

# 环球网图象搜索引擎研究综述

桑梓勤 丁明跃 张天序

(华中理工大学图象识别与人工智能研究所, 武汉 430074)

**摘要** 提出了建立一个 WWW 图象搜索引擎的方案, 搜索引擎在 Internet 上浏览主页, 对遇到的图象进行特征计算, 抽取出索引信息, 存储索引向量, 根据图象内容进行相似图象的查找。分析了颜色直方图、边缘方向直方图、纹理分析和形状不变性等基于图象内容查询的技术, 并对 WWW 图象搜索引擎的发展作出了预测。

**关键词** WWW 图象搜索引擎, 索引向量, 颜色直方图, 边缘方向直方图, 纹理分析, 形状不变性

## 1 引言

随着国际互连网络 Internet 的发展, 特别是环球网(World Wide Web, 以下简称 WWW)的兴起, 网络上传输的信息正按指数式增长。由于现在大学、科研机构、政府机关、公司以及个人的主页绝大部分是手工编制的。信息之间的超级链路也是人工给定的, 因为缺乏必要的索引, 常常导致网络上无用的或冗余的信息大量传输, 造成网络阻塞, 因此有人把 WWW 戏称为“全世界都在等待”(Whole World Waiting)。解决这一问题的一个途径是, 为网上信息建立快捷而有效的索引和搜索机制。

最近一年多以来, 有所谓“爬虫”和“蜘蛛”在 WWW 上为找到的文本文档收集索引信息, 这些搜索引擎可以“阅读”文本, 抽取索引信息, 用于引导以后对实际地址的访问。卡内基—梅隆大学开发的一个 WWW 文本搜索引擎 Lycos, 可支持基于关键词的交互式搜索, 已经投入实际运行([Http://lycos.cs.cmu.edu/](http://lycos.cs.cmu.edu/))。Lycos 每天都要检索、分析成千上万的文本文档, 以保持信息的时效性。我们所需要的是与之相对应的 WWW 图象搜索引擎, 它能在 WWW 上为浏览过的图象建立索引信息。与文本搜索不同, 图象搜索引擎要能够进行图象分解和判别, 为图象加注释, 存储抽取出来的索引信息, 支持基于图象内容的 Lycos 风格的搜索。WWW 图象搜索引擎与普

通的图象数据库搜索引擎亦不同, 它把整个 Internet 当作图象数据库, 但普通图象数据库的搜索策略及图象处理技术, 对于 WWW 图象搜索引擎的建立, 仍然是十分有帮助的。

## 2 有关的工作

最近几年来, 图象库基于内容的查询研究受到了广泛重视, 其中既有学术研究, 也有商业应用。最早出现的系统是 IBM 公司 Almaden 研究中心的 QBIC 系统<sup>[1]</sup>, 它允许用户使用低层的提示(如颜色相似性、纹理相似性等), 给出示例图象, 在图象库中找出相似的图象。如果采用人工预处理的方法, 分割出图象中的物体, 则还可使用形状相似性找出相似图象。麻省理工学院媒体实验室实现了一个类似的项目 Photobook<sup>[2]</sup>, 它采用“语义压缩”, 即构造图象的一种紧密表示, 以保留图象的相似性。密歇根州立大学则开发了一个商标和图标图象库原理系统, 可以根据图象的颜色内容和形状内容进行图象检索<sup>[3]</sup>, 对于颜色内容, 利用 3 个一维的颜色直方图表示图象颜色的 RGB 分布, 通过计算归一化直方图的欧氏距离得到数据库中图象与查询图象的相似性, 而形状内容则采用 Canny 算子提取出边缘点, 用边缘点的方向直方图表示形状。图象检索时要综合考虑颜色内容和形状内容。

而 WWW 文本搜索引擎, 则有 AltaVista、Ex-

cite、Infoseek、Lycos、MetaCrawler、WebCrawler、Yahoo! 等投入了商业运行。在这些流行的搜索引擎中,一般都同时支持关键词查询和主题查询。关键词查询需要考虑主页中的词频、词的组合等因素,而主题查询则往往需要有编辑人员来维护,在这方面,Yahoo! 给人的印象很深刻。而 MetaCrawler 自己没有内部数据库,在收到用户端的查询请求后,MetaCrawler 并行地查询 AltaVista、Excite、Lycos、WebCrawler、Yahoo! 等引擎,然后把返回的参考结果进行验证和排序,把最终结果返回给用户,因此 MetaCrawler 可以称为搜索引擎的搜索引擎<sup>[4]</sup>。

在搜索策略方面,Yuwono 等人提出了建立基于关键词的自动浏览器的思想,采用向量空间检索模型、相对性反馈机制和超文本映射方法,找到 WWW 上的信息资源<sup>[5]</sup>。

就作者目前所知,Internet 上还没有一个成熟的图象搜索引擎,可以像文本搜索引擎那样有效地查找网上图象信息。正如从文本文档中抽取索引信息不需要理解文本的涵义一样,图象的加注、索引和查找也不需要解决图象理解的难题,一般来说,只对图象进行低层的分解和判别。

### 3 主要结构

由于 WWW 图象搜索引擎要完成以下 4 件工作:在网上收集图象,图象相似性计算,维护图象索引库,回答用户的查询。

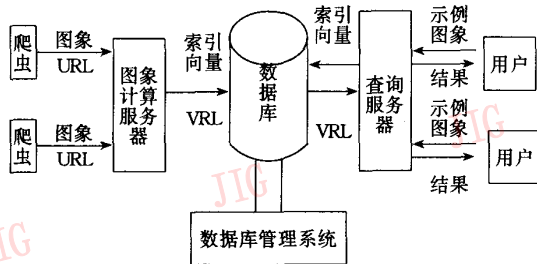


图1 图象搜索引擎系统结构

如图 1 所示,整个系统也按照这 4 个功能分成 4 个部分:(1)网络爬虫,在网上漫游,取回图象及其 URL<sup>[6]</sup>;(2)图象计算服务器,对图象的颜色直方图、边缘直方图、多分辨率纹理特征、平移不变特征向量等图象性质进行计算,得到索引向量;(3)数据库管理系统,对数据库中的索引记录进行周期性的维护与更新,增加新的记录,验证索引记录是否存在,删除重复记录和不再有效的记录;(4)查询服务

器,接收用户的示例图象或特征描述,产生索引向量。由于索引向量一般都是高维向量,可以采用主要构件分析(Principal Component Analysis)的方法,降低维数<sup>[7]</sup>,再进行相似性计算,把相似图象的索引结果按照相似程度排序,返回给用户。

### 4 图象计算

适于图象检索的图象计算方法主要分为 2 类<sup>[3]</sup>:保留空间信息的方法与不保留空间信息的方法。前者在特征计算时保留了图象中的空间信息,可以根据特征集重构图象,代表性的方法有物体的多边形近似、基于物理的建模与主要构件分析。而不保留空间信息的方法则抽取统计特征用于描述所关心的物体,它包括各种基于特征向量的方法,如直方图和不交矩。

WWW 图象搜索引擎的图象计算综合了上述两类方法,即采用颜色直方图、边缘方向直方图、纹理分析、形状不变性表示等。把一幅图象分成  $N$  个子图象,如果有  $M$  种特征计算方法,则图象的索引向量  $X$  由个子向量组成。

$$X = [x_1, \dots, x_n] \quad (1)$$

子向量  $x_i$  的维数是高维的,相似性搜索的时间随向量空间的维数呈指数式增长,随数据库的大小呈对数式增长<sup>[8]</sup>,White 和 Jain 还指出:数据的内在维数要低得多。可以通过最有代表性的特征<sup>[7]</sup>降低维数,把索引向量  $X'$  变为低维的。

索引向量的查询,可以通过计算归一化的 Minkowski 距离而获得<sup>[9]</sup>,

$$\bar{M}(x'_i, y'_i) = \frac{M(x'_i, y'_i)}{\mu_m^{(i)}} \quad (2)$$

和分别是降维的数据库数据记录中的子索引向量和查询子索引向量,归一化因子是整个数据库中索引向量的特征的概率分布。

$$\mu_m^{(i)} = E[M(x'_i, y'_i)] \quad (3)$$

其期望值  $\mu_m^{(i)}$  可以在数据库管理系统中离线计算。

#### 4.1 颜色直方图

颜色直方图是表示图象中颜色分布的一种方法,可以用于检索图象。在颜色直方图中,每个刻度表示颜色空间中的一种颜色,两幅图象直方图的距离定义为它们的直方图分布的相似性匹配,可以用二次式形式表示。直方图颜色空间的一个刻度称为一维,如 256 色的彩色图象,其直方图分布为 256

维。令  $X, Y$  为  $N$  维分布, 则加权的距离测度可以表示为二次式形式:

$$d_{\text{hiss}}^2(X, Y) = (X - Y)^T A (X - Y) \quad (4)$$

这里  $A = [a_{i,j}]$  是一个矩阵, 权值  $a_{i,j}$  说明了颜色  $i$  与  $j$  的相似程度,  $a_{i,j}$  可以归一化为  $0 \leq a_{i,j} \leq 1$ , 且  $a_{i,i} = 1$ 。  $a_{i,j}$  越大说明颜色  $i$  与  $j$  越相似, 反之  $a_{i,j}$  越小说明 2 种颜色越不相似。2 幅图象的颜色直方图分布也可归一化为  $0 \leq x_i, y_i \leq 1$ , 且

$$\sum_{i=1}^N x_i = \sum_{i=1}^N y_i = 1$$

对于自然场景, 256 色就足够了。  $d_{\text{hiss}}^2$  与欧氏距离不同, 其相似程度矩阵  $A$  的每一项  $a_{i,j}$ , 可以代表人对 2 种颜色  $i$  与  $j$  的感观感受。

$N$  维分布的加权距离测度  $d_{\text{hiss}}^2(X, Y)$  的最直观算法, 需要  $O(N^2)$  次运算, 这在实际运算中是非常耗时的。 Hafner 等人提出了采用低维过滤的方法<sup>[10]</sup>, 使得: (1) 低维的距离测度可以无遗漏地过滤图象库; (2) 耗时的距离测度的匹配只限定在过滤出来的图象库的一个小子集中; (3) 图象库可以按低维距离测度建索引。

## 4.2 边缘方向直方图

图象的颜色信息对于光照条件的变化是很敏感的, 而形状特征对于光照是不变的, 采用 Canny 边缘算子<sup>[11]</sup>获得的边缘信息可用于表示形状信息。边缘方向直方图是平移不变量, 但不是旋转不变和比例不变的<sup>[3]</sup>, 通过类似于颜色直方图的比较方法, 可以进行边缘方向的匹配。

## 4.3 纹理特征

图象可以看成是不同纹理区域的组合, 与这些区域有关的图象特征可以用于检索, 用户可给出一小块区域找到整幅图象。纹理分析有很长的历史了, 算法也很多, 如随机场模型如 Markov 随机场/Gibbs 概率分布模型<sup>[12]</sup>、多分辨率过滤技术如纹理金字塔过滤<sup>[13]</sup>及小波变换<sup>[14]</sup>等。

## 4.4 形状不变性

图象中物体形状匹配一般是通过 2 个物体局部的特征点的匹配来完成的, 而找对应点往往是很困难的, 但如果以物体自身中心为坐标系原点, 两个物体的上、下、左、右分别匹配, 则问题会简单一些。但仍然有下面 3 个问题: (1) 取样误差问题, 这是由视点变化引起的, 图象采样的视点稍有变化, 则图象中

的物体特征就会发生极大变化; (2) 参数化误差问题, 图象采样时是先定坐标系再产生物体的特征点的坐标, 而不是先有特征点的几何关系再定坐标系, 因此透视变换中的微小畸变都会引起特征描述的极大变化; (3) 非唯一性问题, 这也是视觉匹配和识别中的常见问题。

Scalaroff 和 Pentland 提出了采用 Galerkin 插值的方法来进行形状描述<sup>[15]</sup>, 以解决上述问题。首先用一个向量  $X = [x_1, \dots, x_n]^T$  来表示图象中的物体形状,  $x_i$  是局部特征点。用这些点构造一个形状的有限元模型, 然后计算有限元模型的特征模式  $\phi$ , 特征模式表示了形状的正交频域描述。模式形状向量  $X_M$  为:

$$X_M = X + a \cdot \phi \quad (5)$$

$a$  为标量。两个物体的形状如果相似, 则其模式形状向量也是相似的。可以考虑一个物体是另一个物体的变形。

3 种特征模式为刚体的平移、刚体的旋转及非刚体模式。非刚体模式的频率阶数由振动频率确定, 低频模式表示全局的变形, 高频模式表示局部的变形, 通过这种变换, 把物体的形状由空域变换到频域, 形状匹配也相应地转换为特征向量的匹配。

## 5 小 结

WWW 图象搜索引擎是网络技术与图象库搜索技术的结合, 其实用性对反应速度提出了很高的要求, 但由于网络带宽总是跟不上网上信息量的增长速度, 因此构造一个索引库是非常必要的, 索引记录的排序与查找仍然值得认真研究, 而且图象内容的语义注释对于索引的建立也是有帮助的。

基于内容的图象查询的研究才刚刚开始, 第二代加上图象理解的图象查询则对视觉理论提出了更高的挑战。随着 Internet 上多种媒体的广泛应用, 如 Internet 电影库、视频库、数字图书馆和艺术馆的出现, 我们深信对 WWW 图象搜索引擎的需求会越来越迫切。

## 参 考 文 献

- 1 Flickner M, Sawhney H, Niblack W, et al. Query by image and video content: The QBIC system. IEEE Computer, 1995, 28(9): 23~32.
- 2 Pentl P, Picard R W, Sclaroff S. Phonobook: Tools for content-based manipulation of image database. International Journal of Computer Vision, 1996, 18(3): 233~254.

- 3 Jain A K, Vailaya A. Image retrieval using color and shape. *Pattern Recognition*, 1996, 29(8): 1233~1244.
- 4 Selberg E, Etzioni O. Multi-engine search and comparison using the MetaCrawler. *WWW Journal*, 1996, 1(1).
- 5 Yuwono B, Lam S L, Ying J H, Lee D L. A World Wide Web resource discovery system. *WWW Journal*, 1996 1(1).
- 6 Uniform resource locators. *Internet Working Draft*, 1994(1).
- 7 Daniel L Swets, Yuyang Weng. Using Discriminant eigenfeatures for image retrieval. *IEEE 1996 PAMI-18(8)*: 831~836.
- 8 White D, Jain R. Algorithms and strategies for similarity retrieval. *Technical Report VCL-96-101*. Visual Computing Laboratory, University of California, San Diego. 1996.
- 9 Duda O, Hart P E. *Pattern Recognition and Scene Analysis*, New York: John Wiley, 1973.
- 10 Hafner J, Sawhney H S, Equitz W, et al. Efficient color histogram indexing for quadratic form distance functions. *IEEE 1995, PAMI-17(7)*: 729~736.
- 11 Canny J. A computational approach to edge detection. *IEEE PAMI-8(11)*, 1986: 679~698.
- 12 Funt B V, Finlayson G D. Color constant color indexing. *IEEE 1995, PAMI-17(5)*: 522~529.
- 13 Freeman W, Adelson E H. The design and use of steerable filters. *IEEE 1991 PAMI-13(9)*: 891~906.
- 14 Manjunath B S, Ma W Y. Texture Features for browsing and retrieval of image data. *IEEE 1996 PAMI-18(8)*: 837~842.
- 15 Sclaroff S, Pentland P. Modal matching for correspondence and recognition. *IEEE 1995 PAMI-17(6)*: 545~561.
- 16 Berners-Lee T. Hypertext transfer protocol. *Internet Working Draft*, 1993(11).
- 17 Gimel'farb, G L, Jain A K. On retrieval texture images from an image database. *Pattern Recognition*, 1996, 29(9): 1461~1483.
- 18 Swain M J, Ballard D H. Color indexing, *International Journal of Computer Vision*, 1991, 7(1).



**丁明跃** 1961年生,华中理工大学图象所副所长,教授,博士生导师。1991年至1993年获德国洪堡研究基金,在德国不伦瑞克工业大学机器人研究所从事博士后研究工作。主要研究领域包括计算机视觉、路径规划、目标识别与跟踪等。《电子学报》与《宇航学报》编委。



**桑梓勤** 1969年生,现为华中理工大学图象所博士研究生,主要研究方向为算法与计算复杂性、光照模型与图象合成、图象相似性。



**张天序** 1947年生。华中理工大学图象所所长,教授,博士生导师。中国人工智能学会计算机视觉与智能控制学会理事,国际应用人工智能协会会员。主要研究领域包括智能图象信号与分析、成像与计算机视觉应用、智能控制与制导。

## The Study of World Wide Web Image Search Engine

Sang Ziqin, Ding Mingyue, Zhang Tianxu

(*Institute of Pattern Recognition and Artificial Intelligence Huazhong University of Science and Technology*)

**Abstract** This papers proposes the development of a world wide web image search engine, which browses homepages on Internet, computes the features of images embedded in the pages, extracts the index information based on image content, stores the index vectors, and answers the users' queries of similar images. The paper also analyzes the technologies of query be image content such as color histograms, edge orientation histograms, texture measures and shape invariants. In the ending discussion the requirement of world wide web image search engine is predicted.

**Keywords** World Wide Web, Image search engine, Index vector, Color histograms, Edge orientation histograms, Texture measures, Shape invariants