

关于多媒体和智能系统的集成研究

On the Integration of Multimedia and Intelligent System

李向阳 鲁东明 潘云鹤

(浙江大学人工智能研究所 杭州310027)

摘要 This paper analyses the background of the integration of multimedia and intelligent system, presents the goal and implement method about integration. It explores the structure of the integration system. It also presents the key technique of the further study on integration.

关键词 Multimedia, Intelligent system, System integration

1 背景和研究现状

多媒体是80年代发展起来的一门技术,虽然目前对多媒体尚未有一个严格的定义,但却有这样的共识:多媒体集成了文本、图形、图象、动画、电视图象、语音及视频等多种媒体。多媒体技术的发展,改变了人机的交互方式;同时,信息媒体的多样化,大大地扩大了计算机应用的范围。

超媒体是将多媒体信息引入超文本技术而构成的信息表示和处理技术,其信息结构表示为元素间以一定的链接关系组成的语义网络。这种非线性关联的网络链将多种媒体联系在一起,它实际上是一个多媒体的信息管理和传输系统。应用导航控制技术,用户可以通过超媒体网络系统对相关联系信息采用点击热点的方式进行浏览和查询,自由形式的导航符合认知心理学的关联记忆模式。

多媒体技术和智能技术的结合导致了智能多媒体技术的产生。基于知识的推理机制和导航技术的结合改进了非线性的超媒体链,另一方面也纳入了多媒体数据库的管理方法。它比传统的智能系统和多媒体系统提供了更加广泛的应用,智能推理机、智能人机交互界面也克服了超媒体的局限性。将智能系统和多媒体集成最早是在1990年7月美国人工智能协会(AAAI)和同年12月的第一届世界专家系统大会被提出的,目前关于多媒体和智能系统的集成在国外研究和应用已经取得了不少成果,国内在这方面的研究基本上处于空白。关于多媒体和智能系统集成的研究主要涉及:集成的基础技术研究、集成的模型结构、人机交互界面的设计、系统设计和开发

的原则、应用领域开发。就多媒体和智能系统的集成应用来说有智能多媒体辅助教学系统、智能多媒体决策支持系统、智能多媒体医疗诊断系统等。

2 研究目标与实现途径

多媒体和智能系统的集成从目标与研究途径上来看,可以划分为两类:智能多媒体与多媒体智能。

所谓智能多媒体是指在多媒体技术的基础上,通过嵌入知识表达与推理机制,来提高或改善多媒体系统在媒体制作、组织存储、检索和显示交互方面的性能。目前,人们对多媒体的研究大多数将注意力都放在存储、压缩和传输多媒体信息上,而对各种多媒体信息的处理则尚未引起足够的重视。多媒体对象通常是无结构的,各部分的内在联系非常复杂,将不同的数据类型(文本、图形、图象)连接成一个语义一致的记录需要嵌套树或图等复杂的表达方法,而且记录之间是非线性排序的,因为几个记录可能同时参与到一个关系中,而且口头描述的可视化信息是自由文本的形式,它们是无结构或半结构的,因而需要新的表达方法。采用智能系统的方法,给不同媒体信息源间获得的各种数据加上语义,从而使得语音和图象自动识别和理解成为可能。采用智能导航链,引导人们只接受需要的信息,而不至于淹没在多媒体数据的汪洋大海之中。多媒体开发者需要对多媒体数据库进行有效地管理是必需的,一个智能多媒体系统可以更加方便用户索引、浏览、检索多媒体数据,智能系统提供了分析能力。

所谓多媒体智能是在智能系统中嵌入多媒体技术,多媒体智能的研究可分为浅层次上的集成和深

层次上的集成。前者主要是从改善人机交互界面上的研究,即在智能系统中采用多媒体技术使交互界面更加自然,特别是空间、时间相关的信息显示,提供多媒体的帮助与培训手册等,这方面的相关文献较多;在深层次上的集成,即在传统智能技术的基础之上,引入多媒体对象的表达与处理机制,拓广智能技术对人类思维形式模拟的范畴,试图采用计算机来模拟人的形象思维,以弥补传统智能系统中只有逻辑推理之不足。综合推理、类比推理研究、多分辨率的表达方法,为多媒体智能的深入研究提供了基本手段,但目前基于形象信息的智能技术研究远未取得实质性的进展。基于计算机的形象思维将涉及计算机科学、认知科学、计算机视觉、人工智能、心理学、脑科学等学科。

3 多媒体和智能系统集成的方式

在集成系统中根据智能系统和多媒体系统的侧重不同可分为:由多媒体支持的智能系统,即多媒体智能;由智能系统支持的多媒体系统,即智能多媒体。

根据多媒体和智能系统集成的耦合程度,可分为松耦合集成和紧耦合集成。所谓松耦合集成是在现有的多媒体系统和智能系统的基础上,利用系统提供的 API 接口函数,开发两者之间的接口程序,实现两系统之间的通信及消息转换(见图1)。优点是

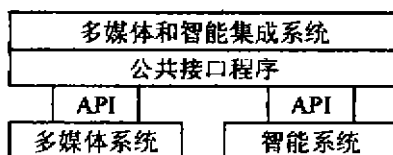


图1 松耦合系统

图1 松耦合系统

可以充分利用现有系统的功能,二次开发相对来说较简单。缺点是两系统之间有数据冗余,维护工作量较大,存在一定通信的开销,操作速度慢等。在紧耦合系统中,智能技术和多媒体技术是集成为同一系统中,充分实现了数据共享和知识表达,系统的健壮性好,资源的利用率和运行的效率较高,提高了解决问题的潜力;缺点是开发的周期长,图2为一个典型的紧耦合系统的集成的结构,多媒体信息管理模块主要对语音库、视频库、模型库、知识库、数据库等多种数据进行统一管理,这些库可以是分布的。多媒体知识表达模型主要根据知识库中的信息对各种媒体

采用多维数据的知识表达方法。建立在多维知识结构上的推理机也将突破符号逻辑推理机的局限。解释功能模块主要是通过多媒体的 I/O 接口由智能多用户界面表达和接受用户的操作。目前来看,集成系统大多处于松耦合阶段。

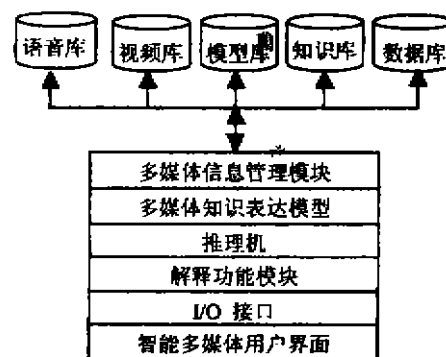


图2 紧耦合集成系统

图2 紧耦合集成系统

4 多媒体和智能系统集成的关键技术

4.1 多媒体数据的分类与管理

多媒体数据的分类与管理主要是从多媒体数据集中提取出特征信息,对相关的数据进行聚类,在此基础上构造索引以及对多媒体数据进行查询等。多媒体和智能集成系统提出了以下新的要求。

1) 多媒体数据需要大的存储空间,对多媒体信息推理将可能同时使用好几种数据,对物理相关度较高的数据进行聚类存储,可以加快访问的速度。

2) 从多媒体数据中提取明显特征属性并开发恰当的表达方法。在信息提取及分类的过程中要考虑采用专家的知识,从多媒体数据集中识别和提取出明显特征将影响多媒体数据的分类,不同的多媒体数据需要不同的分析工具。在应用框架设计中要考虑系统的可扩展性。

3) 传统的数据库检索大多是采用关系数据库的检索方法。多媒体数据的特殊性、组织和多媒体数据检索是近年来研究的一个挑战。如采用面向对象的技术来组织多媒体数据、基于内容的图象检索方法的研究、基于 R-Tree、VP-Tree、X-Tree 的多维数据表达方法等是目前研究的热点问题。

4.2 智能多媒体信息中的推理机

集成系统的推理机应当迎合不同类型的多媒体数据,一些数据是符号型的,另一些是模式,多媒体

数据中的知识嵌套不完整的模糊信息,推理过程可分为三步:

1)处理基于模式的知识(模式层)。在这一层中,推理机采用神经网络方法建立一套训练学习机(后向传播、自组织、Hopfield 网),通过对大量的数据进行模式识别的计算处理,得到一些模式知识。

2)基于规则的推理机(规则层)。这一层的规则主要用来提供推理或用来提供解释。这些特征是自适应的、基于特征的归纳类比过程。规则推理可以分为确定性推理和模糊推理。

3)第三层推理机主要处理第一层和第二层所产生的异常行为,即非符号信息。

根据应用系统的特征,这三层在推理的过程中可以是串行或并行运行。

4.3 基于多媒体信息的形象思维

人类的思维方式可分为逻辑思维和形象思维。逻辑思维方式主要是应用观点、概念、定义等按一定的规则进行推理最后得出结论。形象思维是进行综合、类比推理最终得出结论。将智能系统和多媒体技术结合,利用多种媒体再现现实(如虚拟现实技术)是一种将逻辑知识和形象知识结合的较好方式,以突破仅仅建立在传统符号逻辑推理基础上的局限性,为弥补传统基于符号推理机之不足,将在传统智能技术的基础之上,引入多媒体对象的表达与处理机制,拓广智能技术对人类思维形式模拟的范畴。综合推理、类比推理的研究及多分辨率的表达方法,均为多媒体智能的深入研究提供了基本手段。

4.4 分布式多媒体数据库和分布式人工智能的集成研究

随着计算机网络的应用,分布式人工智能在近几年的研究发展得较快。在一个开放的多智能代理环境中,一组代理之间的交互过程可以定义为:

$$E = \langle F, (A_i)_{i \in I} \rangle$$

其中: E 为多智能代理环境; F 为多智能代理环境中的语义框架模型; A_i 为一组智能代理, I 为代理的数目。

每个代理 A_i 又由一组知识库、自身的信念、消息队列以及相互之间通信的协议组成, A_i 可以表示

成:

$$A_i = (K_i, B_i, M_i, P)$$

K_i 为知识库, B_i 为 A_i 的信念, i 为 A_i 的消息队列, P 为代理之间相互通信的一组协议。

在分布式多媒体数据库系统中,我们可以将不同的媒体信息用不同类型的媒体代理来管理,每一种类型的代理又可以产生其相应的子代理,利用代理之间的协同机制研究多媒体信息的并行或交叉访问。

每个代理的知识库不仅仅局限于符号信息,而且可以赋予媒体信息。利用各种形式的媒体代理可以对媒体对象的语义、时态及空间关系进行模拟。多代理之间共享信息的调度方法可以解决分布多媒体数据库中资源的共享。

结束语 传统的智能系统是在符号化的单维信息空间建立起来的,对抽象的逻辑推理表现出很强的推理能力,但对形象思维的推理和学习显得无力。多媒体技术在多维化的信息空间中表示各种知识,提供了更加自然,更强有力的手段。多媒体和智能系统的结合必将把智能科学的研究和应用推向一个崭新的阶段。

参考文献

- [1] Guest Editor's Note, Multimedia: A New Dimension in Expert Systems Applications, Expert System With Application, 7(3)1994
- [2] AAAI Proc. of the 1991 Workshop on Intelligent multimedia Interface (1991 July), Anaheim, CA: American Association for Artificial Intelligence Publications
- [3] J. M. Ragusa, Models and Applications of Multimedia, Typermedia and Intellimedia Integration With Expert Systems, Same to [1]
- [4] A. D. Narasimhalu, A Framework for the Integration of Expert Systems Multimedia Technologies, Same to [1]
- [5] E. J. Gary & J. C. Siptor, Multimedia as a Vehicle for Knowledge Modeling, Same to [1]