

# 视觉信息管理的数据模型研究

吴炜煜

(清华大学土木工程系, 北京 100084)

吴润泽

(清华大学电机工程系, 北京 100084)

**摘要** 从多媒体信息系统建造目标出发,综述了时基媒体、超媒体文献、视频信息、多媒体复合对象的基本数据模型的研究成果,分析了在实现基于内容检索和特征识别中所采用的视觉数据逻辑表达方法,并就当前图形图象数据库的研究重点和方向发表了见解。

**关键词** 视觉信息 数据模型 特征语义 内容表征

## 0 引言

多媒体信息管理技术研究的一个重要课题,是系统中多种媒体信息数据的有机控制和协调表现。特别是图象图形与音频、文字等多种解释在系统模型中要实现协同工作,使系统能在运行期间进行基于内容的检索,或构成新的特征检测器,使用户能够自然地浏览多媒体对象。

一个好的多媒体信息系统,关键特征是看其如何理解和表达语义信息。它必须把现实世界应用实体与结构化特征语义联系起来,复杂特征应与其各组件的语义相联系。多媒体对象的准确识别依赖其特征集合,即其基于内容的关键值集合。因此,多媒体信息的逻辑表示方法,特别是视觉信息的数据模型,必须充分考虑上述要求<sup>[1-4]</sup>。从这个观点出发,本文将围绕视觉信息表达问题,论述几种基本数据模型。

## 1 时基媒体模型(Time-Based Media Model)

时基媒体模型是在多媒体领域中常用的一种数据模型。多媒体信息是一种包含数字视频、数字音频、文本信息等多种媒体信息集合的综合媒体信息,各种媒体信息之间存在一定的制约关系,最为常见的是时间对应关系。各媒体内部有一定的时间顺序,形成一个独立的时间序列;不同媒体的时间序列之间有一定的对应关系,存在同步问题。在信息的显示

时,必须保证不同媒体的时间序列之间的同步,才能准确地表达信息的内容。

时基媒体模型可以方便地表示这种时间上对应的媒体信息,非常适合多媒体信息的表示。在图形图象信息中,经常用到运动图象序列,即动画。动画中的图象信息与时间有着一一对应关系,是一个时间流。对于一般的图形图象信息,由于存储和查询等操作的需要,常常需要将单独的一幅图象按照其图形特征进行分解,提取出各图形特征,形成层次结构。而且常常需要附加一定的文本、属性等信息,其表现也将会形成一个时间流。因此,用时基媒体模型表示图形图象信息也是合理的。

时基媒体模型可概括为媒体单元(media element)的时间流(timed streams),即媒体单元的时间序列。与其他媒体模型相比较,时基媒体模型具有三大基本结构机制:

### (1) 解释(interpretation)

一个时基媒体数据模型应该有构造时间流的功能,这些构造应该可以从信息中抽象出来,它关心的是媒体单元的物理组织,例如,它们的物理次序和排放位置。同时还要允许多相性、交差存取和填塞调整。时基媒体模型需要存储单元解释表述时间流,复杂的解释包括以下几方面:① 多相性(heterogeneity),含流在内的媒体单元在各方面的差别;② 交叉存取(interleaving),为了在播放过程中简化流的同步,将各媒体单元按一个简单的存储单元交叉存取;③ 填塞调整(padding),为了使存储传输速率与媒体数据速率相匹配,将一些无用的数据填塞到存储单元中;④ 单元混乱(out-of-order elements),一些

压缩技术(如 MPEG),在相邻单元之间发掘其相似之处,识别“关键”单元并通过插值构造出中间单元,因此在解码初期就需要关键单元。为了适应这一要求,在存储时关键单元要优先于中间单元存储;⑤ 可量测性(scale ability),对时基媒体的某些表示,尤其是对数字视频,允许对不同层次内容进行详细表示。

### (2) 派生(derivation)

派生即是从媒体对象集和控制派生的参数集到新媒体对象的映射。其最基本的思想是:表示媒体对象的基本媒体单元,是在需要时计算出来。利用派生,可以从其它项的值中计算出指定项的值(一个属性、一个属性集、一个实体类型等)。通常,派生项的结构是与派生所用项的结构相同的。利用派生这一概念,有以下几个优点:① 不需复制就可构成另一视图,从而减少了空间占用;② 使修改操作进行得更有效。例如,删除一个视频的子序列,可以用复制并重新装配帧数据的方法,但是简单地创建一个派生来表示这种编辑则会更有效;③ 派生通过从存储数据中分离派生对象,给出了独立的物理数据。

### (3) 组合(composition)

复杂的图形图象信息经常是由简单的组件复合构造而成的,组合的形式是其数据模型的基础,其组织原则应该是显式模型。组合是对一组多媒体对象之间的时空关系的说明,组合的结果称作多媒体对象,原始对象称为组件。组合有多种形式,通常可分为两大类:空间组合和时间组合。空间组合处理的是二维和三维空间中对象的位置,如在一场景中,在一页文字内布置一幅图或放置图形对象。时间的组合是指时间媒体合成时的相对时间关系。例如,一视频序列和一属性解析序列组合就属时间组合。

解释、派生和组合给出了一种从简单、无解释的数据得到复杂多媒体集合的途径,它是一种较高层次的概念抽象。数据库的操作可以看作是产生、使用、转换数据流等活动的扩展。在具有支持时基媒体能力的数据库系统的实施中,出现了许多新的问题,如规划安排、同步、资源分配、存储管理等。对时基媒体的支持还要求在数据库体系和数据库/应用程序的界面形式上需要有相应的基本改变。时基媒体模型是近年来视觉信息数据模型研究的一个热点,取得很大进步,有些成功的设计技术和实现方法见文献[1]、[5]、[6]。

## 2 超媒体文献模型<sup>[7]</sup>(Hypermedia Document Model)

一个优秀的图形图象信息系统,除了完善的图形图象信息外,还应包括必要的辅助信息,如文本、声音解析等。多种媒体的共同表现丰富了图形图象的意义,使得图形图象信息系统更具生命力。

为了较好地组织多种媒体信息,不少同行在进行探讨,超媒体文献模型是其中比较成熟的一个。超媒体是把多种媒体形式组合起来,并且各种媒体形式之间保持一定的链接关系。在超媒体表现时,可以方便地从一种媒体转移到另一种媒体,或多种媒体同时显示。对于一个好的超媒体模型来说,要求它能够描述极其复杂的关系,以便能够在从一个平台转到另一个平台时,把表示的要素保留下来。这包括对所用媒体项的详细说明、媒体项之间的时间关系、媒体项的布置方式以及显示中交互的可能性。

一个超媒体表现,可以认为是媒体项位于叶结点上的层次结构,也允许将各媒体项按某种关系构成组,当作一个对象来处理。超媒体的结构关系定义了其各媒体项之间的逻辑连接,包括各媒体项组合在一起显示和组之间链接的说明。各媒体项的组合形式包括:与时间相关组合(time-dependent composition)和与时间无关的组合(time-independent composition)。与时间相关组合就是将两个或更多的项(或者是项组)组合在一起,并指定它们之间的时间关系。时间关系有以下几种:① 串联,即当一个媒体项(媒体组)播放结束时另一个媒体项(媒体组)开始播放;② 并联,媒体项(媒体组)同时开始播放;③ 混合,一个媒体项(媒体组)开始播放后,经过一段时间,在它没有结束之前,第二个媒体项(媒体组)开始播放。而且媒体项常常是没有考虑空间关系就放在屏幕上了,它们的空间关系绝大多数是隐含的。对于与时间相关组合,可以根据各媒体组时间上的串、并联组合关系,计算出超媒体表现的时间约束关系,并记录在数据模型的相应字段中。对于与时间无关的组合,成组的媒体项之间无时间或空间上的关系,在超媒体表现时,播放的多个媒体表现,其时间线是互相独立的,它们之间的空间关系也是独立的,但它们通常不应该占据相同的屏幕位置。可行的方法是:对每一个表现用分离窗口或将各子场景的演播限制在各自专门的子窗口中。

在超媒体文献模型中,引入了锚的概念。锚按照媒体无关的方式在表现中引用一部分媒体项,它是一段摘要,允许相关数据保持在幕后,一个锚值指明媒体项的一部分。锚的主要用途是:当它和链接共同使用时,对表现中或表现之间的链接给出一个源对象或目的对象。锚的使用是提供一个基点,依此基点连接时间关系,使得媒体项中的一部分能够与另一个媒体项进行合成,或者给出空间关系的一个基点。锚可以指向不同形式的媒体:文本的锚常常指明文本媒体项内的字符序列;图象的锚则指出象素图象中的一块区域,该区域可以用矩形区域,也可以涉及图象中的任意对象,只须在锚值的内部说明该区域的数据格式;视频项的锚可被选作一个帧序列,它允许用户在视频的任一帧选择一个链接跳到另一个表现,或更进一步,对帧序列的延伸指明其屏幕上的区域。锚不仅可定义媒体项的扩展,还可以反映与媒体项相关联的属性,因此可用作标签以标明与之相关的媒体项部分。同时,锚作为标签向媒体项附加路径的信息,也可以用作媒体项的检索。

在超媒体文献模型中,除了上述的模型结构问题外,还存在媒体的表现(包括时间和空间位置关系)、交互等问题。空间布置指出了基于屏幕的媒体大小和位置,包括它们相互之间的关系,以及它们与作为整体的表现之间的关系。时间关系指出了可能是存储在不同的地方的媒体项之间的时间依赖关系。交互性包括浏览,给出了最终用户能够在相关信息之间跳转的选择。关于这些部分的内容,可见文献[8]、[9]。

### 3 视频信息模型<sup>[10]</sup>(Video Information Model)

在建模技术中,最关键的问题是为所感兴趣的实体找到一个合适的逻辑表示。在各种视频信息建模方法中,面向对象方法隐藏了实体内部表示及实现的细节,强调了实体间的关系,因而倍受研究人员的关注。以下将要介绍的视频信息模型本质上是面向对象的,旨在统一多种单帧(图象)分析和运动分析技术。

视频信息模型的核心问题是从庞大的视频流数据中析取图象和运动特性。目前,为了满足基于内容检索的要求,视频信息建模常采用以下几种方法来获得图象和运动特性:① 手工标记析取描述信息;② 采用自动检测场景变化(也称视频剪切)的方法

得到图标表示;③ 采用图象分析技术从析取的对象特征中得到静态特性。但是,手工标记的方法对于大型应用程序来说,速度极慢;而图标表示的自动方法没有集中在析取运动对象的重要特征上;采用图象分析技术的方法则着重于单帧的处理。从本质上说,视频建模应该充分考虑视频流的时间本性,即识别出的对象是从相关运动实体得到的。

视频信息模型是基于代数框架的,可以用下列方式来开发代数说明:给定一个包括几类符号及其相关算子(operator)的字母表,也就指定了一个模式(schema)(形式上,模式相当于代数框架中的特征记号的意义)。模式具有全部必要的句法信息和分类规则,该规则确定了什么类型的对象可提供给每一个算子,以及返回什么类型的对象。在视频信息模型中,由系统开发者提供的与域相关的信息和由系统本身所提供的与应用无关的构造组合起来创建模式。本质上,模式是所有感兴趣对象(真实的或概念的)和它们之间关系的形式说明。给定一个模式,就定义了形式表达式集,这些表达式是通过用大量功能强大的提供给每一类的算子所建成的。

视频信息模型的基本组成是数据类型及操作这些数据类型的函数。在视频信息模型中主要使用两种数据类型:

(1) 系统定义的、固定的数据类型,称为“可承载类型”(deliverable types)。包括:字符串、整型、布尔型、文本、图象、音频和视频。

(2) 用户定义的数据类型,称为“实体类型”(entity type)。如“人”,或“学生”,用来表示现实世界中的对象或概念。通常,它们具有可体现在承载类型中的特性。

一种数据类型可以有与之相关的任意多个算子,与用户定义类型相应的算子称为用户定义函数。用户定义函数描述了域相关关系,实体类型间的交叉引用和对象的属性,交叉引用一般表示多值关系。函数类型具有如下一般形式:

$$\phi: \alpha_0 \times \dots \times \alpha_{n-1} \rightarrow \alpha_n$$

这里,每个 $\alpha_i$ 称为一个类型表达式,可归纳定义为:

- (1) 一个数据类型;
- (2)  $\alpha_1 \cup \alpha_2, \alpha_1 \times \alpha_2$  或  $P(\alpha_1)$ , 此处 $\alpha_1$ 和 $\alpha_2$ 是类型表达式。

注意:上述定义允许函数取对象类型、它们的联合、笛卡尔积和超集(powersets)作为自变量。

除用户定义的函数外,还需要按照操作系统中

所定义的数据类型,确定一个与应用域无关的算子集。每一数据类型,如文本、图形、扫描图象、音频和视频,都可以选择与之相关的多种算子。这些算子包括了所有集合论的、布尔型的以及算术算子。例如,AppendPar 算子对文本类型执行一个连接操作。此外,还有多种的联编算子,这些算子的主要特征是它们引起的变化涉及到域中的元素。一个例子是形如  $\{f(x)|P(x)\}$  的构造算子集,这里  $f(x)$  指明所期望的输出对象, $P(x)$  表明这些对象必须持有的检索谓词,一旦  $x$  涉及它的域,只要  $P(x)$  满足, $f(x)$  就加到这个集合中。还有许多与其相象的其它算子,包括逻辑量词等。以下给出一些常用的算子:

(1) 集合操作符号: isin, isSubsetOf, isTrueSubsetOf, union, intersection, difference, Union, Intersection, noOf

(2) 相等算子: is, isnot

(3) 时间同步(对所有媒体类型): sim, before, meets, equals, starts at, finishes

(4) 空间部分(仅应用于图形、图象和视频): left, right, bottom, up, showIn, arrange

(5) 整型操作符号: +, -, \*, /, <, >, <=, >=, min, max, ave, sum, prod

(6) 字符串操作符号: concat, strLen

(7) 逻辑操作符号: and, or, implies, not

(8) 文本: appendPar, cutPar, eqPar, keyword, isKeywordIn, parSim

(9) 图形: insPatch, pictureSum, fill, domain, colors, getPatch, getColor, restriction, scale, translate, dot, lineSeg, box, coincident, contains, disjoint, visible, bounded

(10) 音频: intensity, extract, audioIns, audioLen, audioSim

(11) 图象: shift, zoom, superimpose, overlay, in=ImageSim

(12) 视频: videoLen, pace, videoClip, videoIns

上述所列的数据类型及其相应的算子是所有多媒体信息系统模式(特征)中的一部分。语言的句法是在这个模式的基础上开发的。

在代数设置中,代数与对象集中的每一类型相关,对象集通过类型的说明起到强制的作用。因此,与类型整型(integer)相关的集合,显然是整型数的集合。其它数据类型表明对象具有预先定义好的特性:文本类型(text)的对象是段落;音频(audio)数据类型表明一维信号;图象(image)类型包含两维信

号;视频(video)具有三维信号  $F(x, y, i)$ , 表示为  $F_i(x, y)$ , 此处,  $i$  表示帧记数器,  $(x, y)$  是象素坐标。当不会发生混淆时,可以忽略  $x, y$  坐标或帧记数器  $i$  的引用。当需要时,可以用上标来区别不同的视频流,如  $F_i^1(x, y)$  和  $F_i^2(x, y)$ 。将用到下列标记:

(1)  $F_b(x, y)$ , 视频序列的首帧

(2)  $F_e(x, y)$ , 视频序列的尾帧

(3)  $F_c(x, y)$ , 视频序列的当前帧

代数允许偏函数,只要函数不是定义在域的特定部分,就可以用符号  $\theta$  来表示未定义的值。上述有关内容见文献[11][12]。

#### 4 多媒体复合对象模型<sup>[13]</sup> (Composite Multimedia Object Model)

多媒体复合对象模型由一个类层次结构以及一个基于该结构的复合多媒体对象组成。在该模型中,多媒体对象被看作是通过接口产生和消耗多媒体数据的激活对象。该模型中的多媒体对象被分为3类:

- ① 源对象(source objects),它只有输出接口并产生多媒体数值;
- ② 耗对象(sink objects),它只有输入接口并消耗多媒体数值;
- ③ 过滤器对象(filter objects),它既有输入接口又有输出接口,并且既可产生也可消耗多媒体数值。

在多媒体复合对象模型中,引入了一种图示记号系统,以方便地表示对象间数据流的关系,并为基于视觉组成和编辑复合的对象提供了基础(见图1)。在该图示记号系统中,用圆来表示多媒体对象,用方框来表示接口,且接口附着在多媒体对象上。外部的方框代表输出接口,内部的方框代表输入接口。并且用箭头线来表示数据流。

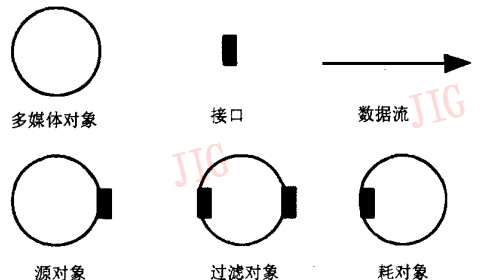


图1 多媒体对象的图示记号系统

多媒体复合对象模型中包含两个时间坐标系:世界时间和对象时间。其中世界时间是在一个应用中的所有多媒体对象所共有的时间坐标,它控制

着它们的时间行为;对象时间对应于给定的多媒体对象。对于每个多媒体对象,都可以规定:① 其对象时间原点在世界时间坐标中的位置;② 其处理多媒体数据的速度;③ 其对象时间相对于世界时间的方向。这些规定通过如下三个时间变换来实现:移动、缩放和反向。

**移动:**在世界时间中移动多媒体对象。

**缩放:**以一个给定的倍数来缩放对象总的持续时间。

**反向:**在“向前”和“向后”之间翻转对象时间的方向。

一个复合多媒体对象包含一组成员多媒体对象,以及关于它们的时间及数据流关系的规定。时间关系定义成员对象的同步性和时间序列。数据流关系定义成员对象的输入输出接口间的联系。

作为例子,让我们来考虑生成一个新的复合对象  $c_1$ ,它执行下面的操作:在时刻  $t_0$  播放一个视频对象,  $video_1$ 。在时刻  $t_1$ ,开始从  $video_1$  渐显第二个视频对象  $video_2$ 。该转变在时刻  $t_2$  时完成,在时刻  $t_3$  时  $video_2$  的播放停止。新的复合对象的成员间的时间

关系可用一个复合时间线图来说明,如图 2 所示。

成员对象间的数据流关系可用先前介绍的图示记号来定义,如图 3 所示。该图说明了对象  $c_1$  在时间间隔  $[t_1, t_2]$  内的数据流关系。在这段时间间隔内,从  $video_1$  到  $video_2$  的视频帧序列被数字视频效果处理器 Dve 处理,该处理器产生一个新的视频帧序列,并将其送给视频播放器对象。

多媒体复合对象模型可以利用二进制大对象 (BLOB: Binary Large Objects) 生成复合对象的模型并存储在多媒体数据库系统中,该模型包括解释和派生。二进制大对象从本质上是位序列,常被用于在数据库系统中存储大量的多媒体数据。解释给出从一个二进制大对象到一个有意义的多媒体对象的映射。例如,从一个动态 JPEG 流到一个 JPEG 帧序列的映射,从一个 MPEG 系统流到一个 MPEG 视频流和到一个 MPEG 音频流的映射等。派生是从一组媒体对象和控制派生的参数到一个新的媒体对象的映射。派生的一个例子是一个视频转换,它实现一个特殊效果,例如渐显和清除。如图 4 所示。

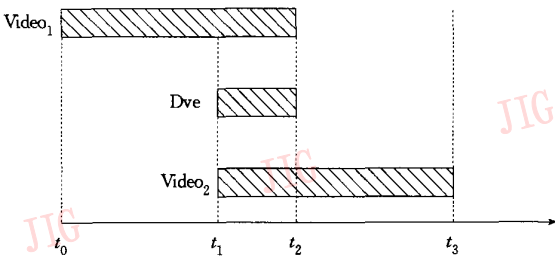


图2 一个复合时间线

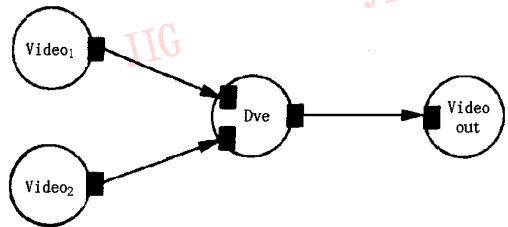


图3 时间组成举例

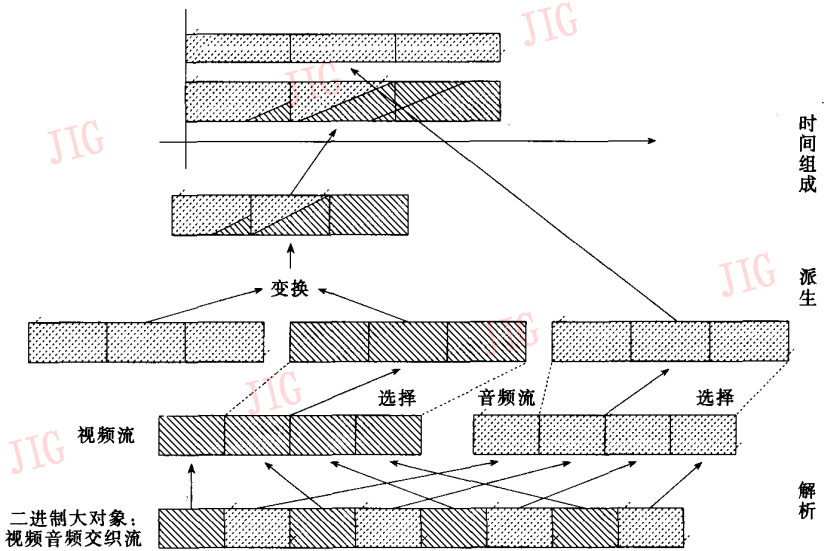


图4 从一个二进制大对象生成一个复合多媒体对象

## 5 结论与展望

多媒体信息中最主要的一种是图形图象,在图形图象信息管理系统中,要解决的问题是:从概念模型入手,针对应用领域的特点和具体要求,构造适合数据库操作,同时又能反映应用对象特征的图形图象表示结构,即解决数据模型的构造问题。管理系统中采用的数据模型,必须适合各种基本的数据库操作,能够较快地响应诸如查询、排序等数据操作。同时,该数据模型还必须适应图形图象的基本特点,图形图象信息与一般的信息不同,蕴涵着丰富的语义特征和特定的限制条件。因此,所采用的数据模型必须能够表示其各类特征,并且能够反映其中的约束关系。此外,图形图象的数据结构还必须紧凑,以便用较小的空间存储最大限度的信息量。

视觉媒体是反映客观世界的原始表现形式,数据量大,抽象程度低,信息表达结构性差,但它是人们认识客观世界的直接感受源。因此,为有效利用视觉媒体,对视觉信息内容的描述与对图形图象的分析抽象、特征提取、检索加工等密切相关。国际标准化组织开展的 MPEG-7 研究,关于多媒体内容描述,其焦点也是视觉信号数据模型。值得注意的是,在基于内容的多媒体信息检索应用需求的推动下, MPEG-7 标准的制定有力地推动和促进了视觉媒体信息管理数据模型研究。

在图形图象信息系统研究中,致力于建立多媒体目标模型的工作在不断取得进步。值得注意的是以下方面:

(1) 加快进行主动型图形图象信息系统的研究。在主动型图形图象信息系统中,输入图形图象之后,系统应可以自动地完成数据的初始化,并形成图形图象相应的知识结构,并生成主动索引等内容。

(2) 图形图象理解专家系统的研究。图形图象信息量大,完全靠人工输入各项属性,工作量大,处理速度慢。建立图形图象理解专家系统,用来进行图形图象信息的知识提取和特征识别,是图形图象信息系统推广应用的重要前提。理想的图形图象专家系统,应当可以根据已有的领域知识库,运用一套完善的推理机制,自动或半自动地进行图形图象信息知识的提取和识别,并可以从已完成的事例中获取知识,扩充自身的知识库,并提高其处理的自动化程度。

(3) 将图形图象信息技术和通讯技术、网络技术及分布式系统相结合,研制传输速度快、效率高、安

全性能好、使用方便的分布式图形图象信息系统。

(4) 从数据模型入手研究全新的图形图象信息系统。除了现有的扩充关系数据库模型、语义数据模型、面向对象数据模型、超文本和超媒体数据模型外,还应当大力研究更能表现图形图象信息内在规律的数据模型,或将多种数据模型相结合,形成语义丰富,操作简单的数据模型。

图形图象数据库技术是一个富有挑战性的新兴技术,具有多学科交汇的特点,集数据库技术和图象处理技术等多种技术于一身,是计算机科学发展中最具活力的技术之一。图象是一种人类视觉所感受到的形象化信息,其信息量丰富、完整、传送快捷,因此,以图象形式描述记录客观世界的努力,贯穿了整个人类的发展史。图象数据管理的应用涉及到我们生活的一切方面,并且随着人类活动范围的不断扩大,其应用亦随之拓展。图形图象信息系统有着广阔的应用前景。

## 参考文献

- [美]Gresky 等著,吴炜煜等译. 多媒体信息管理技术手册. 北京:科学出版社,1998.
- 吴炜煜. 工程数据管理系统. 北京:清华大学出版社,1996.
- McKeown D M, Reddy D J. A hierarchical symbolic representation for image database, In: Proc IEEE Workshop Picture Data Description and Management, 1997.
- Simon Gibbs, Christian Breiteneder, Dennis Tschritzis. Modeling Time-Based Media, the Handbook of Multimedia Information Management. Prentice Hall PTR, 1997.
- Klas W, Neuhold E J, Schrefl M. Using an object-oriented approach to model multimedia data. Computer Communications, 1990, 13(4): 204~216.
- Masunaga Y. Multimedia databases, A formal framework. In: Proc IEEE CS Office Automation Symposium, Gaithersburg, MD, Apr 1987, Washington, DC: IEEE CS Press, pp. 36~45.
- Lynda Hardman, Bulterman D C A. Document Model Issues for Hypermedia, the Handbook of Multimedia Information Management, Prentice Hall PTR, 1997.
- Hardman L, Bulterman D C A, van Rossum G. The Amsterdam hypermedia model, A ding time and context to the Dexter model. Communications of the ACM, 1994, 37(2): 50~62.
- Buchanan M C, Zellweger P T. Automatically generating consistent schedules for multimedia documents. Multimedia Systems, 1993, (1): 55~67.
- Nevenka Dimitrova, Briarcliff Manor, Forouzan Golshani. Video and Image Content Representation and Retrieval, the Handbook of Multimedia Information Management. Prentice-Hall PTR, 1997.
- Chang S K. A methodology for picture indexing and encoding. In: Picture Engineering, Fu K S, Kunii T L (eds), Los Alami-

tos, CA: Springer-Verlag, 1982, pp. 33~53.

- 12 Crosky W, Mehrotra R. Image database management. IEEE Computer, 1989, 22(12):7~8.

- 13 Hamakawa R, Atarashi A. Composite Models, the Handbook of Multimedia Information Management. Prentice Hall PTR, 1997.



吴炜煜 清华大学副教授,长期从事计算机应用科研和教学工作,出版了多媒体技术开发指南,多媒体系统软件及应用设计等十几部著作,在国内外发表学术论文 50 多篇。现任中国图象图形学会理事、中国计算机学会多媒体专业委员会常务委员、多媒体辅助工程学组主任。目前感兴趣的研究领域为多媒体仿真技术,计算机图形学,CAD/CAE/GIS。



吴润泽 1996年毕业于北京航空航天大学大学电子工程系,获学士学位,现在清华大学电机与应用电子工程系生物医学工程专业攻读硕士学位。目前主要研究领域为生物信号处理,计算机图形学,医学电子仪器技术。

## The Research of the Data Model for Vision Information Management System

Wu Weiyu

(Department of Civil Engineering, Tsinghua University, Beijing 100084)

Wu Runze

(Department of Electrical Engineering, Tsinghua University, Beijing 100084)

**Abstract** In order to explain the construction of the multimedia information system, this paper demonstrates and analyses the fundamental data model of time-based media, hypermedia document, video information and composite multimedia object. Furthermore, the logic representation method of the vision data which is used to implement content-based retrieval and feature identification is exposed. And then some opinions about current focuses of graphic's database study and future development are put forward in the paper.

**Keywords** Vision information, Data model, Feature semanteme, Content indication