

多媒体信息系统检索技术研究及其面临的挑战

王小霞

(中国社会科学院民族学与人类学研究所网络信息中心, 北京 100081)

摘要: 介绍了多媒体信息系统国内外发展现状及趋势, 论述了多媒体信息系统的检索技术: 基于文本的检索、基于内容的检索和多媒体的语义表示与语义检索以及多媒体信息系统所面临的挑战。

关键词: 多媒体信息系统; 检索; MPEG

中图分类号: G354.4

文献标识码: A

文章编号: 1674-7720(2010)13-0086-04

Study of multi-media information system and challenges it faces

WANG Xiao Xia

(IT Center, Institute of Ethnology and Anthropology, Chinese Academy of Social Sciences, Beijing 100081, China)

Abstract: This paper gives a brief introduction using of multi-media information system, The paper also discusses text-based retrieval, content-based retrieval and the challenges it face, and the possible measures may be taken to deal with them.

Key words: multi-media information system; searches; MPEG

信息技术以其强大的渗透力深入到社会经济生活的各个方面。在商业金融等领域, 多媒体信息系统作为一种新的手段正在被广泛使用。多媒体技术与计算机、网络技术、通信技术、数字技术的结合, 使学术科研信息的传播不受时间、地点、国界、环境等影响, 这有利于提高世界信息的流通速度, 促进远隔重洋的各民族文化科技的及时交流。

多媒体系统是把声音、图形、图像和计算机系统集成为一个整体, 通过计算机对多媒体进行数字化处理。多媒体不是简单地叠加起来, 而是有机地结合、加工处理并改善信息表示, 从而吸引人的注意力。

多媒体技术, 即是计算机交互式综合处理多媒体信息——文本、图形、图像和声音, 使多种信息建立逻辑连接, 集成为一个系统并具有交互性。简言之, 多媒体技术就是具有集成性、实时性和交互性的计算机综合处理声、文、图信息的技术^[1]。

1 多媒体信息系统国内外发展现状及趋势

早在 1983 年第九届国际超大规模数据库会议上, D.Tsichritzis 等人就提出了多媒体数据库 Multimedia Database 的概念。由于多媒体信息的特殊性, 在信息的收集、存储、检索、管理、处理、传输、表现等方面都需进行新的研究, 采用新的技术和方法。

当前国内外从事多媒体信息检索系统开发工作的机构很多。在这里简单介绍一下国外 IBM 公司开发的 QBIC(Query ByImage Content)系统和国内中国科学院计算技术研究所和国家图书馆开发的 MIRES(Multimedia Information Retrieval System)系统^[2]。

QBIC 系统是 IBM Almaden 研究中心开发的基于内容的图像检索系统, 它可以对图像、视频、文本、声音进行检索。该系统主要由数据库生成部分和查询部分两部分组成。在数据库生成时, 每一个图像对象和视频对象的内容特征, 包括颜色、纹理、形状以及摄像机和对象的移动等, 都被提取出来并存入数据库中。当查询时, 数据库查询部分把用户利用图形化方法提供的对象特征与数据库中存储的对象内容特征进行比较匹配, 找出具有相似特征的图像和视频。

中国科学院计算技术研究所和国家图书馆已成功研制了基于特征的多媒体信息检索系统 MIRES。此系统基于 Internet 的多媒体信息检索系统, 可以实现对图像、视频、声音的基于内容的检索和对文本的全文检索。系统组成如下: (1) 基于形状的图像信息检索。 (2) 基于全局纹理特征的图像检索。 (3) 基于多特征与相关反馈的图像信息检索。 (4) 文本信息检索。除了常用的关键词匹配方式之外, 还提供了组合查询条件、模糊词切分、整句

匹配等多种查询技术,并且具有一定的概念检索功能。(5)基于主体技术的并行信息检索。搜索引擎可将用户的要求进行分解,把相应的查询请求发送到具有特定功能的服务主体,然后把从各个主体返回的结果进行综合,展现给用户。

现在国内最大的多媒体信息系统是新华社多文种多媒体新闻信息处理系统,即多媒体新闻采编系统。该系统集群技术、负载均衡技术实现了大规模用户并发访问条件下的系统整体高响应能力和不间断服务能力,有较强的系统扩展能力。系统采用 UTF-8 编码技术实现了多语种稿件的处理、存储和检索,支持中英西法俄葡日等语言稿件的处理,并且在新系统支持下,实现了西法俄阿葡等文种多媒体稿件的签发功能。该系统 2005 年 7 月开始运行,是集文字、图片、图表、视频等加工处理、存储传输于一体的信息系统。

目前,多媒体技术正向三个方面发展:(1)计算机系统本身的多媒体化;(2)多媒体技术与点播电视、智能化家电、识别网络通信等技术互相结合,使多媒体技术进入教育、咨询、娱乐、企业管理和办公室自动化等领域;(3)多媒体技术与控制技术相互渗透,进入工业自动化测控等领域^[3]。

2 多媒体信息系统检索技术研究

多媒体信息现已成为计算机信息处理系统的主要数据资源,如何对这些信息关联度大、结构复杂、媒体处理要求高的多媒体数据进行有效管理和使用,成为多媒体信息系统的一大技术难题。

多媒体信息系统主要面临以下问题:(1)数据量大,且媒体差异极大,从而影响信息系统的组织和存储方式;(2)媒体种类的增多增加了数据处理的难度;(3)多媒体不仅改变了数据库的接口,使其声、文、图并茂,而且也改变了数据的操作形式,其中,最重要的是查询机制和查询方法;(4)传统的事务一般都短小精悍,但多媒体数据不能满足这个要求,如从动态视频库里提取并播放一部数字化影片;(5)版本控制问题。在具体应用中,往往要涉及到处理对象的不同版本的记录和处理问题。

多媒体信息检索技术是针对用户对数值、文本、图形、图像、动画、声音等各种形式信息的全方位需求而提供的信息查找技术。当前的多媒体信息检索主要包括基于文本的检索(Text-Based Retrieval)和基于内容的检索(Content-Based Retrieval)两种方式^[3]。

2.1 基于文本的检索

基于文本的多媒体信息检索的基本原理是通过通过对多媒体信息的分析找出其要表达的内容,并根据信息的内容给出文字描述,通过检索其描述文字进而找到所需的多媒体信息。例如:通过人工标注的方法给数据库中每个图像标注一个或多个关键字属性,用户通过关键字的匹配对图像或视频进行检索。从本质上说,基于文本

的多媒体信息检索是对关键词的检索。这种检索方式的优点在于:(1)技术简单、标引和检索方便;(2)元数据和数据挖掘的研究将会大大加强文本信息检索的查准率和查全率,使得信息检索的效率大大提高,接近理想的检索效果。

基于文本的检索方式的局限性在于:(1)很难真正揭示多媒体信息的内容。这种检索途径使用文本来表达图像的内容,检索对象的不一致决定了在这种信息传递过程中必定会有大量信息的丢失,这样就不可能完全反映信息的内容。(2)多媒体信息是一种抽象程度很大、随意性很强的信息,缺乏一般意义上的规范性。对于这种信息,不同的人有不同的理解,这样便使得在用文字形式描述多媒体信息时,不同的人对同一多媒体信息的描述会有差异。(3)不能实现对于实时广播流媒体的处理。同时,文本不能有效地表示视频数据的时序特征,也不支持语义关系。(4)基于文本的检索方式中,多媒体信息是人工标引,相对当前数量庞大的多媒体信息而言,是很不现实的^[4]。

2.2 基于内容的检索

基于内容的多媒体信息检索,简单地说,就是根据多媒体信息的内容来进行检索,实际上包含两个方面:多媒体信息的内容描述和检索。要实现基于内容的信息检索,首先要知道信息的内容,然后才能在此基础上进行检索。通常情况下,基于内容的信息检索首先要对媒体信息进行分割,将其分成单独的可供检索的对象,然后再对每个媒体对象进行特征提取,媒体对象特征的集合就构成了它的内容描述。接下来,检索过程就可以根据检索要求从多媒体数据库中返回一组内容描述与检索要求最接近的对象。

与传统的基于关键词的检索手段相比,基于内容的检索是相似度检索和近似检索。

(1)相似度检索

在传统的数据库中,信息是以记录的形式来组织、管理和查询的,每个记录有固定数目的域,每个域都有明确的意义(定性描述),这样,每个记录所包含的语义信息是确定且有限的。所以在传统的数据库检索中,可以通过比较记录中某个域的值(关键词)与提供的值是否相同来进行检索,检索的结果是完全符合检索要求的。而对多媒体信息而言,同样的内容在不同的应用中具有不同的解释,即具有多义性的特点,这样,其内容就很难简单地通过几个属性来充分描述。基于内容的检索只能是一种相似度的检索,即存在一个反映数据库中的对象与检索要求相似程度的量,检索目的也是按照相似程度的大小顺序返回相似度最大的一组对象。

由于基于内容的检索是一种相似度的检索,因此多媒体库中的所有对象与所检索的要求之间都存在一定的相似度,如果不给出检索结果的集合限制,则检索将

返回所有的媒体对象,检索也就没有任何意义了。因此,在基于内容的检索中需要给出对返回结果的集合限制。

(2) 近似检索

基于内容的检索是面向多媒体数据库查询的,而由于媒体对象内容的丰富性以及一般检索表达并不能充分反映检索要求,因此在应用中往往并不要求查询结果一定是多媒体数据库满足相似度和检索结果集合大小限制的所有对象,而允许有所遗漏,其目的是换取其他性能的提高(比如检索速度)。

基于内容的多媒体信息检索具有如下特点:

(1) 直接对图像、视频、音频等媒体内容进行分析,抽取特征和语义建立索引,进行检索;

(2) 放弃常规数据库检索中的精确匹配方法,而采用相似性匹配的方法逐步迭代求精获得检索结果;

(3) 能对大型多媒体数据库进行快速检索;

(4) 采用多种检索手段,除提取多媒体内容特征进行检索外,还提供了其他检索手段,如通过提供样本图像进行相似性检索,或通过人机交互进行浏览检索^[7]。

3 多媒体的语义表示与语义检索

由于多媒体数据库中包含大量的图像、音频、视频等非格式化数据,其查询系统不能只提供基于媒体描述、关键字的检索和查询,而应能对图像或声音等媒体内容进行语义分析,提取相应的元数据,以实现基于语义的检索方法。

基于语义内容检索是一项实用性强的高技术,应用于遥感图像处理 and 空间探测等领域。

QBIC 是由 IBM Almaden 研究中心开发的基于内容检索系统,可以对图像、视频、文本和语音进行检索; VisualSEEK 是美国哥伦比亚大学电子工程系与电信研究中心图像和高级电视实验室共同研究的一种在 WWW 上使用的基于内容的图像/视频检索系统;美国加利福尼亚有限责任公司的 Muscle Fish 是一种正在开发中的音频分析引擎,用于数据库中声音的自动分类和检索。

要实现语义检索,首先要对媒体进行准确地语义表示。媒体语义的表示本质上是一种知识表示,但与一般的知识表示不同。首先媒体包含了大量的语义信息,并且这些信息之间存在着复杂的关系,因此需要有着强大表达能力的方法。目前主要有以下的媒体语义表示方法。

(1) 文本表示法

最简单的语义表示方法是用文本对媒体进行解释。在 IRIS 系统中,使用相似性技术直接从户外图像中推导出场景的自然语言描述。颜色、纹理、区域和空间信息被输入到图像解释器来获得每一图像区域最可能的解释,然后整个场景产生的文本描述,可以利用文本检索技术来检索。在某些情况下,利用词典(WordNet)将文本表示的相关语义概念联系起来,可以获得一定的模糊匹配能力。但是使用文本表示语义对于一些概念之间的复

杂关系缺乏足够的表达能力,因此不具有普遍意义。

(2) 传统的知识表示法

这些表示方法主要采用了人工智能中传统的知识表示,如语义网络、数理逻辑、框架等方法,它们具有表达复杂关系的能力。最近的一些研究者使用了一些不同的语义表示模型,如 Zhuang 等使用了模糊布尔模型、概率布尔模型;Colombo 等使用了形式语言理论表示;Meghini 使用了模糊逻辑语言, Marc Cavazza 等使用了符号语言学方法。这些方法在不同的场合分别显示了它们在语义表达或者模糊匹配方面的能力。但是目前还没有在不同的情况下都能表现出很好效果的通用方法^[4]。

(3) MPEG 技术

以上两种基于关键词或文本的检索方法已不适合于多媒体信息的检索。因此,对海量的多媒体信息进行组织、建库,达到快速、有效地检索的目的,已成为信息时代人们迫切需要解决的问题。基于内容的多媒体信息检索就是在这样的背景下提出的,并成为多媒体研究领域的一个热点。而实现这种基于内容的多媒体信息检索的一个关键性步骤是要定义一种标准的多媒体内容描述接口^[5]。

为了解决在多媒体信息检索方面所面临的问题, MPEG(活动图像专家组)在制定了 MPEG1、MPEG2 标准以及 MPEG4 标准草案后,又在着手制定 MPEG7、MPEG21、MPEG22、MPEG24、MPEG27 等标准。其中 MPEG7 标准的正式名称为“多媒体内容描述接口”(Multimedia Content Description Interface)^[6]。

MPEG7 与以往的 MPEG1、MPEG2、MPEG4 和 MPEG21、MPEG22、MPEG24、MPEG27 等国际标准不同,它旨在对各种不同类型的多媒体信息进行标准化描述,并使之与所描述的内容相联系,以实现快速有效的搜索。

MPEG7 将对各种不同类型的多媒体信息进行标准化的描述,并将该描述与所描述的内容相联系,以实现快速有效的搜索。该标准不包括对描述特征的自动提取,它也没有规定和利用描述进行搜索的工具或任何程序。MPEG7 主要致力于视听数据的信息编码表达上,换句话说就是集中在对多媒体材料的描述的通用接口的标准化上(表达内容的信息而不是内容本身)。正因为如此,MPEG7 致力于在数据资源的交互性与全球化和数据管理的灵活性上。

MPEG 技术是由文本信息时代向多媒体信息时代过渡的必然产物,其基于内容和语义的多媒体搜索引擎将使人们真正置身于随心所欲的多媒体世界。MPEG 标准将会使多媒体技术产生重大的变革。随着 MPEG 标准和网络系统的快速发展,MPEG 的应用也将蓬勃发展,为多媒体检索技术提供更多的方便。

4 多媒体信息系统所面临的挑战

多媒体信息系统需要对图形、图像、音频、视频、动

画等各种媒体进行综合管理,系统的开发日渐重视对这些不同媒体格式信息的集成与整合。如何建立有效的描述和检索机制对多种媒体格式的信息进行整合描述和集成检索,成为当前多媒体信息系统重点要解决的问题之一。多媒体信息系统的集成描述与检索面临以下挑战:

(1)多媒体信息的整合与统一描述。例如,对相同内容的不同媒体形式进行统一描述,以便检索时可屏蔽各种媒体之间的差别,而在浏览时对不同媒体又能区别对待;对同一媒体形式的不同主题内容进行统一描述,以便综合反映同一媒体的不同侧面。

(2)多媒体信息的索引和集成查询,特别是可查找可索引非结构化的图像、音频和视频数据,也就是图像、音频和视频数据的语义检索。如非线性视频的查找和索引,视频的本质是一组连续的图像帧,除了存在时间上的先后顺序外,其本身并不具备任何结构信息。不能像文本一样使用目录、章节、段落、句子和关键词的方式来检索和浏览。因此,如何对视频等非结构数据进行结构化分析与描述成为多媒体信息系统必须解决的问题。

参考文献

- [1] 谢超,陈毓芬.洛阳多媒体旅游信息系统设计与实现[J].测绘科学,2008,33(1).
- [2] 王焱.基于内容的通用视频检索系统框架设计[J].计算机应用研究,2004(3).
- [3] 杨慕莲,张芳芳.论数字图书馆中多媒体信息检索系统的构建.咸宁学院学报,2004,24(5).
- [4] 何立民,万跃华.数字图书馆中基于内容的视频检索关键技术[J].中国图书馆学报,2003(2).
- [5] 卢官明.基于内容的多媒体信息检索与 MPEG7 标准[J].计算机应用,2000(5).
- [6] 范新华,陈宏兵,许满武.基于 MPEG27 的多媒体搜索引擎构建[J].计算机应用研究,2004(11).
- [7] Pramila Gupta, James A Sykes. The conceptual modeling process and the notion of a concept: information modeling in the new millennium [EB/OL]. [2001-08]. <http://www.omg.org/mda>.

(收稿日期:2010-01-15)

作者简介:

王小霞,女,1971年生,硕士,主要研究方向:计算机软件工程与应用。

电子技术应用
APPLICATION OF ELECTRONIC TECHNIQUE
www.chinaAET.com