

# 基于曲率分级的形状编码及识别方法

贾 棋 于美玉 樊 鑫 高新凯 郭 禾

(大连理工大学国际信息与软件学院 辽宁 大连 116621)

**摘 要** 形状识别是计算机识别领域中的基本问题,可以广泛地应用于对象识别、图像检索、图像配准、目标追踪等各个领域.现在的形状识别方法主要利用形状轮廓上采样点的相关性形成特征描述子,在实际应用中由于缩放、旋转、仿射、射影等变换,无法获取采样点之间的对应关系,形状匹配时间长,识别率低.为了克服基本的特征描述子的局限性,该文提出了一种基于曲率分级的形状编码方法.首先,将射影不变量引入到形状的基本表示中,以保证形状描述在各种变换下的稳定性;其次,以形状轮廓段为基本编码单位,对基本的描述子进行聚类编码;最后,为了使编码结果更好地代表形状轮廓信息,作者采用一种对轮廓段曲率分级的方式,将不同曲率级别的编码用 max\_pooling 的方式提取特征作为形状的最终编码.在通用数据库上的实验表明,该方法可以有效地识别在射影变换下的形状,识别率高达 98%,较基本的特征描述子提高了近 10%,与其它基于编码的方法相比也有一定的优势.

**关键词** 形状识别;射影变换;轮廓段;特征编码

**中图法分类号** TP391 **DOI号** 10.11897/SP.J.1016.2018.02453

## Shape Coding and Recognition Method Based on Curvature Classification

JIA Qi YU Mei-Yu FAN Xin GAO Xin-Kai GUO He

(International School of Information Science & Engineering, Dalian University of Technology, Dalian, Liaoning 116621)

**Abstract** Shape is an advanced visual feature of image because it can be used to represent the structure and attitude of the object. These characteristics are not available in object color and texture. Based on these advantages these advantages, shape is widely considered for object recognition. Shape recognition is a fundamental problem in compute vision community and is becoming more practical and getting more attention, it could be used in various areas, such as object recognition, image retrieval, target tracking and image registration. At present, Most of the existing approaches mainly focus on designing low level shape descriptors. Shape recognition is usually considered as a classification problem. Given a set of training shapes and category label of each training shape, shape recognition method needs to determine which category a testing shape belongs to. These traditional approaches have similar classification processes as most of them are based on matching shape descriptors of two different shapes. Thus shape descriptor is a key point in recognition task. For these approaches, they have similar processes to extract shape descriptors, First, the critical points of the shapes are extracted and then the relevance of contour points is computed, in general, the relevance would be regarded as features of these points. Finally, shape features are represented as the features of each point. After getting the shape descriptors, The correspondences between query shape and training shapes are crucial for matching algorithms. Matching strategies, such as Hungarian algorithm, dynamic programming algorithm, are widely used in feature matching.

收稿日期:2016-05-29;在线出版日期:2017-08-29. 本课题得到国家自然科学基金(61402077,61432003,61328206,11171052,61876030)、教育部新世纪优秀人才支持计划(NCET-11-0048)资助. 贾 棋,女,1983 年生,博士,副教授,中国计算机学会(CCF)会员,主要研究方向为数字图像处理、计算机视觉. E-mail: jiaqi@dlut.edu.cn. 于美玉,女,1992 年生,硕士研究生,主要研究方向为计算机视觉. 樊 鑫(通信作者),男,1977 年生,博士,教授,主要研究领域为计算机视觉、图像处理. E-mail: xin.fan@dlut.edu.cn. 高新凯,男,1993 年生,硕士研究生,主要研究方向为计算机视觉. 郭 禾,男,1955 年生,博士,教授,主要研究领域为并行与分布式计算、计算机视觉.

Training shapes based on the matching costs are classified, and the testing shape are matched by the nearest neighbor (NN) classifier. The methods of recognizing shapes without deformations such as scaling, rotation, affine, projective, can achieve high recognition accuracy. However, in our daily life, there are few shapes without deformations. Due to these deformations, the standard recognition process mentioned earlier cannot work as better as we expected. When two shapes are deformed, the critical points of two shapes are different from each other, thus it is difficult to compute the correspondences of the sampling points. Another important reason is that the low-level shape descriptors are point based and the sampling points do not have high level shape information. In order to overcome the shortcomings of the basic shape descriptors, we focus on mid-level shape descriptors and propose a new curvature classification based shape coding method. First, we introduce a new projective invariant and use it to describe the features of sampling points. The projective invariant makes the shape descriptor robust to various deformations. Then, basic shape descriptors are clustered and coded based on contour fragments. Finally, contour fragments are graded by curvature and the final codes are extracted by max-pooling strategy. The proposed method are tested on several public shape sets, and the results show that our method can recognize projective shape with accuracy up to 98%. Compared with basic shape descriptor, our method increases the accuracy about 10% with some unique advantages. The experiments results shows that our method is more powerful in recognizing shape with projective deformations.

**Keywords** shape recognition; projective transformation; contour fragment; feature coding

## 1 引 言

形状是描述对象、理解对象的重要线索,不仅可以描述目标物的拓扑结构,对颜色、纹理和光照的变化也具有一定稳定性.因此,形状在很多领域中得到了应用,比如目标识别、图像检索、目标追踪等.随着轮廓提取技术的发展,基于形状的认识问题变得越来越有实用价值,同时受到了业内外更多的关注.通常来说,形状识别可以被视作一种分类问题.给定一组带有类别的形状集合,然后通过分类算法对测试的形状进行分类.传统的形状识别算法通常基于两个不同形状的形状特征的匹配程度,这些传统形状识别算法有着相似的过程:提取形状特征,使用形状匹配算法如动态规划算法计算形状特征之间的匹配程度,然后对各个形状的匹配程度进行排序,最后根据排序的结果,使用最近邻算法完成形状识别分类.从形状识别的过程可以看出,形状特征是完成形状识别工作的基础,而后面的识别工作都需要依赖形状特征,所以形状特征的提取是目标识别中的关键步骤,优秀的描述子应该具有很强的区分性,因为一些形状具有局部相似性,差异度很小.而且在实际应用中,目标形状可能发生一系列几何变换,如平移、

旋转、射影变换等.因此如何提取形状在几何变换前后的不变特征也成为了形状识别领域的重要问题.

现有的形状描述子,多以轮廓采样点的位置或角度关系作为特征,其中最经典的是形状上下文(Shape Context, CS)<sup>[1]</sup>,通过统计采样点的分布信息作为描述子.后续一系列的方法在其基础上进行改进,如内部距离上下文(Inner Distance Shape Context, IDSC)<sup>[2]</sup>,近年来比较有代表性的是 Height Function<sup>[3]</sup>,利用采样点之间的距离作为描述子,计算简单,识别率也有一定程度的提高.这类方法可以解决基本形状的认识问题,但是无法使形状描述在实际应用中的仿射、射影等变换下保持稳定.针对这一问题,文献[4]将交比引入到形状的认识中,文献[5]利用二次曲线之间的系数关系作为射影不变特征描述子.我们在之前的工作中也通过引入新的仿射及射影不变量特征比<sup>[6]</sup>和基于特征数的层次化上下文<sup>[7]</sup>对形状进行描述,提高了形状在各种变换下识别的正确率.然而这类方法依然需要通过采样点的对齐来实现特征之间的比较和匹配,在实际应用中采样点之间严格的对应关系是无法获取的.因此,如何突破基本特征描述子的局限是形状识别中亟待解决的问题.

人类在观察 2D 形状时,很容易通过简单的线

条了解形状各部分之间的结构关系所代表的含义. 然而, 目前计算机视觉领域对形状特征表示处于用相应的特征值来表示基本元素位置关系的水平, 不利于形状的描述及识别. 认知学的研究表明高层次的特征更接近图像、目标物在人类大脑中的印象, 更利于识别和理解<sup>[8]</sup>. 在纹理信息的高层次特征表示方面, 最具代表性的就是“词袋”模型<sup>[9]</sup>. 受该思想的启发, 文献[10]和文献[11]也分别通过对角度信息和轮廓段进行编码实现对形状的高层次的理解和表示. 然而如何将视觉不变特征与高层表示进行融合, 以及如何选取更好的编码对形状进行表示仍然是形状表示的重点和难点.

本文以特征数为基础, 通过对特征编码与曲率分

级的结合形成形状描述子, 实现了形状射影不变特征的提取及形状的识别. 图 1 是本文方法的流程图, 我们首先用离散轮廓进化算法 (Discrete Contour Evolution, DCE)<sup>[12]</sup> 将形状分解成轮廓段, 然后使用层次化的方法获取轮廓段的射影不变的特征表示, 并进一步对每个轮廓段采用基于局部约束的线性编码 (Local-constraint Linear Coding, LLC)<sup>[13]</sup> 方式进行编码. 考虑到具有不同曲率程度的轮廓段往往具有不同的信息, 所以在计算轮廓段特征的同时, 还要计算轮廓段的曲率, 并且排序. 本文选择曲率位于所有轮廓段曲率前 60% 的轮廓段, 划分曲率级别, 在每个曲率级别内使用 max\_pooling 获得形状的特征描述. 在形状识别阶段, 采用线性支撑向量机达到识别的目的.

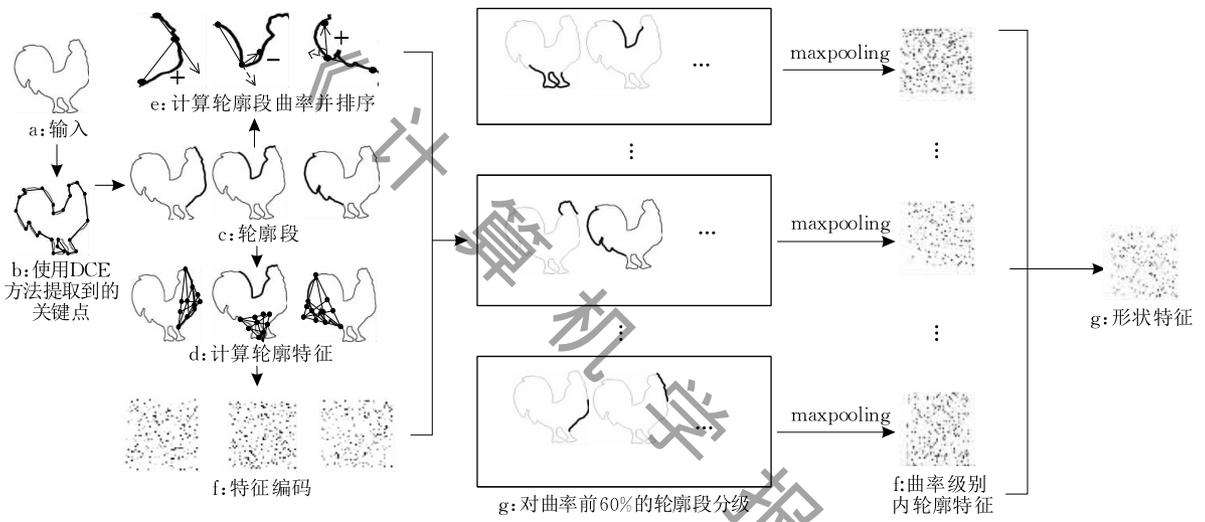


图 1 算法流程图

本文的方法主要有以下贡献:

- (1) 将射影不变特征描述子引入形状的高层次特征表示中.
- (2) 对基本描述子进行特征编码, 突破采样点的对应关系, 形成形状的更高层次描述.
- (3) 利用曲率分级使提取到的特征区分性更强.

本文第 2 节中介绍本领域的相关工作; 第 3 节以基本的特征描述子为基础进行编码及曲率分级, 形成对形状的最终表示; 第 4 节是实验结果及分析; 最后总结全文.

## 2 相关工作

对形状特征的研究由来已久, 从其构造的方法上来分基本可以分为基于度量信息的特征描述子、基于仿射、射影不变量的特征描述子及基于编码的

特征描述子.

### 2.1 基于度量信息的特征描述子

基于度量信息的特征描述子主要利用距离、角度等度量信息.

形状上下文<sup>[1]</sup> 是这类描述子中最经典的描述子之一. 该方法需要对轮廓均匀采样, 统计其他点在以该点为中心的极坐标系中的分布情况, 获得该点的统计直方图, 用直方图之间的差值表示两个形状的相似度. 这种描述子具有旋转、伸缩、平移不变性. 但是无法解决类内部形变问题而且对噪声鲁棒性差. 为了解决这些问题, 文献[2]中使用内部距离代替欧氏距离, 形成了内部距离形状上下文描述子, 这种方法对形状绞合关节的变性有很好的保持作用. 文献[14]使用多项式拟合提取特征点的方法增强了形状上下文描述子对于平移、旋转和尺度变换的鲁棒性. 文献[15]提出了一种基于多弦长函数的傅里

叶描述子,这种方法首先在轮廓上取点,然后将轮廓分为等弧长的段,轮廓函数由轮廓中提取的点到弧各个端点的距离表示.这种方法计算简单而且可以描述形状的总体和细节信息这种方法具有旋转、伸缩、平移不变性.文献[3]中介绍了高度函数描述子,一个点的特征由其他点到该点的距离表示,这种方法计算速度快,识别率也有一定的提升.文献[16]利用方向直方图提出方向上下文直方图(Histogram Of Orientation Shape Context, HOOSC),这种描述子可以在图像反转的情况下保持不变.文献[17]在形状上下文的基础上添加了符号串表示方法,并且为了提高形状匹配的精度,作者将该工作的重心放在形状元素的对齐即对应关系的确定方面上.该方法首先通过动态规划对采样点进行预匹配,去除无法匹配的异常点,然后使用 Procrustes 分析方法进行精确的匹配,确定形状之间的对应关系.最后通过符号串对轮廓点的编码与编辑距离结合的方式得到形状之间的匹配程度.这种方法不仅获得了两个形状元素之间的匹配关系也显著地提高了形状匹配的效果.文献[18]提出名为距离集的描述子,这种描述子利用形状轮廓点之间的距离描述形状信息.这种方法采用局部描述的方式,即使用某一样本点  $P$  与  $S$  个邻居样本点之间的距离作为样本点  $P$  的局部描述子.这种方法可以更好地描述形状局部的细节信息,但是无法在形状发生整体变形时保持不变.文献[19]提出的信息特征结合了形状局部特征和全局特征,并且使用角度信息作为度量信息,全局特征也将角度信息作为衡量的方法,考察目标轮廓点和其他轮廓点的夹角关系,也就是将距离约束转换为角度约束.

## 2.2 基于仿射、射影不变量的特征描述子

在实际应用中,由于不同拍摄角度的影响,形状均存在不同程度的仿射及射影变换,如何使形状特征描述子在各种情况下保持稳定,成为了亟待解决的问题,所以人们又提出了一些具有仿射、射影不变性的描述子.交比是一种基本的射影不变量,它具有构造简单,计算速度快等优点,它的透视不变性已经广泛的应用在经过各种透视变换后图像的认识工作中.二进小波<sup>[20]</sup>、余弦转换<sup>[21]</sup>是比较典型的仿射不变描述子.文献[22]给等仿射不变量增加了相似度不变性并且提出了仿射不变半度量.早在1998年,具有射影不变性的交比就被用于构造平面形状的描述子<sup>[23]</sup>,文献[4]用共线4点的交比作为描述子,一些其他射影不变量如共面圆锥曲线对<sup>[24-25]</sup>也被用

于构造形状描述子.我们在之前的工作中以特征比<sup>[6]</sup>作为描述子,实现了对内部有复杂结构形状的识别;文献[7]中,我们构造基于特征数的层次上下文描述子,继承了特征数射影不变性的优点,同时避免了特征数无法描述形状的情况.

## 2.3 基于特征编码的形状描述子

基于特征编码的形状描述子是一种新兴的描述子,它使用基本的描述子进行特征描述,然后进行特征的聚类和编码,最后用编码后的特征进行形状匹配.文献[11]等提出了基于轮廓段的编码方法,首先将形状分为多个轮廓,然后使用形状上下文描述子描述轮廓段,再使用空间金字塔和 max\_pooling 对特征进行选择,形成形状的描述子.最后使用分类器进行分类.文献[10]介绍了另一种编码方式:特征角编码.首先提取轮廓点,然后计算每个点与其相邻点连线间的夹角,为了解决尺度问题,又计算该点与两侧等间距点连线的夹角,然后给每个尺度的夹角一个权重,再进行编码表示.这类方法的优势是不需要采样点之间的对应关系,匹配速度快.但是如何选择合适的编码表示形状仍然是需要进一步研究的内容.

## 3 基于曲率分级的形状编码方法

为了用更高层次的信息表示形状,本文采用聚类编码的方式对底层的特征表示进行进一步处理,从而避免因采样点对齐而带来的精度损失和匹配时间消耗.总体来说本文方法分为以下5个部分:轮廓段提取、轮廓段特征描述、特征编码、曲率分级及编码表示,基于SVM的形状分类.接下来对每个部分进行详细介绍.

### 3.1 轮廓段提取

轮廓段的提取有多种方法,比如在文献[26]中使用的密集采样和基于曲率的采样,但是两种方式往往无法提取形状轮廓上有意义的轮廓段.因此采用离散轮廓进化(Discrete Contour Evolution, DCE)的方法<sup>[12]</sup>,它可以将形状的完整轮廓分解成多个有意义的轮廓段. DCE方法首先以形状轮廓为基础构造一个多边形,这个多边形是对原轮廓的近似.多边形的顶点集合为

$$s = \{s_1, s_2, \dots, s_v\} \quad (1)$$

$v$  为顶点的数目.顶点的数目是由输入参数  $\tau$  决定.对于一个轮廓  $S$ ,它的轮廓段集合为  $C(S)$ ,轮廓段  $c_{ij}$  为顶点  $s_i$  和  $s_j$  之间的轮廓.所以有

$$C(S) = \{c_{ij}\} \quad (2)$$

$s_i$  和  $s_j$  为顶点中的任意两点, 注意轮廓  $c_{ji}$  和  $c_{ij}$  是两个不同的轮廓段. 图 2 给出每个形状的不同轮廓段表示, 形状分段之后每一段可以基本代表形状的局部信息.

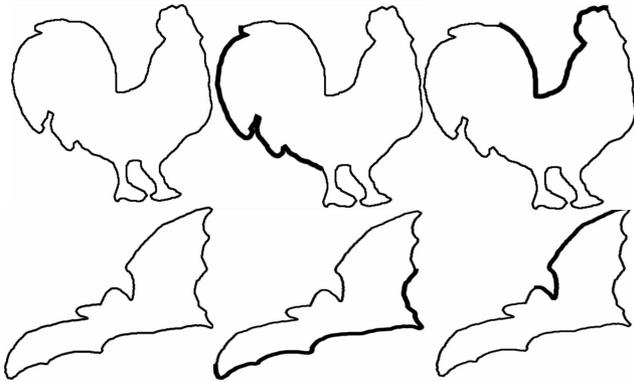


图 2 轮廓段举例

### 3.2 轮廓段特征描述

为了描述轮廓段的射影不变特征, 我们使用一种新的特征描述子——基于特征数的层次上下文描述子 (Hierarchical Characteristic Number Contexts, HCNC)<sup>[7]</sup>. 特征数是一种可以反映给定点固有几何关系的射影不变量<sup>[27]</sup>.

对于每一个轮廓段  $c_{ij}$ , 我们在轮廓段上均匀选取由 12 个采样点组成的集合  $P = \{P_1, P_2, \dots, P_{12}\}$ , 针对轮廓段上的任意一个样本点  $P_i$ , 我们可以从剩余样本点中任选 4 个点, 按照如图 3 所示的方法使用 5 个样本点构造并计算特征数; 为描述点  $P_i$  与轮廓段  $c_{ij}$  整体的关系, 我们使用文献[7]中的方法, 在轮廓段上由粗到精地选取一系列 5 点的组合, 得到以点  $P_i$  为基础的层次化特征向量即  $HCNC_i$ , 最后每个轮廓的特征向量  $x_{ij}$  可由 12 个样本点的特征向量连接成的一个向量表示:

$$x_{ij} = [HCNC_1, HCNC_2, \dots, HCNC_{12}] \quad (3)$$

$x_{ij} \in \mathbb{R}^{d \times 1}$ ,  $d$  是特征向量的维度.

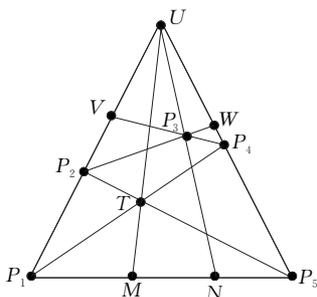


图 3 5 点特征数三角形构造法

### 3.3 特征编码

考虑到不同形状中可能存在部分点的位置关系一致, 仅仅以点的位置关系作为特征表示并没有体现形状之间的区别和联系, 因此我们需要对形状做中高层次的表示, 这里我们通过编码来实现. 获得每个轮廓段的特征之后, 首先进行码书的学习, 码书的学习方法有很多种, 其中包括有监督方法<sup>[28-29]</sup>和无监督方法<sup>[30]</sup>. 其中  $k$ -means<sup>[31]</sup> 方法既简单又稳定, 所以本文中选取  $k$ -means 作为码书的学习方法. 我们随机地选取每幅图像一定数量的特征, 然后用  $k$ -means 对选取的特征进行聚类, 获得聚类中心的集合即为码书.

$$B = [b_1, b_2, \dots, b_M] \in \mathbb{R}^{d \times M} \quad (4)$$

码书中的每一列代表一个聚类中心,  $M$  为聚类中心的数目.

根据学习得来的码书对轮廓段特征向量  $x_{ij}$  进行编码; 编码之后, 每个轮廓段特征由  $w_{ij}$  表示.

传统的编码方法 VQ<sup>[32]</sup> 只使用码书中与特征  $x_{ij}$  距离最近的码字来表示特征  $x_{ij}$ , 这种方法速度快但是可能会造成严重的量化错误, 特别是当特征  $x_{ij}$  位于一些码字边缘时. 为了避免这种错误, 我们使用局部限制线性编码 (Local-constraint Linear Coding, LLC)<sup>[13]</sup>, 这种方法不仅速度快, 而且十分有效. 为了表示特征  $x_{ij}$ , 首先找到码书中与  $x_{ij}$  距离最近的  $k$  个码字, 这  $k$  个码字形成了局部坐标系, 由  $B_{\pi_{ij}} \in \mathbb{R}^{d \times k}$  表示, 其中  $\pi_{ij} = \{\pi_{ij}^1, \dots, \pi_{ij}^k\}$  为码书  $B$  中对应特征向量的下标集合. 图 4 空心原点代表特征, 实心圆点代表码字, LLC 方法为每个特征选择离它最近的 5 个码字, 形成局部坐标系. 所以  $B_{\pi_{ij}}$  是一个由码书  $B$  中第  $\pi_{ij}^1, \dots, \pi_{ij}^k$  的列构成的矩阵. 特征向量  $x_{ij}$  可以由  $B_{\pi_{ij}}$  重构. 求解重构系数的过程即对式(5)求解, 并且所求解需要满足式(6).  $w_{\pi_{ij}}$  描述了  $x_{ij}$  与  $B_{\pi_{ij}}$  之间的局部几何关系即  $B_{\pi_{ij}}$  中各个向

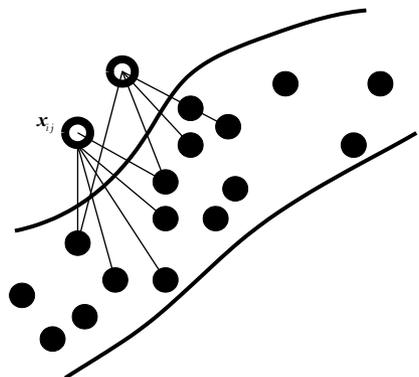


图 4 用 LLC 选取离  $x_{ij}$  最近的 5 个码字

量对  $\mathbf{x}_{ij}$  重构的贡献度.  $\mathbf{x}_{ij}$  编码后的表示为  $\mathbf{w}_{ij} \in \mathbb{R}^{d \times 1}$ , 其中  $\pi_{ij}^1, \dots, \pi_{ij}^k$  对应位置的值为  $\omega_{\pi_{ij}}$  中的值, 其他位置的值设置为 0.

$$\min_{\omega_{\pi_{ij}}} \|\mathbf{x}_{ij} - B_{\pi_{ij}} \omega_{\pi_{ij}}\|^2 \quad (5)$$

$$\mathbf{1}^T \omega_{\pi_{ij}} = 1 \quad (6)$$

### 3.4 曲率分级及编码表示

在一个形状中, 轮廓段存在很多相似性, 如果都用来表示形状就会造成数据的冗余, 同时一些明显的形状轮廓段也无法在编码过程中很好的代表该形状的特征. 因此, 本文通过对轮廓段进行曲率分级, 将不同曲率层次的轮廓段集中在一起, 并用统一的编码方式来代表. 同时考虑到不同形状曲率分布的不同, 我们采用了多种分级的方式, 目的是能够从多尺度多侧面地反映可能代表每个形状的不同曲率的段的编码表示.

曲率衡量了物体的弯曲程度, 是形状包含的基本描述量. Cartan 等人<sup>[33]</sup>证明了对于平面可微曲线, 曲线的曲率和曲率的一阶微分可以完全决定曲线. 而且曲率具有旋转和缩放、平移不变性. 由于曲率的重要作用, 本文使用曲率作为形状最后特征提取的基本准则.

下面给出获取形状最终表示的方法. 在获取每个形状的轮廓段集合  $C(S)$  后, 计算每个轮廓段的曲率, 计算轮廓段曲率的方法如图 5 所示, 图中的黑点表示轮廓段上的样本点, 轮廓段的曲率是图中所有角度值的和. 然后将所有轮廓段曲率从大到小进行排序. 选取曲率位于所有曲率前 60% 的轮廓段, 再使用不同的曲率间隔划分轮廓段. 本文方法中曲率间隔选取 10%, 12%, 15%, 20%, 30%, 60%. 根据不同的曲率间隔, 可以将选取的轮廓段划分为 6, 5, 4, 3, 2, 1 个子集合, 一共 21 个子集合. 然后在每个子集合  $SS_r$  ( $1 \leq r \leq 21$ , 且为整数) 中使用 pooling 获得每个子集合的特征. 轮廓段特征向量  $\mathbf{w}_{ij}$  中每个数字大小表示对应码字的响应, max-pooling 可以获得每个码字上最大的响应, 这样获取的特征更具有代表性<sup>[30-36]</sup>. max-pooling 后子集合特征的向量的大

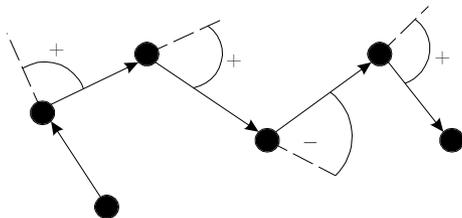


图 5 计算轮廓段的曲率

小与  $\mathbf{w}_{ij}$  相同. 每个子集合最后的特征表示为

$$f(S, r) = \max(\omega^s \mid z \in SS_r) \quad (7)$$

每个子集合特征向量连接成一个向量后即为一个形状的最终特征向量:

$$f(S) = [f(S, 1)^T, \dots, f(S, 21)^T]^T \quad (8)$$

不同曲率的轮廓段可以描述不同的形状特征, 有的轮廓段曲率大, 但是对于识别形状的作用很小, 而有的轮廓段曲率小, 但是对于识别一个形状很重要. 本文方法选取曲率前 60% 的轮廓段考虑了具有不同曲率的轮廓段. 这对于描述一个形状很有帮助. 而且在每个曲率间隔内使用 max-pooling 技术, 更加准确的提取了各个曲率间隔内形状的特征信息.

### 3.5 基于 SVM 的形状识别方法

在获得每个形状的最终表示向量后, 本文使用支撑向量机对各个形状进行分类. 因为形状有很多种, 所以需要多类别支撑向量机<sup>[37]</sup>. 给定用于训练的集合  $\{f_i\}$ , 集合中的每个形状都有自己的类别  $y_i \in [1, \dots, N]$ .  $N$  是形状类别数目. 文献[37]中的多类别支撑向量机可以用于解决如下优化问题:

$$\min_{\omega_1, \dots, \omega_N} \sum_{n=1}^N \|\omega_n\|^2 + \lambda \sum_i \max(0, 1 + \omega_i^T f_i - \omega_{y_i}^T f_i) \quad (9)$$

这个优化问题的左侧是正规项, 右侧是损失函数参数  $\lambda$  控制正规项之间的相对权重. 文献[38]中提出的 off-shelf 向量机解决器 LibLinear 可以解决上面的优化问题. 在测试阶段, 测试形状类别使用式(10)进行预测.

$$\hat{y} = \arg \max_{n \in [1, \dots, N]} \omega_n^T f \quad (10)$$

## 4 实验及分析

为了验证本文提出的形状表示方法的射影不变性及区分性, 我们在多个标准数据集上进行了测试. 根据实验结果, 我们又进行了对应的性能分析.

在本节中我们首先对实验的参数选择方法提供一定的验证, 其次通过和现有方法的比较验证本文方法的有效性, 然后通过在射影变换数据库上的实验证明本文方法在该方面的优势, 并进一步通过噪声实验证明本文方法通过编码后可以有效地抑制噪声的影响, 最后分析对比方法及本文方法形状提取部分的时间复杂度.

### 4.1 实验参数设置

本小节对影响分类准确率的参数进行了实验分析, 其中涉及到的参数有: (1) 构造特征数样本点数; (2) 轮廓段曲率百分比; (3) 计算轮廓段特征样本点

数; (4) 计算轮廓段曲率的间隔点数; (5) 码书大小。

为了选出可以平衡分类准确率及效率的各个参数值, 我们在通用数据集 MCD<sup>[39]</sup> 上使用两种测试方法: 一半训练一半测试和只留一个测试<sup>[11,26,40]</sup>。这两种测试方法被用于测试各个参数取不同值时, 本文方法的分类准确率, 然后根据实验结果选取合适的参数值。MCD 数据集包含从 MPEG-7<sup>[41]</sup> 数据集选择的 40 个形状目录, 每个目录包含的 14 个样本由同一个基本形状经过变换得来, 其中 7 个样本是使用数码相机从 7 个不同的视角拍摄得来, 另外 7 个样本为前 7 个样本的镜面反射。所以 MCD 一共有 560(40×14) 个形状。图 6 是 MCD 数据集一个目录的示例形状。一半训练一半测试, 即在每个目录下随机选取形状数目的一半用于训练, 另一半用于测试, 这类测试执行十次, 然后取十次的平均值; 只留一个测试, 采取经典的“留一法”, 即每类形状中只留一个测试, 其他的形状用于训练。

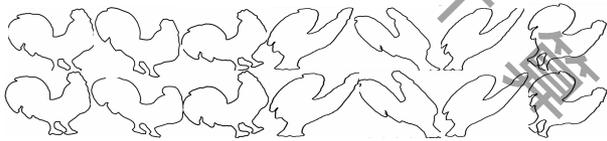


图 6 MCD 数据集示例形状

#### 4.1.1 构造特征数样本点数

由附录 1 可知, 构造射影不变量至少需要共面 5 点, 但是并没有具体样本点数, 所以我们在 MCD 数据集上测试构造特征数样本点数对形状识别性能的影响, 然后选取合适的样本点数构造特征数。文献<sup>[42]</sup>使用 6 点特征数作为射影不变量, 如图 7 所示, 点 A, I, C, H, B, J 被用于构造 6 点特征数, 具体的构造方法见文献<sup>[42]</sup>。所以我们使用 5 点和 6 点分别构造特征数, 然后计算层次上下文, 最后在数据集 MCD 上测试它们的性能。如图 8 所示, 5 点特征数和 6 点特征数之间的性能差很小, 几乎可以忽略。但是当使用 6 点特征数构造层次上下文时, 形状描述子的平均计算时间为 2.1 s; 使用 5 点特征数时, 时

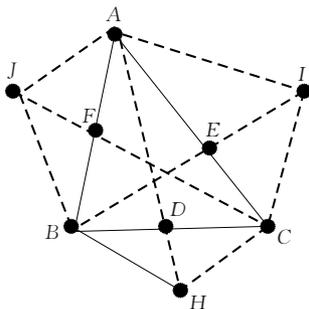


图 7 6 点特征数

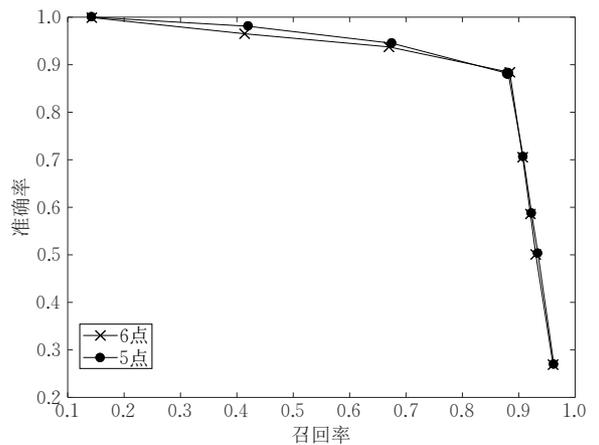


图 8 5 点特征数构造层次上下文和 6 点特征数构造层次上下文在 MCD 数据集上准确率-召回率曲线

间为 1.63 s。与 5 点特征数相比, 6 点特征数可以提取更多的几何信息, 但是层次上下文的构造方法即可充分的描述形状轮廓信息。所以本文选取 5 个样本点构造特征数。

#### 4.1.2 轮廓段曲率百分比

不同曲率百分比的轮廓段所包含特征的区分性不同。因此我们测试了不同曲率百分比的轮廓段对分类准确率的影响, 然后选出了最佳的曲率百分比。如图 9 所示, 轮廓段曲率百分比分别取 20%, 40%, 60%, 80%, 100%。从图中可以看出, 当曲率百分比从 20% 增加到 60% 时, 两类测试的分类准确率都显著提高, 一半训练一半测试提高了近 3%, 只留一个测试提高了 2%。曲率百分比从 60% 增加到 100% 时, 两类测试分类准确率都没有显著提高; 而且在 80% 时, 一半训练一半测试实验中还出现了准确率下降的情况。由于本文方法根据曲率对所有轮廓段进行排序, 曲率大的轮廓段包含的特征更具区分性, 随着曲率的降低, 轮廓段所包含特征的区分性也随之下降。所以当轮廓段百分比从 20% 增加到 60%

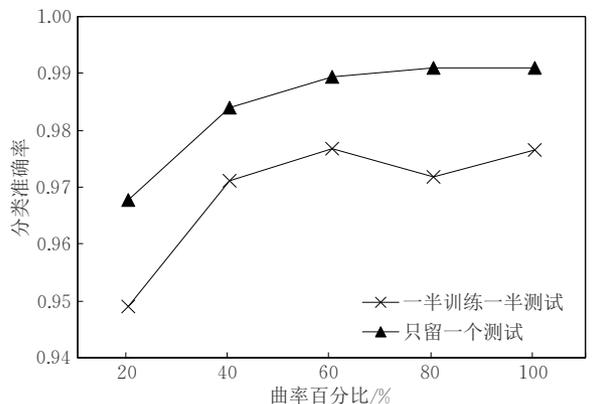


图 9 轮廓段曲率百分比对分类准确率的影响

时,分类准确率有显著提高,但是从 60% 增加到 100% 时,分类准确率没有显著提升. 所以我们选取曲率前 60% 的轮廓段计算形状特征.

#### 4.1.3 计算轮廓段特征样本点数

计算轮廓段特征样本点数越多,所描述的轮廓段信息越丰富,但是所需的计算时间也随之增加,为了平衡二者,我们测试了在不同样本点数的情况下,分类准确率的变化,然后根据实验结果选出合适的样本点数. 样本点数分别取 5、7、10、12、15、18. 如图 10 所示,当样本点数为 12 时,本文方法在两类测试中都可以获得较高的分类准确率. 当样本点数少于 12 点时,分类准确率均低于 12 点的分类准确率; 样本点数为 18 点时,在一半训练一半测试实验中,分类准确率为 97.78%,稍高于 12 点的 97.68%,但是在只留一个测试实验中,18 点的分类准确率为 98.39%,低于 12 点的 98.93%. 增加样本点数将会增加计算量而且不能显著提高分类准确率. 因此在计算轮廓段特征时,选取 12 个样本点.

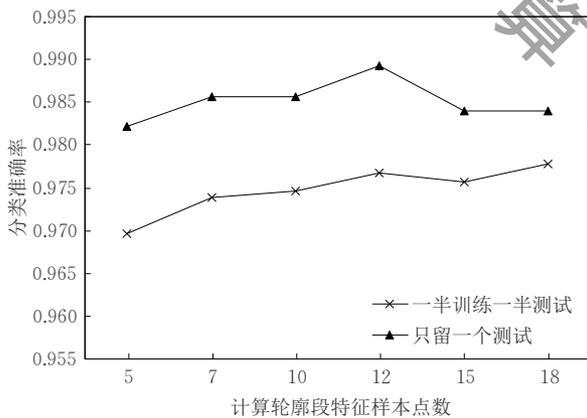


图 10 计算轮廓段特征样本点数对分类准确率影响

#### 4.1.4 计算轮廓段曲率的间隔点数

由 3.4 节可知,我们需要根据轮廓段的曲率对轮廓段进行分级,而计算轮廓段曲率的间隔点数影响着轮廓段的曲率,所以我们需要选择合适的间隔点数. 本小节测试了不同间隔点数对本文方法分类准确率的影响,然后根据实验结果,选择合适的间隔点数. 如图 11 所示,分别选取 3、5、7、9、11 点作为计算轮廓段曲率的间隔点数. 当取 5 点时,本文方法在两类测试中均可以获得较高的分类准确率,一半训练一半测试实验中的分类准确率为 97.68%,只留一个测试实验中的分类准确率为 98.93%. 取 7 点时,只留一个测试的分类准确率为 99.1%,稍高于高于 5 点的 98.93%,但是一半训练一半测试实验中,7 点的

分类准确率为 97.28%, 低于 5 点的分类准确率. 所以在计算轮廓段曲率时,我们选取 5 点作为间隔.

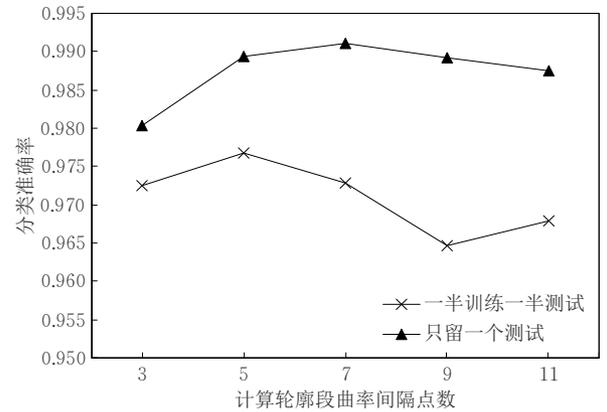


图 11 计算轮廓段曲率的间隔点数对分类准确率的影响

#### 4.1.5 码书大小

最后,我们在 MCD 数据集上测试了码书大小对分类准确率的影响. 如图 12 所示,码书大小为 1800 时左右时,本文方法在两类测试中均有很高的准确率,所以本文将码书的大小设置为 1800.

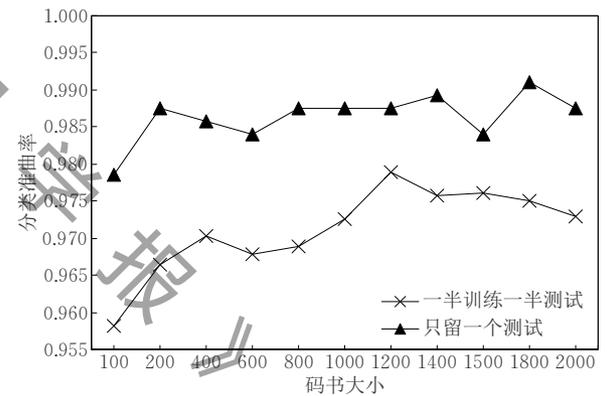


图 12 码书大小对分类准确率的影响

## 4.2 MPEG-7 数据集

MPEG-7 数据集在形状分析领域有着广泛的应用. 它有 70 个目录, 每个目录包含一类形状, 每类形状有 20 个不同的形状, 所以 MPEG-7 一共包含 1400 (70×20) 个形状. 图 13 是 MPEG-7 数据集中的一些形状.

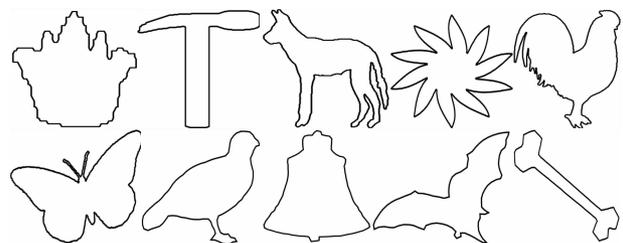


图 13 MPEG-7 数据集示例形状

由表 1 可以看出, 本文的识别率与其他方法相当, 略高于 Contour segments<sup>[40]</sup> 和 Skeleton paths<sup>[40]</sup>. 其中 BCF<sup>[11]</sup> 的识别率最高达 96.91%, 本文方法识别率略低的主要原因是本文采用几何不变关系对点的位置进行表示, 这种约束性较强, MPEG-7 数据库中每一类有很多变化比较大的形状, 它们之间的变化不满足射影不变性, 因此可能会对识别率略有影响.

表 1 MPEG-7 数据集上的分类准确率对比

方法	分类准确率/%	
	一半训练一半测试	只留一个测试
Class segment set <sup>[26]</sup>	90.90	97.93
Contour segments <sup>[40]</sup>	91.10	—
Skeleton paths <sup>[40]</sup>	86.70	—
ICS <sup>[40]</sup>	96.60	—
Polygonal multi-resolution <sup>[43]</sup>	—	97.57
String of symbols <sup>[44]</sup>	—	97.36
BCF <sup>[11]</sup>	96.91	98.28
本文方法	92.77	96.29

为了验证本文方法在各种变换下的稳定性, 我们进一步测试本方法在其他的数据集上的稳定性, 如 MCD、PLD<sup>[45]</sup>. 这两类数据集中由受到不同变换的形状组成, 所以常被用于测试方法在多种变换中的稳定性.

#### 4.3 MCD 数据集

由表 2 的结果可以看出, 在一半训练一半测试实验中, 我们的方法比 Class segment set 的准确率高 30%, 比 Contour segments 高 37%, 比 BCF 方法的分类准确率高 19%, 准确率高达 97%; 在每类形状只留一个测试的实验中, 我们的方法比 Class segment set 高近 25%, 比 Contour segments 高 30%, 比 BCF 高 12%, 准确率高达 98%. 为了证明本文方法对于各类射影变换形状的分类均具有较高的准确性, 我们测试了本文方法及其对比方法在 MCD 数据集各个类别的分类准确性. MCD 数据集共有 40 个类别, 数目较多, 所以本文中我们随机选取了 10 个类别作为对比. 图 14 为 10 个类别的示例形状. 如表 3 的实验结果所示, 本文方法在 9 个类别中均获得了 100% 的分类准确率, 只有在类别 12 中获得了 78.57% 的分类准确率, 但是仍然高于方法 Class segment set, Contour segments 和 BCF 在类别 12 中的分类准确率; 方法 BCF 只在 5 个类别中获得了 100% 的准确率, 在其他 5 个类别中均获得不超过 86% 的准确率, 其中类别 12 的准确率仅为 50%; 方法 Class segment set 仅在类别 22, 30, 25, 20 中获得了超过 90% 的准确率, 其他 6 个类别的准确率均低于 80%, 在类别 35 的分类准确率仅为 7.14%;

Contour Segments 在 4 个类别中的分类准确率高 于 85%, 其余 6 个类别的分类准确率均不高于 78.57%. 这说明我们的方法对于射影变换不变特征的提取好于 Class segment set, Contour segments, BCF. 图 15 分别给出了类别 12 和类别 35 中的 4 个形状示例. 我们方法对于这两个类别的识别准确率明显高于其他方法.

表 2 MCD 数据集上的分类准确率对比

方法	分类准确率/%	
	一半训练一半测试	只留一个测试
Class segment set	67.82	75.36
Contour segments	60.89	68.39
BCF	78.57	87.32
本文方法	97.68	98.93



图 14 MCD 数据集对比类别示例形状

表 3 MCD 数据集中部分类别分类准确率对比

类别号	方法/%			
	Class segment set	Contour segments	BCF	Our method
22	100.00	35.71	100.00	100.00
30	92.86	92.86	100.00	100.00
12	64.29	7.14	50.00	78.57
24	78.57	78.57	78.57	100.00
25	100.00	85.71	100.00	100.00
6	64.29	100.00	100.00	100.00
1	42.86	57.14	100.00	100.00
20	92.86	71.43	85.71	100.00
35	7.14	50.00	78.57	100.00
23	78.57	85.71	85.71	100.00

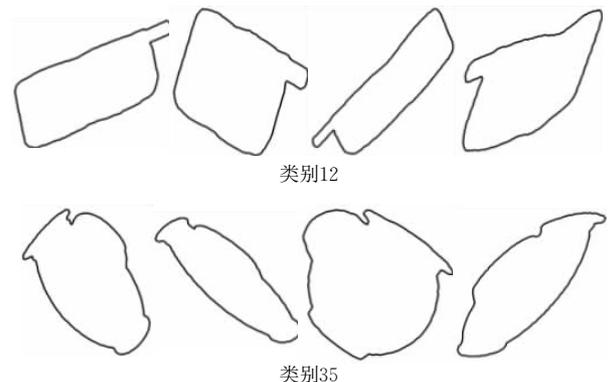


图 15 MCD 数据集类别 12 和类别 35 的形状示例

#### 4.4 PLD 数据集

PLD 数据集<sup>[45]</sup>从 MPEG7 中选取 10 类形状,然后在每类形状上随机的应用 9 种射影变换,最后生成 100 个形状(10×10).虽然 PLD 中的形状数目比 MCD 中的少许多,但是应用在每个形状上的射影变换比应用在 MCD 数据集上的射影变换更加剧烈.图 16 为 PLD 数据集中类别 1 的所有形状.可以看出有的形状受到十分剧烈的射影变换,如第 2 行的第 2 和第 4 个形状.

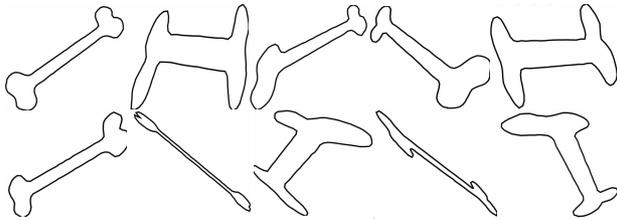


图 16 PLD 数据集类别 1 全体形状

由表 4 可以看出,本文的方法在两类测试中均比 BCF 高 12%左右.在一半训练一半测试实验中,本文方法的分类准确率比 Class segment set 高 33%,在只留一个测试实验中,本文方法的分类准确率比 Class segment set 高 25%;而在两类测试中,Contour segments 的分类准确率只有 19%和 21%.同样的,我们测试了本文方法及其对比方法在 PLD 数据集各个类别中的分类准确率.图 17 为各个类别的示例形状.实验结果如表 5 所示,本文方法对于各个类别形状的分类准确率均不低于 90%;方法 BCF 只在类别 6~8 中获得了不低于 90%的分类准确率,在其它类别中的分类准确率均不高于 80%;方法 Class segment set 只在类别 6 获得了 90%的分类准确率,其他类别的分类准确率均不高于 80%;而方

表 4 PLD 数据集上的分类准确率对比

方法	分类准确率	
	一半训练一半测试	只留一个测试
Class segment set	61.8	70
Contour segments	19.0	21
BCF	80.0	81
Our method	94.4	95

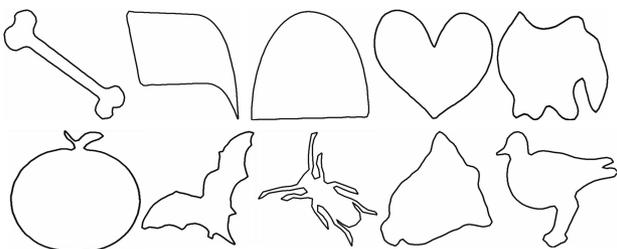


图 17 PLD 数据集所有类别示例形状

法 Contour segments 在各个类别的分类准确率均不高于 40%.图 18 给出了 PLD 数据集中类别 2 和类别 9 的形状示例,对于这两类形状的识别,我们方法的准确率明显高于其他对比方法.

表 5 PLD 数据集中各类别分类准确率对比

类别号	方法			
	Class segment set	Contour segments	BCF	Our method
1	80	20	80	100
2	70	0	80	100
3	50	20	80	90
4	80	20	60	90
5	60	30	70	90
6	90	30	100	90
7	60	40	90	90
8	80	20	100	100
9	50	0	70	100
10	80	30	70	100

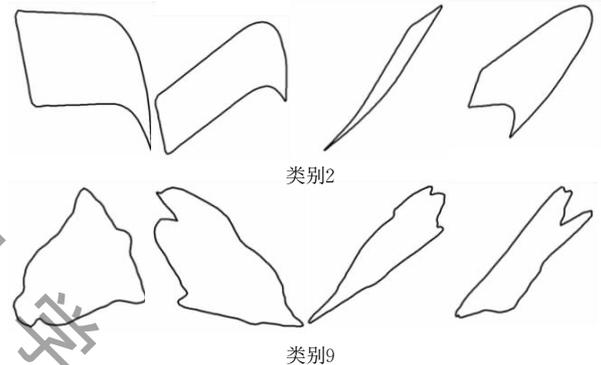


图 18 PLD 数据集类别 2 和类别 9 形状示例

与 BCF 相比,本文的方法有两个优势.首先,因为使用 HCNC,所以本文方法在受到射影变换图像的分类中有更好的表现.另外,BCF 中使用了空间金字塔,这种方法旋转不变性较差.而本文的方法使用的是曲率,曲率具有旋转不变性.这两个优势的相结合使得本文方法在 MCD 数据集及 PLD 数据上的分类准确率比 BCF 的分类准确率高.

Class segment set 和 Contour segments 均是基于轮廓段的方法.两种方法首先将所有的轮廓段归一化,然后使用所有点的坐标描述轮廓段.这种特征表示方法虽然具有旋转、平移、尺度不变性,但是无法描述射影不变特征.所以当形状发生射影变换时,这两种方法的分类准确率低于本文方法.

#### 4.5 对噪声的鲁棒性实验

之前的实验中,数据集中的所有形状都是平滑的.为了评估本文方法在有噪声情况下的性能,我们给形状添加噪声,测试数据集为 MCD,然后再次测试本文方法的分类准确率.噪声被添加在每个形状

的样本点上, 每个样本点的  $x$  坐标和  $y$  坐标都使用高斯随机变量进行变化, 高斯随机变量的均值是 0 方差为  $\sigma$ , 随着  $\sigma$  的增加, 轮廓上的噪声也在增加, 图 19 为一个形状添加噪声后的形状.

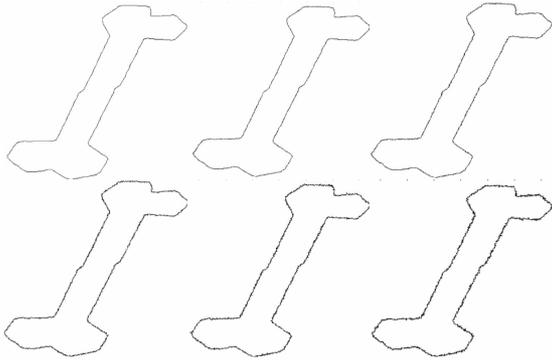


图 19 添加噪声后的形状(从左到右, 从上到下噪声率  $\sigma$  从 0 变化到 1, 间隔为 0.2)

图 20、图 21 表示在不同的噪声下, 本文方法与 BCF 分类准确率的变化. 由图 20 和图 21 可以看出, 当  $\sigma$  从 0 变化到 1, 本文方法的分类准确率在一半训练一半测试实验中下降了 8%, 只留一个测试实验中下降了 5%; BCF 的分类准确率均下降了 5%, 从对

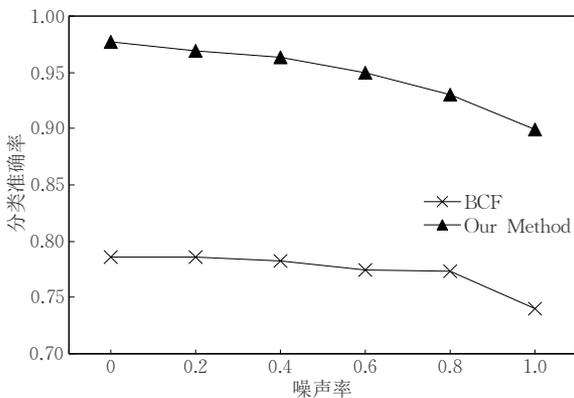


图 20 MCD 数据集上添加噪声后各个方法在一半训练一半测试实验中的分类准确率

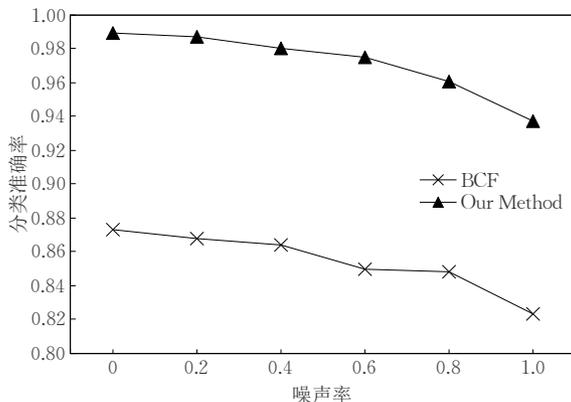


图 21 MCD 数据集上添加噪声后各个方法在只留一个测试实验中的分类准确率

比结果可以看出, 本文方法对噪声具有鲁棒性. 这是因为我们使用的轮廓提取方法 DCE 和描述子 HCNC 对噪声具有鲁棒性.

#### 4.6 时间复杂度

根据表 2~表 5 的实验结果, 我们可以看出本文方法在射影变换形状的分类问题中, 分类准确率高与其他方法. 为了更进一步证明本文方法的高效性, 我们分析了本文方法及其对比方法形状特征提取部分的时间复杂度. 假设在每个轮廓段上取  $N$  个样本点. BCF 方法中形状上下文的时间复杂度为  $O(N^2)$ , Class segment set 方法中形状特征提取的时间复杂度为  $O(N^3)$ ; Contour segments 方法中形状特征提取的时间复杂度为  $O(N)$ ; 本文方法中层次上下文的时间复杂度为  $O(NL)$  ( $L$  为计算层次上下文的尺度). 因此本文方法时间复杂度较低.

## 5 结 论

本文提出了一种基于曲率分级的形状编码方法, 这种方法将新近发现的射影不变量——特征数引入基本特征描述子的构造过程中, 并进一步利用聚类 and 编码的方式获取形状的更高层次的表示, 有效地避免了基本特征描述子需要采样点匹配这一弊端最后利用曲率分级方法使获得的形状特征区分性更强. 实验表明本文方法在射影变换数据库下的识别率提升明显, 比目前最好的方法高出 10% 以上.

将低层次的几何属性度量分析转向中、高层次的特征关联与结构分析一直是我们的研究方向. 未来的研究中我们将继续探索二者之间的关系.

## 参 考 文 献

- [1] 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
- [2] Ling H, Jacobs D W. Shape classification using the inner-distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, 29(2): 286-299
- [3] Wang J, Bai X, You X, et al. Shape matching and classification using height functions. *Pattern Recognition Letters*, 2012, 33(2): 134-143
- [4] Li L, Tan C L. Recognizing planar symbols with severe perspective deformation. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2009, 32(4): 755-762

- [5] Srestasathiern P, Yilmaz A. Planar shape representation and matching under projective transformation. *Computer Vision and Image Understanding*, 2011, 115(11): 1525-1535
- [6] Jia Q, Fan X, Luo Z, et al. A new geometric descriptor for symbols with affine deformations. *Pattern Recognition Letters*, 2014, 40: 128-135
- [7] Jia Q, Fan X, Liu Y, et al. Hierarchical projective invariant contexts for shape recognition. *Pattern Recognition*, 2016, 52: 358-374
- [8] Pylyshyn Z W. Mental imagery: In search of a theory. *Behavioral and Brain Sciences*, 2002, 25(2): 157-182
- [9] Fei-Fei L, Perona P. A Bayesian hierarchical model for learning natural scene categories//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. San Diego, USA, 2005: 524-531
- [10] Hu R X, Jia W, Ling H, et al. Angular pattern and binary angular pattern for shape retrieval. *IEEE Transactions on Image Processing*, 2014, 23(3): 1118-1127
- [11] Wang X, Feng B, Bai X, et al. Bag of contour fragments for robust shape classification. *Pattern Recognition*, 2014, 47(6): 2116-2125
- [12] Latecki L J, Lakämper R. Convexity rule for shape decomposition based on discrete contour evolution. *Computer Vision and Image Understanding*, 1999, 73(3): 441-454
- [13] Wang J, Yang J, Yu K, et al. Locality-constrained linear coding for image classification//*Proceedings of the 23rd IEEE Conference on Computer Vision and Pattern Recognition*. San Francisco, USA, 2010: 3360-3367
- [14] Zhao L, Peng Q, Huang B. Shape matching algorithm based on shape contexts. *IET Computer Vision*, 2015, 9(5): 681-690
- [15] Wang Bin. A Fourier shape descriptor based on multi-level chord length function. *Chinese Journal of Computers*, 2010, 33(12): 2387-2396 (in Chinese)  
(王斌. 一种基于多级弦长函数的傅立叶形状描述子. *计算机学报*, 2010, 33(12): 2387-2396)
- [16] Roman-Rangel E, Pallan Gayol C, Odobez J M, et al. Searching the past: An improved shape descriptor to retrieve Maya hieroglyphs//*Proceedings of the 19th ACM International Conference on Multimedia*. Scottsdale, USA, 2011: 163-172
- [17] Daliri M R, Torre V. Robust symbolic representation for shape recognition and retrieval. *Pattern Recognition*, 2008, 41(5): 1782-1798
- [18] Grigorescu C, Petkov N. Distance sets for shape filters and shape recognition. *IEEE Transactions on Image Processing*, 2003, 12(10): 1274-1286
- [19] Tu Z, Zheng S, Yuille A. Shape matching and registration by data-driven EM. *Computer Vision and Image Understanding*, 2008, 109(3): 290-304
- [20] Rube I E, Ahmed M, Kamel M. Affine invariant multiscale wavelet-based shape matching algorithm//*Proceedings of the 1st Computer and Robot Vision*. Boston, USA, 2004: 217-224
- [21] Fang C H, Lin W S. Two-dimensional shape recognition with cosine transform-based synthesized affine invariant function//*Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*. Montreal, Canada, 2007: 1072-1076
- [22] Raviv D, Kimmel R. Affine invariant geometry for non-rigid shapes. *International Journal of Computer Vision*, 2015, 111(1): 1-11
- [23] Roh K S, Kweow I S. 2-D object recognition using invariant contour descriptor and projective refinement. *Pattern Recognition*, 1998, 31(4): 441-455
- [24] Tripathi V, Gangodkar D, Latta V, et al. Robust abnormal event recognition via motion and shape analysis at ATM installations. *Journal of Electrical and Computer Engineering*, 2015, 2: 1-10
- [25] Tarel J P, Wolovich W A, Cooper D B. Covariant-conics decomposition of quartics for 2d shape recognition and alignment. *Journal of Mathematical Imaging and Vision*, 2003, 19(3): 255-273
- [26] Sun K B, Super B J. Classification of contour shapes using class segment sets//*Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. San Diego, USA, 2005: 727-733
- [27] Luo Z X, Zhou X C, Gu D X F. From a projective invariant to some new properties of algebraic hypersurfaces. *Science China Mathematics*, 2014, 57(11): 2273-2284
- [28] Moosmann F, Nowak E, Jurie F. Randomized clustering forests for image classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008, 30(9): 1632-1646
- [29] Yang J, Yu K, Huang T. Supervised translation-invariant sparse coding//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. San Francisco, USA, 2010: 3517-3524
- [30] Chatfield K, Lempitsky V S, Vedaldi A, et al. The devil is in the details: An evaluation of recent feature encoding methods//*Proceedings of British Machine Vision Conference*. Dundee, Scotland, 2011, 2(4): 8
- [31] Duda R O, Hart P E. *Pattern Classification and Scene Analysis*. New York, USA: Wiley, 1973
- [32] Lazebnik S, Schmid C, Ponce J. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. New York, USA, 2006: 2169-2178
- [33] Levy H, Cartan E. *La Méthode de Repère Mobile, La Théorie des Groupes Continus et Les Espaces Généralisés*.

Bulletin of the American Mathematical Society, 1935, 41(11): 774

- [34] Huang Y, Wu Z, Wang L, et al. Feature coding in image classification: A comprehensive study. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, 36(3): 493-506
- [35] Boureau Y L, Ponce J, LeCun Y. A theoretical analysis of feature pooling in visual recognition//*Proceedings of the 27th International Conference on Machine Learning*. Haifa, Israel, 2010: 111-118
- [36] Perronnin F, Sánchez J, Mensink T. Improving the fisher kernel for large-scale image classification//*Proceedings of the European Conference on Computer Vision*. Heraklion, Greece, 2010: 143-156
- [37] Crammer K, Singer Y. On the algorithmic implementation of multiclass kernel-based vector machines. *The Journal of Machine Learning Research*, 2002, 2: 265-292
- [38] Fan R E, Chang K W, Hsieh C J, et al. LIBLINEAR: A library for large linear classification. *The Journal of Machine Learning Research*, 2008, 9: 1871-1874
- [39] Zuliani M, Bhagavathy S, Manjunath B S, et al. Affine-invariant curve matching//*Proceedings of the 2004 Interna-*

tional Conference on Image Processing. Singapore, 2004: 3041-3044

- [40] Bai X, Liu W, Tu Z. Integrating contour and skeleton for shape classification//*Proceedings of the International Conference on Computer Vision Workshops*. Kyoto, Japan, 2009: 360-367
- [41] Latecki L J, Lakämper R, Eckhardt U. Shape descriptors for non-rigid shapes with a single closed contour//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Hilton Head Island, USA, 2000: 424-429
- [42] Fan X, Wang H, Luo Z, et al. Fiducial facial point extraction using a novel projective invariant. *IEEE Transactions on Image Processing*, 2015, 24(3): 1164-1177
- [43] Attalla E, Siy P. Robust shape similarity retrieval based on contour segmentation polygonal multiresolution and elastic matching. *Pattern Recognition*, 2005, 38(12): 2229-2241
- [44] Daliri M R, Torre V. Shape recognition and retrieval using string of symbols//*Proceedings of the International Conference on Machine Learning and Applications*. Florida, USA, 2006: 101-108
- [45] Bryner D, Klassen E, Le H, et al. 2D affine and projective shape analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, 36(5): 998-1011

## 附录 1. 少于 5 点不能构造射影不变量.

**定理 1.** 假设  $b^{(i)} (i=1,2,3,4)$  为共面 4 点, 其中任意 3 点不共线; 另取 4 个共面点  $\tilde{b}^{(i)} (i=1,2,3,4)$  满足任意 3 点不共线, 那么有且仅有一个射影变换矩阵  $T$  满足  $\tilde{b}^{(i)} = T b^{(i)}$ .

证明. 首先证明存在性, 再证明唯一性.

存在性证明: 4 点  $b^{(i)} (i=1,2,3,4)$  可以确定一个射影坐标系  $\sigma = [b^{(1)}, b^{(2)}, b^{(3)}; b^{(4)}]$ ; 同理,  $\tilde{b}^{(i)} (i=1,2,3,4)$  可以确定另一个射影坐标系  $\sigma' = [\tilde{b}^{(1)}, \tilde{b}^{(2)}, \tilde{b}^{(3)}; \tilde{b}^{(4)}]$ . 在射影平面的齐次坐标中, 令  $b^{(1)} + b^{(2)} + b^{(3)} = b^{(4)}$  及  $\tilde{b}^{(1)} + \tilde{b}^{(2)} + \tilde{b}^{(3)} = \tilde{b}^{(4)}$ . 令  $\sigma^* = [O^{(1)}, O^{(2)}, O^{(3)}, e]$  为辅助坐标系, 那么从  $O^{(i)} (i=1,2,3)$  到  $b^{(i)} (i=1,2,3,4)$  的射影变换可以表示为

$$T_1: \lambda \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1^{(1)} & b_1^{(2)} & b_1^{(3)} \\ b_2^{(1)} & b_2^{(2)} & b_2^{(3)} \\ b_3^{(1)} & b_3^{(2)} & b_3^{(3)} \end{pmatrix} \begin{pmatrix} x_1^* \\ x_2^* \\ x_3^* \end{pmatrix} \quad (1)$$

其中  $\det(b_j^i) \neq 0$ . 同理, 可以将从  $O^{(i)} (i=1,2,3)$  到  $\tilde{b}^{(i)} (i=1,2,3,4)$  的射影变换表示为

$$T_2: \mu \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{pmatrix} \tilde{b}_1^{(1)} & \tilde{b}_1^{(2)} & \tilde{b}_1^{(3)} \\ \tilde{b}_2^{(1)} & \tilde{b}_2^{(2)} & \tilde{b}_2^{(3)} \\ \tilde{b}_3^{(1)} & \tilde{b}_3^{(2)} & \tilde{b}_3^{(3)} \end{pmatrix} \begin{pmatrix} x_1^* \\ x_2^* \\ x_3^* \end{pmatrix} \quad (2)$$

其中  $\det(\tilde{b}_j^i) \neq 0$ . 假设  $T = T_2 \cdot T_1^{-1}$ , 此时有

$$T: \rho \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{pmatrix} \tilde{b}_1^{(1)} & \tilde{b}_1^{(2)} & \tilde{b}_1^{(3)} \\ \tilde{b}_2^{(1)} & \tilde{b}_2^{(2)} & \tilde{b}_2^{(3)} \\ \tilde{b}_3^{(1)} & \tilde{b}_3^{(2)} & \tilde{b}_3^{(3)} \end{pmatrix} \begin{pmatrix} b_1^{(1)} & b_1^{(2)} & b_1^{(3)} \\ b_2^{(1)} & b_2^{(2)} & b_2^{(3)} \\ b_3^{(1)} & b_3^{(2)} & b_3^{(3)} \end{pmatrix}^{-1} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad (3)$$

因此, 存在射影变换矩阵  $T$  使得  $\tilde{b}^{(i)} = T b^{(i)}$ .

唯一性证明: 假设存在另一个射影变换矩阵  $T'$  满足等式  $\tilde{b}^{(i)} = T' b^{(i)} (i=1,2,3,4)$ , 因此  $T'$  满足如下等式:

$$T' \cdot T_1 [O^{(1)}, O^{(2)}, O^{(3)}; e] = T' [b^{(1)}, b^{(2)}, b^{(3)}; b^{(4)}] \\ = [\tilde{b}^{(1)}, \tilde{b}^{(2)}, \tilde{b}^{(3)}; \tilde{b}^{(4)}] \quad (4)$$

因为已经有

$$T_2 [O^{(1)}, O^{(2)}, O^{(3)}; e] = [\tilde{b}^{(1)}, \tilde{b}^{(2)}, \tilde{b}^{(3)}; \tilde{b}^{(4)}] \quad (5)$$

而且  $T_2$  具有唯一性, 所以  $T$  和  $T'$  之间的关系可以表示为

$$T' \cdot T_1 = T_2 \Rightarrow (T' \cdot T_1) \cdot T_1^{-1} = T_2 \cdot T_1^{-1} \\ \Rightarrow T' = T_2 \cdot T_1^{-1} = T \quad (6)$$

等式(6)可以证明射影变换矩阵  $T$  具有唯一性.

由定理 1 我们可以得出结论: 任意 3 点不共线的共面 4 点可以唯一确定一个射影变换矩阵. 换句话说, 这 4 点在射影变换下没有任何区分性, 不能定义用于形状识别的射影不变量. 因此我们至少需要 5 点才能构造射影不变量, 并且可以将构造的不变量用于形状识别.



**JIA Qi**, born in 1983, Ph. D. , associate professor. Her current research interests include image processing and computer vision.

research interest is computer vision.

**FAN Xin**, born in 1977, Ph. D. , professor. His research interests include computer vision and image processing.

**GAO Xin-Kai**, born in 1993, postgraduate. His main research interest is computer vision.

**GUO He**, born in 1955, Ph. D. , professor. His research interests include distributed computing and computer vision and parallel.

**YU Mei-Yu**, born in 1992, postgraduate. Her main

## Background

A new shape feature extraction and recognition method is proposed in this paper. Feature extraction and object recognition is a fundamental problem in computer vision. Currently, most of the existing works about this problem are focusing on basic shape descriptor. The relevance between sample points are usually used as descriptors, such as angle, distance and so on. These methods are easy for computing, but can not keep robustness under affine and projective transformations. There are also some works that introduce invariant to shape descriptor. Our previous works [7] and [6] also provide some projective invariant descriptors. But these methods still need to align points before matching. When doing alignment, we always use Dynamic Time Warping Algorithm. In this way, matching process needs a lot of time but have low accuracy.

In this paper, we use coding-based method to resolve problems occurred in alignment process of basic shape descriptor. At the same time, we introduce a projective invariant descriptor to our method, and then take the influence of contour curvature into consideration, remove redundancy of coding, improve representativeness of code, finally achieve high accuracy in shape recognition.

This task is one part of a project called “Research on Complex Shape Description and Abstract Character Expression Based on 3D/2D Internal Constraints”, this project (No. 61402077, 61432003, 61328206, 11171052, 61876030) is funded by the Youth Program of National Natural Science Fund. At present, most of the existing works describe a shape in a low level in geometry way. The progress made in this project is a meaningful research in field of 3D/2D shape analysis. Moreover, it is an important extension and valid complementary to existed shape analysis methods.

Most of researches in this filed are only use 2D or 3D shape descriptors to solve the problem on each side. How to use 3D constraint to analyze 2D shape is an important part of our project. And our team has make some progress about this problem.

The main contribution of this paper is forming a new shape descriptor in middle/high level, so that we can obtain a more discriminative shape descriptor and achieve a high accuracy in recognition. The progress made in this paper is a solution to some problems in our project.