

基于内容的图像检索技术

徐 庆, 杨维维, 陈生潭

(西安电子科技大学 机电工程学院, 陕西 西安 710071)

摘 要:随着多媒体技术的迅速发展,传统的基于关键字的信息检索技术已逐渐不能满足要求,因此,基于内容的图像检索技术成为当今的一个研究热点。介绍了基于内容图像检索系统的基本组成。综述了基于颜色、纹理、形状、语义等图像检索技术的主要方法。分析和比较了现有的各种图像检索技术的方法。讨论了相关反馈技术、检索性能的评价等 CBIR 研究中的关键问题。同时指出了 CBIR 研究中存在的问题,以及未来的发展趋势和研究方向。

关键词:基于内容的图像检索;多媒体技术;特征;相关反馈

中图分类号:TP391.3;G354.4

文献标识码:A

文章编号:1673-629X(2008)01-0126-03

Content - Based Image Retrieval Technology

XU Qing, YANG Wei-wei, CHEN Sheng-tan

(School of Electromechanical Engineering, Xidian University, Xi'an 710071, China)

Abstract: With the rapid development of technology of multimedia, the traditional information retrieval techniques based on keywords are not sufficient, content - based image retrieval (CBIR) has been an active research topic. The basic components of content - based image retrieval system are introduced. Image retrieval methods based on color, texture, shape and semantic are discussed, analyzed and compared. Some key methods such as the relevance feedback, the performance evaluation and so on are also discussed. From evaluating them, have presented the current challenges and future trends for CBIR.

Key words: content - based image retrieval; multimedia technology; feature; relevance feedback

0 引言

随着计算机技术和网络技术的发展,以及多媒体技术的应用,数字图像的数量正以惊人的速度增长,传统的基于关键字的信息检索技术已逐渐不能满足要求,如何有效地组织、管理和检索大规模的图像数据库,正成为当前研究的一个热门课题。基于内容的图像检索(CBIR)是解决这一问题的关键技术之一。

CBIR 不同于传统的检索手段,它是利用图像的颜色、纹理、形状、对象的空间关系等基本特征进行检索,并把这些量化特征与图像存储在一起。它使用的是基于相似度量的示例查询方法。目前国内外许多机构都在进行相关的研究,并推出了以 IBM 的 QBIC^[1] 和 MIT 的 PhotoBook^[2] 为代表的一系列成功产品。

1 基于内容的图像检索系统

一般可把 CBIR 系统看作是介于信息用户和(多媒体)数据库之间的一种信息服务系统。这类系统的通用框架如图 1 所示。在此系统中,用户发出查询要求,系统将查询要求转化为计算机内部描述,并借助这些描述与数据库中的信息进行匹配,提取出需要的信息数据,用户可通过人机交互界面利用相关反馈技术改进查询条件进行新一轮检索。

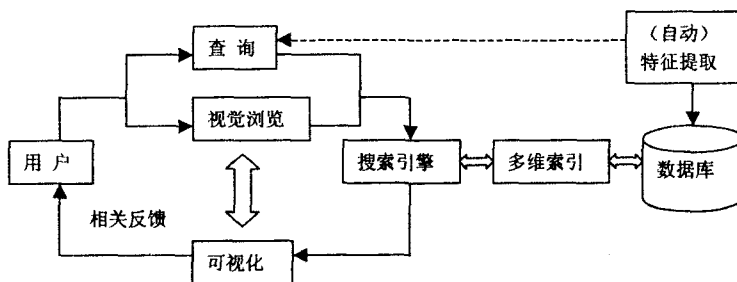


图 1 基于内容的图像检索系统

收稿日期:2007-04-05

作者简介:徐 庆(1982-),男,山东新泰人,硕士研究生,主要研究方向为图像处理、模式识别和多媒体技术;陈生潭,副教授,主要研究方向为图像处理、视频压缩编码以及多媒体数据通信等。

2 基于内容的图像检索技术

基于内容的描述是基于内容的图像检索的前提。图像内容既包括了颜色、纹理、形状等低层次视觉特

征,又包括了高层语义特征。基于低层特征的图像检索,其特点是图像特征可以从图像中直接获得,而基于高层语义特征的图像检索不仅要解决数学上的建模、描述、搜索算法等问题,还要考虑语义特征所具有的模糊性、不确定性,以及对自然语言描述的依赖性。

2.1 基于颜色特征的检索

颜色特征是图像最直观而明显的特征,一般采用直方图来描述。颜色直方图法具有运算速度快、存储空间要求低的特点,并且对图像的尺度及旋转变化的不敏感,因此得到了广泛关注。目前基于颜色特征的检索已成为检索的一个重要手段,并提出了许多改进方法。总体来说,主要分为两类^[3]:全局颜色特征检索和局部颜色特征检索。

2.1.1 全局颜色特征检索

全局颜色特征检索方法目前采用最多的是色彩直方图的方法,Swain的主要思想是根据色彩直方图统计每种色彩在图像中出现的概率,然后采用色彩直方图的交来度量两幅图像色彩的相似性,其最大的缺点是完全丢失了图像色彩的空间信息。Pass^[4]等人提出以图像的色彩聚合矢量(CCV, Color Coherence Vector)来作为图像的索引,它是图像直方图的一种演变,聚合矢量中的聚合信息在某种程度上保留了图像色彩的空间信息。Stricker和Orengo^[5]提出了累计色彩直方图方法,并且提出了色彩矩的方法,认为色彩信息集中在图像色彩的低阶矩中。他们主要对每种色彩分量的一阶、二阶和三阶矩进行统计。

2.1.2 局部颜色特征检索

由于全局颜色特征检索捕获了整幅图像色彩分布的信息,丢失了许多局部的色彩空间信息。目前从划分局部区域的角度来说可分为:基于固定块的图像分割、基于手工的区域分割、采用交互的半自动的区域分割以及一些自动的色彩分割方法。局部区域中的色彩信息可以表示为平均色彩、主色彩、色彩直方图和二进制色彩集。Hsu^[6]等人试图结合图像的色彩信息和图像色彩的部分空间信息对色彩的直方图进行检索。Smith和Chang^[7]采用色彩的自动分割方法,形成一个二进制的色彩索引集,在图像匹配中,比较这些图像色彩集的距离和色彩区域的空间信息。

2.2 基于纹理特征的检索

对纹理图像的描述常借助纹理的统计特性或结构特性进行,对基于空域的性质也常可转换到频域进行活动,所以常用的纹理描述方法是:统计法、频域法、结构法,这些方法也常结合使用。

纹理统计分析方法主要有共生矩阵分析法、马尔可夫分析法、多尺度自回归MRSAR模型以及遗传算

法等。Tamura等从人类感知心理学的角度提出纹理的6种视觉特性:粗粒度、对比度、方向性、线性度、规则度和粗糙度。该纹理特征表示具有实际的视觉意义,可以为图像检索提供更好的用户交互性。同时,由于纹理很少能提供语义信息,描述比较困难。因此通常作为检索过程的辅助手段或者和其他特征结合使用。

2.3 基于形状特征的检索

在基于形状的图像检索中有3个问题需要解决。首先,形状通常与图像中的特定目标对象有关,因此形状比颜色和纹理的语义性更强,要获得目标的形状参数,先要进行图像分割,所以形状特性会受到图像分割效果的影响。其次,描述目标的形状是一个非常复杂的问题,人对形状的感觉是视网膜感受和现实世界的知识相结合的结果。事实上,目前还没有找到与人的主观感觉相一致的形状模型。第三,从不同的视角获得的形状可能会有很大差别,为了准确进行形状匹配,需要解决平移、缩放、旋转中的不变性问题^[8]。

形状的代表方法有两类:基于边界的表示和基于区域的表示。这两种表示方法的典型代表分别是傅里叶描述子和不变矩。傅里叶描述子的基本思想是用对图像进行傅里叶变换得到的边界作为形状描述。其中一个优点就是把二维问题简化为一维问题。傅里叶变换的高频分量对应细节而低频分量对应总体形状,所以可以只用一些对应低频分量的傅里叶系数来近似描述边界形状。文献[9]比较了基于区域的方法、基于边界的方法和两种方法的综合使用,实验结果表明,两种方法综合使用效果更好。

2.4 基于语义特征的检索

与基于低层物理特征查询不同,语义特征查询是基于文字的查询,包含了自然语言处理和传统图像检索技术。这种检索方法的目标是最大限度地减小图像简单视觉特征与丰富的语义之间的语义鸿沟(semantic gap)。缩小语义鸿沟的办法有2种:由高层语义导出低层特征和由低层特征向高层语义的转换。

图像语义具有模糊性、复杂性、抽象性,一般包括3个语义层次:特征语义、目标和空间关系语义、高层语义。特征语义就是图像的颜色、形状、纹理等低级视觉特征,与视觉感知直接相连;目标语义和空间关系语义需要识别和提取图像中的目标类别、目标之间的空间位置等关系,涉及到模式识别和逻辑推理的相关技术;高层语义主要涉及图像的场景语义(如海滨、街道、室内等)、行为语义(如表演、超越、进攻等)和情感语义(如平静、和谐、振奋等)。一般而言,高层的图像语义往往建立在较低层次的语义获得的基础上,并且层次

越高,语义越复杂,涉及的领域知识越多。基于语义的图像检索一般指的是基于目标和高级语义的图像检索方法。

2.5 图像检索中的相关反馈技术

基于相关反馈的图像检索方法的基本思路是:在检索过程中,允许用户对检索结果进行评价和标记,指出结果中哪些是与查询图像相关的,哪些是不相关的,然后将用户标记的相关信息作为训练样本反馈给系统进行学习,指导下一轮检索,从而使得检索结果更符合用户的需要。

应用较广的一种相关反馈方法一方面修改查询向量,另一方面,利用反馈信息修改距离公式中各特征向量的权值,突出查询向量中较为重要的分量。

基于支持向量机(SVM)的相关反馈方法^[10,11]是在每次反馈过程中对用户标记的正例和反例样本进行学习,建立 SVM 分类器作为模型,并根据该模型进行检索。在每次反馈过程中,如果用户标记的样本是在特征空间中属性相近的图像,即作为支撑向量的样本,则因为距离较远的样本对 SVM 分类器没有影响,所以虽然用户标记的反馈图像有限,但是对于建立 SVM 分类器已经足够,可以有效地控制机器学习的推广能力。SVM 方法的主要难点在于核函数的选取,核函数的选取直接影响到分类器的泛化能力。但是,SVM 方法还有一个缺点,就是用于反馈的图像必须在标记边界图像和最优图像中作出折中选择。

基于 Bayes 准则的相关反馈方法的基本思想是根据用户反馈的信息进行统计判断。理论上讲,Bayes 算法与其他分类算法相比具有最小的出错率,然而实践中由于对应用假设的不准确性,缺乏可靠的概率估计。

3 基于内容图像检索中的效果评价

在基于内容的检索中,由于检索算法比较多,需要对各种算法的检索结果进行评判,以比较其优劣。从研究情况来看,对于系统的响应时间和吞吐率的评价论述的较少,对检索效果评价更多地放在对检索结果的正确与否,主要使用的是查准率(precision)和查全率(recall)两个指标。

$$\text{查全率} = \frac{\text{有关联的正确检索结果}}{\text{所有有关联的结果}} \times 100\%$$

$$\text{查准率} = \frac{\text{有关联的正确检索结果}}{\text{所有检索结果}} \times 100\%$$

用户在评价算法的时候,可以预先选定含有特有目标的图像作为一组相关的图像,然后根据返回的结果计算查全率和查准率。查全率和查准率越高,说明该检索算法的效果越好。查全率反映系统检索相关图像的能力,而查准率则反映系统拒绝无关图像的能力。

4 研究中的主要问题及研究方向

通过对基于内容检索的图像数据库的主要研究技术的分析以及对目前大量的实验原型系统的广泛调查,支持基于内容检索的图像数据库研究中仍存在如下主要问题:

(1)基于图像色彩特征的索引存在的主要问题是人对色彩特征的视觉感知方面考虑的仍然不够。

(2)基于图像纹理特征的索引目前存在的主要问题是各种方法所选择的纹理特征集依赖于具体的纹理图像。

(3)对于形状特征的图像检索,形状边界的自动提取一直是困扰图像处理领域多年的难题,在目前的检索系统中大多采用手工勾勒的方式,形状特征提取是一件非常繁重的工作,对于大批量图像数据此问题将显得更为突出。

(4)目前流行的大多算法,不论是基于那种特征都有计算量大这一缺点,现今的很多图像检索技术或多或少地采用了智能方法,但还远没有达到人们的要求。

(5)当前的很多检索算法也用到反馈技术,但大部分算法都是将其作为辅助手段,这样,反馈技术就不能很好地优化检索算法。

综上所述,可以看出基于内容的图像检索仍然是一个开放性的研究课题,其研究将涉及认知科学、人工智能、模式识别、图像处理、信息检索等多个领域。对于基于内容的图像检索中存在的重要问题,在图像处理和模式识别方面今后较长的时间内仍需要继续完善图像特征的提取、表达、相似性度量。目前国际上 MPEG 标准化组织正在制定的 MPEG-7 标准,其目标就是实现集高层语义特征和低层视觉特征的基于内容的多特征综合检索,今后研究的热点之一将是高层的基于语义内容的图像检索。寻找快速智能的检索算法和如何设计用户的反馈方式并充分反馈信息也是今后一个重要的研究方向。

参考文献:

- [1] Flickner M. Query by image and video content[J]. The QBIC System IEEE Computer, 1995, 28(9): 23-32.
- [2] Pentland A P, Picard R, Sclaroff S. Photobook: Content-based manipulation of image databases[J]. International Journal of Computer Vision, 1996, 18(3): 233-254.
- [3] 李向阳, 庄越挺, 潘云鹤. 基于内容的图像检索技术与系统[J]. 计算机研究与发展, 2001, 38(3): 344-354.
- [4] Zabih G R, Miller J. Comparing images using color coherence vectors[C]//In: Proc of ACM Intern Conf Multimedia. Boston, MA: [s. n.], 1996.

(下转第 131 页)

```

<p> 学生 信息 </p>
<p>学校 <input name="text" type="text"></p>
<p>院系 <input name="text2" type="text"></p>
<p>专业<input name="text3" type="text"></p>
<p>
  <input type="submit" name="Submit" value="提交">
  <input type="submit" name="Submit2" value="重填">
</p>
</form>

```

图 4 是根据该页面的表单代码逆向得到的任务模型。在该页面的代码中, input 为输入任务。表单中选择按钮 submit 和 reset 分别表示提交信息(send form)和重填(reset)两个任务。

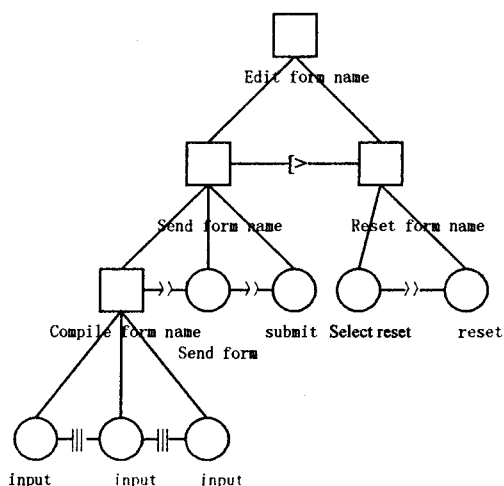


图 4 任务模型

笔者对 Web 系统的逆向工程解决方案还存在一些问题。如该方案对站点部署的逆向还只限于静态网站,没有涉及动态网站的逆向。这些都还要在以后的工作中做深入研究。

4 总 结

对 Web 系统维护中的逆向工程技术进行了分析,对软件维护当中的问题,逆向工程基本方法进行了简要介绍。最后用一个 Web 页面逆向的例子提出了针对 Web 系统逆向工程的方法。由于 web 系统应用的越来越广泛,技术越来越复杂,因此也越来越有必要对 Web 系统逆向工程进行更深一层次研究和发展。

参考文献:

- [1] Bennet K H. The Software Maintenance of large Software Systems: Management, Methods and Tools[M]. [s. l.]: Elsevier Science Publishing Co., Inc., 1990: 88-98.
- [2] 普瑞斯曼. 软件工程 - 实践者的开发方法[M]. 第 5 版. 黄柏素, 梅 宏译. 北京: 机械工业出版社, 1999.
- [3] Biggerstaf T J. Design Recovery for Maintenance and Reuse[J]. IEEE Computer, 1989, 22(7): 36-49.
- [4] 张 亮, 赵文耕, 彭 鑫, 等. 一种支持时序图恢复的逆向工程方法[J]. 南京大学学报, 2005, 41(10): 386-387.
- [5] Elliot C, James C. Reverse Engineering and Design Recovery: a Taxonomy[J]. IEEE Software, 1990, 7(1): 13-17.
- [6] 刘宗田. 软件维护与逆向工程评述[J]. 计算机软件与应用, 1995(1): 1-8.
- [7] Jamesrunbaugh, Javobson J, Booch G. UML 参考手册[M]. 第 2 版. UMLCHINA 译. 北京: 机械工业出版社, 2005.
- [8] Merlo E. Reverse Engineering[J]. IEEE, 1993, 41(3): 171-173, 176-178.
- [9] Milner R. Communication and Concurrency[J]. IEEE Software, 1989, 37(4): 268-280.
- [10] Estiévenart F, Francois A, Henrard J. A tool supported method to extract data and schema from web sites[C]//Proc of the 5th IEEE Int'l Workshop on Web Site Evolution. New York: IEEE Press, 2003: 3-11.

(上接第 128 页)

- [5] Stricker M A, Orengo M. Similarity of color images[C]//In: Proc of SPIE: Storage and Retrieval for Image and Video Databases III. San Jose, CA: [s. n.], 1995: 381-392.
- [6] Hsu W, Chua T S, Pung H K. An integrated color - spatial approach to content - based image retrieval[C]// In: Proc ACM Multimedia '95 Conference. San Francisco: [s. n.], 1995: 305-313.
- [7] Smith J R, Chang S F. Tools and techniques for color image retrieval[C]//In: Proc of SPIE: Storage and Retrieval for Image and Video Database. San Jose, CA: [s. n.], 1996: 426-437.
- [8] 章毓晋. 图像理解与计算机视觉[M]. 北京: 清华大学出版社, 2000.
- [9] Mehre B M, Anhalii M K, Lee W F. Shape Measures for

Content Based Image Retrieval: a Comparison[J]. Information Processing and Management, 1997, 33(3): 319-337.

- [10] ZHANG Lei, LIN Fu-zong, ZHANG Bo. Support vector machine learning for image retrieval[EB/OL]. International Conference on Image Processing. (2001-08-16)[2006-12-15]. <http://research.microsoft.com/users/leizhang/Paper/ICIP01.pdf>.
- [11] HONG Peng-yu, TIAN Qi, HUANG T S. In corporate support vector machines to content - based image retrieval with relevant feedback[EB/OL]. Proceedings of International Conference on Image Processing. (2000-03-10)[2006-12-15]. <http://www.ifp.uiuc.edu/~qitian/e-paper/icip00svm/icip00-svm.pdf>.