

基于距离比上下文的形状描述与识别方法

束 鑫 唐 楠 邱 源

(江苏科技大学计算机科学与工程学院 镇江 212003)

摘 要 基于形状轮廓上的采样点到形状质心的距离,提出了一种距离比上下文形状描述符,用于形状识别和检索。该描述符计算简单,能有效区分不同形状,本质上具有平移、缩放不变性,且在一定程度上能抗部分遮挡和形变。用动态规划算法度量形状比上下文之间的距离,解决了对起始轮廓点的选择问题。在 kimia's-99 形状图像数据库中的实验结果表明,该方法在单目标封闭轮廓的形状图像检索中取得了良好的效果。

关键词 形状识别,形状检索,距离比上下文,动态规划

中图法分类号 TP391.41 **文献标识码** A

Shape Description and Recognition Approach Based on Distance Ratio Context

SHU Xin TANG Nan QIU Yuan

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

Abstract We suggested a shape descriptor named Distance Ratio Context (DRC) which is based on the distance between sampled points and the centroid of the object. This descriptor has the properties of invariant to scaling and translation essentially and can be calculated easily. It's also invariant to deformation and distortion in some ways and can discriminate different shapes effectively. The dynamic programming algorithms were used to measure the distance between DRCs and this mechanism solved the problem of start point choosing on the object contour. The experiments in kimia's-99 shape dataset show that this approach, used in image retrieval of shape with a single closed contour, can get favorable results.

Keywords Shape recognition, Shape retrieval, Distance ratio contexts, Dynamic programming

1 引言

图像内容包含的视觉特征主要包括颜色、纹理、形状等,其中形状是非常重要的一种特征,对目标的理解与识别起着非常关键的作用。对于自然界的物体,形状可能比颜色和纹理包含更多的语义信息,人们往往大多通过形状来识别不同的物体。例如,人们阅读文章时,本质上是从形状上来识别文字的。因此,对形状进行深入研究,提出一些能满足平移、缩放、旋转不变性要求的形状描述符,并以此为基础应用于图像图形识别、字符识别、商标分类与检索等领域都具有重要的应用价值和现实意义。

在二维坐标平面中,形状通常被认为是一条封闭轮廓线所包围的区域。对物体形状的表示和描述已经有了许多方法^[1,2],通常可分为两大类:基于区域的方法和基于轮廓的方法。基于区域的表示方法使用形状区域中所有像素来描述形状,典型的方法有基于矩(Moment Based)的描述方法,由 Hu 于 20 世纪 60 年代首先提出^[3]。常用于形状描述的矩还包括:Legendre 矩、Zernike 矩和 Fourier-mellin 矩^[4-6]等。但基于矩的方法通常缺乏明确的物理意义,且计算量过大,因此在实际应用中效果不太理想。基于轮廓的表示方法利用边界信息来描述形状,常用方法有:简单几何特征(如边界长度、边界

半径、曲率、离心率、狭长度、挠度能量等^[7])、弹性匹配^[8](Elastic Matching)、曲率尺度空间 CSS^[9](Curvature Scale Space)、傅里叶描述子^[10](Fourier Descriptor)、小波描述子^[11](Wavelet Descriptor)、形状上下文 SC^[12](Shape Contexts)等。SC 是近年来提出的性能较为优秀的轮廓描述符,在二维形状和字符识别中取得了不俗表现。在 SC 方法的基础上发展起来的方法已有很多^[13-16],本文也在一定程度上受其启发。

本文基于形状轮廓上的采样点到形状质心的距离,提出了一种基于领域距离比上下文(Distance Ratio Contexts)的形状描述方法,它与动态规划算法(Dynamic Programming Algorithm)相结合,用于形状识别与检索。其特点在于:(1)距离比上下文(DRC)本质上具有平移、缩放不变性;(2)距离比上下文(DRC)的旋转不变性(即轮廓上起始点的选择问题)可由动态规划算法(DPA)加以解决;(3)距离比上下文(DRC)反映了轮廓的结构信息,与人的视觉感知特性一致。

2 特征提取

2.1 形状表示

对于目标的轮廓来说,轮廓上某像素点与其领域内的其他像素点的位置关系反映了该点的局部形状特征。因此,可

到稿日期:2010-12-21 返修日期:2011-03-24 本文受国家自然科学基金项目(60973094),江苏科技大学本科生创新计划项目资助。

束 鑫(1979-),男,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。

以通过像素点相对于目标质心的位置分布对轮廓形状进行描述。本文首先采用轮廓跟踪算法对图像目标轮廓进行提取,但由于目标形状轮廓的点数过多,为了减少计算量,采用等间隔方法对轮廓点进行采样,得到轮廓的部分点集 $P=\{p_1, p_2, \dots, p_n\}$, $p_i \in \mathbb{R}^2$, n 表示点的数量,来描述物体的形状(见图 1 轮廓提取效果图)。图 1 中(a)为原始图片;(b)为目标的轮廓;(c)为采样得到的轮廓点及质心(*表示目标的质心)。

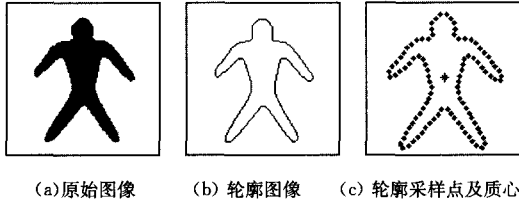


图 1 形状图像及其轮廓

2.2 DRC 的构建

如图 2 所示,形状由轮廓上的 N 个采样点表示(P_0 为质心),对于某点 P_i 来说,其与质心的距离为 $|P_0 P_i|$,在 P_i 相邻的两个方向上分别取 m 个相邻点,可得到 $2m$ 个轮廓点到质心的距离 $|P_0 P_j|$,其中 $j=1, \dots, 2m$,那么,对于点 P_i 可构建其领域距离比上下文 DRC 记为: $C_i = \{|P_0 P_j| / |P_0 P_i|\}$, $j=1, \dots, 2m$, C_i 为一个 $1 \times 2m$ 的向量,简记为: $C_i = \{r_1, r_2, \dots, r_{2m}\}$ 。对于轮廓上的每个采样点,都可按照上述方法构建一个 $1 \times 2m$ 的向量,共 N 个。而这 N 个向量又可构成一个 $N \times 2m$ 的矩阵,记为 DRC,该矩阵称为距离比上下文矩阵。

$$DRC = \begin{bmatrix} r_{1,1} & r_{1,2} & \dots & r_{1,2m} \\ r_{2,1} & r_{2,2} & \dots & r_{2,2m} \\ \vdots & \vdots & \dots & \vdots \\ r_{n,1} & r_{n,2} & \dots & r_{n,2m} \end{bmatrix}$$

矩阵 DRC 即为该形状的描述符。

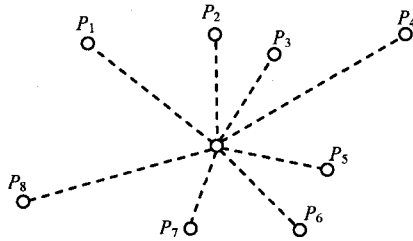


图 2 DRC 构建示意图

3 DRCs 间的距离度量

图像之间的相似性度量是直接影响识别与检索效果的重要因素,人们也一直在致力于研究如何用数值来有效地表示图像在特征上的相似程度。本文采用被广泛用于模式识别领域的 DP^[17-19] 算法来度量目标形状的 DRCs 之间的距离。在 DP 求解过程中,使用一张 DP 表来寻找两个 DRC 之间的最佳匹配,并把搜索路径限定在对角线两侧固定宽度的窗口内,这样可以提高 DP 算法求解的效率,具体如图 3 所示。DRC₁ 中的一行,即一个向量 C_i (对应于轮廓上的一个采样点的领域距离比上下文特征)对应于 DP 表的一行, DRC₂ 中的一行,即一个向量 C_j 对应于 DP 表的一列。DP 表中的初始值可用式(1)求得。

$$DPT_{initial}(n, m) = \begin{cases} 0, & \max(1, n-w+1) \leq m \leq \min(N, n+w-1) \\ \infty, & \text{otherwise} \end{cases} \quad (1)$$

式中, $n, m \in \langle 1, N \rangle$, N 为 DP 表的行数或列数(即为轮廓上的采样点数目), w 为 DP 表对角线的窗口宽度(实验中, $w=3$)。在 DP 算法求解过程中,只有 DP 表中的特定区域(图 3 中的灰色对角线区域)的值得到更新,更新后的值用式(2)求得。

$$DPT(n, m) = D(n, m) + \min \begin{cases} DPT(n-1, m) \\ DPT(n-1, m-1) \\ DPT(n, m-1) \end{cases} \quad (2)$$

$$D(n, m) = \|C_n - C_m\| \quad (3)$$

式中, $D(n, m)$ 为向量 C_n 和 C_m 之间的欧式距离(C_n 和 C_m 分别来自 DRC₁ 和 DRC₂)。不失一般性, DP 算法求解过程从图 3 中标记为(1,1)的单元开始,沿着指定的对角线从左到右、从下到上根据式(2)依次更新, $DPT(N, N)$ 中的值即为该次动态规划所求结果。为了弥补 DRC 对于旋转不变性要求的不足,可采用重复求解而部分地解决。即固定 DRC₁ 不动,把 DRC₂ 中的各列向量循环左移一列,重复一次 DP 算法,共 N 次,取其最小值为图像之间的相似性度量值,记为:

$$Dist1 = \min_{i=1, \dots, N} \{DPT_i(N, N)\} \quad (4)$$

同样,固定 DRC₁ 不动,把 DRC₂ 对称翻转一次,重复上述步骤,可得到具有对称不变性的相似性度量值,记为:

$$Dist2 = \min_{i=1, \dots, N} \{DPT_i(N, N)\} \quad (5)$$

两幅图像的 DRC 之间的最终相似度定义为:

$$Dist = \min(Dist1, Dist2) \quad (6)$$

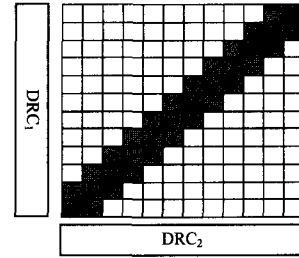


图 3 用于匹配 DRCs 的 DP 表

4 实验结果

(1) DRCs 提取结果示例

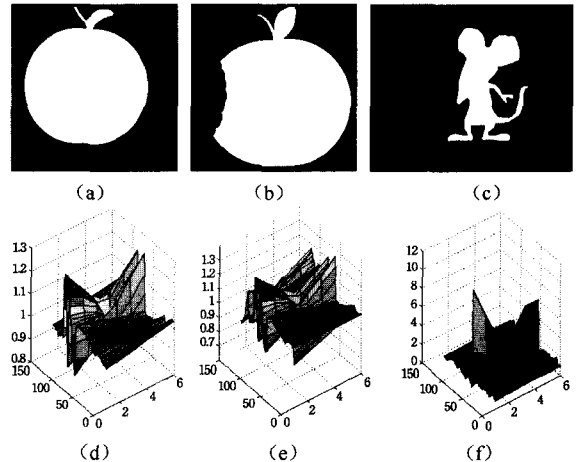


图 4 用于匹配 DRCs 的 DP 表

依据 2.2 节所述方法对图像进行距离比上下文(DRC)提取,效果如图 4 所示,其中(a),(b),(c)为原始图像,而(d),(e),(f)为对应于(a),(b),(c)的距离比上下文 DRC。从图 4

中可以看出,由于图(a)与图(b)中的均为苹果图像,形状非常相似,故它们对应的 DRC 图像也比较相似;而图(a)、(b)与图(c)中的形状存在较大差异,则提取到的 DRC 图(f)与图(d)、(e)差异也较大,所以可以通过 DRC 来识别不同的形状。

(2)DRCs 与 SC 的比较

根据上述算法,把 Kimia-99(包含 9 类不同形状,每类 11

幅共 99 幅图像)形状库中的每幅图像都作为查询图像,分别统计返回的前第 $i(i \leq T, T$ 为每个类别的图像数目,本实验中 $T=10$)幅图像中与查询图像是同一类别的图像数目。表 1 给出了本文方法与 SC 方法的比较。从表 1 中的数据来看,本文算法要略好于 SC^[12]算法(本文方法使用的参数为 $N=100, m=9$)。

表 1 Kimia-99 数据中的检索结果比较

Method	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	Total
Shape Contexts	97	91	88	85	84	77	75	66	56	37	756
本文方法	97	91	90	90	87	74	68	68	61	48	774

(3)查询结果示例

图 5 给出了在 Kimia-99 形状库中的检索结果示例,每行第一列为查询图像,其余为查询结果图像,按照相似性大小排序,带有红框的为误检图像。从图 5 来看,本文方法检索的结果比较符合人眼的视觉感受,能取得较为满意的结果。

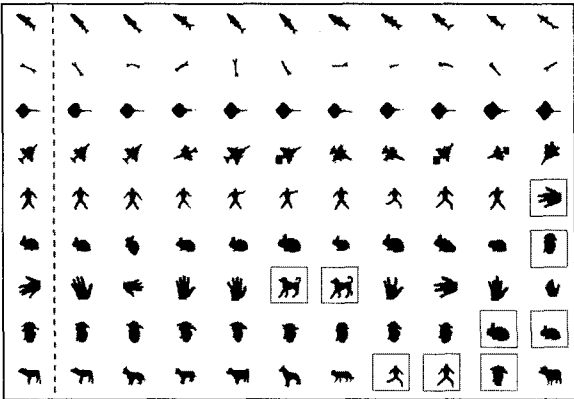


图 5 在 Kimia-99 形状库中的查询结果

(4)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^[20]。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,实验中把数据库中每幅图像都作为查询图像,统计平均检索精度与回召率,图 6 给出了本文方法与两种典型的形状描述方法的 PVR 曲线比较。

从图 6 来看,本文方法要优于 Hu 矩和 Zernike 矩方法,较基于 Hu 矩的方法优势尤为明显。在返回的前 22 幅图像中统计,平均检索精度为 80.17%,平均识别率为 97.98%。

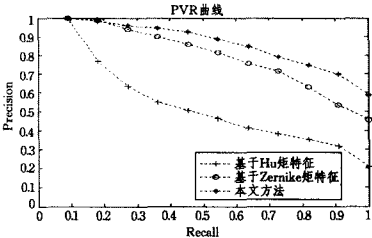


图 6 PVR 曲线比较

结束语 本文提出了一种距离比上下文形状描述符,用于形状识别和检索。该描述符计算简单,能有效区分不同形状,具有平移、缩放不变性,且在一定程度上能抗部分遮挡和形变。结合动态规划算法来度量形状比例上下文之间的距离,在一定程度上解决了对旋转不变性的要求。本文仅考虑

了轮廓点与相邻点之间的一个比例约束关系,今后考虑加上角度等其它可利用的信息,构建鉴别性能更加优越的形状描述方法。还可尝试将该方法推广到 3D 图像的描述与识别中。

参 考 文 献

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

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

[3] Hu M K. Visual pattern recognition by moment invariants [J]. IRE Transactions on Information Theory,1962,8:179-187

[4] Khalid M H. Exact Legendre moment computation for gray level images [J]. Pattern Recognition,2007,40:3597-3605

[5] Papakostas G A, Boutalis Y S, Karras D A, et al. A new calss of Zernike moments for computer vision applications [J]. Information Sciences,2007,177:2802-2819

[6] Zhang H, Shu H Z, Haigron P, et al. Construction of a complete set of orthogonal Fourier-Mellin moment invariants for pattern recognition applications[J]. Image and Vision Computing,2010, 28(1):38-44

[7] [美] Calsteman K R. 数字图像处理[M]. 朱志刚,等译. 北京:电子工业出版社,1998

[8] Attalla E, Siy P. Robust shape similarity retrieval based on contour segmentation polygonal multiresolution and elastic matching[J]. Pattern Recognition,2005,38(12):2229-2241

[9] Mokhtarian F, Ung Y K, Wang Z T. Automatic fitting of digitised contours at multiple scales through the curvature scale space technique[J]. Computers & Graphics, 2005, 29(6):961-971

[10] Zhang D S, Lu G J. A comparative study of curvature scale space and Fourier descriptors for shape-based image retrieval [J]. Journal of Visual Communication and Image Representation, 2003,14(1):39-57

[11] Yadav R B, Nishchal N K, Gupta A K, et al. Retrieval and classification of shape-based objects using Fourier, generic Fourier, and wavelet-Fourier descriptors technique; A comparative study [J]. Optics and Lasers in Engineering,2007,45(6):695-708

[12] Belongie S, Malik J. Shape Matching and Object Recognition Using Shape Contexts [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2002,24:509-522

[13] 陈实,马天骏,黄万红,等. 基于形状上下文描述子的步态识别 [J]. 模式识别与人工智能,2007,20(6):794-799

数或分块数的关系以及安全性进行实验。

为了取得最大的隐写容量,分块采用大小为 $1 \times 1, 2 \times 2, \dots, 20 \times 20$, 共进行 20 次实验, 隐写算法采用文献[11]的方法。分块数与隐写容量的关系如图 3 所示。分块越大, 载体隐写容量越小。采用 2×2 分块将取得最大的隐写容量。

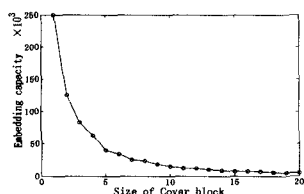


图 3 载体分块的数量与隐写

当 ω 取不同值时, 隐写系统的嵌入容量呈现规律性变化, 如图 4 所示。当 ω 取值一定时, 隐写的最大有效载荷与载体数的平方根成正比; 当 ω 取值越大, 最大载体有效也随之增大, 系统的安全性将变低。

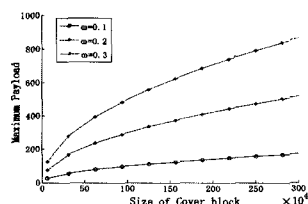


图 4 批量二值图像的隐写容量与载体数的相关性

采用图 2 的载体设计批量隐写系统, 载体数采用隐写载体总数计算, 隐秘信息采用文本信息, 基于 D_{KL} 的理论描述公式, 计算批量二值图像的载体数变化下的 D_{KL} , 实验如图 5 所示。批量隐写满足定理 1 的两个假设条件。

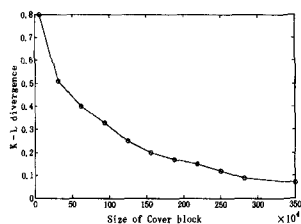


图 5 隐写系统安全性变化曲线

从图 5 可看出, 隐写系统随着载体数的变大, 其安全性将更高, 满足批量隐写的需求。当然, 图 5 中的曲线也会随 ξ 取值的不同而发生变化。

结束语 本文提出了一种批量二值图像的隐写模型, 并

且基于一定的假设条件, 证明了隐写系统的安全性, 分析了系统隐写容量与载体数变化的相关性, 进行了相应的实验, 分析了批量二值图像隐写。本文从理论安全方面论证秘密信息的存在性, 它区别于密码学意义上的安全性分析, 有助于隐写术安全性研究。今后将对本文的隐写安全性证明过程的两个假设条件进行适当放松, 进一步探讨多二值图像多隐秘信息的理论规律。

参考文献

- [1] 傅德胜, 谢永华. 基于图像的信息隐藏检测技术[J]. 信息安全, 2008, 91(7): 72-74
- [2] 周琳娜, 杨义先, 郭云彪, 等. 基于二值图像的信息隐藏研究综述[J]. 中山大学学报, 2004, 43(2): 72-75
- [3] Wu M, Liu B. Data hiding in binary for authentication and annotation [J]. IEEE Transactions on Multimedia, 2004, 6(4): 528-538
- [4] Tseng Y-C, Chen Y-Y, Pan H-K. A secure data hiding scheme for two-color images[J]. IEEE Transactions on Communications, 2002, 50(8): 1227-1231
- [5] 丁一军, 郑学峰, 于桂荣. 一种图像隐藏通信的安全模型[J]. 计算机科学, 2010, 37(2): 120-122, 130
- [6] Ker A D. Batch steganography and pooled steganalysis [C]// Processing of the 8th Information Hiding Workshop. Springerlink LNCS, 2007(4437): 265-281
- [7] Pevny T, Fridrich J. Benchmarking for steganography [C]// Information Hiding 2008. Santa Barbara, CA USA, 2008: 251-267
- [8] Cachin C. An Information-theoretic Model for Steganography [C]// 2nd International Workshop on Information Hiding. Oregon, USA, 1998, 1525: 306-318
- [9] Ker A D, Pevny T, Kodovsky J. The Square Root Law of Steganographic Capacity [C]// Proceedings of the 10th ACM workshop on Multimedia and Security. Oxford U K, 2008: 107-116
- [10] 吕欣, 马智, 冯登国. 安全隐写系统的信息理论分析[J]. 计算机科学, 2006, 33(6): 140-142
- [11] 郭萌, 张鸿斌, 魏磊. 二值图像中的数据隐藏算法[J]. 电子学报, 2009, 37(11): 2409-2415

(上接第 266 页)

- [14] 曹治国, 鄢睿丞, 宋喆. 利用模糊形状上下文关系的红外与可见光图像匹配方法[J]. 红外与激光工程, 2008, 37(6): 1095-1110
- [15] 唐俊, 王年, 梁栋, 等. 一种结合形状上下文分析的 Laplace 谱匹配算法[J]. 系统仿真学报, 2009, 21(14): 4345-4350
- [16] 束鑫, 吴小俊, 潘磊. 一种新的基于形状轮廓点分布的图像检索[J]. 光电子·激光, 2009, 20(10): 1385-1389
- [17] Alajlan N, Elrube I, Kamel M S, et al. Shape retrieval using triangle-area representation and dynamic space warping [J]. Pat-

tern Recognition, 2007, 40(7): 1911-1920

- [18] Petrakis E G M, Diplaros A, Milios E. Matching and retrieval of distorted and occluded shapes using dynamic programming [J]. IEEE Trans. Pattern Anal. Mach. Intell., 2002, 24(11): 1501-1516
- [19] Adamek T, O'Connor N. A multiscale representation method for nonrigid shapes with a single closed contour [J]. IEEE Trans. on CSVT, 2004, 14(5): 742-753
- [20] 章毓晋. 基于内容的视觉信息检索[M]. 北京: 科学出版社, 2003