

基于最小外接圆补偿机制的形状图像检索

束 鑫^{1,2} 吴小俊¹ 潘 磊¹

(江南大学信息工程学院 无锡 214122)¹ (江苏科技大学计算机科学与工程学院 镇江 212003)²

摘 要 提出了一种新的基于最小外接圆补偿机制的二值形状图像检索算法,其不仅提取了目标区域的特征,而且补偿了目标最小外接圆下的背景区域特征。提取常用于描述区域特征的 Hu 不变矩和 Zernike 不变矩作为图像特征,不变矩特征具有平移、缩放、旋转不变性。图像相似度用归一化特征向量的欧氏距离表示。该方法不仅计算简单,而且可有效补偿人眼的视觉感受。通过大量实验表明,该算法较仅基于目标区域的检索算法取得了更好的检索精度和回召率。

关键词 背景补偿, Hu 矩, Zernike 矩, 形状, 图像检索

中图法分类号 TP391.41 文献标识码 A

Shape Image Retrieval Based on Compensation Mechanism under Minimum Circumscribed Circle

SHU Xin^{1,2} WU Xiao-jun¹ PAN Lei¹

(School of Information Technology, Jiangnan University, Wuxi 214122, China)¹

(School of Computer Science and Engineering, Jiangsu University of Science and Technology, Zhenjiang 212003, China)²

Abstract A novel shape image retrieval algorithms based on compensation mechanism under minimum circumscribed circle was proposed in this paper, which not only extract the features of the target area, but also compensate the background characteristics of the object under minimum circumscribed circle. Extracting the Hu invariant moments and Zernike moment invariants, commonly used in the description of the regional characteristics, as image features. Moment invariants have the unchanged characteristics of translation, scaling and rotating. Similarity is given by the Euclidean distance between two images' normalized moment vectors. The approach is implemented simply, and compensates the human eye's visual experience effectively. A large number of experiments shown that the approach can achieve better accuracy and recall rate than that based on an object area only.

Keywords Background compensation, Hu moments, Zernike moments, Shape, Image retrieval

1 引言

随着多媒体时代的到来,人们越来越多地接触到大量的图像信息。如何有效地组织、管理和检索大规模的图像数据库,成为当前一个热门的研究课题。基于内容的图像检索(Content Based Image Retrieval, 简称 CBIR)^[1,2] 技术是解决这一问题的关键性技术之一。

形状是描述图像内容的一个重要特征,对形状分析和分类的方法技术有许多^[3,9]。通常形状的描述可以分为基于轮廓和基于区域两类。其中,基于轮廓的形状描述技术已相当成熟,但不足之处在于没有考虑形状区域内的像素。基于区域的形状描述方法主要有区域的面积、形状的纵横比、区域拓扑、区域频谱等^[3]。但基于矩的描述方法更适合图像的形状特征,因为它们不仅仅计算轮廓上的像素点,而且还计算构成形状的所有像素点^[4-8]。早在 1962 年, M. K. Hu^[4] 就提出了

基于矩不变量的模式识别思想。近年来国内也有很多学者对用矩来描述形状特征作了深入研究,姚玉荣等^[10]提出了利用小波和矩进行基于形状的形状图像检索,夏永泉等^[11]提出了不变矩方法在区域匹配中的应用,云挺等^[12]提出了基于 Zernike 矩的区域匹配方法。但上述方法仅考虑了图像中目标区域的特征匹配,忽视了背景在视觉感受中的作用。

本文提出了一种新的以形状目标的最小外接圆为特征提取区域的目标、背景补偿机制的匹配方法,该方法不仅提取目标区域的特征,而且综合考虑了在该外接圆下的背景区域特征,其特点在于:(1)形状目标与背景反映了图像的结构信息,与人的视觉感知特性一致;(2)以目标轮廓的最小外接圆为特征提取区域的补偿方法不仅很好地反映了图像形状的目标区域信息,也考虑了背景区域对视觉感知的补偿特性;(3)以常用的 Hu 不变矩和 Zernike 不变矩作为图像特征进行相似性度量,具有平移、缩放、旋转不变等特性。该方法不仅计算简

到稿日期:2009-02-13 返修日期:2009-04-27 本文受国家自然科学基金项目(60472060,60572034),江苏省自然科学基金项目(BK2006081),教育部新世纪优秀人才支持计划基金项目(NCET-06-0487)资助。

束 鑫(1979-),男,博士研究生,讲师,主要研究方向为模式识别、计算机视觉、基于内容的信息检索, E-mail: ecsishu@yahoo.com.cn; 吴小俊(1967-),男,博士,教授,博士生导师,主要研究方向为神经网络、模式识别和人工智能; 潘 磊(1980-),男,博士研究生,主要研究方向为模式识别、计算机视觉。

单,而且直观地反应了人眼的视觉特性。通过大量实验表明,该算法较仅基于形状目标的检索算法取得了更好的检索精度和召回率。

2 Hu 矩与 Zernike 矩

2.1 Hu 矩的定义

不变矩理论是模式识别和计算机视觉中的一个重要内容,是对图像的一种统计形式,它的计算要用到图像或区域中所有相关的像素点。常用的基于区域的矩不变量理论最早由 Hu 提出,并获得不断发展和广泛应用。

对一个数字图像 $f(x, y)$, 如果它分段连续且只在 XY 平面上的有限个点不为零, 则可证明它的各阶矩存在。 $f(x, y)$ 的 $p+q$ 阶矩定义为:

$$m_{pq} = \sum_x \sum_y x^p y^q f(x, y)$$

可以证明, m_{pq} 惟一地被 $f(x, y)$ 所确定, 反之, m_{pq} 也惟一地确定了 $f(x, y)$, $f(x, y)$ 的 $p+q$ 阶中心矩定义为:

$$\mu_{pq} = \sum_x \sum_y (x-x_0)^p (y-y_0)^q f(x, y)$$

其中, $x_0 = m_{10}/m_{00}$, $y_0 = m_{01}/m_{00}$, 即质心坐标。 $f(x, y)$ 的归一化中心矩可以表示为:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma}$$

其中, $\gamma = \frac{p+q}{2} + 1, p+q=2, 3, \dots$

将一些归一化的二阶和三阶中心矩进行组合可得到 7 个对平移、旋转和尺度变化不变的矩, 即有:

$$\phi_1 = \eta_{20} + \eta_{02}$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{33})^2$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{33})^2$$

$$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{33})^2] + (3\eta_{21} - \eta_{33})(\eta_{21} + \eta_{33})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{33})^2]$$

$$\phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{33})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{33})$$

$$\phi_7 = (3\eta_{21} - \eta_{33})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{33})^2] + (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{33})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{33})^2]$$

2.2 Zernike 矩的定义

Zernike 矩是一种正交复数矩^[13], 它利用的正交多项式集是一个在单位圆 ($D^2: x^2 + y^2 \leq 1$) 内的完备正交集, Zernike 矩定义为:

$$P_{p,q} = \frac{p+1}{\pi} \iint_{D^2} f(x, y) W_{p,q}^*(x, y) dx dy$$

其中, p 是正整数或零, q 是整数, 并且 $|q| \leq p$, $f(x, y)$ 是图像函数, $*$ 表示复数共轭。

在极坐标中, p 阶 q 重 Zernike 基函数表示为:

$$W_{p,q}(\rho \cos \theta, \rho \sin \theta) = S_{p,q}(\rho) \exp(iq\theta)$$

其中, $S_{p,q}$ 是由下式给出的径向多项式。

$$S_{p,q}(\rho) = \sum_{s=0}^{p-q} \frac{(-1)^s (2p+1-s)!}{s! (p-q-s)! (p+q+1-s)!} \rho^{p-s}$$

对于数字图像, 设 $f(x, y)$ 是当前像素, 则 Zernike 矩可表示为:

$$P_{p,q} = \frac{p+1}{\pi} \sum_x \sum_y f(x, y) W_{p,q}^*(x, y), x^2 + y^2 \leq 1$$

3 最小外接圆补偿机制描述

3.1 补偿机制描述

传统的基于形状的图像检索方法仅考虑了图像中目标(轮廓封闭区域)的匹配, 忽视了背景(本文指最小外接圆减去目标所得)在人的视觉感受中的作用。本文提出的以最小外接圆为特征提取区域的目标、背景补偿机制的匹配方法, 不仅提取目标区域的特征, 而且综合考虑了在该外接圆下的背景区域特征。

首先把原始图像中的目标统一到其最小外接圆下, 如图 1(a) 为原始图像 ($N \times M$), 图 1(b) 为包含目标的最小外接圆的图像 ($(2R+1) \times (2R+1)$, R 为半径), 图 1(c) 为图 1(b) 中目标的最小外接圆下的背景(白色为背景区域)图像, 分别提取目标与背景的 7 个 Hu 不变矩和若干阶 Zernike 矩特征进行相似性度量, 通过大量实验表明这种目标与背景的互补机制能有效提高检索精度和召回率。

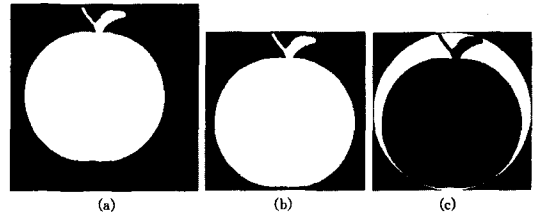


图 1 目标与背景补偿机制示例

3.2 基于最小外接圆面积的特征归一化

为了减小目标大小对相似性匹配的影响, 文中选用基于最小外接圆面积的特征归一化。令 $F_{hu} = \{f_1, f_2, \dots, f_7\}$ 为目标 Hu 不变矩特征, $B_{hu} = \{b_1, b_2, \dots, b_7\}$ 为目标对应的最小外接圆下背景的 Hu 不变矩特征; $F_{zernike} = \{f_1, f_2, \dots, f_n\}$ 为目标 Zernike 不变矩特征, $B_{zernike} = \{b_1, b_2, \dots, b_n\}$ 为目标对应的最小外接圆下背景的 Zernike 不变矩特征, n 为矩的阶数。目标最小外接圆的半径为 R , 则其面积 $A_{circle} = \pi \times R^2$ 。归一化后的目标和背景的 Hu 与 Zernike 不变矩分别为:

$$\tilde{F}_{hu} = F_{hu} / A_{circle}$$

$$\tilde{B}_{hu} = B_{hu} / A_{circle}$$

$$\tilde{F}_{zernike} = F_{zernike} / A_{circle}$$

$$\tilde{B}_{zernike} = B_{zernike} / A_{circle}$$

3.3 图像相似性度量

设 $I^{(q)}$ 和 $I^{(o)}$ 分别为查询图像和数据库图像, $\tilde{F}_{hu}^{(q)}, \tilde{F}_{hu}^{(o)}$ 为查询图像和数据库图像中目标归一化的 Hu 不变矩特征, $\tilde{B}_{hu}^{(q)}, \tilde{B}_{hu}^{(o)}$ 为归一化的背景 Hu 不变矩特征; $A_{hg}^{(q)}, A_{hg}^{(o)}$ 分别为查询图像中目标和背景的面积。则 $I^{(q)}$ 和 $I^{(o)}$ 的 Hu 不变矩相似度可以定义为:

$$Dist(I^{(q)}, I^{(o)}) = \frac{A_{hg}^{(q)}}{A_{circle}^{(q)}} \times \| \tilde{F}_{hu}^{(q)} - \tilde{F}_{hu}^{(o)} \| + \frac{A_{hg}^{(o)}}{A_{circle}^{(o)}} \times \| \tilde{B}_{hu}^{(q)} - \tilde{B}_{hu}^{(o)} \|$$

其中, $\| * \|$ 为特征距离算子。为简便起见, 本文实验中采用欧式距离。 $I^{(q)}$ 和 $I^{(o)}$ 的 Zernike 不变矩相似度可以类似定义。

4 实验结果与分析

4.1 形状数据库

实验中使用了3个常用的形状测试图像库,分别是:(1) Kimia's shapes database(9个形状类别,每类11幅,共99幅);(2)Haibin Ling's^[14] shapes database(18个形状类别,每类12幅,共216幅);(3)MPEG7_CE-Shape-1_Part_B shapes database(70个形状类别,每类20幅,共1400幅)。其中, Kimia's shapes database如图2所示,限于篇幅,其他数据库图示不再给出。

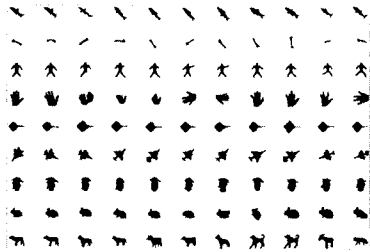


图2 Kimia's 形状数据库

4.2 实验结果

(1)基于本文提出的以目标形状的最小外接圆为特征提取区域的目标、背景补偿机制的匹配方法,图3给出了以 Haibin Ling's shapes database 中部分图像作为查询图像检索返回的前12幅图像的结果,图像上面的数字标识该图像在图像库中的位置,第一幅为查询图像,后面以相似性大小排列,检索结果令人非常满意。图3(a)基于 Hu 不变矩作为图像特征,图3(b)基于 Zernike 不变矩作为图像特征,实验中 Zernike 矩的阶数取自文献[15],从图上返回的结果来看基于 Zernike 不变矩的检索结果要优于 Hu 不变矩。

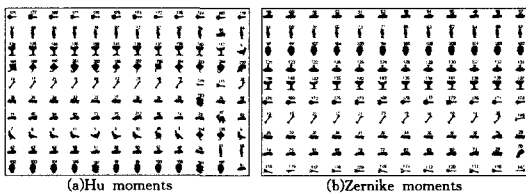


图3 部分检索结果基于 Hu 和 Zernike 不变量特征

(2)检索精度对比。表1列出了3个数据库中的检索精度数据(在返回图像数目是各数据库中每类形状图像数目的2倍情况下统计所得)。从表中数据来看,基于目标和背景补偿机制的检索方法要优于仅基于目标的检索方法,特别是在采用 Hu 不变矩作为图像特征的情况下。这是因为 Zernike 矩本质上较 Hu 矩有更好的图像表示效果,故在采用补偿机制后的检索效果不如采用 Hu 不变矩明显。

表1 检索精度对比表

数据库	Hu moments		Zernike moments	
	仅基于目标	基于目标和背景	仅基于目标	基于目标和背景
Database1 (返回22幅)	0.6143	0.7539	0.8402	0.8549
Database2 (返回24幅)	0.6470	0.7269	0.8430	0.8573
Database3 (返回40幅)	0.3840	0.4736	0.6645	0.6660

(3)PVR 曲线

通常,人们使用检索精度和回召率来评价检索系统的性能^[16]。检索精度定义为返回的相关图像数与检索返回的图像总数的比率,检索回召率定义为返回有效图像数与图像库中有效图像总数的比率。根据 PVR 指标的定义,实验中把数

据库中的每幅图像都作为查询图像来统计平均检索精度与回召率,详细结果如下。

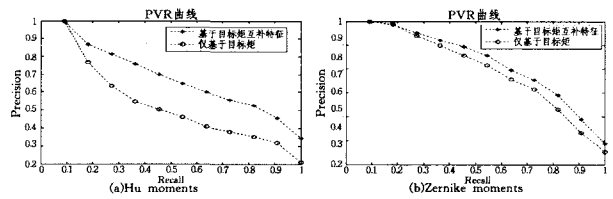


图4 Kimia 形状数据库的 PVR 曲线

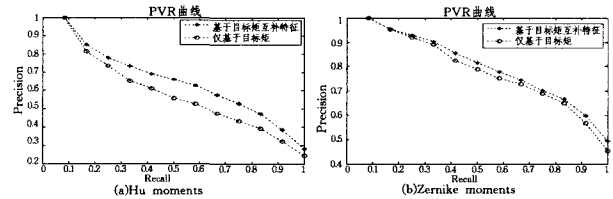


图5 Haibin Ling 形状数据库的 PVR 曲线

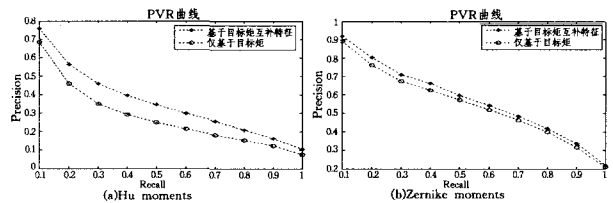


图6 MPEG7_CE-Shape-1_Part_B 形状数据库的 PVR 曲线

图4、图5、图6给出了本文算法与传统算法在3个典型的形状数据库中的比较实验,从以上的 PVR 曲线来看,本文提出的目标与背景的互补机制在任一数据库中都能显著提高检索精度与回召率,无论是采用 Hu 不变矩特征还是 Zernike 不变矩特征。在 Zernike 矩上的实验效果不如 Hu 矩明显,这是因为单用 Zernike 矩进行检索的结果已经较 Hu 矩取得了更好的效果,所以这种互补机制用于 Zernike 矩的效果不如 Hu 矩明显,但较传统的仅提取目标区域特征的效果有了较大改善。

结束语 本文提出了一种新的以形状目标的最小外接圆为特征提取区域的目标、背景补偿机制的封闭边界形状图像匹配方法,该方法不仅提取目标区域的特征,而且综合考虑了在该外接圆下的背景区域特征。以目标轮廓的最小外接圆为特征提取区域的补偿方法不仅很好地反映了图像形状的目标区域信息,也考虑了背景区域对视觉感知的补偿特性;通过大量实验表明,该算法较仅基于形状目标的检索算法取得了更好的检索精度和回召率。

参考文献

- [1] Worring A W M, Santini M S, et al. Content-based image retrieval at the end of the early years[J]. IEEE Trans. Pattern Analysis and Machine Intelligence, 2000, 22(12): 1349-1380
- [2] Liu Ying, et al. A survey of content-based image retrieval with high-level semantics[J]. Pattern Recognition, 2007, 40(1): 262-282
- [3] Costa L F, Cesar R M. Shape Analysis and Classification: Theory and Practice[M]. CRC Press, 2001
- [4] Hu M K. Visual Pattern Recognition by Moment Invariants[J]. IRE Trans. Information Theory, 1962, 8: 179-187

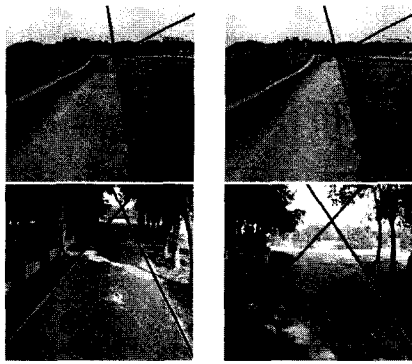


图6 普通道路实验环境

结束语 主动导航车辆中利用视觉系统算法理解车辆的环境非常重要。为了在图像中更精确地检测因为较多光照阴影、边缘信息模糊等因素造成影响的道路边缘,本文提出了一种新的方法。该方法以直线变形模型为基础,在先验知识的辅助下定义后验概率分布描述道路特征结构,利用粒子群优化算法搜索与图像匹配的最优模板。同时利用车辆运动的动态模型和粒子滤波算法,综合预测信息,达到了对道路边缘检测较为准确的估计。实验证明,本文算法实现了较为准确的道路边缘或者行道线的检测,能够有效地减少噪声因素的影响,具有较好的实用性。

参考文献

[1] Chapuis R, Aufrere R, Chausse F. Accurate road following and reconstruction by computer vision[J]. IEEE Transactions on Intelligent Transportation Systems, 2002, 3(4): 261-270

[2] Li Qing, Zheng Nan-ning, Cheng Hong. Springrobot: a prototype autonomous vehicle and its algorithms for lane detection[J]. IEEE Transactions on Intelligent Transportation Systems, 2004, 5(4): 300-308

[3] Wang Yue, Teoh E K, Shen Dinggang. Lane detection using B-Snake[J]. Image and Vision Computing, 2004, 22(4): 269-280

[4] Park J W, Lee J W, Jhang K Y. A lane-curve detection based on an LCF[J]. Pattern Recognition Letters, 2003, 24(14): 2301-2313

[5] Bcher T, Curio C, Edelbrunner J, et al. Image processing and behavior planning for intelligent vehicles[J]. IEEE Transactions on Industrial Electronics, 2003, 50(1): 62-75

[6] Crisman J D, Thorpe C E. UNSCARF: A Color Vision System for the Detection of Unstructured Roads[C]// Proc. of IEEE Int'l Conf on Robotics and Automation. 1991: 2496-2501

[7] Crisman J D, Thorpe C E. SCARF: A Color Vision System that Tracks Roads and Intersections[J]. IEEE Trans. on Robotics and Automation, 1993, 9(1): 49-58

[8] 孙涵. 基于红外图像的道路识别与运动目标跟踪[D]. 南京, 2005

[9] Thorpe C, et al. Vision and navigation for the Carnegie-Mellon NAVALAB[J]. IEEE Trans. Pattern Analysis and Machine Intelligence, 1988, 10(3): 362-73

[10] Zhang Jin-you, Naqel H. Texture - based segmentation of road images[C]// IEEE Intelligent Vehicles apos 94 Symposium. Oct 1994: 260- 265

[11] Eberhart R C, Kennedy J. A new optimizer using particle swarm theory[C]// Proc. 6th Int. Symp. Micro Machine and Human Science. 1995: 39-43

[12] Engelbrecht A P, Ismail A. Training product unit neural networks[J]. Stability Control: Theory Appl. , 1999, 2(1): 59-74

[13] Isard M, Blake A. Condensation-condition density propagation for visual tracking[J]. International Journal of Computer Vision, 1998, 29: 5-28

[14] Arulampalam M S, et al. A tutorial on particle filters for online nonlinear/non-gaussian Bayesian tracking[J]. IEEE Trans. Signal Processing, 2002, 50(2): 174-188

[15] Doucet A, et al. On sequential Monte Carlo sampling methods for Bayesian filtering[J]. State Compute, 2000, 10(3): 197-208

(上接第 254 页)

[8] 张华平, 刘群. 基于 N-最短路径方法的中文词语粗分模型[J]. 中文信息学报, 2002, 16(5): 1-7

[9] Goh C-L, Asahara M, Matsumoto Y. Chinese Word Segmentation by Classification of Characters[J]. Computational Linguistics and Chinese Language Processing, 2005, 10(3): 381-396

[10] Wang Zhuoran, Liu Ting. Chinese Unknown Word Identification Based on Local Bigram Model[J]. International Journal of Computer Processing of Oriental Languages, 2005, 18(3): 185-196

[11] Sebastiani F. Machine learning in automated text categorization [J]. ACM Computing Surveys, 2002, 34(1): 1-47

[12] 苏金树, 张博锋, 徐昕. 基于机器学习的文本分类技术研究进展[J]. 软件学报, 2006, 17(9): 1848-1859

[13] 李荣陆, 胡运发. 基于密度的 kNN 文本分类器训练样本裁剪方法[J]. 计算机研究与发展, 2004, 41(4): 539-545

[14] 庄东, 陈英. 基于加权近似支持向量机的文本分类[J]. 清华大学学报: 自然科学版, 2005, 45(S1): 1787-1790

[15] Kim S-B. Some Effective Techniques for Naive Bayes Text Classification[J]. IEEE Transactions on Knowledge and Data Engineering, 2006, 18(11): 1457-1466

[16] 钟将, 温罗生, 冯永, 等. 基于近似支持向量机的 Web 文本分类研究[J]. 计算机科学, 2008, 35(3): 167-169

(上接第 281 页)

[5] Liao S X, Pawlak M. On Image Analysis by the Methods of Momnets[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1996, 18(3): 254-266

[6] Flusser J. Moment Invariants in Image Analysis[J]. Proceedings of World Academy of Science, Engineering and Technology, 2006, 11(2): 196-201

[7] Chong C W, Raveendran P, Mukundan R. Translation invariants of Zernike moments[J]. Pattern Recognition, 2003, 36(9): 1765-1773

[8] Xu Dong, Li Hua. Geometric moment invariants [J]. Pattern Recognition, 2008, 41(1): 240-249

[9] 丁险峰, 吴洪, 张宏江. 形状匹配综述[J]. 自动化学报, 2001, 27(5): 678-694

[10] 姚玉荣, 章毓晋. 利用小波和矩进行基于形状的图象检索[J]. 中国图像图形学报, 2000, 5(3): 206-210

[11] 夏永泉, 刘正东, 杨静宇. 不变矩方法在区域匹配中的应用[J]. 计算机辅助设计与图形学学报, 2005, 17(10): 2152-2156

[12] 云挺, 顾磊, 吴慧中. 基于 Zernike 矩的区域匹配方法[J]. 中国图像图形学报, 2008, 13(8): 1517-1524

[13] 叶斌, 彭嘉雄. 伪 Zernike 矩不变性分析及其改进研究[J]. 中国图像图形, 2003, 8(3): 246-252

[14] <http://www.ist.temple.edu/~hbling/>

[15] Zhang D S, Lu G. Review of Shape Representation and Description Techniques[J]. Pattern Recognition, 2004, 37(1): 1-19

[16] Iqbal Q, Aggarwal J K. Applying Pecepture Grouping to Content-based Image Retrieval; Building images[C]// IEEE Workshop on Content-based Aecess of Image and Video Libraries. Fort Collins, CO, 1999: 42-48