

多媒体计算机技术的现状和发展趋势(上)

清华大学计算机系 钟玉琢

摘要:本文简要介绍多媒体计算机技术的发展现状,论述了研制多媒体计算机要解决的关键技术:视频和音频数据的压缩和解压缩技术;多媒体计算机硬件体系结构——视频音频引擎和微码引擎(AVE和MCE);以及多媒体计算机软件系统——视频和音频子系统和视频音频核(AVSS和AVK)。最后介绍了多媒体技术的发展趋势。

美国计算机咨询委员会估计到1994年多媒体计算机技术的销售额将是140亿美元。1991年10月在美国Las Vegas召开的“Comdex 91”,参加的展商超过1900家,出席人数超过10万人,在这次展览会上,IBM和Intel公司共同研制的DVI(Digital Video Interactive)系统Action media 750 II,荣获了最佳多媒体产品奖和最佳展示奖(Best Multimedia Product and Best of Show)。Comdex的最佳奖相当于计算机工业的奥斯卡(Oscars)奖。

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,那么什么是多媒体呢?媒体(Medium)在计算机领域有两个含义,一是指用以存储信息的实体,如磁盘、磁带、光盘和半导体存储器;一是指信息的实体,如数字、文字、声音、图形和图象。在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杂志上两篇文章中,分别给出了不太严格定义,概括起来是:计算机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。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。

一、利用多媒体是计算机技术发展的趋势

在计算机发展的初期,人们只能用数值这种媒体承载信息,当时只能通过0和1两种符号表示信息,即在纸带和卡片打孔不打孔表示信息,纸带机和卡片是主要输入输出设备,0和1很不直观,很不方便,输入输出的内容很难理解,而且容易出错,出错了也不容易发现。这

一时代是使用机器语言的时代,由于有上述种种缺点,使得计算机应用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,开始用文字做为信息的载体,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水平的科技人员。这时的输入输出设备主要是打字机、键盘和显示终端。使用文字同计算机交往,对于文化水平较低、特别是非英语国家,仍然是件困难的事情。

80年代开始,人们便致力于研究声音、图形和图象做为新的信息媒体,输入输出计算机,这将使计算机的应用更为直观、容易。1984年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驱动的窗口技术和图符,(Window and icon),受到了广大用户的欢迎,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采用发展多媒体技术,扩大用户层的方针,使得它在个人计算机的市场是唯一能同IBM公司相抗衡的力量。七年后的今天,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可做低成本、大容量PC机的只读存储器(可换的5寸盘片,每片容量为600MB)。
- 双通道VRAM的引进;
- 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器结构的改进,促使七年前单

色文本图形子系统转变成今天彩色丰富、高清晰度显示子系统,同时能够做到全屏幕、全运动的视频图象、高清晰度的静态图象、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。

从以上介绍可以看出,无论从半导体还是计算机技术进步的角度,还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。

人与计算机信息交流最方便最自然的途径是使计算机具有视觉和听、说能力。但是由于图象和语言的识别和理解都属于约束不充分(Under constrained)的问题。也就是说,只根据图象和声音本身不能提供充分的约束求得唯一解,还必须有知识的导引,这就涉及到人工智能的研究。目前这方面的研究尽管已取得很大进展,但远未达到实用阶段,多媒体技术可以被看成是完成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速度快、综合处理等优点,用交互式技术来弥补用计算机实现图象和语音理解和识别的不足。多媒体计算机技术就目前来看它的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,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。

二、多媒体计算机技术的发展现状

目前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,其中卓有成效的公司和系统如下:

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Amiga 500,1000,1500,2000,2500以及3000等型号的产品,它们可分别配置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。为了提高视频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;Agnus(8370),Paula(8364)以及Denise(8362)。

Amiga系统的结构与68000微机系统以及后面介绍的CD-I系统非常相似,只是在系统总线上连接了很有特色的三个专用芯片,下面重点介绍三个专用芯片的

结构:

(1)Agnus(8370)是专用的动画制作芯片,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处理连在一起。因为在Agnus有较多的控制寄存器,所以还有寄存器地址译码器以及寄存器地址存储器译码器,此外还有系统总线的接口电路:缓冲器、多路开关、以及时钟发生器等。

概括起来它的功能是:

- 用硬件显示移动数据,允许高速的动画制作;
- 显示同步协处理器;
- 控制25个通道的DMA,使CPU以最小的开销处理盘、声音和视频信息;
- 从28MHz振荡器产生系统时钟;
- 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)是专用音响处理及外设接口芯片,芯片中音响处理器、盘控制器、异步通讯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。它可以通过DMA的方式和Amiga系统的存储器以及其它设备交换音响信息,在Paula内的音响处理器处理音响信息,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系统中存储器的数据,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,还有异步通讯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四路,二个立体声道,九个八音阶,使用音频放大和频率调制,还有异步通讯接口,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;优先排队控制逻辑以及位平面排队的控制寄存器;彩色选择译码器以及位置比较逻辑;优先排队控制逻辑以及位平面排队的控制器;彩色选择译码器以及32位彩色输

出寄存器; Mouse 计数器。上述可见,它就是多功能彩色图形控制器,它可以控制不同分辨率的输出,从 320X200 到 640X400;在电视机和 RGB 彩色监视器的屏幕上可同时显示 4096 种颜色;有八个可重复使用“硬件游标”控制器。

较新的 Amiga 3000 型,采用了 25MHz 的 68030 做为 CPU,配有协处理器,内存最大容量为 16MB, 9×100MB 硬盘以及任选 Ethernet, Novell Netware 和 unix 网络和软件。

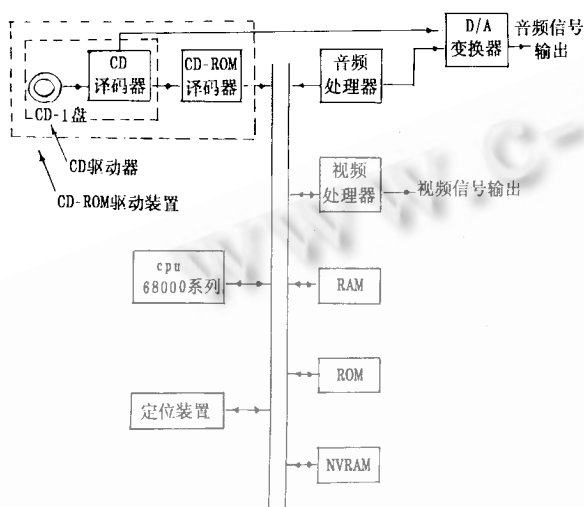


图 1 CD-I 基本系统结构

为了适应不同用户对多媒体技术的需要, Commodore 公司提供了一个多任务 Amiga 操作系统,它有上下拉的菜单、多窗口、图符(icon)以及 PM(Presentation Manager)等功能。同时配备了大量应用软件,如能绘制动画,制做电视片头及作曲等专用软件。最近该公司又推出了一个 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 5 英寸的只读光盘上。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 系统进行通讯,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。

CD-I 基本系统的结构如图 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,它是由音频信号处理器、视频信号处理器、68000 微处理器、RAM、ROM 不发挥的 RAM 以及定位装置组成。

从 CD-I 系统结构图很容易看出,它有两种工作方式,一种是不需要其它计算机,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,并把声音和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音机,用户可通过鼠标或操纵杆等定位装置移动显示屏幕的游标,向 CD-I 系统发出指令,运行交互式的培训教育节目。另一种方式是,CD-I 基本系统可以做为多媒体控制器连接到其它微型计算机、工作站以及小型计算机上。

CD-I 基本系统提供四种不同音质的运行方式,如表 1 所示,一种是选用 CD-DA 光盘,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 转换器,再送给音响设备输出,会得到超级 HiFi 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。基本系统还定义了三种不同的图象分辨率(如表 2 所示)以及几种视频工作方式(如表 3 所示)。

表 1 CD-I 音频方式

级别	采样率 (KHZ)	位数 / 样本	频率响应 (KHZ)	数据率 (字节 / 秒)	通道数	数据流 百分数	播放时间 (小时)
CD-DA PCH 超级 HiFi	44.1	16	20	1171,100	1 立体声	100%	1
CD IADPCM	A (LV)	8	17	85,100	2 立体声	50%	2
				42,500	4 单通道	25%	4
	B (FM)	4	17	42,500	4 立体声	25%	4
				21,300	8 单通道	12.5%	8
	C (AM)	4	8.5	21,300	8 立体声	12.5%	8
				10,600	16 单通道	6.25%	16

表 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
高分辨频率	720 × 480	768 × 480

表 3 CD-I 频率方式

方式	倍数 / 象数	显示颜色 数目	数据量 / 幅画面(4 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	86.4	—	345.6 *
CLUT4	4	16	—	86.4	172.8
RL7		128	86.4 以下	—	345.6 *
RL3		8	—	86.4 以下	172.8

* 表示 CD-I 参基本系统不提供 < 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, Song 公司为 CD-I 扩充硬件, 增强功能的 CD-I 系统如图 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 新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最复杂的专用芯片, 它处理视频信号的压缩编码和解压缩问题, 为 TV 提供全屏幕的运动图象; 视频系统控制器主要用于内存

管理; 视频合成器主要是处理位映射图象; 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。其它还有: EPROM, 非易失性的 RAM 以及视频帧存储器 DRAM。从上述结构可以清楚地看到, 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。

3.Apple 公司的 HyperCard

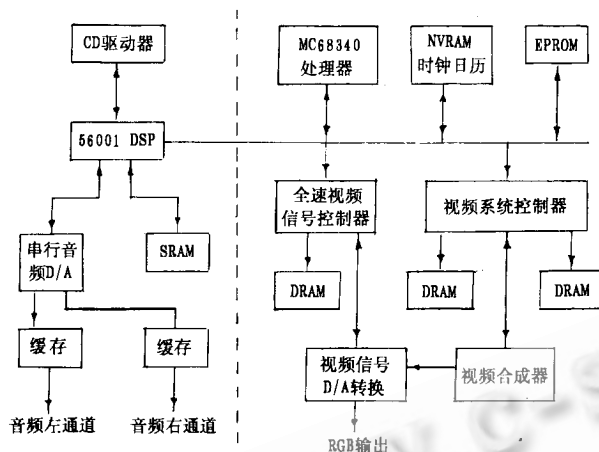


图 2 增强功能的 CD-I 系统

Apple 公司的 Macintosh 系统具有许多国家都公认的良好图形特性,它是桌上印刷和桌上展示系统的先驱。Apple 公司的多媒体系统也有人称它为桌上媒体,实质上它是把高质量的音响及活动的视频图象加到原来的 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(Card)为节点的超文本(Hypertext)系统,基本的信息单元是卡或称节点,一个卡可充满整个屏幕。一组卡片称为卡堆(Stack),可以认为卡堆是 HyperCard 中的文件,同类和相关的卡可在一个卡堆内。每个卡不仅是字符,还包括图形、图象和声音。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,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的标准集,叫做 AMCA(Apple Media Control Architecture)。AMCA 是系统级的结构,它用来访问视频光盘,音频光盘以及录象带的信息,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 Mse I 做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 中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它能够代替键盘、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响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,第一次公布了 DVI 技术的研究成果,1988 年 10 月 Intel 公司从 GE 公司买来 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 II。

DVI 技术第一代的产品 Action Media 750 I 的结构图如图 2 所示,它的核心是三块专用的 DVI 接口板;DVI 视频板、DVI 音频板以及 DVI 多功能板,IBM PC/AT、386、486 或其兼容型计算机做为工作平台,同时配有 CD-ROM 驱动器,带有放大器和音响效果好的 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 音频板),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、录象机、音响设备、监视器以及摄像机或扫描仪等组成的 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 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,它采用微码编程,

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,不同象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是 AVSS(Audio / Tideo Support 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和运行音响视频系统。AVK 是在 Window 环境下运行,因此它就不局限在 DOS 操作系统环境,可以在其它种类操作系统环境下运行。AVSS 和 AVK 最主要的任务是:为声音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,管理控制它们并送到显示缓冲器等。

Intel 公司 DVI 产品发展规划如表 4 所示,1989 年到 1991 年它推出第一代产品 AM750 I,1991 年到 1993 年它要推销第二代产品 AM750 II,1992 年到 1995 年它要推销第三代产品 S3(性能指标详见表 3)。第二代的产品性能指标比第一代产品提高了一倍,价钱仅是原来的三分之一。Intel 公司副总裁 David House 预测,

表 4 Intel 公司 DVI 产品发展的规划

	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	1750A	1750B	V3
硬件板	AM750-series I	AM750-series I	S3
软件	AVSS-DOS(2.13)	AVSS-DOS,AVK-Window(2.2)	AVK
压缩 算法	PLV1.0 / 1.5 PTV1.5 PIC1.0	PLV2.0 PLV2.0 PTV2.0 PIV2.0 PIC2.0 (JPEG) (JPEG)	MPEG PX64 PIC3.0(JPEG) PLV3.0
工具	MDT-Limited MPT.Lumena Authology	MOTlimited MOTlimited Third party Tools	V3 Tools

目前 DVI 技术有丰富的软件支持,它有工具软件、库函数。到 1991 年底,全世界又有 80 多家厂商,为 DVI 技术编制开发工具和各种应用软件,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 图象编辑加工工具,还有很多著作语言,都为用户提供了极大的方便。

全世界还有很多厂商和公司正在从事多媒体计算机

1985 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 中晶体管数目可达到 0.5-1 亿。”他还说:“在本世纪末,Intel 公司生产的 Intel 786,运算速度可达 2000 MIPS,包括 DVI 运动视频,它和 386 结构兼容。”

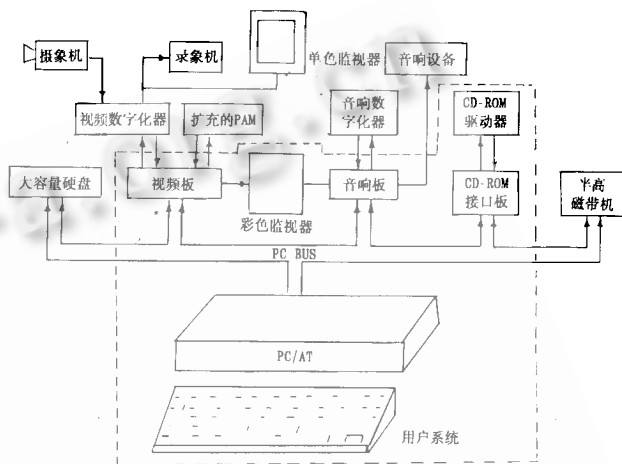


图 3 DVI 系统结构图

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”的史蒂夫·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。它利用 Pest Script 实现高级绘图功能,采用 DSP 进行高保真度的声音信号处理,提供图象和声音的 Next mail 邮件系统等。

(下期续)