



杰出青年学者研究文丛
Distinguished Young Scholars

面向 互联网的多媒体大数据 信息高效查询处理

Internet Oriented Efficient Query Processing of
Multimedia Big Data Information

◎ 庄 毅 著



ZHEJIANG UNIVERSITY PRESS
浙江大学出版社



杰出青年学者研究文丛
Distinguished Young Scholars

面向 互联网的多媒体大数据 信息高效查询处理

Internet Oriented Efficient Query Processing of
Multimedia Big Data Information

◎ 庄 毅 著



ZHEJIANG UNIVERSITY PRESS
浙江大学出版社

图书在版编目 (CIP) 数据

面向互联网的多媒体大数据信息高效查询处理 / 庄毅著. —杭州: 浙江大学出版社, 2015. 6

ISBN 978-7-308-14413-1

I. ①面… II. ①庄… III. ①互联网络—情报检索
IV. ①G354.4

中国版本图书馆 CIP 数据核字 (2015) 第 032982 号

面向互联网的多媒体大数据信息高效查询处理

庄毅 著

丛书策划 张凌静 许佳颖

责任编辑 张凌静 (zlj@zju.edu.cn)

封面设计 刘依群

出版发行 浙江大学出版社

(杭州市天目山路 148 号 邮政编码 310007)

(网址: <http://www.zjupress.com>)

排 版 杭州中大图文设计有限公司

印 刷 杭州杭新印务有限公司

开 本 710mm×1000mm 1/16

印 张 20

字 数 349 千

版 印 次 2015 年 6 月第 1 版 2015 年 6 月第 1 次印刷

书 号 ISBN 978-7-308-14413-1

定 价 75.00 元

版权所有 翻印必究 印装差错 负责调换

浙江大学出版社发行部联系方式: 0571-88925591; <http://zjdxcbbs.tmall.com>

前 言

目前,互联网已经进入大数据时代,其包含的多媒体信息资源呈现出以下特点:①数量巨大,增长迅速;②内容丰富,形式多样;③结构复杂,分布广泛;④无序混乱,杂乱无章。面对这些浩瀚的多媒体信息资源,如何对其进行快速准确的检索以及高效的处理已经成为一个很重要的研究课题。

本书是笔者博士论文的扩展,力求从数据库层面对互联网环境下多媒体大数据信息的查询、索引及并行化处理技术进行较为系统的介绍,为网络多媒体大数据处理提供一个初步的解决思路,起到一定的“抛砖引玉”作用。同时结合最新的网络多媒体研究现状、发展趋势和最新应用,进行深入阐述和分析。全书共分为10章,从深度和广度上对高维多媒体大数据信息的高效查询处理技术进行阐述:第1章对网络大数据多媒体信息的查询、索引及管理进行综述;第2~4章从语义、内容及两者混合的角度,介绍多媒体查询及其索引机制;第5、6章结合网络多媒体技术最新发展趋势,分别介绍跨媒体检索和社会化媒体检索及其索引;第7章针对在单机环境下多媒体大数据信息查询性能低下的问题,分别介绍了数据在网格环境下和移动云计算环境下的可扩展并行查询技术;第8章针对用户的频繁查询请求,提出分布式网络环境下的高维相似及连接查询的多重查询优化方法,进一步提高查询密集条件下多媒体大数据查询的并发性;第9章介绍多媒体技术在大数据环境下的若干成功应用,如数字图书馆、在线网络购物等;第10章对网络多媒体大数据信息查询处理技术进行总结,并展望其未来发展。

本书得以顺利完成,首先要衷心感谢我在攻读博士学位期间遇到的两位老师,一位是我的博士论文导师庄越挺教授,另一位是吴飞教授。感谢他们在我读博期间所给予的悉心指导和提供的良好科研学术环境。同时,本书在编写过程中,得到了香港城市大学李青教授的支持和鼓励,在此向他致以衷心的感谢。特别感谢浙江大学出版社编辑在本书出版过程中给予的支持。最后感谢在本书撰写过程中给予我无私帮助的同事和朋友,他们是浙江工商大学的琚春华教授、凌云教授和姜波教授,杭州电子科技大学的胡华教授和胡海洋教授,南京财经大学的伍之昂副教授等,及为本书部分章节写

作做出贡献的学生曹建通、诸葛菲、陈一泉、王晓晴等。

书中的相关研究得到了国家自然科学基金(项目编号:61003074)、教育部人文社科基金项目(项目编号:14YJC2H235)、浙江省自然科学基金(项目编号:LY13F020008)和浙江省科技厅重点创新团队(计划编号:2010R50041及项目编号:2012R10041-06)的支持,在此一并表示感谢!

最后,谨以本书献给我敬爱的父母和妻子,没有他们的支持,本书很难顺利出版,在这里特别表示感谢。

书中部分图片下载自网络,在此对这些图片的原作者表示诚挚的谢意。如您看到本书使用了您的图片,敬请与我(zhuang@mail.zjgsu.edu.cn)或者本书的责任编辑张凌静(zlj@zju.edu.cn)取得联系。

由于网络多媒体及大数据处理技术发展日新月异,加上本人学识浅陋,必有许多不足之处,望老师和同学提出意见并指正。

庄毅

2014年12月于杭州

目 录

第 1 章 绪论	1
1.1 网络多媒体与大数据	1
1.2 多媒体查询技术	4
1.2.1 查询模式	4
1.2.2 计算模式	8
1.3 大数据与云计算技术	9
1.4 典型的网络多媒体大数据系统	12
1.5 本书的内容结构	14
参考文献	15
第 2 章 基于元数据的多媒体查询及索引	17
2.1 文本度量模型	17
2.1.1 布尔模型	17
2.1.2 向量空间模型	18
2.1.3 其他模型	23
2.2 文本索引	23
2.2.1 倒排文件	24
2.2.2 签名文件	25
2.3 小结	26
参考文献	26
第 3 章 基于内容的多媒体查询及索引	28
3.1 基于内容的图像查询与索引	28
3.1.1 图像视觉特征	28
3.1.2 图像相似度量	37
3.1.3 图像检索中的相关反馈	39

3.1.4	高维图像索引	40
3.1.5	实例:交互式局部距离图书法字索引	49
3.2	基于内容的音频查询与索引	61
3.2.1	音频特征	61
3.2.2	音频相似度量	66
3.2.3	高维音频索引	69
3.3	基于内容的视频查询与索引	70
3.3.1	预处理技术	70
3.3.2	视频相似度量	73
3.3.3	高维视频索引	76
3.4	小结	76
	参考文献	76
第4章	基于多特征的多媒体查询及索引	82
4.1	基于多特征的图片查询及索引	82
4.1.1	基于语义和视觉特征的图片查询及索引	82
4.1.2	基于客观和主观特征的图片查询及索引	83
4.1.3	实例:个性化书法字图片查询系统	97
4.2	基于多特征的音频查询及索引	108
4.3	基于多特征的视频查询及索引	110
4.3.1	多特征视频查询	110
4.3.2	多特征视频索引	111
4.4	小结	112
	参考文献	112
第5章	跨媒体查询及索引	114
5.1	引言	114
5.2	交叉参照图模型	116
5.3	异构媒体对象相关性挖掘	118
5.3.1	语义标注方法	118
5.3.2	链接分析方法	119
5.3.3	异构特征分析方法	121
5.3.4	其他方法	123
5.4	跨媒体索引	124

5.5 索引支持下的跨媒体查询	130
5.6 实验分析	131
5.7 小结	135
参考文献	135
第6章 社会化媒体查询及索引	137
6.1 引言	137
6.2 基于混合特征的社会化图片概率查询及索引	140
6.2.1 语义特征概率建模	140
6.2.2 基于选择性模型的查询策略选择	141
6.2.3 基于标签和视觉特征的社会化图片概率查询及索引	147
6.2.4 基于位置的社会化图片复合查询	151
6.3 社会化图片偏好查询	153
6.3.1 用户偏好向量	154
6.3.2 排序算法	159
6.3.3 用户偏好传递	164
6.3.4 偏好向量的增量维护	169
6.3.5 查询算法	170
6.4 基于多特征的社会化图片连接查询及索引	170
6.4.1 问题定义	170
6.4.2 连接查询算法	172
6.4.3 实验分析	186
6.5 小结	197
参考文献	197
第7章 分布式并行环境下多媒体大数据信息相似查询处理	199
7.1 引言	199
7.1.1 网格计算	199
7.1.2 点对点计算	200
7.1.3 云计算	201
7.2 网格环境下高效相似查询处理	203
7.2.1 预备工作	204
7.2.2 支撑技术	205

7.2.3	查询算法	211
7.2.4	实验分析	215
7.2.5	应用实例:数据网格环境下书法字并行检索	217
7.3	云计算环境下社会化图片相似连接查询处理	219
7.3.1	问题定义	219
7.3.2	数据分布	220
7.3.3	并行连接处理	221
7.3.4	实验分析	222
7.4	移动云计算环境下大尺度医学图像查询处理	224
7.4.1	预备工作	225
7.4.2	支撑技术	227
7.4.3	查询算法	247
7.4.4	实验分析	250
7.4.5	应用实例:移动云计算环境下的书法字并行检索	255
7.5	小结	256
	参考文献	257

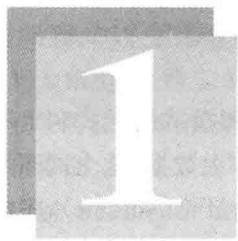
第8章 分布式并行环境下多媒体大数据信息多重查询优化

		259
8.1	引言	259
8.2	分布式并行环境下的多重相似查询优化	260
8.2.1	预备工作	261
8.2.2	动态查询层次聚类	264
8.2.3	MDSQ 算法	266
8.2.4	实验分析	267
8.3	云计算环境下社会化图片多重相似连接优化	269
8.3.1	问题定义及动机	269
8.3.2	连接盒聚类	271
8.3.3	并行多重连接查询	273
8.3.4	实验分析	276
8.4	小结	278
	参考文献	278

第9章 多媒体技术在大数据环境下的应用

9.1	多媒体数字图书馆	279
-----	----------------	-----

9.1.1 国内外数字图书馆的发展	280
9.1.2 多媒体检索在数字图书馆中的重要性	282
9.1.3 代表性的数字图书馆系统	283
9.2 基于众包模式的个性化自助旅游平台	286
9.2.1 关键技术	288
9.2.2 系统结构	289
9.2.3 功能模块设计	290
9.3 基于视觉感知的网络商品搜索	292
9.3.1 基于图片的网络商品搜索	292
9.3.2 移动商品视频搜索	295
9.4 小结	300
参考文献	300
第 10 章 挑战及发展趋势	301
10.1 面临的挑战	301
10.2 发展趋势	303
参考文献	306
索 引	309



绪 论

1.1 网络多媒体与大数据

当今人类正同时身处两个时代:互联网(Internet)时代和大数据(big data)时代。

互联网的出现使得一个全球范围的信息基础设施逐步形成,人类社会的信息化进程大大加快,同时人与人之间的距离被“缩短”了。作为一个巨大的数字资源库,互联网包含了大数据量的文字、图像、音频和视频等数字化资源。人类可以通过互联网方便地查询和共享所需的信息,同时也可以借助它方便地与远在异地的朋友进行实时通信和交流,因此它的出现确实极大地改变了人类的生活方式。人们接触信息的器官“眼”和“耳”在无形中被延伸。

纵观互联网的发展史,不难发现在 50 多年时间里,其规模得到飞速发展。据统计^①,截至 2012 年 3 月,全球活跃网站数量已达到 5.55 亿个,比 2010 年翻了一倍。网页数目在 5000 亿以上,为全球几十亿计的网民提供了各种服务。同时,作为全球信息资源的重要载体之一,互联网的作用越来越凸显。截至 2013 年年底,互联网总数据量已突破 5000 亿 GB。除此之外,据

^① <http://www.techweb.com.cn/news/2012-03-09/1164055.shtml>.

互联网数据中心的报告,全球产生的数据量于 2012 年已达到 2.7ZB^①,预计到 2020 年,全球的数据总量将达到 35ZB^②。如“图灵奖”获得者 Jim Gray 博士所断言:互联网环境下,每 18 个月产生的数据量等于有史以来数据量之和。

为此,2008 年 9 月的《自然》(*Nature*)杂志以社论和报道的形式刊出了一期专刊(*Big Data: Science in the Petabyte Era*)^[1]，“大数据”这个词开始被广泛传播。IBM 将大数据的特点总结为三个 V,即大量化(volume)、多样化(variety)和快速化(velocity)^[2]。首先,网络数据不断增长,数据规模已从“超大规模”的 GB 级,到“海量”的 TB 级,再到“大数据”的 PB 级及以上级别;其次,网络大数据类型繁多,包括结构化数据、半结构化数据和非结构化数据。在当前互联网应用中,以多媒体数据(如图片、音频等)为代表的非结构化网络数据已成为互联网中的重要信息类型,数据量已呈现出爆炸式增长的趋势,占有比例将达到互联网数据量的 75%以上^[2]。同时,社交网络(Social Network)和移动互联网(Mobile Internet)的兴起使得大量的用户生成内容(user generated content, UGC),如文本、图片、音频、视频及位置信息等非结构化数据也急剧增长。据统计^[2],一分钟内,Flicker 上会新上传 3125 张照片,Facebook 上新发布 70 万条信息,YouTube 上有 200 万次观赏^③。除了大数据量和数据类型复杂以外,网络多媒体数据动态性明显,随时都在变化。

从商业、经济及其他领域到国家的决策行为,大数据管理与分析都发挥着日益重要的积极作用。各国都开始密切关注和重视对大数据的研究。2012 年瑞士达沃斯世界经济论坛上的一个特色主题为“大数据,大影响”。同年 3 月,美国总统奥巴马宣布投资 2 亿美元,启动“大数据研究和发展计划”^[1]。可以说互联网、物联网和移动互联网催生和加速了大数据时代的到来。

一般来讲,当前互联网中的多媒体信息资源具有以下主要特点。

(1) 数量巨大,增长迅速。

根据权威机构国际数据公司(International Data Corp., IDC)最新的研究报告《数字宇宙膨胀:到 2010 年全球信息增长预测》中的统计数据^④,2006

① 1ZB(zetta byte)=1024EB,1EB(exa byte)=1024PB,1PB(peta byte)=1024TB(tera byte),1TB=1024GB(giga byte)。

② <http://news.sciencenet.cn/htmlnews/2013/5/277742.shtm>。

③ <http://www.chinairn.com/news/20120917/132442.html>。

④ <http://www.dostor.com/article/2011-03-28/9370234.shtml>。

年全球每年制造、复制出的数字信息量共计 1610 亿 GB,而人类开始记录历史以来,到 2006 年为止全部的书本文字加起来大约为 50PB,显然当年信息产生量大约是图书信息总量的 300 万倍。如果将书籍排列起来,总长度为地球到太阳距离(约 1.5 亿千米)的 12 倍。IDC 报告同时显示,至 2010 年,这个数字将猛增 6 倍,达 9880 亿千兆(GB),年复合增长率为 57%。与这个报告相佐证,AT&T 的网络每天流动 16PB 的数据,Google^[3] 每天处理 20PB 的数据,Facebook^[4] 每天存储 1PB 的照片,Opera 浏览器每个月处理多于 1PB 的数据,而 BBC 的 iPlayer 每个月有大约 7PB 的数据流,YouTube^[5] 存储了 31PB 的流媒体数据。Cisco 公司统计:到 2012 年每个月网络上视频流大约为 5EB(5000PB)。

(2) 内容丰富,形式多样。

海量(大数据)多媒体数据资源不仅来源于互联网,还来源于其他行业,包括不同领域(如物联网、传感器网络和遥感领域等)、不同学科(生物信息学、情报学等)、不同地区和不同语言的各种信息,其内容非常丰富,且以文本、图像、音频、视频等多种形式存在。为了实现这些多媒体资源信息的统一共享与查询,需要对这些媒体对象的跨媒体及跨语言处理技术进行研究。

(3) 结构复杂,分布广泛。

大多数多媒体信息(图像、音频、视频和 Flash 等)为非结构化信息,具有复杂的结构和不同的文件格式,增加了处理的难度。同时,互联网是开放性的,通过 TCP/IP 协议将不同的网络互联起来,对网络信息资源的组织管理并无统一的标准和规范,各种多媒体信息以网页的形式分布在不同地区的服务器上。

(4) 无序混乱,杂乱无章。

互联网改变了信息发布和评价的程序,使得网络信息的分布具有很大的随意性,数据的质量参差不齐。同时,由于缺乏有效的网络搜索技术,查询返回的数字资源显得杂乱无章,难以有效实现资源共享的目的。具体表现在两方面:①二义性,同一个关键词可能会对应多个不同类型的对象,如“苹果电脑”和“苹果水果”,需要实时地对搜索结果进行聚类处理;②重复性,互联网中多个网站可能包含相同内容的搜索对象,这样会使搜索结果产生冗余,需要对其进行去重检测(near-duplicate detection)。

综上所述,尽管已经有许多商用的搜索引擎,如 Google^[3] 和 Yahoo^[6] 等,以及数字图书馆所提供的在线查询服务已得到了成功的应用,但这些查询服务大都针对文本信息。面对互联网中的多媒体大数据(包括图像、音频、视频、跨媒体及社会媒体等),如何对其进行有效的数据组织和索引,以

实现高效、精准的查询处理,给传统数据库、信息检索、机器学习以及计算机视觉等领域的研究人员提出了新的挑战,需要我们进一步探索和研究。因此,本书重点介绍网络多媒体大数据信息的高效查询处理的若干关键技术。

1.2 多媒体查询技术

如上节所述,网络及数字媒体技术的迅猛发展使得互联网的规模及所包含的信息量都发生了巨大的变化,它已成为人们日常生活中获取信息的重要来源。如图 1.1 所示,作为互联网中数据的主要信息载体,多媒体信息广泛出现在各类网站和网页中,从简单的文本到图像、声音、视频等,并且其规模还在飞速增长中。本节将从多媒体信息的查询模式和计算模式两个角度来介绍多媒体信息查询技术的发展及其局限性。

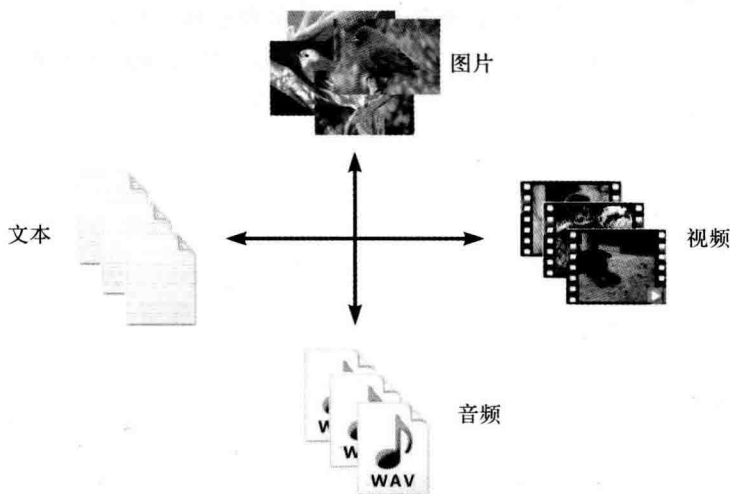


图 1.1 不同类型的多媒体数据

1.2.1 查询模式

多媒体查询技术经历了 40 多年的发展,如图 1.2 所示,其查询模式经历了从 20 世纪 70 年代到 80 年代的基于元数据(文本)的多媒体查询^[22],发展到 90 年代初的基于内容的多媒体查询^[7~11],再发展到 90 年代末的语义和

内容相结合的多媒体查询^[12,13],最后发展到目前热门的跨媒体查询^[14~16]及社会化媒体查询^[17,18]。其中,前两者是针对单一类型媒体对象的查询,后者则是针对多种类型媒体对象的综合查询。从整个多媒体查询技术的发展趋势可以看出,它从过去的支持单一类型(简单)媒体对象的查询发展到了目前的支持多种类型(复杂)媒体对象的查询。下面对这五种多媒体查询模式作具体阐述。

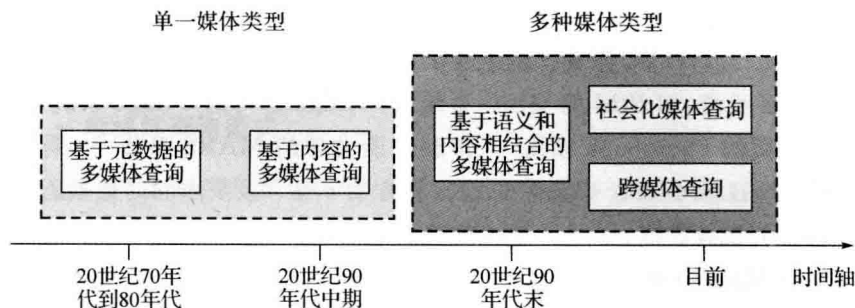


图 1.2 多媒体查询模式的发展

• 基于元数据(文本)的多媒体查询模式

基于元数据(文本)的多媒体查询方式是将成熟的信息查询(information retrieval, IR)技术应用于多媒体信息的查询。正如我们所知,文本信息查询技术在过去的几十年中得到了充分的研究,并已成功运用于诸如 Google^[3]、Lycos^[19]等商用搜索引擎中。在20世纪70年代末,该技术首次用于图像查询中,这就是基于关键字(元数据)的图像查询。这种方法的通用流程是:首先人工对图像用关键字进行注释,然后通过匹配用户查询(关键词)和图像的注释来搜索相关图像。Tamurn 和 Yokoya^[20]中对使用这种方法的系统进行了综述。该方法的优点在于基于关键字匹配的多媒体查询的查询效率要远远优于基于内容的查询模式。然而这种方法的缺陷也是显而易见的:首先,随着多媒体信息数据量的增加,人工注释需要大量的劳动力和时间,因此这种方法不适用于大规模的数据集合;其次,由于完全依赖人工来标注图像(或视频)中的对象、事件等所有信息,它所支持的查询的复杂程度也完全取决于人工标注的详尽程度;最后,由于不同的人对同样的图像/视频有不同的理解,甚至可能出现错误理解,导致图像注释的不精确性,从而引起查询过程中的错误匹配。

• 基于内容的多媒体查询模式

以基于内容的图像查询(content-based image retrieval, CBIR)^[7~11]为代表的基于内容的多媒体查询技术的研究始于20世纪90年代初。该技术通过分析和提取多媒体对象的底层视觉(如颜色、纹理和形状等)和听觉(如MFCC系数)等特征,利用相似度量方法(如欧式距离(Euclidean distance)、曼哈顿距离(Manhattan distance)及度量时序数据的动态时间规整(dynamic time warping, DTW)距离等进行多媒体对象的相似匹配。早期最有代表性的CBIR系统,如IBM的QBIC系统^[7]、Virage公司的Virage系统^[8]、麻省理工学院的Photobook系统^[9]、伊利诺斯大学的MARS系统^[11]、哥伦比亚大学的Visual SEEK (Web SEEK)系统^[21,22]等。该技术同时也被运用到基于内容的视频查询,如卡内基·梅隆大学的Informedia系统^[10]和哥伦比亚大学的VideoQ系统^[23]等,基于内容的音频查询^[24]、基于内容的3D模型查询^[25]等。

一般来说,基于内容的查询方法的局限性在于,它所用来描述图像或其他多媒体数据的特征是一些底层的视觉/听觉特征,而人们却习惯于在语义层次上衡量查询结果的相关与否。以目前的计算机视觉技术,我们还很难从多媒体内容的底层特征中准确得到其高层语义,因此到目前为止,基于内容的多媒体查询技术的查询效果往往不太理想,离实际应用还有很长的距离。为了克服“语义鸿沟”的问题,Rui等人首次将相关反馈思想应用于基于内容的图像查询^[26],通过交互式的手段改变高维空间中不同图像对象之间的距离权重,以提高基于内容多媒体查询的准确度。最近,Kelly等人提出隐形相关反馈(implicit relevance feedback)的概念^[27],这在用户层面上弥补了原来显式相关反馈(explicit relevance feedback)的不足。

• 基于语义和内容相结合的混合查询模式

混合多媒体查询模式是一种结合了基于语义查询模式和基于内容的查询模式两者的优点而提出的一种改进的查询模式^[12,13,28]。混合查询模式通过对基于语义和基于内容两种查询方式得到的查询结果进行融合分析,使得到的查询结果既能反映语义层次上的相似性,又能体现底层特征上的相似性,从而可以进一步提高查询效率,包括查全率和查准率。Zhou等人^[12]提出了一种新的基于混合模式的互联网图片查询方法。对于一个给定的查询,首先进行基于语义的查询得到一组候选图像集,然后对得到的候选图像集进行基于内容的求精处理。实验表明,通过该方法得到的查询结果要明

显优于基于单一特征的查询。混合查询模式在视频查询领域也得到了广泛的应用,这是因为原始视频数据包含多种类型的信息,如语义信息可以通过声音识别或提取字幕信息等方式得到,视觉的底层信息可以通过视频关键帧获得,听觉信息则可通过提取视频中伴随的音频信息获得。例如,Rong等人^[28]提出一种将不同查询结果进行融合分析的概率模型,其中包括文本查询,基于底层视觉信息的相似比较和语义的概念,可以得到与给定例子相似的视频镜头。实验结果表明混合的多媒体查询方式确实能大大提高查询结果的查全率和查准率,是一种非常有望的查询技术。

• 跨媒体查询模式

近年来,互联网与多媒体技术的迅猛发展使得多媒体数据呈现爆炸式增长的趋势,多种异构的多媒体数据(如图像、视频、文档等)在 Web、数字图书馆以及其他多媒体应用中大量涌现,它们有着相似的语义表达。例如,如果用户在数字百科全书中查询“Great Wall”(长城),他可能会希望得到有关长城的文字介绍、图片,甚至是反映当地风土人情的视频短片。但是,几乎所有现有的查询系统或方法都只是针对某种特定媒体对象的查询,它们在上述这些应用中有很大的局限性:首先,它们局限于某种单一类型的媒体(如单纯的图像查询方法);其次,它们仅依赖多媒体数据的某种特定的特征(如关键字的 $TF \times IDF$ 或图像的颜色、纹理特征等),因此难以提供在语义层面上的相关查询结果。现有的基于单一类型媒体对象的查询技术无法满足大量应用中人们对多媒体信息查询的新需求。

跨媒体(cross media)查询机制能够“兼容”属于各种不同模态(类型)的多媒体数据(如文本、图像、视频等)。同时,它能够表达并利用多种类型信息,包括多媒体数据的底层特征、文本中的关键字、数据之间的超链接等,并且能够综合运用多种查询方法得到查询结果。与基于内容的查询方法相比,这种查询机制不但能获得更为丰富的查询结果,而且尽可能运用多方面的知识进一步提高查询结果的相关度,是一种非常主动(相对于保守而言)的查询机制。早在 1976 年,上述结论已被麦格克等人^[29]所揭示,他们认为人脑对外界信息的认知需要跨越和综合不同的感官信息,以形成整体性的理解。近期,认知神经心理学方面的研究也进一步验证了人脑的认知过程呈现出跨媒体的特性,认知结果来自视觉、听觉等不同感官的信息相互刺激、共同作用。基于以上分析,跨媒体查询具有坚实的理论依据和现实意义,将成为今后多媒体领域的一个新的研究方向。

目前,跨媒体查询研究已取得了一些研究成果^[14~16]。为了解决异构媒