

INTERNET CHAZHAO TUXIANG

怎样在Internet上

查找图像

社会广角类 商务与金融类 社会广角类 商务与金融类 社会广角类 商务与金融类
自然类 动物类 人物类 社会生活类

储荷婷 敖毅 著

宇宙与天体类 工业类

人物类 社会生活类

旅游渡假类 运动与医疗保健类

科技类 建筑类 世界名胜类

社会广角类 商务与金融类

抽象设计类 艺术类

自然类 动物类

农业类 交通与运输类

宇宙与天体类 工业类

类 静物素材 运动类

科技类 建筑类 世界名胜类

人物类 社会生活类

人物类 社会生活类

运动与医疗保健类

旅游渡假类 运动与医疗保健类

商务与金融类

社会广角类 商务与金融类

自然类 动物类

自然类 动物类

宇宙与天体类 工业类

宇宙与天体类 工业类

建筑类 世界名胜类

科技类 建筑类 世界名胜类

抽象设计类 艺术类

社会广角类 商务与金融类

农业类 交通与运输类

宇宙与天体类 工业类

环保类 静物素材 运动类

抽象设计类 艺术类

人物类 社会生活类

农业类 交通与运输类

旅游渡假类 运动与医疗保健类

金融类 环保类 静物素材

社会广角类 商务与金融类 社会广角类 商务与金融类

科技类 建筑类 世界名胜类

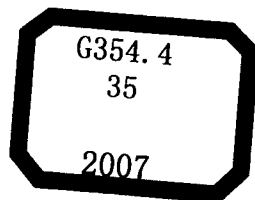
自然类 动物类 自然类 动物类 自然类 动物类

宇宙与天体类 工业类 宇宙与天体类 工业类



辽宁科学技术出版社

LIAONING SCIENCE AND TECHNOLOGY PUBLISHING HOUSE



怎样在 Internet 上 查找图像

储荷婷 敖 毅 著

 辽宁科学技术出版社

沈 阳

图书在版编目 (CIP) 数据

怎样在Internet上查找图像/储荷婷, 敖毅著. —沈阳: 辽宁科学技术出版社, 2007.1

ISBN 978-7-5381-4915-9

I. 怎... II. ①储... ②敖... III. 因特网—基本知识 IV. TP393.4

中国版本图书馆CIP数据核字 (2007) 第148978号

出版发行: 辽宁科学技术出版社

(地址: 沈阳市和平区十一纬路25号 邮编: 110003)

印刷者: 沈阳市北陵印刷厂

经销者: 各地新华书店

幅面尺寸: 175mm × 234mm

印 张: 16

字 数: 300千字

印 数: 1 ~ 3000

出版时间: 2007年1月第1版

印刷时间: 2007年1月第1次印刷

责任编辑: 杨廷郊

封面设计: 刘冰宇

版式设计: 胡 妍

责任校对: 王晓秋

定价: 38.00 元

联系电话: 024-23284360

邮购热线: 024-23284502

E-mail: lkzsb@mail.lnpgc.com.cn

<http://www.lnkj.com.cn>

作者简介

储荷婷 于1982年获北京大学图书馆学系文学学士学位,后赴加拿大和美国,分别攻读图书馆信息学专业的硕士和博士学位。求学生涯结束之后,开始走上美国大学讲台,教授图书馆信息学专业。现在,是美国长岛大学(Long Island University)图书馆信息学学院教授,主要从事信息检索,特别是网络信息检索方面的教学和研究。她先后就信息检索、科学计量学以及信息技术在图书馆信息管理领域的应用等课题发表了几十篇论文。同时,还出版了《当代最新信息技术:Internet使用指南》《Internet网络信息检索——原理 工具 技巧》《Information representation and retrieval in the digital age》三本专著。

敖毅 男,1976年9月出生,四川乐山人。现在南京政治学院上海分院信息管理系攻读博士学位,主要研究方向是信息组织与检索。曾参加过军内外多项大型项目的研究开发,并在国内期刊上发表十余篇相关论文。目前正进行语义网和内容管理方面的理论与技术研究。

前言

本书旨在向读者介绍怎样利用互联网上的各种检索系统，查找可以免费获取的图像。全书共分十章，其中，第一章主要介绍图像的基本种类和相关概念，第二章着重概述检索图像信息的途径和方法。第三章到第九章，分门别类地对五十余个图像信息检索系统的使用进行了描述和解释。第十章讨论了图像检索和使用的版权等问題，同时也展望了网络图像检索的未来。

本书涉及的图像检索系统大多数源自美国，主要是对美国在网络图像检索领域中占主导地位这一现实的反映。随着时间的推移，发展迅速的信息检索领域的图像检索系统也必将发生显著的变化。因此，尽管所有在本书中引用或提及的网站都于2005年7月至8月间经过检查核对，但本书介绍的内容还会出现过时或与您所见所用不同的情况。应对这些变化的最佳方法是经常关注和使用网络图像检索系统。

另外，有些网络服务器为了防止垃圾邮件或病毒等的攻击，对特定的网站进行屏蔽，无意中可能会封堵本书介绍的图像检索系统。如若出现这种状况，您不妨换个网络服务器，然后重新访问。

在本书付梓之际，笔者要特别感谢辽宁科学技术出版社编审杨廷郊老师在选题、定题以及写作过程中给予的鼓励、支持和帮助；武汉大学信息管理学院前任院长、教授和博士生导师马费城老师也在百忙之中为促成本书的写作出谋划策，笔者对此深表谢意。另外，笔者还感谢美国长岛大学 Palmer 图书馆情报学院博士生 Selenay Aytac 在搜集资料数据方面提供的帮助。

在信息技术、特别是网络技术迅速发展的今天，图像信息量激增，人们对检索图像信息的需求也日益增加。因此，尽管我们在本书中讨论了五十多个图像信息检索系统，但还是有许多系统（如<http://www.flickr.com>）未能提及。而且，可以肯定，今后还会有更多、更先进的图像检索系统出现。

最后，希望您在掌握 Internet 图像信息检索的方法和技巧上，能灵活地将它们运用于不同的系统和环境，以达到图像信息检索的目的。

储荷婷

美国纽约长岛

2006年5月

目 录

第一章 图像信息检索概述

- 一、Internet 与图像信息·····001
- 二、图像的基本种类·····001
- 三、关于图像的几个基本概念·····003
- 四、图像类型的选择·····004

第二章 图像信息检索的主要途径和方法

- 一、图像信息检索的三种途径·····005
 - 1. 图像搜索引擎·····005
 - 2. 图像目录·····011
 - 3. 图像专类收藏·····013
 - 4. 三种图像检索途径的比较和归纳·····013
- 二、检索图像信息的两种方法·····014
 - 1. 基于描述的图像检索方法·····014
 - 2. 基于内容的图像检索方法·····014
- 三、各种检索途径和方法的小结·····016

第三章 利用综合型检索系统检索图像

- 一、AltaVista Image Search (AltaVista 图像搜索)
<http://www.altavista.com/image/>·····018
- 二、Ditto
<http://www.ditto.com>·····020
 - 1. Ditto 的图像检索·····021
 - 2. Ditto 的常见专题图像检索·····023
 - 3. Ditto 的图像目录·····023
- 三、Google Image Search (Google 图像搜索)
<http://images.google.com>·····025
- 四、Picsearch
<http://www.picsearch.com>·····028

五、WebCrawler Images	
http://webcrawler.com	031
六、Yahoo! Image Search	
http://search.yahoo.com/images	034

第四章 利用专类图像检索系统检索图像 (一)

——艺术、医学、地理环境及建筑类图像的检索

一、艺术图像	039
1. The Art Museum Image Consortium (AMICO, 艺术博物馆图像联合)	
http://www.amico.org	039
2. NYPL Picture Collection Online (PCO, 图像收藏联网)	
http://digital.nypl.org/mmpco/	044
3. Artcyclopedia (艺术百科)	
http://www.artcyclopedia.com	049
4. 其他艺术图像收藏	058
二、医学图像	058
1. University of Texas Health Science Center at San Antonio (UTHSCSA)	
Medical Image Databases on the Internet (得克萨斯大学医学科学中心	
Internet 医学图像数据库)	
http://www.library.uthscsa.edu/internet/ImageDatabases.cfm	058
2. Duke University Medical Center Library	
Medical Images (杜克大学医学中心图书馆医学图像)	
http://www.mcilibrary.duke.edu/subject /medimages	061
三、地理环境图像	064
1. TerraServer	
http://terraserver-usa.com	064
2. National Environment Satellite, Data, and Information Service (NESDIS,	
美国环境卫星、数据和信息服务)	
http://www.nesdis.noaa.gov	070
四、建筑图像	082
1. Digital Archive of American Architecture (美国建筑图像数字专藏)	
http://www.bc.edu/bc_org/avp/cas/fnart/fa267/	082
2. Rochester & Monroe County (NY) Library System Rochester Images——	
Architecture Pathfinders (美国纽约 Rochester 和 Monroe 郡图书馆系统——	
Rochester 图像专藏建筑部分)	
http://www.rochester.lib.ny.us/rochimag/architecture/ArchIntro.htm	087
3. Wikipedia——List of Images/Architecture	
(Wiki 百科全书——建筑部分)	
http://en.wikipedia.org/wiki/Wikipedia:List_of_images/Architecture	089

第五章 利用专类图像检索系统检索图像 (二)

——动物、植物及图形和标志、漫画类图像的检索

一、动、植物图像	091
1. US Fish and Wildlife Service Digital Library Systems (美国鱼类和野生动物机构数字图书馆系统)	
http://images.fws.gov	091
2. National Biological Information Infrastructure Digital Image Library (美国生物信息基础数字图像图书馆)	
http://images.nbi.gov	095
3. Burpee	
http://www.burpee.com	097
二、图形、标志图像	101
1. Free Graphics Store (免费图形收藏)	
http://ausmall.com.au/freegraf/freegrafa.htm	101
2. Bells & Whistles	
http://www.bellsnwhistles.com	103
3. Airline Logos (航空公司标志)	
http://www.aerosite.net/logos.htm	105
4. Google Holiday logos (Google 节日标志)	
http://www.logoogle.com/holidaylogos.html	107
三、照片图像	109
1. FreeFoto	
http://www.freefoto.com/index.jsp	109
2. Webshots	
http://www.webshots.com/homepage.html	112
四、漫画、剪贴画图像	117
1. Illingworth Cartoons (Illingworth 漫画专藏)	
http://www.llgc.org.uk/illingworth/index_s.htm	117
2. Hirschfeld Archive (Hirschfeld 漫画选藏)	
http://theater.nytimes.com/ref/theater/hirschfeld/index.html?rf=index.html	122
3. Clipart Connection (剪贴画联合)	
http://www.clipartconnection.com	124

第六章 利用图像专藏检索图像 (一)

一、Library of Congress Prints & Photographs Division (美国国会图书馆图片图像部)	
http://www.loc.gov/rr/print/	128
1. Prints & Photographs Online Catalog (PPOC, 图片图像联机目录)	
http://lcweb2.loc.gov/pp/pphome.html	129

2. List of Images on Popular Topics (常见专题图像列表)	
http://www.loc.gov/rr/print/list/listguid.html	133
二、Smithsonian Photographic Services (Smithsonian 图像部)	
http://photos.si.edu	136
三、Seattle Municipal Archives Photograph Collection (西雅图市政府档案馆照片专藏)	
http://clerk.ci.seattle.wa.us/~public/phot1.htm	146
四、Hawaii Image Bank (夏威夷图像库)	
http://www.hvcb.org/imagelibrary/Standard/index.jsp	148
五、Newsimages.com	
http://www.newsimages.com	155
六、Getty Images	
http://creative.gettyimages.com/source/home/home.aspx	158
1. Creative (非新闻图像)	159
2. Editorial (新闻图像)	164
第七章 利用图像专藏检索图像 (二)	
一、Astronomical Picture of the Day (APOD, 每日天文图像)	
http://antwrp.gsfc.nasa.gov/apod/archivepix.html	169
二、London's Transportation Museum Photographic Collection (伦敦交通博物馆图片收藏)	
http://photos.ltmcollection.org	174
三、Boeing Photo Store (波音照片商店)	
http://www.boeingphotostore.com/source/Home.aspx	180
四、Scale Models (环球模型网)	
http://scalemodel.net	185
五、小型专藏	190
1. Museum of Bad Art (劣质艺术博物馆)	
http://www.glyphs.com/moba/	190
2. Toothpaste World (牙膏世界)	
http://www.toothpasteworld.com	191
第八章 现存照片图像的检索	
一、FREEIMAGES.co.uk	
http://www.freeimages.co.uk	194
二、FreeStockPhotos.com	
http://freestockphotos.com	199

三、Geek Philosopher Free Stock Photos (Geek Philosopher 免费现存照片)	
http://geekphilosopher.com/MainPage/photos.htm	202
四、Bigfoto	
http://www.bigfoto.com	206
五、Stock.xchng (现有照片交换)	
http://www.sxc.hu	209
六、Barry' s free photos	
http://www.barrysfreephotos.com	213

第九章 图像检索的非常规方法

一、基于内容的图像检索方法	216
1. Visual Retrieval Ware (可视检索系统)	
http://vrw.convera.com:8015/cst	216
2. PicToSeek	
http://zomax.wins.uva.nl:5345/ret_user/	220
3. Ikona	
http://www-rocq.inria.fr/cgi-bin/imedia/ikona	222
4. PictureFinder	
http://134.102.206.80:8080/demo/frameset.jsp?lang=en	227
二、图像的混合检索法	232
1. WebSEEk	
http://persia.ee.columbia.edu:8008/	232
2. Berkeley Project (伯克利项目)	
http://elib.cs.berkeley.edu/kobus/famsf/model_2/text_and_blobs/bbox.html	240

第十章 图像信息检索的相关问题和未来

一、检得图像的版权和合理使用	243
二、图像检索系统的登记注册	244
三、免费图像的质量问题	245
四、图像信息检索的展望	245

第一章 图像信息检索概述

一、Internet 与图像信息

自从 Internet 在 20 世纪 90 年代面世以来,愈来愈多的平民百姓成为网民,利用网络进行交流和查询各类信息。据 Nielsen//NetRatings (2001) 对分布于各大洲的 27 个国家的调查统计,全世界的网民数在 2001 年的第一季度已达到 4.29 亿。该机构于 2004 年的调查结果表明,仅美国就有 2.04 亿多人具备上网的条件,接近美国人口总数的 75% (Nielsen//NetRatings, 2004)。同样,根据 Nuttall (2004) 的报道,中国的网民数已达 8000 万之多,且有快速增加的趋势。

当然,网民数量的持续和快速增长与 Internet 自身的功用和所含信息密切相连。关于 Internet 的功能和用途,已有许多文章、书籍进行了专门论述。虽然 Internet 所含信息数量由于其覆盖面宽、变化快而难以精确计算,但据 IBM 所建立的 WebFountain (万维网喷泉——<http://www.almaden.ibm.com/webfountain/>) 统计,万维网每日以新添加或被更新 5000 万网页的速度递增 (Cass, 2004)。深受网民们喜爱的搜索引擎 Google 宣称,目前它收录的信息记录已超过 80 亿条 (Google, 2005)。由此,我们可对 Internet 载运的信息量之巨大窥见一斑。

Internet 的其他组成部分 (如电子邮件和 FTP——File Transfer Protocol) 也传送大量信息,然而,上述提及的 WebFountain 和 Google 所收录的内容均以万维网信息为主,并且这些信息中又包括大量的图像。按使用率衡量,万维网位于电子邮件之后,居所有 Internet 应用工具的第二位。但是,万维网的特点之一是它装载有大量的图像信息,以万维网作为载体的图像信息,不仅传送快捷,而且获取和浏览欣赏方便。

对于 Internet,尤其是万维网上这浩如烟海、丰富多彩的图像信息,我们如何进行查找搜寻?我们应采用什么方法、途径检索 Internet 图像信息?在 Internet 图像信息检索方面,我们应注意哪些问题?本书将对这些问题一一进行专门介绍和讨论。

二、图像的基本种类

图像的种类包罗万象,大致可分为静态图像 (still images) 和动态图像 (moving images) 两大类。其中静态图像包括绘画、图片、照片、招贴画等;动态图像则泛指

电影 (movie)、录像 (video)、动画 (animation) 等含有图像的多媒体。动态图像还可以进一步细分为无声图像 (如无声电影) 和有声图像 (如有声动画), 后者实质上由声音 (sound, 包括讲话、音乐和其他声音) 和图像两种信息组成。根据计划, 本书专门介绍讨论如何查找 Internet 上的静态图像, 而不涉及动态图像的检索。为了简略起见, 我们在书中以图像代替静态图像一词。

Internet 运载的图像根据制作其的不同软件标准而类分为不同类型的图像文件。从 Internet 获取的最常见的图像文件类型有 gif 和 jpeg 两种, 其中, gif 是 Graphic Interchange Format 的首字母缩写, 它由 Compuserve Information Service 在 20 世纪 80 年代研制开发; jpeg 代表 Joint Photographic Experts Group, 它问世的时间晚于 gif, jpeg 也常被缩写为 jpg。

诸如 gif 和 jpg 这些图像种类的简称也充当指明图像文件种类的扩展名, 如 greatwall.gif 中的 gif 或 YellowRiver.jpg 中的 jpg。文件扩展名指文件名后小圆点后面的字母, 长度一般为 3 个字母, 但近年来也常见 4 个字母的文件扩展名, 如 jpeg。通过图像的文件扩展名, 我们可以获知检得图像的种类型。

除去 gif 和 jpeg 图像之外, 在万维网上常见的图像类型还有 png (Portable Network Graphics), 但不是任何一种万维网浏览器 (browser) 都能显示 png 图像, 这在一定程度上限制了 png 图像的广泛采用。表 1.1 列出了其他图像类别以及相应的文件扩展名。如

表 1.1 图像类别一览表

图像扩展名	图像类别简介
bmp	Bitmap, 适用于 Windows 软件
cgm	Computer Graphics Metafile 适用于 Canvas 和 ClarisDraw
cvs	适用于 Canvas 软件
pbm, ppm, pgm	Portable Bitmap, Portable Pixmap, and Portable Graymap, 适用于 Unix 软件和 Adobe 的 PhotoShop
pcx	Pictures, 适用于 Windows 软件
pdf	Portal Document Format, 适用于 Adobe 的 Acrobat
pict	Pictures, 适用于 Macintosh 软件
ps, eps, epsi	Postscript, Encapsulated Postscript, and Encapsulated Postscript with preview image, 适用于 Postscript 打印机
psd	适用于 Adobe 的 Photoshop
tga	Targa, 适用于 Adobe 的 Photoshop
tiff, tif	Tagged Image File Format, 适用于 Macintosh 软件
wmf	Windows MetaFiles, 适用于 Canvas 软件
wpg	WordPerfect Graphics, 适用于 Lview 软件
xpm, xbm	X PixMap, X bitmap for, 适用于 X Windows 软件

果您想了解更多的图像种类, 请参阅 Alison Zhang (1996) 书中的附录, 该附录全文刊载在 <http://www.is-edu.hcmuns.edu.vn/WebLib/Books/Internet/multimed/appen.htm>。

如前所述, 由 Internet 运载显示的图像以 gif 和 jpeg 两种类型居多, 这主要是因为与其他类型的图像相比, 它们所占的空间较小。顺便提一下, 有些 gif 文件属于动态图像中的动画 (即 animated.gif), 这是 gif 图像的一个特点。

三、关于图像的几个基本概念

人们在描述和使用图像时, 常常会涉及一些基本概念, 如像素 (pixel)、分辨率 (resolution)、原色 (true color)、有损压缩 (lossy compression)、无损压缩 (lossless compression), 等等。对这些基本概念的了解, 将有助于更便捷地索取自己所需的图像信息。

像素是 picture elements 的简称, 它们是组成图像的最小计量单位, 实际上就是我们眼睛能看到的小点。Internet 上的每幅图像都由成千上万个像素组成, 像素的多少决定图像的清晰度。换言之, 构成图像的像素数量越多, 图像的细节便能清晰地得以展示, 图像的清晰度也就越高。

分辨率 (resolution) 是与像素密切关联的一个概念, 它通过计算机屏幕显示的像素纵横比来表示。例如分辨率 640×480 表示屏幕的纵向为 640 个像素, 横向为 480 个像素, 全屏幕共显示 307 200 个像素。其他常见的分辨率有 1024×768 (共 786 432 个像素)、 1280×1024 (共 1 310 720 个像素), 等等。分辨率愈高, 图像所占的空间就愈大, 或者说, 分辨率的高低决定图像所占空间的大小。正是因为如此, 人们往往把图像的分辨率和所占空间大小作为同义词 (见本书将要介绍的一些图像检索系统)。图像所占空间通常以字节 (byte, 简称为 B) 表示, 如 125kB (k 为 kilo 的简称, kB 代表千字节)。

表示每个像素的基本单元比特 (bit) 决定图像颜色的种数。如果计算机用 8 比特 (8 bits) 表示每个像素, 那该图像就能显示 256 (即 2^8) 种不同的颜色, gif 图像就属于这种类型。如果计算机以 24 比特表示每个像素, 那此类图像则可显示 16 000 000 (2^{24}) 余种颜色, 即所谓的原色 (true color) 图像, jpeg 属于原色图像。

由于图像所占空间一般较大, 因此出现了与之相应的压缩技术, 以略去图像中一些人眼观察不到而又占据空间的信息。不过, 并不是所有类型的图像经过压缩质量 (如清晰度) 都不受影响。导致图像质量降低的压缩称为有损压缩 (lossy compression), 如 jpeg 图像; 反之, 不影响图像质量的压缩叫做无损压缩 (lossless compression), 如 gif 图像。

四、图像类型的选择

我们知道从 Internet 检索获取的图像主要有 gif 和 jpeg 两种, 虽然 png 图像时有所见, 但在目前还不经常。因此, 我们在此仅指出和比较 gif 的 jpeg 图像的特性, 以利于对图像类型的选择。图像选择的基本标准有两条: 图像所占空间和图像质量。一般而言, 图像所占空间越大, 其质量也越高。具体地讲, 我们在选择图像时应考虑以下几个方面。

第一, 图像所占空间。如果我们需要的图像将放置在万维网上, 或通过电子邮件和 FTP 传递, 那我们可以采用占据空间较大的 jpeg 格式。如果我们所需的图像大小只含有少于几百个单位的像素或几千字节, 我们则应使用占据空间较小的 gif 格式。

第二, 需要图像的目的。Chastain (2001)指出, jpeg 图像文件每经过一次编辑存档就会降低其质量。因此, jpeg 图像文件最好一次生成, 而需要经常修改变动的图像或需要作为档案保留存放的图像应采用 gif 格式。另外, 如果所制作的图像是照片, 应当采用 jpeg 格式。如果所制作的图像为标识图 (logo)、图标 (icon)、屏幕摄取 (screen shot)、含有文字、包括轮廓分明的线条或大面积色彩, 都应使用 gif 格式。

第三, 图像压缩是否有损。gif 属于无损压缩图像, 而 jpeg 归类于有损压缩图像。如果我们期望压缩图像体积而不影响图像的质量, 那我们应选择 gif 图像。反之, 则使用 jpeg 图像。

第四, 图像的颜色种数。根据前面的介绍, gif 图像只含有 256 种颜色, 相比之下, jpeg 图像则支持 16 000 000 余种颜色。因此, 如果我们所需的图像对颜色种数要求不高, 或者图像中包括许多单色部分, 那么我们可选 gif 图像。反之, 我们则应选择 jpeg 图像。

由此可见, 图像类型的选择涉及数个方面的因素, 在选择时应从全面考虑, 以选用适合于自己需求的图像类型。

参考文献

- 储荷婷. 当代最新信息技术: Internet 使用指南. 北京: 科学出版社, 1997.
- Cass, Stephen. 2004. A fountain of knowledge. IEEE Spectrum Online.
<http://www.spectrum.ieee.org/WEBONLY/publicfeature/jan04/0104comp1.html>.
- Chastain, Sue. 2001. Myth and facts about JPG. In: Your guide to graphics software.
http://www.computerbuddies.us/myths_and_facts_about_jpg.htm.
- Google. 2005. Corporate information; Quick profile. <http://www.google.com/corporate/facts.html>.
http://www.google.com/pr/pr_010611_2.pdf
- Nuttall, Chris. 2004. Why Internet need to break down the language barriers.
 Financial Times, p. 7.
- Alison. Zhang. 1996. Multimedia File Formats on the Internet: A Beginner's Guide for PC Users. http://www.is_edu.hcmuns.edu.vn/WebLib/Books/Internet/multimed/table.htm.

第二章 图像信息检索的主要途径和方法

一、图像信息检索的三种途径

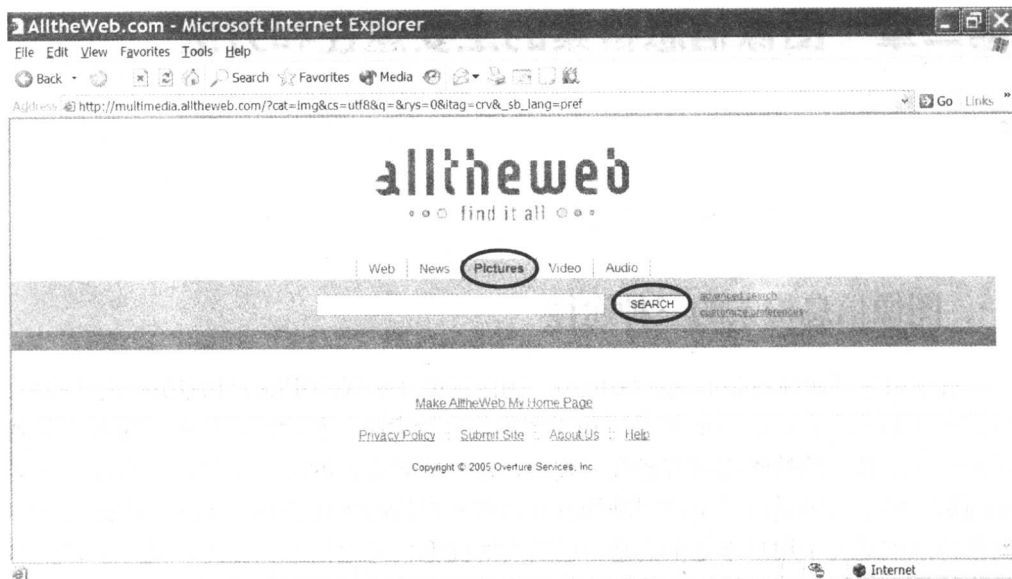
搜索引擎（如Google Image Search）、图像目录（如WebSEEk）和专类收藏（如Art Museum Image Consortium）是可用于图像信息检索的三种途径。其中，搜索引擎要求用户提出具体的检索提问方能进行查询，这一检索途径的特点还包括检索过程直接和简短。搜索引擎适合于用户明确知道自己所要查询的图像信息的情况。相比之下，图像目录则适宜于用户不是很清楚自己所要的图像信息之情形，在此，用户不必提出具体的检索提问，而是按事先编制的目录类别或预先设置的超链接等，浏览自己想要查找的图像。这种检索途径一般比采用搜索引擎花费更长的时间，但它也往往会取得原先意想不到的收获。第三种图像检索途径从专类收藏入手，再借助于搜索引擎和图像目录，达到搜寻某些专题图像（如艺术图像、建筑图像）的目的。我们在分别具体介绍每种图像检索途径之后，再对它们作进一步的比较和归纳。

1. 图像搜索引擎

经过数十年的探索和实践，用于查找文本信息的搜索引擎出现了长足的进步。尽管图像信息检索的历史较短，但它可部分地借鉴历史较长、经过检验且使用广泛的文本信息搜索引擎技术。因此，图像搜索引擎与文本信息搜索引擎之间存在着许多相似之处，图例 2.1 列出了图像搜索引擎之一 AlltheWeb Multimedia Search (<http://multimedia.alltheweb.com>) 的首页。

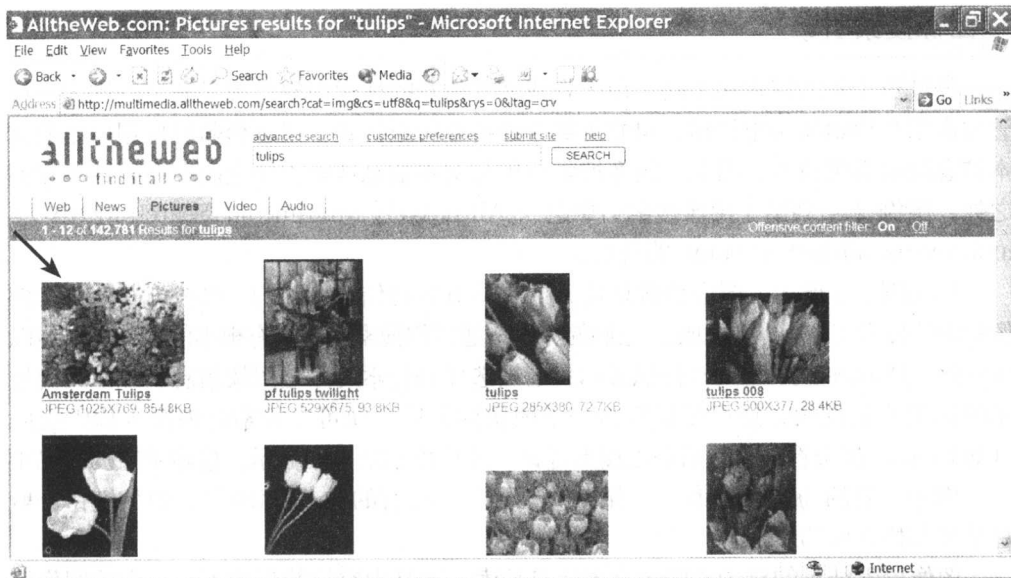
如图例 2.1 所示，图像检索只是 AlltheWeb 各种信息检索的一部分，其他支持图像检索的搜索引擎也基本如此。它们的界面与通常的搜索引擎没有根本的区别，一般以 Images、Photos 或 Pictures 标识表示目前被选择的检索对象是图像信息。如果当前显示的标识不是图像，您用鼠标轻击一下图像标识后，便可在检索框内键入检索提问（如 tulips，郁金香），然后点击搜索按钮，等待检索结果的显示。搜索按钮的命名不一，例如，有的为“Search”，有的叫“Go”，也有的称“Fetch”，但该按钮一般总是列在检索框的右方。

图像检索结果的显示不同于文本信息检索，而且内容详简程度不一，有些图像检索系统提供详细的描述信息，有些系统则只列出题名。AlltheWeb 在这方面的情况居中，



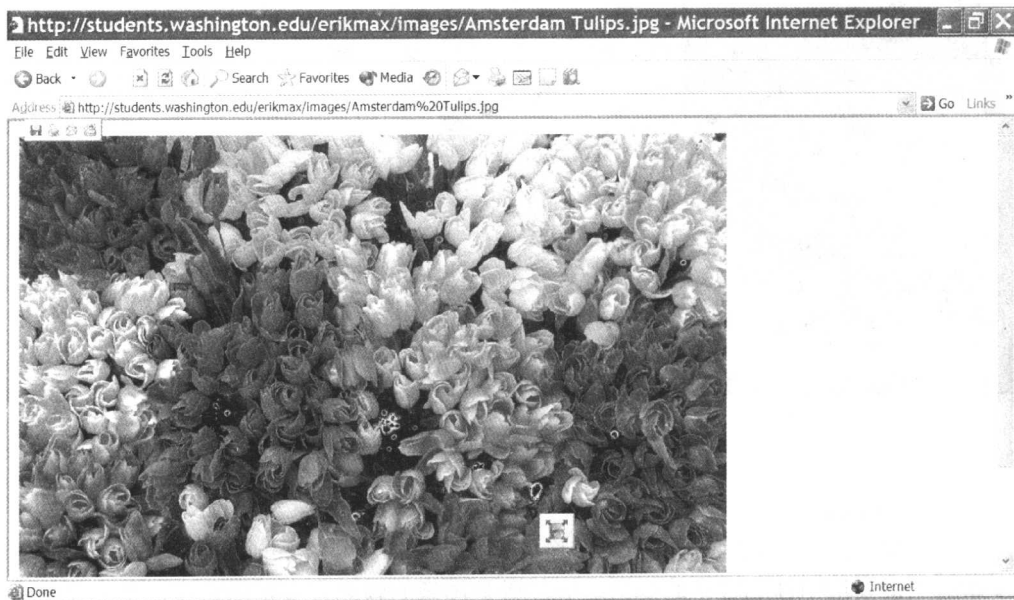
图例 2.1 图像搜索引擎 AlltheWeb 检索界面

给出了检得图像的略图 (thumbnail)、题名 (如 Amsterdam Tulips)、图像文件类别 (如 JPEG)、分辨率 (如 1025×769) 和图像大小 (如 854.8KB) 信息 (见图例 2.2)。



图例 2.2 图像检索结果

如果您对某幅检得图像感兴趣,可直接点击该图像略图或其他相关的超链接,以获得图像全图。如果您使用的图像检索系统还提供有关该图像的其他信息,它们也会一并列出。假设我们选择了图例 2.2 中左起第一幅略图,屏幕上立即出现了该图像的全图(见图例 2.3)。



图例 2.3 检得图像全图

如果您想获得该图像文件,只需把鼠标移至图像的任何一个位置,然后轻击鼠标的右键,选择“Save Picture As...”(见图例 2.4),就可把该图像文件下载并存放到自己的计算机上备用。如果您使用的个人电脑是 Apple 生产的苹果牌电脑,那只需把图像拉到自己准备存放下载图像的文件夹即可。无论您使用哪一个图像检索系统,获取图像文件的操作过程都如上所述,非常简单。因此,我们在本书的其他部分,将不再讲述图像下载的操作步骤。

顺便提一下,下载图像的文件类型取决于检得图像,我们不能在下载过程中加以变换。但是,您如果拥有 Photoshop 这类图像编辑软件,则可将下载后的图像转变为软件支持允许的任何类型的图像文件。

由此可见,通过搜索引擎查找图像的过程并不复杂。从用户的角度看,检索结果的好坏主要取决于检索提问的构成,也就是说,我们该选用什么词语来表达自己的所要查询的图像,不同的词语类型代表了不同的图像检索方法。具体地讲,利用搜索引擎查找图像信息的方法有关键词法、文件扩展名法和图像特性法。