

基于形状特征的图像检索技术研究

付 玮, 曾接贤

(南昌航空大学 无损检测技术教育部重点实验室, 江西 南昌 330063)

摘 要:随着图像信息数量的飞速增长,人们迫切需要对大量的图像信息进行快速、有效的检索。因此,研究者们提出了许多图像检索新技术。阐述了图像检索技术的发展过程;分析了基于形状特征的图像检索技术研究现状和局限性;着重从特征提取、高层语义和形状特征的关联、高维索引技术等方面研究了基于形状特征的图像检索的发展趋势。

关键词:形状特征;图像检索;研究进展

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2007)11-0228-05

Research of Shape - Based Image Retrieval

FU Wei, ZENG Jie-xian

(Ministry of Education Key Laboratory of Nondestructive Test,
Nanchang Institute of Aeronautical Technology, Nanchang 330063, China)

Abstract: With the rapid growth in the number of image information, it is urgent to retrieve a great deal image information effectively. Therefore, researchers put forward a lot of new image retrieval technology. Summarizes the image retrieval development firstly, and then analyzes research status and the localization of shape - based image retrieval. Lastly, its primary aim is to study in the future development direction of shape - based image retrieval from several aspects, such as feature extraction, high semantic with the connection of shape feature, high - dimension index etc.

Key words: shape feature; image retrieval; review and prospect

0 引 言

随着互联网和多媒体信息技术的飞速发展,图像信息正变得越来越重要;而且许多行业的需求对专用图像的检索提出了新的要求,例如多媒体数字图书馆系统、卫星遥感图像、医疗图像的分析和检索。图像数据库的日益庞大使得对图像内容的组织、表达、存储、管理、查询和检索面临前所未有的挑战。如何对大量的图像数据组织、建库以及对大容量的图像信息进行快速、有效的检索成为信息时代人们迫切需要解决的问题。1992年,基于内容的图像检索一词由 Toshikazu Kato^[1]提出,随后,国内外许多学者开展了图像检索技术的研究,使这方面的研究和应用得到长足的进展。随着图像检索技术的发展,1994年 Scasselati B^[2]提出了基于形状特征图像检索概念。介绍了图像检索技术

的发展过程,着重研究基于形状特征图像检索技术的现状和发展趋势。

1 图像检索技术发展过程

图像检索技术发展到现在大致经历了两个阶段:基于文本的图像检索和基于内容的图像检索^[3]。

1.1 基于文本的图像检索

基于文本的图像检索,可以追溯到 20 世纪 70 年代,图像数据研究者们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。其基本步骤是先对图像文件建立相应的关键字,并将图像的存储路径与该关键字对应起来,然后用基于文本的数据库管理系统来进行图像检索。该方法实质是把图像检索转换为对该图像对应的文本检索,Gudivada 等对该技术进行了较为全面的综述^[4]。

随着图像信息数量的飞速增长,基于文本标注的图像检索的问题就显现出来了:其一是文本描述是一种定性描述,当图像中包含多个物体时,难以用简短的描述全面体现其意义,难以表达图像的空间关系;其二是大量描述性文字的手工输入需要花费大量的人力物

收稿日期:2007-01-25

基金项目:国家自然科学基金(60675022);江西省自然科学基金(0311019);江西省教育厅科技计划项目

作者简介:付 玮(1975-),男,江西南昌人,硕士研究生,研究方向为计算机视觉与图像处理;曾接贤,教授,硕士,研究方向为计算机视觉、工程图学,计算机图形学等。

力,无法满足大型的多媒体信息库,如 Web 网络资源,手工处理是完全无法适应的;其三是图像的内容非常丰富,而人们对图像的理解具有主观性。由于理解的不同,不同的人对同一幅图像的文字描述可能不同。这种理解的主观性和描述的不确定性在图像检索过程中可能会导致误匹配。显然,直接采用传统的基于关键字的方法对图像数据进行检索不能满足人们的需要,于是就提出了基于内容的图像检索(CBIR, Content-based image retrieval)^[3]。

1.2 基于内容的图像检索

基于内容的图像检索是一门 20 世纪 90 年代兴起的新技术,因其直观、高效、通用等特点,近年来在国内外均是一个热门研究课题。CBIR 的基本思想是通过分析图像的视觉特征和上下文联系来进行检索。其中,图像内容是通过图像的特征来反映的。通常把图像的特征分为两大类,即低层物理特征(如:颜色、纹理、形状、轮廓、图像内容的空间关系等)和高层语义特征(是人们对图像内容概念级的反映,一般是对图像内容的文字性描述)。CBIR 突破了传统的基于关键字的表达方式检索的局限,直接对图像内容进行分析和特征提取,利用这些描述图像内容的特征建立索引。目前,国外许多科研机构都在进行相关的研究,并推出了以 IBM 的 QBIC^[5]和 MIT 的 PhotoBook^[6]为代表的一系列产品。

2 基于形状特征的图像检索

由于许多图像检索系统都把重点放在基于颜色或者纹理的方法上。但是对于某些图像来说,纹理和颜色信息不够丰富,如一些商标图像等,这时基于颜色和纹理的方法就无法满足检索需要,而必须从图像的形状着手。形状特征是图像的核心特征之一,图像的形状信息不随图像颜色的变化而变化,是物体的稳定特征。用形状特征区别物体非常直观,是人们分类不同图像的主要特征之一。因此,利用形状特征检索图像可以提高检索的准确性和效率。基于形状特征的图像检索^[4]主要是检测出目标的轮廓线或分割出目标的轮廓,并针对其进行形状特征的提取或直接针对图像寻找适当的矢量特征。形状描述应该在尽可能区别不同目标的基础上对目标的平移、旋转和尺度变化不敏感。目前,虽然已经提出了许多形状分析方法,但要将其有效地运用于图像检索上还有一些问题亟待解决,如算法的效率和复杂性,形状特征的提取与描述等。

2.1 基于形状特征的图像检索系统结构

基于形状特征的图像检索技术主要是利用图像中目标的形状特征对图像进行检索,这种检索过程融合

了图像处理、模式识别、计算机视觉和数据库等技术。Hongjiang Zhang^[7]提出了如图 1 所示的基于形状特征的图像检索结构框架。

在基于形状特征图像检索结构中,图像库中的图像经过“形状特征提取与描述”模块分析之后,得到每幅图像的特征向量,这些特征向量经过索引组织之后生成一个特征库。用户提交的查询也经过“形状特征提取与描述”模块分析,得到查询的特征向量 V_q 。 V_q 与特征库中的特征向量 V_i 进行相似性计算,计算的结果往往是一个反映 V_q 与 V_i 之间距离的数值,这个数值的大小反映了待查询图像与被检索图像之间的相似程度,最终输出检索结果。

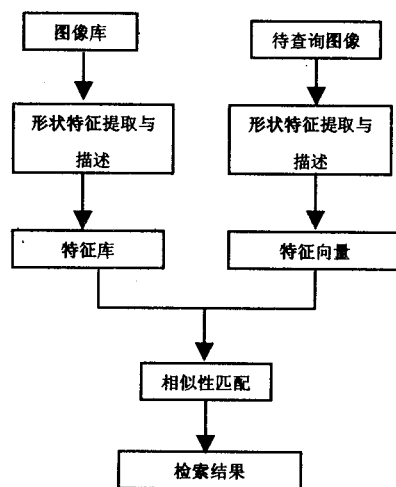


图 1 基于形状特征的图像检索系统结构

2.2 基于形状特征的图像检索技术研究现状

基于形状特征的图像检索通过提取形状特征来描述图像,一类是局部的特征,另一类是全局的特征。局部特征即针对图像中的目标进行分析。因此这类方法对图像分割算法有很大的依赖性,图像分割效果的好坏直接影响到后续的分析工作。对目标的形状进行分析有很多方法,主要分为两类:基于边界的方法和基于区域的方法,前者只利用形状的外部边缘,而后者利用形状的全部区域。基于边界的形状特征提取关键在于边缘检测的研究^[8~10],在提取边缘的基础上,定义边缘的特征描述,常见的有 Fourier 描述子、几何常量、多边形顶点数、孔径数、曲率等。基于区域的特征提取关键在于图像分割的研究,使用 7 个(或更多)不变矩作为目标区域的特征量进行图像匹配^[11]。

在形状描述方面,近年来提出的一些方法,主要包括边界特征法、傅里叶形状描述符法(Fourier shape descriptors)^[12]、几何参数法、形状不变矩法^[13]、有限元方法(Finite Element Method, FEM)、小波描述子等。边界特征法主要通过对边界特征的描述来获取图像的形状参数。其中,利用 Hough 变换^[14]检测平行直线方法

和边界方向直方图方法是经典方法。傅里叶形状描述符法基本思想是用物体边界的傅里叶变换作为形状描述,利用区域边界的封闭性和周期性,将二维问题转化为一维问题^[15]。在几何参数法研究中,B. Scasselati^[2]等人将代数矩、参数曲线距离、参数曲线和一阶导数距离、参数曲线和二阶导数距离、转角、曲率符号,以及改进的霍斯道夫距离等多种几何形状参数与人的视觉相似性判断进行了实验比较,结果表明,转角、曲率符号和代数矩对形状的描述与人的视觉特性最接近。其中,转角方法的结果最好,代数矩对像素占主要地位的目标描述能力很强,但对目标的轮廓不敏感,曲率符号对那些有中心主干且内主干向四周有突起的形状区别能力差。形状不变矩法是常用的形状描述方法,它利用目标所占区域的矩作为形状描述参数。Huilin Xiong 等提出了一种具有尺度、平移不变的自适应小波变换^[16],该方法的基本思想是:先求出图像的均值和方差,然后分别用均值和方差作为某一母小波的位移和尺度因子,这样形成的小波再作用于同一图像,变换后,小波系数具有尺度、位移不变性。有限元法^[6]定义了一个稳定性矩阵,用来描述物体上的每一个点与其它点之间的关系。这个稳定性矩阵的特征向量被称作特征空间的模合基。所有的形状都首先映射到这个特征空间,再在特征值的基础上计算形状相似性。类似于傅里叶描述符的思路,Arkin 等人提出了旋转函数,用来比较凹面或凸面多边形的相似性^[17]。

另外,F. Mahmoudi 等^[18]提出了一种边缘方向相关图来描述物体的形状。这种方法比较有效,但是处理时间很长。T. Bernier^[19]也进行了这方面的研究,他使用相对于物体质心的距离和角度来描述物体的轮廓点,这种方法是平移、尺度、旋转变换不变的。但是,这种描述方法对形变很敏感并且边缘检测结果较差。H. Nishida^[20]提出了一种对轮廓具有较好的抗形变的描述方法,但是,它的结构特征对噪声很敏感。

通常,形状匹配方法有:欧氏距离、街区距离、马氏距离等。在众多的形状匹配算法中,Chamfer 匹配方法吸引了不少研究者的兴趣。Barrow 等人首先提出了 Chamfer 比较法^[21],该方法能够以线性的时间复杂度比较两个图像的形状块集合。为加快匹配的速度,Borgerfos 提出了分层 Chamfer 匹配算法^[22]。这种匹配算法可以在不同的精确层次上进行,从粗糙到精确。尽管计算上述的形状特征并不复杂,但发明一种符合人们主观判断的形状相似度量算法还是一个有待解决的难题。J. Zhang^[23]使用形状空间方法来处理噪声和遮挡问题。同时,要在图像检索中充分使用形状特征,还必须要有稳健的图像自动分割通用算法。近年来,有

人提出了适用于处理大规模图片库的图像自动分割算法^[24]。

3 基于形状特征的图像检索技术发展趋势

未来的图像检索技术必定是以提高图像的查准率、查全率和查询速度为目标,因此,对基于形状特征的图像检索技术的研究还需要在下面几个方面进一步研究。

3.1 形状特征的有效提取与描述

特征提取是 CBIR 技术的关键。以往的研究大多从颜色、纹理和形状的模式匹配角度处理这个问题。在计算机视觉中,相对于颜色或纹理等底层特征而言,形状特征属于图像的中间层特征。形状特征作为刻画图像中物体和区域特点的重要特征,是描述高层视觉特征(如目标、对象等)的重要手段,而目标、对象对获取图像语义尤为重要。要把图像的底层特征与高层特征有机地结合起来,必须有好的形状特征描述与提取算法的支持。形状特征的有效表达必须以对图像中物体或区域的分割为基础。但是,现有的形状描述方法还不成熟,至今还没有找到与人的视觉特性一致的有关形状的确切数学描述,难以做到准确而鲁棒的自动图像分割和形状特征提取与匹配,使得图像识别或检索中的形状特征往往只能在某些特殊的应用场合中使用,或者需要借助某些领域知识来提高形状识别或检索的精度,如人的脸谱、虹膜和指纹等生物特征识别。此外,在主观上,由于人们对物体形状的平移、旋转和缩放不太敏感,合适的形状特征提取与匹配必须满足对平移、旋转和缩放无关,这给形状特征的相似性度量带来了难度,增加了算法的复杂性。

将形状特征提取和所涉及到的应用领域联系起来也有利于提高效率,因为不同应用领域的图像或图形会有一些特有的形状特征,例如黑白商标^[25]图案的区别主要在于图案形状,这样就必须根据不同的特点设计不同的匹配算法,这将涉及到相关领域知识。

在形状描述研究中还有一个很重要的问题就是怎样评价一个描述算法的好坏。对于形状描述算法的评价还没有一个统一的标准^[26]。MPEG-7 是仍在制定的多媒体内容描述标准,称为多媒体内容描述接口(Multimedia Content Description Interface)。其目标是制定一组标准的描述符及其描述模式,将内容描述与媒体内容结合,使用户快速准确地进行检索。MPEG-7 的处理链包含有三个部分:特征提取、标准描述和检索引擎。特征的自动分析和提取对 MPEG-7 是至关重要的,抽象程度越高,自动提取也越困难,而且不是都能够自动提取的。这个标准中已经提出了几个形

状描述符的评价标准^[27,28],这些准则是:检索精确(good retrieval accuracy)、特征压缩(compact features)、通用性(general applicaton)、低复杂度(low computation complexity)、鲁棒的检索性能(robust retrieval performance)以及由粗到精的描述(hierarchical coarse to fine representation)。在研究中一般使用旋转不变性、平移不变性、尺度变换不变性作为形状描述的一个基本的评价标准,另一个评价标准是查准率,即用某种描述符对形状进行识别的效果。总之,MPEG-7 的研究成果将给基于形状特征图像检索技术的研究带来巨大的影响。

3.2 高层语义与形状特征的关联

以形状特征来描述图像一直是图像检索领域的热点,新颖的算法层出不穷,这确实提高了图像检索的精确度和速度,但与用户的要求还有差别。对用户来说,判断两幅图像是否相似具有主观因素,它根据的是人的感知能力。另外,通过选择待查询图像,比较两者特征来寻找相似图像的方法固然比较简单,但是,概念性的描述更利于准确表达需求。实际检索中用户常用特定的概念来描述物体的内容,如用“楼房,汽车,花”等来表达检索对象,这属于图像数据高层语义的范畴。当前研究形状检索的重点和难点集中在如何将提取的形状特征和图像内容所表示的语义特征间建立很好的联系^[29],也就是说,由于我们所拥有的形状特征并不能很好地体现出图像真正的语义信息,以致于检索的结果往往不能令人满意。Zhou 和 Huang 在他们的文章中就提出了语义层的思想,并提出了一种使用统计学方法将图像归类并自动产生映射辞典的算法^[15]。

与许多期望获得对目标精确测量结果的图像分析工作不同,在图像检索中,获得精确的测量结果并不总是最重要的目标。这是因为图像检索具有一定的主观性(包括主观的需求和主观的判定结果),分割的目的更主要的是辨识感兴趣的目标区域。由于类似的原因,基于形状特征的图像检索也常仅隐式地对图像进行分割。

目前,对图像的形状特征方面的研究已经做了很多工作,如果能够将形状特征与高层语义较好地关联起来,就可以实现计算机自动提取图像的语义了。但在高层语义与形状等底层特征的关联方面仍存在着以下二个问题^[30]:

- ①如何从视觉的形状特征中提取出语义特征;
- ②如何判断不同注释的图像之间是否有相似的语义。

3.3 基于形状特征图像检索的索引技术

对于大型的图像数据库,如何在系统建立的庞大

索引库中检索是一个难题。因为在基于形状特征的检索过程中,特征矢量维数可能高达上百维,大大多于一般数据库的索引能力,另一方面,由于现有的多维数据的索引技术是针对传统数据库设计的,所有相似索引的基本思想都是根据特征数据集在特征空间中的分布特性,数据切分成子数据集,并对子数据集建立描述,如 SS 树, R * 树等^[31]。但是,由于图像索引本身的近似性、无序性以及图像属性的多重性等特点,传统数据库的许多索引结构都无法用来描述图像的索引组织。因此,非常需要研究新的索引结构和检索算法,以实现快速检索。

目前,提高索引的效率有两种方法:缩减特征向量的维数和改进索引算法^[5]。但是,维数的降低不可避免地会造成信息的丢失,为更好地支持多方面的图像检索,必须研究更加有效的高维索引技术。由于图像的检索是相似性匹配,因而可以采用聚类方法来构造索引。聚类是在研究大量图像特征的基础上通过学习产生出类别,然后按此类别再对未处理的图像进行分类。Charikar 等人讨论了聚类技术,并由此提出了动态图像分类^[5]。

4 结 论

基于形状特征的图像检索是目前 CBIR 中的难点。虽然由于形状特征本身的特殊性使其具体实现有一定难度,但形状特征本身良好的检索性能又在 CBIR 中起着重要的作用,所以进一步加强对基于形状特征的图像检索的研究势在必行。目前还有很多探索性研究,如将其他领域的思想和成果用于基于形状特征的图像检索中,如利用数学形态学、神经网络及遗传算法等进行基于形状特征的图像检索,值得进一步探讨。同时与基于形状特征的图像检索紧密相关的图像边缘检测及分割的进一步发展也将会促进其发展。

参考文献:

- [1] Kato T. Database architecture for content - based image retrieval[J]. SPIE, 1992(1662):112 - 123.
- [2] Scasselati B, Alexopoulos S, Flickner M. Retrieving images by - 2D shape: a comparison of computation methods with human perceptual judgments[J]. SPEF, 1994(2185):2 - 14.
- [3] Rui Yong, Huang T S, Chang Shih - Fu. Image Retrieval Current Techniques, Promising Directions and Open Issues [J]. Journal of Visual Communication and Image Representation, 1999, 10(3):39 - 62.
- [4] Gudivada V N, Raghavan V V. Content - Based Image Retrieval System[J]. IEEE Computer, 1995, 9: 18 - 22.
- [5] IBM's Query By Image Content. [EB/OL]. 2003 - 06.

- http://www.qbic.almaden.ibm.com
- [6] Pentland A P, Picard R, Sclaroff S. Photobook: Content - based manipulation of image databases[J]. International Journal of Computer Vision, 2002, 18(3): 233 - 254.
- [7] Zhang Hongjiang. Relevance Feedback in Content - based Image Retrieval [EB/OL]. 2001 - 08. http://research.microsoft.com/asia.
- [8] Hou Z, Koh T S. An adaptive contour model for highly Irregular boundaries [J]. Pattern Recognition, 2003 (34): 323 - 331.
- [9] 章毓晋. 图像分割[M]. 北京: 科学出版社, 2001.
- [10] Lin H J, Kao Y T. A prompt contour detection method[C]// International Conference on the distributed Multimedia Systems. Proceedings of 6th Joint Conference on Information Sciences. USA: [s. n.], 2002: 15 - 18.
- [11] 章毓晋. 基于内容的视觉信息检索[M]. 北京: 科学出版社, 2003.
- [12] Zhang Dengsheng, Lu Guojun. Generic Fourier descriptor for shape - based image Multimedia and Expo[J]. IEEE, 2002 (8): 425 - 428.
- [13] Hu M K. Visual pattern recognition by moment invariants [J]. IRE Trans on Information Theory, 1962(8): 179 - 187.
- [14] Sim L C, Schroder H. Fast line detection using major line removal morphological Hough transform[J]. Neural Information Processing, 2002(4): 18 - 22.
- [15] Zhou Xiang Sean, Huang Thomas S. CBIR: From Low - level Features to High - level Semantics[C]//Proc: SPIE Image and Video Communication and Processing. [s. l.]: [s. n.], 2000.
- [16] Daubechies I, Sweldens W. Factoring wavelet transforms into lifting steps[J]. Fourier Anal, 1998, 4(3): 245 - 267.
- [17] Gart J E, Mehrotra R. Shape Similarity - Based Retrieval in Image Database Systems[J]. SPIE, 1992(1662): 2 - 8.
- [18] Mahmoudi F, Shanbehzadeh J, Eftekhari - Moghadam A M. Image retrieval based on shape similarity by edge orientation autocorrelogram[J]. Pattern Recognition, 2003 (36): 1725 - 1736.
- [19] Bernier T, Landry J A. A new method for representing and matching shapes of natural objects[J]. Pattern Recognition, 2003(36): 1711 - 1723.
- [20] Nishida H. Structural feature indexing for retrieval of partially visible shapes[J]. Pattern Recognition, 2002(35): 55 - 67.
- [21] Barrow H G. Parametric correspondence and chamfer matching[C]//in Proc. 5th Int. Joint Conf. Artificial Intelligence. Cambridge, MA: [s. n.], 1997: 37 - 42.
- [22] Borgerfos G. Hierarchical chamfer matching: A parametric edge - matching algorithm[J]. IEEE Trans Patt Recog And Mach Intell, 1998, 10(6): 849 - 865.
- [23] Zhang J, Zhang X, Krim H, et al. Object representation and recognition in shape spaces [J]. Pattern Recognition, 2003 (36): 1143 - 1154.
- [24] Ma W Y, Manjunath B S. Edge flow: a framework of boundary detection and image segmentation[C]// Proc. Computer Vision and Pattern Recognition. Puerto Rico: [s. n.], 1997: 744 - 749.
- [25] Zhao Tong, Tang H L, Horace H S, et al. Content - Based Trademark Recognition and Retrieval Based on Discrete Synergetic Neural Network [M]. USA: Idea Group Publishing, 2002: 58 - 72.
- [26] Loncaric S. A Survey of Shape Analysis Techniques[J]. Pattern Recognition, 1998, 31(8): 983 - 1001.
- [27] Zhang Dengsheng, Lu Guojun. Review of Shape Representation and Description Technique[J]. Pattern Recognition, 2004 (37): 1 - 19.
- [28] Kim H, Kim J. Region - based shape descriptor invariant to rotation, scale and translation[J]. Signal process Image Commun, 2000(16): 87 - 93.
- [29] Jain R. Semantics in Multimedia Systems[C]//The 9th International Conference on Multi - media Modeling. USA: [s. n.], 2003: 7 - 10.
- [30] 杭燕, 杨育彬, 陈兆乾. 基于内容的图像检索综述[J]. 计算机应用研究, 2002(9): 12 - 14.
- [31] Zhou X S, Huang T S. Relevance Feedback in Content - Based Image Retrieval: Some Recent Advances Information Science - Application [J]. IEEE Transactions on Multimedia, 2002 (3): 129 - 137.

(上接第 227 页)

规范性、Agent 间协作模型、通信机制等方面, 还有待继续深入研究。

参考文献:

- [1] 黄雨田. 高校教师教学评价系统的研究与实现[D]. 太原: 太原理工大学, 2005.
- [2] 邱晖, 孙正顺. 知识管理系统的构建及其策略[J]. 计算机工程与应用, 2001(1): 52 - 54.
- [3] 郭红, 陈锐, 胡黎明. 基于 XML、CORBA 和 Agent 技

术的集成模型研究[J]. 小型微型计算机系统, 2003, 24 (9): 1646 - 1649.

- [4] Tang Chao, Xu Li Da. An agent - based geographical information system[J]. Knowledge - Based System, 2001, 14: 233 - 242.
- [5] 张礼华, 卢道华. 基于多 Agent 的网络课程协作学习平台的构建与研究[J]. 计算机技术与发展, 2006, 16(9): 126 - 128.