

●陈传义 余传明

因特网声像信息资源及利用

摘 要 介绍了因特网网上声像信息的类型,并对搜索方法和音频、视频播放工具进行探讨。参考文献 5。

关键词 因特网 声像信息 音频视频 检索

分类号 G255.73

ABSTRACT The authors classify A/V information on the Internet, and discuss search tools and audio-visual playing tools 5 refs

KEY WORDS Internet A/V Information Retrieval

CLASS NUMBER G255.73

近年来,全球互联网络以惊人的速度发展,因特网上声像信息成级数增长。我国远程通讯设施和技术不断完善,为网络多媒体利用创造了先决条件。声像信息资源及其类型如何,在浩瀚的信息海洋中又如何快速查找和正确播放这类媒体?这是信息用户网上搜寻时常遇到的问题。本文正是针对这些问题展开一些初步探讨,并希望借此能引出更深入的研究。

1 因特网声像信息资源定义及类型

从因特网信息资源所包含的媒体形式来看,声像信息资源可分为以下几类:文本信息,这是最为基本的一种媒体存储形式;图片信息,指以 GIF (Graphics Interchange Format), JPEG (Joint Photograph Experts Group) 等文件格式存储的信息;音频信息,主要指以 WAV (wave), A IFF (Audio Interchange File Format), M I D I (Musical Instrument Digital Interface), MP3 (MPEG-1 Layer3) 等文件格式存储的信息;视频信息,主要指以 QuickTime, AV I (Audio Video Interleave), 以及 MPEG (Moving Picture Expert Group) 等形式存储的信息;三维虚拟影像,信息模型是以 VRML (Virtual Reality Modeling Language) 组织,以立体三维形式呈现的信息。

声像信息资源,有人也称之为多媒体信息资源,以笔者之见,这两者小有区别。前者是指以声像为主要媒体形式的信息,它并不排除少量必要的文字说明,较之单纯的声画表达的效果更胜一筹。然而,声像信息是在网上用超文本标记语言(HTML)安插(Plug-in)的,因此它具有交互性和超媒体链接能力,而这个特性也正是多

媒体信息具备的属性,它们在本质上是一致的,只是声像信息媒体这一概念强调了声像在信息呈现窗口的主导形式以及利用它们的特殊性。由此,我们认为声像信息媒体与多媒体信息是有区别的,进而认为,它们的信息资源的存在也略有区分。

从因特网声像信息资源的使用角度来看,可将其分为以下几类:一是声像数据库;二是多媒体辅助教学;三是声像电子出版物,与传统的电子出版物相比,文本已不再是出版物的全部,图像、音频、视频成为出版物的重要组成部分;四是多媒体公众服务系统,它的内容相当广泛,包括网络多媒体广告、多媒体旅游信息系统、多媒体商品导购信息系统、网络会议(Netmeeting)、网络广播(Webbroadcasting)、网络电视(WebTV)、视频点播(VOD)等等。

2 因特网声像信息搜寻

随着因特网万维网的超媒体连接方式的普遍利用,网络声像信息媒体的优点得到了充分展现。由于多数万维网站点 24 小时开放,全球范围均可访问,而且声像信息丰富多彩,易于理解,便于人机交互,因而,因特网尤其是万维网中声像信息的查询和搜索显得日益重要。现在,人们主要借助以下一些方法:

(1) 借助于资源指南。有关因特网的资源指南相当多,详尽地介绍了资源类目、网址大全、参考信息数据库等。它们往往以印刷品形式出版或录入 CD-ROM 发行,有的则直接在网上发布。它简便易行,但实效性较差,完备性不高。

(2) 借助于搜索引擎。搜索引擎是收集、整理网上

信息资源并按一定规则加以整理和组织, 提供人们按相应的规则提取信息线索, 并能直接链接到相关站点的网上信息搜索工具。可用于查找声像信息的搜索引擎包括 Hotot, Lycos 等。这些搜索引擎通常都是提供两种方法检索相关站点及声像信息: 一种是按主题范畴检索, 即按类目查找; 另一种是关键词查找。

按主题范畴检索, 方式比较简单, 搜索引擎的首页一般都有一个主题范畴表, 有的类目甚至还分为几个层次, 按照这些层次从上级类目至下级类目依次选择即可。不过它的筛选量过大, 为了获得较准确和较完整的信息, 还必须与关键词检索结合起来。

所有搜索引擎都提供关键词检索功能, 可以直接进入搜索引擎的首页输入关键词(包括布尔逻辑、位置逻辑以及特定字段的限定等)进行检索, 也可以在进入首页以后, 按照主题范畴进入指定的类目后输入关键词进行查找。如果要查找图片信息, 则直接进入 hotbot (<http://www.hotbot.com>), 将“网页必须包括(Page Must Included)”中的“图片选项(image)”做上标记, 再输入关键词就得到了含有所需图片的站点, 按址访问即可。还可以在进入指定类目后, 再进行关键词查找。以 altavista 为例, 先进入站点(<http://www.altavista.com>), 然后选择“特殊搜索(Special Searches)”之中的“AV Photo Finder”, 进入“AV Photo Finder”页面之后, 输入相应的关键词进行查找, 便能得到相关的图片信息。与检索文本信息资源相类似, 检索到相关的声像信息之后, 可根据检索结果调整检索式, 用“and”、“or”以及“not”来限制检索条件; 用“near”来限制检索词的逻辑位置, 用“or”放宽检索条件等等。

利用搜索引擎可以迅速查找到相关网址, 再利用这些网址进入相关站点, 就很容易得到相关声像信息。不过, 由于这些搜索引擎都只是附带的一些声像信息搜索功能, 建立的数据库也仅仅是利用排除了词频特高的词之后的关键词建立倒排文档, 它们均不是叙词、标题词等受控词, 因而查全率和查准率都不可能很高。由此对检索式的表达提出了较高的要求。

(3) 直接利用声像信息搜索系统。图像、音频、视频信息日益丰富, 专门针对这类信息的检索系统也应运而生。例如美国的 scour 网(<http://www.scour.net>)就是。用户查询时首先在检索媒体(Image, Audio, Video)中选择一种, 然后再按关键词查找或者按类目查找。与搜索引擎不同的是前者得到的是相关网址, 而后者直接就可以得到所需的声像信息和相关网址, 进入相关网址便能了解到更多的信息。利用该站点能很方便地得到有关的声像信息, 但查准率同样很低, 大量的娱乐性信息混杂在检索结果之中。这一点在检索过程中要尤其注意。哥

伦比亚大学开发的 Webseek 系统(<http://disney.columbia.edu/webseek>), 也提供了按关键词和类查找的两种途径, 按此搜寻可以查找到声像信息。

(4) 其他搜索方法。例如可以按照颜色、形状、纹理在图像中的位置来对图像进行索引和查找, 用颜色分布来索引和查询纺织品的图案, 用边缘轮廓结合其他图像处理技术来索引和查询人脸照片等。哥伦比亚大学电力工程系与远程通讯研究中心研制的 CBQS 系统(Content Based Image Query System) (<http://disney.columbia.edu/safe>), 就是这样一种与文本查询完全不同的图像特征查询系统。不过, 目前这些方法还不够成熟, 大多处于试用研究阶段, 仍有待于进一步开发。

不管借助何种查询工具、以哪种查询方式, 用户进行声像信息搜索时应当注意时刻警觉自己的查找目标, 因为因特网是一个集科教、文化广告、娱乐、休闲于一体的庞大系统, 其中五花八门、将人诱入歧途的画面很多, 一旦深入, 待你幡然醒悟已是失之千里。如果在检索过程中不明确紧扣检索主题, 盲目激活一些无用的超文本链接, 进入不相关站点, 最终可能迷失在信息流之中。

3 声像信息的下载及播放

(1) 对于传统文本信息而言, 其传递与下载需要时间很短, 进行浏览也只需借助一些简单的文字转换系统即可。而声像信息所含的数据量相当大, 并且有着特殊的存储格式, 不仅传输需要一定的时间, 浏览播放也得借助于特定的播放工具才能完成。因此, 有必要研究音频及视频播放的有关技术。

(2) 影响网上声像媒体播放的有两个重要因素: 数据率与带宽。数据率, 即数据传输速率, 是单位时间内传送给用户的信息量的量度, 一般情况下, 以每秒千字节来衡量(Byteps)。带宽(Band Wide), 是表示单位时间内通过一根电缆或光缆能传送的数据量, 一般情况下, 以每秒千比特来衡量(Kbps)。对于单独用户而言, 实际所用的带宽取决于网络负载。

WEB 的带宽相当于 CD-ROM 的 1/10~1/100。正是由于带宽的限制带来缓慢的数据传输, 使得声像媒体的播放变得断续乃至停顿。动态图像更是如此。因此不得不对声像信息数据进行大幅度地无损乃至有损压缩。由于压缩软件工具和压缩方式不一样, 就形成了多种文件格式。为了在网络终端正确再现声音图像, 则应采用对应的解压软件(播放工具)进行播放, 以此来解决数据传输的问题。

(3) 音频格式与视频格式。

MIDI 格式, 一般用于存储音乐文件, 以高度压缩形

式存储,并可用于MIDI设备播放。

WAV 格式,基于Windows平台的多数音频剪辑都是以WAV格式保存,其压缩比率并不大,但易于制作,因而也有着较广的应用。

AIFF 格式,主要用Apple Macintosh平台上,信息以8位或16位的格式存储。

MP3 格式,是MPEG-1 Layer3的缩写,是按照MPEG-1算法对声音进行压缩的第3层,是在保证CD音质的条件下,将声音压缩到原来大小的1/12,它只能借助于专门的MP3播放器播放。

网络中常用的视频格式包括:

AVI 格式,主要是指对视频信息的一种压缩格式,其主要方法是减少帧频。正常的视频信息播放时的帧频大约是30帧/s,在AVI格式下则缩小至15帧/s甚至更小,除此之外,每帧图像的像素也减少了,使得视觉效果受到一些影响。

MPEG 系列格式,也是对视频信息的一种压缩格式,但比AVI更为有效,它使用了帧差异(frame differencing)等压缩技术,使得在不必保存每个帧的同时,能在视频播放器上解压后产生质量较好的运动视频图像。

QuickTime 格式,是一种成熟的、标准的、跨平台多媒体文件格式,具有标准的播放界面以及压缩—解压特性。一个QuickTime文件可以包括32个单独视频轨道和同步声音信息。

还有ActiveX Streaming 格式等,主要用于存放ASF或ASX为扩展名的文件,播放ASF文件需要使用Netshow 播放器。

(4) 声像信息播放的类型:非实时播放与实时播放。非实时播放是多数声像信息媒体播放器所采取的形式,使用这种方式,用户必须整个音频或视频文件全部下载之后,才能进行播放。随着流处理技术(Streaming)的发展,与非实时播放相对应的实时播放也开始应用。采用这种方式,播放器接收到部分视频信息之后就可以同步播放,但这对于用户的因特网连接状况(包括提供的数据率与带宽等)以及计算机硬件的速度都提出了较高的标准,要求数据传输速率以及硬件能与播放的速度相适应。对于多数用户而言,即使是采用第一种播放方式,由于受传输线路的限制,往往也要等待一段较长时间,因而要实现实时播放,还具有一定的难度。

4 两种浏览器及其播放多媒体的工具

微软公司的探索者IE(Internet Explore)是浏览多媒体信息的有效工具之一,它提供两种类型的播放器ActiveMovie 与Netshow。ActiveMovie 用来播放WAV

格式的音频,以及MPEG,QuickTime 和AVI 格式视频信息。Netshow 不仅能支持纯音频文件,还可用来播放ASF 和ASX 文件。

网景公司的导航者也是浏览多媒体的有效工具。NC 提供Live audio 与Live video,当用户激活以超文本方式链接的声音和影像时,Live video 能自动播放以AVI, SND, WAV, MIDI 格式存储的音频文件,以及以AVI, QuickTime 格式存储的视频文件。

除此以外,还有用来播放各种流格式的音频与视频的媒体播放器Realplayer(可从<http://www.real.com/products/player/index.html>处下载),VDOlive(可从<http://www.vdolive.com/download>下载);播放MPEG 格式视频媒体的MPEGPlay,播放MP3 格式音频信息的Winamp, Winplay3, Jetaudio, Unreal Player 等。

因特网是一个庞大的网络信息系统,在这个系统中,多媒体所占的比重越来越大。各种更好的声像信息格式以及相应的播放器也不断推出,用户只有经常搜索和浏览,并及时获取这些信息和工具加以利用,才能使声像信息完美的表达效果充分呈现出来。

参考文献

- 1 Gene Weiskopf. Internet Explorer4 从入门到精通, 邱仲潘等译. 北京: 电子工业出版社, 1998
- 2 David Miller. Web 多媒体开发指南. 京京翻译组译. 北京: 机械工业出版社, 1998
- 3 Phil Jame. Navigation Communicator4 使用指南. 甘特工作室译. 北京: 机械工业出版社, 1998
- 4 张明敏. 网络多媒体技术与使用. 北京: 清华大学出版社, 1998
- 5 陈光祚等. 计算机情报检索系统导论. 北京: 书目文献出版社, 1993

陈传义 武汉大学传播与信息学院讲师。通讯地址: 湖北武汉。邮编 430072。

余传明 武汉大学传播与信息学院 99 级研究生。通讯地址同上。

(来稿时间: 1999-07-13。编发者: 刘喜申)