

基于内容的WEB图像检索引擎

袁玉宝 老松杨 谢毓湘 韩智广 (国防科技大学人文与管理学院多媒体技术研究中心 410073)

摘要: 现在的图像搜索引擎大多是基于关键字的检索, 怎样从internet的海量数据中检索到自己想要的图像信息, 是一个亟待解决的问题。本文提出了基于内容Web图像检索引擎及其体系结构, 研究其数据库结构设计, 基于内容检索中索引和相关反馈的问题。

关键词: 基于内容图像检索 CBIR 索引过滤 相关反馈 检索引擎

1 引言

纵观网络的发展历程, 从最早的 ARPA NET, 到如今的 WWW, 网络呈现了飞速发展的趋势。在于网络的应用越来越多, 网络信息量不断以几何级数爆炸式增长, 网上的多媒体信息急剧增加。1999年估计至少有1600万台主机接入互联网, 网上的网页数量至少达到10亿, 有人估计web页面的数量每隔100-120天就会翻一番。互联网上的多媒体信息以图像、音频为主, 据市场调查公司统计, 现在整个互联网上大约保存有500亿张高质量的照片。人们对多媒体信息的检索需求越来越多, 面对着这样一个巨大的、实时扩展的、变化的

数据库, 如何从中找到我们关心的图像信息成为一个亟待解决的问题, 然而, 由于因特网上图像种类繁多, 数量巨大, 很少或基本不用语义来描述, 因此人们很难从海量数据中迅速定位所需信息。随着基于内容图像检索技术发展, 互联网上基于内容的图像搜索引擎技术将成为研究的热点之一。

1994年Yahoo公司创立, 它提供基于目录的信息检索服务, 是Web搜索引擎的雏形, 而真正意义上的搜索引擎是创建于同年的Lycos, 当时Michael Mauldin将John Leavitt的“网络蜘蛛”(spider)程序接入到其索引程序中。在随后的几年里, 互联网和web技术进一步发展, 搜索

引擎也就被广泛使用。

由于很难表达图像自身包含的信息, 现有的图像搜索引擎大部分是通过对图像的文字表达来进行检索。文字信息不能充分表达图像的信息, 用户也不能对查询要求用文字合理描述, 所以图像的检索效率和精度非常低。现在流行的搜索引擎如: Google, Yahoo, OmfoSeek, WebSEEK, 或新浪, 263等都是文本的基础上检索信息的, 图像检索从早期的借用文本索引技术到后来的基于内容检索, 再发展下去将是高层语义和内容结合的检索技术, 最终要达到从海量数据中按需查找图像的目的。互联网上的图像检索引擎就是一个具体的、重要的应用。

本文第二部分提出基于内容检索引擎的体系结构和数据库结构设计, 第三部分详细说明检索匹配子系统的工作原理, 第四部分介绍图像搜索子系统, 第五部分介绍索引过滤机制, 第六部分是我们所作的检索系统, 最后介绍基于内容图像检索引擎的应用和发展趋势。

2 基于内容的Web图像检索引擎的体系结构

从Web搜索引擎的现状可以看出, 现在还很少有一种搜索引擎能够提供基于内容的图像检索, 能够提供给用户一种直观易用的检索界面, 能够让用户可以轻松地找到自己需要的图像。我们提出基于内容的Web图像检索引擎, 正

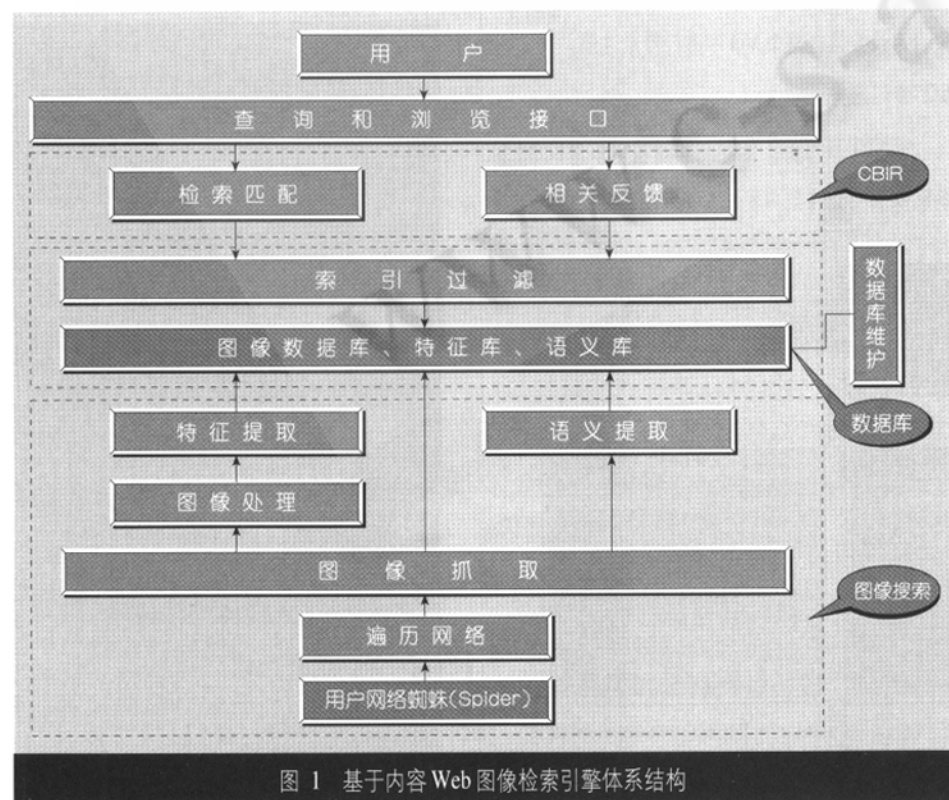


图1 基于内容Web图像检索引擎体系结构

是这一应用的需要。

基于内容的图像检索引擎综合了基于内容检索技术和网络搜索引擎技术, 提供一种在Internet上按照用户所需检索图像的方法。为了保证用户查找信息的精度和实时性, 搜索引擎要建立并维护一个庞大的数据库和索引, 基于内容的web图像检索引擎也和通用的搜索引擎一样要由网络蜘蛛、索引以及检索引擎等部分组成。

基于内容的图像检索引擎包括以下4个模块(如图1所示):

2.1 用户查询和浏览接口

提供给用户方便易用的交互界面, 提供多特征融合的查询方式, 支持关键字的语义检索(该部分要在经过一定使用次数的反馈并且语义网络已经建立的基础上), 允许用户以示例查询和模糊描述等方式对整幅图像、特定对象或多种方式的组合进行查询, 返回用户指定大于一定相似度的若干幅图像, 提供分类浏览功能。

2.2 检索匹配子系统

对用户提交的查询请求进行相似度计算, 并返回符合条件的结果, 对用户的相关反馈作出反映, 调整数据库中相关参数和权重。

2.3 图像搜集子系统

网络蜘蛛不间断的遍历网络, 抓取最新的图像并提取其特征和语义, 存入数据库。

2.4 数据库及其索引和过滤器

索引/过滤模块可以提高检索的速度, 从而可以应用到少量图像数据库中。在进行向量运算之前先滤除那些不符合要求的图像, 过滤出的数据集合再用高维特征匹配来检索, 索引则用于低维特征, 可以利用B+树方式以加快检索, 过滤器作用于全部数据, 索引结构根据图像特征数据的高维特征来建立。

3 检索匹配子系统

检索匹配子系统是用户查询接口的具体实现, 也是基于内容图像检索的重要部分, 它的算法的好坏直接影响检索的结果和精度。

对于低层特征利用特征之间的距离来进行相似性匹配, 对于文本的查询直接搜索关键字与图像描述, 对于色调和示例的查询通过计算和用户提交的特征之间的距离, 查找出符合条件的图像集。匹配引擎中包括一个较为有效、可靠的相似性测度函数集, 对于多种方式的组合检索, 通过一定的排序过滤找到符合条件的图像。

用户评判检索出图像的差与好, 并返回给系统, 系统根据相关反馈对图像数据库中的相关部分参数和权重作出调整, 从而提高下次检索的精度, 通过一定次数的反馈, 检索精度会达到一个令人满意的程度, 这也是系统的一个自学习过程。

3.1 检索与相似度计算

根据用户提交的查询条件, 分别在数据库中匹配计算, 把大于一定相似度的符合条件的结果返回给用户, 最多使用的是基于主色调和基于示例的查询以及语义的查询。

主色调能够代表一幅图像的基本概貌, 如大海可以和蓝色相关连, 要想检索大海则指定蓝色为主色调进行检索, 对选定主色调进行适当扩展来描述图像的颜色特征是非常有效的, 有的图像中可能会有两种以上的主色调, 则可以用多色调来描述。

对于用户提交的示例, 用一种特征来计算相似度可能会使检索精度不高, 我们可以用多特征的融合来检索和示例图像最相似的图像。

图像的分类信息和从网页上通过分析得到的语义信息作为初始的语义存储在语义库中, 通过用户的相关反馈来添加关键字并与相关的图像之间建立联系, 经过一定次数的反馈, 语义



库可以达到满足关键字查询的要求。

3.2 用户相关反馈

根据用户对感兴趣的图像的点击,或是主动交互:输入、编辑等操作的返回,系统根据返回作相应的参数调整,对本次返回和以后的检索都可以得到更精确的结果,这也是系统的自学习过程。

相关反馈和图像的自动标注技术是重要的突破。相关反馈是在信息检索系统中的一种指导性学习的技术,用以提高系统的检索能力。在基于内容的图像检索中,通常检索的结果不能令人满意,这时用户可以告诉系统哪些是符合用户需求的正反馈图片,哪些是不符合检索内容的负反馈图片,系统根据用户提交的指导信息,对内部检索参数进行调整,从而优化检索结果并提供给用户新的检索结果。对于图像的标注问题,可以手工标注很小一部分图,利用相关反馈的方法,将用户因为反馈而标定的信息不断加入数据库,同时对这些信息进行扩展,使具有相似内容的图像因为这些图的相似性而得到标注,这种方法使检索精度得到了很大提高。

4 图像搜索和特征提取子系统

为了搜索在Internet上的很多图像,我们用网络蜘蛛(Spider或Crawler)自动的遍历

网站和网页,把所有相关图像或链接抓回来存入本地数据库。首先建立图像的大致分类,我们把图像分为人物、运输工具、体育、风景、建筑等大类,再在下面建立次级分类如运输工具下面建立汽车、飞机、船只等。其次给网络蜘蛛预设几个比较主要的链接,以便它能够遍历整个网络,不遗漏主要的页面,Spider搜索每个页面,搜集图像,对图像提取内容特征、语义特征,分别存入本地数据库的特征库和语义库中。

4.1 图像低层特征提取

对图像进行去噪声和增强处理,从包含大量信息的图像中分解出不同种类的低层特征信息,主要包括视觉特征和统计特征,视觉特征包括具有直观意义的图像的形状与颜色特征,统计特征是对图像像素、纹理等特征的统计,特征的提取可以是整幅图像,也可以是图像某个区域或具体的内容对象。按照给定的描述算法提取它的颜色、形状、纹理及空间位置等特征。

4.2 语义特征的提取

由于现在的图像理解技术还不足以提取出真正意义上的人们理解的信息,图像的语义信息很难从图像本身的内容中获得,但由于图像处于Internet环境,有很多关联的信息可以提供这一语义。我们从以下几方面提取语义信

息:文件名和URL:通常网页编辑人员会给图像赋一个有意义的名字,URL通常代表网页的层次信息,如分类信息“动物”、“鸟类”等,这也可以给我们提供语义信息。替换文本(ALT):为了在不能显示图像的页面浏览中,给读者明确的信息,用一段文本代替图像显示。图像周围的文本:很多时候图像是用来说明文字要表达的内容的,因此在图像上下文中都包含着图像的语义信息。

5 索引过滤机制

基于内容检索所依赖的图像数据库、特征库、语义库之间有着复杂的联系(如图2),同时为了保证检索的效率和信息的及时更新,对数据库要实时地建立索引,并建立索引库。

有效的索引过滤机制配合图像库的检索引擎,可以极大地提高查询的效率。和其他类型数据库不同,图像库需要采用多维索引技术,对于大型图像库,还必须对图像特征和结构建立高效的索引机制,具体包括索引的表达、索引的组织以及索引的提取等。图像库的索引比字符数据库的索引更加复杂,由于它所处理的信息具有高维特征,因而更具近似性、无序性、多重性等特点。

早期研究的索引算法,结构比较简单,且不适合于目前的多媒体数据,于是提出了多维索引的方法,多维索引在多个属性上进行,一个索引结构用于检索K个属性,这种结构可以大大缩短响应时间。比较流行的多维索引方法有R树、线性四叉树以及栅格文件,其中R树及其变体是最为有效的多维索引方法,但是当向量维数超过20时,以上方法的效率均迅速降低,几乎等价于顺序查找。另外,由于R树的构造基于几何意义上的覆盖关系,它只限于以欧氏距离作为相似度量的查询,为了能够有效地利用以上的

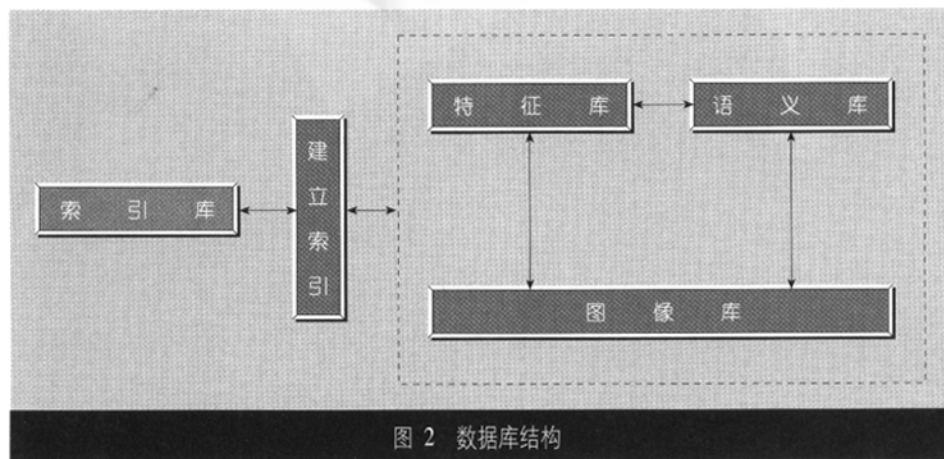


图2 数据库结构

索引方法, 必须将 n 维特征向量转换为 20 以内的 m 维。维数的降低不可避免地带来信息的丢失, 导致查询结果中有较多的错误记录。

在多媒体数据基于内容的索引算法中, 聚类是经常采用的方法。聚类的目标是将数据聚集成类, 使得类间的相似性尽可能地小, 而类内的相似性尽可能地大。聚类的这种思想对基于内容的检索很有借鉴意义的。通过对海量图像库采用聚类算法并按某种相似度原则进行分类, 可以把相似的图像聚合到一起, 从而大大缩小图像搜索的范围, 达到快速准确检索到目标图像的目的。

另外值得一提的是与检索过程相反的过滤机制。网上的图片信息错综复杂, 网站管理者迫切需要行之有效的过滤机制来对图片信息进行管理。例如, 用户上传的图片信息进行过滤, 以剔除掉那些含有不健康或者反动言论的图片。由于需求已经上升到图像内容这一层次,

因此对应的过滤机制也必须结合基于内容检索技术来进行。

6 结束语

随着技术的发展, 用户可以越来越精确地搜索到自己所需的图片。基于内容 Web 图像检索技术可用于商务网站中商品的搜索。有些产品如椅子等, 用户比较难以用关键词描述出明确的查询。但是利用基于内容的图像搜索技术, 却可以解决此类问题: 用户首先点选画面中随机产生的产品图片中较符合需求的产品图片, 利用基于内容的搜索技术, 找到与用户选择的图片较为相似的图片, 经过数次交互后, 用户可以愈来愈接近所要的产品。

另外一种与图像搜索技术密切相关的技术是敏感图像的过滤。这种技术可以在大量的图片数据库中, 搜索出敏感图片(如色情图片过滤)。这种技术非常适合在许多允许用户上传图

片的环境中, 如免费个人网页、个人拍卖网站的产品图片搜索, 个人相册等网络服务; 而政府部门、ISP、公安部分及教育单位等也都需要自动化追踪用户使用不当的情况。LookThatUp 及 Cobion 都推出了这种相关的产品。

基于内容的 Web 图像检索引擎未来的发展, 重点将是提供多种特征融合的、低层特征与高层语义相结合的检索方式, 能够支持自然语言的查询、多语言的自适应性、自我学习的功能、个性化和多媒体交互的输入输出界面, 使用户的使用更加方便和简单, 检索的结果更加准确。 ■



参考文献

- 1 Zheng Chen, Liu Wenyin, Feng Zhang, Mingjing Li, Hongjiang Zhang, Web Mining for Web Image Retrieval, <http://www.microsoft.com/china/research>.
- 2 Ye Lu, Chunhui Hu, Xingquan Zhu, etc, A Unified Framework for Semantics and Feature Based Relevance Feedback in Image Retrieval Systems.
- 3 Zi jun Yang, C.C. Jay Kuo, Learning Image Similarities and Categories from Content Analysis and Relevance Feedback.
- 4 李国辉、王辰、薛峰, 几种典型的基于内容检索系统简介。
- 5 张卫丰、徐宝文等, Web 搜索引擎综述, 计算机科学, Vol.28, No.9, 2001.
- 6 李向阳、庄越挺、潘云鹤, 基于内容的图像检索技术与系统, 计算机研究与发展, Vol.38, No.3, 2001.
- 7 段立娟、高文、马继惠, Rich Get Richer——图像检索中的一种自适应的相关反馈方法, 计算机研究与发展, Vol.38, No.8, 2001.

