一步压缩的目的。对于运动图像划分为“I 图像”，“P 图像”和“B 图像”三种。I 图像仅作帧内压缩编码，P 图像仅作正向帧间预测，而 B 图像则作双向帧间预测。用 MPEG-1 标准平均压缩比为 50:1。

(4) 高质量数字影像标准：MPEG-2

MPEG-2 数字影像压缩和解压缩技术对于未来的发展，具有重大的意义。MPEG-2 是一系列设计出来的标准的集合，MPEG-2 标准和相应的新的编码和解码芯片将使多媒体技术发生革命性的飞跃，多媒体信息高速公路的发展很可能比人们预料的快得多。

未经压缩的数字化影像带有比需要的更多的信息。将模拟影像数字化时，对每一位独立象素的颜色和亮度都要精确地标明，颜色可达 1600 万种以上，亮度的等级也差不多有同样的数量。因而每一象素都带有大量的不必要的负担，其中大部分可以进行组合和简化表达。另外，未经压缩的音频和视频信息并非全部为人们感知，人们的大脑还具有补足某些短暂的影像缺少的数据能力。

MPEG-2 标准综合了各方面因素用来除去不必要信息，并系统化地减少必要的、作为编码基础的空白。在 MPEG-2 中的每一个字节可以表达原始影像中 200 个字节的信息，是一个效率极高的压缩系统。

在 MPEG-2 标准中，引入了十分重要的 I-帧，P-帧和 B-帧的概念。I-帧是一种减少内部编码的基本方法，I-帧结构以“可接受的量化”方式使损失降至最小。MPEG 编码器自动地分析输入的影像比特流，并以最优的方式进行压缩。如一个蓝色背景，不需要对每一个象素标识出 32 位字长的颜色、24 位字长的色彩和亮度，只用一种方法表达此 8×8 的方块是某一种蓝色和某一亮度即可。再如蓝色背景在下帧若不发生变化，那就无须记录所有象素的数据，可以只说明这一蓝色方块以同样的颜色在下一帧出现即可。

一帧图像的某些部分常常在 100 帧中维持不变。实现这一复杂的数学技巧是一名“二维离散余弦变换”的数字处理。这一处理过程跟踪每一个象素的变化，并得到运动的向量，估计哪里的运动最可能在下一帧中出现。通过向前和向后的观察，可以非常精确地构造出中间的各帧画面。

在 MPEG 中，由计算机完成中间帧的复制工作。P-帧（预计的帧）的构造，要分析前一帧中的运动方向，估计出将处于下一帧中的何处。B-帧（双向预计的帧）的构造，要分析前后两个方向的运动情况。这样传输的不是象素的细节，而是关于帧内容的操作描述，从而使传输带宽达到最小化。

MPEG-2 的基本分辨率为 720×480 ，每秒传输速度为 30 帧，并且有 CD 质量的声音。MPEG-2 允许在一定范围内改变压缩比，以便在画面质量、存储容量和带宽之间作出权衡。MPEG-2 有 30:1 或更低的压缩比时具有电视广播级的质量，但为了经济可达到 200:1 的压缩比。MPEG-2 的解码器通常支持 MPEG-1 和 MPEG-2 两种标准。

MPEG-3 标准已被放弃。而 MPEG-2 和 MPEG-1 用于高清晰度电视的效果更好。