

# 多媒体模型设计

## Multimedial Model Design

吴玉军

Wu Yujun

(广西计算中心广西软件新技术实验室 南宁市星湖路 32 号 530022)

(Guangxi New Software Technology Lab., Guangxi Computing Center,  
32 Xinghu Road, Nanning, Guangxi, 530022)

**摘要** 介绍了多媒体的一些基本知识和相关探索。其中的例子是麻省理工学院的雅典娜项目。

**关键词** 多媒体 超文本 链 锚 模型 创作系统 应用系统

**Abstract** Expresses some of the basic knowledge and research work in the field of multimedia. The example is Project Athena of MIT.

**Key words** multimedia, hypertext, link, anchor, model, authoring system, application system

### 1 概述

#### 1.1 多媒体简介

所谓媒体 (Media), 指的是图 (Image)、文 (Text)、声 (Audio)、像 (Video) 等信息载体。多媒体就是将这些信息媒体溶合起来。通过计算机技术进行获取、管理、通讯和处理, 最终让人们可以用更加自然、高效的方法来生活、工作和研究。有人把多媒体技术的出现称为“人类处理信息手段继纸张印刷、电报电话、广播电视、计算机之后的第四次飞跃”<sup>[1]</sup>。

#### 1.2 多媒体模型

多媒体系统与以往计算机应用系统的一个显著差别, 就是信息量大, 信息之间的关系复杂。特别是图像和视频。数据量非常庞大, 传统的存储和处理手段已经不能适应这种新情况, 一些压缩和解压缩的标准也就应运而生, 如 JPEG、MPEG-1、MPEG-2 等。

由于多媒体系统需要和各种媒体打交道, 因此, 驱动和控制这些硬设备, 并提供软件接口, 也就是必须的了。这一部分可称为多媒体 I/O 控制及接口层。

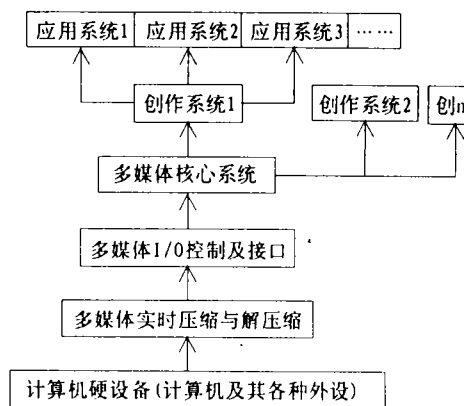
为了全面、系统、高效地管理这些复杂的硬设备, 特别是为了能够更方便、有效地应用这些丰富多彩的软硬件资源, 还需要相当于多媒体操作系统的多媒体核心系统层。如 Intel 和 IBM 公司的 AVSS (视频/音频支撑系统) 和 AVK (音频核心部分), Microsoft 的 Windows NT 和 Apple 公司的 QuickTime, IBM 的

OS/2 也有相应的多媒体扩充版。

有了底层工具后还需要一些真正的开发工具, 它们是为了更加方便地开发应用系统而设立的。这个部分可称为创作系统。它的典型技术就是超文本 (Hypertext) / 超媒体 (Hypermedia)。超文本/超媒体是一种信息管理技术, 它的表现形式是, 在节点 (Node) 和节点之间, 通过各种各样的类型链而连接起来的复杂的、非线性的综合信息网络。节点中的数据可以是图、文、声、像、程序或它们的组合, 节点之间经由链互相连接, 并可以根据链的类型和用户的意图按照一定的路径进行游动。

多媒体的最终应用就是应用系统。

归纳起来, 可以建立如下的多媒体模型:



### 2 超文本/超媒体设计

多媒体模型涵盖的内容很广泛, 但其中最关键、

最引人注目的部分,无疑是尚处于婴儿状态的超文本/超媒体技术。

事实上,至今并没有正式的超媒体模型,大多数超媒体的开发者们都仅仅在外观和界面上做文章,即软件设计者们常常挂在嘴边的“looks and feels good”。

没有任何技术是“天外来客”,一般地,一种新的技术往往都是在旧的技术基础上补充、发展起来的。超媒体技术也不例外。因此,我们可以将超媒体设计的论点分为两个部分。一个是超媒体所特有的,而另一个则是一般软件设计所共有的。例如,对于任何软件工程来说,软件的版本控制都是必不可少的。

一个超媒体系统,往往包含复杂的媒体结构,而且存有大量的数据信息。如何能够既保存信息的变化而又不必生成多个副本,并且可以允许同一个初始文件的两个不同修改版进行合并,这在一个大型的超媒体系统中是一个严重而且关键的问题。对于这种情况,我们可以开发一个版本控制系统,通过定义对节点进行增加或者删除的操作符,并且还可以经由一些继承链和定义新的节点,从而实现对节点的版本控制。

大型的超媒体项目必须要有一支包含各种专业的设计队伍。事实上,这种需求对于一般的软件工程也是需要的,但是在超媒体设计中显得更为突出。如果采用自顶向下的设计方法,那么超媒体设计将由以下方面组成:

- 图形设计及界面设计
- 快速节点原型
- 节点—链结构
- 视觉隐喻
- 辅助功能

超媒体技术是超文本技术的延伸。在超文本设计中,一个必须解决的问题就是标记(Mark)问题。因为所谓超文本,就是允许思维的任意跳跃,以非线性的方式来组合信息间的相互关系。但是,这种跳跃不是一种盲动,在跳跃过程中会形成某种序列。这些序列对于改善超文本技术本身的某些缺陷(如迷路)是大有裨益的。我们可以将这些标记称为“超文本书签”,在实现上可以通过建立一些链来完成。

成熟的技术往往都有它们自己的标准,如机械制造业、印刷业等等,都制订了很多的标准,这是人类追求合理、合作和效率的一个必然产物。比方说,我们日常接触到的书籍、报刊、杂志,它们往往都有自己统一的外观(纸张、页格式、字体安排、目录位置、图表形式等),这样做可以使读者能更快地接近

“主题”,而不必花费时间在各种各样的“搜寻”上,这或许可以用生物学上的“条件反射”原理来证明它的好处。对于超文本系统来讲,它的界面及操作必须要前后统一,尽量使用基本的软硬件资源,这样可以简化超文本应用在不同软硬件系统上的安装。这一点对于超文本设计来说也是至关重要的。

超媒体可以利用多姿多彩的图、文、声、像等媒体手段,设计出精致醒目的屏幕界面来。超媒体系统由于有结构复杂、信息量大、媒体之间的关系繁杂等特点,因此,屏幕设计要特别重视,不太重要的信息,最好不要放在显著的位置,以免用户在浏览时会“走”到另外的偶然链上,造成超媒体系统中最令人头疼的问题——迷路。超媒体最吸引人之处就是它的“链”,链可以充分体现思维的非线性特征。对于超媒体设计来说,如何选择合适的联系链的相关阈值,是对开发者的一个严峻的挑战。从一个极端情况来看,只有当节点间是直接相属或相等时,它们才可以有机地链接起来;而从另外一个极端来看,节点之间几乎可以随机地链接在一起。因此,应该设置一些选择项,让用户可以去控制那些可见联系的数量和强度,尽量减少盲目性,提高浏览的质量。

超媒体系统一般是面向大型项目的,在这个浩瀚的信息海洋中,需要有“望远镜”、“指南针”、“航海日记”和“有经验的水手”。只有明确了自己当前的位置,才能向系统提下面的一些问题:[2]

- 我需要看一些新玩意儿
- 我想换一条更有趣的路线,再去重温以前看到的美景
- 我想看看自己是怎么闯到这里来的,我走过的是怎样的一条曲折之路
- 我想回到上星期二的那个地方去,那里的风光真是美不胜收
- 我终于到达了目的地,我想使用我一贯选择的链类型来漫游这个地方

超文本技术之所以区别于传统的书本,就是因为它可以回答不知疲倦的用户所提出的上述问题,始终保持用户思维的连续性,让用户在浏览和获取所需信息的过程中得到方便、快捷和准确的服务。这是超文本技术的优势所在,所以在超文本设计中,必须保证这些问题都会得到满意的回答。

作为一个信息集散地,超文本的信息过滤功能必须要快速和有效。超文本中的过滤,有的相当于传统数据库中的 Views,有的则相当于修剪搜索树(这是一个从人工智能术语中借用过来的隐喻)。索引也是信息过滤的重要方式。众所周知,检索语言在数据库

应用中取得了极大的成功,因此在超媒体设计中也有必要提出这种需求。另一方面,基于超媒体的信息检索系统应该根据用户的需要来过滤信息,以避免无谓的操作和提高系统的响应速度,这些都是由超媒体系统的特性决定的,在设计中必须得到重视。

一个文档的结构往往比它的内容要有用,特别是对那些叙述复杂,内容包罗万象的文档来说。这也是大纲和目录得以大展身手的原因所在。超文本的一个重大贡献,就是它将信息表达和信息结构合二为一。

一般来讲,我们所说的基本超文本,是指超文本的基本结构和用来构造超文本应用的创作工具。而应用超文本指的是包含在超文本中的信息集合。广西软件新技术实验室的超媒体产品 GBH (基于图形的超媒体),就分有创作系统和浏览系统两部分。其中,创作系统提供了一系列创作工具和直观的操作界面,引导创作者根据他(她)的需要生成他(她)自己的应用系统,应用系统可以聚集大量的图、文、声、像等媒体信息,通过不同的类型链,这些信息可以根据需要相互链接,并形成一定的信息单元(节点或标志)……最终凝聚成一个信息的有机体。这个有机体一般是用作浏览,即浏览系统。

超媒体巨大的信息量,使得在进行超媒体设计中,不得不考虑海量存储设备问题。

Apple 公司 1988 年推出了苹果学习盘(CD-ROM),不久,包含超文本特性的格罗列学术全美百科全书的 CD-ROM 也出来了。CD-ROM (只读光盘)是一种大容量的只读存储介质,一般为 600M。超媒体一旦保存在 CD-ROM 上,它将不会被删除或修改。这真是安装超媒体浏览系统的好方法!它解决了覆盖超媒体的这个难题。然而,即使超媒体的基本结构在只读介质上无法改变,我们还是可以添加节点及注解的。

静态超媒体和动态超媒体有着本质的差别。GBH 的创作系统,则相当于动态超媒体,它们是共同写作和在线出版等创作活动的工具,而静态超媒体(如 GBH 的浏览系统),是一种固定媒体,它们传递的信息,无需读者向作者进行反馈。

### 3 多媒体应用系统

下面我们将主要介绍麻省雅典娜项目中的多媒体应用系统,及其作为设计文档载体的试用情况。<sup>[3]</sup>

大型、协作的项目(如汽车设计或计算机芯片设计等)会不可避免地遇到大量的有关通讯交流方面的问题,因为这种项目需要数以百计的管理人员、科研开发人员,在不同的地点,不同的时间下进行分工协

作。要克服这种瓶颈,方法之一就是编写有关设计进程的详细的技术文档。

研究表明,交流的有效性,55%依赖于面对面的视觉效果,38%依赖于声音,仅仅有 7%依赖于文字。所以,传统的文字档案已经不太适应当前竞争激烈、高科技飞速发展的状况。为了保存开发过程中的关键设计信息,需要借助有效的多媒体技术。

雅典娜项目是数据设备公司和 IBM 公司资助的一个研究开发期达八年之久的项目。它主要应用在麻省理工大学的计算机课程上。我们今天介绍的是雅典娜项目中的一部分——雅典娜视觉工作站项目。这个项目的目的是往文字文档中增添设计过程中收集的音频、视频材料。

#### 3.1 计算机硬设备及各种辅助装置

VCG (视觉计算组)是这个项目的一部分。它改造一台标准的雅典娜工作站,让它支持 256 色的图像和全动数字视频。

具体配置如下:

- IBM RT-PC 或者 DEC Micro Vax I 工作站
- 视差图像板
- Panasonic 的 OMDR (只写一次光盘)
- VHS 摄像机
- 视频转换设备

其中,IBM RT-PC (或 DEC MicroVax I)是作为基本系统;视差图像板是作为显示子系统,它将每秒 30 帧的标准 NTSC 视频信号数字化;Panasonic 的 OMDR 是做为设计文档的载体,它可以存储大约 13 分钟的全动视频或 24,000 帧的静止图像和双声道音频;VHS 摄像机用来采集材料,如录制会议、解说、交谈、介绍、访问,还有个人的设计日记等;视频转换设备(录像机、校园有线电视网等),用来将视频材料转移到光盘上,并使之能够在工作站上应用。

#### 3.2 计算机软件系统

开发一个多媒体设计文档系统,首先要解决多媒体信息的认知、存储、传输、显示和动态管理等技术问题。其中最关键的就是所谓的动态管理技术,它意味着多媒体设计文档可以按照时间顺序、标题、形式、功能等方式来进行指导和相互关联。VCG 早期开发的多媒体工作站,包含了管理文本、图象、视频和音频等基本功能,但是,如果仅仅提供图、文、声、像等媒体信息而没有相应的高层次的信息处理工具。那么,这个系统是很不好用的。在工作站上,如果信息管理只是通过 Unix 层、C 层、X 层或 X Toolkit 接口层的话,那么这个信息系统充其量只能算作是一个演示版,作用有限,达不到实际工程的应用标准。为此

目的, VCG 于 1988 年开发了一个原型创作系统——雅典娜缪斯创作系统。

缪斯提供了对文本、图像、视频和声频的信息处理支持。它允许生成包括所有这些不同信息的结构文档。通过借助从超媒体或其他“低度结构”系统而来的动态配置,人们甚至可以将这些文档建成复杂的信息组织,也可以使用从传统数据库而来的“高度结构”信息。

由于雅典娜项目是用来为麻省教员开发复杂的教育软件服务的,所以,从教育的角度出发,雅典娜缪斯提供了对三种主要软件类型的支持:

- 交互介绍
- 模拟
- 参考功能

介绍,组成了讲解教学材料的基础;模拟则可以让用户控制基于图形或视频图像的动态系统模型;参考功能则可以生成让学生存取的大型信息体,其中的内嵌工具允许用户从信息体中提取材料,并集成为他们自己的笔记和说明。在缪斯结构中,所有这些功能都集成在一个相关环境里,这样,用户或作者就可以自由地将不同的教学法组合进单个的应用系统中。

雅典娜缪斯创作系统,不仅提供生成教材的丰富基础,而且也可以用来生成设计文档。

视频,在各种媒体信息中占据了最重要的角度。因为它最直观,最符合人类的认知习惯。在计算机科学中,教学视频技术是最激动人心,最具挑战性的了。

视频,作为一种文档形式,它不仅在时间上延伸,而且还包含着内在的“图像数据”。为了有效地使用它,首先要求能够容易地移动这样一个文档内的图象数据。缪斯系统的一个最基本特征,就是使用了能够组织不同信息的“维”。对于视频,很直观地,我们可以生成一个时间维。

图 1 显示了视频文档在工作站屏幕上的典型布置情况。它使用了一组类似录像机控制按键的按钮(button),有“反放”、“停止”和“前放”三种,并且可以用不同的速度(这多少受限于光盘机的性能参数)来操作。

视频窗的正下方是可以表示时间维的滚动条。滑块的位置指示了视频文档的当前位置,滚动条的长度代表了视频文档的长度。滑块可以控制视频,用户可以用鼠标将滑块拖前或拖后,这将让视频很快地跳到适当的位置。通过这种控制,很容易在视频文档中搜索到感兴趣的位置,或者快速重放片段,或者让图像静止。



图 1 视频文档界面

Fig. 1 Interface for a video document

搜索问题,是在计算机环境中使用视频信息的一个难题。因为它不能沿用在文本或数据库中的特定模式搜索。你很难告诉计算机如何去找视频文档中“王先生站起来”的那个位置或片段。

作为这个难题的部分解决方法,可以依靠时间轴这个概念。因为时间轴事实上是独立的,在系统设计时可以很容易地将文本或图形材料分配到相同的时间轴上,使文本、图形、视频的组合同一个时间轴来定位。前后移动时间轴不仅引起视频内容的变化,而且还将导致文本、图形的变化。

雅典娜缪斯的维还提供了对动态及静态视频添加注解的机制。文本注解可以用来克服视频独自搜索时的一些问题。类似于人们在电视译制片中采取的方法那样,用户可以在视频场景的特定位置加注解,或者加一个完全的文本副本,当视频中的人说话时,它也同步显示在屏幕上。

这项功能原来是为麻省本科课程中的外语材料而设计的。此时,学生可以随时中断视频文档,阅读显示在一个独立窗口里的副标题,这个独立窗在视频窗之下。使用鼠标,学生可以高亮显示副标题流中的任何单词或短语,并将在词典中查到的相应解释显示在屏幕的独立窗口上。

相反的,学生也可以从其他文本中选出一个单词或短语,然后在副标题中搜索,找出视频中的相应位置。有了这种功能,就可以找出“王先生站起来”的位置,前提是,这些位置需预先做好注解。

在缪斯创作系统,用户可以为维添加任意数量的注解流。不同的用户可以添加自己的注解而互不干扰,这通常是由生成带有用户名的注解链来实现的。视频此时也可以根据不同的注解集来搜索。一个这样的例子就是用一种以上的语言来给视频添加注解。

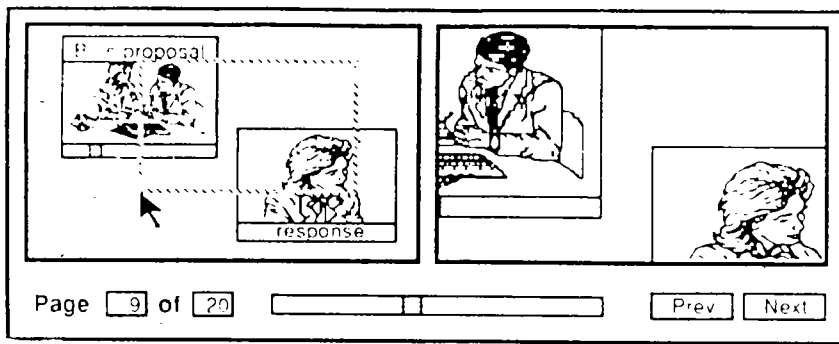


图2 电子影集

Fig. 2 An electronic photo album

使用时间结构体制,我们可以象字处理软件那样“CUT”和“PASTE”。方法是,打开几个视频文档,从时间轴上提取片段,并移动到某个新位置上。

如何管理视频片段也是视频编辑的难题之一。这里,缪斯创作系统采用了“影集”的手法。影集提供了这么一个页空间,在页空间上可以组织一定数量的图象。

图2显示这样一页的外观。右边部分显示放大的视图,用户可以控制看每页的任一部分,选择单张图像或一组图像的某一部分后,它将会显示在放大屏上。

电子影集与影集的最大差别是,它不仅限于静止的图像,而且还可以是动态的系列图像(如视频)。电子影集的每页可以包含文章和图象,而且包括可以控制放大倍数的滚动条。借助这个工具,用户可以剪辑视频文档里的片段,然后以缩图的形式存到自己的个人电子影集中,再经过组织和标注后,就可以把它们挑出来并放进其他的文档或应用中去了。

这套工具勾勒出了在缪期创作系统中管理视频文档的基本方法。作为一个创作系统,它也包括将把这些文档组合成主要信息网的基本机制。这个主要信息网事实上就是一个设计进程的历史记录。

在多媒体系统中,信息数据采集、组织和存取都是相当重要的。

其中,视频数据的生成必须要依据下面三点:

第一点,生成设计文档材料的过程必须最简单。对于生成设计文档材料来说,录制视频当然要比用手写来得快。因此,应该最大限度地利用VHS摄像机里的时间日期等附加信息,让视频里的信息量加大。

第二点,选择适当的数据形式。比如说,表达设想时,我们可以通过讲、手势和说明,而不必在如何构造文字上花心思。视频可以传递信息的方方面面的内容。例如,一个工程师可以在现场借助于发动机的

声音来指出它的设计问题,如果换成文字的话,这将是非常困难的。当然,适合视频的材料不见得适合文字文档,而适合文字文档的材料也不见得适合视频文档。应该发挥它们各自的长处。

第三点,有了好的文档材料,还需要经过归纳、精炼,使之更加细致、有效,表达更加紧凑。视频材料有一个局限,它必须被实时观看。而对于阅读能力很强的人,他(她)可以快速地浏览一页文字,这一点对于视频是行不通的。因此必须让视频内容紧凑。缪斯开发人员试验了三种表达方式:个人说明,两人交谈,设计研讨会。最终发现,两人交谈是较理想的表达方式。

#### 4 总结

多媒体系统设计是一个复杂的系统工程,它没有固定的模式和标准,但是,作为一门继承和发展中的新兴技术,它可以借鉴比较成熟的其它计算机技术,通过计算机硬件技术、通讯技术和人工智能等技术的不断进步而发展、突破。尽管现在的多媒体技术距离人们的要求还较远,但是它已经在很多领域、很多部门发挥了极其重要的作用。凭着多媒体天然的应用优势,多媒体技术已经迎来了它的网络时代,我们期待着多媒体全面应用的真正来临。

#### 参考文献

- 1 王利芳, 须德. 多媒体技术发展的现状及其趋势. 微型机与应用, 1995, (1): 2~4.
- 2 Parsaye K, Chignell M, Khoshafian S et al. Hypermedia. Intelligent Databases. P258~261.
- 3 Matthew E. Hodges, Ben H. Davis, Russell M. Sasnett. Investigations in Multimedia Design Documentation. The Society of Text Complexity. P78~89.