

用于形状识别的目标轮廓无序点集描述与匹配*

王斌^{1,2,3}



¹(南京财经大学 信息工程学院, 江苏 南京 210023)

²(江苏省电子商务省级重点实验室(南京财经大学), 江苏 南京 210023)

³(江苏省现代粮食流通与安全协同创新中心, 江苏 南京 210023)

通讯作者: 王斌, E-mail: wangbin@njue.edu.cn

摘要: 将目标形状的轮廓看成一个无序的点集, 从中抽取形状特征, 用于快速而有效的目标识别是形状分析任务中的挑战性问题. 针对该问题, 提出了一种基于复杂网络模型的形状描述和识别方法. 该方法提出用一种自组织的网络动态演化模型构成一个分层的描述框架, 在网络动态演化的每一个时刻, 对网络分别进行局部测量和全局测量, 抽取网络的无权特征和加权特征. 在形状匹配阶段, 用获得的局部描述子和全局描述子分别进行局部匹配(基于Hausdorff距离)和全局匹配(基于 L_1 距离), 组合两种匹配的距离值构成对形状的差异度度量. 用标准的测试集对所提出的方法进行性能测试, 实验结果表明, 所提出的算法能够快速而又鲁棒地完成较高精度的形状识别任务.

关键词: 目标识别; 复杂网络; 形状描述; 形状匹配

中图法分类号: TP391

中文引用格式: 王斌. 用于形状识别的目标轮廓无序点集描述与匹配. 软件学报, 2016, 27(12): 3131-3142. <http://www.jos.org.cn/1000-9825/5101.htm>

英文引用格式: Wang B. Shape recognition using unordered point-set description and matching of object contour. Ruan Jian Xue Bao/Journal of Software, 2016, 27(12): 3131-3142 (in Chinese). <http://www.jos.org.cn/1000-9825/5101.htm>

Shape Recognition Using Unordered Point-Set Description and Matching of Object Contour

WANG Bin^{1,2,3}

¹(School of information engineering, Nanjing University of Finance and Economics, Nanjing 210023, China)

²(Key Laboratory of Electronic Business (Nanjing University of Finance and Economics), Nanjing 210023, China)

³(Collaborative Innovation Center for Modern Grain Circulation and Safety, Nanjing 210023, China)

Abstract: Treating the shape contour as an unordered point set and extracting shape features from it for fast and effective shape recognition is a challenge task of shape analysis. To address this issue, a complex-network based shape description and recognition method is proposed in this paper. In this method, a self-organized dynamic network-evolution model is built for providing a hierarchical description framework. In each moment of the dynamic evolution of the complex network, local and global measurements are performed against the network such that both weighted and un-weighted features are extracted from the network. At the shape matching stage, the

* 基金项目: 国家自然科学基金(61372158); 江苏省自然科学基金(BK20141487); 江苏省“333 工程”高层次人才资助项目(BRA2015351); 江苏高校科研成果产业化推进工程项目(JHB2012-18); 江苏高校优势学科建设工程资助项目(PAPD); 江苏省政策引导类计划(产学研合作)-前瞻性联合研究项目(BY2016009-03)

Foundation item: National Natural Science Foundation of China (61372158); the Natural Science Foundation of Jiangsu Province (BK20141487), the “333” Foundation for high level talents of Jiangsu Province (BRA2015351); The industrialization of scientific research achievements in Universities of Jiangsu Province (JHB2012-18); the Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD); and the Policy guidance program (Cooperation of Industry, Education and Academy) of Jiangsu Province (BY2016009-03)

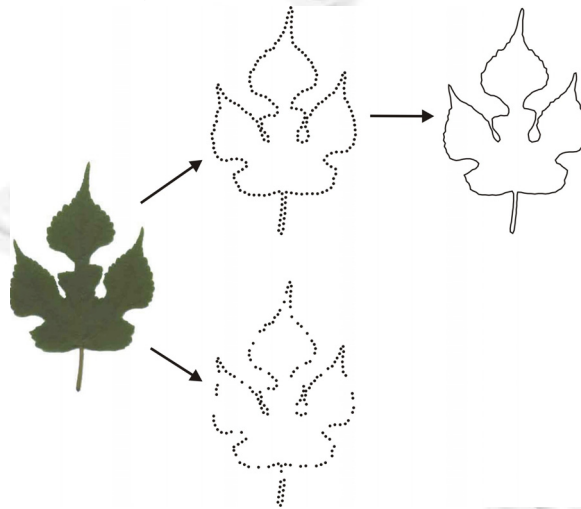
收稿时间: 2015-11-16; 修改时间: 2016-03-22; 采用时间: 2016-04-27; jos 在线出版时间: 2016-07-30

CNKI 网络优先出版: 2016-08-01 09:38:54, <http://www.cnki.net/kcms/detail/11.2560.TP.20160801.0938.003.html>

local matching (based on Hausdorff distance) and global matching (based on L_1 distance) are conducted using the obtained local descriptor and global descriptor respectively. The dissimilar value between two shapes is determined by combining the two distance measures. Several standard test sets are used to evaluate the performance of the proposed method, and the experimental results show that the proposed method can provide robust and fast shape recognition in high accuracy.

Key words: object recognition; complex network; shape description, shape matching

形状是物体非常重要的一类视觉属性,是人的视觉系统对目标进行识别和分类的一个重要依据.在计算机视觉中,形状分析的焦点之一就是如何有效地抽取目标具有辨识力的形状特征并用于后续的认可、分类、检索等任务.人的视觉对目标的轮廓非常敏感,甚至目标的二维轮廓也能给人的视觉系统传递足够的形状信息对其进行准确的识别^[1].图 1 给出了一个树叶图像的完整轮廓线(图 1 右上图)和不完整的轮廓线(图 1 右下图)的一个示例.从该图可以看出:人的视觉仅需借助于目标的轮廓线,甚至在得到的目标的轮廓不完整的情况下,也能正确地识别目标.因此在很多情况下,我们仅需对目标的边界像素点进行分析,就可以抽取出具具有强的辨识能力的形状特征,由此产生了一类形状分析方法,该类方法通常被称为基于轮廓线的形状分析方法^[2].



上图:树叶图像的完整轮廓及其参数化表示 下图:树叶图像的不完整轮廓

Fig.1 A leaf image sample

图 1 树叶图像的一个样本

在轮廓线形状分析方法中,获得大量研究的是参数化曲线的形状分析方法^[1].该类方法将目标外轮廓上的像素点按照它们在图像中的邻接关系(8-邻接或4-邻接)进行有序化表示,形成一条像素点的序列,从而构成一条参数化的简单曲线.图 1 右上图最右边的树叶图像给出了参数化轮廓线的一个示例,这种有序表示要通过轮廓追踪算法^[3]获得.近年来,参数化曲线形状分析得到广泛和深入的研究,产生了大量的形状描述和识别方法^[4-11].这类将目标外轮廓表示为一条参数化曲线的形状分析方法,在形状描述阶段可以对曲线的几何性质如凸凹性进行分析,从而抽取到具有辨识能力的形状特征;而在形状匹配阶段,可以借助像素点的有序信息进行轮廓线的精确匹配,如采用动态规划的匹配方法.这些方法在标准测试集上(如 MPEG-7 形状测试集)都取得了非常高的识别正确率.但这类方法因严重依赖于目标外轮廓像素点的有序化信息,从而存在很大的局限性.当从目标图像中抽取的边界像素点不完整时(如图 1 右下图),则这种对外轮廓像素点的有序化操作将无法完成,即,参数化曲线无法获得.在一些情况下,即使这种参数化曲线能够获取,也相当不准确,如当目标图像的轮廓线出现自交叉现象时,获取的参数化曲线显然无法正确表示原来的形状,图 2 给出了这一问题的示例.此外,由于边界噪声的干扰,也使得获取的参数化曲线非常不稳定.

基于轮廓线的形状分析方法的另一个思路是:将目标边界(包括内边界和外边界)看成一个无序的点集,直

接抽取点集的几何特征作为形状的描述子,我们称为基于边界无序点集的形状分析方法.该类方法因不需要预先对边界像素点进行有序化处理,所以在抽取的边界轮廓不完整的情况和边界噪声干扰的情况下,都能够有效地工作,也当然避免了对曲线进行有序化操作所引入的误差和不稳定性,因此相较于参数化曲线形状分析方法更具鲁棒性和一般性.但是到目前为止,该类方法的研究还相对较少.其面临的挑战在于:

- (1) 相较于参数化曲线形状分析方法,从无序的点集中提取具有强的形状辨识能力的特征将更困难.在参数化曲线方法中,点的有序性这一非常有价值的信息,在基于无序点集的形状分析中将无法使用.在参数化曲线分析方法中被大量使用的数学工具,如小波变换、傅里叶变换、微分几何等,在基于边界无序点集的形状分析中一般也无法使用.因此,该类方法虽具鲁棒性和一般性,但在高辨识能力的形状特征抽取方面具有挑战性;
- (2) 该类方法在形状匹配阶段,为获得较高的识别正确率,往往需要以非常高的形状匹配的计算耗费为代价,如用于点集匹配的 Hungarian 方法^[30],在已知点到点的匹配的耗费情况下,完成两个点集匹配的时间复杂度达到了 $O(N^3)$,这里, N 是点集中点的个数.

因此,该类方法只适合在小规模目标边界采样点集上的匹配任务.而目标边界的小规模采样点集又忽略了目标形状的很多细节信息,所以现有的基于边界点集的较高计算复杂度,也限制了该类方法在对时间要求很高的识别任务中的应用.而本文的研究工作则专注于基于边界无序点集的形状分析新方法的研究,旨在从目标边界的无序点集表示中抽取鲁棒的、具有强的辨识能力的形状特征,在较低的计算代价下获得理想的目标识别效果.

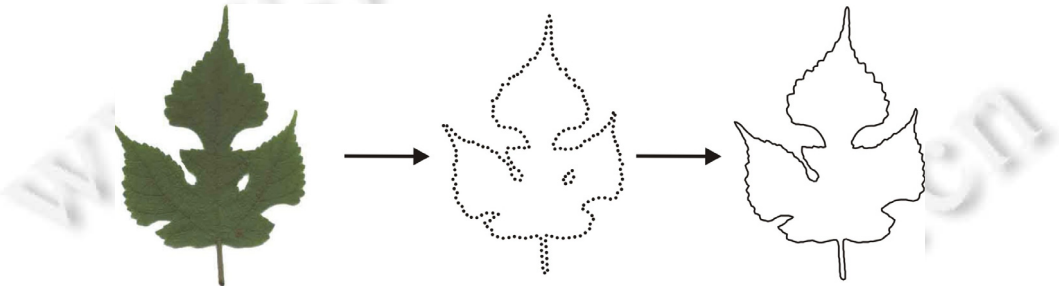


Fig.2 A leaf image sample of self-intersection (left) with its contour (middle) and parametric representation (right)
图2 发生了自交叉的树叶图像的一个样本(左图)及其轮廓(中图)和参数化曲线表示(右图)

1 相关工作

到目前为止,在形状的参数化曲线分析方面,现有的工作已经非常丰富,但在基于边界无序点集的形状分析方面的研究工作却相对较少.本节对这方面的代表性的研究工作进行综述.

形状上下文(shape contexts)^[13]是经典的基于边界无序点集的形状分析方法,该方法将边界像素点集中的每一个点作为参考点,通过在该点放置一个极坐标栅格来统计点集中的其他各点相对于该点的空间分布,并产生直方图作为该点的描述子,也称为该点的形状上下文.在形状匹配阶段,用两个形状的点集中各点的形状上下文来建立两个点集之间的优化 0 的匹配关系,然后计算各对匹配点的匹配误差的和,并将其作为两个形状的距离.该方法有效抽取了轮廓线上点的空间分布信息和相对位置关系,抽取的形状特征具有较强的辨识能力,在 MPEG-7 形状测试集上取得了 76.51% 的识别准确率.但该方法的计算量非常可观,根据文献[13]的报告,执行一次两个形状的匹配计算,在每个形状仅用 100 个采样点的情况下,仍然要耗费 0.2s.为提高计算的效率,文献[14]提出了一种称为轮廓点分布直方图的改进的形状描述子.不同于形状上下文方法将极坐标栅格放置于每个轮廓点的局部描述机制,该方法将极坐标栅格放置于整个形状的质心,从而得到一个描述整个形状轮廓像素点的空间关系的轮廓点分布直方图.形状相似性用 EMD(earth mover's distance)距离进行度量.该方法实质上是一种全局描述子,因不需要为点集中的每一个点准备描述子,因此无论在形状的描述还是在形状的匹配阶段,都减少

了计算量.该方法在 MPEG-7 形状测试集报告的识别准确率与形状上下文方法相当.

距离集(distance sets)是文献[15]提出的另外一种基于边界无序点集的形状分析方法.该方法对大小为 n 的点集中中的每一个点,计算点集中与其最邻近的 $N \ll n$ 个点到它的距离构成一个距离集合,经过归一化处理的距离集合构成关联于该点的局部描述子,用以描述该点与其邻近各点的空间关系.而由各点的距离集构成的集合,即距离集的集合,构成了描述整个点集的空间安排.对两个形状的相似性比较,则通过求解两个距离集的集合的优化匹配关系,该优化问题可归结为经典的二次指派问题.因此,距离集形状分析方法跟形状上下文方法一样具有很高的计算复杂度.该方法报告,在 250 个采样点的情况下,执行两个形状的匹配要耗费 0.7s.但该方法的描述能力非常强,在 MPEG-7 形状测试集获得了 78.38% 的识别准确率,高于形状上下文方法报告的结果.

近年来,随着对复杂网络研究的兴起^[16,17],一些研究工作将复杂网络模型、理论与方法应用于形状分析.文献[18]提出将目标边界建模为一个小世界复杂网络.在该模型中,形状边界像素点对应网络中的节点,像素点间的连线对应网络中的边,像素点间归一化的欧式距离作为边的权重.通过网络的动态演化,抽取各时刻网络的度分布和联合度分布特征组合成一个特征向量,作为形状的描述子,而形状的差异度则通过计算其相应的特征向量的欧式距离来确定.基于该模型,文献[19]提出用演化阈值和度分布特征(包括最小度、最大度和平均度)计算复杂网络的多尺度分形维数描述子,以抽取用于形状识别的不规则模式.文献[20]用内部距离^[5]替代该模型中的欧氏距离来建立形状的复杂网络模型,用度特征和聚类系数特征,再结合内部距离上下文特征来度量形状之间的距离.文献[21]则在用内部距离建立的复杂网络模型的基础上设计了一种 k 近邻的演化机制,得到不同时刻的有向子网络,然后提取有向网络的特征来描述形状.

用复杂网络建模形状,为基于边界无序点集表示的形状分析带来了便利.边界点集中的节点很自然地对应复杂网络中的节点,点集中点的相互联系则用网络的边的有关信息来表达,而网络的动态演化行为也为多尺度形状特征的抽取创造了条件.本文提出一种用于边界无序点集形状分析的复杂网络模型,虽然用复杂网络模型进行形状分析的想法并不是新的,前述工作^[18-21]都是基于该想法而展开的,但本文做出了如下创新和贡献:

- (1) 提出了一种新的网络动态演化机制.现有的用于形状分析的复杂网络模型需要预先设定一系列的阈值作为参数(一般设置阈值的初始值、步长值和终止值)来控制网络的动态演化过程.这种策略一方面因为需要对多个参数进行设置,从而不便于算法的一般使用;另一方面,对不同形状的复杂网络的动态演化都采用相同的阈值参数进行控制,这种没有考虑网络内部特性的阈值选择机制并不易于具有辨识力的形状特征的抽取.本文提出了一种自组织的网络动态演化机制,网络的动态演化不需要预先设定阈值参数,完全交由网络根据其内部信息确定其当前演化的阈值.该机制使得算法更易使用,且抽取形状特征更具识别能力;
- (2) 提出了一种不同的复杂网络特征抽取和匹配方法.现有的基于复杂网络的形状分析都只考虑了复杂网络的无权网络特征(如度分布、联合度分布特征、聚集系数等),使得抽取的特征仅能反映形状的拓扑结构,而不能描述边界像素点之间的空间关系.本文提出了一种带权的复杂网络模型,既抽取了网络的无权特征,又抽取网络的加权特征,同时对网络既进行了全局测量,又进行了局部测量,从而更为全面地描述形状.而在形状匹配阶段,对形状既进行了全局特征的比较,又进行了局部的特征匹配(点到点的匹配),使得对形状差异的度量更为精细.

通过采用公认的标准测试集所进行的性能评估以及与现有的基于边界无序点集表示的形状分析方法的比较,验证了本文提出的方法的有效性.

2 形状复杂网络建模

从统计物理学看,网络是由大量的相互作用、相互依赖的个体组成的系统.据钱学森给出的定义,复杂网络是具有自组织、自相似、吸引子、小世界、无标度中的部分或全部性质的网络.复杂网络是呈现出高度复杂性的网络,可以灵活表达任何自然的结构以及拓扑结构的动态变化特性^[17].

同文献[13]一样,我们首先将形状表示为离散点集 $V = \{v_1, v_2, \dots, v_N\}$, 这些点取自于目标形状的内外轮廓.由于

通过边界检测算子获得的像素点个数太多,需要重新采样,我们采用文献[13]中的方法,对获得的边界像素点采样成大小为 N 的像素点集,这里,正整数 N 是预先指定的参数.跟文献[13]中的方法略有不同的是:如果一个目标形状通过边界检测算法获得的像素点的个数已经小于指定的参数 N ,则不需要进行重新采样.

由形状的离散点集 V 和点集中任意两点的欧式距离 d_{ij} 可构造一个加权的无向图 $G=(V,E)$,这里, E 是图 G 中边的集合,由邻接矩阵 $A=\{a_{ij}\}_{N \times N}$ 给定, $a_{ij}=1$ 表示有一条边连接点 v_i 和 v_j ;而 $a_{ij}=0$ 则表示两个点之间无边相连,图中的每一条边带有权重,其值由矩阵 $W=\{w_{ij}\}_{N \times N}$ 确定.这里, $w_{ij} = d_{ij} / \max_{1 \leq i, j \leq N} (d_{ij})$.显然有 $w_{ij} \leq 1$.复杂网络不同于一般的图的是它的动态演化特性,即,网络的节点或边在不同的时刻可以增加或减少.为此,我们为形状复杂网络建模设计了动态演化机制.令 $A^{(t)}$ 为 G 在 t 时刻的邻接矩阵,此时,网络的平均权重 $\bar{w}^{(t)}$ 定义为

$$\bar{w}^{(t)} = \frac{\sum_{ij} w_{ij} a_{ij}^{(t)}}{\sum_{ij} a_{ij}^{(t)}} \tag{1}$$

则网络 G 在 $t+1$ 时刻的邻接矩阵变化为

$$a_{ij}^{(t+1)} = \begin{cases} a_{ij}^{(t)}, & \text{if } w_{ij} < \bar{w}^{(t)} \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

网络 G 中,那些权重大于或等于当前网络平均权重的边将被去掉.我们定义网络在 $t=0$ 时刻的邻接矩阵 $A^{(0)}=\{1\}_{N \times N}$,即:此时网络中的每两个节点都有边连接, G 构成一个完全图.那么依据公式(1)和公式(2),后续的 $t=1,2,\dots$ 各个时刻的网络图都可以唯一地确定.图 3 给出了狗的形状复杂网络模型(60 个节点)的从 $t=0$ 到 $t=4$ 的各个时刻的网络演化图(从左到右列出的是狗的形状复杂网络模型在 $t=0,1,2,3,4$ 时刻演化出的网络图).值得指出的是,本文提出网络演化机制不同于文献[18–20]提出的方法:本文提出的演化机制是由当前形状网络中内部特性(平均权重)来决定下一个时刻的演化图,因此是一种自组织的演化机制,不需要预先设置一系列的阈值参数来控制网络的动态演化.这也是本文提出的复杂网络模型与文献[18–20]给出的复杂网络模型不同的地方.根据该演化机制,后一个时刻的图是前一个时刻的子图,从 $t=0$ 时刻开始,演化产生的各个时刻的图一起构成了一个分层的形状描述框架,基于该框架,我们将对各个时刻的网络图进行测量,以分层的抽取形状的结构特征.

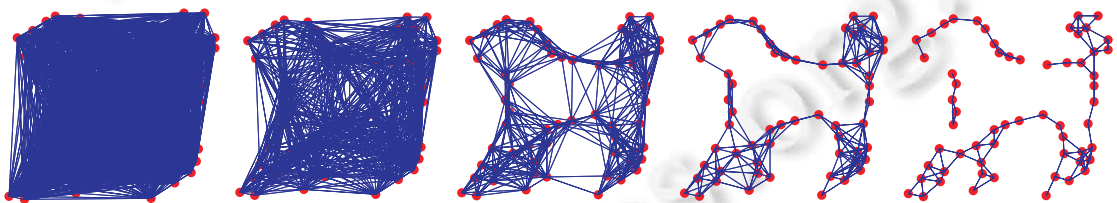


Fig.3 An example of graphically illustrating the proposed dynamic evolution scheme for the complex network
图 3 本文提出的复杂网络动态演化的一个示例

3 形状特征抽取

本节对前面建立的复杂网络进行局部和全局的测量来抽取形状的分层结构特征,以用于后面的形状识别任务.

3.1 网络的局部测量

这里,网络的局部测量的主要任务是,针对每一时刻的网络中的每一节点构造描述子.对网络中的每一个节点 v_i ,我们用其在 t 刻的度特征 $K_i^{(t)}$ 、单位权重特征 $U_i^{(t)}$ 、聚集系数特征 $C_i^{(t)}$ 、加权聚集系数特征 $\tilde{C}_i^{(t)}$ 来构造描述子,定义这 4 个特征的数学公式如下:

$$K_i^{(t)} = \sum_j a_{ij}^{(t)} \tag{3}$$

$$U_i^{(t)} = \frac{\sum_j w_{ij} a_{ij}^{(t)}}{\sum_j a_{ij}^{(t)}} \quad (4)$$

$$C_i^{(t)} = \frac{\sum_{k>j} a_{ij}^{(t)} a_{ik}^{(t)} a_{jk}^{(t)}}{\sum_{k>j} a_{ij}^{(t)} a_{ik}^{(t)}} \quad (5)$$

$$\tilde{C}_i^{(t)} = \frac{\sum_{k>j} w_{ij} w_{ik} w_{jk} a_{ij}^{(t)} a_{ik}^{(t)} a_{jk}^{(t)}}{\max_{k>j} (w_{jk} a_{jk}^{(t)}) \sum_{k>j} a_{ij}^{(t)} a_{ik}^{(t)}} \quad (6)$$

对网络中的每个节点 v_i , 组合 $t=0, \dots, T-1$ 个时刻的度、单位权、聚集系数、加权聚集系数这 4 类特征, 构成一个 $4 \times T$ 维的特征向量:

$$\varphi_i = \{K_i^{(0)}, \dots, K_i^{(T-1)}, U_i^{(0)}, \dots, U_i^{(T-1)}, C_i^{(0)}, \dots, C_i^{(T-1)}, \tilde{C}_i^{(0)}, \dots, \tilde{C}_i^{(T-1)}\} \quad (7)$$

作为节点 v_i 的描述子, 这里, T 是给定参数, 用以控制网络演化的代数. 描述子 φ_i 既抽取了网络的无权特征—度和聚集系数, 又抽取了网络的加权特征——单位权和加权聚集系数, 从而对网络的拓扑结构和节点的空间分布都进行了描述.

3.2 网络的全局测量

除了对网络进行局部测量, 用无权特征和有权特征为网络的每一个节点构造局部描述子之外, 我们还对网络进行全局测量, 对网络中的所有节点对的无权特征和有权特征的数字值分别进行比较, 并进行组合统计, 由此得到 4 类反映网络全局特性的统计特征: $R_{KU}^{(t)}$, $R_{KC}^{(t)}$, $R_{CU}^{(t)}$ 和 $R_{CC}^{(t)}$, 其数学定义如下:

$$R_{KU}^{(t)} = \frac{|\{(v_i, v_j) : (K_i^{(t)} - K_j^{(t)})(U_i^{(t)} - U_j^{(t)}) < 0\}|}{N(N-1)} \quad (8)$$

$$R_{KC}^{(t)} = \frac{|\{(v_i, v_j) : (K_i^{(t)} - K_j^{(t)})(\tilde{C}_i^{(t)} - \tilde{C}_j^{(t)}) < 0\}|}{N(N-1)} \quad (9)$$

$$R_{CU}^{(t)} = \frac{|\{(v_i, v_j) : (C_i^{(t)} - C_j^{(t)})(U_i^{(t)} - U_j^{(t)}) < 0\}|}{N(N-1)} \quad (10)$$

$$R_{CC}^{(t)} = \frac{|\{(v_i, v_j) : (C_i^{(t)} - C_j^{(t)})(\tilde{C}_i^{(t)} - \tilde{C}_j^{(t)}) < 0\}|}{N(N-1)} \quad (11)$$

这里, $|\cdot|$ 表示集合的势, $\langle v_i, v_j \rangle$ 是由节点 v_i 和 v_j 构成的节点对, N 是网络中节点的个数.

令 $\xi = \{R_{KU}^{(0)}, \dots, R_{KU}^{(T-1)}, R_{KC}^{(0)}, \dots, R_{KC}^{(T-1)}, R_{CU}^{(0)}, \dots, R_{CU}^{(T-1)}, R_{CC}^{(0)}, \dots, R_{CC}^{(T-1)}\}$ 为组合 $t=0, \dots, T-1$ 时刻的这 4 类统计特征而成的向量, 该 $4 \times T$ 维的特征向量 ξ 构成了形状的全局描述子.

4 形状距离

4.1 点到点的特征匹配

在第 3.1 节, 我们为形状边界点集中的每一个点构造了一个描述子, 通过比较两个点集中的点的描述子来度量两个形状的差异度. 令形状 A 和 B 形状的局部描述子分别为 $\{\varphi_i^A, i=1, \dots, M\}$ 和 $\{\varphi_j^B, j=1, \dots, N\}$, 这里, M 和 N 分别是形状 A 和形状 B 的边界点集表示中点的个数. 定义两个点集的距离的常用方法是 Hausdorff 距离, 这里我们用文献[22]提出的增强 Hausdorff 距离来定义两个形状的差异度. 我们首先定义形状 A 点集中的任意一点 v_i^A 和形状 B 点集中的任意一点 v_j^B 之间的距离为

$$d(v_i^A, v_j^B) = \|\varphi_i^A - \varphi_j^B\| \quad (12)$$

这里, $\|\cdot\|$ 表示 1 的范数, 则点集 A 到点集 B 的距离定义为

$$h(A, B) = \frac{1}{M - \eta_A} \sum_i \min_j d(v_i^A, v_j^B) \quad (13)$$

这里, η_A 的计算步骤为

- (1) 设置一个数组 $\zeta(1:N)$, 将其每一个元素的初始值设置为 -1;
- (2) 对点集 A 中的每一个点 v_i^A , 在点集 B 中寻找其最邻近点 v_j^B , 满足 $d(v_i^A, v_j^B) = \min_k d(v_i^A, v_k^B)$, 并置:

$$\zeta(j) \leftarrow \zeta(j) + 1;$$

- (3) 将数组 $\zeta(1:N)$ 中所有大于 0 的元素进行求和, 并赋值给 η_A .

该计算实质上是将点集 A 中的每一个点用公式(11)中定义的距离按最邻近原则映射到 B 中的某个点, 然后对 B 中所有被映射两次或两次以上的点的被映射次数减 1 后进行求和, 即得到 η_A 的值. $\eta_A=0$, 表明 A 中的所有点都映射到了 B 中不同的点; 而 $\eta_A=M-1$, 则表明 A 中的所有点都映射到了 B 中的同一个点.

同样, 定义点集 B 到点集 A 的距离为

$$h(B, A) = \frac{1}{N - \eta_B} \sum_j \min_i d(v_j^B, v_i^A) \quad (14)$$

这里, η_B 的计算步骤同 η_A , 即: 将点集 B 中的每一个点用公式(11)中定义的距离, 按最邻近原则映射到 A 中的某个点, 然后对 A 中所有被映射两次或两次以上的点的被映射次数减 1 后进行求和. 点集 A 和点集 B 的 Hausdorff 距离定义为

$$D_H(A, B) = \max(h(A, B), h(B, A)) \quad (15)$$

增强 Hausdorff 距离^[22]是对文献[23]提出的 Hausdorff 距离的一个改进, 在距离的计算中加入了两个点集之间的映射信息(公式中的变量 η_A 和 η_B 记录了这方面的信息).

4.2 全局特征匹配

这里用第 3.2 节给出的全局特征向量, 通过比较两个形状的全局特征的差异来计算形状的全局匹配距离. 记形状 A 和形状 B 的全局特征向量分别为 ξ_A 和 ξ_B , 定义形状 A 和 B 之间的全局匹配距离为

$$D_g(A, B) = \|\xi^A - \xi^B\|_1 \quad (16)$$

4.3 形状差异度量

组合前面给出的局部匹配距离 $D_L(A, B)$ 和全局匹配距离 $D_g(A, B)$, 得到如下形状 A 和形状 B 的差异度量:

$$D(A, B) = D_L(A, B) + W \cdot D_g(A, B) \quad (17)$$

这里, $0 \leq W \leq 1$ 是权重因子, 用以平衡两种距离在形状差异度量中的贡献, 在实际应用中经验给定.

5 实验结果和讨论

为评估本文提出的形状分析方法的性能, 我们将 Kimia^[24]和 MPEG-7^[25]这两个被广泛采用的形状图像测试集作为基准测试集, 通过 3 组实验来验证本文提出的形状分析方法的识别精确率和鲁棒性.

值得指出的是: 本文提出的是基于边界无序点集的形状分析方法, 因此在比较对象的选择方面, 那些参数化曲线的形状分析方法^[4-11], 尽管它们在这两个测试集上都报告了较高的识别精确率, 由于都利用了目标边界的有序信息, 与本文提出的方法都属于不同类方法, 因此不被列入比较对象. 在实验中, 我们选取了经典的形状上下文方法(shape context)^[13]、距离集方法(distance set)^[15]、轮廓点分布直方图(CPDH)^[14]和两种基于复杂网络的形状分析方法^[18, 20]作为比较对象. 这些方法都是基于边界无序点集的形状分析方法, 与本文提出的方法同类. 本文提出的方法在特征抽取阶段将形状边界重采样成不超过 300 个点, 设置的网络动态演化的迭代次数为 $T=5$, 在特征匹配阶段的权重因子为 $W=0.25$.

5.1 Kimia形状测试集

Kimia^[24]图像数据库是一个被广泛采用的标准形状测试集, 该库有 9 类形状, 每类 11 个样本, 同类形状存在

边缘扭曲、局部缺损或遮挡、类内形变等复杂的变化(如图 4 所示).在该测试集上,标准的性能评估方法是将 99 个样本中的每一个作为检索形状,与图像库中的其余形状进行差异性比较(共产生 99×98 次比较),对每一个检索形状,将比较的结果按差异度由小到大排序,检查排在前 10 的形状,统计在每一个排名位置上匹配正确的形状个数.显然,在前 10 的排名位置上,每一个最多会有 99 个正确的匹配(每一个待匹配的形状在该位置产生 1 个).

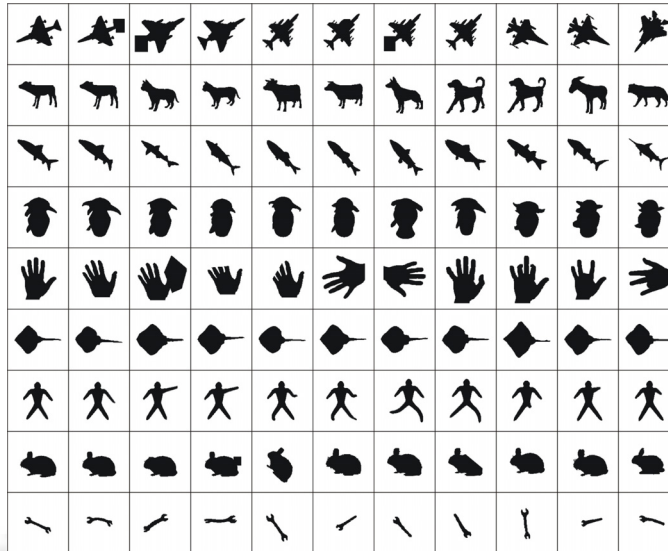


Fig.4 All the images in the Kimia shape database^[24], each row corresponds to one object category

图 4 Kimia 形状数据^[24]的所有图像,图中的每一行是同一类形状的所有样本

表 1 给出了本文提出的方法和参与比较的各方法在该测试集上的形状检索结果.检索结果给出了前 10 个排名位置的正确匹配的个数.在第 1 个排名位置,即最近邻匹配,采用本文给出的方法,每一个检索形状都找到了其正确的匹配,而 Shape Contexts^[13],CPDH^[14],Complex Networks^[18]分别产生了 2 个、1 个和 2 个错误的匹配;在后面的各个排名位置,本文提出的方法匹配正确的形状个数都大于参与比较的各方法.在前 10 个位置,本文提出的方法共产生了 932 个正确的匹配,平均检索正确率(AP)可以计算为 $932/990=94.14\%$,比 Shape Contexts($775/990=78.28\%$),CPDH($849/990=85.76\%$),Complex Networks($759/990=76.67\%$)这 3 种方法分别高出了 15.86%,8.38%,17.47%.该实验结果表明,本文提出方法的检索正确率要优于参与比较的各种方法.

Table 1 Retrieval results on Kimia shape dataset

表 1 在 Kimia 形状测试集上的检索结果

Algorithm	1	2	3	4	5	6	7	8	9	10	AP (%)
Shape Contexts ^[13]	97	88	83	79	81	78	73	73	65	58	78.28
CPDH ^[14] (表示该方法的实验数据直接引自相关文献)	98	94	95	92	90	88	85	84	71	52	85.76
Complex Network ^[18]	97	90	82	84	86	73	73	69	56	49	76.67
The proposed	99	99	98	98	97	98	95	92	78	78	94.14

表 2 则给出了本文提出的方法与 Shape Contexts^[13]和 Complex Networks^[18]的计算效率的比较.这里,检索时间定义为一个检索形状与图像库中的所有形状进行匹配的时间,对于本实验,则是进行 99 次匹配的时间.从该表可以看出:本文提出的算法耗费了 0.16s 就完成了一次形状的检索,而经典的 Shape Contexts^[13]则要耗费 19.58s,其耗时超过本文提出的方法 122 倍.值得指出的是,本文提出的方法要慢于 Complex Networks^[18]方法.这是因为本文提出的方法既有全局特征匹配,又执行了点到点的特征匹配,而后者仅进行了全局特征匹配.从综合平衡检索效率和检索精度两个方面,本文提出的方法要优于参与比较的各方法.

Table 2 Retrieval time on Kimia shape dataset

表 2 在 Kimia 形状测试集上的检索时间

Algorithm	Retrieval time (s)
Shape contexts ^[13]	19.58
Complex network ^[18]	0.05
The proposed	0.16

5.2 MPEG-7形状测试集

MPEG-7 形状图像数据库^[25]是另一个用于形状检索性能评估的标准测试集,该测试集由 70 类形状(每类 20 个样本),总共 1 400 幅图像组成(部分样本如图 5 所示).在该测试集上,被广泛采用的形状检索性能评估方法是“bull’s-eye test”测试^[4-11,13].采用该测试,将图像库中的每一个样本作为一个查询样本,跟测试集中的所有样本进行差异性比较,然后根据差异度值进行由小到大排序,统计前 $2 \times 20 = 40$ 幅图像中与查询样本属于同一类的样本数(记为 r),显然有 $r \leq 20$,计算 $r/20 \times 100\%$ 的值作为该查询样本的检索率,计算 1 400 个检索样本的检索率的平均值,作为整个测试的检索率(bull’s-eye test score).

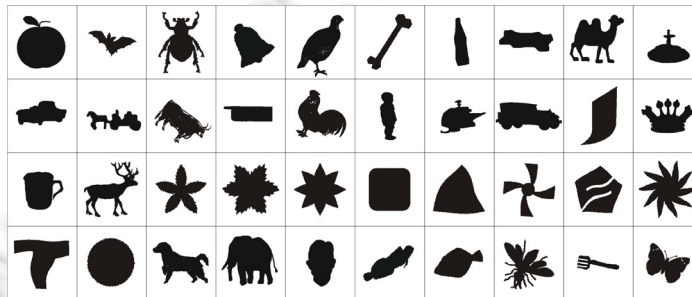


Fig.5 Part samples from the MPEG-7 image database

图 5 MPEG-7 图像库中的部分样本

表 3 列出了本文提出的方法和其他参与比较的 4 种方法的检索结果.本文提出的方法获得 78.18%的检索精确率,在表格列出的所有方法中排名第 2.排名第 1 的方法 Distance Sets 获得了 78.38%,比本文提出的方法仅高出 0.2 个百分点.但该方法跟 Shape Contexts 方法^[13]一样,在形状匹配阶段都采用了时间复杂度非常高的 Hungarian^[12]方法(该算法的时间复杂度为 $O(N^3)$ ^[13],这里, N 是点集的大小).也就是说,该方法较高检索精确率的取得是以高的计算时间复杂度为代价.而本文提出的方法在形状匹配阶段采用计算复杂度较低的 Hausdorff 距离^[26](其时间复杂度为 $O((M+N)\log(M+N))$ ^[22,26],这里, M 和 N 分别是待匹配的两个点集的大小,如两个点集大小相等,则时间复杂度为 $O(N\log N)$,要远低于 Distance Sets 方法^[15]和 Shape Contexts 方法^[13].据文献[15]报告:采用 Distance Sets 方法,执行一次两个形状的匹配(250 个采样点),在一台 Pentium III/667MHZ 的计算机上大约耗时 7×10^{-1} s.根据文献[13]报告的结果,采用 Shape Contexts 方法,执行一次两个形状的匹配计算(100 个采样点),在一台 Pentium III/500MHZ 的计算机上需要耗时 2×10^{-1} s.文献[14]则报告,采用 CPDH^[14]方法,执行一次两个形状的匹配(200 个采样点),在一台 Pentium IV/2.66GHZ 的计算机上大约耗时 1.4×10^{-2} s.而采用本文提出的方法,在一台 CPU 为 Intel Core i7-4770/3.4GHZ 的计算机上,执行一次两个形状的匹配计算(300 个采样点),仅用时 9.4×10^{-4} s.在表 4 中,我们给出了本文提出的方法和经典的 Shape Contexts^[13]方法和 Complex Network^[18]的计算时间的比较.从表中可以看出:本文提出的方法检索时间仅为 1.91s,而 Shape Contexts^[13]要耗时 328.77s,其检索速度要原低于本文提出的方法.尽管 Complex Network^[18]的计算速度要快于本文提出的方法,但其检索精确率要比本文提出的方法低 15.78 个百分点.因此,本文提出的方法有着相对较少的计算耗费和更好的检索精度,更适合图像的在线检索的需要.

Table 3 Retrieval rates on MPEG-7 shape dataset**表 3** 在 MPEG-7 形状测试集上的检索率

Algorithm	Bulls-Eye test score (%)
Shape Contexts ^[13]	76.51
CPDH ^[14]	76.56
Distance Sets ^[15]	78.38
Complex network ^[18]	62.40
The proposed	78.18

Table 4 Retrieval time on MPEG-7 shape dataset**表 4** 在 MPEG-7 形状测试集上的检索时间

Algorithm	Retrieval time (s)
Shape Contexts ^[13]	328.77
Complex Network ^[18]	0.16
The proposed	1.91

行了比较.该方法在报告的实验中选取了 MPEG-7 形状测试集中的 18 类形状,构成了 $18 \times 20 = 360$ 幅图像的测试集,然后在该测试集中进行图像检索实验,采用“bulls-eye test”检索性能评估方法,报告的检索率为 86.47%. 为便于与该方法的比较,我们对本文提出的方法在该测试集上进行了同样的测试,在实验中,我们取得了 95.26% 的检索率,高于文献[20]提出的方法超过 8 个百分点,进一步证明了本文提出的用于形状分析的复杂网络模型的优越性.

5.3 鲁棒性测试

为评估本文提出的方法的鲁棒性,我们进行了另外两组实验:第 1 组实验测试当形状的轮廓线上的点发生了缺失后,算法的鲁棒性;第 2 组实验测试算法对噪声的鲁棒性,实验方案类似于文献[18]介绍的方法.

针对第 1 组实验,将前面 Kimia 测试中用到的形状轮廓线随机地分别移去 20%,40%和 60%的点(图 6 给出了对狗的形状轮廓进行该操作的一个例子),由此得到 3 个测试集.在得到的这 3 个形状测试集中分别进行 Kimia 测试,实验结果在表 5 中列出(轮廓线上的点发生缺失).



Fig.6 Three contours (from the second left Figure to the fourth left Figure) obtained by randomly removing 20%, 40% and 60% of the points of the original dog shape contour (the first left Figure), respectively

图 6 从狗的形状轮廓(最左图)分别随机的移去 20%,40%和 60%的点的轮廓图(从左到右第 2 图~第 4 图)

Table 5 Experimental results of robust performance test**表 5** 在 Kimia 上进行鲁棒性测试的实验结果

Degradation levels (r)	Algorithm	1	2	3	4	5	6	7	8	9	10	AP (%)
$r=20\%$	Shape contexts ^[13]	95	89	84	79	78	77	73	66	62	47	75.76
	The proposed	99	98	96	97	95	92	89	88	77	70	91.01
$r=40\%$	Shape contexts ^[13]	96	86	81	76	76	75	70	62	55	41	72.53
	The proposed	99	97	97	93	93	84	76	79	69	59	85.45
$r=60\%$	Shape contexts ^[13]	89	81	78	74	74	62	57	54	52	42	66.97
	The proposed	94	88	87	88	76	79	73	74	63	55	78.49

从表 5 可看出:当移去 20%的点时,匹配正确的形状个数达到 901 个,平均检索正确率 AP 达到了 $901/990 = 91.01\%$,比原来的 93.43% 的检索率仅降低了 2.42 个百分点;而当移去的点达到 60%(观察图 6 最右边的图可以看出:在这种情况下,有多个成段的轮廓线被去掉,此时轮廓线上的点已变得非常稀疏),本文提出的方法仍然检索到了 777 个正确的形状,检索率达到了 78.49%,表明算法仍然能工作.在该表中,我们还列出经典算法 Shape Contexts^[13]的相应的实验结果,以资对比.通过对比可以看出:对于每一组实验,本文提出的方法平均检索正确率比 Shape Contexts^[13]都高出了至少 11 个百分点以上.该组实验结果表明了本文提出的算法对轮廓线上点缺失的鲁棒性.

在噪声鲁棒性测试实验中,将前面 Kimia 测试中用到的每一个采样后的形状轮廓的所有点的 x 坐标和 y 坐标分别看成一维信号,并在其中加入信噪比(SNR)分别为 40dB,35dB,30dB,25dB 的随机高斯噪声(加入噪声后的形状样本如图 7 所示),由此得到 4 个加入不同强度噪声的形状测试集.在得到的这 4 个形状测试集中分别进行 Kimia 测试,实验结果在表 6 中列出.



从左到右:原形状、分别加入 SNR=40dB,35dB,30dB,25dB 随机高斯噪声的形状

Fig.7 Shape samples distorted by random Gaussian noises of different intensities

图 7 加入了不同强度高斯噪声的形状

Table 6 Experimental results of robust performance test (noise interference) on the Kimia shape dataset

表 6 在 Kimia 上进行鲁棒性测试(噪声干扰)的实验结果

Noise intensity (dB)	Algorithm	1	2	3	4	5	6	7	8	9	10	AP (%)
SNR=40	Shape contexts ^[13]	96	91	84	79	78	75	77	68	60	55	77.07
	The proposed	99	97	97	97	95	96	92	89	84	66	92.12
SNR=35	Shape contexts ^[13]	95	90	86	79	79	75	73	67	63	50	76.46
	The proposed	99	97	97	94	94	92	90	85	82	68	90.71
SNR=30	Shape contexts ^[13]	94	88	82	81	79	73	73	64	66	49	75.66
	The proposed	99	97	94	96	94	91	90	85	75	65	89.49
SNR=25	Shape contexts ^[13]	94	86	84	72	77	70	69	59	54	46	71.82
	The proposed	94	92	91	94	91	88	86	75	67	55	84.24

从表 6 可看出:在 4 种强度的噪声影响下,本文提出的算法仍然分别取得了 $912/990=92.12\%$, $898/990=90.71\%$, $886/990=89.49\%$, $834/990=84.24\%$ 的平均检索正确率.特别是在高强度噪声干扰下(SNR=25dB),本文提出的方法仍然取得了较高的检索率 84.24%.在与经典算法 Shape Contexts^[13]的比较中,对于 4 组实验,本文提出的算法每一组实验都要比之高出 12 个百分点以上,表明了本文提出的算法对噪声干扰的鲁棒性.

6 结 论

本文提出了一种新的基于复杂网络模型的目标轮廓无序点集形状的分析方法.该方法用自主的网络动态演化机制,分层地抽取形状特征,通过对网络的局部测量和全局测量获取网络的无权特征和加权特征,构成形状的局部描述子和全局描述子.该描述子为目标形状的认识提供了强辨识能力的特征,仅用低计算复杂度的 Hausdorff 距离(局部匹配)和 L_1 距离(全局匹配),就能有效地进行形状的认识.用基准的图像测试集和标准性能评估方法对本文提出的方法进行了测试,实验结果表明:本文提出的方法具有较高的识别精确率和鲁棒性,其综合性能要优于同类方法——形状上下文^[13]、距离集^[15]、轮廓点分布直方图(CPDH)^[14]和其他两种基于复杂网络的形状分析方法^[18,20],能够有效地用于解决边界无序点集的形状识别这一形状分析任务中的具有挑战性的问题.将该方法拓展到一般的点集模式的描述与匹配,将是下一步的研究工作.

References:

[1] Costa L da F, Jr. Cesar RM. Shape Analysis and Classification: Theory and Practice. 2nd ed., CRC Press LLC., 2001. 1–25.
 [2] Zhang D. Review of shape representation and description techniques. Pattern Recognition, 2004,37(1):1–19. [doi: 10.1016/j.patcog.2003.07.008]
 [3] Freeman H. On the encoding of arbitrary geometric configurations. IRE Trans. on Electronic Computers, 1961,10(2):260–268. [doi: 10.1109/TEC.1961.5219197]
 [4] Adamek T, O'Connor NE. A multiscale representation method for nonrigid shapes with a single closed contour. IEEE Trans. on Circuits and Systems for Video Technology, 2004,14(5):742–753. [doi: 10.1109/TCSVT.2004.826776]

- [5] Ling H, Jacobs DW. Shape classification using the inner-distance. *IEEE Trans. on Pattern Analysis Machine Intelligence*, 2007,29(2): 286–299. [doi: 10.1109/TPAMI.2007.41]
- [6] Alajlan N, Rube IE, Kamel MS, Freeman G. Shape retrieval using triangle-area representation and dynamic space warping. *Pattern Recognition*, 2007,40(7):1911–1920. [doi: 10.1016/j.patcog.2006.12.005]
- [7] Xu C, Liu J, Tang X. 2D shape matching by contour flexibility. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2009,31(1):180–186. [doi: 10.1109/TPAMI.2008.199]
- [8] Biswas S, Aggarwal G, Chellappa R. An efficient and robust algorithm for shape indexing and retrieval. *IEEE Trans. on Multimedia*, 2010,12(5):372–384. [doi: 10.1109/TMM.2010.2050735]
- [9] Bai X, Yang X, Latecki LJ, Liu W, Tu Z. Learning context-sensitive shape similarity by graph transduction. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2010,32(5):861–874. [doi: 10.1109/TPAMI.2009.85]
- [10] Wang J, Bai X, You X, Liu W, Latecki LJ. Shape matching and classification using height functions. *Pattern Recognition Letters*, 2012,33(2):134–143. [doi: 10.1016/j.patrec.2011.09.042]
- [11] Wang B, Gao Y. Hierarchical string cuts: A translation, rotation, scale and mirror invariant descriptor for fast shape retrieval. *IEEE Trans. on Image Processing*, 2014,23(9):4101–4111. [doi: 10.1109/TIP.2014.2343457]
- [12] Papadimitriou C, Steiglitz K. *Combinatorial Optimization: Algorithms and Complexity*. New York: Dover Publications, 1998. 248–254.
- [13] Belongie S, Malik J, Puzicha J. Shape matching and object recognition using shape contexts. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2002,24(4):509–522. [doi: 10.1109/34.993558]
- [14] Shu X, Wu X.J. A novel contour descriptor for 2D shape matching and its application to image retrieval. *Image and Vision Computing*, 2011,29(4):286–294. [doi: 10.1016/j.imavis.2010.11.001]
- [15] Grigorescu C, Petkov N. Distance sets for shape filters and shape recognition. *IEEE Trans. on Image Processing*, 2003,12(10): 1274–1286. [doi: 10.1109/TIP.2003.816010]
- [16] Albert R, Baravási A. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 2002,74(1):47–97. [doi: 10.1103/RevModPhys.74.47]
- [17] Costa LDF, Rodrigues FA, Traverso G, Villas Boas PR. Characterization of complex networks: A survey of measurements. *Advances in Physics*, 2007,56(1):167–242. [doi: 10.1080/00018730601170527]
- [18] Backes AR, Casanova D, Bruno OM. A complex network-based approach for boundary shape analysis. *Pattern Recognition*, 2009, 42(1):54–67. [doi: 10.1016/j.patcog.2008.07.006]
- [19] Backes AR, Bruno OM. Shape classification using complex network and multi-scale fractal dimension. *Pattern Recognition Letters*, 2010,31(1):44–51. [doi: 10.1016/j.patrec.2009.08.007]
- [20] Tang J, Chen ZZ, Luo B, Sun DD. Shape descriptor and matching based on complex network and OSB. *Acta Electronica Sinica*, 2011,39(8):1757–1765 (in Chinese with English abstract).
- [21] Tang J, Zhi DP, Jiang B, Luo B. Shape description and recognition based on directed complex network. *Journal of Computer-aided Design & Computer Graphics*, 2014,26(11):2039–2045 (in Chinese with English abstract).
- [22] Gope C, Kehtarnavaz N. Affine invariant comparison of point-sets using convex hulls and Hausdorff distances. *Pattern Recognition*, 2007,40(1):309–320. [doi: 10.1016/j.patcog.2006.04.026]
- [23] Dubuisson M, Jain A. A modified hausdorff distance for object matching. In: *Proc. of the 12th Int'l Conf. on Pattern Recognition*. IEEE, 1994. 566–568. [doi: 10.1109/ICPR.1994.576361]
- [24] Sebastian TB, Klein PN, Kimia BB. Recognition of shapes by editing their shock graphs. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2004,26(5):550–571. [doi: 10.1109/TPAMI.2004.1273924]
- [25] Latecki LJ, Lakamper R, Eckhardt U. Shape descriptors for non-rigid shapes with a single closed contour. In: *Proc. of the 2000 IEEE Conf. on Computer Vision and Pattern Recognition*. IEEE, 2000. 424–429. [doi: 10.1109/CVPR.2000.855850]
- [26] Alt H, Behrends B, Blomer J. Measuring the resemblance of polygonal shapes. In: *Proc. of the 7th ACM Symp. on Computer Geometry*. New York: ACM Press, 1992. 102–109.

附中中文参考文献:

- [20] 汤进,陈展展,罗斌,孙登第.基于复杂网络和最优子序列双射的形状描述与匹配方法. *电子学报*,2011,39(8):1757–1765.
- [21] 汤进,鄧大鵬,江波,罗斌.基于有向复杂网络模型形状描述与识别. *计算机辅助设计与图像学学报*,2014,26(11):2039–2045.



王斌(1969—),男,博士,教授,硕士生导师,CCF高级会员,主要研究领域为计算机视觉,图像处理,模式识别.