

多媒体计算机技术的现状和发展趋势(下)

钟玉琢(清华大学计算机系)

三、研制多媒体计算机要解决的关键技术

1. 视频和音频数据压缩和解决压缩技术

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计算机实时地综合处理声、文、图信息。数字化的图象和声音信号数据是非常大的。一幅 640×480 中等分辨率的彩色图象(24bit/象素)数据量约为 7.37Mbit/帧,如果是运动图象要以每秒 30 帧或 25 帧的速度播放时,则视频信号传输速度为 220Mbit/s。如果存放在 600MB 的光盘中,只能播放 20 秒钟。对于音频信号,以激光唱盘 CD-DA 声音数据为例,如果采样频为 44.1KHz,量为 16bit 两通道立体声,600MB 光盘只能存放 1 小时的数据。它的传播速度为 1.4Mbit/s。综上所述视频和音频数字信号数据量大,要求传输速度快。考虑到 IBM PC/AT 或与其兼容型计算机,总线传输速率为 150KB/S。如果要处理上述视频和音频数据所需数据压缩比将近 200 倍,否则 IBM PC/AT 就无法完成任务。

从 1984 年 O.Liver 提出 PCM(脉冲编码调制)编码理论开始,至今已有四十余年的历史,编码技术日臻成熟。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(空域子采样,时域子采样,自适应);统计编码(哈夫曼,算术编码,香农-费诺);此时还有矢量量化,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,它可把上述几种方法混合使用。在研究选用编码时一定要注意下述两个问题:

- 该编码的方法一定能用计算机或 VLSI 电路快速实现;

- 一定要符合当前的国际标准。

目前最时髦的关于压缩编码的国际标准是:

(1) JPEC(Joint Photographic Expert Group)标准“联合图片专家组”(JPEG)经过五年的细致工作后,于

1991 年 3 月提出了 ISO CD 10918 号建议草案:多灰度静止图象的数字压缩编码”。主要内容如下: A) 基本系统(baseline system)提供顺序扫描重建的图象,实现信息有丢失的图象压缩,而重建图象的质量要达到难以观察图象损伤的要求。它采用 8×8 象素自适应 DCT 算法、量化、以及哈夫曼型的熵编码器。 B) 扩展系统(extended system)选用累进工作方式,编码过程采用具有自适应能力的算术编码。 C) 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),可保证重建图象数据与原始图象数据完全相同(即均方误差等于零)。

(2) 电视电话/会议电视 $P \times 64\text{kb/s}$ (CCITT H.261)标准

CCITT 第 15 研究组积极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。其中图象压缩采用“帧间条件修补法”的预测编码,变字长编码,及梅花形(Quincunx)亚抽样/内插复原等技术。该研究组又在 1988 年提出电视电话在/会议电视 H.261 建议 $P \times 64\text{kb/s}$, P 是一个可变参数,取值为 1 到 30, $P=1$ 或 2 时支持四分之一中间格式(Quarter Common Intermediate Format-QCIF)每秒帧数较低的视频电话,当 $P > 6$ 时可支持通用中间格式(Common Intermediate Format-CIF)每秒帧数较高的电视会议。

$P \times 64\text{kb/s}$ 视频编码压缩算法,是采用混合编码方法,即基于 DCT 的变换编码以及带有运动预测的差分脉冲编码调制(DPCM)预测编码方法的混合,在低速时($P=1$ 或 2, 即 64 或 128kbit/s)除采用 QCIF 外,还可采用并帧(Supl-frame)技术,即隔一或(2,3)帧处理一帧,压缩比可达 48:1。

(3) 运动图象专家组的 MPEG-1 标准。“用于数字存储媒体运动图象及其伴音率为 1.5Mbit/s 的压缩编码简称为 MPEG-1,做为 ISO CD 11172 号建议,于

1992年通过。它包括三部分,MPEG视频、MPEG音频和MPEG系统。由于视频和音频需要同步,所以MPEG压缩算法,应该对视频和音频联合考虑,最后产生一个电视质量的视频和音频压缩形式的位速率约为1.5Mkbit/s的MPEG单一位流。

MPEG视频压缩算法,采用两个基本技术:运动补偿既予测编码和插补编码及变换域(DCT)的压缩技术,在MPEG中DCT不仅用于帧内压缩,对于帧间预测误差再作DCT变换,可减少空域冗余,以达到进一步压缩的目的。对于运动图象,把它们划分为“I图象”,“P图象”和“B图象”三类;其中I图象仅作帧内压缩编码,P图象仅作正向帧间预测,而B图象仅做反向帧间预测。用MPEG-I标准平均压缩比为50:1。

2.多媒体计算机硬件体系结构的关键是专用芯片

多媒体计算机需要快速、实时完成视频和音频信息的压缩和解压缩,图象的特技效果(改变比例尺,淡入淡出,马赛克)、图形处理(图形的生成和绘制)、音频信息处理(抑制噪声,滤波)。满意地完成上述任务,一定要用专用的芯片。半导体的厂商,清楚地看到这点,它们纷纷参与了这一竞争。

多媒体计算机专用芯片可归纳为二种类型:一种是固定功能的芯片,一种中可编程的处理器。第一批固定功能的芯处,目标瞄准了图象数据的压缩处理。LSILogic公司,SGS-Thomson公司以及C-cube公司都设计制造了一批这样的芯片。

SGS-Thomson公司设计制造的STI 3220,它采用了Systolic结构,做了256个处理器,可以在 31×31 窗口条件下执地 16×16 像素块全搜索算法,可做为运动预测处理器。INMOS分公司设计制造的IMS-A121,它采用流水线处理器能够执行 8×8 像素的DCT运算。此外还有STU 3200(4×4 到 16×16 DCT运算)以及STU 3208(Z字形扫描器)。

C-Cube公司设计制造的CL-550是世界上第一个把执行JPEG标准算法做在一个专用芯片上。现在已在国际市场上广泛销售。最近该公司又推出了CL-550B,它的速度比第一型提高了一倍。用CL-550专用芯片能够对静态图象实时地进行压缩和解压缩处理。与此同时,C-cube公司正在考虑采用灵活的方式,即可编程的结构设计制造执行MPEG标准的专用芯

片,设计时已考虑MPEG标准的扩充和修改。目前其它公司SGS-Thomson,LSI Logic等也正在设计制造执行MPEG标准的专用芯片。

集成信息技术公司(Integrated Information Technology)最近推出了一种多功能芯片,即最新的视频压缩芯片VP(Vision Processor),为了支持多种压缩编码标准,VP采用微码可编程技术,VP可直接被主处理器控制执行JPEG标准算法,同时它还能执行MPEG和 $P \times 64$ 标准算法。VP有高速像素I/O通道,微码控制器,VLC以及缓冲区系统到主计算机接口。VP建立了高速数据逻辑通道,它把微码控制器和MPEG及 $P64$ 压缩和解压算法数据流相结合,同时VP可并行执行4个 16×16 乘法器及8个16bitALU运算,它们能够并行工作。

VP还有Risc处理器和64个64bit寄存器以及可扩展的大容量的存储器,它能够执行30Fp JPEG压缩编码和解码,MPEG和 $P \times 64$ 的解码以及15fps MPEG和 $P \times 64$ 的编码算法。

前面已经介绍了一批多媒体系统使用的专用芯片,如Amiga系统中的Agnus,Paula以及Denise三个专用芯片,增强型CD-I系统使用的视频系统控制器、全运动视频控制器、视频合成器以及视频D/A变换器,还有一个56001(DSP)做为音响信号处理器,在DVI系统中使用的VDP(82750PA,82750PB)和VDP:(82750DA,82750DB)实践证明它们都是实用而有效的。

现在Intel公司正在开发下一代750芯片,它能为多媒体新技术的应用提供足够的计算能力,下一代的750芯片采用了微码技术。使用单指令多数据流多执行单元的结构,希望达到I Bops的运算速度。

下一代的750芯片在微码级可编程,这与上述的像素处理器750兼容,Intel公司希望用一个单片解决MPEG压缩算法以及为PC或工作站解决其它图象处理的任務。下一代的750芯片能够实时地完成(每秒30帧的速度) $P \times 64$ 压缩编码和解码的任务,但是压缩编码的时间是不一样的。同时,它还可执行标准图象格式的MPEG解压缩的算法。Intel公司计划制造可用的750系列,高档的750芯片还能够处理音频信号以及在视频编码时增加特技效果,低档产品可做为MPEG的解码器。

TI公司也在开发一个高效可编程模式的多媒体处

最低层是微码引擎,它不用主机 CPU 完成严格的时间操作任务:实时和任务调度,实时的数据压缩编码和解码,以及图象的拷贝和改变比例尺。微码引擎实际上是可编程的微码子程序集合,把它们直接加载到专用处理器芯片上的存储器中,可支持广泛变化的操作任务。可编程的方法是高性能的专用硬件与基于软件的灵活性相结合解决多媒体计算机中严格的时间同步问题较好的一种方案。

在多媒体计算机中很多关键的操作之一是多任务调度,它包括:建立位映射,管理它们并把它们送到显示缓冲区。多媒体环境中的任务调度方法是运动视频中语音和视频数据流相关同步控制特别关键问题。有些运动视频技术中使用“差帧”技术,以便减小数字式视频数据存储容量。在这里使用了在相邻帧间的逻辑差,代替每一帧的存储数据。反过来,在实时播放时多媒体软件必须能够重构原来的帧数据。

AVK 的多任务微码调度方法确保视频和音频的同步问题。它不是简单的保持视频和音频步伐一致,AVK 的调度器保证每一个数据流的额定的速度播放,因上,它用所有媒体都能接受的平均速度提供给用户。特殊情况时,处理速度或输入输出速度需要较慢的视频数据流,AVK 调度器继续让其它数据流以正常额定速度播放,但是它会建立一个误差数据流信息给同步机构。在这个层次每个数据的操作都是独立的,AVK 还具有更高层次的同步机制,即视频和音频数据的开始或停止播放的控制,它很像录象机中开始或停止的按钮。

为了支持灵活的同步机制,在微码中采用关键的多任务调度算法,AVK 和专用处理芯片保证开发者和用户具有实时处理视频和音频数据计算机的能力,而不需要考虑主计算机的能力。这是因为现在操作系统的调度器是在主机操作环境之上,用户和开发者仍然保留主机其它的操作功能。

(4)在 AVSS 或 AVK 的上层是 AVL (Audio / Video Library),开发工具、语言以及各种应用程序。

AVL 支技为管理运动视频和音频需要的特殊型的数据。这些数据规格化为数据流,多个数据流的集合称为数据流组。数据流组具有同步控制功能:例如播放、暂停、停止、前后倒带等,它们都是数据流组的操作。

AVL 还为获取和显示数据缓冲区执行该读写功能。AVL 还能控制这些数据流数据类型的属性,即能够改变音频数据流的音量以及视频数据流的色调。/此外,它还提供音频视频混合器、模拟设备接口、显示管理器、采样器以及数据流管理器多种功能。

为了使开发者和用户编制应用程序方便,不少厂商为多媒体计算机系统编制了工具软件,例如,一型 DVI 系统的工具软件软件包有下述工具:图形工具、静态图象工具、图象文件工具、实时视频工具、AVSS 文件工具、音频工具以及误差打印工具等。

四、结束语

多媒体计算机技术是当前计算机工业热点课题之一,在硬件方面 IBM / Intel 公司要在 1995 年把多媒体技术做到母板上,到 2000 年时要把这种技术做到芯片上。在软件方面,正在或已经研制了多媒体操作系统,多媒体窗口系统以及多媒体数据库管理系统。由于多媒体技术的引进,改善了人机接口的界面,扩大了计算机的应用领域,促进了全新产品和服务的出现,其中有:

1.桌上视频演播系统(Desktop Video Studio)。与此有关的是视频绘图和数字视频特技系统;计算机艺术制作系统(计算机音乐,计算机电影等)。

2.桌上出版和演示系统(Desktop Publishing and Presentation)。

3.新型办公室自动化系统,美能达(Minolta)公司最近研制的综合信息和影像管理系统采用了先进的数字影像和计算机技术,把文件扫描仪、图文传真机以及文件资料微缩系统等现代办公设备综合管理起来,以影像代替纸张,用计算机代替人工操作,这是当今办公室自动化一个新的发展方向。

多媒体计算机技术还有一个重要的应用领域是通讯工程中的多媒体终端和多媒体通讯系统。包括声、文、图在内的多媒体的电子邮件将会受到更多用户的欢迎,在此基础上发展起来的可视电话,视频会议系统将为人类提供全面的信息服务。再如计算机支持的协同工作环境(Computer Supported Cooperative Work)将使不同地点的设计人员如同面对面的讨论方案,讨论结果都存在计算机的存储器中。以上这些都将在 90 年代得到迅速发展,多媒体技术的发展将是一幅绚丽多彩的画卷。