

文章编号:2095-0365(2013)04-0071-06

剪纸艺术多媒体交互平台中的 图像检索技术研究

韩立华¹, 王晓芬², 王玉梅³

(1. 石家庄铁道大学 继续教育学院, 河北 石家庄 050043;

2. 石家庄铁道大学 建筑与艺术学院, 河北 石家庄 050043;

3. 石家庄铁道大学 图书馆, 河北 石家庄 050043)

摘 要:利用信息技术手段保护非物质文化遗产资源是目前研究的热点之一,图像检索技术的研究是构建剪纸艺术多媒体交互平台的关键环节。分析了基于文本的图像检索(TBIR)和基于内容的图像检索(CBIR)的思路及算法,并在剪纸艺术多媒体交互平台中进行了实现。深入对比了两者的优缺点,探讨了两种检索方式相结合的可能途径,在此基础上提出利用自动语义标注技术解决图像检索中存在的“语义鸿沟”问题,有利于提高图像检索的效率和准确率。

关键词:剪纸艺术;基于文本图像检索;基于内容图像检索;特征提取;语义标注

中图分类号:TP37 **文献标识码:**A

我国民间剪纸艺术是世界级非物质文化遗产之一,为了更好地对其进行数字化保护与传承^[1],构建了基于Web的河北剪纸艺术多媒体交互体验平台,对散落民间的剪纸作品进行搜集、分类整理、数字化、矢量化、建库、检索以及展示等,目前已完成10大类75小类约8000余幅剪纸作品的数字化和入库。面对日益庞大的剪纸数据库,为使用户快速找到自己感兴趣的作品,研究和开发高效率、人性化的剪纸图像检索系统势在必行。

图像检索是指根据用户提交的查询请求,从图像数据库中提取与查询相关的一幅图像或图像集合,目前图像检索技术主要有两种^[2]:基于文本的图像检索TBIR(Text Based Image Retrieval)和基于内容的图像检索CBIR(Content Based Image Retrieval)。

一、基于文本的图像检索技术(TBIR)

基于文本的图像检索研究开始于20世纪70年代末,它是在对图像进行文本标注的基础上,对图像进行基于关键字的检索^[3]。其基本步骤是先对图像文件建立相应的关键词或描述字段,通过对图像的名称、编号、内容描述、图像大小、来源、作者、创建时间、存储地点等关键性的信息采用自动标引或进行人工注释,进行图像的文本特征抽取,建立图像索引数据库,然后按全文数据库管理,采用全文数据库检索方法。该方法实质是把图像检索转换为对与该图像对应的文本检索。

(一)建立TBIR索引数据库

TBIR技术中最关键的是索引数据库的建立,

收稿日期:2013-08-30

基金项目:2011年国家社科基金项目(11BG078);2013年国家自然科学基金项目(61304176)

作者简介:韩立华(1978-);女,副教授,研究方向:多媒体技术、图像检索、数据库系统开发。

建立索引数据库首先需要提取图像的有关信息,图像的文件信息可以自动识别与提取,如图像格式、大小、创建时间等,但图像内容信息的提取不太容易,一般有两种方法:人工输入与基于上下文环境的自动识别。人工输入方式建立索引数据库与早期文本的人工标引和分类是相同的,它先由专业人员负责选择图像,然后对每幅选定的图像内容进行描述,给出关键词,逐个图像进行审核和标引,它的查准率是相当高的,但它的查全率较受限制,其检索范围仅限于人工标注完成的部分。基于上下文环境的自动识别是根据图像所处的上下文 Web 环境来判断图像主题内容,目前对 Web 文档的主题提取算法不少,如著名的 PageRank 算法、HITS 算法^[4]等,结合这一成熟技术可以简化图像主题人工标引的消耗,百度、谷歌的大部分图像库就用这种方式建立的。在本课题的研究中,部分典型代表性剪纸图像采用了人工标注的方式,其他图像可采用这种自动识别方式进行全站范围检索。

(二) TBIR 系统实现

建立 TBIR 索引数据库后,图像检索系统的实现就变得较为简单,考虑任意多个条件的“与”和“非”随意组合,通过 SQL 语句的多项连接,实现了如图 1 所示的 TBIR 系统。



图 1 基于 TBIR 的剪纸艺术检索平台

(三) TBIR 特点

TBIR 的优点主要有两个:一是技术相对成熟,易于实现;二是查准率比较高,能满足大多数查询需求。但同时基于文本的检索存在着两大困

难,特别是当图像数量非常大时更为突出:其一,内容丰富的图像特征难以用文本描述全面表达,如剪纸图像中的纹样形状、多彩颜色、特殊纹理等,而且由于图像内容的丰富性加上用户的兴趣点和理解的不同,导致内容描述的建立具有一定的主观性,由此带来内容标注上的歧义;其二,文本描述难以实现基于图像视觉特征的相似性检索,如查询颜色相近或形状类似的剪纸图样。此外,文本描述一般需要专业分析和手工输入,效率较低,由于数据规模不断膨胀,人工标注的开销越来越大,难以满足大容量数据库的要求。

二、基于内容的图像检索技术(CBIR)

基于内容的图像检索(Content-Based Image Retrieval, CBIR),是指利用图像本身的特征(颜色、纹理、形状等)作为索引,克服了上述基于人工标注的检索系统缺点。典型的 CBIR 系统包括 QBIC、VisualSeek、WebSeek、ImgRetr 以及百度识图等,这些系统利用从图像中提取的数字特征(也称视觉特征)来比较图像的相似性^[5]。通过计算机自动提取图像库和示例图像的颜色、形状、纹理、结构、位置等特征信息进行比对,找出与用户提交的图像在某方面相类似的结果。在剪纸多媒体交互平台中,例如查找一幅图是否已经加入数据库,查找相似颜色的图像,查找具有某些典型纹理的图像以及查找具有某些特定纹样或形状的剪纸图像等需求都可以借助于 CBIR 来实现。

(一) CBIR 系统架构

基于内容的图像检索系统一般由输入模块、数据库、查询模块和检索模块等组成,其结构与各部分的功能如图 2 所示。

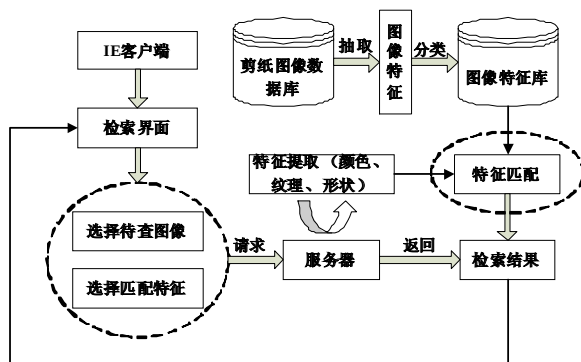


图 2 CBIR 系统架构

(二) CBIR 关键技术

CBIR 的关键技术包括颜色、纹理及形状等底层特征的提取、特征相似度测量等。

1. 颜色特征提取

颜色特征由于其计算简单、较好的鲁棒性以及几何变换的不变性成为机器可自动提取的图像内容中最重要的特征^[6]。用于检索的图像特征在颜色方面有直方图法、累积直方图法、分块直方图法、颜色聚类法、主色调法、颜色矩和颜色集等,本系统采用了基于颜色—空间信息的图像分块直方图特征提取方法,该方法首先选用符合人类视觉特性的色彩空间模型 HSV,并对 HSV 空间进行非等间隔量化得到 72 种代表颜色,然后构造一维特征矢量。按照量化级,把 3 个颜色分量合成为一维特征矢量 $G=9H+3S+V$, G 为 72 级的一维直方图。然后对图像空间按照矩形重叠分块策略进行划分,在 HSV 颜色空间中,统计各个分块区域内的 72 维颜色直方图,得到一个 5×72 的二维颜色—空间信息直方图。最后计算出各分块的权值 w_i ,配合得到的 5×72 二维颜色—空间直方图即可作为提取的特征进行检索。

2. 纹理特征提取

纹理特征是一种不依赖于颜色或亮度的反映图像中同质现象的视觉特征,它是所有物体表面共有的内在特征。纹理特征主要包括粗糙度、方向性、对比度以及规则性。在纹理特征检索方面一般有统计分析法、频谱分析法、结构分析法等,本文主要采用 Gabor 小波分析的方法^[7]来提取图像的纹理特征。对一幅给定的 $P \times Q$ 大小的图像 $I(x, y)$,其离散 Gabor 小波变换为:

$$G_{nm}(x, y) = \sum_s \sum_t I(x-s, y-t) \psi_{nm}^*(s, t) \quad (1)$$

式中, s, t 是滤波器模板大小变量; ψ_{nm}^* 是 ψ_{nm} 的共轭复数。它由基函数 $\psi(x, y)$ 经过伸缩和旋转得到:

$$\psi(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right] \cdot \exp(j2\pi Wx) \quad (2)$$

式中, W 称为级向中心频率; $\psi(x, y)$ 是经过复数正弦函数调制的 Gaussian 函数。对图像从不同方向和尺度进行 Gabor 变换后,得到一系列系数:

$$E(m, n) = \sum_x \sum_y |G_{nm}(x, y)| \quad (3)$$

由变化系数计算出的均值 μ_{nm} 和标准方差 σ_{nm} 可以作为图像的纹理特征:

$$\mu_{nm} = \frac{E(m, n)}{P \times Q}, \sigma_{nm} = \sqrt{\frac{\sum_x \sum_y (|G_{nm}(x, y)| - \mu_{nm})^2}{P \times Q}} \quad (4)$$

采用 6 个方向, 5 个尺度的 Gabor 滤波器,得到特征向量为

$$g = [\mu_{00}, \sigma_{00}, \mu_{01}, \dots, \mu_{45}, \sigma_{45}] \quad (5)$$

3. 形状特征提取

形状是图像的重要视觉特征之一。目前,基于形状特征的检索主要是围绕着从形状的轮廓特征和形状的区域特征建立图像索引,关于前者的描述主要有:直线段描述、样条拟合曲线、傅立叶描述子以及高斯参数曲线等;对于后者主要有形状的无关矩、区域的面积、形状的纵横比等。本课题采取基于不变矩的形状特征提取方法^[8],该方法通过 Canny 算子对图像进行边缘提取,计算用二值图像表示的物体 R 形状的 $p+q$ 阶中心矩 μ_{pq} 及图像的归一化中心矩 η_{pq} ,将这些二阶和三阶中心矩进行组合得到 $\varphi_1 \sim \varphi_7$ 个对平移、旋转和尺度无关性的不变矩,并将其作为描述形状的特征向量。

4. 相似度测量

常用的相似度度量方法是向量空间模型,即将视觉特征看作是向量空间中的点,通过计算两个点之间的接近程度来衡量图像特征间的相似度。本系统采用 Minkowski 距离及其加权变形的二次距离来计算特征之间的相似度,在一个 d 维的特征空间中,给定查询 $q = (q_1, q_2, \dots, q_d)^T$ 和图像 i 的特征 $x_i = (x_{i1}, x_{i2}, \dots, x_{id})^T$,加权的 Minkowski 距离为:

$$D_p(x_i, q, w) = \left(\sum_{j=1}^d w_j |x_{ij} - q_j|^p\right)^{1/p} \quad (6)$$

二次距离为

$$D^2(x_i, q, M) = (x_i - q)^T M (x_i - q) = \sum_j \sum_k m_{jk} (x_{ij} - y_j)(x_{ik} - q_k) \quad (7)$$

其中 $w = (w_1, w_2, \dots, w_d)^T$ 为权值向量; $M = [m_{ij}]$ 为一实对称矩阵。

(三) CBIR 模型实现

根据以上关键技术,分别采用颜色特征、纹理特征以及形状特征对剪纸图库进行基于图像内容的检索,建立了如图 3 所示的 CBIR 剪纸图像检

索模型系统。以“颜色特征选择”为例,当用户选取了一副本地图像文件后,系统自动提取其直方图,判断是一副红色的剪纸图像,则根据特征向量从特征库中查找主色调为红色的图片,并给出结果。



图 3 基于 CBIR 的剪纸艺术检索测试平台

(四) CBIR 特点

CBIR 的优点主要有四个:一是直接从图像内容中提取特征线索,无需通过图像的相关文本注释;二是特征提取和索引建立可由计算机自动实现,大大提高了检索效率;三是具有较强的交互性,即用户能够参与检索过程,评估和改进检索结果;其四,具有一定的客观性,其检索结果能突破关键词主观性标注的限制。基于图像视觉特征的 CBIR 系统的主要缺点是算法复杂,实现成本高,难以建立从底层特征到高层语义的对应关系。

三、TBIR 与 CBIR 对比分析及结合运用

(一) TBIR 与 CBIR 对比分析

TBIR 与 CBIR 各有优势和不足,表 1 在描述方式、技术实现、查准率、交互性等方面对比了两种检索技术。

表 1 TBIR 与 CBIR 的比较

比较项目	TBIR	CBIR
描述方式	手工标注关键词	自动提取特征
检索方式	精确匹配	相似度对比
技术实现	简单	复杂
应用范围	较广	较窄
前期工作量	很大	较小
查准率	较高	一般
检索效率	取决于图像标注的完备性	较高
用户交互	很少	有

TBIR 发展较早,技术成熟,应用广泛,而且符合人们熟悉的检索习惯,实现简单,因此仍然是现在大多数系统的主要检索方式。但由于图像在手工标注时有太多主观性,缺乏统一标准,费时费力。与之相反,CBIR 主要利用可视化特征来标引图像,具有一定的客观性,而且可以利用相关反馈指导用户逐步逼近真实检索意愿,但是 CBIR 实现算法较为复杂,难以建立从底层图像特征到高层语义的联系。

(二) TBIR 与 CBIR 结合运用

从以上分析可知 CBIR 和 TBIR 各自有优缺点,如果能将二者结合起来取长补短,把高层文字描述和低层图像特征组合利用,图像检索系统的性能将更加优化,功能亦更人性化。TBIR 和 CBIR 的结合可以有两种情形:

(1)简单组合。即检索系统同时具备 TBIR 和 CBIR 的功能,例如用户通过关键词开始一个检索,而返回结果后再选择可以作为查询样图的图像进行基于内容特征的检索,然后由用户相关反馈筛选出符合意愿的结果;或者用户先进行图像内容的检索,在检索结果中可以利用关键词再缩小范围,逐步找出符合的图像。这种简单组合方式是对检索结果的筛选处理,虽然能弥补各自的不足,但仍然需要繁重的手工标注工作为其先期基础。

(2)自动语义标注。早期的 CBIR 系统由于仅注重依靠图像低层特征进行检索,无法解决低层特征与高层语义之间的“语义鸿沟”问题,因此人们的研究重点转向了基于自动语义标注的图像检索^[10],它也是 TBIR 与 CBIR 的结合,只不过是图像特征描述上采用了更为智能的自动标注语义技术,既避免了人工标注的效率低下、主观性强的缺陷,同时能通过对图像视觉特征的分析来提取高层语义用于表示图像的含义,一定程度上能够解决“语义鸿沟”问题,是目前较为理想的图像检索技术。自动语义标注通过视觉特征提取、图像分割、图像识别、对象空间关系分析等步骤建立语义自动标注数据库^[11],结合人工辅助标注,形成图像的综合语义描述,进而可以实现基于文本和内容的图像检索。语义提取的过程如图 4 所示。

四、结语

TBIR 图像检索技术已经在各类大中小型管

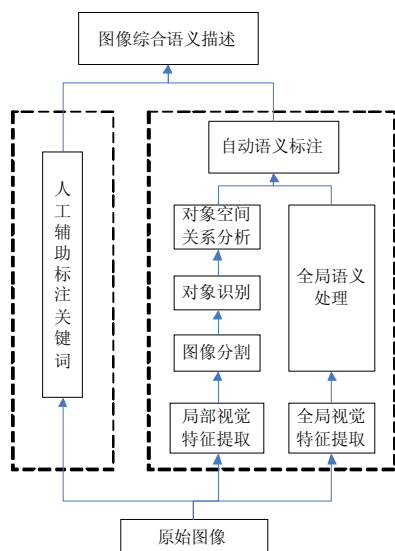


图4 图像自动语义提取过程

理系统中广泛应用,CBIR 图像检索技术目前流行的算法比较多,但成熟的商业化应用并不多见,许多算法还仅停留在实验和改进阶段。基于两者各自的优缺点,将两者结合的检索技术是今后研究和应用的主流方向。本文在传统 TBIR 的技术基础之上研究并引入了 CBIR 技术,在剪纸艺术交互平台中对两者均进行了实现,达到了用户基本应用需求,但要想大幅度提高剪纸艺术图像的检索效率,满足用户检索图像的多样化需求,必须在传统基于文本的图像检索技术基础上,加强对基于内容的图像检索技术的研究,特别是应在自动语义标注前提下寻求一种与人的感知更为符合的图像语义特征描述模型。在此基础上,研究更为有效的算法性能评价准则和全面的图像数据库测试,从而将底层特征与高级语义更好的结合,实现图像语义的自动标注,这将是本课题下一步重点研究的内容。

参考文献:

- [1]彭冬梅,刘肖健,孙守迁.信息视角:非物质文化遗产保护的数字化理论[J].计算机辅助设计与图形学学报,2008,20(1):117-123.
- [2]周明全,耿国华.基于内容图像检索技术[M].清华大学出版社,2007.
- [3]Lim J H, Jin J S. Image indexing and retrieval using visual keyword histograms[J]. Proc. IEEE conference on ICME, 2002(1):213-216.
- [4]郑莉霞.基于文本的 Web 图像检索技术研究[D].西宁:广西大学,2007.
- [5]刘颖,范九伦.基于内容的图像检索技术综述[J].西安邮电学院学报,2012,17(2):1-7.
- [6]郭士会,杨明.基于颜色的图像检索方法的研究[J].西南大学学报:自然科学版,2012,34(1):128-133.
- [7]DUNN D, HIGGINS W E. Optimal Gabor Filters for Texture Segmentation[J]. IEEE Transactions on Image Processing, 1995,4(7):227-237.
- [8]韩立华,王学军,王晓芬.多特征融合及 SVM 相关反馈技术在教育资源图像检索中的应用[J].河北科技大学学报,2010,31(6):240-244.
- [9]穆莹,王学军.基于提升小波变换的医学图像融合算法[J].石家庄铁道大学学报:自然科学版,2010(4):58-60,71.
- [10]Tsai C, Hung C. Automatically annotating images with keywords: a review of image annotation systems [J]. Recent Patents on Computer Science, 2008(1):55-68.
- [11]吴楠,宋方敏.一种基于图像高层语义信息的图像检索方法[J].中国图像图形学报,2006,12(11):1774-1780.

(下转第 80 页)

- [2]汪丽君,寒梅. 杂交与共生——建设节约型社会背景下对东西向住宅适应性的优化设计研究[J]. 建筑学报, 2010(5): 30.
- [3]李琳. 紧凑城市中“紧凑”概念释义[J]. 城市规划学刊, 2008(3): 41-45.
- [4]韩笋生,秦波. 借鉴“紧凑城市”理念,实现我国城市的可持续发展[J]. 国外城市规划, 2004(6): 263-268.
- [5]窦强. 新形势下关于东西向住宅的讨论[J]. 住区, 2012(02): 128-132.
- [6]许爱兰,刘煜. 多层集合住宅中的“空中四合院”设计初探[J]. 低温建筑技术, 2008(4): 30-32.

To Explore the Residential Modes for the Aged with Chinese Characteristics: On the Residential Mode of Compact

SUN Hong-yan, SHAN Cheng-li

(School of Architecture & Art Design, Hebei University of Technology, Tianjin 300401, China)

Abstract: With the increasing aged tendency of population and the increasing shortage of land resource, it is needed to explore the residential modes for the aged with Chinese characteristic at the present stage of China. Through defining the concept of compact resident and reading the relevant theories, a compact residential mode for the aged which is suitable the current situation in China was proposed, moreover, the overall layout, the construction of the monomer solution and the feasibility of the residential mode were studied deeply, and the range of the servicing object was defined. The compact residential mode was put forward to save land resource, improve the resident environment, inherit the traditional culture of China and improve the resident quality for the aged.

Key words: residential modes of settlement for the aged; compact; Chinese characteristic

(上接第 75 页)

Research on Image Retrieval Technique in Paper-cut Art Multimedia Interactive Platform

HAN Li-hua¹, WANG Xiao-fen², WANG Yu-mei³

(1. College of Continuing Education, Shijiazhuang Tiedao University, Shijiazhuang 050043, China;

2. College of Architecture and Art, Shijiazhuang Tiedao University, Shijiazhuang 050043, China;

3. Library, Shijiazhuang Tiedao University, Shijiazhuang 050043, China)

Abstract: Using information technology to protect intangible cultural heritage resources is one of the hot topics of current research, and the research on image retrieval technology is critical in the process of constructing paper-cut art multi-media interactive platform. The ideas and algorithms of text-based image retrieval (TBIR) and content-based image retrieval (CBIR) were analyzed which were implemented on the paper-cut art multi-media interactive platform. The advantages and disadvantages of TBIR and CBIR were compared with and the approach of combining the two kinds of retrieval methods were discussed, based on which the solution using automatic semantic annotation technology to solve the “semantic gap” existing in the image retrieval problem was proposed to improve the efficiency and accuracy of image retrieval.

Key words: paper-cut art; TBIR; CBIR; feature extraction; semantic annotation