

● 钟玉琢 王 琪 贺玉文 编著

基于对象的多媒体数据 压缩编码国际标准



MPEG-4 及其校验模型



科学出版社

基于对象的多媒体数据 压缩编码国际标准

——MPEG-4 及其校验模型

钟玉琢 王 琪 贺玉文 编著

科学出版社

2000

内 容 简 介

本书较全面地介绍了新一代压缩编码中的基于对象的压缩编码方法。全书分三部分,第一部分有三章,主要概述了多媒体计算机技术、多媒体数据压缩编码技术以及多媒体数据压缩的国际标准;第二部分有三章,主要讲述 MPEG-4 标准中的系统流、视频流和音频流;第三部分有两章,主要讲述 MPEG-4 校验模型中编码器和解码器的结构。

本书可供大专院校有关专业师生阅读,也可供从事多媒体数据压缩编码技术研制、开发和应用人员学习、参考。

图书在版编目(CIP)数据

基于对象的多媒体数据压缩编码国际标准——MPEG-4 及其校验模型/钟玉琢,王琪,贺玉文编著. - 北京:科学出版社,2000

ISBN 7-03-008775-5

I. 基... II. ①钟... ②王... ③贺... III. 多媒体-数据压缩-编码-国际标准 IV. TN919.8-65

中国版本图书馆 CIP 数据核字(2000)第 68813 号

科学出版社 出版

北京东黄城根北街 16 号
邮政编码:100717

新蕾印刷厂印刷

科学出版社发行 各地新华书店经销

*

2000 年 10 月第 一 版	开本:787 × 1092 1/16
2000 年 10 月第一次印刷	印张:30 1/2
印数:1—5 000	字数:708 000

定价:46.00 元

(如有印装质量问题,我社负责调换(环伟))

前 言

21 世纪的人类社会将是信息化社会,数字化后的信息,尤其是数字化后的视频和音频信息具有数据海量性,它给信息的存储和传输造成较大的困难,成为阻碍人类有效地获取和使用信息的瓶颈问题之一。因此,研究和开发新型有效的多媒体数据压缩编码方法,以压缩的形式存储和传输这些数据将是最好的选择。

传统的压缩编码是建立在香农(Shannon)信息论的基础上,它以经典的集合论为基础,用统计概率模型来描述信源,但是,它未考虑信息接受者的主观特性及事件本身的具体含义、重要程度和引起的后果。压缩编码的发展历程实际是以香农信息论为出发点,不断完善的过程,编码方法可以分成下述三类:

- 考虑信源的统计特性:预测编码方法,变换编码方法,矢量量化编码方法,子带-小波编码方法,神经网络编码方法等;
- 考虑人眼的视觉特性:基于方向滤波的图像编码方法,基于图像轮廓-纹理的编码方法;
- 考虑图像传递的景物特性:分形编码,基于内容的编码方法。

也有人将图像编码方法分为两代:第一代是基于数据统计,去掉的是数据冗余,称为低层压缩编码方法;第二代是基于内容,去掉的是内容冗余,其中基于对象(Object-Based)方法称为中层压缩编码方法,其中基于语义(Syntax-Based)方法称为高层压缩编码方法。基于内容压缩编码方法代表新一代的压缩方法,也是目前最活跃的领域,最早是由瑞典的 Forchheimer(1983 年)提出的,随后日本的 Harashima 等人也展示了不少研究成果。

标准化是产业化活动成功的前题,国际标准化组织于 1986 年成立了联合图片专家组 JPEG(Joint Photographic Expert Group),主要致力于制定静态图像的帧内压缩编码标准 ISO/IEC 10918,1994 年 2 月 15 日正式第一次编辑出版。国际标准化组织于 1986 年成立了运动图像压缩编码组织 MPEG(Moving Picture Expert Group),主要致力于制定运动图像的压缩编码标准,1993 年 8 月正式编辑出版了 MPEG-1 ISO/IEC 11172;1994 年又编辑出版了 MPEG-2 ISO/IEC 13818。上述三种压缩编码国际标准主要采用了第一代压缩编码方法,如预测编码、变换编码、熵编码及运动补偿。

MPEG 组织于 1999 年 1 月正式公布了 MPEG-4(ISO/IEC 14496)V 1.0 版本,1999 年 12 月又公布了 MPEG-4 V2.0 版本。MPEG-4 的制定初衷主要针对视频会议、可视电话超低比特率压缩编码的需求,在制定过程中,MPEG 组织深深感受到人们对媒体信息,特别是对视频信息的需求由播放型转向基于内容的访问、检索和操作。MPEG-4 与前面提到的 JPEG、MPEG-1 和 MPEG-2 有很大的不同,它为多媒体数据压缩编码提供了更为广阔的平台,它定义的是一种格式、一种框架,而不是具体算法,它希望建立一种更自由的通信与开发环境。于是 MPEG-4 新的目标就定义为:支持多种多媒体的应用,特别是多媒体信息基于内容的检索和访问,可根据不同的应用需求,现场配置解码器。编码系统也是开放的,可随时加入新的有效的算法模块。

为了支持对视频内容的访问,MPEG-4 引入了视频对象(Video Object)的概念,所谓对象就是在一个场景中能够访问(搜索和浏览)和操作(剪贴)的实体。对象的分割可以根据其纹理、形状、运动、2D 模型、3D 模型和高层的语义为依据。MPEG-4 采用了基于对象的压缩编码方法,它把图像和视频分割成不同的对象,分别处理,基于对象的编码除了能提高数据压缩比,还能实现许多基于内容的交互功能,它可以广泛地用在基于对象的多媒体存取、网上购物和电子商店、远程监控、医疗和教学等。

在 MPEG-4 之后,现在 MPEG 组织正在制定 MPEG-7“多媒体内容描述接口”,同时,还在酝酿制定 MPEG-21,1999 年 8 月 MPEG 主席 Leonardo Chiariglione 提出了“Technologies for E-Content”的报告,引起了各国代表团的关注,1999 年 10 月日本代表团在 MPEG 国际会议上提出制定 MPEG-21 标准的提案,以支持电子内容传输和电子贸易(N3002),1999 年 12 月 MPEG 会议通过了征集 Multimedia Framework 技术报告的议案(N3162),2000 年 3 月 MPEG 会议将成立 MPEG-21 工作组。

为了较全面介绍新一代压缩编码中的基于对象的压缩编码方法,我们编写了“基于对象的多媒体数据压缩编码国际标准——MPEG-4 及其校验模型”一书,全书分成三部分:第一部分有三章,主要概述多媒体计算机技术、多媒体数据压缩编码技术以及多媒体数据压缩的国际标准;第二部分有三章,主要讲述 MPEG-4 标准中的系统流、视频流和音频流;第三部分有两章,主要讲述 MPEG-4 校验模型中编码器和解码器的结构。全书由钟玉琢、王琪和贺玉文编写,参加编写工作的还有:杨品、蔡连红、李树青、祁卫、袁昱等人。在编写过程中,作者参考了国内外有关多媒体数据压缩编码的书刊及文献资料,还有清华大学计算机系的论文及科研成果报告。但是多媒体数据压缩编码技术正处在蓬勃发展阶段,新的文献资料我们搜集得还不完整,限于作者的学识水平,书中不足和错误之处,恳请读者给予批评指正。

本书编写过程中得到作者所在研究所及本组其他成员的大力支持,在此表示衷心感谢。

作 者

2000 年 3 月 28 日

目 录

第一部分 多媒体的关键技术:数据压缩

第一章 多媒体计算机技术	(3)
1.1 多媒体计算机的定义、分类及其发展趋势	(3)
1.1.1 利用多媒体是计算机技术发展的必然趋势	(3)
1.1.2 多媒体计算机的定义和分类	(4)
1.2 多媒体技术促进了通信、娱乐和计算机的融合	(4)
1.2.1 多媒体技术是解决常规电视数字化及高清晰度电视(HDTV)切实可行的方案	(4)
1.2.2 用多媒体技术制作 V-CD 及影视音响卡拉 OK 机	(7)
1.2.3 个人信息通信中心(PIC——Personal Information Communication Centre)	(7)
1.3 多媒体计算机技术的发展和應用	(8)
1.3.1 多媒体数据库	(9)
1.3.2 多媒体通信	(10)
1.3.3 多媒体创作工具及其应用	(11)
1.3.4 多媒体计算机的发展趋势	(12)
第二章 多媒体数据压缩的基本技术	(14)
2.1 多媒体数据压缩的必要性和可能性	(14)
2.2 统计编码	(15)
2.2.1 统计编码原理	(15)
2.2.2 霍夫曼(Huffman)编码	(17)
2.2.3 算术编码	(18)
2.2.4 行程编码(RLC)	(19)
2.3 预测编码	(20)
2.3.1 预测编码的基本原理	(20)
2.3.2 DPCM 的基本原理	(20)
2.3.3 最佳线性预测	(21)
2.3.4 自适应预测编码	(22)
2.4 变换编码	(23)
2.4.1 变换编码的基本原理	(23)
2.4.2 最佳的正交变换——K-L 变换	(24)
2.4.3 次最优正交变换——DCT 变换	(28)
2.4.4 小波变换	(31)
第三章 多媒体数据压缩标准	(37)
3.1 JPEG	(37)
3.1.1 引言	(37)
3.1.2 基于离散余弦变换(DCT)的编码	(38)

3.1.3	基于 DPCM 的预测编码	(50)
3.1.4	基于 DCT 的累进操作方式编码	(50)
3.1.5	基于 DCT 的分层编码操作方式	(51)
3.1.6	结束语	(51)
3.2	MPEG-1	(52)
3.2.1	引言	(52)
3.2.2	MPEG-1 视频	(52)
3.2.3	MPEG-1 音频	(77)
3.2.4	结束语	(78)
3.3	MPEG-2	(78)
3.4	AC-3	(80)
3.4.1	引言	(80)
3.4.2	AC-3 编码原理	(81)
3.4.3	AC-3 编码的技术要点	(82)
3.4.4	结束语	(85)
3.5	MPEG-4	(86)
3.6	MPEG-7	(87)
3.6.1	概述	(87)
3.6.2	MPEG-7 的制定工作	(89)

第二部分 MPEG-4 标准

第四章	MPEG-4 系统流	(93)
4.1	系统体系结构	(93)
4.1.1	系统终端模型	(94)
4.1.2	传输层定义	(95)
4.1.3	同步层定义	(95)
4.1.4	压缩层定义	(95)
4.2	系统解码器模型	(97)
4.2.1	系统解码器模型的基本概念	(97)
4.2.2	定时模型规范	(98)
4.2.3	缓冲区模型规范	(100)
4.3	对象描述框架	(101)
4.3.1	基本数据结构	(103)
4.3.2	属性智能管理和保护	(105)
4.3.3	对象内容信息	(106)
4.3.4	对象描述符流	(108)
4.3.5	对象描述符的部件	(111)
4.3.6	对象描述框架的使用规则	(129)
4.3.7	IPMP 系统接口的使用	(134)
4.4	场景描述	(138)
4.4.1	场景描述概念	(138)
4.4.2	场景的二进制格式(BIFS)语法	(150)
4.4.3	结点语义	(203)

4.5 基本码流的同步	(206)
4.5.1 同步层概念	(207)
4.5.2 基本码流接口	(212)
4.6 基本码流的多路合成	(213)
4.7 句法构造描述语言	(216)
4.7.1 基本的数据类型	(217)
4.7.2 句法流控制	(217)
4.8 附录:基本术语	(218)
第五章 MPEG-4 视频流	(221)
5.1 视频流的语法和语义	(221)
5.1.1 视频流中的数据结构	(221)
5.1.2 视频流语法	(230)
5.1.3 视频流语义	(271)
5.2 视频流解码	(321)
5.2.1 视频解码过程	(322)
5.2.2 高级语法结构	(322)
5.2.3 视频对象平面(VOP)重建	(323)
5.2.4 纹理解码	(323)
5.2.5 形状解码	(333)
5.2.6 运动补偿解码	(346)
5.2.7 交织视频解码	(360)
5.2.8 Sprite 解码	(368)
5.2.9 一般扩展性解码	(374)
5.2.10 静态纹理对象解码	(380)
5.2.11 网格对象解码	(392)
5.2.12 人脸对象解码	(397)
5.2.13 解码过程的输出	(401)
5.3 视频系统合成	(401)
5.3.1 时域扩展性合成	(401)
5.3.2 Sprite 合成	(403)
5.3.3 网格对象合成	(403)
5.4 框架和级	(404)
5.4.1 视觉对象类型	(404)
5.4.2 视觉框架	(405)
5.4.3 视觉框架和级	(407)
第六章 MPEG-4 音频流	(409)
6.1 概述	(409)
6.1.1 MPEG-4 音频	(409)
6.1.2 MPEG-4 音频中的新概念	(409)
6.1.3 MPEG-4 音频的性能	(410)
6.2 技术特点	(413)
6.2.1 MPEG-4 音频对象类型	(413)
6.2.2 音频框架和级	(415)

6.3 语音编码	(418)
6.3.1 CELP	(418)
6.3.2 HVXC	(421)
6.4 通用音频编码	(423)
6.4.1 编码器和解码器框图	(423)
6.4.2 编码器和解码器工具	(423)
6.5 结构化音频	(428)
6.5.1 目标	(428)
6.5.2 主要元素	(428)

第三部分 MPEG-4 校验模型

第七章 编码器结构	(433)
7.1 MPEG-4 校验模型编码	(433)
7.1.1 MPEG-4 校验模型(VM)的编码结构	(433)
7.1.2 VM 中的数据结构	(434)
7.1.3 基于 VOP 的编码	(434)
7.1.4 视频对象平面的编码结构	(437)
7.2 形状编码	(438)
7.2.1 VM 中的形状编码方法	(438)
7.2.2 二值形状编码方法	(439)
7.2.3 灰度级的形状编码	(444)
7.3 运动估计和补偿	(445)
7.3.1 填充技术	(446)
7.3.2 基本的运动估计技术	(447)
7.3.3 无限的运动估计	(449)
7.3.4 高级预测模式	(449)
7.3.5 重叠的运动预测	(450)
7.4 纹理编码	(451)
7.4.1 DCT 变换	(451)
7.4.2 量化方法	(452)
7.4.3 I-VOP 和 P-VOP 帧内编码 DC、AC 的预测	(454)
7.4.4 量化变换系数的 VLC 编码	(455)
7.5 B-VOP 的编码	(456)
7.5.1 B-VOP 编码模式	(456)
7.5.2 编码模式选择	(457)
7.5.3 运动矢量编码	(458)
7.6 容错和码率控制	(458)
7.6.1 MPEG-4 容错方式	(458)
7.6.2 码率控制	(460)
7.7 可扩展性编码	(464)
7.7.1 空域扩展编码	(465)
7.7.2 时域扩展编码	(466)
7.8 sprite 编码	(468)

7.8.1	sprite 介绍	(468)
7.8.2	sprite 的生成方法	(469)
7.8.3	sprite 编码	(470)
第八章	解码器结构	(472)
8.1	视频对象平面解码器结构	(472)
8.2	时域预测结构	(472)
8.3	可扩展性解码	(473)
8.3.1	空域扩展解码	(474)
8.3.2	时域扩展解码	(474)
8.4	合成器定义	(475)
参考文献		(478)

第一部分 多媒体的关键技术:数据压缩

第一章 多媒体计算机技术

应用多媒体技术是 20 世纪 90 年代计算机的时代特征,是 90 年代计算机的又一次革命。多媒体是针对单媒体而言的。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;一是指信息的载体,如数字、文字、声音、图形和图像。多媒体技术中的媒体是指后者。

1.1 多媒体计算机的定义、分类及其发展趋势

1.1.1 利用多媒体是计算机技术发展的必然趋势

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 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。

20 世纪 50 年代~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,对于文化水平较低的人,特别是非英语国家的人,仍然是件困难的事情。

从 80 年代开始,人们致力于研究将声音、图形和图像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。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(Windows 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唯一能同 IBM 公司相抗衡的力量。今天,国际上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;
- 超大规模集成电路的速度增加了;
- CD-ROM 可作为低成本、大容量 PC 机的只读存储器(可换的 5 英寸盘片,每片容量为 600MB;以及 DVD,单面 4.7GB);
- 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 10 年前单色文本/图形子系统转变成今天的色彩丰富、高清晰度显示子系统,同时能够做到全屏幕、全运动的视频图像,高清晰度的静态图像,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。

综上所述,无论从半导体的发展还是从计算机进步的角度,或者从普及计算机应用、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。

1.1.2 多媒体计算机的定义和分类

1. 多媒体计算机(Multimedia Computing)的定义

计算机综合处理多种媒体信息:文本、图形、图像、音频和视频,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。

即多媒体计算机具有信息载体多样性、集成性和交互性。

要把一台普通的计算机变成多媒体计算机要解决的关键技术是:

- 视频音频信号获取技术;
- 多媒体数据压缩编码和解码技术;
- 视频音频数据的实时处理和特技;
- 视频音频数据的输出技术。

2. 多媒体计算机的分类

从开发和生产厂商以及应用的角度出发,可以将多媒体计算机分成两大类:

一类是家电制造厂商研制的电视计算机 Teleputer,是把 CPU 放到家电中,通过编程控制管理电视机、音响,有人称它为“灵巧”电视(Smart TV);

另一类是计算机制造厂商研制的计算机电视(Compuvision),是采用微处理器(80x86、68xxx)作为 CPU,其他设备还有 VGA 卡、CD-ROM、音响设备以及扩展的多窗口系统,有人说它的发展方向是 TV-Killer。

1.2 多媒体技术促进了通信、娱乐和计算机的融合

所谓通信、娱乐和计算机的融合,即把消费类电子产品(电话、电视、图文传真机、音响、录像机等)与计算机融为一体,由计算机完成视频音频信号的采集、压缩和解压缩,实时处理视频和音频及其特技、视频的多窗口显示及音频的立体声输出,从而形成新一代的产品,为人类的生活和工作提供全新的信息服务。

1.2.1 多媒体技术是解决常规电视数字化及高清晰度电视(HDTV)切实可行的方案

近几年,在美国成立了一个高级电视研究集团(ARTC),它采用 MPEG 压缩编码标准、同时播出方案、打包数据结构以及双层传输技术,较之早些时候日本推出的模拟式的高清晰度电视,这是一个切实可行的方案。

目前研制的 HDTV 有下述几个特点:

(1) 采用国际标准的压缩编码算法 MPEG-2,这意味着它能与以 MPEG, JPEG 压缩编码算法为基础的多媒体计算机兼容,并与其互连通信;

(2) 采用打包数据结构,当电视信号在视频通道传输时,图像和声音数据分成不同分量,在大多数情况下,这些分量要遵循大小和次序的限制。HDTV 将图像和声音信息以及用于多媒体服务的附加数据以包的方式传送。这些数据可任意大小,只要它们符合频道特性,能以随机次序传送,这些数据包能够动态分配,使 HDTV 能与计算机、多媒体娱乐、教育系统及录像机通信,打开了将电视机、计算机和通信融为一体,通向更灵活服务领域的大门。

(3) 采用双层传输技术,双层传输技术保证 HDTV 的可靠性和抗干扰性。它将信息分开传送,最重要的数据放到具有高优先级的载波上传输,其他数据则放到具有标准优先级的载波上传输。

采用多媒体计算机技术制造的 HDTV,可以支持任意分辨率的输出,输入输出分辨率可以独立,输出分辨率可以任意变化,可以用任意窗口尺寸输出。与此同时,它还能赋予 HDTV 很多新的功能,如图形功能、视频音频特技以及交互式功能。

常规电视数字化技术及交互式电视技术(包括点播电视技术 VOD)都是当前世界上的热点课题,最佳的解决办法是采用数字式视频、数字式音频及 MPEG 压缩编码算法,以便于数据传输、存储及计算机控制和管理。世界上很多大的公司都在从事这方面的开发和研究。几年前,汤姆逊(Thomson)消费电子公司制定的战略目标是,做常规电视数字化的先驱。具体做法是通过休斯银河(Hughes Galaxy)601 卫星,开创世界首次全数字直接到户的卫星广播业务(DSS——Digital Satellites System 及 DBS——Direct Broadcast Service)。它能传送激光视盘和激光音盘的质量,使消费者很容易获得 120~150 个频道最受欢迎的电视节目。用户端只需投资 600~800 美元,购置一个易于安装的 18 英寸或常规碟形天线、一个与录像机体积差不多的接收机/解码器以及一个易于控制和操作的遥控器。汤姆逊公司的 DSS 系统的销售目标是北美的每个家庭都选购一台,使他们有机会在家中观看卫星数字电视,这些新产品将改变用户娱乐、采购、学习甚至工作的方式。我国中央电视台也租用了两个卫星信道,开始向全世界播放常规电视数字化节目。同时在我国“九五”科技攻关计划中,也安排了卫星数字电视机顶盒的设计和制造任务。

如何解决常规电视和高清晰度电视同播(Simulcast)的问题,就像彩色电视采用 YUV 方案和黑白电视兼容同等重要。最近推出的国际标准 MPEG-2,采用了分层的编码体系(Hierarchic Coding),提供了较好的可扩充性(Scalability)及互操作能力(Interoperability)。MPEG-2 整个视频比特流由逐级嵌入的若干层组成,这样不同复杂度的解码器可根据自身的能力从同一比特流中抽出不同层进行解码,得到不同质量、不同时间/空间分辨率的视频信号,分层编码使同一比特流能适应不同特性的解码器,极大地提高了系统的灵活性、有效性,同时也为视频通信系统向更高时间/空间分辨率过渡提供了技术保证。为了实现分层编码, MPEG-2 提供了四种工具:空间可扩展性(Spatial Scalability)、时间可扩充性(Temporal Scalability)、信噪比可扩充性(SNR Scalability)及数据分块(Data Partitioning)。为了支持灵活的性能价格比, MPEG-2 还提供了框架(Profile)与等级(Level)的概念,给出了丰富的编码方法、灵活的操作模式,以适合不同场合的需要。

最近几年的热点课题是交互式电视技术(ITV),因为交互式电视技术有较好的发展环

境、较好的经济、社会效益及广阔的应用前景。从美国宣布“信息高速公路”计划后,全球掀起了信息高速公路的热潮,纷纷投巨资建设国家信息基础设施(NII)。我国也在积极、慎重地开展 CNII 计划,1994 年 9 月正式建成开通了我国公用数字数据网(CHINADDN),它可为用户提供 $N \times 64\text{Kbps}$ ($N = 1 \sim 31$), $2.4 \sim 19.6\text{Kbps}$ 数字专线业务,用户可以进行单向、双向及 N 向的广播、会议电视等点对点、点对多点以及多点对多点的传输业务;1993 年 9 月正式建成开通了我国公用分组交换数据网(CHINAPAC),该网已覆盖所有省会城市及地、县、乡 2000 多个点,总容量已达 10 万多个。1995 年 6 月正式开通了我国公用计算机互联网(CHINANET),它是在 CHINADDN 和 CHINAPAC 基础上的增值网。此外,广州、深圳、上海等城市正在积极筹建 VOD(Video On Demand)系统。

未来信息高速公路上,传递最多的信息是交互式电视和其他视频信息,交互式电视有最广大的用户,潜在的用户量可以是几亿或数十亿。交互式电视用户可以坐在家里的机顶盒(STB——Set Top Box)前,通过单键遥控器和菜单选择自己喜欢的电影、电视和新闻,它还可以提供交互式电视教育、电视采购、视频游戏以及各种方便的电视、电话和数据信息服务。

交互式电视系统和分布式多媒体数据管理系统从机理上是完全一样的,交互式电视台把新闻和其他节目,经过视频和音频的压缩存储到数据库中,用户可以通过机顶盒而不是多媒体工作站,通过网络点播各种广播节目。交互式电视最常用的是节目间的交互,即 VOD 系统,典型的 VOD 系统主要由下述四部分组成:

- 视频服务器;
- 编码器/路由器;
- 用户请求计算机和记帐计算机;
- 机顶盒。

多媒体计算机技术遮常规电视和高清晰度电视、影视节目制作中的应用可以分成两个层次。一是影视画面的制作,影视画面的生成可以采用计算机软件生成二维、三维动画画面;摄像机摄制真实的影视画面后采用数字图像处理技术制作影视画面,最后是采用计算机生成和实拍结合后用图像处理技术制作影视特技画面。这方面的成功示例已经很多:美国惊险科幻影片《侏罗纪公园》(Jurassic Park)中史前动物恐龙的许多精彩镜头都是用计算机制作,它荣获了奥斯卡最佳视觉效果奖;电影《刚果》、《真实的谎言》、《阿甘正传》及《狮子王》中很多画面都是用计算机制作的,它们都产生了极佳的影视效果。另一个层次是影视的后期制作,例如现在常用的数字式非线性编辑器,实质上是一台多媒体计算机,它需要有广播级质量的视频音频的获取和输出、压缩和解压缩、实时处理和特技以及编辑功能。目前,美国、加拿大、德国的一些公司研制生产了一些较好的产品在世界销售,例如:苹果机为平台的非线性编辑器以 Avid 产品占主流;Truevision 公司的 TARGA2000 是非线性编辑器的核心板卡;此外还有 Miro、Ops 以及 Fast 等公司的产品。非线性编辑器长远的发展方向,一方面是提高质量、速度,降低成本,更重要的发展是网络化的编辑器,它的好处是多个编辑协同工作,还可以连通“素材库”、“节目库”和“播放库”,可以大大提高编辑的效率和质量。

1.2.2 用多媒体技术制作 V-CD 及影视音响卡拉 OK 机

多媒体数据压缩和解压缩技术是多媒体计算机系统的关键技术,全世界很多半导体厂商都在积极开发、生产、销售图像压缩和解压缩芯片。美国 C-Cube 公司从 1989 年成立以来,已投资数千万美元,先后开发生产了 CL-550、CL-560 及 CL-460,它们是 JPEG 静态图像压缩和解压缩处理器以及 MPEG-1 解码器,同时还提供板级产品及开发系统。经过几年努力宣传和推销,虽然连续几年在 Comdex 博览会上深受观众的欢迎和青睐,但该公司的产品在多媒体计算机领域中的销量欠佳。为了打开芯片的市场销路,该公司的董事长决定把芯片用到 VCD 的播放机中。

VCD 是 JVC、Philips、Matsushita 及 Sony 联合制定以数字电视视盘技术标准,它于 1993 年问世。我国安徽合肥万燕(Wyai)公司(与美国、韩国合资)在世界上首先利用 MPEG 国际标准和 CD 光盘技术,研制了全功能影视音响卡拉 OK 机 CDK-320。

CDK-320 采用 VCD 的标准,一片 VCD 盘片可以存放 70 分钟影视节目,利用 MPEG-1 的音频编码技术将声音压缩到原来的六分之一,再利用帧内和帧间动态图像压缩技术,使图像的分辨率可达:NTSC 制为 $352 \times 240 \times 30$, PAL 制为 $352 \times 288 \times 25$,声音质量可达 CD-DA 的质量。目前国内 VCD 的售价为 2 000 元人民币左右。VCD 盘片的售价将是 LD(激光影碟)的五分之一,是录像磁带的三分之一。

目前 VCD 播放系统主要有下述两种产品:

(1) MPEG 播放卡,目前见到的最简单的 MPEG 播放卡是由 Mediamatics 公司研制的基于 Mx501™芯片的 MPEG 播放卡,全卡只有三块芯片, Mx501, 4Mb DRAM 及 40MHz OSC。售价为 30 美元。它的硬软件支撑环境是:具有 PCI 总线的 PC 机, CD-ROM、图形显示适配器及 Windows 3.0。

(2) VCD 播放机,它由 CD-ROM 驱动器、MPEG 解压卡及控制操作电路组成。

下一步应该是 DVD 播放卡和 DVD 播放机,它的工作原理与 VCD 播放机一样,只是视频和音频编解码标准不是 MPEG-1,而是 MPEG-2 或 AC-3。

1.2.3 个人信息通信中心(PIC——Personal Information Communication Centre)

由于采用多媒体技术,使一台个人计算机具有录音电话、可视电话机、图文传真机、立体声音响设备、电视机和录像机等多种功能,有人称它为 PIC。如果计算机再配备丰富的软件并联网上, PIC 还能翻阅旧的传真文件。草拟编辑文件并控制发送,同时还能具有多媒体邮件功能。

PIC 也有人称它为 PAC——Personal Activity Centre 和 PDA——Personal Digital Assistant,早期的 PAC 产品,如深圳英达利信息通信产品有限公司于 1993 年首先推出了 MPEG-K100 PAC 产品,它的硬件平台用 JT 486 SMV 主板, CPU 为 Intel 486 DX-33, VESA LOCAL BUS, 256 KB CACHE, 8MB RAM(可扩展为 32MB), I/O 总线速率 8.33MHz。采用 SVGA、视频处理和音频处理功能合为一体的 Ecomedia I 多媒体卡,接上 Ecovision 电视调谐卡,接上天线即可以收看电视。MPC-K 100 还配有 Sound Power 语音卡,该卡具有先进的数字化语音放音及数字化语音录音功能。

中国长城计算机集团公司生产制造的 GW-OA5IN1 长城五合一办公通信系统:一根电