

信息媒体数字视频图书馆 ——开创多媒体数字图书馆先河

□ 刘燕权 / 美国南康涅狄格州立大学 美国康州纽黑文 06515
孙秀良 / 北京化工大学图书馆 北京 100029

摘要: 信息媒体数字视频图书馆 (The INFORMEDIA Digital Video Library, 简称IDVL) 项目是美国数字图书馆创始计划 (Digital Libraries Initiative, 简称DLI) 最初的六个项目之一, 由卡内基·梅隆大学于1994年创建。该项目一直在深入地开发数字视频及其相关技术和工具, 致力于建立一个庞大的在线数字视频图书馆, 为用户提供多媒体信息服务。文章对该数字图书馆的项目概况、资源组织、技术特征和服务特点等进行了综合介绍和评价。

关键词: 数字图书馆, 卡内基·梅隆大学, 美国数字图书馆创始计划(DLI), 信息媒体项目, 数字视频, 美国国家科学基金会 (NSF)

DOI: 10.3772/j.issn.1673—2286.2014.03.010

1 项目概述

信息媒体数字视频图书馆 (The INFORMEDIA Digital Video Library, 简称IDVL) 项目由美国卡内基·梅隆大学计算机学院于1994年创建, 是美国数字图书馆创始计划 (Digital Libraries Initiative, 简称DLI) 项目之一, 也是当时美国唯一研究视频信息的项目。Informedia项目的参与机构以卡内基·梅隆大学为主, 合作机构包括Microsoft公司、数字设备公司、贝尔大西洋网络服务公司、QED广播公司、英格兰米尔顿·基尼斯的开放大学、温彻斯特Thurston中学和费尔法克斯县的公立学校等, 目标是建立一个联机交互的在线数字视频图书馆系统, 使用户能够通过网络查询、检索和获取科学与数学方面的视听资料。Informedia项目的网址为<http://www.informedia.cs.cmu.edu/>, 主页如图1所示。

Informedia项目的开发经历了两个阶段: 1994—1998年为第一阶段, 称为Informedia-I; 1999年以后为第二阶段, 称为Informedia-II。Informedia-I是DLI-I的一个组成部分, 接收美国国家科学基金会 (NSF)、美国国防部高级研究计划署 (DARPA) 和美国宇航局 (NASA) 480万美元的资助, 目标是赋予音频、视频

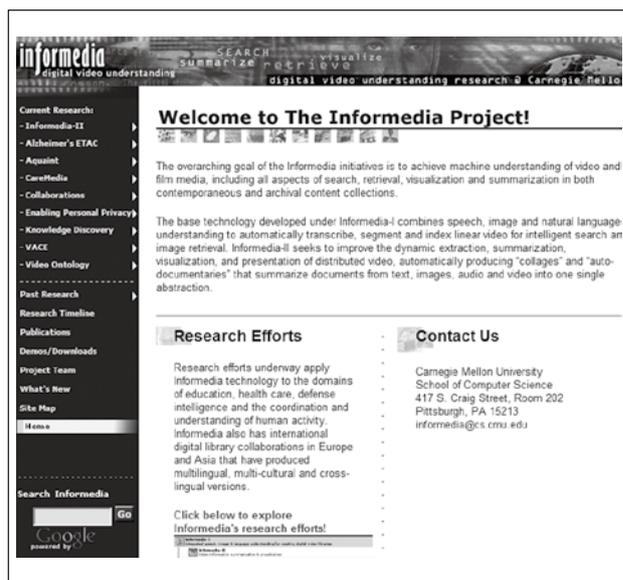


图1 Informedia项目主页

在查询、检索、分类和摘要各方面与文字文献相同的功能。Informedia-II在DLI-II中继续得到400万美元资助, 除DLI-I的三个赞助单位 (NSF、DARPA和NASA) 外, 赞助单位还增加了美国医学图书馆、美国国会图书馆、美国人文科学基金会、美国联邦调查局, 目标是继续在视频信息处理方面进行探索, 增强视频信息提取

的速度和准确性,加强信息的显示能力,有效处理大量视频文献,提高系统的访问效率。

2 资源组织

2.1 资源范围和类型

Informedia系统提供对现在和过去的电视、广播新闻、纪录片和广播的全文查询及检索。自1994年开始,卡内基·梅隆大学的信息媒体项目对来自CNN的新闻视频信息、英国开放大学、QED通信机构(QED Communications)、发现频道(the Discovery Channel)以及美国许多政府机构(如美国宇航局、国家公园服务处National Park Service、美国地质调查局U.S. Geological Survey等)的纪录片及教育录像制品进行数字化。Informedia项目的开始阶段,该系统通过网络向用户提供1000小时的录像资料;2000年,视频数字图书馆已拥有2000小时5000个片段的视频信息,并且继续以每周10小时的速度增长;到2002年,Informedia系统拥有2TB的每日新闻和公众电视、政府机构的纪录片,视频信息已经增加到大约3000小时。

2.2 元数据方案

Informedia系统以全自动化的方式对每天的内容进行捕获和信息抽取,并把它们存储在在线档案中;利用广播电视节目中的脚本信息构造文本索引,通过语音识别器将伴音转换为文本;使用视频分析模块提取人脸特征、叠加在屏幕上的文本和视频镜头边界等,然后将这些信息全部绑定形成最终的索引结构,支持快速多媒体相似检索;同时,允许根据声道、封闭字幕和屏幕上的文本对个别视频段落进行快速检索,支持任意说出的或文字输入的主题词提问;它还能够对相似的脸和图像进行匹配。该系统能将文本和视频信息融合起来,尤其是把底层和中层的视频特征,包括镜头边界、人脸等,与语音识别出来的伴随文本和视频的字符识别(Optical Character Recognition,简称OCR)出来的镜头标题和字幕集成起来使用。

为了提高视频信息的检索效果,Informedia视频数字图书馆对录像制品进行自动编码处理,并从中提取有关信息生成地点附注。在该检索界面中,可以将特定事件与其发生地建立关联,实现从事件、时间到地点的论

题检索,或从地点入手检索视频信息的空间查询方法。在信息媒体图书馆视频信息中,描述性的文本信息是地点附注(Geographic Reference)的主要来源。通过对视频文件中的文本元数据进行信息源处理,就能产生地理编码。Informedia图书馆项目所使用地名词典是从环境系统研究所(Environmental Systems Research Institute, ESRI)的世界地名词典中派生出来的数据子集。该地名数据子集包括全世界约300个国家、州、行政实体以及17000个主要城市。经过处理,最后将文本元数据和视频的播放时间联系起来,并根据与之对应的地理编码簿标出该地的经度和纬度。同时,对文本式的元数据进行语法分析,从中抽取包含地名的句子。通过对来源元数据的上下文分析,系统能够辨别它们的不同含义,确定它是人名还是地名,并按专有名词归类,从而更精确地表示其地理方位。

3 技术特征

3.1 界面设计

卡内基·梅隆大学Informedia数字视频图书馆的主页设计清晰实用(图1),使用户能够在最短的时间内方便快速地找到Informedia项目的具体内容。页面整体可分为三大部分,最上面是页面的logo;下左方是主菜单模块,各个主题菜单依次顺延,可以直接浏览进入Informedia各个子项目以及该项目的出版物、团队和更新情况等,主菜单的下面还提供了谷歌搜索,并设有检索框;下正中模块包括项目的基本内容介绍、研究成果介绍和联系方式,还提供了Informedia的大事年表。这种设计风格让读者易接近,方便读者对该项目进行了解、评价。

3.2 技术分析

Informedia-I把语音、视频、图像与自然语言理解技术相集成,将语音、图像和自然语言理解相互融合,以实现智能搜索和图像检索为目的,对线性视频进行自动转录、分段和索引。采用高质量的语音识别系统把视频录像带上的音频内容转换成文本文件,然后通过语音识别系统的分析和组织,把文件保存到全文信息检索系统中。同样,在视频流的自动切割中,应用了图像识别技术来确定边界。这些工具也同样运用在智能视频监控、导航和选择性检索中。这个过程自动为每个故

事片段生产各种摘要,如标题、幻灯影片的故事概要和视频速览。

在Informedia-II系统中,采用了视频文件集的摘要、多模型查询、自适应总结和基于本体的视频检索(video indexing ontology)等技术,提高了系统的访问效率。将各种音视频信息拼贴在一起是该阶段的关键技术,这一技术是通过加强对信息显示能力的研究来实现的,能够更有效地处理大量视频文献。拼贴技术可以让用户方便查询数字视频图书馆中不同方面的信息,具有可扩展和基于语义缩放的特点。基于本体的视频检索技术利用本体在概念级建模多媒体语义内容,使用构建的概念本体作为多媒体语义内容标注、索引和用户检索概念匹配的统一术语集,以提高多媒体内容标注的有效性和检索的准确性。在Ontology的应用中基于MPEG-7国际标准,解决对多媒体信息描述的标准问题。同时,采用XML可拓展标记语言与Ontology结合应用。Informedia-II系统还可以从语音识别系统和视频流的OCR输出中抽取命名实体,并综合多个特征来获取更好的查询结果。

Informedia数字视频图书馆系统结构原型是在一个商业化的关系数据库系统上建立的,并把所有的分析模块和元数据转移过去。目前的系统使用Informix,但是可以在很多标准的关系数据库系统上运行(如Oracle、Sybase)。为了减少系统迁移时客户端的改变,系统提供了新的应用程序接口,可以分离数据库、查询模块和视频文件,这种结构还提供了函数和数据的复制功能。

4 信息服务

4.1 目标用户

Informedia项目致力于将多媒体的信息应用于教育、信息和娱乐的环境中,主要目标用户是广大的教育工作者,包括小学、中学和大学的学生、老师和教授,使用户能进入、开发、检索利用可视文献,并通过公共学校系统为函授大学设计可视教程,满足用户对录像和视频信息的检索需求。

4.2 检索功能

在视频信息的检索方面,用户最关心的是快速找

到符合自己需求的视频情节或片段。Informedia项目开发的信息可视化界面,可以让用户浏览整个检索结果,而不对结果集进行耗费的重新排序和在结果集内往返移动。同时,在录像片段的收集过程中自动提供文摘,即根据丰富的信息资源自动生成元数据,并将自动生成的视频文摘与揭示人物、事件、机构、论题以及地点和时间的“视频拼贴画”联系起来。这些拼贴画(关键视频帧)是从众多的录像片段中提取出来的信息。

Informedia数字视频检索系统(IDVL)如图2所示,允许用户访问、挖掘和检索海量的数字视频库。例如,检索厄尔尼诺效应时,1997年初它所影响的地区和1998年影响的地区是不同的,在地图上显示的“点”分布的情况也不同。图2展示了IDVL接口显示12个从“El Niño”查询返回的文档以及不同多媒体从这些文档中提取的信息。左上屏幕是检索词输入栏,左下屏幕是查询返回的文档,右侧屏幕是不同多媒体从这些文档中提取的信息。当选中一个视频时,在右下屏幕将播放该视频,并且在其下面,也将显示真正的文本,其中“El Niño”一词会突出显示。关键视频帧汇集了最能反映录像中心内容的画面,同时也可以让用户聚焦于自己感兴趣的某些特殊细节。

Informedia-II项目增强了底层信息提取的速度以



图2 基于内容的Informedia数字视频检索系统

及准确性,还包括对名字、地点、日期和时间引用的解释,并且增加了对动态事件片段、语音以及人脸识别和视频事件特征的识别。概要信息而不是文档本身成为

叙述的单元,如图3所示。视频源可以在这些概要的上下文中看到,以显示事件如何随时间跨越地理边界发展,从而允许从时间和空间视角来展现。

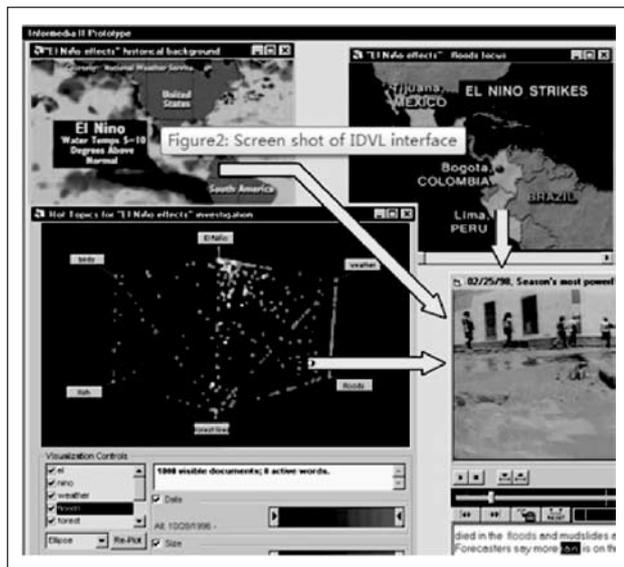


图3 Informedia-II接口对于“El Niño effects”查询提供的附加的视图

5 评价及建议

卡内基梅隆大学的Informedia项目是六个初始NSF/DLI-I项目之一,是其中唯一一个着眼于视频媒体的项目,开创了一个专注于方便用户访问以及检索视频信息的时代。自1994年至今,该项目一直在深入地开发数字视频及其相关的技术和工具。Informedia数字视频图书馆使用户能够从视频数字图书馆里接近、探索和检索科学资料,已成为数字图书馆建设中一个比较成功的范例。其优势主要体现在:一、资金和设备基础方面,Informedia项目起步早,获得机构的资助合作较多,对信息资源采集整理有较好的设备和技术支持;二、资源建设方面,以视频、音频多媒体资源建设为重点,资源优势独特,满足教育、信息和娱乐用户对录像和视频的检索需求;三、检索方面,不仅赋予视频检索具有文本检索中的所有功能,自动进行内容获取,信息提取以及在线存档,还增强了视频信息提取的速度以及准确性,提高了系统的访问效率;四、服务方面,Informedia系统的视频音频自动标引、导航、可视化、搜索与检索,以及在此基础上的多模型查询、视频

的自动提供摘要等功能给用户提供了莫大的便利,是很有针对性的特色服务。视频信息除了用日期、时间、论题检索外,还可以通过定位等方式进行检索。文字检索的结果可显示在一幅地图上,并选择自己喜欢的方式进行揭示,这些都改进了信息的展现方式。授权的用户可以按时间的模式对视频图书馆中的所有元数据进行管理。

虽然Informedia数字视频图书馆有很多优势,但在发展建设中也有一些应注意的问题:一、文本、视频、音频三者融合的方法适用于带有非常清晰的语音、文字的视频,如教育片、纪录片。但是对于包括复杂类型声音的视频,如故事片,这种文字或关键词的方法就可能难以达到令人满意的效果。二、静态视频摘要和全景拼接图方式实现的方法相对比较简单,在一定程度上能够满足人们浏览信息的需要。但是静态视频摘要和全景拼接作浏览时把视频中的很多东西都丢失了,造成浏览信息的片面。三、基于文字的视频摘要方式,由于需要有清晰的人工文字,且OCR系统也并非十全十美,因而限制了摘要的应用场所。动态视频摘要和多媒体信息摘要可以满足人们日益增长的需求,但是实现起来比较困难。四、Informedia系统还未对公众开放,需要授权。五、页面上的站内Google搜索链接暂时不能访问。需要Informedia进一步完善,以进一步促进其数字图书馆的优化服务。

Informedia项目基于内容的多媒体信息检索是一种新兴的网络信息处理和检索技术,它包含了自然语音处理、图像处理、视频处理、语音识别、数据库与数据挖掘、模式识别、人工智能、机器学习等众多学科,是一个长期的、复杂的、富有挑战性的研究课题。

参考文献

- [1] 朱强.数字图书馆:21世纪图书馆的原型——美国“数字图书馆创始”计划简介[J].大学图书馆学报,1995(4):50-54.
- [2] 沈正华.数字视频图书馆的交互式地图[J].现代图书情报技术,2004(2):4-9.
- [3] 苏苏宁.信息检索理论与技术[M].科学技术文献出版社,2004.
- [4] WACTLAR H D. Multi-Document Summarization and Visualization in the Informedia Digital Video Library [C]// Invited speaker at New Information Technology 2001 Conference, Tsinghua University, Beijing, May 29-31, 2001.
- [5] WACTLAR H D. Exploring Cultural Origins and Evolution

- by Navigating Searchable Video and Film Archives [C]// International Conference on Digital Libraries, IT Opportunities and Challenges in the New Millenium (invited speaker), Beijing, China, July 8-12, 2002.
- [6] Informedia Digital Video Library [EB/OL]. [2014-01-10]. <http://www.informedia.cs.cmu.edu/>.
- [7] Research Informedia-I [EB/OL]. [2014-1-10]. <http://www.informedia.cs.cmu.edu/dli1/index.html>.
- [8] ResearchInformedia-II [EB/OL]. [2014-01-10]. <http://www.informedia.cs.cmu.edu/dli2/index.html>.
- [9] ResearchOntology [EB/OL]. [2014-01-10]. <http://www.informedia.cs.cmu.edu/ontology/index.html>.
- [10] Research Timeline [EB/OL]. [2014-01-10]. <http://www.informedia.cs.cmu.edu/timeline/index.html>.

作者简介

刘燕权，男，博士，毕业于美国麦迪逊大学信息图书馆学院，现任南康涅狄格州立大学教授。近年的出版物主要集中在数字图书馆、数据挖掘、多媒体技术、软件管理、国家信息基础结构、信息存储、图书馆统计及管理等方面。E-mail: liuscsu@gmail.com

孙秀良，男，博士，北京化工大学图书馆副研究员。E-mail: sunxl@buct.edu.cn

The INFORMEDIA Digital Video Library——The Forerunner of Multimedia Digital Libraries

Yan Quan Liu / Southern Connecticut State University, New Haven, CT, USA, 06515
Sun Xiuliang / Beijing University of Chemical Technology Library, Beijing, 100029

Abstract: Carnegie Mellon University's INFORMEDIA Digital Video Library is one of six original Digital Libraries Initiative projects in USA since 1994. Its overarching goal is to enable users to achieve machine understanding of video and film media, including all aspects of search, retrieval, visualization and summarization in both contemporaneous and archival content collections. This review examines various aspects of the library, including the project's background, organization of resources, technologies, and service features. Authors' comments are also given.

Keywords: Digital library, Carnegie Mellon University, Digital Libraries Initiative (DLI), INFORMEDIA, Digital video, National Science Foundation (NSF)

(收稿日期: 2014-02-08)