

实用多媒体技术及其

C语言实现

郝 杰 吴元清 郑 榕



电子工业出版社

3.1.2
456

实用多媒体技术及其 C 语言实现

郝杰 吴元清 郑榕 著



電子工業出版社

9610083

内容提要

本书详细地剖析了各种多媒体实用技术的基本原理和算法实现。主要内容包括电话二次拨号识别、语音压缩编码 ACPCM 实现,图像压缩编码 JPEG 实现,图像的无损压缩编码,自适应彩色量化技术,微机扩展内存统一管理。书中提供大量实用 C 语言源程序,可让读者学习如何有效地使用多媒体技术,以及如体制作多媒体实用程序。

实用多媒体技术及其 C 语言实现

郝杰 吴元清 郑榕 著

责任编辑:卢 美

*

电子工业出版社出版

北京市海淀区万寿路 173 信箱(100036)

电子工业出版社发行 各地新华书店经销

电子工业出版社计算机排版室排版

北京大中印刷厂印刷

*

开本:787×1092 1/16 印张:34 字数:833千字

1995年11月第一版 1995年11月北京第一次印刷

印数:5000册 定价:40.00元

ISBN 7-5053-3131-0/TP·1114

前 言

近几年来,随着超大规模集成电路技术、光存储技术与计算机技术的发展,以及计算机网络技术特别是 Internet 的广泛使用,多媒体技术得到了迅速的发展。多媒体系统是对文本、语音、图像、影像、动画等多种媒体进行综合处理的集成信息系统,多媒体系统具有友好的人机界面。多媒体技术的不断发展最终将推动人工智能技术的进步。

多媒体系统的一个重要特点是其具有很强的交互性,用户可通过多种信息接口,如话筒、摄像机、扫描仪、电写设备以及通常的键盘、鼠标、触摸屏等与系统进行通信。多媒体技术的最新发展方向是分布式多媒体信息系统,即网络多媒体技术,如 Internet 就是一个最大的分布式多媒体系统。分布式多媒体系统的一个优点是信息查询与信息获取不再受地域的限制。

实现多媒体系统,最关键的技术是语音、图像的处理技术,本书介绍几种在多媒体制作过程中必不可少的语音、图像处理的实用技术,并且围绕国际标准化组织(ISO)及国际电报电话咨询委员会(CCITT)的有关标准具体讲述实现方法,而且给出 C 语言源程序,供读者参考。

本书共分十章。第一章主要介绍多媒体技术的基本概念,以及发展方向。第二章介绍 Internet 上的多媒体电子邮件标准 MIME。第三章介绍语音处理的一些基本技术。第四章介绍电话的二次拨号,即拨号脉冲识别技术。第五章介绍电话语音的自适应差分脉冲编码调制(ADPCM)技术。第六章介绍图像处理的基本技术及基本概念。第七章介绍静止图像的压缩编码 JPEG 标准。第八章介绍二值图像的无损压缩编码技术。第九章介绍彩色图像的自适应量化技术。第十章给出一个微机内存管理程序。

本书的内容比较具体,其特点是实用性强,可供通信、图像、计算机、信息处理等专业的高年级学生及科技人员参考。

本书的编写过程中得到清华大学电子工程系陆大纶教授、李星教授的关心与支持以及严程老师和刘今晖同学的帮助,特在此表示感谢。另外,特别感谢电子工业出版社的铲美在本书出版过程中给予的帮助与支持。

由于多媒体技术的发展日新月异,本书定有不尽之处,而且作者水平有限,错误在所难免,欢迎广大读者给以批评指正。

作者

1995. 4. 27

目 录

第一部分 多媒体技术原理

第一章 多媒体技术基础	(3)
1.1 分布式多媒体系统的特点及其应用	(3)
1.2 超大规模集成电路技术对多媒体的影响	(7)
1.3 通信网的发展	(8)
1.4 多媒体 CD-ROM 技术	(9)
1.5 数字电视和多媒体会议	(11)
1.6 多媒体电子邮件	(13)
1.7 Internet 信息鼠 Gopher	(15)
1.8 Internet 广域信息服务 WAIS	(18)
1.9 Internet 环球信息网 WWW 与 Mosaic	(20)
第二章 多媒体电子邮件	(24)
2.1 多媒体电子邮件的信体扩展	(24)
2.2 多媒体电子邮件的信头扩展	(41)

第二部分 语音媒体处理技术

第三章 语音媒体处理技术基础	(47)
3.1 语音编码技术基础	(47)
3.2 语音合成技术基础	(55)
3.3 语音识别技术基础	(56)
第四章 电话的二次拨号技术	(59)
4.1 拨号脉冲识别的用途	(59)
4.2 二次拨号的原理	(59)
4.3 二次拨号识别算法	(62)
4.4 二次拨号识别源程序	(65)
第五章 语音信号的自适应差分脉冲编码技术	(74)
5.1 ADPCM 编解码原理	(74)
5.2 ADPCM 编解码算法实现	(82)
5.3 ADPCM 编解码源程序	(96)

第三部分 图像媒体处理技术

第六章 图像媒体处理技术基础	(137)
6.1 图像编码的方法	(137)
6.2 二值图像编码技术基础	(138)
6.3 彩色静止图像编码技术基础	(140)
6.4 序列图像编码技术基础	(142)

6.5 图像编码技术的发展方向	(144)
第七章 Windows 下的具有标准接口和快速算法的 JPEG 编码	(146)
7.1 JPEG 编码技术基础	(146)
7.2 JPEG 算法的基本原理	(148)
7.3 快速 DCT 算法	(162)
7.4 Huffman 算法	(186)
7.5 JPEG 编解码标准接口	(216)
7.6 JPEG 编解码源程序	(350)
第八章 静止图像无损压缩编码技术	(398)
8.1 二值图像编码的应用	(398)
8.2 几种常用的二值图像编码方法	(400)
8.3 改进的 READ 编码	(418)
8.4 二值图像编解码源程序	(426)
第九章 自适应彩色量化技术原理与算法	(466)
9.1 矢量量化技术原理及其算法	(466)
9.2 彩色量化技术的原理	(468)
9.3 Heckbert 彩色量化算法	(471)
9.4 增量距离计算的原理	(474)
9.5 彩色量化算法的 C 语言实现	(477)
附 录	
第十章 方便有效地使用微机扩展内存编程	(515)
10.1 微机内存结构	(515)
10.2 微机扩展内存统一管理的 C 语言源程序	(515)

第一部分

多媒体技术原理

第一章 多媒体技术基础

第二章 多媒体电子邮件

第一章 多媒体技术基础

多媒体(Multimedia)技术是通过多种媒体(Media)如文本(Text)、语音(Voice)、图像(Image)、图形(Graphics)、影像(Video)及音频(Audio)等形式,利用计算机对信息进行加工、存储、恢复及传输等处理,以此实现通信的一种技术。通信最初只局限于声音、纸张等简单媒体,这个世纪则出现了电话、电视、计算机网等多种通信媒体,而多媒体系统就是多种通信媒体的集成;可以简单到文本与图形的集成,也可以复杂到动画(Animation)、音频及影像的集成。

近几年来,随着超大规模集成电路(VLSI, Very Large Scale Integration)技术的进步,CD-ROM 的广泛使用及海量(Mass Capacity)存储设备的出现,高速综合业务数字网(ISDN, Integrated Service Digital Net)性能的提高,以及工作站(Work Station)的价格下降,多媒体技术得到了迅速的发展。多媒体技术在许多领域,如医学、教育、旅游、金融、保险、管理、出版、广告等方面得到了广泛的应用;与此同时,各种多媒体技术相继成立其专家小组,并已经或将要制定其国际标准,这种标准化又将推动多媒体工业的发展。

随着 Internet 的广泛应用,分布式多媒体系统(Distributed Multimedia System)越来越受到更多的重视。本章各节依次介绍分布式多媒体系统的特点及其应用,超大规模集成电路技术对多媒体技术的影响,通信网的发展,只读光盘存储器(CD-ROM),数字电视与多媒体会议,多媒体电子邮件及 Internet 的几个信息查询/获取工具。

1.1 分布式多媒体系统的特点及其应用

分布式多媒体系统是指由一个或多个网络连接的多台工作站构成的多媒体系统,与此相对的是仅由一台工作站或微机构成的多媒体系统。本节介绍多媒体信息技术的发展趋势,分布式多媒体技术的应用,分布式多媒体系统的简单模型,以及现有的分布式多媒体系统及其有关标准。

1.1.1 多媒体信息技术的发展趋势

为了认识分布式多媒体技术的重要性,我们从终端用户(End User)和整个系统这两个角度来观察多媒体信息技术的发展趋势。

终端用户对多媒体信息系统的认识可以分为两个方面:信息的媒体形式和利用这些媒体形式进行有效通信的能力。近几十年来,信息媒体种类不断增加,通信不再局限于简单的声音或文字,这种增长趋势在本世纪末更加明显。这种信息媒体的增加为教育、商业等众多领域提供了便利,但是信息的不断增长也带来一些问题,比较突出的是“信息爆炸”(Information Overload)。在“信息爆炸”的情况下,调查、分析繁多的信息成了商业、政府及教育事业等领域的决策者的沉重负担。

另一方面,许多现代化的组织机构覆盖相当大的地域,组织成员们共享某些信息,并且

相互交换信息。因此,有效的通信对这类组织是至关重要的。若没有可靠、有效的通信,决策者们就不能获得正确的信息,甚至有“信息贫乏”(Information Stavation)的危险。

因此,多媒体信息系统的终端用户有可能面临两个问题:信息爆炸和信息贫乏,为了帮助用户获得正确的信息,必须实现信息管理的自动化。

从整个系统的角度来看,多媒体信息系统的-一个发展趋势是各种媒体越来越集成化。信息技术可看作是计算机技术与通信技术的组合。最初计算机只处理数值运算,随后,文本处理成了计算机的重要用途;同时,通信技术对文字、数字数据的传输的支持也得到了发展。现在,图形、图像已被集成到多媒体信息系统中,而更吸引人的是音频、影像的集成。

1.1.2 分布式多媒体技术的应用

多媒体技术具有非常广泛的应用领域。多媒体的引入可以加强现有系统的可操作性,而且多媒体的出现会导致一些新的领域的产生。以下简单讨论多媒体的技术在各领域内的应用,并突出分布式多媒体技术的重要作用。

办公自动化(Office Automation)

在办公室中,大量的、各种形式的信息需要被处理,因此有许多多媒体系统是为办公自动化而开发的。

实现办公自动化,首先要解决多媒体文档(Multimedia Document)的产生、浏览、存储与获取。集成了文字、图形的字处理商用软件已经流行了数年,并且推动了桌面排版(Desktop Publishing)的发展,这类应用的发展方向是将影像、音频集成到文档中。信息的浏览是多媒体环境中的一种重要应用,但是信息的种类多、数目大给浏览造成了困难,因此多媒体信息的表示及查询方法需要认真研究,在这方面,超媒体(Hypermedia)技术是一种有力的工具。多媒体信息的数据量是相当大的,因此多媒体信息的存储必须综合考虑文档的性质、所含信息的结构以及用作存储设备的物理媒介的特点。多媒体文档通常由若干个部分组合而成,每个部分对应一种媒体,存储多媒体文档时可将各部分分别存储到适当的物理媒介中,如将文本存储在磁盘中,而将影像存储在高速光存储设备中。多媒体文档的获取是指将文档的各个组成部分查询出来并且传送给用户的工作站,这种信息获取过程涉及到查询语言(Query Language)、文档分类知识及文档建模等技术,因此比较复杂。在这方面,已产生了一些标准,如办公文档结构标准(ODA,Office Document Architecture)。多媒体文档获取的难点在于存储介质输入输出带宽的限制及文档各部分传输时的同步要求。

在办公自动化中,多媒体电子邮件(Multimedia Electronic Mail)是一种重要的应用。不同于文本邮件,多媒体邮件的数据量非常大,而且结构非常复杂。例如一个多媒体邮件由许多兆字节构成,且包含多种媒体数据。多媒体邮件要求接收端的存储能力足够大,但是有的多媒体邮件系统不把数据量太大的邮件成分如影像在接收端复制,只允许接收端对其进行查询和浏览。有关多媒体电子邮件的标准有国际电报电话咨询委员会(CCITT)的X.400建议和Internet多媒体邮件扩展标准(MIME,Multipurpose Internet Mail Extensions)。

公共事业

教育、金融和医疗保健这三种事业都需要多媒体信息技术的支持。实际上信息技术的许

多重大发展是从这三个领域的应用中产生的。

在教育中使用多媒体技术已经有很多年的历史了,文本、图形、图像及声音的引入补充并增强了传统的教学方法,使得学习变成一项更有趣的活动。光盘(Optical Disk)技术的出现使多媒体教学得到突破性发展,许多多媒体教学软件可以通过光盘发行,供学校或其他培训机构使用。但是,制作光盘的成本比较高,因此需要开发分布式的多媒体系统,使一个信息中心可以为许多个远程用户提供多媒体教学服务。随着教育系统的自动化,各教育团体都将建立其信息库,而且各团体之间的信息交流也日益增加,因此将来的教育系统需要制定相应的标准,来协调各教育组织间的协同工作。

金融业被看作世界上主要的信息消费行业,一个金融组织的成败依赖于它的信息收集与通信系统的效率。金融市场的性质要求决策者以最快的速度获得相关信息,根据市场行情决定投资方案。决策者得到的信息往往来自不同的渠道,如一份公司报告或一条政府声明等,这些信息具有各种媒体类型,因此信息的接收、通信及表示必须能够适应各种媒体类型。为了支持信息的处理和通信,需要建立分布式的金融信息系统。

医疗保健事业中,各种医疗保健设备产生大量的、各种形式的信息。如扫描仪输出的影像,X光透视机输出的高质量图像等。利用分布式多媒体技术可以建立医疗信息系统,可以使医疗机构不受地域的限制,方便、快捷地获取所需的信息,这种分布式的信息系统也有利于医疗保健事业的行政管理。

科学技术

科学技术领域也需要多媒体技术的支持,但是多媒体在科学技术中的应用与其在办公自动化中的应用有些差异。办公自动化要求多媒体技术提供人与人之间的通信,如音频、影像及由此构成的电视会议等,在这种情况下,人直接接收多媒体信息,并对信息进行处理,因此可以降低信息的质量,以减轻处理器及存储设备的负担。科学技术方面的多媒体应用则强调对信息进行分析,根据其内容得到某些结论,因此对信息的质量有较严格的限制。

计算机支持的协同工作(CSCW, Computer Supported Co-operative Wordking)

协同工作系统的发展已有五年,协同工作研究涉及到计算机科学、网络技术、社会学及心理学等诸多方面。协同工作要求有多媒体电子邮件、多媒体会议的支持。协同工作的例子有协同编辑(Co-authoring)及讨论系统(Argumentation System)等。

其他应用

分布式多媒体技术还在出版、旅游、房地产、家庭生活等方面有重要用途。

1.1.3 分布式多媒体系统的模型

为了对分布式多媒体技术有一个初步认识,这里介绍一个分布式多媒体系统的简单模型。

在这个模型中,信息空间由结点(Node)和连接(Link)组成。一个结点有若干个属性(Attribute),最重要的属性是该结点的内容(Content),其次是结点的媒体类型(Media Type),含有多种媒体类型数据的结点称为“合成结点”(Compound Node)。

连接是两结点间或两结点的部分内容间的路径,连接由信息指针(Anchor)来实现。连

接的起始端称为源指针(Source Anchor),连接的末端称为目标指针(Target Anchor)。一个连接只能有一个源指针,却可以有多个目标指针。不含源指针的结点称为叶结点(Leaf Node)。

在分布式多媒体系统中,用户可以“传递”(Traverse)一个连接,使目标结点得到显示。通常这种系统采用“客户机-服务器”模式(Client-server Mode),用户工作站上的客户程序完成结点内容的获取与显示。

“可搜索结点”(Searchable Node)是一类特殊的结点,它可以指示用户输入关键字(Keywords),并根据关键字对数据库进行搜索,并将搜索结果传送给用户的客户程序。

1.1.4 现有系统及标准

现在,Internet 上的主要的分布式信息查询及获取系统有信息鼠(Gopher)、广域信息服务(WAIS, Wide Area Information Service)和环球信息网(WWW, World Wide Web),这三种系统都工作于“客户机-服务器”模式,并且都对多媒体数据有不同程度的支持。

Gopher 系统的信息空间称为“Gopher 空间”(Gopherspace)。Gopher 空间内有三种结点:目录结点,其内容是指向其他结点的若干个连接的菜单;叶结点,含有文本、图像及其他媒体类型的数据;可搜索结点,允许用户对结点进行关键字搜索。对用户来说,Gopher 空间的组织方式类似于计算机中的文件系统的分级结构。Gopher 协议支持的媒体类型不太多,它的扩展即 Gopher+对多媒体数据有较好的支持。Gopher 服务器易于建立,因此 Gopher 在 Internet 上应用较广泛,已有 1000 多个 Gopher 服务器。

WAIS 系统允许用户对远程数据库进行搜索。用户可以在“WAIS 空间”内选择一个或多个数据库进行关键字搜索,搜索的结果是一组指向与关键字匹配的结点的连接构成的菜单,用户可以在此菜单内选择所需的信息。在 WAIS 空间中,有一个特殊的数据库,称为“数据库的数据库”(Database of Database),包含所有 WAIS 数据库的有关信息。WAIS 支持图像媒体数据,但对图像信息的索引只限于图像文件名。WAIS 不是一个通用信息系统,而是一个针对文本信息的专用检索工具。

WWW 是一个大规模的分布式超媒体系统。WWW 信息空间的内结点可以具有多种媒体类型,大多数文本结点用“超文本描述语言”(HTML, Hypertext Markup Language)表示。WWW 支持可搜索结点。WWW 的寻址方式使得 Gopher 服务器、WAIS 服务器、无名氏文件传输协议(Anonymous File Transfer Protocol, Anonymous FTP)服务器及使用其他协议的信息源都可在 WWW 信息空间内查询,因此 WWW 是一个集成的分布式超媒体系统。美国伊利诺斯州立大学(University of Illinois)的国家超级计算中心(NCSA, National Centre for Supercomputing Applications)开发的 WWW 客户软件 Mosaic 在 Internet 上得到了普遍应用。Mosaic 通过“外部播放器”(External Viwer)支持图像、语音、影像等多媒体数据的播放。

目前已有一些关于超媒体技术的标准。标准生成描述语言(SGML, Standard Generalized Markup Language)是用来表示超媒体信息的工具,它只说明结点的媒体类型,而不对其具体表示方法做规定。另一个基于 SGML 的、用来表示集成超媒体信息的标准是 ISO 10744 标准,即 HyTime 标准(Hypermedia Time-based Structure Language)。“多媒体、超媒体专家小组”(MHEG, Multimedia and Hypermedia Expert Group)制定的多媒体与超媒体信息的编码标准即 ISO MHEG 标准是独立于平台(Platform)的编码标准。MHEG 标准最初是为分布

式物理媒介如 CD-ROM 上的应用制定的,但是它运用到分布式的网络多媒体应用中也是很合适的。其他标准还有 Internet 的多媒体邮件 MIME,以及 CCITT 的“交互式视听服务”(AVIS Audio-Visual Interactive Services)建议 T. 17x。

1.2 超大规模集成电路技术对多媒体的影响

多媒体技术,尤其是视频压缩技术的发展受益于以下三方面的发展:标准(Standards),网络技术,超大规模集成电路技术。现有的影像编码标准采用一系列技术分离去除信号的冗余度,这些技术在硬件实现时对超大规模集成电路技术有许多要求。制造商可以选用多种方法来设计集成的编码器,以下将讲述各种方法的优缺点。本节还将介绍 AT&T 的 AMP4000 芯片系列,突出超大规模集成电路技术对多媒体技术的促进作用。

1.2.1 视频压缩技术

视频压缩方面的标准有 ISO MPEG 标准(Moving Pictures Expert Group)及 CCITT 的 H. 261 建议。MPEG 与 H. 261 都在编码过程中采取了一系列的分离、去除图像存列的冗余度的技术,这些技术在硬件实现时有的相对容易些,有的技术则对超大规模集成电路技术提出很高的要求。下面先介绍一下 MPEG 与 H. 261 的编码原理,框图见 1-1。

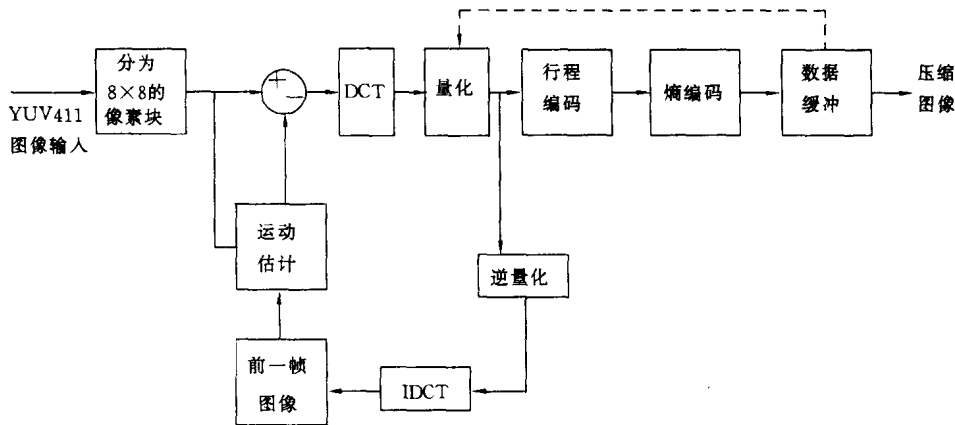


图 1-1 MPEG、H. 261 编码框图

由于人眼对高频的彩色信息不太敏感,图像序列首先做色度分量的亚抽样(Subsample),色度分量 Gr、Cb 在空间两个方向上都做 2:1 的亚抽样,这种亚抽样也称为 YUV411 编码。图像被分为许多个 8×8 的像素块,每个像素块与其预测值相减后,对差值做二维离散余弦变换(DCT, Discrete Cosine Transform),这可看作是一种变换编码。DCT 系数可进行熵编码,由于有许多 DCT 系数为零,这种编码可以显著地压缩数据量。为了去除时间轴上的即图像帧间的冗余性,需要进行运动补偿(Motion Compensation)即对一个 16×16 的像素块在相邻帧内寻找与其匹配的块,然后对两者差值进行编码。MPEG 还使用帧内插(Frame Interpolation)技术,即根据前后相邻帧预测当前帧。图 1-1 中的熵编码可采用霍夫曼(Huffman)编码。

1.2.2 视频压缩的实现方法

用集成电路实现视频编码,有多种方法可供选择。对于尚未形成标准,需作多次改动的编码方案最好采用可编程芯片或数字信号处理器(DSP, Digital Signal Processor);也可采用单元电路设计,即DCT、运动补偿、熵编码等分别由一个芯片完成,设计者的任务是把这些芯片连接成一个整体。对于已形成标准的编码方案可用专用芯片来实现,如C-Cube公司的CL451芯片可完成MPEG与H261的编解码。

为了说明现有超大规模集成电路技术对多媒体的支持,我们介绍AT&T的AVP4000芯片系列,此系列包括三个芯片及一片DSP处理器:视频编码器、视频解码器、系统控制器与音频编解码器。AVP4000系列支持以下功能:

- (1) MPEG、H. 261 视频编解码;
- (2) JPEG (Joint Photograph Expert Group, 联合图像专家小组) 图像编解码;
- (3) CCITT G. 711 (64Kb/s)、G. 722 (64kb/s)、G. 728 (16kb/s) 音频接口, 及 MPEG 高保真音频接口。

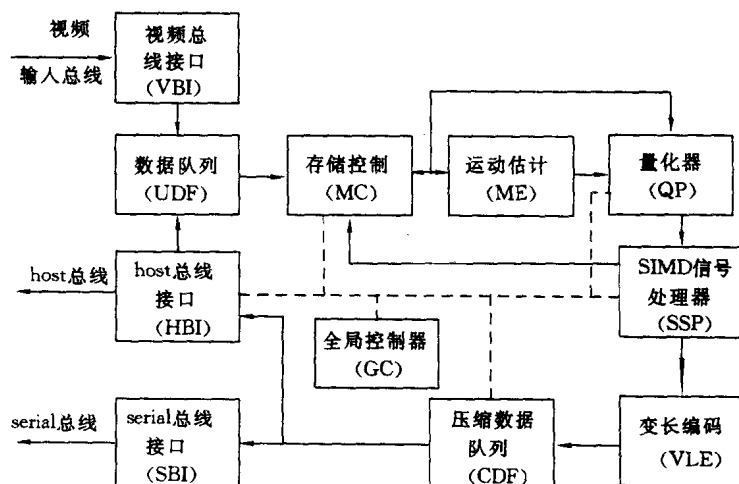


图 1-2 AVP4000 视频编码器框图

图 1-2 为 AVP4000 视频编码器框图组成。此编码器每秒完成 70 亿次运动估计与 5 亿次算术操作。整个芯片面积为 132mm^2 , 采用 0.9 微米 CMOS 工艺, 工作在 31MHz 时, 功耗为 3 瓦。

1.3 通信网的发展

传输(Transmission), 交换(Switching)技术的进步使通信网发生了重大的变化。将来的通信网要使用光纤(Fiber Optics)技术, 光纤的使用改善了通信网的带宽。宽带综合业务数字网(B-ISDN, Broadband ISDN)将促进分布式多媒体技术的迅速发展。

近年来, 随着办公自动化中信息传播的需求的增长以及 IEEE802 标准的制定, 局域网(Local Area Net, LAN)技术迅速发展, 而高速局域网的发展将进一步强化信息传播的能力。

通信网的最终发展方向是一个通用综合业务数字网(UISDN, Universal ISDN), UISDN

的建立涉及到一系列有关终端用户之间点对点的数字连接的标准。

ISDN 的传输速度可以从几个 kb/s 到 600 Mb/s, 据此可将 ISDN 分为两类: 窄带 ISDN, (N-ISDN, Narrow band ISDN,) 与宽带 ISDN。窄带 ISDN 可以提供 1.536/2.048 Mb/s 以下的传输速率, 可以传播各种格式的信息, 包括数字语音、图像及数据等。现在已定义的窄带 ISDN 接口有两种: 一种接口由两个 64 kb/s 的 B 信道和一个 16 kb/s 的 D 信道构成, 其总的传输速率是 144 kb/s, 另一种接口由 23/30 个 B 信道和一个 64 kb/s 的 D 信道构成, 总的传输速率为 1.536/1.984 Mb/s。

宽带 ISDN 可以为使用宽带通信如视频通信设备的用户提供 150/600 Mb/s 的传输速率。由于高速数据包 (High Speed Data Packet) 以及电视会议 (Video Conference) 和电视电话 (Video Telephony) 等都要 2 Mb/s 以上的带宽, CCITT 制定了宽带 ISDN 的 150/600 Mb/s 的接口标准, 而且采用异步传输方式 (ATM, Asynchronous Transfer Mode) 作为宽带 ISDN 的传输机制。在 ATM 方式中, 各种信息都被组合或切割为长度固定的信元 (Cell) 来传输, 每个信元由一个 5 字节的信头 (Header) 和 48 字节的数据组成, 信头中的虚拟信道标识符 (VCI, Virtual Channel Identifier) 用作错误检验。ATM 是分组交换 (Packet Switching) 技术中的概念, 与此相对的同步传输方式 (STM, Synchronous Transfer Mode) 主要用于线路交换 (Circuit switching) 中。

1.4 多媒体 CD-ROM 技术

只读光盘存储器 (CD-ROM, Compact Disc-Read Only Memory) 是存储数字信息的介质, 最初是为音乐发展起来的。1985 年, Philips 与 Sony 两家公司发布了在光盘上记录计算机数据的黄皮书, CD 便在计算机领域得到了应用。1985 年 11 月苹果计算机公司、DEC、Microsoft 等公司制定了 High Sierra 标准, 此标准被 ISO 采纳, 并于 1988 年制定了关于 CD-ROM 的逻辑格式的 ISO9660 标准。CD-ROM 的存储容量之大、数据不易被损伤或污染的优点为多媒体制作提供了很大帮助。

1.4.1 光盘的物理结构

一个光盘是由三部分组成的圆盘: 底面, 即未标记的一面是聚碳酸酯, 是一种极硬的塑料; 中间层是一个薄的反射金属层, 通常是铝; 顶层涂有保护漆, 可以在上面印刷 CD 标鉴, 通常采用绢网印花法。

这三层整体厚度为 1.2 mm, CD 的直径为 12 cm (4.72"), 中间有一个直径 15 mm 的孔, 不带包装时平均重量 18 克。

CD 的数据存放在聚碳酸酯上, 凹下的部分叫作“皮特” (Pit), 突出的部分叫作“陆地” (Land)。陆地到皮特的交界处表示信道比特“1”, “1”的个数与皮特的长度成正比; 反之, 皮特到陆地的交界处表示信道比特“0”, “0”的个数正比于陆地的长度。

1.4.2 光盘的逻辑格式

制作光盘前, 首先需要产生一个 ISO 映像, 即将要装入光盘的所有文件和应用程序集合的一个字节-字节的映像, 它们被按照出现在 CD 轨迹上的顺序放置, 包括计算机操作系

统控制数据存取所需的信息。将 ISO 映像记录在硬盘或磁带上,就可以送到工厂进行光盘制作了。

在 ISO9660 标准之下,单张的 CD-ROM 被定义为一个“卷”(Volume),如果一个应用程序或单一的文件不能被一张光盘装下,它可以扩充到任意多张光盘上,称此光盘集合为“卷集”。

CD-ROM 上的存储空间可以用分钟、秒和逻辑扇区来描述,每分钟由 60 秒组成,每秒由 75 个逻辑扇区组成,通常用记号 mm:ss:dd(分:秒:扇区)来标定光盘上的具体位置。逻辑扇区的大小为 2^{n+11} 个字节, n 被设为零,即每个扇区有 2048 个字节,每个扇区由唯一的“逻辑扇区序号”(LSN)标识,序号从零开始顺序增加。逻辑扇区可分为若干个逻辑块,逻辑块的长度为 2^{n+9} 个字节, n 可设为 0,1,或 2,即逻辑块的大小可以为 512、1024 或 2048 字节,每个逻辑块由唯一的“逻辑块序号”(LBN)标识,按从零开始增加的顺序分配,光盘上所有地址设定实际上是用逻辑块序号完成的。

所有逻辑扇区的集合被称为“卷集”的“卷空间”(“卷集”可以是一张或几张光盘),“卷空间”被分为两个区域:系统区和数据区。系统区占据了从 00:02:00 到 00:02:15 的物理空间,有固定的 16 个扇区(LSN0~LSN15)。从地址 00:02:16 处开始是数据区,所有的程序、数据、图像、语音等都按逻辑文件格式存放在数据区。

1.4.3 光存储技术的发展

音响 CD 是音乐工业的巨大成就,音响 CD 机的巨大市场推动了它们的技术发展并且迅速降低了价格。数字音响 CD-DA 提供采样率为 44.1KHz 的高保真音响。音响 CD 有一种原始的逻辑结构,它最多可以分为 99 条轨迹,通常每条轨迹含有一首单曲。

混合模式光盘既可以包含计算机数据,又可以包含音响信息,通常这两种信息记录在分离的轨迹上,比如计算机数据占据轨迹 1,而音响数据占用其后的轨迹,因为声音轨迹是按常规 CD 音响形式记录的,它们可以在任何 CD 放音机或 CD-ROM 驱动器中播放。

Philips 与 Sony 公司认识到了 CD-ROM 在交互式多媒体应用中的巨大潜力,于 1986 年发布了一个多媒体光盘格式及其相关硬件的标准,称为 Compact Disc-Interactive,即 CD-I。CD-I 机为消费产品,是一种可以与电视机相连的,便宜,可独立应用的电脑/音响光盘阅读机,CD-I 专门为实时运行的多媒体应用程序而设计,这些应用将声音、图片、动画和影像组合在一起,构成一个交互式多媒体系统。

CD-ROM XA 是 CD-ROM 的扩展,是 Philips、Sony 和 Microsoft 公司对 ISO9660 标准的扩展。CD-ROM XA 在一般情况下与 ISO9660 逻辑格式一致,但在其设计中加入了更优越的音频和视频性能,更易于为多媒体应用服务。CD-ROM XA 与 CD-I 提供采样率为 37.8KHz、18.9KHz 的压缩声音数据,采用自适应差分脉冲编码调制(ADPCM)对声音编码。

可重写光盘(CD-MO, CD-Magnetic-Optical)被用作高密度可重写存储介质,其存储容量可达 650M 字节。

一次性写入光盘(CD-WO, CD-Write Once)是 CD-ROM 向一次性写入介质的演变。CD-WO 的底层也是聚碳酸酯,其上绿色的可记录材料,然后是一个全反射层,因此这种光盘也被称为“金盘”(Golden Disc)。

Philips 与 Kodak 公司联合制定了照相 CD 标准,与现有的 CD-ROM XA 及 CD-I 机兼容,使得数字技术被引入了摄影领域。利用普通的 35mm 相机拍完一卷(100 张)胶片,经 Kodak 照相 CD 系统可输出照相 CD,内含 100 幅照片的数字图像,插入照相 CD 机或其他 CD 机中,就可以在电视或计算机屏幕上欣赏相片了。一个照像 CD 价格约为 20 美元,全世界约有 2.5 亿部 35mm 照像机,因些照相 CD 的市场潜力很大。照相 CD 也可用于将高质量模拟相片转为高分辨率数字图像,这种图像可以通过 CD-ROM XA、CD-I 或照相 CD 机直接放入计算机,而不用昂贵的数字化设备。

1.5 数字电视和多媒体会议

人们普遍认为所有的电视设备都将数字化,包括高清晰度电视(HDTV, High Definition TV)。交互式多媒体技术要求数字电视与其他视频设备间具有最大的互操作性(Interoperability),包括分辨率、传输方式、存储机制及传输速率等方面。因此,将来的数字电视会具有“智能性”与“交互性”。

多媒体会议系统(Multimedia Conference System)是主要基于视频技术,并综合了音频接口、图像接口、数据库及 ISDN 信道的集成信息系统。研究表明,办公室中的管理人员 40% 的工作时间花在开会上,而且这些人员要求有快速、准确的信息获取的途径,因些多媒体会议系统将得到广泛的应用。

数字电视与多媒体会议的实现中,需要解决的基本问题是数字影像巨大的数据量对传输信道和存储设备的要求。现在已有视频编码的标准,这些标准的制定降低了硬件实现的价格,并增强了系统的互操作性。

1.5.1 视频编码标准

CCITT 建议 H. 261 是一个电视会议标准,也称为“Px64kbit/s 电视会议标准”,P 可以取 1 到 30,以得到变速率的视频压缩信号。

ISO 的运动图像专家小组制定的 MPEG 标准可分为两个阶段,分别称为 MPEG-1 与 MPEG-2。MPEG-1 实现对数字存储介质(如 CD-ROM)上的视频信息的压缩编码,其比特率约为 1.5Mb/s。

H. 261 与 MPEG 都采用 DCT 变换编码与变长编码(VLC, Variable Length Coding)技术,以及运动估计与运动补偿技术。MPEG 标准还引入了帧间内插技术。

运动图像专家小组从 1993 年 9 月起已研究低速率视频/音频编码方法并计划于 1997 年制定相应标准,即 MPEG-4。这个标准将采用计算机视觉、图像合成及视觉模型等技术。MPEG-4 将用于无线的窄带通信如蜂窝电话及其他个人通信设备,而且 MPEG-4 标准的形成也将带来一些新的应用,如交互式多媒体移动通信及在普通电话线上传输的电视电话。

1.5.2 数字电视与多媒体会议技术的发展

美国、日本及欧洲都在努力研究电视的数字化问题,数字电视系统的设计要考虑到电视接收终端的价格及各种存储设备及传输机制间的互操作性。最早的数字电视期望采用直播卫星(DBS, Direct Broadcast Satellite)来传输,并采用 MPEG-2 编码标准;其次是用电缆传输