

科学数据库基于内容的多媒体检索系统

陈峰莲 阎保平 黎建辉 胡良霖

(中国科学院计算机网络信息中心 北京100080)

摘要 科学数据库中存在大量的多媒体数据,为了实现对多媒体内容的有效存储、管理和检索,基于内容的多媒体综合检索技术将成为技术研究的重点。本文首先分析了科学数据库多媒体资源的特点和对多媒体内容管理的需求,然后探讨了基于内容的多媒体检索技术的原理、特点和检索方法。最后提出了科学数据库多媒体检索系统的一套设计方案,并说明了该体系的结构和功能。

关键词 多媒体元数据,内容检索

Content-Based Multimedia Indexing and Retrieval System of Scientific Databases

CHEN Feng-Lian YAN Bao-Ping LI Jian-Hui HU Liang-Lin

(Computer Network Information Center, Chinese Academy of Sciences, Beijing 100080)

Abstract There are a large number of multimedia data in the scientific databases. In order to implement effective content-based multimedia storage, management and retrieval, content-based multimedia indexing and retrieval technique will be the emphasis of the work. The authors firstly analyze the characteristic of multimedia in the scientific databases and the requirements of content-based multimedia management, then introduce the principle, characteristic and method of content-based multimedia management technology. The authors finally present multimedia system framework, implement and function in the scientific databases.

Keywords Multimedia metadata, Content-based retrieval

1 需求分析

中国科学院科学数据库通过几十个科学院研究所近20年的共同努力,现已建成了三百多个不同数据类型、不同学科的科学数据库专业子库,形成了一个庞大的、分布式的、异构的数据库群。科学数据库中的很多数据库包含了大量的多媒体数据,如植物图像数据库、植物图谱数据库和遥感影像数据库等。数字标本馆里也包括大量的图像、3D和2D模型、视频和音频数据。如何对大规模的多媒体内容进行有效存储、管理和检索是我们需要解决的问题。因此,基于内容的多媒体综合检索技术将成为技术研究的重点。

如图1所示,我们试图通过对元数据技术、领域本体(Ontology)和多媒体数据库技术的研究,确定多媒体内容描述框架,研究和制定多媒体内容的存储模型,开发多媒体内容存储和管理的示范系统,实现对科学数据库中多媒体内容的有效管理。在此基础上,研究基于元数据、本体和内容的多媒体综合检索技术,开发多媒体综合检索试验系统。

2 基于内容的多媒体检索技术

2.1 基于内容的多媒体检索技术的特点

随着计算技术和互联网的飞速发展,以及新的图像/视频压缩技术的不断出现,出现了大量的多媒体数据。如何对大量的多媒体数据进行分析、存储和检索是一个严峻的问题。显然,传统的基于文本的检索已不能满足这种多元化需求,于是人们提出了基于内容的多媒体检索。基于内容的检索是根据媒体对象的内容及上下文环境,从中提取指定的特征(如颜色、纹理、形状等),在媒体数据库中检索、发现相似的多媒体资源。基于内容的检索突破了传统的基于文本检索技术的局限,直接对图像、视频、音频内容进行分析,抽取内容特征,利用这些内容特征建立索引并进行检索。在这一检索过程中,它主要以图像处理、模式识别、计算机视觉、图像理解等学科中的一些方法为部分基础技术,是多种技术的合成。

基于内容的多媒体检索技术具有如下特点:

(1)直接从资源中提取内容特征:基于内容的检索突破了传统的基于关键词检索的局限,直接对图像、视频、音频进行分析,抽取特征,使得检索更加接近媒体对象。以图像的特征

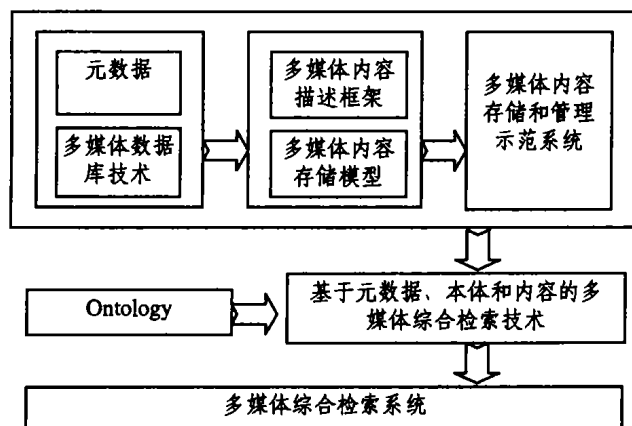


图1 项目目标

陈峰莲 在读硕士,主要研究方向:多媒体元数据,多媒体内容管理。阎保平 博士生导师,主要研究方向:大规模数据库应用技术,多媒体数据库,智能搜索引擎,下一代互联网,大型网络及系统集成工程化规范管理。黎建辉 硕士生导师,主要研究方向:大规模科学数据库共享与集成应用,多媒体内容管理,XML与Metadata。胡良霖 主要研究方向:数据库技术,多媒体内容管理,元数据及其互操作。

提取为例,可以提取形状特征、颜色特征、纹理特征、轮廓特征等。

(2)相似性检索:在检索过程中,采用近似匹配和逐步求精的办法,每一层的中间结果是一个集合,不断减小集合的范围,直到定位到目标。摒弃了传统的精确匹配技术,避免了因采用传统检索方法所带来的不确定性。

(3)满足用户多层次的检索要求:基于内容的多媒体检索系统通常由媒体库、特征库和知识库组成。媒体库包含多媒体数据,如文本、图像、音频、视频等;特征库包含用户输入的特征和预处理自动提取的内容特征;知识库包含领域知识和通用知识,其中的知识表达可以更换,以适应各种不同领域的应用要求,而且可以根据用户的查询交互学习和优化。

(4)大型数据库(集)的快速检索:针对数量巨大、种类繁多的多媒体数据库,能够实现对多媒体信息的快速检索和定位。

基于内容的多媒体检索技术涉及内容广泛,需要利用图像处理、模式识别、计算机视觉、图像理解等领域的知识作为基础,还需从认知科学、人工智能、数据库管理系统、人机交互、信息检索等领域引入新的媒体数据表示和数据模型,从而设计出可靠、有效的检索算法、系统结构以及友好的人机界面。

2.2 基于内容的多媒体检索方法

(1)基于内容的图像检索:对于图像检索,常用的内容特征主要包括:颜色、纹理、形状等。

·颜色检索。在颜色检索中运用的关键技术包括:颜色的表示、颜色特征的提取和基于颜色的相似度量。颜色特征的提取主要有互补颜色空间直方图、直方图交叉法、直方图距离比较法、二次型距离算法等。因为颜色具有一定的稳定性,所以是基于内容的图像检索中应用最广泛的特征。

·纹理检索。纹理特征主要包括粗糙性、规则性、线条相似性、凸凹性、方向性和对比度。纹理检索的方法主要有共生矩阵法、K-L变换、多分辨率分析、纹理谱分析等方法。由于难以描述,对纹理的检索一般采用示例查询方法。纹理检索在图像分析和识别中起着重要的作用,被广泛应用于气象云图、卫星遥感图像等。

·形状检索。形状检索的基础是图像边缘的提取。常用的形状检索方法主要有两种:针对图像边缘轮廓线进行的检索和针对图形矢量特征进行的检索。

(2)基于内容的视频检索:视频又称动态图像,是一组图像按时间的有序连续表现,它的表示与图像序列、时间关系有关。视频序列主要由镜头组成;镜头由一系列连续的帧组成;帧是一幅静态的图像,是组成视频的最小单位;一般场景含有多个镜头,而幕是由一系列相关的场景组成的,表达一个完整的事件。视频检索的方法主要有三种:基于关键帧的检索、基于运动的检索和浏览。

·基于关键帧的检索。因为关键帧通常是一幅图像,因而可以采用与图像检索相似的方法。

·基于运动的检索。它基于镜头和视频对象的时间特征来检索,包括检索摄像机的移动、场景移动、运动方向和运动幅度等特征。

·浏览。对视频进行分类分组,然后通过浏览选择。

(3)基于内容的音频检索:音频数据一般用音量、音调、音强、带宽、音长和音色等属性来描述,其中音量、音调、音强、带宽和音长属性易于通过技术手段进行信息化建模,而对音色

的处理较为复杂。在检索前,首先对音频数据建立索引,索引可以基于韵律、和音、旋律以及其它的感知或声学特征。基于内容的音频检索就是将输入的字符序列和音频数据库中的字符序列相匹配。常用的音频检索方法有:赋值查询、示例查询和组织浏览查询。

·赋值查询。用户指定某些声学特征的值或范围来查询。

·示例查询。用户提交一个示例声音,针对一个或多个特征,查出所有与示例相似的音频数据。

·组织浏览查询。对声音进行分类分组,然后通过浏览选择。

3 科学数据库多媒体检索系统体系结构

科学数据库多媒体内容管理与检索系统采用了典型的三层体系结构,如图2所示,自下而上依次是资源与加工层、数据层和服务层。其中资源及加工层实现多媒体资源的收集、归类等基础支持,是整个系统的最基础的部分;数据层的任务是实现多媒体资源的有效管理,以及根据检索命令把匹配结果反馈给服务层;服务层是用户的访问界面,是用户访问资源、获取服务支持的窗口。

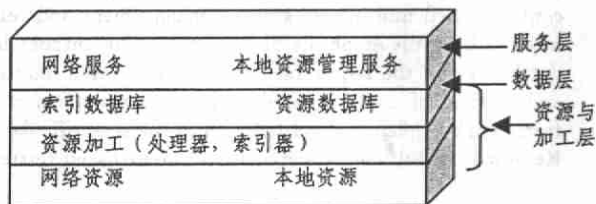


图2 系统三层体系结构

本系统将基于多媒体元数据实现对多媒体内容的管理和综合检索服务。基于系统的三层体系机构,我们设计了系统的框架图,如图3所示。

(1)资源与加工层。资源与加工层包括规范化预处理、元数据抽取和编辑、参考Ontology分类体系实现资源和元数据的入库及索引。

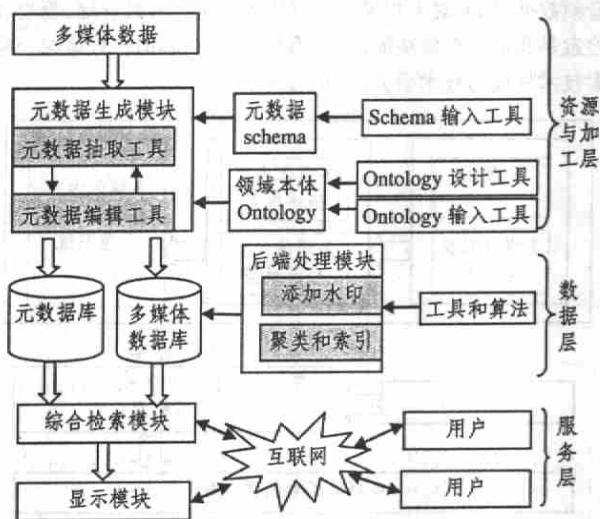


图3 系统框架图

元数据的生成模块将直接基于多媒体文件生成元数据。其中元数据抽取工具将实现自动抽取与多媒体内容直接相关的元数据(如图像的颜色、大小、视频的长度等),元数据编辑工具将实现对多媒体内容的上下文和语义进行描述,领域本

体可以帮助编辑人员更规范、更清楚地描述其所要表达的概念和知识。

Ontology 领域本体是本系统的一个技术关键。一般说来,本体是对一个领域中的公认的知识及其属性、它们之间的关系的一种明确的、形式化的描述。本体明确定义了一个领域中不同的人 and 计算机都能理解的基本概念(Concept)以及概念之间基本关系。通过本体可以实现人与人之间、人与计算机之间,计算机与计算机之间信息共享。

(2)数据层。数据层包括多媒体库及其索引库、元数据库及其索引库。

多媒体数据库实现对多媒体内容的存储和管理,多媒体数据库将采用“对象-关系”模型来实现。在多媒体数据库后面将有一个可以根据需要不断扩展的后端处理模块,负责对多媒体文件添加水印,利用成熟的工具或者算法实现基于多媒体内容的聚类索引,直接为未来的检索服务,提高检索效率。

元数据库实现对元数据的存储和管理。从便于元数据互操作的角度出发,元数据可以采用 Native XML 数据库。但是,考虑到目前该技术不是非常成熟,在我们以前的试应用中效果不太理想。因此,我们采用传统的关系数据库进行管理,但提供元数据转换成 XML 文件的工具。

(3)服务层。服务层包括综合检索模块和显示模块。

综合检索模块是本系统最复杂和最具特色的一项功能。综合检索模块将实现基于用户提交的关键词(概念)、基于元数据属性以及基于多媒体内容的综合检索。基于关键词的检索将以领域本体为基础,研究和基本实现以本体为基础的信息检索方式。基于内容的检索将以图像为基础,实现基于图像

颜色、形状、纹理等的特征检索。视频和音频的基于内容检索技术作为研究的方向。

显示模块实现对多媒体内容的浏览和显示,包括动态生成缩略图(Thumbnail)等。

结论 通过研究基于内容的多媒体检索技术,本系统将实现对多媒体内容元数据的抽取、编辑和管理等功能,利用数据库实现对多媒体内容的集中管理,利用元数据实现对多媒体内容的查询、检索和定位。在此基础上,研究和引入领域本体,研究和开发基于概念、属性和内容的多媒体综合检索系统。

经过初步的研究和实践,本论文提出了多媒体检索系统的框架设计,在具体实现过程中还有许多问题需要深入研究,比如:多媒体数据存储和管理技术、多媒体元数据互操作技术、多媒体元数据自动分析和抽取技术、领域本体的研究等。同时,这套设计方案本身还需要进一步改进和完善。

参考文献

- 1 Kiranyaz S, Caglar K, Guldogan E, et al. MUVIS: A Content-based Multimedia Indexing and Retrieval Framework[J]. IEEE, 2003(2)
- 2 Chang S-F, Chen W, Meng H J, Sundaram H, Zhong Di. VideoQ: An Automated Content Based Video Search System Using Visual Cues[J]. ACM, 1997(2)
- 3 Smith J R, Chang S-F. VisualSEEK: a fully automated content-based image query system[J]. ACM, 1996(1)
- 4 Weber R, Bolliger J, Gross T, Schek H J. Architecture of a Networked Image Search and Retrieval System[J]. ACM, 1999(11)
- 5 MPEG-7 [S]. <http://ipsi.fhg.de/delite/Projects/MPEG7/>
- 6 Virage. www.virage.com

(上接第89页)

3) 语义分析:依据语法分析的结果,以及规则语句中每个词语的语义关系(在业务对象模型中反映),确定规则语句的语义。语义分析为词语添加语义标识,并将中文表示转换成系统内部引用时的英文标识。

4) 转述:根据语义分析的结果,将〈模式〉和〈动作〉翻译成 OBJRLL 形式,完成整个中文规则语句的转换。

3.3.2 面向对象业务规则编程语言 OBJRLL 开发人员具备计算机对象设计和编程知识,为了方便他们定制规则,加速业务规则的开发,我们采用类 Java 语言的文法,定制了面向对象业务规则编程语言 OBJRLL (Object-oriented Java-like Business Rule Language)。

在 OBJRLL 中,业务规则的基本形式为:IF 〈Pattern〉 THEN 〈Action〉。在〈Pattern〉和〈Action〉中采用类 Java 的规范访问对象。部分语法如下:

```
(RuleStatement)::=[ruleHeader] IF 〈Pattern〉 THEN 〈Action〉
(ruleHeader)::= ruleName = 〈string〉; [priority = 〈integer〉];
(argumentList)::= argument 〈parameterList〉
(Pattern)::= 〈objectExpression〉 { 〈logicLinkWord〉 〈objectExpression〉 } *
( objectExpression ) :: = [ 〈 variable 〉 ; ] 〈 object 〉 (
  (conditionExpression) { 〈object〉 }
(Action)::= 〈ActionExpression〉 +
( ActionExpression )::= 〈variable〉. 〈object〉 ( 〈par〉 ; ) { 〈object〉. 〈object〉 ( 〈par〉 ; ) }
```

该语言支持面向对象的功能,并根据业务规则的特点,具备变量指代、集合运算等功能。例如下面这条网上购物时的规则:“如果购物车中有鱼,则在购物车中添加鱼片赠品。”

```
ruleName = "ruleAddSampleFlakedFish";
IF ?sc:ShoppingCart { 'fish' IN Items[] }
```

```
THEN ?sc.addItem("SampleFlakedFish");
```

3.3.3 业务规则执行语言(Rule Execution Language)

是规则引擎的执行语言。我们系统采用的规则引擎是 JESS6.0,相应的执行语言是 CLIPS。用 OBJRLL 编写的规则需要编译成 CLIPS 的形式才能被规则引擎执行。我们定制了词法、语法文件,采用相关编译器构造技术生成了 OBJRLL 的编译器。

总结 本文介绍了业务规则管理系统的原理,并设计了一个三层业务规则语言模型,实现了一个轻型业务规则管理系统。系统为业务分析人员定制了受限中文业务规则语言,提供实用方便的规则开发工具,将企业的业务逻辑及其实现相分离,并具备良好的集成性,使得企业的信息系统能够更快更好地适应业务需求的变化。在今后的工作中,我们将进一步完善该系统的功能,如为业务规则引擎提供更多的封装方式,方便系统的应用集成;完善中文支持功能,为业务人员定制灵活方便的中文业务规则模板,方便规则的录入。

参考文献

- 1 Morgan T. Business Rules and Information Systems. Boston: Addison-Wesley, 2002
- 2 Business Rule Group. <http://www.businessrulesgroup.org/>, 2004
- 3 Giarratano J, Riley G 著. 印鉴, 刘星成, 汤庸译. 专家系统原理与编程. 北京: 机械工业出版社, 2000
- 4 Giarratano J C. CLIPS User's Guide. v6. 20. Boston: PWS, 2002. 3
- 5 JESS. <http://herzberg.ca.sandia.gov/jess/>. 2004. 3