

# 形状匹配方法研究与展望

周瑜<sup>1</sup> 刘俊涛<sup>1,2</sup> 白翔<sup>1</sup>

**摘要** 形状匹配及分类是计算机视觉中的重要问题。近年来,以形状上下文为代表的基于轮廓的形状匹配方法和以奇点图为代表的基于骨架的形状匹配方法获得了长足的发展。本文介绍了形状匹配问题的基本概念,分析了形状匹配问题的难点,按照基于轮廓和基于骨架的分类方法对近年来最新出现的形状表示与形状匹配的方法进行了详尽的介绍,并介绍了基于度量学习的形状检索方法,本文还详细介绍了近年来形状匹配研究领域常用的一些测试数据库,之后对局部形状匹配和形状分类等有潜力的研究方向进行了展望。最后对形状匹配的整体框架及其应用前景进行了总结。

**关键词** 形状表示, 形状匹配, 形状分类, 轮廓, 骨架

**引用格式** 周瑜, 刘俊涛, 白翔. 形状匹配方法研究与展望. 自动化学报, 2012, 38(6): 889–910

**DOI** 10.3724/SP.J.1004.2012.00889

## Research and Perspective on Shape Matching

ZHOU Yu<sup>1</sup> LIU Jun-Tao<sup>1,2</sup> BAI Xiang<sup>1</sup>

**Abstract** Shape matching and classification are important issues in computer vision. In recent years, contour-based shape matching approaches (e.g., shape context) and skeleton-based shape matching methods (e.g., shock graph) both have a lot of developments. In this paper, we introduce the basic concept of shape matching, give the difficulties of this topic, and provide a detailed review on the most recent approaches about shape representation and matching for both contour-based and skeleton-based methods. We also give a brief introduction about metric learning based shape retrieval. Moreover, we introduce some widely used benchmarks for shape matching in details, along with some hot topics including partial shape matching, shape classification, etc. Finally, this paper concludes with the whole framework of shape matching and the application perspective of this topic.

**Key words** Shape representation, shape matching, shape classification, contour, skeleton

**Citation** Zhou Yu, Liu Jun-Tao, Bai Xiang. Research and perspective on shape matching. *Acta Automatica Sinica*, 2012, 38(6): 889–910

形状匹配是计算机视觉中的重要问题,已经被应用在多个领域,如目标识别、图像检索、文字识别、人脸识别、医学图像分析、机器人导航、传感器网络等。形状匹配,作为计算机视觉的基本问题,尽管已经被研究了几十年,但近几年来国际上对此问题的研究依旧十分活跃。不少关于形状匹配的文章仍然陆续在国际顶级期刊和顶级会议上发表,例如国际顶级期刊: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (PAMI); *International*

*Journal of Computer Vision* (IJCV); *IEEE Transactions on Image Processing* (TIP); *Computer Vision and Image Understanding* (CVIU) 等, 和国际顶级会议: ICCV, ECCV, CVPR, NIPS, IJCAI 等。最近, PAMI 还就形状识别问题组织了专刊征文 (*Special Issue of IEEE Transactions on Pattern Analysis and Machine Intelligence on Shape Analysis and Its Applications in Image Understanding*, October 2008)。近年来,研究者在形状匹配问题上已经取得了显著的进步,涌现了很多新的形状匹配算法。这就要求对形状匹配算法进行归纳和总结,尤其是目前形状匹配中的热点问题和重要发展方向。本文不但介绍了形状匹配的基本概念,而且对重要形状匹配算法进行了介绍,同时还对形状匹配领域的未来发展方向进行了预测和讨论。

## 1 形状匹配的基本概念

在计算机视觉和模式识别中,形状通常由目标范围的二值图像所表示。不同于纹理、颜色等其他基本特征,形状信息是一种高级别 (High level) 的视

收稿日期 2011-11-19 录用日期 2012-02-10  
Manuscript received November 19, 2011; accepted February 10, 2012

国家自然科学基金 (60873127, 60903096), 河北省自然科学基金资助项目 (F2009001435) 资助

Supported by National Natural Science Foundation of China (60873127, 60903096) and Natural Science Foundation of Hebei Province (F2009001435)

本文责任编辑 姚力

Recommended by Associate Editor YAO Li

1. 华中科技大学电子与信息工程系 武汉 430074 2. 军械工程学院计算机工程系 石家庄 050003

1. Department of Electronics and Information Engineering, Huazhong University of Science and Technology, Wuhan 430074

2. Department of Computer Engineering, Mechanical Engineering Institute, Shijiazhuang 050003

觉信息. 如图 1 所示, 尽管这些二值图表示的形状失去了纹理、颜色等其他特征, 但这并不影响人眼去识别它们的类别.



图 1 二值图表示的物体形状

Fig. 1 Object shapes represented by binary images

所谓形状匹配, 就是按照一定的准则来衡量形状间的相似程度. 两个形状间的匹配结果一般用一个数值来表示, 称为形状相似度 (Shape similarity). 形状相似度的值越大, 表示两个形状越相似. 有些形状匹配方法是求取形状间的非相似度 (Shape dissimilarity), 通常它被称为形状间的距离 (Shape distance). 与形状相似度相反, 距离值越大, 表示两个形状越不相似. 无论是计算形状相似度还是形状距离, 衡量形状匹配方法性能的最重要的指标在于以它们的值为排序依据的形状检索的精度. 如文献 [1] 所述, 形状匹配方法分为两个大类: 一类是计算各种变换下的不变量的差值; 另一类是通过寻找目标形状和模型之间的局部对应关系, 使匹配误差最小, 获得的最小值就是形状距离. 早期的形状匹配方法如不变矩、傅里叶描述子等都属于第一类方法. 但由于此类方法往往只考虑全局形状特征, 丢失了很多重要的形状细节, 其效果往往不能令人满意. 而第二类方法是整体形状和局部形状的紧密结合, 最近十年发表的形状匹配方法大多属于此类. 由于文献 [1] 已经对 2000 年以前的方法作了详细的介绍, 这里不再重述. 本文介绍的重点为近十年来形状匹配的进展和未来有潜力的发展方向, 希望能有助于读者完整地认识形状匹配问题并获得启示.

### 1.1 形状匹配的基本流程

形状匹配的主要过程包括形状表示、特征点匹配和度量学习三个步骤, 其主要应用于数据库形状的检索和识别. 图 2 给出了传统形状匹配、传统形状检索以及基于度量学习的形状检索的流程图, 下面将对这一流程图给出具体解释.

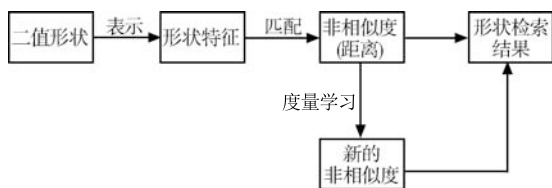


图 2 形状匹配流程图

Fig. 2 Flow chart of shape matching

对于一个二值形状, 先是要提出形状特征描述子. 该描述子的形状区分能力的强弱将直接影响形状识别的结果, 因此大多数形状匹配的工作都集中在形状特征提取方面. 匹配过程是找到一对不同形状之间的整体及局部的对应关系, 对应关系的准确性也将会对之后得到的非相似度的区分能力产生直接影响, 因此也有不少工作在匹配算法方面展开了深入的研究. 度量学习方法是通过对已知的数据库形状的上下文 (Contextual) 信息来改进原有距离度量的方法, 这一步骤能将原有距离的区分性能大幅度提升.

### 1.2 形状匹配的基本性质

对于一个形状集合  $S$ , 其距离函数  $d: S \times S \rightarrow \mathbf{R}$ . 假设  $A, B$  和  $C$  是  $S$  中的形状, 它们将满足如下性质:

- 1)  $d(A, B) \geq 0$  (非负性).
- 2)  $d(A, A) = 0$  (等值性).
- 3) 当  $d(A, B) = 0$ , 说明  $A = B$  (唯一性).
- 4)  $d(A, B) + d(B, C) \geq d(A, C)$  (三角不等式).

与欧氏空间的情况不同, 在形状空间中三角不等式并非一直适用. 例如局部相似 (Partial similarity) 的情况, 当  $A$  的一部分与  $B$  的一部分十分相似的时候,  $d(A, B)$  的值会比较小. 如图 3 中的马 (horse)、人 (man) 和人马 (centaur), 局部相似性造成  $d(\text{man}, \text{centaur})$  和  $d(\text{centaur}, \text{horse})$  的值比较小, 而人和马是完全不相似的, 此时  $d(\text{man}, \text{centaur}) + d(\text{centaur}, \text{horse}) \geq d(\text{man}, \text{horse})$  不再成立. 因此需要将 4) 改写成:

- 5)  $c(d(A, B) + d(B, C)) \geq d(A, C)$ , 其中  $c$  为大于 1 的常数 (松弛三角不等式).

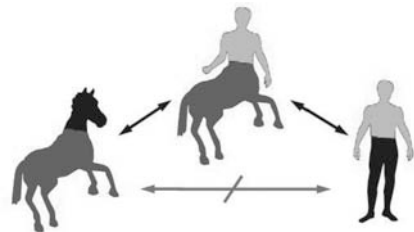


图 3 局部相似不满足三角不等式的情况

Fig. 3 A case that partial similarity does not meet triangle inequality

- 6)  $d(A, B) = d(B, A)$  (对称性).

对称性并不是需要一直满足的性质, 因为人类视觉感观往往不认为  $A$  与  $B$  的相似程度等于  $B$  与  $A$  的相似程度.

- 7) 对任一种变换方式  $g$ , 都有  $d(g(A), g(B)) =$

$d(A, B)$  (不变性).

在目标识别中, 往往要求形状相似度不受仿射变换的影响.

### 1.3 形状匹配的难点问题

形状匹配中的难点问题主要包括: 1) 由各种变换 (例如仿射变换等) 带来的形状差异; 2) 物体的自遮挡和互遮挡; 3) 非刚性物体的运动 (如: 马奔跑时带来的形状变化); 4) 同类物体间的变化 (如: 人有高矮胖瘦的区别); 5) 图像噪声和分割错误. 这些问题都会给形状匹配带来困难. 尤其是问题 2) 和 5), 现有的方法依然无法处理好有严重遮挡的情况. 而噪声和分割错误是由于现有的图像分割算法远未达到从真实图像中准确获得物体边缘造成的.

### 1.4 形状匹配方法的分类

形状匹配方法有很多不同的分类方法. 一种比较直观的方式根据形状表示的方式来对其进行分类. 形状通常可以被点集 (Point set), 轮廓 (Contour) 和骨架 (Skeleton) 等形状描述符所表示. 相比点集, 轮廓多了物体边界的顺序关系, 是一种含更多信息量的形状表示方式. 而骨架, 又被称为物体的中轴 (Medial axis)<sup>[2]</sup>, 不但包含了物体形状的几何特征, 而且表示了物体的拓扑结构 (物体各个部分的连接关系). 由于不同的形状描述方法具有不同的特点, 基于它们的形状匹配方法也会有所区别. 为便于更加清晰地介绍各种形状匹配方法的贡献, 这里按照图 2 中流程图的三个步骤: 形状表示、特征点匹配及形状度量学习来对各类方法进行归类. 其中形状表示又分成基于轮廓 (包括点集和区域) 的匹配方法和基于骨架的形状匹配方法来进行具体的介绍.

## 2 形状表示方法

### 2.1 基于轮廓边界点集的形状匹配方法

近年来, 在形状匹配领域, 人们习惯于将物体看作是一系列离散点的集合, 物体的形状信息就可以通过表示物体的离散点集合或其子集来获得. 基于轮廓的形状匹配方法一般包括以下两个步骤: 1) 对轮廓序列集合上的点进行描述 (表示); 2) 对经过表示的点集序列进行匹配, 获得形状之间的非相似度量 (距离). 下面, 我们首先对基于轮廓的形状表示进行介绍. 传统的基于轮廓的形状表示方法有很多, 本节不再一一介绍, 本节主要关注于最近十年来一些引起广泛关注的新技术、新理论. 为此, 本节对基于轮廓的形状表示的分类进行了重新的划分, 主要分为四大类方法, 分别是基于轮廓点空间位置关系的方法、多尺度方法、变换域方法和基于轮廓区域的新进展. 图 4 给出了基于轮廓的形状表示方法的

分类以及各种方法所属的类别. 下面就这四类方法进行详细的介绍.

#### 2.1.1 基于空间位置关系的形状表示方法

基于空间位置关系的形状表示方法是近年来最为重要的方法, 这类方法通过统计轮廓序列上的点的空间位置分布关系来对形状进行描述, 取得了很好的表示效果. 这类方法主要涉及三个问题: 1) 在什么样的范围内进行统计; 2) 以怎么样的标准和方法来进行统计; 3) 统计信息的表达能力是否强, 信息量是否丰富. 研究者们围绕这三个问题设计出了多种基于空间位置关系的形状表示方法.

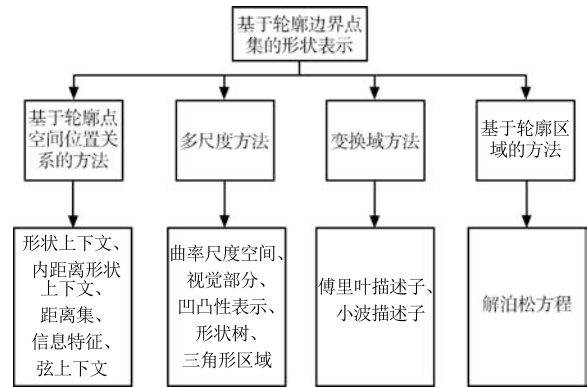


图 4 基于轮廓的形状表示分类

Fig. 4 Classification of shape representation based on contour

Belongie 等<sup>[3]</sup> 提出的形状上下文 (Shape context, SC) 描述方法, 这种描述方法着重考察轮廓序列上的某个点与该轮廓序列上的其他所有点的空间分布关系, 对于轮廓序列上的每一个点, 都用一个向量来描述, 这样的表示方法含有丰富的信息, 对形状的描述能力非常强. 考虑形状轮廓序列上的某一个点  $p_i$  和另一个形状上的某个点  $q_j$ , 用  $C_{i,j} = C(p_i, q_j)$  表示两个点的匹配损耗, 因为形状上下文信息是点分布的统计直方图表示, 点与点之间的相似度用  $\chi^2$  距离表示:

$$C_{i,j} = C(p_i, q_j) = \frac{1}{2} \sum_{k=1}^K \frac{[h_i(k) - h_j(k)]^2}{h_i(k) + h_j(k)} \quad (1)$$

形状上下文使用一组直方图来表示整个形状, 每个轮廓上的抽样点都用一个统计直方图来表示, 如图 5 所示, 图 5(a) 中以圆圈表示的点和图 5(b) 中以菱形表示的点分别使用图 5(d) 和图 5(e) 所示的统计直方图表示来进行描述, 统计直方图的横坐标是角度  $\theta$ , 表示其他轮廓点在待表示轮廓点的哪个空间方向上. 纵坐标是  $\log r$ , 其中  $r$  表示形状轮廓上的点与待表示的轮廓点之间的欧氏距离. 由图 5

可以看出处在两个形状中的大致相同位置的两个点的统计直方图是非常相似的. 而图 5 (a) 中用三角形表示的点的统计直方图 (图 5 (f)) 与上述的两个点有非常大的区别.

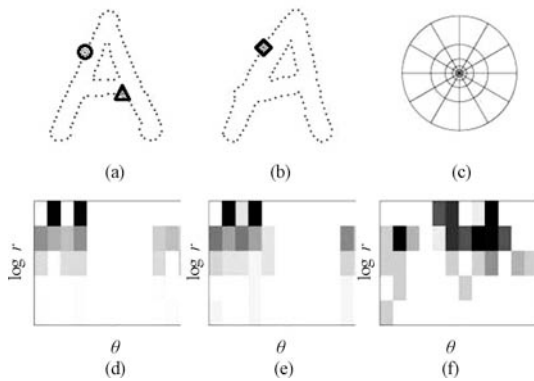


图 5 形状上下文描述子  
Fig. 5 Shape context descriptor

一种好的形状描述方法应该满足以下一些不变性: 1) 在尺度 (Scaling) 变化和发生平移 (Translation) 时保持不变; 2) 在发生小的几何扭曲 (Geometrical distortions), 存在异常点的时候具有较好的鲁棒性; 3) 在发生旋转 (Rotation) 变换时保持不变, 或者至少得到一组仿射变换 (Affine transformations). 形状上下文在这些方面均具有优良的性质, 下面进行简单的介绍.

平移不变性是形状上下文的本质属性, 因为这种描述方法是考察了整个形状上的点的空间相对位置关系, 平移变化不会改变形状轮廓点的空间相对位置关系. 为了获得尺度不变性, 该方法用形状中  $n^2$  个点对的平均距离  $\alpha$  归一化了所有的径向距离 (Radial distance). 形状上下文具有完全的旋转不变性, 这是因为在计算描述子的过程中使用的不是绝对坐标, 而是相对坐标, 该方法把每个点的切向量 (Tangent vector) 方向作为  $x$  轴的正方向, 在这种度量方式下, 参考坐标系和切角度发生同样的转动, 所以得到的结果是一个具有完全的旋转不变性的描述子. 这种度量方法也具有一定的局限性, 例如, 当要识别的对象是 6 和 9 的时候, 旋转不变性是完全不适合的. 另外一个问题是许多点没有定义良好的切线 (Tangent). 形状上下文是一种信息丰富的描述方法, 对局部噪声点是不敏感的, 文献 [3] 通过实验证明了形状上下文描述子对小的非线性形变以及存在异常点的情况具有良好的鲁棒性.

Daliri 等<sup>[4]</sup> 将形状上下文描述方法与符号串表示 (String of symbols) 结合, 提高了形状匹配的效果, 该方法首先用形状上下文描述方法表示形状, 使用动态规划算法找到两个形状轮廓序列上点的对应

关系, 以此为依据去除轮廓序列中的异常点, 然后使用 Procrustes 分析的方法来校准两个形状的对应关系, 最后使用符号串来表示轮廓上的每个点, 用编辑距离 (Edit distance) 计算两个形状之间的距离, 最终获得两个形状之间的匹配关系.

Ling 等<sup>[5-6]</sup> 在形状上下文统计直方图表示的基本思想指导下, 用轮廓点之间的内距离代替了形状上下文算法中使用的轮廓点之间的欧氏距离, 并提出了内距离形状上下文 (Inner-distance shape context, IDSC) 方法. 内距离定义为在形状内部连接两个轮廓序列上两个点的最短路径的长度, 该方法在对非刚性物体的表示上取得了较好的效果. Biswas 等<sup>[7]</sup> 使用内距离形状上下文作为形状描述子, 结合自己提出的检索框架, 在形状检索方面取得了不错的效果.

Grigorescu 等<sup>[8]</sup> 同样利用空间分布关系信息得到一组称为距离集 (Distance sets) 的局部描述子, 该方法使用形状轮廓序列上某点  $p$  与其  $i$  个邻居轮廓点之间的距离作为点  $p$  的局部描述子. Tu 等<sup>[9-10]</sup> 定义了一组信息特征 (Informative feature), 这种表示方法由局部特征和全局特征两部分组成, 局部特征在一个小的局部范围内考察目标轮廓点及其局部邻居轮廓点之间的夹角的统计关系, 全局特征在整个形状轮廓上考察目标轮廓点及其全局邻居轮廓点之间的夹角的统计关系.

Yang 等<sup>[11]</sup> 定义了弦上下文描述子 (Chord context), 对于一个给定的形状, 弦上下文描述了不同方向上的弦长的频率分布, 该方法不需要定义采样点或者关键点, 也不需要知道轮廓的顺序信息, 能够同时获得轮廓信息和区域信息.

基于空间位置分布的形状上下文描述方法自发表以来引起了广泛的关注, 并被引入到很多其他具体应用当中. Mori 等<sup>[12]</sup> 基于形状上下文描述子, 给出了两种快速修剪 (Fast pruning) 的方法. 一种叫做形状上下文代表 (Representative shape contexts), 这种方法只计算待检索形状的一部分轮廓序列点的形状上下文特征来表示整个形状, 部分轮廓序列点的选择是使用拒绝采样 (Rejection sampling) 的方法. 第二种快速修剪算法叫做 Shapemes, 这种方法是对形状上下文描述子进行向量量化 (Vector quantization). 这两种快速修剪算法能够有效地在一个大的候选数据库中找到一个小的候选子集. 这就大大提高了形状或者图像检索的效率. Thayananthan 等<sup>[13]</sup> 将形状上下文描述子与倒角匹配 (Chamfer matching) 相结合, 来对复杂场景中的物体进行匹配. Zhang 等<sup>[14]</sup> 使用形状上下文距离学习得到了一组判决式分类器, 该方法能有效地进行物体识别.

### 2.1.2 基于多尺度理论的形状表示方法

基于多尺度理论的形状表示也是近年常见的方法, 这类方法主要使用了平面曲线的凹凸特性. 曲线的凹凸特性是人类视觉识别的重要特性. 代表方法有 Adamek 等<sup>[15]</sup> 提出的 MCC 描述子, Mokhtarian 等<sup>[16]</sup> 提出的 CSS 描述子, 以及 Alajlan 等<sup>[17-18]</sup> 提出的 TAR 描述子等等. 这类方法首先提取轮廓上的点, 然后以不同的描述方法来获得曲线的凹凸特性, 然后通过这些凹凸特性来判定形状轮廓的性质.

Mokhtarian 等<sup>[16]</sup> 使用曲率尺度空间的描述方法来进行有效地形状检索, 该方法将形状的弧长参数  $C(s)$  与一维高斯核进行卷积运算, 核函数的窗口宽度为  $\sigma$ , 随着窗口宽度值  $\sigma$  的增加, 形状的轮廓将变得平滑, 轮廓上的曲率零交叉点 (Zero-crossing) 的数量将会不断减少, 直到整个形状轮廓是凸的. 随着窗口函数  $\sigma$  的增大, 曲率零交叉点的位置会在轮廓上连续的移动, 直到两个这样的曲率零交叉点相遇并消失. 计算每个轮廓的曲率尺度空间图 (CSS image), 然后用最大的 CSS 轮廓作为每个形状的描述子. 这种形状的描述方法对边界噪声, 尺度变化和旋转变换具有很好的鲁棒性, CSS 形状描述子是 ISO/IEC MPEG-7 标准中的两个形状描述子之一. 文献 [19] 对基于曲率值的方法进行了系统的介绍.

Latecki 等<sup>[20]</sup> 使用视觉部分 (Visual part) 对形状进行描述, 与 CSS 不同之处在于该方法使用离散曲线演化的方法来对形状进行简化, 该方法用一条直线代替两个点之间的连续线段, 从而对形状进行简化, 该方法的关键性质是上述代替的顺序, 该方法定义了一个轮廓段重要性函数, 通过该重要性函数对轮廓段的重要性进行评估, 然后依次删除对整个形状贡献小的点, 从而对整个形状进行简化.

Adamek 等<sup>[15]</sup> 提出了多尺度凹凸表示 (MCC) 的方法来对形状进行描述, 与文献 [16] 类似, 该方法同样使用弧长参数  $C(s)$  与一维高斯核进行卷积运算的方法对形状轮廓进行简化, 不同之处在于对形状的凹凸特性的定义上, 该方法用轮廓上某一个点在两个相邻尺度上的欧氏距离  $d(u, \sigma)$  来度量形状的凹凸特性, 形状的凹度或凸度越大,  $d(u, \sigma)$  值就越大, 通过对该欧氏距离赋予符号权重 +1 (代表凸形状) 或者 -1 (代表凹形状) 来区分形状是凸的还是凹的, 当点  $p(u, \sigma - 1)$  在其上一个尺度的轮廓  $C$  的内部时, 曲线是凸的, 否则曲线是凹的. 与用曲率描述形状的凹凸性质相比, 这种描述方法能够有效地抵抗局部噪声的影响. 使用上述定义的形状凹凸度, 集合点的位置和尺度变化, 得到了三维的 MCC 形状描述子, 第一维表示轮廓序列上的所有点, 第二维表示尺度变化, 第三维表示形状的凹凸度量. 这种形状描述方法对平移、旋转、尺度变化以及轻微的

几何形变具有不错的鲁棒性. 该方法的计算复杂度较高, 为  $O(KN^2)$ , 其中  $K$  表示有选择的尺度层数,  $N$  表示轮廓点的个数.

Alajlan 等<sup>[17-18]</sup> 使用三角形区域表示 (TAR) 的方法来对形状进行描述, 与 CSS 表示相比, 这种描述方法定义了新的描述形状凹凸特性的方法, 同时不再是通过轮廓进行演化来获得多尺度性质, 该方法使用由轮廓序列上的点组成的三角形来描述形状, 并通过控制三角形的边长来获得多尺度性质, 也就是说通过一个可变边长的三角形作为形状描述的基础. 而对于形状的凹凸性质的分辨, 该方法使用三角形的面积  $TAR(n, ts)$  来表征. 按轮廓的顺时针方向来看, 当  $TAR(n, ts)$  为正值时表示轮廓是凸的, 当  $TAR(n, ts)$  为负值时表示轮廓是凹的, 当  $TAR(n, ts)$  值为零时表示是一条直线. 该方法是在各个尺度上分别对  $TAR$  描述子进行归一化, 采用这样的归一化方法的性能好于其他的在整体上对描述子进行归一化的方法. 该方法能够有效地获得形状局部和全局的信息, 对于平移、旋转和尺度都有较好的不变性. 该描述子的计算复杂度为  $O(N^2)$ , 其中  $N$  表示轮廓点的个数. 文献 [18] 在此基础上, 使用曲率树 (Curvature trees) 来表示一个多对象二值图的结构, 即物体对象和洞的包含关系, 曲率树在第零层有唯一的一个根节点, 表示背景, 第一层存储形状的外轮廓, 在第二层存储形状的洞, 等等. 曲率树上的每一层的表示均使用  $TAR$  描述子. 曲率树的计算复杂度为  $O(nN^2)$ , 其中  $n$  表示对象的个数,  $N$  表示轮廓点的个数.

Attalla 等<sup>[21]</sup> 使用多分辨率多边形 (Multi-resolution polygonal) 来对形状进行描述, 这种方法首先对形状轮廓进行等分, 通过控制等分点的个数  $n$  来获得多尺度特性, 然后对等分后的每个轮廓段提取一组三元的描述子: 中心延长线与弦之间的起始角度  $a$ , 轮廓段起始点到中心的距离  $lw$ , 弦长与弧长的比值  $m$ , 其中参数  $m$  考察了形状的平滑性质. 三元描述子  $(a, lw, m)$  都用各自的最大值来进行归一化, 使得三种描述对形状表示具有相同的贡献. 这种描述子对尺度、旋转和平移具有不变性. 该方法的计算复杂度为  $O(N)$ , 其中  $N$  表示轮廓点的个数.

Felzenszwalb 等<sup>[22]</sup> 使用形状树 (Shape tree) 作为形状描述的方法, 该方法使用层次化的表示来表示开曲线, 这种层次化的表示能够在不同的分辨率上获得形状信息, 所以该方法也是一种多尺度的形状描述方法. 对于一条给定的开曲线, 用一系列曲线上离散点表示该曲线, 然后找到这些离散点的中点, 用中点对整个开曲线进行二分, 得到的两个子开曲线, 再用类似的方法进行划分, 迭代进行这样的

划分,将每次划分得到的新线段放在上次划分的下一个层次,从而将一个连续的线段转化成为一个离散的树状结构,这就是形状树.对于闭合曲线,形成形状树的方法与开曲线相同,是开曲线的一种特殊情况,即曲线的起点等于曲线的终点.形状树描述方法对曲线上的点进行稠密抽样,从而能够完全地捕获形状的几何形变.

### 2.1.3 基于轮廓区域的现代表示方法

基于轮廓区域的方法是形状表示的一种重要方法,近年来又发展起了一些全新的基于轮廓区域的方法,取得了很好的效果. Gorelick 等<sup>[23]</sup>通过解泊松方程来对形状进行表示,是一种基于区域的方法.该方法对于每个轮廓内部的点给定一个值,这个值表示使用随机行走的方法从该点出发到达边界点所需的时间,可以通过求解泊松方程得到该值,用轮廓作为边界判决条件.该方法能够有效提取不同的形状性质,包括局部结构 (Part structure), 粗糙的骨架 (Rough skeleton), 局部方向 (Local orientation), 不同局部的纵横比 (Aspect ratio of different parts), 分辨出轮廓的凸的部分和凹的部分.

### 2.1.4 基于变换域方法的新进展

小波分析、傅里叶变换等传统的变换域方法在近年来也取得了新的发展. Peter 等<sup>[24-25]</sup>使用最大似然小波密度估计 (Maximum likelihood wavelet density estimation) 来进行图像和形状的匹配. 该方法能够精确地估计一个非负的密度,估计的过程被结合在最大后验框架之下,能够有效地用于图像识别和点集匹配. Bartolini 等<sup>[26]</sup>提出了一种基于傅里叶变换的描述子 WARP, 该描述子主要使用了傅里叶系数的相位 (Phase) 作为形状的描述信息,相比于单独使用傅里叶系数的振幅 (Amplitude) 信息,相位信息给出了对物体边缘的更加精确的描述. 该方法是基于傅里叶变换的方法中比较好的方法.

### 2.1.5 其他方法

Rodrigues 等<sup>[27]</sup>使用解析签名 (Analytic signature) 的方法来表示形状,这种方法是一种重置不变 (Permutation-invariant) 的表示方法,待检索形状被映射成一个解析函数,这个解析函数就是解析签名. 当对一个形状的轮廓上的点进行重排列 (Re-ordering) 时,形状的解析签名表示保持不变,解析签名的方法能够有效地表示形状,对平移、旋转、尺度改变具有很好的鲁棒性.

Veltkamp 等<sup>[28]</sup>对各种形状描述方法的性能进行了详细的比较.

### 2.1.6 基于轮廓的形状表示方法综合评述

本节重点论述了基于轮廓的形状表示方法,这

类方法与传统的基于区域的形状描述子有一定的相关性. 传统的基于区域的描述子在近几十年来获得了长足的发展,包括矩、傅里叶描述子、小波描述子、形态学描述子等在形状表示领域都取得了不错的效果. 这些基于区域的方法包含着形状区域内所有像素的信息,所以这一类描述子适用于处理复杂形状,例如商标、汉字等. 但是这类方法的局限性在于它们仅仅包含了形状的全局信息,丢失了许多重要的形状细节. 而形状细节对形状表示起着至关重要的作用. 因此,这类方法在大型复杂数据库 (包含严重的同类变形) 上的性能十分有限.

以形状上下文为代表的基于空间位置关系的形状表示方法是本节论述的重点,形状上下文表示方法的提出可以被认为是形状表示方法研究的一个里程碑,是近十年来最为重要的方法,大量的新方法,新理论都建立在形状上下文表示的基础之上. 与传统方法相比,这类方法的缺点在于它们仅仅利用了物体的边界信息,无法反映物体内部的形状内容;同时,这些方法也很难处理包含有两个或更多分离部分的复杂形状. 这类方法的优势在于它们将形状轮廓的整体信息与局部信息有机地融合在一起,描述形状时更加准确,与一些匹配算法相结合,能够获得比较准确的形状局部上的对应关系,进而获得更为合理的形状相似性度量,大幅提升形状匹配与检索的性能. 同时这类方法的描述子设计简单有效,具有良好的几何不变性,并能自然地与其他很多方法结合在一起提升形状表示的能力. 该类方法还能有效地应用于自然图像的物体检测和物体识别,具有良好的实用性. 虽然基于轮廓的描述子基本上只能处理具有简单闭合轮廓的形状. 但因为其描述的精确性,在今后的形状表示研究中,该类方法依然是研究界的研究重点,会在研究领域占据主导地位.

本节论述的另一个重点是多尺度方法,这类方法的优势在于它们充分挖掘了形状在多个尺度上的特性,逐级揭示形状从局部细节到全局概貌上的信息. 这类方法的问题在于:首先,多尺度分析方法的复杂度往往较高,在多个尺度上的分析工作,大大加重了处理的工作量和负担;其次,多尺度难以定义. 无论是基于高斯平滑滤波的方法 (CSS, MCC 等),还是基于轮廓演化的方法 (DCE 等),都必须通过多个参数来控制尺度等级的选取. 而选取多少个尺度等级,每个尺度等级如何定义,其稳定解尚无定论.

总体上来看,将形状的整体特性与局部特性有效地结合来进行形状描述已成为基于轮廓的形状表示方法的主流思路,如何将局部与全局的几何或分布特性融合在一种描述方法中,如何在保证表示能力的前提下有效地降低描述子的特征维度将会是下一个阶段研究界关注的重点.

### 2.2 基于骨架的形状表示方法

#### 1) 常见方法

基于骨架的形状匹配方法首先需要计算待匹配形状的骨架, 据此生成 ARG (Attributed-relation graphs). 典型的 ARG 有奇点图 (Shock graph)<sup>[29-31]</sup>、Bone graph<sup>[32]</sup>、骨架树 (Skeleton tree)<sup>[33-37]</sup> 等. ARG 具有图结构, 通常用图匹配的方法来计算形状的相似程度. 而图匹配是一个 NPC 问题, 因此, 采用匹配 ARG 的方法匹配骨架往往只能得到近似解. 下面对几种有代表性的基于骨架的形状匹配方法进行介绍.

奇点图 (Shock graph) 中的奇点 (Shock point) 是按照烧草模型<sup>[2]</sup> 对骨架点的别称, 在烧草模型中火熄灭的点即为奇点 (骨架点). Siddiqi 等<sup>[31]</sup> 按照奇点及其邻接奇点最大圆的半径变化趋势将奇点分为 4 类, 如图 6 所示, 第 1 类奇点处最大圆的半径单调变化; 第 2 类奇点处最大圆的半径为局部最小值; 第 3 类奇点处最大圆的半径不变化; 第 4 类奇点处最大圆的半径为局部最大值. 可以发现第 1 类和第 3 类奇点不会单独出现, 往往会与 1 些同类奇点相邻; 而第 2 类和第 4 类奇点都是孤立的点.

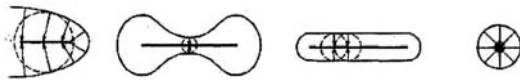


图 6 奇点的分类

Fig. 6 Classification of shock points

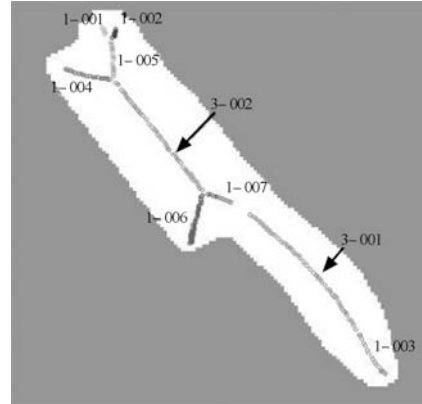
依据奇点的分类, Siddiqi 等<sup>[31]</sup> 还定义了奇点语法 (Shock grammar). 根据奇点语法可以把所有的奇点映射到奇点图 (Shock graph) 中. 如图 7 所示, 图 7(a) 是物体的奇点及其分类, 图 7(b) 是按照奇点语法组织的奇点图. 从图中可以发现, 同类相邻的节点在奇点图中为一个节点, 故而匹配奇点图只需建立图中节点的对应关系.

在两个待匹配的 Shock graph 中, 不同类节点间的距离为无穷大, 对于两个同类节点则比较以其为根节点的子图的差异以及该节点代表的骨架段的差异, 以此作为两个节点之间的距离.

匹配两个 Shock graph  $G_1, G_2$  时, 先在一个最小匹配中找到距离最近的两个节点  $P_1, P_2$ , 建立匹配关系, 将  $G_1$  分割为以  $P_1$  为根节点的子图  $g_1$ , 和不含  $g_1$  的子图  $h_1$ . 同样的将  $G_2$  分成两个子图  $g_2, h_2$ ; 对两个子图对  $(g_1, g_2), (h_1, h_2)$  用同样的方法匹配, 直到所有节点都找到匹配对象为止. 最后两个 Shock graph 的距离 (即非相似度) 为所有匹配的节点对的距离之和.

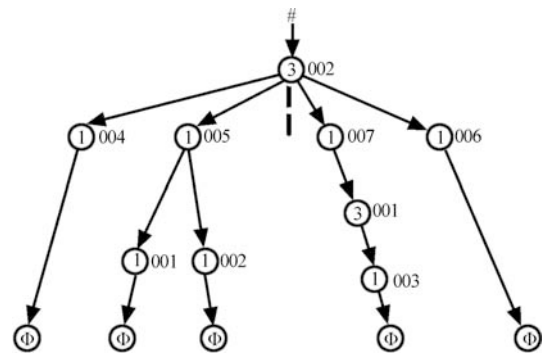
Sebastian 等<sup>[29]</sup> 采用计算编辑距离的方法来度

量两个奇点图的不相似程度. 其中定义了适用于奇点图的编辑操作 (Edit operation) 以及它们的代价. 这些编辑操作包括收缩 (Contract)、接合 (Splice) 和融合 (Merge). 编辑距离即为一个奇点图经过编辑操作变成另一个奇点图的最小代价之和. 显然, 距离近的奇点图有更高的相似度. 可以使用动态规划的方法得到具有最小代价和的编辑操作序列.



(a) 物体的奇点

(a) Shock points of an object



(b) 奇点图

(b) Shock Graph

图 7 物体的奇点及奇点图

Fig. 7 Shock points of an object and its shock graph

Shokoufandeh 等<sup>[30]</sup> 用 TSV (Topological signature vector) 表示 Shock graph 的拓扑结构. TSV 是一个向量, 其中的每一个元素是根节点下子树邻接矩阵特征值之和. 这种表示方法能够处理拓扑结构的微小变化.

在 Shock graph 框架下, Torsello 等<sup>[38]</sup> 用骨架长度变化率与对应边界长度变化率来表示骨架点的特征, 这种表示方法对形状部分弯曲是不变的.

Bone graph<sup>[32]</sup> 是 Shock graph<sup>[31]</sup> 方法的改进. Macrini 等<sup>[32]</sup> 对骨架分支分为两类: 连接段 (Ligature) 和非连接段. 其中, 连接段又分为全连接段 (Full-ligature) 和半连接段 (Semi-ligature). 在全

连接骨架点处, 最大内切圆与两个 (或两个以上) 边界凹点相切. 而半连接骨架点处的最大内切圆仅与一个边界凹点相切. 连接段和非连接段通过计算边轴率 (Boundary-to-axis ratio) 判断. 边轴率指的是与骨架分支相关联的边界长度与该骨架分支长度的比值. 连接段的骨架分支的边轴率至少有一个小于 1. 对于骨架的连接点, 该方法中分为端到边连接 (End-to-side, ES) 和端到端的连接 (End-to-end, EE). 端到边的连接点处有 2~3 个骨架分支连接在上面, 其中的两个分支表示了物体的同一部分, 这两个骨架分支被称为 Broken, 而另一个连接的骨架分支被称为 Attached. 端到边的连接有两种: 复杂的和简单的. 简单的端到边的连接有一个连接段是由一个或两个边界凹点导致的, 而复杂的端到边的连接段由三个或三个以上的边界凹点形成的. 判断一个连接点是 EE 还是 ES 的方法是比较连接点处最大圆切点间的距离和最大圆的半径, 小于则是 ES, 否则是 EE. 对于多个连接段连接到同一个连接点的情况, 比较每个连接段的内切圆半径, 把最大的一对标记为 Broken. 该方法中提到也可用其他方法标定 Broken, 但没有指明具体的方法.

对于连接在一起连接段, 断开标记为 Broken 的连接段. 这样相当于去掉了物体的一部分, 对缺失的边界用平滑的曲线填补. 重复这样的操作, 直到骨架中只剩下简单的连接关系.

构造 Bone graph 的方法是, 每一段剩下的骨架分支都是 Bone graph 的一个节点, 如果两个骨架通过一个连接点相连接, 则在这两个节点间有一个有向边, 边的方向由两个骨架分支的最大圆的半径大小确定. 也允许一些边是无方向的. Bone graph 匹配的方法与匹配 Shock graph<sup>[31]</sup> 的方法相同.

Ruberto<sup>[39]</sup> 的方法使用了 ASG (Attributed skeletal graphs) 来表示物体, ASG 由物体骨架转化而来, 是一个三元组  $G(V, E, R)$ ,  $V$  是骨架中分支点和端点组成的集合,  $E$  是骨架分支的集合,  $R$  是距离变换. 此外,  $E$  中的每条边  $e_i$  与向量  $(m_i, \theta_i, \lambda_i, s_i, \delta_i, \rho_i)$  关联起来, 该向量表示了骨架分支所代表的物体某一部分的特征, 其中,  $e_i$  代表的是骨架分支的形状上的特性 (具体的计算方法见文献 [39]);  $m_i$  表示骨架分支曲率的变化;  $\theta_i$  表示骨架分支起始点到物体起始点的距离;  $\lambda_i$  是骨架分支所占有的区域面积与物体面积的比值;  $s_i$  是骨架分支曲线长度与其端点间距离的比值;  $\delta_i$  是骨架分支上距离变换值的变化;  $\rho_i$  是骨架分支所在区域面积与物体面积的比值. 此时, 两个不同 ASG 中两条边的距离就是上述与之对应的向量的差异.

ASG 的匹配问题仍然化为一个最小匹配问题, Ruberto<sup>[39]</sup> 使用渐近分配 (Graduated assigned) 的

方法确定两个 ASG 之间的匹配矩阵. 通常, 匹配矩阵  $M_{m \times n}$  是一个 0-1 矩阵,  $m_{i,j} = 1$  表示一个 ASG 中第  $i$  条边和另一个 ASG 中的第  $j$  条边之间建立了匹配关系. 渐近分配把匹配矩阵中的元素化为连续的变量, 以此防止问题陷入局部最小. 经过多次调整匹配矩阵的值使问题得到满意的结果, 最后匹配矩阵还需还原为 0-1 矩阵. 渐近分配算法完成后往往不能得到一个 0-1 矩阵, 需要用启发式的方法得到最终的匹配矩阵, 该方法寻找矩阵中的最大元素, 将该元素置 1, 同一行, 列的其他元素置 0; 重复这个过程直到所有的元素都为 0 或 1. Ruberto<sup>[39]</sup> 为了避免噪声的影响, 对骨架进行了预处理: 删除较短的外部分支, 合并较短的内部分支.

Skeletal context<sup>[40]</sup> 方法首先依据骨架对轮廓进行采样, 轮廓到对应骨架点的距离越大采样的密度越小, 距离越小采样的密度越大. 接着, 利用骨架的端点对物体的轮廓进行分解, 文献 [40] 认为骨架的端点大都处于物体的轮廓上, 并以此为轮廓分解的依据. 事实上, 并不是所有的端点都会落在物体的轮廓上. 仿照 Shape context 定义了 Skeletal context, 即在对数极坐标系中, 计算骨架端点到所有轮廓采样点的坐标之差的直方图. 并利用与 Shape context 一样的方法计算两个骨架端点之间的距离. 同样, 利用 Hungarian 算法在两组骨架端点之间建立匹配关系. 一个骨架端点可以看作是物体的一段轮廓, 因此对骨架端点的匹配实际得到了不同轮廓段的对应关系. 接下来, 建立匹配的轮廓段内轮廓点之间的对应关系. 两个轮廓点之间的距离计算方法与两个骨架端点之间距离计算方法类似.

路径相似性 (Path similarity) 方法<sup>[41-42]</sup> 使用修剪后的骨架<sup>[43]</sup> 进行匹配, 修剪后的骨架的端点都位于物体的边界上. 该方法主要通过比较骨架端点间最短路径之间的几何相似性来匹配骨架图并用于形状识别. Path similarity 方法在两个物体骨架的端点间建立对应关系, 物体之间的不相似程度是所有对应的骨架端点之间距离的和. 计算骨架端点的距离需要计算任意两端点之间最短路径的距离, 以此构成距离矩阵, 利用 OSB 算法计算骨架端点之间的距离. 进行骨架端点匹配时, 计算两个骨架中任意两个端点之间的距离, 构成距离矩阵, 用 Hungarian 算法建立匹配. 与传统的基于树和图的匹配方法不同, 这里不考虑拓扑图的结构, 而是将所有骨架路径转换成特征向量序列的形式, 从而克服了骨架分叉点不稳定的问题.

由于图匹配具有较高的复杂度, 一些文献中把骨架简化为树形的 ARG 来进行比较与匹配. 这种树形的 ARG 称为骨架树. 在骨架树中, 根节点往往是靠近形状中心的骨架点或者具有最大内切圆



半径的骨架点. 而骨架树的叶子节点则是靠近形状边界的骨架点. 骨架树可以由连续完整的骨架得到<sup>[35-37]</sup>, 也可以由不连续的骨架得到<sup>[33-34]</sup>.

Liu 等<sup>[36-37]</sup> 根据连续的线型骨架得到骨架树, 骨架树的根节点是具有最大内切圆半径且靠近形状中心的骨架点, 骨架的连接点和端点按照连接关系构成骨架树中的节点. 在匹配骨架树时, 文献 [37] 用树描述符来表示骨架树的拓扑结构, 利用在树描述符中搜索最长相同字符串的方法搜索骨架树中的最大同构子树. 建立两个同构子树的匹配关系后, 在剩下的未匹配的骨架树中进行同样的操作, 直到所有的节点之间都建立了匹配关系. 两个骨架树之间的距离是匹配的骨架树节点之间距离的和.

Chen 等<sup>[35]</sup> 用树形结构表示骨架, 树的根节点为骨架中的脊柱基元, 骨架树的匹配归结为搜索最大同构子图的问题, 该方法采用整数规划的方法求其近似解.

一些方法匹配不连续的骨架, 例如, 中轴被分为正、负两类, 正的中轴终止于物体的突出部分, 而负的中轴终止于物体的凹进部分<sup>[33-34]</sup>. 选择通过物体中心的中轴的方向为极轴, 建立极坐标系. 以每一段中轴的终止点的极坐标表示该段中轴的位置. 以中轴段为单位进行匹配, 匹配的时候除考虑中轴段的极坐标位置, 归一化长度外, 还要考虑中轴段的类型 (正负), 不同类型的中轴段相似度为 0. 两个形状的相似度是它们匹配的中轴段的相似度的加权和. 在此基础上, 文献 [44] 把文献 [33-34] 中的形状表示看作是一个深度为 1 的树结构, 提出了用编辑距离匹配形状的方法.

Goh 等<sup>[45]</sup> 利用多分辨率的梯度向量获得骨架, 该方法首先利用一阶高斯滤波处理各个分辨率下的图像, 利用扩展运算, 求得各个分辨率下的梯度向量, 进行高斯滤波后得到局部离散图. 局部离散图中的局部极大值指示了骨架点的位置. 骨架被分为结构骨架和纹理骨架, 结构骨架具有缩放不变的特点, 纹理骨架大多由边界上的扰动造成的, 并且与结构骨架之间是不连通的. 该方法认为保证这种不连通性是有必要的, 特别是为了保证缩放不变性. 该方法首先在连接点按照极大-极小曲率准则或者最大分离度准则将骨架分解成基本骨架枝. 然后, 依据方向连续度和分离连续度把基本骨架合并. 把合并后的骨架段组合成任意的骨架分支, 称为二次骨架 (Second skeleton). 通过方向直方图差异, 宽度差异和骨架曲率差异表示两个骨架之间的距离. 进而, 通过计算空间距离和角度距离得到两个二次骨架之间的全局距离. 在匹配两个形状的骨架时, 计算它们所有可能的二次骨架之间的距离, 形成距离矩阵, 据此找到一个最小距离的匹配结果.

骨架还用于 3D 形状的识别, Cornea 等<sup>[46]</sup> 利用 3D 形状的线型骨架进行匹配. 该方法用 EMD (Earth mover's distance) 表示两个骨架部分的距离, 可以用于部分匹配. Sundar 等<sup>[47]</sup> 把 Shock Graph 匹配<sup>[31]</sup> 的方法应用于 3D 形状线型骨架的匹配.

## 2) 基于骨架的形状表示方法综合评述

骨架是形状的一种简洁表示. 相对于其他形状表示方法, 基于骨架的形状表示方法不但能很好地表示形状的几何特征, 还能表示形状的拓扑结构, 即形状各个部分之间的连接关系. 为了充分利用形状的拓扑结构信息, 基于骨架的形状表示方法通常会把骨架映射为一种图结构. 而基于骨架的形状表示方法的一个重要问题是图的匹配问题. 由于图匹配问题的复杂性, 基于骨架的形状表示方法往往采用一些启发式的方法或近似算法进行形状匹配. 本节重点介绍了几种骨架到图结构的映射方法及其匹配方法. 此类方法虽然能获得较高的匹配准确度, 但同时具有较高的计算复杂度, 因此现有的方法需要在保证匹配准确程度的基础上提高计算效率. 另外, 利用骨架对形状进行多尺度的描述, 能够在不同层次上对形状进行匹配, 有助于提高形状匹配的性能.

## 3 形状匹配方法

基于轮廓的匹配方法是在轮廓上找到一组点的最佳对应关系, 该问题常常被转化为最优化问题来进行求解. 一个优化问题包括一组候选解的集合, 以及一个目标函数, 解空间一般包含大量的候选解, 目标函数一般定义为对候选解质量的评估. 通常可以把优化问题看作是一个能量最小化的问题, 目标函数被称为能量函数. 目前很多的形状匹配的问题都归结为最小化一个能量函数的问题, 而形状的最佳匹配可以看作是在一个解空间中寻找一个最优解, 使得能量函数最小化. 基于骨架的匹配算法往往是通过权值图 (或树) 匹配, 由于第 2.2 节已经包括了此内容, 这里不再赘述. 基于轮廓的形状匹配方法的分类见图 8. 下面对近年来一些常见的形状匹配方法进行介绍.

### 1) 轮廓序列点集顺序关系未知的形状匹配方法

Chui 等<sup>[48]</sup> 提出了用 TPS-RPM 的方法对非刚性点进行匹配. TPS 算法是一种插值方法, 最初是用来产生一个平滑的函数映射, 它寻找一个通过所有的控制点的弯曲最小的光滑曲面, 弯曲最小由一个能量函数来确定. TPS 算法是用来获得两组点集之间的空间映射关系, 该方法的局限性在于要求已知点与点之间的对应关系, 这大大限制了 TPS 在序列匹配中的应用. TPS-RPM 扩展了 TPS, 使用迭代优化的方法来解决点的对应问题. 损失函数定义

为一个形状和另一个经过变换的形状上的点之间的欧氏距离的和. 该方法将对应问题和物体形变结合起来考虑, 对应关系问题是用 Softassign 算法来进行考虑, 物体形变问题仍使用 TPS 算法进行考虑, 从而将形状匹配问题转化成为一个非凸优化问题. 该问题的解是很难精确得到的, 一种有效的解法是使用确定性模拟退火算法迭代进行计算. 在每次迭代过程中, 首先固定对应关系, 求解物体的形变量, 然后固定形变量, 求解对应关系. 以此类推进行迭代, 直到求得全局最优解. TPS-RPM 算法在基于点集的形状匹配中是一种非常重要的方法, 很多其他的方法都是受它的启发而得到.

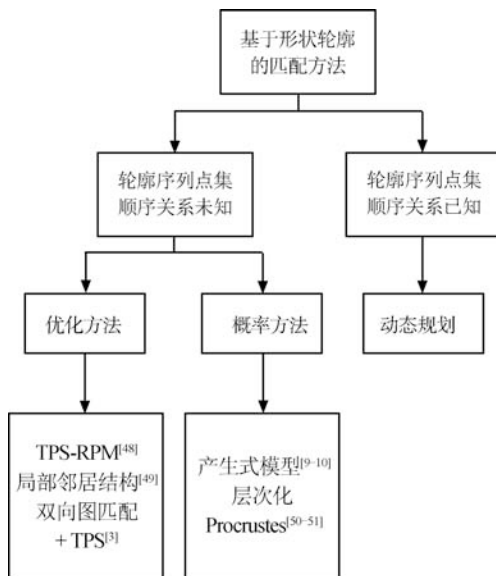


图 8 基于轮廓的形状匹配方法分类

Fig. 8 Classification of shape matching methods based on contour

Zheng 等<sup>[49]</sup> 使用邻居节点的局部关系来对非刚性物体进行匹配, 该方法将点匹配问题公式化为一个优化问题, 在匹配过程中存储局部的邻居结构. 该方法使用图结构来表示局部的邻居关系, 每个待匹配的点对应图中的一个顶点, 如果两个顶点是邻居节点, 那么在它们之间有一条边, 两个图之间的最优匹配是最大化匹配上的边的数量. 为了求得最优解, 该方法首先使用形状上下文距离来初始化图匹配, 然后使用松弛标记 (Relaxation labeling) 算法求解全局最优解. 该方法对噪声点、异常点、物体旋转和局部遮挡具有很好的鲁棒性.

Belongie 等<sup>[3]</sup> 在用形状上下文表示形状的基础之上, 也是用了 TPS 对形状进行匹配, 为了获得点与点之间的对应关系, 该方法使用双向图匹配 (Bipartite graph matching) 算法来计算形状上采样点的对应关系. 双向图匹配算法能够有效地找到两个

点集之间的对应关系. 该算法在最小化匹配损耗的时候不需要有附加约束. 在获得两个形状之间点的对应关系之后, 根据形状之间的对应关系, 使用 TPS 算法估计两个形状的非线性变换的校正, 最终的形状之间的距离 (不相似性) 由物体轮廓的对应关系以及校准形变的变形度量来共同决定. 该方法能够有效地应用于自然图像中的形状匹配.

Tu 等<sup>[9-10]</sup> 使用概率的方法来进行形状的匹配, 该方法受到 TPS-RPM 算法<sup>[48]</sup> 以及 DDMCMC 算法的启发, 首先使用允许形变的产生式模型将形状匹配的问题公式化为贝叶斯推理的问题, 然后使用 EM 算法的自由能量 (Free energy) 框架将产生式模型与上文介绍的信息特征<sup>[9-10]</sup> 相结合, 而计算出两个形状之间的相似程度. 信息特征可以认为是真实分布的一种近似, 类似于 DDMCMC, 信息特征的使用是一种数据驱动的思想. 在此基础上, McNeill 等<sup>[50]</sup> 将形状的局部轮廓点对应问题转化成为一个产生式模型, 以此来对轮廓点进行匹配. 之后 McNeill 等<sup>[51]</sup> 又将软对应 (Soft correspondence) 与估计形状形变的 Procrustes 分析方法统一到 EM 算法的框架之下对模型参数进行估计, 从而完成对形状的匹配.

## 2) 轮廓序列点集顺序关系已知的形状匹配方法

动态规划 (Dynamic programming) 算法是近年来形状匹配领域中最常用的一种算法, 该方法本质上是一种离散最优化的方法, 最早被用于时间序列的匹配. 动态规划算法的基本思想是将一个较复杂的问题转化为一系列同一类型的更容易求解的子问题, 由于是对每个子问题进行最优化求解, 就删去了所有的非最优组合方案, 使得计算量比穷举法大大的减少. 动态规划算法的另外一个优势在于单个子问题的解常常被多次使用来求解一些更大的子问题, 通过保存各个单个子问题的解, 能有效避免在求其他子问题时的重复计算, 这也在一定程度上提升了计算的效率. 用动态规划算法进行形状匹配类似于一个填表的过程, 每一个表格对应于要求解的一个子问题, 动态规划算法在整个表中迭代计算每个表格中的值, 当前表格值的计算依赖于已经计算得到的表格的值. 常见的用动态规划算法对形状进行匹配的方法一般是给出一个简单的迭代方程来定义当前表格和已计算得到的表格的值之间的关系. 如图 9 中的例子, 该例子以动态规划算法对 MCC 描述子进行匹配为例, 说明了动态规划算法的计算过程, 图中的黑线部分即为所求得两个轮廓序列的最优匹配结果. Felzenszwalb 等<sup>[52]</sup> 对动态规划算法在计算机视觉中的应用进行了详尽的综述.

与前述的用 TPS 等算法进行匹配相比, 动态规划算法需要将轮廓的顺序关系作为先验, 需要预先

定义轮廓的起始点和终止点, 因此在形状匹配的过程中拥有更多的有用信息, 这也是这类方法能大幅度提高形状检索精度的重要原因之一. 使用动态规划及其变形算法求解形状匹配问题的众多算法的不同之处在于定义形状之间的距离的方法不同, 在给出形状距离的定义之后, 都是使用动态规划算法求得全局的最优解. 以下对基于动态规划算法的形状匹配进行介绍.

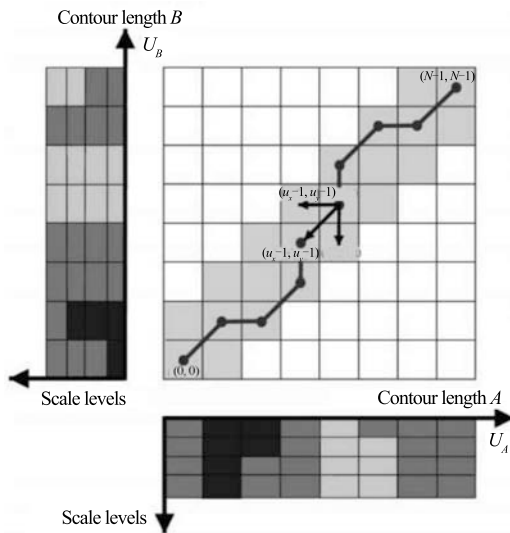


图9 使用动态规划算法对 MCC 形状描述子进行匹配的示意图

Fig. 9 Matching MCC by dynamic programming

Adamek 等<sup>[15]</sup> 定义两个轮廓点之间的距离为这两个点在不同尺度上的 MCC 描述子的平均距离, 根据这个距离定义使用动态规划算法对两个形状进行匹配. Alajlan 等<sup>[17]</sup> 的方法与此类似, 定义两个轮廓点之间的距离为这两个点在不同边长的三角形区域上的描述子距离的平均距离, 然后使用动态时间规整算法 (Dynamic time warping) 对两个形状上的轮廓序列点集进行匹配. Bartolini 等<sup>[26]</sup> 基于傅里叶描述子, 也是使用动态时间规整算法来对形状进行匹配. 这类算法的计算复杂度为  $O(N^2)$ . Alajlan 等<sup>[18]</sup> 基于曲率树还给出了一个匹配两个树结构的方法, 该方法迭代的寻找两个树之间的所有可能的匹配关系, 然后选择一个具有最大相似度的树作为最佳匹配结果. 该方法的计算复杂度为  $O(N)$ . Felzenszwalb 等<sup>[22]</sup> 基于形状树的表示, 定义两个形状之间的距离是最小化经过中点分割的两个子轮廓段的相似度之和再加上用来分割子轮廓段的中点的相似度, 对于中点的相似度赋予一个非均匀的权重, 越靠近树的底部的点有越大的权重, 因为处于底部的点反映形状细节, 不会影响整个形状的几何表达. 该距离表达也是通过动态规划算法来求解. 整

个算法的计算复杂度为  $O(nm^3)$ . Ling 等<sup>[5-6]</sup> 计算内距离形状上下文统计直方图距离作为度量形状不相似度的标准, 然后使用动态规划算法对轮廓点的集合序列进行匹配. McNeill 等<sup>[53]</sup> 使用层次化的 Procrustes 分析方法对形状进行匹配, 用动态规划算法来求得最优解.

Sebastian 等<sup>[54]</sup> 使用曲线编辑距离 (Curve edit distance) 来匹配开曲线或者是闭合曲线, 两条曲线上的点之间的对应关系由它们的空间位置关系 (Spatial location) 和曲率来控制, 然后定义一个匹配函数来最小化对应关系损耗, 为了找到最佳的对应关系, 该方法将每条曲线进行分段, 然后用动态规划算法求得最优解. 在该算法中, 分解后的曲线段是允许进行合并, 删除和附加操作的, 这样是为了有效地应对形状的变形. 该算法匹配一个闭合曲线的计算复杂度为  $O(N^2 \log N)$ . Daliri 等<sup>[4]</sup> 首先用动态规划算法对形状轮廓序列进行匹配, 以此为依据去除异常点, 然后通过 Procrustes 分析和符号串表示对形状进行二次描述, 定义两个符号串之间的距离为它们之间的编辑距离, 然后再次使用动态规划算法求得两个形状之间的最佳匹配关系. Attalla 等<sup>[21]</sup> 定义两个轮廓点之间的距离为这两个点的差分距离之和, 然后使用轮廓段的弹性比较来进行形状间的匹配.

Biswas 等<sup>[7]</sup> 给出了一种可扩展的快速形状检索方法, 该方法借鉴了特征包 (Bag-of-feature) 的思想, 将表示形状的特征向量 (由前文所述的基于轮廓的形状描述方法得到) 映射到一张哈希表当中, 哈希表中的每一格存储了某个形状  $S_k$  映射到该格中的特征向量的数量  $n_{ki}$ , 这样哈希表的每一格就包含了一组二元组  $\langle S_k, n_{ki} \rangle$ , 在进行形状检索的时候, 首先对待检索形状提取特征, 然后将得到的特征向量同样映射到哈希表当中, 通过解析比对待检索形状与数据库中的形状之间的哈希表分布, 可以得到待检索形状与数据库中的形状的非相似度度量 (距离). 该方法的检索效率非常高.

### 3) 形状匹配算法综合评述

轮廓点集的顺序关系对形状匹配的性能有着很大的影响, 这是近年来形状匹配领域最为重要的发现, 基于轮廓序列点集顺序关系的匹配算法也成为近年来最为常用的方法, 以动态规划为代表, 其他方法大多是该方法的变种. 与轮廓点集顺序关系未知的匹配算法相比, 这类方法利用了轮廓点集的顺序信息, 所以可以获得更准确的形状局部的对应关系, 进而改善形状的相似性度量结果. 其问题在于获得点集的顺序关系是这个方法的基本前提假设, 目前动态规划算法仅仅只能用于理想的二值图像的匹配, 而对于更有价值的点集匹配和自然图像中的形状匹

配等轮廓关系无法获得的问题, 这类方法则不适用. 所以这种匹配算法具有很强的局限性.

在下一阶段的形状匹配研究中, 如何在轮廓关系未知的情况下提升形状匹配的性能将会继续成为研究的热点和难点, 而如何更有效地利用轮廓顺序关系提高二值图像的匹配也是一个很有潜力的研究方向.

### 4 基于度量学习的形状检索进展

形状检索一直以来都是形状匹配的重要应用, 同时也是衡量不同形状描述方法匹配性能的重要标准. 大多数形状匹配的研究者都公布了自己在 MPEG-7 CE-Shape-1 Part B 以及其他一些重要形状数据库上的检索精度. 近年来, 新兴起了基于学习的形状检索方法, 研究者通过研究发现形状相似性度量形成的形状距离矩阵中存在着不满足三角形不等式的情况, 这就会引起一定的检索错误, 从而影响最终的检索精度. 如图 10 所示, 在图 10(a) 中, 左边的马和中间的狗的相似度要高于与右边的马的相似度. 在图 10(b) 中, 任意相邻的两匹马之间的相似度都很高, 但处在最左边的马和处在最右边的马的相似度很低, 因为它们之间的形态变化非常大. 基于学习的形状检索方法针对这一问题, 对每一个形状隐性的或显性的在形状相似性度量空间中找到一条最优的检索路径. 如图 10(b) 中通过中间的一些马的传递从最左边的马找到最右边的马便是一条检索路径. 在 2008 年的 ECCV 上, Yang 等<sup>[55-56]</sup> 首次提出这个问题, 并使用半监督学习中的标记传播算法 (Label propagation) 很好地解决了这个问题, 提出了图转导算法 (Graph transduction), 该方法将待检索形状作为标记数据, 数据库中的形状作为未标记数据, 使用标记传播算法传递待检索形状的标记信息, 从而隐性地找到一条传播路径, 该路径即为检索结果. 图转导算法成功将 MPEG-7 CE-Shape-1 Part B 形状数据库的检索正确率提高到了 91%, 这在当时引起了广泛的关注, 也开始了学者们对基于学习的形状检索的研究.

Yang 等<sup>[57]</sup> 使用局部约束传播过程 (Locally constrained diffusion process, LCDP) 的方法学习形状的传播路径, 并使用添加一些构造点 (Ghost point) 的方法使得形状相似性度量空间稠密化, 从而很好地提高了检索的精度. 接着, Kontschieder 等<sup>[58]</sup> 提出了相互  $k$ -最近邻图 (Modied mutual  $k$ -NN graph) 方法, 该方法将相似性度量空间中的点映射到图中, 每个点代表一个顶点, 当两个顶点中的每一个顶点的  $k$  个最近邻也是对方  $k$  个最近邻时, 这两个顶点就有一条边连接, 得到的图是一个稀疏表达的图, 对每个待检索形状, 在这个稀疏表达图

上使用求最短路径的 Dijkstra 算法找到一条最优路径, 该最优路径即为检索结果. 与学习内容敏感的相似性算法<sup>[59-60]</sup> 中使用的迭代算法相比, 该算法的计算效率非常高, 能够显性的获得检索路径, 检索精度也有所提高.

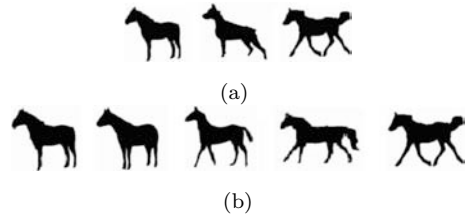


图 10 形状相似性度量空间中存在不满足三角形不等式的示例

Fig. 10 An example that triangle inequality is violated in shape similarity metric space

最近, Bai 等<sup>[61]</sup> 再次成功地将半监督学习中的协同学习 (Co-training) 方法引入到了形状检索领域, 与图转导算法相结合提出了协同转导 (Co-transduction) 的方法, 该方法使用两个相似度矩阵, 对每个矩阵使用图转导算法迭代求得最优路径, 经过一次计算之后, 将与待检索形状最相近的形状作为标记数据加入另外一个度量矩阵中, 以此类推. 协同转导算法有效的利用了检索过程中的置信度高的数据. 如图 11 所示, 图 11(a) 由于关节旋转变化造成的影响, 在查询形状  $A$  与形状  $B$  或  $C$  之间的 SC 距离较大. 然而, 在图 11(b) 中, IDSC 得到了理想的结果, 因为 IDSC 在关节旋转变化上比 SC 更稳定. 图 11(c) 中形状  $B$  和  $C$  之间的 SC 距离较小, 因为它们具有相同的姿态. 即便  $C$  比  $B$  更粗, SC 距离仍然在它们之间获得了好的匹配. 正因为 SC 和 IDSC 各具优势, 协同转导算法考虑将两者充分结合起来, 从而获得更好的检索结果. 具体来说, 先使用 IDSC 检索到  $B$ , 然后把  $A$  和  $B$  放在一起共同作为已标记的数据 (查询形状). 一个根据  $A$  和  $B$  训练而成的, 基于 SC 距离的新分类器会对  $C$  给出高的响应, 因为  $B$  和  $C$  之间的 SC 距离较小, 如图 11(d) 所示.

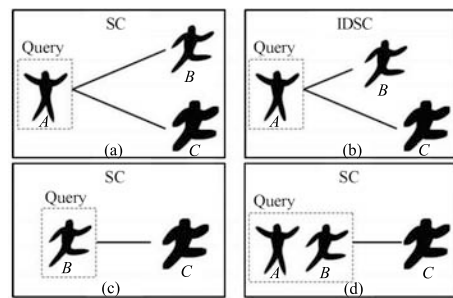


图 11 协同转导算法

Fig. 11 Co-transduction algorithm

该方法成功地将 MPEG-7 CE-Shape-1 Part B 的检索正确率提高到了 97.72%, 目前该方法是在 MPEG-7 CE-Shape-1 Part B 数据库上形状检索精度最高的方法.

## 5 常用数据集及评价方法

评价形状匹配性能的数据集非常多, 比较常用的数据集主要包括: MPEG-7, ETH80, Kimia, Rutgers Tools, Swedish Leaf, Tari 数据集.

MPEG-7 CE-Shape-1 数据库由 Latecki 等<sup>[62]</sup>发布. 该数据库包含三个部分, 第一部分是 CE-Shape-1 Part A, 主要用来衡量描述子对物体的尺度变化 (Scaling) 和旋转变化 (Rotation) 的鲁棒性. 第二部分是 CE-Shape-1 Part B, 主要用来衡量基于相似性的方法的检索精度 (Similarity-based retrieval), 这个部分是当前衡量形状描述子的重要参考指标之一, 部分 MPEG-7 CE-Shape-1 Part B 数据库的形状见图 12, 在该图中, 从每个形状类别中选出了两个代表形状. 该部分共有 1 400 个形状, 分为 70 类, 每类包含 20 个形状, 在进行检索时, 每个形状都作为待检索形状, 然后在整个数据库中找到最相似的 40 个形状, 如果这个 40 个形状中包含了和待检索形状同类的所有 20 个形状, 那么该类的检索正确率为 100, 统计每个待检索形状的最佳匹配形状中和该待检索形状是同类的形状的个数, 将每个形状的正确个数相加然后除以 28 000 (该值表示每个待检索形状能找到的同类形状之和的上限), 即得到了形状描述子在该数据库上的检索精度. 这种度量形状检索精度的方法叫做 Bull-eye. 第三部分是 CE-shape-1 Part C, 这个部分主要用来度量描述子对行为 (Motion) 以及非刚性变形 (Non-rigid deformation) 的鲁棒性. 表 1 给出了部分重要方法在 MPEG-7 CE-Shape-1 Part B 数据库上的检索精度.



图 12 MPEG-7 CE-Shape-1 Part B 数据库中的部分形状  
Fig. 12 Part of MPEG-7 CE-Shape-1 Part B database

表 1 部分重要方法在 MPEG-7 CE-Shape-1 Part B 上的检索精度

Table 1 Some important methods' precision on MPEG-7 CE-Shape-1 Part B

方法	检索率 (%)
CSS <sup>[16]</sup>	75.44
Visual part <sup>[20]</sup>	76.45
SC + TPS <sup>[3]</sup>	76.51
Curve edit distance <sup>[54]</sup>	78.16
Distance sets <sup>[8]</sup>	78.38
Prob.Approach <sup>[50]</sup>	79.19
Chance probability functions <sup>[63]</sup>	79.36
Skeletal context <sup>[40]</sup>	79.9
Generative model <sup>[9]</sup>	80.03
Optimized CSS <sup>[19]</sup>	81.12
Contour segmentation <sup>[21]</sup>	84.33
Multiscale representation <sup>[15]</sup>	84.93
Shape L'Ane Rouge <sup>[25]</sup>	85.25
Fixed correspondence <sup>[64]</sup>	85.40
IDSC + DP <sup>[6]</sup>	85.40
Symbolic representation <sup>[4]</sup>	85.92
Hierarchical Procrustes <sup>[53]</sup>	86.35
IDSC + EMD-L <sub>1</sub> <sup>[65]</sup>	86.56
Triangle area <sup>[18]</sup>	87.23
Shape tree <sup>[22]</sup>	87.70
Aspect shape context <sup>[66]</sup>	88.30
Layered graph <sup>[67]</sup>	88.75
Contour flexibility <sup>[68]</sup>	89.31
IDSC + LP <sup>[56]</sup>	91.61
IDSC + LCDP <sup>[57]</sup>	92.36
IDSC + Modied mutual $k$ -NN graph <sup>[58]</sup>	93.40
Aspect shape context + LP <sup>[66]</sup>	95.96
SC + IDSC + Co-transduction <sup>[61]</sup>	97.72

Leibe 等<sup>[69]</sup>给出了 ETH-80 数据库, 这个数据库包含有 8 类, 每类 10 个物体共 80 个物体, 每一个物体都有 41 个不同角度的图像, 整个数据库包含 3 280 张图片, 主要用来考察不同的基于 Appearance 的方法和基于轮廓的方法在物体识别上的性能. 评价方法是 Leave-one-object-out, 对于每一个待测试物体, 用数据库中的其他 79 个物体作为训练集来训练分类器对待测试物体进行分类, 当测试图片被赋予正确的类别标记时, 就认为是识别正确, 以此类推, 测试所有的图片, 最终的识别率由所有测试图片的正确率取平均得到. 图 13 给出了部分 ETH-80 数据库中的物体, 每个物体取了一个角度

的图片. 表 2 给出了部分方法在该数据库上的识别率.

表 2 部分方法在 ETH-80 数据库上的识别率

Table 2 Some methods' recognition rate on ETH-80

方法	识别率 (%)
SC + DP <sup>[6]</sup>	86.40
MDS + SC + DP <sup>[6]</sup>	86.80
IDSC + DP <sup>[6]</sup>	88.11

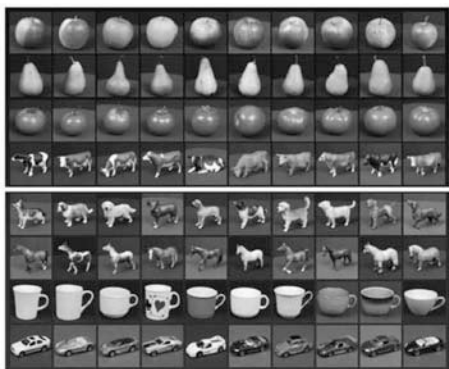


图 13 ETH-80 数据库的部分物体

Fig. 13 Part of ETH-80 database

Sebastian 等<sup>[70]</sup> 给出了 Kimia 数据库, 该数据库由三个数据集组成, 分别是 Kimia 25, Kimia 99, Kimia 216. Kimia 25 数据集由 6 类形状组成, 前 5 类每类有 4 个形状, 最后一类有 5 个形状, 共 25 个形状. 评价检索精度的方法是将数据库中所有的形状依次作为待检索图像, 计算其他形状到待检索形状的距离, 把距离最小的形状作为第一相似形状, 以此类推, 统计所有与待检索图像同类的第一相似, 第二相似和第三相似的形状数量. 图 14 给出了 Kimia 25 数据库中的形状, 表 3 给出了部分重要方法在该数据库上的检索精度. Kimia 99 数据库由 9 类组成, 每类有 11 个形状, 共 99 个形状. 评价检索精度的方法是将数据库中所有的形状依次作为待检索图像, 计算其他形状到待检索形状的距离, 把距离最小的形状作为第一相似形状, 以此类推, 统计所有与待检索图像同类的第一相似到第十相似的形状数量. 图 15 给出了 Kimia 99 数据库中的形状, 表 4 给出了部分重要方法在该数据库上的检索精度. Kimia 216 是由从 MPEG-7 CE-Shape-1 Part B 数据库中选出的部分形状组成, 共有 18 类形状, 每类有 12 个形状. 评价检索精度的方法是将数据库中所有的形状依次作为待检索图像, 计算其他形状到待检索形状的距离, 把距离最小的形状作为第一相似形状, 以此类推, 统计所有与待检索图像同类的第一相似到第十一相似的形状数量. 图 16 给出了

Kimia 216 数据库中的形状, 表 5 给出了部分重要方法在该数据库上的检索精度.



图 14 Kimia 25 数据集

Fig. 14 Kimia 25 database

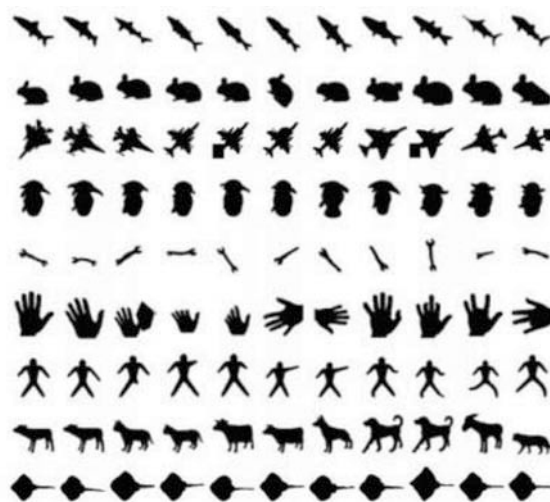


图 15 Kimia 99 数据库

Fig. 15 Kimia 99 database



图 16 Kimia 216 数据库

Fig. 16 Kimia 216 database

表 3 部分重要方法在 Kimia 25 数据库上的检索结果  
Table 3 Results of some important methods on Kimia 25

算法	Top 1	Top 2	Top 3
SC <sup>[3]</sup>	25/25	24/25	22/25
MDS + SC + DP <sup>[6]</sup>	23/25	20/25	19/25
IDSC + DP <sup>[6]</sup>	25/25	24/25	25/25
Efficient indexing <sup>[7]</sup>	25/25	25/25	23/25
Symbolic representation <sup>[71]</sup>	25/25	25/25	25/25

表 4 部分重要方法在 Kimia 99 数据库上的检索结果  
Table 4 Results of some important methods on Kimia 99

算法	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
SC <sup>[3]</sup>	97	91	88	85	84	77	75	66	56	37
Shock edit <sup>[31]</sup>	99	99	99	98	98	97	96	95	93	82
Generative model <sup>[9]</sup>	99	97	99	98	96	96	94	83	75	48
Efficient indexing <sup>[7]</sup>	99	97	98	96	97	97	96	91	83	75
Path similarity <sup>[41]</sup>	99	99	99	99	96	97	95	93	89	73
MDS + SC + DP <sup>[6]</sup>	99	98	98	98	97	99	97	96	97	85
IDSC + DP <sup>[6]</sup>	99	99	99	98	98	97	97	98	94	79
Triangle area <sup>[18]</sup>	99	99	99	98	98	97	98	95	93	80
Shape tree <sup>[22]</sup>	99	99	99	99	99	99	99	97	93	86
Symbolic representation <sup>[4]</sup>	99	99	99	98	99	98	98	95	96	94
IDSC + graph transduction <sup>[56]</sup>	99	99	99	99	99	99	99	99	97	99
IDSC + modied mutual kNN graph <sup>[58]</sup>	99	99	99	99	99	99	99	99	99	99

表 5 部分重要方法在 Kimia 216 数据库上的检索率  
Table 5 Results of some important methods on Kimia 216

算法	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th
SC <sup>[3]</sup>	214	209	205	197	191	178	161	144	131	101	78
Shock edit <sup>[31]</sup>	216	216	216	215	210	210	207	204	200	187	163
Path similarity <sup>[41]</sup>	216	216	215	216	213	210	210	207	205	191	177

Rutgers Tools 数据库共有 8 类 25 个形状, 如图 17 所示, 其评价方法与评价 Kimia 数据集的方法相同. 由于该数据集中的许多类形状肢体 (Articulation) 变化较大, 因此该数据集较适合于评价基于骨架的形状匹配方法的性能.

Söderkvist<sup>[72]</sup> 给出了 Swedish Leaf 数据库, 该数据库主要用来评价算法对轮廓的细小变化的鲁棒性, 图 18 给出了 Swedish Leaf 数据库的部分物体.

Baseski 等<sup>[44]</sup> 给出了 Tari 数据库, 该数据库主要包括两个子部分: 第一部分包括 180 个形状, 第二部分包括 1000 个形状, 该数据库主要用来测试算法

在非刚性物体上的性能, 常用来进行对骨架的匹配算法进行测试. 图 19 给出了 Tari180 数据库, 图 20

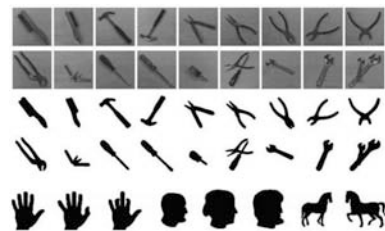


图 17 Rutgers Tools 数据库  
Fig. 17 Rutgers Tools database



图 18 Swedish Leaf 数据库的部分物体  
Fig. 18 Objects from Swedish Leaf database



图 19 Tari 180 数据库  
Fig. 19 Tari 180 database

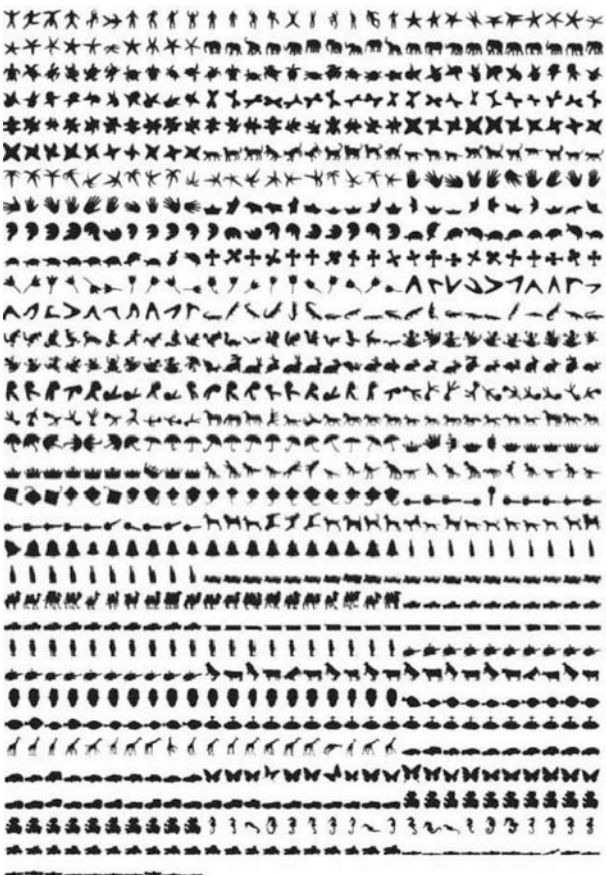


图 20 Tari 1000 数据库  
Fig. 20 Tari 1000 database

给出了 Tari 1000 数据库. 该数据库上的评价没有统一的标准, 以精度召回 (P-R) 曲线评价为主.

### 6 有潜力的研究方向

关于形状匹配的相关研究内容, 除以上回顾的这些方法外还有几个方面值得注意. 主要包括局部形状匹配、形状分类以及相似性度量空间的挖掘, 这

些方向目前已有初步的发展, 也出现了一些有代表性的方法, 但是还很不成熟, 值得研究界进行深入的研究. 下面就对这些研究方向及其已有的代表性方法进行介绍.

#### 6.1 局部形状匹配

与全局信息相比, 形状的局部 (Part) 信息在人眼对物体的识别中往往占据着重要的地位, 甚至只观察物体的部分信息就能够完成对物体的识别. 近年来, 局部形状匹配引起了广泛的关注, 是一个重要的研究方向. 局部形状匹配问题可以定义为给定一个开轮廓段, 在目标形状中找到最佳匹配的部分. 如图 21 所示, 图 21 (a) 给出了两个开轮廓段, 图 21 (b) 为目标形状, 其中的形状有一部分相似于图 21 (a) 中的开轮廓段, 图 21 (c) 中给出了经过局部形状匹配算法得到的目标部分形状. 下面对已有的一些局部形状匹配方法进行一个简单的介绍.

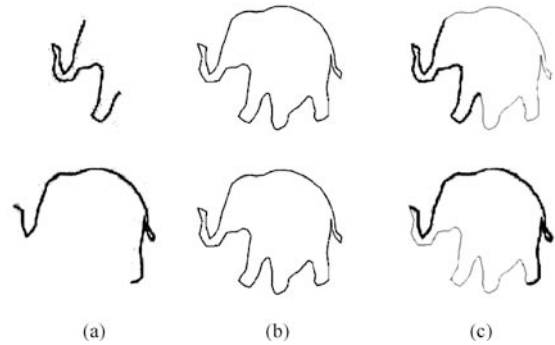


图 21 局部形状匹配  
Fig. 21 Part matching

Latecki 等<sup>[73]</sup> 对最佳局部相似性 (Optimal partial similarity) 进行了定义, 并给出了最佳局部形状相似性的计算方法. Lakaemper 等<sup>[74]</sup> 使用概率方法-粒子滤波器来求解形状的对应该问题, 并以此来解形状匹配的问题, 该方法使用局部的形状特征描述子来计算对应概率, 然后用附加约束来计算全局的对应结构. Chen 等<sup>[59]</sup> 使用史密斯-沃特曼算法 (Smith-Waterman algorithm) 来进行有效地局部形状匹配. Tanase 等<sup>[75]</sup> 使用相关反馈机制 (Relevance feedback mechanism) 来进行基于局部的形状检索. 基于局部形状匹配的物体识别也是物体识别中一类重要的方法, Saber 等<sup>[76]</sup> 使用子矩阵 (Sub-matrix) 匹配的方法来进行物体识别, Bai 等<sup>[77]</sup> 给出了一个用于在数字图像中检测和识别物体局部轮廓的系统.

局部形状检索是局部形状匹配的重要应用之一, 如图 22 所示, 给定一个目标物体 (锤子), 通过提取部分轮廓, 进行部分匹配, 检索出数据库中的同类



物体. 由图可知, 虽然待检索的图像只是一个开轮廓段, 但是依然能够检索出大部分正确的同类物体. 在自然图像中, 由于背景噪声的影响, 大部分情况下轮廓提取算法无法得到完整闭合的物体轮廓, 但是获得其中一个连续的轮廓段是可行的, 这就使得局部形状匹配显得尤为重要. 它为了解决自然图像中的物体检索问题提供了一条行之有效的途径. 同时, 对于自然图像中经常存在的局部遮挡问题, 局部形状匹配也能够自然地解决.

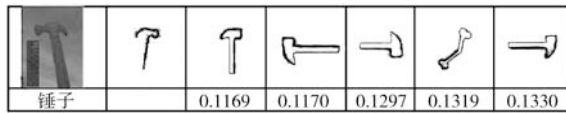


图 22 局部形状检索

Fig. 22 Part shape retrieval

基于局部形状匹配的物体识别自身而言还是一个病态问题. 如何定义物体的显著部分 (Visual part), 如何度量各个不同部分的视觉重要程度, 如何描述各个部分在空间的尺度及位置关系, 这些都是值得在今后的研究中继续深入探讨的问题. 但总体而言, 由于这种方法能够从本质上解决自然图像检索和识别中存在的很多问题, 其应用前景将非常广阔.

## 6.2 形状分类

形状分类也是形状分析领域中一个重要的问题, 主要思想是用数据库中的一部分形状作为训练样本, 提供一些先验信息, 然后通过一些数学模型 (产生式模型或判决式模型等) 对待分类图像的分类进行推理. 在 MPEG-7 CE-Shape-1 Part B 形状分类的评价标准是 Leave-one-out, 在每一类中当一个形状作为测试样本时, 该类中其他所有的形状均作为训练样本, 依次类推, 对每一个形状都进行测试, 最终的评价是所有分类正确的形状个数和数据库中形状总数的比值.

形状分类问题不同于传统的形状匹配问题, 它不再通过使用点集或图匹配算法来进行两两形状之间的匹配, 不是基于两两形状之间的非相似度 (距离) 来进行形状检索. 而是将所有的形状划分为两个集合: 训练样本集合和测试样本集合, 在训练样本集合中, 所有的形状都具有类别标记, 而测试样本集合中的形状都不包含类别标记. 形状分类问题的任务就是对测试样本集合中的形状赋予类别标记.

形状分类问题的基本框架见图 23, 形状分类问题的研究重点在于两个方面: 1) 如何表示图像以适用于形状分类的任务; 2) 如何设计有效的分类器来进行分类. 目前常见的形状分类的方法都是在这两个方面进行探讨, 并取得了一些阶段性的成果. 下面

对这一类方法进行一个简单的介绍.

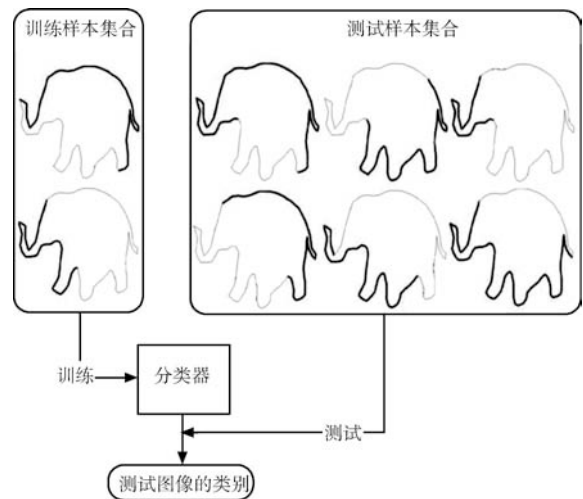


图 23 形状分类整体框架

Fig. 23 Shape classification framework

Sun 等<sup>[78]</sup> 使用产生式模型对形状进行分类, 该方法首先将形状表示成轮廓段 (Contour segments), 这些轮廓段相互之间是允许有重叠的, 然后使用贝叶斯模型将待分类图像与数据库中的图像相比较, 以最大化后验概率为标准找出待检索图像最有可能的类别. Bicego 等<sup>[79]</sup> 使用隐马尔科夫模型 (Hidden Markov models) 来对二维形状进行分类. Bai 等<sup>[37, 80]</sup> 将轮廓段 (Contour segments) 和骨架路径 (Skeleton path) 转化为向量特征的形式, 用高斯混合模型来判定输入形状的分类. Chen 等<sup>[81]</sup> 通过学习形状相似度对形状进行分类, 这与 Daliri 和 Torre 的方法<sup>[60, 82]</sup> 中的主要思想十分相似, 区别在于他们分别选择了不同的相似度计算方法和不同的分类器. Erdem 等<sup>[83]</sup> 采用一种二层的骨架联合树模型进行形状分类. 容易发现, 形状分类方法与形状匹配方法存在着共性, 例如都需要具有较好区分能力的形状特征; 而本质区别在于形状匹配的对象是一对形状, 形状分类只是给定输入形状的分类信息, 并不提供形状间的整体或局部对应关系.

形状分类的应用前景十分广阔, 在当训练形状只有一个的时候, 形状分类问题就简化为形状检索问题, 而在形状检索的应用中, 计算两两形状之间的非相似度 (距离) 耗费了大量的时间复杂度和空间复杂度, 由于形状分类使用分类算法替代了这一过程, 从而能大大地提高形状检索的效率. 同时, 形状分类是更一般性的问题, 形状检索只是它的一种特殊情况. 形状分类与自然图像的分类有着自然的联系, 通过对形状分类的研究能大大地促进自然图像的分类与检索的发展. 因此该方向极具研究潜力.

从目前来看, 基于局部轮廓段的表示是形状分

类中的有效表示方法, 如何获得更有效的表示方法, 以及如何结合有效的形状表示来设计有效的分类器将是下一个阶段的研究重点.

### 6.3 相似性度量空间挖掘

在第 4 节中介绍了基于度量学习的形状检索的最新进展, 这一类方法的提出为形状匹配、形状检索提供了新的研究思路, 相似性度量空间中的流形结构的发现具有决定性的意义, 在此思路下, 如何有效地挖掘相似性度量空间中的流形结构将会成为形状匹配和形状检索中的重要研究方向. 目前该研究方向还处于初步发展阶段, 并没有形成一套完整的框架, 常见方法主要参见一些具体算法的应用, 建立一套完整的理论框架将成为该方向下一步发展的重点.

相似性度量空间挖掘目前主要是关注在度量空间中的流形结构, 通过挖掘这些潜在的结构信息, 能够大大地提升形状检索的性能. 如图 24 所示, 这些数字形状都是未标示的形状, 可以认为是形状检索应用中的待检索形状, 在这些形状之间存在的特定的结构信息, 我们称之为上下文信息, 在形状检索的应用中, 除了初始给出的一个检索目标图像, 待检索形状之间的这种上下文信息也是非常关键的信息, 通过图传递或者最大间隔分离等方法, 这种上下文的信息是可以有效地被捕捉到的.

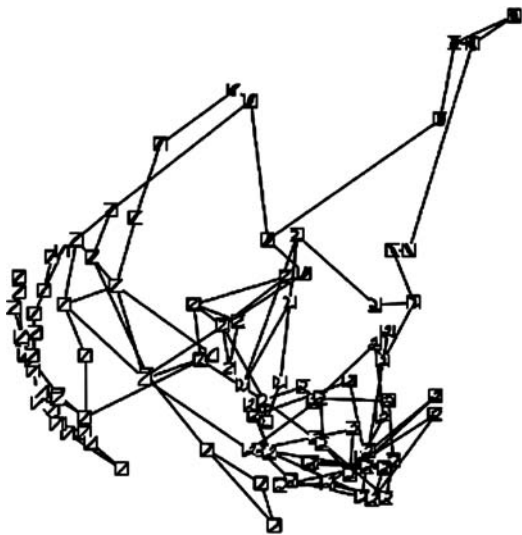


图 24 度量空间中的流形结构<sup>[84]</sup>

Fig. 24 Metric space structure of the manifold<sup>[84]</sup>

在传统的机器学习问题中, 监督学习、非监督学习已经获得了充足的发展, 但是这些方法在形状匹配、形状检索领域中的应用还近乎于空白, 我们相信这些传统的算法是能够有效地用于形状分析领域的. 近年来, 一些新兴的机器学习理论引起了广泛的注意, 半监督学习、集成学习、多示例学习、多标记学习、主动学习等全新的领域已经涌现出了很多优秀的方法, 目前大部分用于形状检索的算法都属于半

监督学习的框架, 如何将其他算法引入到形状分析领域也值得我们关注和研究.

另外, 现有的方法仅仅局限于在度量空间中挖掘潜在的流形结构, 从而提升形状检索的性能, 但是在匹配空间中进行结构挖掘也是处于空白阶段.

总之, 在下一个阶段的研究中, 传统的机器学习算法将会更多的应用于形状分析领域, 并会对自然图像分析等其他领域产生重要的影响, 对于这个发展趋势, 希望读者能够引起足够的重视.

### 6.4 综合评述

本节主要介绍了在形状匹配问题大框架下的一些有潜力的研究方向, 但是这些研究方向并不是孤立的, 它们与传统的形状匹配问题有着紧密的联系, 例如局部形状匹配是传统匹配算法的扩展, 形状分类是形状检索问题更一般化的形式, 相似性度量空间挖掘能大大地提升形状检索的性能. 同时这些算法之间也存在着紧密的联系. 目前形状分类的主要方法的描述子与局部形状匹配的描述子紧密相关, 局部描述子的提出将会促进形状分类的发展, 相似性度量空间的挖掘也会对形状分类和局部形状匹配产生重要的影响. 而相似性度量空间挖掘的研究思路也将形状匹配问题与传统的机器学习问题相联系, 从而打破了传统的形状匹配中仅仅使用几何或者概率方法进行求解的局限性, 为形状匹配的研究提供了全新的方向.

本节介绍的研究方向也会对其他领域有着重要的影响, 自然图像中的物体匹配、检索、识别、检测等方向将会获得全新的研究思路, 应用前景十分广阔.

## 7 总结

本文介绍了近年来形状表示, 形状匹配领域的最新进展, 并对有潜力的研究方向进行了介绍和分析, 由此, 我们提出了一个更为一般意义下的形状匹配及其应用的研究框架, 如图 25 所示.

在图 25 中, 白色的部分主要是传统形状匹配所囊括的内容, 而深灰色的部分则是在传统匹配的基础之上近年来的新进展以及有潜力的研究方向, 基于轮廓和基于骨架的形状表示将成为今后形状表示研究的主流方向, 而轮廓点集空间分布关系和多尺度表示是基于轮廓的形状表示的主要研究方向, 基于骨架的形状表示还将以图表示为主. 基于点集的形状匹配和图匹配将继续成为形状匹配的研究重点. 形状分类、局部形状识别、相似性度量空间挖掘都包含在这个大的框架之下, 成为了新的形状匹配框架下的重要组成部分, 并将成为下一个阶段的研究热点. 在图 25 中的浅灰色区域则给出了形状匹配大框架下的广泛应用.

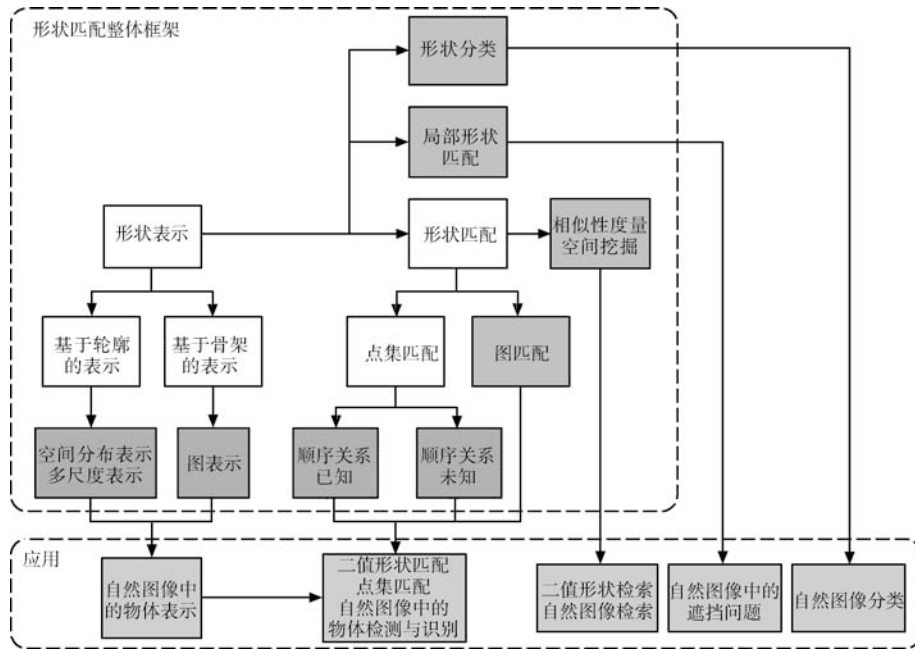


图 25 形状匹配及其应用整体框架

Fig. 25 Shape matching and its application framework

基于本节提出的全新的形状匹配及其应用的框架, 现总结目前存在的问题如下:

- 1) 基于轮廓序列空间位置关系分布的形状表示方法缺失了形状内部信息, 在应对复杂形状, 如商标数据等, 表示能力有限;
- 2) 基于轮廓多尺度特性的形状表示方法计算复杂度一般过高, 对于尺度的控制主要靠人工调整, 这就加大了这类方法的局限性;
- 3) 基于骨架的图匹配算法一般具有较高的计算复杂度;
- 4) 轮廓点集序列顺序关系未知的情况下匹配精度较低, 所用的优化算法或概率方法一般比较复杂, 轮廓点集序列顺序关系已知情况下的形状匹配虽然已有成熟的成果, 但是轮廓顺序关系已知的假设成为了其不能应用于自然图像的瓶颈;
- 5) 局部形状的表达还不成熟, 局部形状匹配的算法还有待发展;
- 6) 形状分类中的有效形状表示途径过少, 分类器的设计目前还较为单一;
- 7) 相似性度量空间挖掘还处于起步阶段, 目前还没有一个完整的理论框架, 大量的机器学习算法在形状分析中的可行性还有待讨论;
- 8) 由于形状匹配本身存在的缺陷, 其应用目前还十分的有限.

针对以上问题, 我们提出以下一些可能的解决思路:

- 1) 在轮廓表示的基础之上, 将传统的基于区域

的形状表示方法与基于轮廓的表示方法有效结合, 以提升算法对复杂形状的表达能力;

- 2) 在保持表示能力的基础之上设计低复杂度的形状描述子;
- 3) 考虑在轮廓点集顺序关系未知的情况之下如何提高形状的匹配精度;
- 4) 设计有效的局部形状表示方法和局部匹配算法;
- 5) 设计有效的表示方法来有效地结合分类器设计以提升形状分类的性能;
- 6) 对相似性度量空间挖掘进行理论分析, 建立完整的理论框架, 并加强现有机器学习算法在形状匹配中的研究.

### 8 结束语

形状匹配是计算机视觉中的一个重要问题, 最近十年来发展起了很多的新理论、新方法, 本文对这些新理论、新方法进行了总结, 按照基于轮廓和基于骨架的分类方法对形状表示和形状匹配进行了详细的介绍, 对基于度量学习的形状检索新进展进行了介绍和分析, 在此基础上对形状匹配及其应用的发展方向进行了预测, 并给出了一个全新的形状匹配的框架结构, 希望能对读者的研究工作有所帮助.

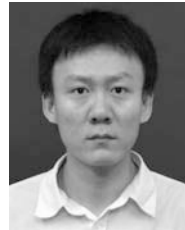
### References

- 1 Ding Xian-Feng, Wu Hong, Zhang Hong-Jiang, Ma Song-De. Review on shape matching. *Acta Automatica Sinica*, 2001, 27(5): 678-694

- (丁险峰, 吴洪, 张宏江, 马颂德. 形状匹配综述. 自动化学报, 2001, **27**(5): 678–694)
- 2 Blum H. Biological shape and visual science (Part I). *Journal of Theoretical Biology*, 1973, **38**(2): 205–287
  - 3 Belongie S, Malik J, Puzicha J. Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, **24**(4): 509–522
  - 4 Daliri M R, Torre V. Robust symbolic representation for shape recognition and retrieval. *Pattern Recognition*, 2008, **41**(5): 1782–1798
  - 5 Ling H B, Jacobs D W. Using the inner-distance for classification of articulated shapes. In: Proceedings of the 2005 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Washington, DC, USA: IEEE, 2005. 719–726
  - 6 Ling H B, Jacobs D W. Shape classification using the inner-distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, **29**(2): 286–299
  - 7 Biswas S, Aggarwal G, Chellappa R. Efficient indexing for articulation invariant shape matching and retrieval. In: Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Minneapolis, MN: IEEE, 2007. 1–8
  - 8 Grigorescu C, Petkov N. Distance sets for shape filters and shape recognition. *IEEE Transactions on Image Processing*, 2003, **12**(10): 1274–1286
  - 9 Tu Z W, Yuille A. Shape matching and recognition: using generative models and informative features. In: Proceedings of the 8th European Conference on Computer Vision (ECCV). Prague, Czech Republic: Springer, 2004. 195–209
  - 10 Tu Z W, Zheng S F, Yuille A. Shape matching and registration by data-driven EM. *Computer Vision and Image Understanding*, 2008, **109**(3): 290–304
  - 11 Yang M Q, Kidiyo K, Joseph R. Shape matching and object recognition using chord contexts. In: Proceedings of the International Conference Visualisation. London, UK: IEEE, 2008. 63–69
  - 12 Mori G, Belongie S, Malik J. Efficient shape matching using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, **27**(11): 1832–1837
  - 13 Thayananthan A, Stenger B, Torr P H S, Cipolla R. Shape context and chamfer matching in cluttered scenes. In: Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR). Washington, DC, USA: IEEE, 2003. 127–133
  - 14 Zhang H, Malik J. Learning a discriminative classifier using shape context distances. In: Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR). Washington, DC, USA: IEEE, 2003. 242–247
  - 15 Adamek T, O'Connor N E. A multiscale representation method for nonrigid shapes with a single closed contour. *IEEE Transactions on Circuits and Systems for Video Technology*, 2004, **14**(5): 742–753
  - 16 Mokhtarian F, Abbasi S, Kittler J. Efficient and robust retrieval by shape content through curvature scale space. In: Proceedings of the International Workshop on Image Databases and Multi-Media Search. Amsterdam, The Netherlands: IAPR, 1996. 35–42
  - 17 Alajlan N, El Rube I, Kamel M S, Freeman G. Shape retrieval using triangle-area representation and dynamic space warping. *Pattern Recognition*, 2007, **40**(7): 1911–1920
  - 18 Alajlan N, Kamel M S, Freeman G H. Geometry-based image retrieval in binary image databases. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008, **30**(6): 1003–1013
  - 19 Mokhtarian F, Bober M. *Curvature Scale Space Representation: Theory, Applications, and MPEG-7 Standardization*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 2003
  - 20 Latecki L J, Lakaemper R. Shape similarity measure based on correspondence of visual parts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, **22**(10): 1185–1190
  - 21 Attalla E, Siy P. Robust shape similarity retrieval based on contour segmentation polygonal multiresolution and elastic matching. *Pattern Recognition*, 2005, **38**(12): 2229–2241
  - 22 Felzenszwalb P F, Schwartz J D. Hierarchical matching of deformable shapes. In: Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Minneapolis, MN, USA: IEEE, 2007. 1–8
  - 23 Gorelick L, Galun M, Sharon E, Basri R, Brandt A. Shape representation and classification using the Poisson equation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, **28**(12): 1991–2005
  - 24 Peter A M, Rangarajan A. Maximum likelihood wavelet density estimation with applications to image and shape matching. *IEEE Transactions on Image Processing*, 2008, **17**(4): 458–468
  - 25 Peter A, Rangarajan A, Ho J. Shape  $L^{\hat{A}}_{\infty}$  rouge: sliding wavelets for indexing and retrieval. In: Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Anchorage, AK, USA: IEEE, 2008. 1–8
  - 26 Bartolini I, Ciaccia P, Patella M. Warp: accurate retrieval of shapes using phase of Fourier descriptors and time warping distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, **27**(1): 142–147
  - 27 Rodrigues J J, Aguiar P M Q, Xavier J M F. ANSIG — an analytic signature for permutation-invariant two-dimensional shape representation. In: Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Anchorage, AK, USA: IEEE, 2008. 1–8
  - 28 Veltkamp R C, Latecki L J. Properties and performance of shape similarity measures. In: Proceedings of the 10th International Conference on Data Science and Classification (IFCS). Ljubljana, Slovenia, 2006. 47–56
  - 29 Sebastian T B, Klein P N, Kimia B B. Recognition of shapes by editing their shock graphs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, **26**(5): 550–571
  - 30 Shokoufandeh A, Macrini D, Dickinson S, Siddiqi K, Zucker S W. Indexing hierarchical structures using graph spectra. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, **27**(7): 1125–1140
  - 31 Siddiqi K, Shokoufandeh A, Dickinson S J, Zucker S W. Shock graphs and shape matching. *International Journal of Computer Vision*, 1999, **35**(1): 13–32
  - 32 Macrini D, Siddiqi K, Dickinson S. From skeletons to bone graphs: medial abstraction for object recognition. In: Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Anchorage, AK, USA: IEEE, 2008. 1–8

- 33 Aslan C, Tari S. An axis-based representation for recognition. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV). Beijing, China: IEEE, 2005. 1339–1346
- 34 Aslan C, Erdem A, Erdem E, Tari S. Disconnected skeleton: shape at its absolute scale. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008, **30**(12): 2188–2203
- 35 Chen Xiao-Fei, Wang Run-Sheng. Multi-scale trees representation for object's skeletons. *Chinese Journals of Computers*, 2004, **27**(11): 1540–1545  
(陈晓飞, 王润生. 目标骨架的多尺度树表示. *计算机学报*, 2004, **27**(11): 1540–1545)
- 36 Liu Wen-Yu, Liu Jun-Tao. Objects similarity measure based on skeleton tree descriptor matching. *Journal Infrared Millimeter and Wave*, 2005, **24**(6): 432–436  
(刘文予, 刘俊涛. 基于骨架树描述符匹配的物体相似性度量方法. *红外与毫米波学报*, 2005, **24**(6): 432–436)
- 37 Liu J T, Liu W Y, Wu C H. Objects similarity measurement based on skeleton tree descriptor matching. In: Proceeding of 10th IEEE International Conference on Computer-Aided Design and Computer Graphics. Beijing, China: IEEE, 2007. 91–101
- 38 Torsello A, Hancock E R. A skeletal measure of 2D shape similarity. *Computer Vision and Image Understanding*, 2004, **95**(1): 1–29
- 39 Di Ruberto C. Recognition of shapes by attributed skeletal graphs. *Pattern Recognition*, 2004, **37**(1): 21–31
- 40 Xie J, Heng P A, Shah M. Shape matching and modeling using skeletal context. *Pattern Recognition*, 2008, **41**(5): 1756–1767
- 41 Bai X, Latecki L J. Path similarity skeleton graph matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008, **30**(7): 1282–1292
- 42 Bai X, Yang X W, Yu D G, Latecki L J. Skeleton-based shape classification using path similarity. *International Journal of Pattern Recognition and Artificial Intelligence*, 2008, **22**(4): 733–746
- 43 Bai X, Latecki L J. Discrete skeleton evolution. In: Proceedings of the 6th International Conference on Energy Minimization Methods in Computer Vision and Pattern Recognition (EMMCVPR). Berlin, Heidelberg: Springer-Verlag, 2007. 362–374
- 44 Baseski E, Erdem A, Tari S. Dissimilarity between two skeletal trees in a context. *Pattern Recognition*, 2009, **42**(3): 370–385
- 45 Goh W B. Strategies for shape matching using skeletons. *Computer Vision and Image Understanding*, 2008, **110**(3): 326–345
- 46 Cornea N D, Demirci M F, Silver D, Shokoufandeh A, Dickinson S J, Kantor P B. 3D object retrieval using many-to-many matching of curve skeletons. In: Proceedings of the 2005 International Conference on Shape Modeling and Applications (SMI). Washington, DC, USA: IEEE, 2005. 368–373
- 47 Sundar H, Silver D, Gagvani N, Dickinson S. Skeleton based shape matching and retrieval. In: Proceedings of the 2003 Shape Modeling International Conference. Washington, DC, USA: IEEE, 2003. 130–139
- 48 Chui H, Rangarajan A. A new algorithm for non-rigid point matching. In: Proceedings of the 2000 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Hilton Head Island, SC, USA: IEEE, 2000. 44–51
- 49 Zheng Y F, Doermann D S. Robust point matching for non-rigid shapes by preserving local neighborhood structures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, **28**(4): 643–6496
- 50 McNeill G, Vijayakumar S. 2D shape classification and retrieval. In: Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI). San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2005. 1483–1488
- 51 McNeill G, Vijayakumar S. Part-based probabilistic point matching using equivalence constraints. In: Neural Information Processing Systems (NIPS) Informatics Publications, 2006. 969–976
- 52 Felzenszeal P F, Zabih R. Dynamic programming and graph algorithms in computer vision. *IEEE Transactions Pattern Analysis and Machine Intelligence*, 2011, **33**(4): 721–740
- 53 McNeill G, Vijayakumar S. Hierarchical procrustes matching for shape retrieval. In: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR). Washington, DC, USA: IEEE, 2006. 885–894
- 54 Sebastian T B, Klein P N, Kimia B B. On aligning curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, **25**(1): 116–125
- 55 Yang X W, Bai X, Latecki L J, Tu Z W. Improving shape retrieval by learning graph transduction. *European Conference on Computer Vision (ECCV)*, 2008, **32**(5): 788–801
- 56 Bai X, Yang X W, Latecki L J, Liu W Y, Tu Z W. Learning context-sensitive shape similarity by graph transduction. *IEEE Transactions Pattern Analysis and Machine Intelligence*, 2010, **32**(5): 861–874
- 57 Yang X W, Koknar-Tezel S, Latecki L J. Locally constrained diffusion process on locally densified distance spaces with applications to shape retrieval. In: Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Miami, FL: IEEE, 2009. 537–564
- 58 Kotschieder P, Donoser M, Bischof H. Beyond pairwise shape similarity analysis. In: Proceedings of the 9th Asian Conference on Computer Vision (ACCV). Berlin, Heidelberg: Springer-Verlag, 2009. 655–666
- 59 Chen L B, Feris R S, Turk, M. Efficient partial shape matching using Smith-Waterman algorithm. In: Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Anchorage, AK, USA: IEEE, 2008. 1–6
- 60 Daliri M R, Torre V. Classification of silhouettes using contour fragments. *Computer Vision and Image Understanding*, 2009, **113**(9): 1017–1025
- 61 Bai X, Wang B, Wang X G, Liu W Y, Tu Z W. Co-transduction for shape retrieval. In: Proceedings of the 11th European Conference on Computer Vision (ECCV). Berlin, Heidelberg: Springer-Verlag, 2010. 328–341
- 62 Latecki L J, Lakaemper R, Eckhardt T. Shape descriptors for non-rigid shapes with a single closed contour. In: Proceedings of the 2000 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Hilton Head Island, SC, USA: IEEE, 2000. 424–429
- 63 Super B J. Learning chance probability functions for shape retrieval or classification. In: Proceedings of the 2004 IEEE Conference on Computer Vision and Pattern Recognition Workshop (CVPRW). Washington, DC, USA: IEEE, 2004. 93

- 64 Super B J. Retrieval from shape databases using chance probability functions and fixed correspondence. *International Journal of Pattern Recognition and Artificial Intelligence*, 2006, **20**(8): 1117–1138
- 65 Ling H B, Okada K. EMD- $L_1$ : an efficient and robust algorithm for comparing histogram-based descriptors. In: Proceedings of the 9th European Conference on Computer Vision (ECCV). Graz, Austria: Springer-Verlag, 2006. 330–343
- 66 Ling H B, Yang X W, Latecki L J. Balancing deformability and discriminability for shape matching. In: Proceedings of the 11th European Conference on Computer Vision (ECCV). Berlin, Heidelberg: Springer-Verlag, 2010. 411–424
- 67 Lin L, Zeng K, Liu X B, Zhu S C. Layered graph matching by composite cluster sampling with collaborative and competitive interactions. In: Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Miami, FL, USA: IEEE, 2009. 1351–1358
- 68 Xu C J, Liu J Z, Tang X O. 2D shape matching by contour flexibility. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, **31**(1): 180–186
- 69 Leibe B, Schiele B. Analyzing appearance and contour based methods for object categorization. In: Proceedings of the 2003 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Madison, Wisconsin, USA: IEEE, 2003. 409–415
- 70 Sebastian T B, Klein P N, Kimia B B. Recognition of shapes by editing their shock graphs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, **26**(5): 550–571
- 71 Demirci M F, Shokoufandeh A, Keselman Y, Bretzner L, Dickinson S. Object recognition as many-to-many feature matching. *International Journal of Computer Vision*, 2006, **69**(2): 203–222
- 72 Söderkvist O. Computer vision classification of leaves from Swedish trees [Master dissertation], Linköping University, Sweden, 2001
- 73 Latecki L J, Lakaemper R, Wolter D. Optimal partial shape similarity. *Image and Vision Computing*, 2005, **23**(2): 227–236
- 74 Lakaemper R, Sobel M. Correspondences between parts of shapes with particle filters. In: Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Anchorage, AK, USA: IEEE, 2008. 1–8
- 75 Tanase M, Veltkamp R C. Part-based shape retrieval with relevance feedback. In: Proceedings of the 2005 IEEE International Conference on Multimedia and Expo (ICME05). Amsterdam, Netherlands: ACM, 2005. 936–939
- 76 Saber E, Xu Y W, Murat Tekalp A. Partial shape recognition by sub-matrix matching for partial matching guided image labeling. *Pattern Recognition*, 2005, **38**(10): 1560–1573
- 77 Bai X, Yang X W, Latecki L J. Detection and recognition of contour parts based on shape similarity. *Pattern Recognition*, 2008, **41**(7): 2189–2199
- 78 Sun K, Super B J. Classification of contour shapes using class segment sets. In: Proceedings of the 2005 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Washington DC, USA: IEEE, 2005. 727–733
- 79 Bicego M, Murino V. Investigating hidden Markov models' capabilities in 2D shape classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, **26**(2): 281–286
- 80 Bai X, Yang X W, Yu D G, Latecki L J. Skeleton-based shape classification using path similarity. *International Journal of Pattern Recognition and Artificial Intelligence*, 2008, **22**(4): 733–746
- 81 Chen L B, McAuley J J, Feris R S, Turk M. Shape classification through structured learning of matching measures. In: Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Miami, FL, USA: IEEE, 2009. 365–372
- 82 Daliri M R, Torre V. Shape recognition based on Kernel-edit distance. *Computer Vision and Image Understanding*, 2010, **114**(10): 1097–1103
- 83 Erdem A, Tari S. A similarity-based approach for shape classification using Aslan skeletons. *Pattern Recognition Letters*, 2010, **31**(13): 2024–2032
- 84 Zhu X J. Semi-Supervised Learning with Graphs [Ph. D. dissertation], Carnegie Mellon University, USA, 2005



**周 瑜** 华中科技大学电子与信息工程系博士研究生。主要研究方向为形状匹配, 形状检索与分类, 视频跟踪以及半监督学习。

E-mail: zhouyu.hust@gmail.com

(**ZHOU Yu** Ph. D. candidate in the Department of Electronics and Information Engineering, Huazhong University of Science and Technology. His research interest covers matching, retrieval and classification visual tracking and semi-supervised learning.)



**刘俊涛** 军械工程学院计算机工程系讲师。现为华中科技大学电信系博士研究生。主要研究方向为计算机视觉, 计算机图形学和模式识别。本文通信作者。

E-mail: ropobox@21cn.com

(**LIU Jun-Tao** Lecturer in the Department of Computer Engineering, Mechanical Engineering Institute.

Now, he is a Ph. D. candidate in the Department of Electronics and Information Engineering, Huazhong University of Science and Technology. His research interest covers computer vision, computer graphics, and pattern recognition. Corresponding author of this paper.)



**白 翔** 华中科技大学电信系副教授。主要研究方向为计算机图形学, 计算机视觉, 模式识别。

E-mail: xiang.bai@gmail.com

(**BAI Xiang** Associate professor in the Department of Electronics and Information Engineering, Huazhong University of Science and Technology. His

research interest covers computer graphics, computer vision, and pattern recognition.)