

分类号 _____

密级 _____

UDC _____

编号 _____

中国科学院研究生院 硕士学位论文

基于形状的图像匹配

张翠

指导教师 _____ 普林特副研究员

中国科学院自动化研究所

申请学位级别 硕士 学科专业名称 模式识别与智能系统

论文提交日期 2011年5月 论文答辩日期 2011年6月

培养单位 _____ 中国科学院自动化研究所

学位授予单位 _____ 中国科学院研究生院

答辩委员会主席 _____

Typeset by L^AT_EX 2 ϵ at May 8, 2011

With package C_ASt_hesis v0.1j of C_TE_X.ORG

Image Matching Based on Shape

Cui Zhang

Supervisor:

Prof. V.Prinè

Institute of Automation
Chinese Academy of Sciences

May, 2011

*Submitted in total fulfilment of the requirements for the degree of Master
in Pattern Recognition and Intelligent System*

独创性声明

本人声明所成交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确地说明并表示了谢意。

研究生签名： _____ 日期： _____

关于论文使用授权的声明

本人完全了解中国科学院有关保留、使用学位论文的规定，即：中国科学院自动化研究所有权保留送交论文的复印件，允许论文被查阅和借阅；可以公布论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存论文。

(保密的学位论文在解密后应遵守此协议)

研究生签名： _____ 导师签名： _____ 日期： _____

摘 要

在人的视觉感知、识别和理解中,形状是图像中目标的基本内在特性,是用于目标识别的重要特征,因此基于形状的图像匹配、目标识别方法研究具有重要意义。针对不同的研究目的,人们已经提出了许多种不同的形状匹配方法,如:k邻接分割(kAS) [11],金字塔匹配[21],形状上下文[6-8],基于内积的形状上下文[22, 23]等。形状的描述及其表达依然是一个开放性问题,其中最困难的问题是缺乏对于形状这一概念的清晰的定义[42]。虽然,形状的轮廓能够通过参数曲线,例如:傅立叶系数等来表示,但是,这些数学方法难以解决我们在描述和比较不同形状时所遇到的问题。例如:谱图理论在图像分割中得到了广泛的应用,但是,在应用到3D图像匹配时却需要对尺度进行归一化[19, 20];而形状上下文描述子也受到各位学者的推崇,但是却依赖于对目标轮廓的检测。

本文的目的在于探索新方法来描述和比较目标的形状,并且应用到形状匹配和识别等方面。因此,本文首先系统地分析了目前较流行的几类形状描述算子,并比较了它们的优缺点。接下来详细介绍了我们基于谱图理论,以及形状上下文描述子的工作。本文的主要工作和贡献可以概括如下:

1. 提出了基于特征向量直方图匹配的形状匹配算法

这部分工作主要基于D.Knosow等人基于谱图理论的模糊匹配[19, 20]的工作展开的,由于图的Laplacian矩阵同时包含了形状轮廓的局部和全局信息,我们通过Laplacian矩阵的特征向量直方图来表示2维形状轮廓。但是,谱分析本身存在许多难点,例如,如何有效的进行特征值排序,为了克服上述问题,我们引入子矩阵匹配的方法来对由不同尺度图形的特征谱定义的特征空间匹配。通过设定恰当的距离度量方法来比较两个形状之间的相似性。通过在图像数据库MPEG-7CE-Shape-1上的实验,证明了方法的有效性。

2. 提出了基于夹角直方图匹配的形状匹配算法

为了避免上一章方法中因计算特征值和特征向量而引起的高计算量,这一章我们基于S.Belongie等人的形状上下文描述子[6-8],进行了相关的

改进。首先我们选取轮廓上的局部曲率最大值作为特征点，形状轮廓的中心作为参照点，计算给定特征点相对其他所有点的夹角，从而每个特征点可以表示为相对其他点的角度直方图。所有特征点的直方图可以作为形状的全局描述子，进行形状相似性比较。最后，我们在图像数据库MPEG-7CE-Shape-1上的实验结果验证了该方法的有效性。

关键词： 形状匹配，形状描述子，谱分析，形状上下文

Abstract

Shape is an intrinsic and discriminant property of object for human visual perception. Therefore, shape is very critical for object recognition and image understanding. To address different application cases, a large variety of shape descriptors have been proposed such as: K-Adjacent Segmentation[11], Spatial Pyramid Matching[21], Shape Context[6–8], Shape Context Based on Inner-Distance[22, 23] and so on.

Shape description and representation is an old and still open problem. The intrinsic difficulty is the absence of a clear definition of "what is a shape" [42]. Though classically, a shape contour can be represented as a parametric curve or fourier coefficients for example, these mathematical approaches are far from being satisfying when dealing with the problem of describing and comparing different shapes.

The objective of this master is to explore new ways to describe and compare the shape of objects. Applications of this work are concerned with shape matching and recognition. In the first part of the thesis, we introduce and compare several shape descriptors recently proposed in the computer vision literature, analyze their advantages and limitations. In the rest of the thesis, we present and detail the main work of this master. Contributions can be summarized as follows:

1. A shape matching method based on eigenvector histogram matching is proposed

Inspired by recent researches on spectral graph theory for 3D geometry matching [19, 20], we use histograms of eigenvectors of the Laplacian matrix to represent the 2D shape contour. The motivation for doing so is that the Laplacian of a graph defined over the shape contour encompasses both local and global properties. In this process however, we are facing an intrinsic difficulty of spectral analysis, ie the appropriate ordering of eigenvalues. In order to overcome this problem and to allow the alignment of the basis of

two shapes with different scale defined by their eigenspectrum, we introduce a "block matching" approach. We then define an appropriate 'distance measure', making possible the comparison of two shapes. We illustrate results on the MPEG-7CE-shape1 benchmark database (publicly available).

2. Proposed shape matching method based on included angle histogram descriptor

To avoid the high computation cost required to estimate the eigen-decomposition, we introduce a new shape descriptor. We first select key points with high local curvature along the contour. Using the centroid of the shape as a 'reference point', we then compute the pairwise angles between a given keypoint and all others. This enables us to describe each keypoint with regard to all others by a histogram of angle. The set of histograms computed at all keypoints provide us with a global descriptor of the shape. Application to shape matching are illustrated on the MPEG-7CE-shape1 database.

Keywords: shape matching, shape descriptor, spectral analysis, shape context

目 录

摘要	i
Abstract	iii
目录	v
第一章 绪论	1
1.1 选题背景与意义	1
1.2 国内外研究状况	3
1.3 本文组织结构	4
第二章 基于形状的图像匹配技术综述	9
2.1 引言	9
2.2 基于全局轮廓特征	9
2.2.1 金字塔匹配	9
2.2.2 基于Laplacian特征向量的图匹配	12
2.2.3 形状上下文	13
2.2.4 基于内距的形状分类	17
2.3 基于局部轮廓特征	19
2.3.1 k邻接分割	19
2.3.2 基于轮廓框架的形状识别	24
2.4 小结	26
第三章 基于谱图理论的形状识别	29
3.1 引言	29
3.2 谱图理论回顾	30
3.2.1 图划分	30

3.2.2 图的Laplacian矩阵及其属性	31
3.3 基于特征向量直方图的形状匹配算法	32
3.4 实验结果及分析	34
3.5 总结	37
第四章 基于夹角直方图的形状匹配	41
4.1 引言	41
4.2 基于夹角直方图的形状上下文的形状匹配	41
4.3 实验结果及分析	45
4.4 总结	47
第五章 结束语	51
5.1 本文主要工作	51
5.2 未来工作展望	52
参考文献	55
发表文章目录	61
简历	63
致谢	65

插 图

1.1	两幅杯子图像，虽然，他们的颜色、纹理等内容很难判断他们属于同一类，但是，从形状可以判断很显然他们属于同一类。	2
1.2	农业中的形状匹配	3
1.3	指纹匹配	4
1.4	检索图像（左侧），和从数据库中检索到的图像[?]。	5
1.5	左侧图像通过将检测到的特征点两两相连构成图来重定义，如右图。	5
1.6	将目标形状轮廓表示为梯度方向（横轴）和两点之间长度（纵轴）的二维直方图。右侧为目标轮廓，左侧为对应的DFH直方图表示。	6
1.7	由两条直线组合成的2邻接分割的特征袋。	7
2.1	通过用金字塔匹配核函数计算局部特征的金字塔直方图的交叉，估计集合间特征的最佳匹配[15]。	10
2.2	本例[15]用两个一维特征集合来创建两个直方图金字塔，其中每一行对应一个分辨率。图2.2(a)中左侧为集合 y ，右侧为集合 z ，其中浅色虚线为仓的边界，粗体虚线为在本分辨率下的匹配，粗实线为在更高分辨率下已有的匹配。图2.2(b)为多分辨率直方图。图2.2(c)为图2.2(b)中直方图之间的交集，对应每个分辨率的 $N_i = 2, 2, 1$ ，权重为 $\omega_i = 1, \frac{1}{2}, \frac{1}{4}$ 。	11
2.3	图像的Laplacian矩阵的前4个最小的非零特征值对应的特征向量对应的直方图[19]。	14
2.4	左侧为两幅比较的图像，右侧为对应两图的特征向量直方图，从图中可以看出，他们非常的相似[19]。	14

- 2.5 形状上下文描述子的计算和匹配[6]。其中(a,b)为原始图像, (c,d)为边缘抽样点, (e,f,g)分别为图(c,d)中抽样点 \circ , \diamond , \triangleleft 的形状上下文描述子。每个描述子都是其他点相对参照点的坐标的log - polar直方图。(f)为两幅图像抽样点的匹配结果。..... 15
- 2.6 计算形状上下文描述子的log - polar直方图的图表[7]。..... 16
- 2.7 3个物体, 虚线表示形状内部连接两个点的最短路径[22]。..... 17
- 2.8 图为人造形状给出的例子[22], 其中黄色分割为图形的组成部分, 蓝色为两个或者多个分割的连接处。..... 19
- 2.9 图中给出的四个不同的形状, 提取同样的样本点, 任意两点之间的欧式距离分布完全相同, 但内距的分布却体现出了各自的不同[22]。..... 19
- 2.10 从图中的4个形状中抽取同样数量的样本点, 他们的欧式距离分布基本都不同, 但是, 内距分布除了前两个相似外, 其他的都明显不同[22]。..... 20
- 2.11 实验中, 特征袋中出现频率最高的35个2AS特征[11]。..... 21
- 2.12 边缘链之间的链接[11], 如图(a-c)。其中(a)为终点与终点连接, (b)为T形连接, (c)为切线不连续连接, (d)瓶子形状的边缘链的8个分割, (e)为连接b的分割桥, 标记为弧形。..... 22
- 2.13 创建轮廓分割的6个规则[11], 通过这些规则连接分割与分割桥(标记为弧形), 其中规则2-6连接分割与边缘链 c_i 。..... 22
- 2.14 特征袋中出现频率最高的35个3AS特征[11]。..... 23
- 2.15 a), 示例图像, 3个2AS和CSN连接; b), 由5个分割和他们中间的连接组成的3个边缘链; c), 2个2AS (B,C) 和 (D, E), 每个分割的序号标记在其旁边; d), 3AS (C,A,E); e), 4AF (E,B,C,D); f), 在d)中的AS描述所涉及到的向量 \vec{r}_i [11]。..... 24
- 2.16 在边界框 b 内(绿色), 轮廓框架 \bar{T} (黑色轮廓线)分布在目标中心(绿色十字)周围[36]。..... 25
- 2.17 方向倒角距离[36]: 模板 T 上的边缘点 x_1 的方向倒角距离为从 x_1 到边缘图 E 上的最近邻边缘点 x_2 的距离 d , 他们之间的边缘梯度差值为 $|\phi(x_1) - \phi(x_2)|$ 。..... 26

3.1	特征向量的匹配：低维图的特征向量组与邻域内的高维图的特征相量组进行无序匹配，查找最佳匹配	35
3.2	不同形状的两幅人首马身原始图。	36
3.3	图3.2对应的Laplacian矩阵的特征值曲线，为了图像的简洁，其中大部分特征值为1曲线被省略掉了。	36
3.4	图3.2对应的前9个特征向量直方图，通过直方图匹配，两个特征相量集合的匹配结果为：(11 : 21), (12 : 22), (13 : 23), (14 : 24), (15 : 25), (16 : 26), (17 : 27), (19 : 28)，从而可得到下述结论：两幅图像为同类图像。	38
3.5	不同目标的两幅原始图像：蜥蜴与人首马身。	39
3.6	图3.5对应的前9个特征向量直方图，通过直方图匹配，只有一对特征向量直方图能够匹配(14 : 21)，从而可得到下述结论：两幅图像不是同类图像。	40
4.1	图形的局部曲率最大值点（标记为绿色），以及轮廓拟合直线（标记为红色）。	42
4.2	锤子和苹果图像在不同的位置，具有不同数量的曲率特征点。 ...	43
4.3	图中绿色点为角点，中心的红点为轮廓重心，重心与角点之间的蓝色线为对应角点的表示向量。	44
4.4	同类图像的两张图片，相对与图二，图一轮廓内包含大量的噪声—内部边缘。	46
4.5	图4.4填充后的对应图像。	46
4.6	计算夹角直方图形状上下文描述子时，角度量化图表。	47
4.7	在MPEG-7 CE-Shape-1图像库上的匹配结果。第一列为目标图像，后面的10个为匹配到的结果图像。	48
4.8	在图像库KIMIA[18]的匹配结果。第一列为目标图像，后面为其前11个最佳匹配结果。	49
4.9	由于包含大量的圆弧和直线，他们检测到同样数量角点，并且相对位置相同，特征描述完全相同，但是他们完全是不同的两个形状。	49

- 4.10 属于同一目标的两个形状，但是，由于毛刺的存在，他们的角点数量相差很大，在预处理中被删除。 50

第一章 绪论

1.1 选题背景与意义

视觉是人类从大自然中获取信息的主要途径，外部世界丰富多彩的信息有60%的信息都是通过视觉感知的，因此视觉信息对人类是非常重要的。随着科技的高速发展，人们尝试利用计算机来模拟人类视觉的各种功能，通过一幅或者多幅图像获取周围环境的信息，由此形成了计算机科学这门学科。它是与图像处理、模式识别、计算机图形学、人工智能、人工神经网络、认知学、神经心理学、生理心理学等多门学科密切联系、交叉的综合性学科。目前，已经被广泛应用到遥感图像处理、文字识别、医学图像处理、多媒体技术、图像数据库、工业检测与军事等各个方面。随着研究的不断深入，以及计算机性能的快速提高，计算机视觉将应用到更加复杂的场合。其中目标识别是计算机视觉的核心部分。

在计算机视觉领域，目标识别是进行图像检索、目标跟踪等研究的基础，因此受到越来越多的研究者的关注。目前目标识别已经被广泛应用到我们生活的各个方面，从交通、工业到农业等很多方面，例如，门禁系统、指纹识别、交通监控。目标识别的重点和难点主要集中在：图像特征描述，分类模型，以及分类器学习。相对于外部特征，通过内容进行图像匹配成为一种重要的方法，而这种方法的难点在于如何比较图像之间的相似性。目前图像匹配主要有两种方法：基于密度的特征描述子与基于几何的特征描述子，其中，图像的密度主要指图像的纹理、颜色、亮度等信息；而几何特征则主要指目标的形状、轮廓信息。通常在观察周围环境时，人们首先注意到的是物体及其周围环境的颜色、纹理、形状和空间关系等等，其中相对于颜色、纹理，形状发挥着更加重要的作用[31]。但是，通过形状等几何特征进行图像匹配仍然是基于内容的方法中最艰难的问题之一。事实上，IBM's Query By Image Content, QBIC[29]大概是目前最先进的图像检索系统，通过颜色和纹理进行检索效果较好，而相对的通过形状进行检索的效果则稍逊一筹。目前最新的Alta Vista photo finder[1]也表现出了相似的结果。基于图像密度的图像匹配方法，通常能够取得很好的效果，但是，在极端的光照条件，或者颜色、纹理发生急剧变化

的情况下，这种方法的效果就会大打折扣。例如，如图1.1中，包含两个水杯的图像，仅仅从图像的密度特征—颜色、纹理等特征很难判断，他们属于同一类。相对而言，基于形状等几何特征的图像匹配则表现出了极大的优越性，如图1.1，通过轮廓信息很容易判断两者属于同一类—杯子。



图 1.1: 两幅杯子图像，虽然，他们的颜色、纹理等内容很难判断他们属于同一类，但是，从形状可以判断很显然他们属于同一类。

在视觉信息系统、计算机视觉、模式识别和机器人技术中，形状匹配是其核心技术。形状匹配广泛应用在工业检测、指纹匹配和基于内容的图像检索中。图1.2，1.3和1.4分别展示了需要解决的典型问题：

1. 图1.2展示了一个在农业检测中的应用，其中一个重要的问题是如何进行匹配转化，基于形状特征从一幅水果图像匹配到另一幅的转化法则；
2. 图1.3展示了一个在指纹认证中点集匹配的应用。在提取特征点之后，两个点集必须进行匹配，其中的难点在于两个集合之间并不存在典型的一一对应关系，并且应该对噪声具有鲁棒性；
3. 图1.4展示了一个在多媒体检索中的应用。在给点左侧的检索形状后，如何从数据库中找到包含相似形状的所有图片，其中的问题在于检索图像只有部分信息包含在图像数据中的部分图片的某些部分中。

本文我们的工作主要集中在基于几何形状的图像匹配的方法上，更加注重在有限点集、曲线等几何模型的描述和匹配。

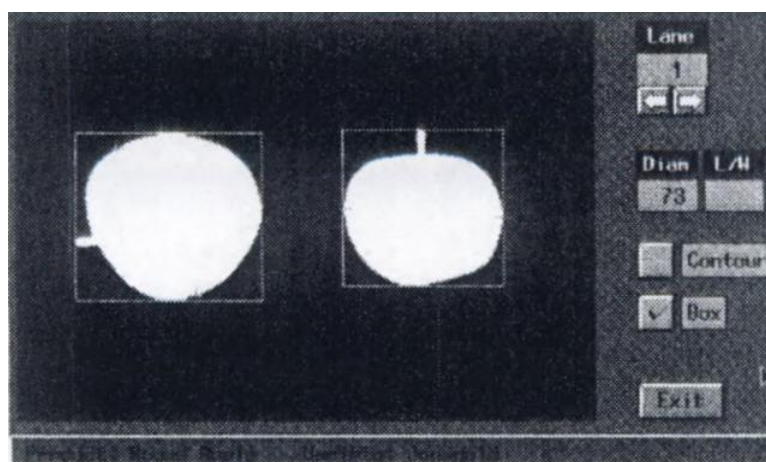


图 1.2: 农业中的形状匹配

1.2 国内外研究状况

目前大量的图像匹配方法已经被提出, 包括: 树剪枝[41]、广义的Hough变换和姿态聚类[3, 39]、几何哈希表[43]、匹配方法[16]、统计[38]、可变形模板[32]、松弛标记[30]、傅立叶描述子[25]、小波变换[17]、曲率尺度空间[27]和神经网络[14]等。每年在各种大型的模式识别会议和杂志, 如Computer Vision and Pattern Recognition (CVPR)、International Conference on Shape Modeling and Applications (ICSMA)、Pattern Analysis and Machine Intelligence (NIPS)、International Journal of Computer Vision (IJCV)、Pattern Analysis and Machine Intelligence (PAMI)、International Conference on Computer Graphics and Interactive Techniques (GRAPH) 等, 都会发表各种关于几何形状匹配在2维和3维图像上的应用。在2000年S.Belongie等人提出形状上下文描述子[6]之后, 由于其简洁的思路及较好的匹配效果, 迅速被相关领域的研究人员H.B.Ling[23]和X.Bai[2]所引用和发展, 可见基于几何形状的图像匹配在模式识别中的重要性, 及其研究的深远意义。

目前, 基于形状的图像匹配主要分为两类: 基于全局形状特征的图像匹配与基于局部形状特征的图像匹配。其中, 全局形状特征是基于图形的整体形状特征来描述目标, 实现目标的匹配或者检测; 而局部形状特征是基于图形中特征点相对局部区域内部点集的特征描述, 作为整体图像相似性度量的基本组成部分。但是, 基于全局形状特征的方法不能显式的表达形状信息, 整个图像轮廓、颜色、纹理信息都被嵌入其中, 不能准确的描述出两幅不同图

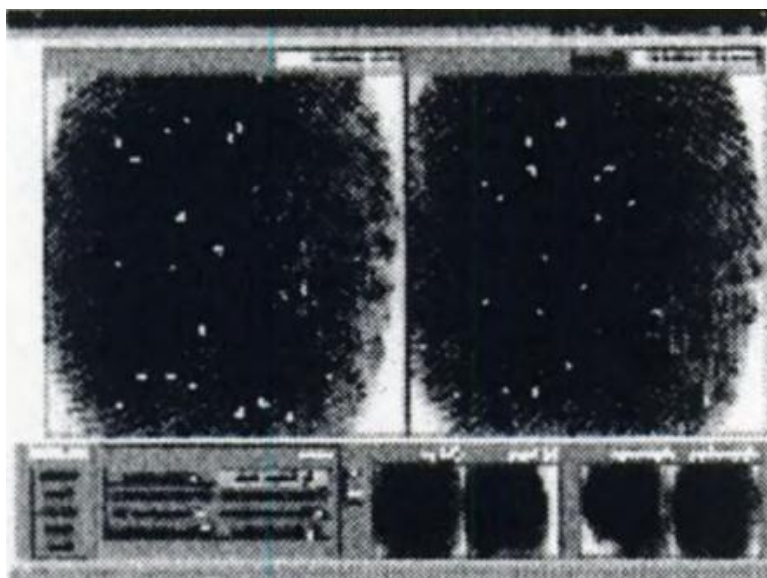


图 1.3: 指纹匹配

像的形状相似性；且特征的全局特性，也决定了它不能用于比较两个子图的相似性。但是，全局特征也有其自身的优势：易于实现分类、识别、模糊匹配，运算简单、有效。例如，U.V.Luxburg等人[26]将图像表示为图，如图1.5，利用图的拉普拉斯矩阵特征向量表示图像，在降维的同时，保留图像的主要信息；I.Yabiaoui等人[44]提出的一种基于形状的方向框架直方图（Directional Fragment Histogram—DFH）方法对图像进行识别，如图1.6，将图像轮廓分割成一系列的分割单位，用分割方向和分割长度的二维直方图来表示目标轮廓。基于局部特征的形状匹配，因其描述的准确性，得到越来越多的研究人员的关注。例如，V.Ferrari等人[11]提出的k邻接分割（k-AS）描述子：通过建立轮廓分割特征袋，如图1.7，将目标轮廓表示成为这些特征的组合。本文综合全局特征与局部特征的优缺点，进行了相应的探索，并提出了我们的改进算法。

1.3 本文组织结构

本文在掌握了目前的几种经典形状匹配算法的基础上，分析了每种方法的优势和不足，结合自己的研究背景和工作，以经典图像库MPEG-7CE-Shape-1为基础，围绕寻找更加高效的形状描述子这一主题进行了研究。并对以下两种算法进行了改进：基于谱图理论的Laplacian特征向量直方图匹配算法、基于

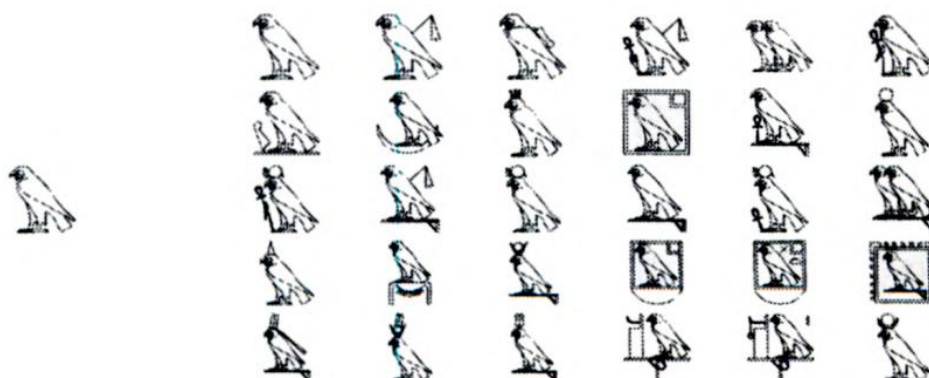


图 1.4: 检索图像（左侧），和从数据库中检索到的图像[?]。

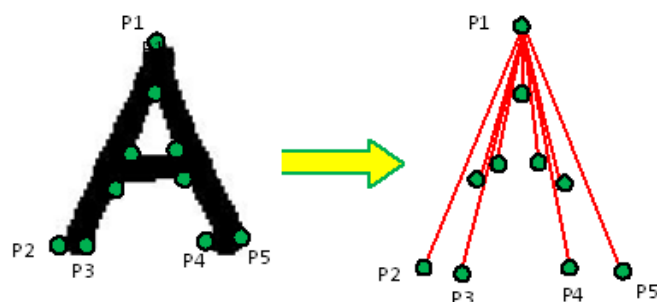


图 1.5: 左侧图像通过将检测到的特征点两两相连构成图来重定义，如右图。

夹角直方图的形状上下文的形状匹配算法。本论文共分五章。

本章主要介绍了本学位论文的研究背景和研究意义以及论文的组织结构。

第二章是对当前主要的基于形状的形象匹配技术的综述。本章主要介绍了几种目前比较流行的形状描述子：V.Ferrari的基于k邻接分割的形象分类，S.Lazebnik等人的金字塔匹配，S.Belongie等人的基于形状上下文的形状匹配算法，J.Shotton等人的基于轮廓框架的形状识别算法，H.B.Ling等人的基于内距的形状分类算法。对于上述每种算法我们都详细介绍了他们的形状描述子、匹配算法和距离计算，并分析了其优缺点。

第三章首先回顾了谱图理论，之后介绍了D.Knossow等人的基于图的拉普拉斯矩阵特征向量的图的模糊匹配算法，最后重点介绍了我们提出的基于特征向量直方图匹配的形象匹配算法，给出了实验结果，并进行了详细的分析和总

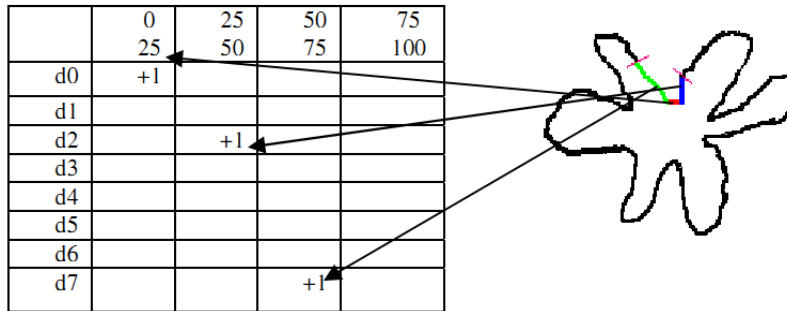


图 1.6: 将目标形状轮廓表示为梯度方向（横轴）和两点之间长度（纵轴）的二维直方图。右侧为目标轮廓，左侧为对应的DFH直方图表示。

结。

第四章主要介绍了我们基于形状上下文描述子提出的基于夹角直方图的形状上下文描述子，最后给出了实验结果，并且根据实验结果进行了详细的分析和总结。

第五章是小结。对本文的主要工作和贡献进行了总结，指出了存在的问题，并且对将来进一步的研究工作进行了展望。

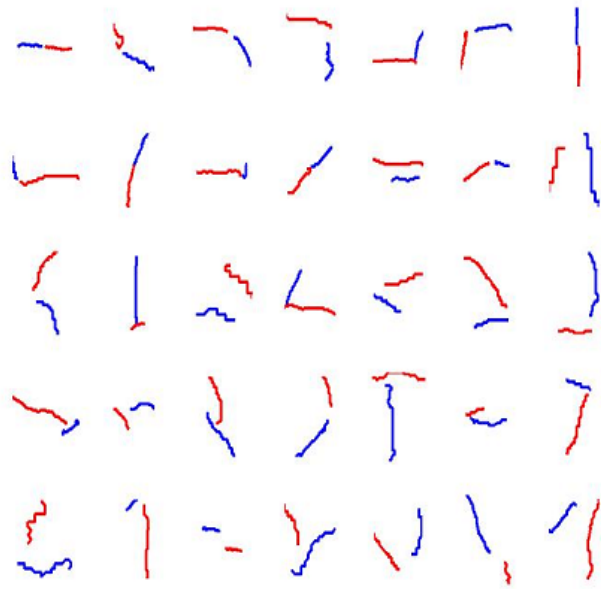


图 1.7: 由两条直线组合成的2邻接分割的特征袋。

第二章 基于形状的图像匹配技术综述

2.1 引言

目前，基于形状轮廓的目标识别算法因为对周围环境、以及自身的颜色、光照、纹理的稳定性，已经受到越来越多研究人员的亲睐。根据图像和目标的不同，例如：图像中边缘噪声的多少、轮廓提取的难易程度等，研究人员分别提出了不同的形状描述子和算法。在本章我们主要介绍几种近来较流行的几种算法所应用的描述子、距离矩阵和分类方法，并且分析了各自的优缺点。

本章的组织结构如下：2.1节为引言，列出章节组织；2.2节介绍四种基于全局轮廓特征的图像匹配算法；2.3节则介绍了两种基于局部轮廓特征的图像匹配算法；最后在2.4节总结了各个算法的特点。

2.2 基于全局轮廓特征

2.2.1 金字塔匹配

金字塔匹配核函数[15, 21]是基于无序特征集的核函数，在基于核函数的学习方法中得到有效应用。每个特征集都被映射为多分辨率的直方图，其中最高分辨率的直方图保留了每个特征的不同信息。通过权重直方图的交叉运算比较金字塔直方图之间的距离，如图2.1

金字塔匹配核函数算法的核心思想是将特征集合映射为多分辨率的直方图，通过权重直方图的交叉度量比较直方图，估计特征集合间最佳部分匹配的相似度。给定有 d 维特征向量集合构成的空间 X ，且 X 为一个直径为 D 的球体有限空间，向量间内积距离的最小值为 $\frac{\sqrt{d}}{2}$ ：

$$X = \{\vec{x} | \vec{x} = \{[f_1^1, \dots, f_d^1], \dots, [f_1^{m_x}, \dots, f_d^{m_x}]\}\} \quad (2.1)$$

其中 m_x 随空间 X 的变化而变化。

特征提取函数 Ψ 定义如下：

$$\Psi(x) = [H_{-1}(x), H_0(x), \dots, H_L(x)] \quad (2.2)$$

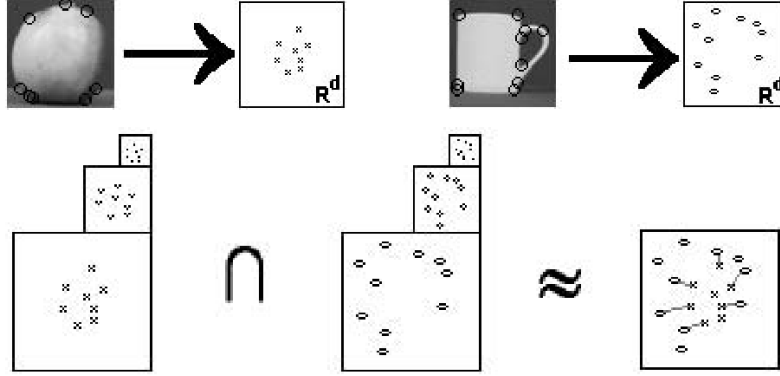


图 2.1: 通过用金字塔匹配核函数计算局部特征的金字塔直方图的交叉, 估计集合间特征的最佳匹配[15]。

其中 $L = \lceil \log_2(D) \rceil$, $\vec{x} \in X$, $H_i(x)$ 为由 d 维的数据 x 得到的直方图向量, 且每个仓的长度为 2^i , $H_i(x)$ 的维度为 $r_i = (\frac{D}{2^i \sqrt{d}})^d$ 。换句话说, $\Psi(x)$ 是一系列直方图串联构成的向量, 其中每个子直方图在任意维上的仓宽都是前一个的 2 倍。分辨率最高的直方图为 H_{-1} , 其仓宽足够小, 使得集合 X 中的任意数据都分别占有一个仓; 且仓宽随着分辨率 i 增大而增大, 当分辨率为 L 时, 集合 X 中所有的数据点都落入同一个仓中。

取金字塔匹配核函数 K_Δ 来衡量两个点集之间基于多分辨率直方图匹配的相似性。两个集合的相似性定义为由函数 Ψ 构造的金字塔的每个分辨率上特征匹配数量的权重和, 如下式:

$$K_\Delta(\Psi(y), \Psi(z)) = \sum_{i=0}^L \omega_i N_i \quad (2.3)$$

其中 N_i 为在分辨率为 i 时新增匹配对的数量。这里新的匹配指在任何更高的分辨率时, 相应的特征对都不匹配。

若两个点位于相同的直方图仓时, 认为两个点匹配, 则计算核函数只需要计算两个点集之间的匹配。上述匹配过程等价于下面的按等级划分的过程: 在高分辨率时不能匹配的向量, 在低分辨率时有可能匹配。例如, 如图 2.2, 在最高分辨率时, 有 2 个点匹配; 在中间分辨率时有两个新的匹配; 在最低分辨率时仅有一个新匹配。 K_Δ 的输出反映了匹配的总相似性, 其中分辨率 i 的每个新匹配对都与分辨率成比例贡献一个值 ω_i 。

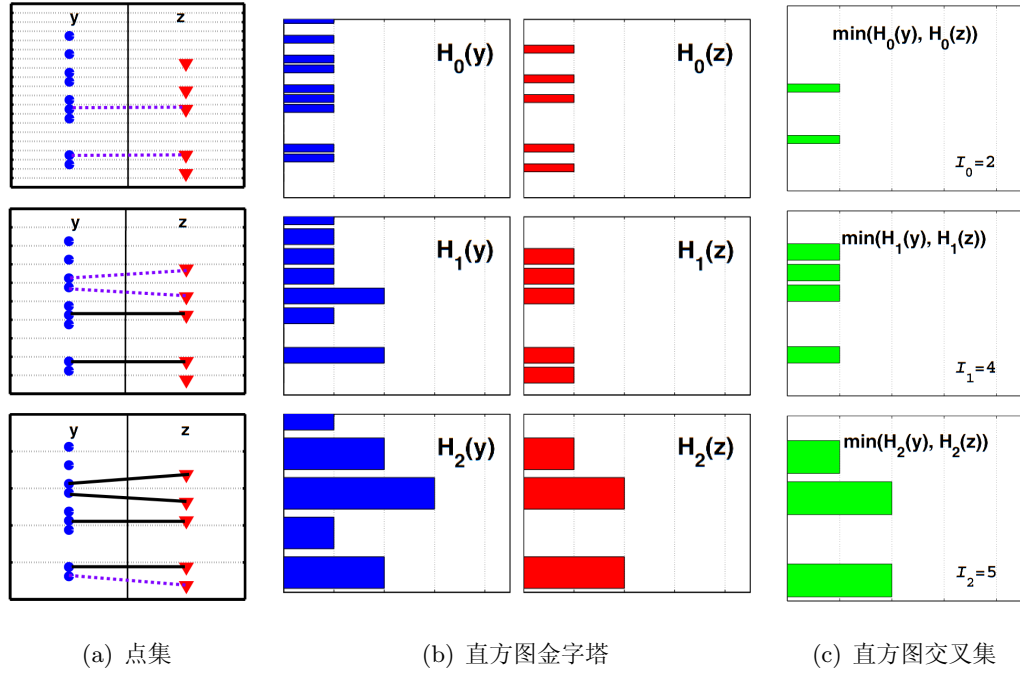


图 2.2: 本例[15]用两个一维特征集合来创建两个直方图金字塔, 其中每一行对应一个分辨率。图2.2(a)中左侧为集合 y , 右侧为集合 z , 其中浅色虚线为仓的边界, 粗体虚线为在本分辨率下的匹配, 粗实线为在更高分辨率下已有的匹配。图2.2(b)为多分辨率直方图。图2.2(c)为图2.2(b)中直方图之间的交集, 对应每个分辨率的 $N_i = 2, 2, 1$, 权重为 $\omega_i = 1, \frac{1}{2}, \frac{1}{4}$ 。

为了计算 N_i , 核函数利用了直方图交叉函数 I , 如下:

$$I(A, B) = \sum_{j=1}^r \min(A^{(j)}, B^{(j)}) \quad (2.4)$$

其中 A 和 B 为 r 维的直方图, $A^{(j)}$ 表示 A 的第 j 个仓的值。在分辨率 i 的最新匹配对数 N_i , 即为两个连续分辨率下匹配数的差值:

$$N_i = I(H_i(y), H_i(z)) - I(H_{i-1}(y), H_{i-1}(z)) \quad (2.5)$$

其中 H_i 为由 Ψ (2.2) 产生的第 i 个子直方图。

综合式??可以得到归一化的金字塔匹配核函数:

$$K_{\Delta} = \frac{1}{\sqrt{C}} \sum_{i=0}^L \frac{1}{2^i} (I(H_i(y), H_i(z)) - I(H_{i-1}(y), H_{i-1}(z))) \quad (2.6)$$

上式可以简记为 $K_\Delta = \frac{1}{\sqrt{C}} \tilde{K}_\Delta(P, Q)$, 其中 $y, z \in X$, $H_i(x)$ 为 $\Psi(x)$ 的第 i 个子直方图, $C = \tilde{K}_\Delta(P, P) \tilde{K}_\Delta(Q, Q)$ 为归一化因子, 防止输入过大。

金字塔匹配算法是基于特征的统计特性, 对于噪声比较稳定。但是, 这种方法依赖于特征的选取, 一方面要保证特征点的数量, 令一方面当特征不包含点的分布信息时, 这种描述子将丢失特征点的空间信息。

2.2.2 基于Laplacian特征向量的图匹配

这一小节我们主要介绍D.Knossow等人[19, 20]基于Laplacian特征向量的稀疏图的模糊匹配算法, 作为后续改进的基础。考虑将一幅图像表示为无向权重图 $G = \{V, E\}$, 其中顶点集合: $V = \{v_1, \dots, v_N\}$, 边集: $E = \{e_{ij}\}$ 。通过公式(3.1, 3.2, 3.4)计算图 G 对应的权重矩阵、度量矩阵和非归一化Laplacian矩阵, 考虑归一化矩阵: $\mathcal{L} = D^{-\frac{1}{2}}(D - W)D^{-\frac{1}{2}}$, 这是一个半正定的对称矩阵, 其特征值满足: $0 = \lambda_0 \leq \lambda_1 \leq \dots \leq \lambda_{N-1}$ 。矩阵对应的零空间为其常数特征向量: $U_0 = (1, \dots, 1)^T$ 。

\mathcal{L} 的特征向量构成正交基: $U_0^T U_i = 0, \forall i \in \{1, \dots, N-1\}$, 即, $\sum_k U_{ik} = 0, \forall i \in \{1, \dots, N-1\}$ 。由 \mathcal{L} 的运算公式可得到 $\mathcal{L} = I - \mathcal{W}$, 其中 $\mathcal{W} = D^{-\frac{1}{2}}(D - W)D^{-\frac{1}{2}}$ 为归一化的权重矩阵, 且 \mathcal{L} 与 \mathcal{W} 具有相同的特征向量。忽略零特征值及其对应的常值特征向量, 可以得到截尾特征分解: $\mathcal{L}_t = U \Lambda U^T$ 。

给定两幅图 G_x 和 G_y , 分别有 N 和 M 个顶点。当 $N = M$ 时, 上述两图的匹配为精确匹配, 转化为求解 $N \times N$ 的交换矩阵 P 最小化下述方程:

$$P^* = \arg \min_P \|\mathcal{W}_x - P \mathcal{W}_y P^T\|_F^2 \quad (2.7)$$

其中 $\mathcal{W}_x - P \mathcal{W}_y P^T = \mathcal{L}_x - P \mathcal{L}_y P^T$ 。令 \mathcal{L}_x 和 \mathcal{L}_y 为对应Laplacian矩阵的截尾特征分解, 即 $\mathcal{L}_x = U_x \Lambda_x U_x^T$ 与 $\mathcal{L}_y = U_y \Lambda_y U_y^T$ 。 $U = [U_1, \dots, U_{N-1}]$ 的列对应特征向量, 构成一组正交基, $\Lambda = \text{Diag}[\lambda_1, \dots, \lambda_{N-1}]$ 。

根据谱图匹配理论[40]: 如果 \mathcal{L}_x 与 \mathcal{L}_y 的特征值各不相同, 且有序, 那么式(2.7)的最优解为:

$$Q^* = U_x S U_y^T \quad (2.8)$$

其中 $S = \text{Diag}[s_1, \dots, s_{N-1}]$, $s \in \{+1, -1\}$ 对应特征分解符号特性。为了减少运算量, 将特征空间降维到 K 维, 其中 $K < (N, M)$ 。通过特征值排序, 每个矩阵可以表示为 K 个有序特征向量构成的特征空间。但是, 由于特征值的几何重

数，基于特征值的排序并不是绝对的，为了解决上述问题，作者提出了新的特征变化矩阵 $P_{K \times K}$ ：

$$Q = \bar{U}_x \bar{S} P \bar{U}_y^T \quad (2.9)$$

其中 \bar{U}_x 和 \bar{U}_y 为通过提取 U_x 与 U_y 的前 K 列，然后进行行归一化后得到的 $(N-1) \times K$ 维矩阵，并且这 $N-1$ 行向量位于 K 维的超球上；同式(2.8)， \bar{S} 为 $K \times K$ 维的符号对角矩阵；矩阵 Q 不再是正交矩阵，秩为 K 。

\bar{U}_x 的每一列 U_{x_i} (\bar{U}_y 的每一列 U_{y_j}) $1 \leq i, j \leq K$ 对应一个 $N-1$ 维的特征向量。因此，向量 U_{x_i} (与 U_{y_j}) 可以认为是将图 \mathcal{G}_x 的顶点映射到实线的函数： $U_{x_i} : \mathcal{R}^{N-1} \rightarrow \mathcal{R}$ ($U_{y_j} : \mathcal{R}^{N-1} \rightarrow \mathcal{R}$)。在 \bar{S} 和 \bar{P} 都已知的情况下，两幅同构图中第一幅图的特征向量与第二幅图的特征向量应该是一一对应的。但是，正如前面所述，并不能保证两个特征基的扩展相同的欧式空间。因此，考虑特征向量直方图 $h([U_{x_i}])$ 和 $h([U_{y_j}])$ ，其中 $h([U])$ 对应特征向量 U 的直方图，如图2.3，并且，特征向量直方图具有下述特点：

1. 直方图对于节点变换具有不变性，即，对于特征向量元素的排序是稳定的；
2. $h([U]) = h(B - [U])$ ，其中 B 为直方图向量的维度；
3. 直方图对于图中的顶点数量不敏感，因此，可以用于比较由不同维度的图得到的直方图；

此时，计算矩阵 \bar{S} 与 \bar{P} 的问题转化为求解基于图之间直方图比较的最佳匹配： $\{U_{x_i} \Leftrightarrow \pm U_{y_j}, 1 \leq i, j \leq K\}$ ，即最大二分匹配问题。由于特征向量的符号特性，每个特征向量对应两个直方图。设 $C(h_i, h_j)$ 为两个直方图之间的相似性，通过计算所有特征向量直方图队之间的相似性，得到两图的相似性矩阵 $\bar{A}_{K \times K}$ ：

$$A_{ij} = \min\{C(h([U_{x_i}]), h([U_{y_j}])), C(h([U_{x_i}]), h([-U_{y_j}]))\} \quad (2.10)$$

通过匈牙利算法即可求解最优变换矩阵 \bar{P} ，以及 \bar{S} ，如图2.4。

2.2.3 形状上下文

通过将文本处理中的上下文特征信息统计的方法引入到形状识别中来，形成了现在在形状识别与匹配中被广泛应用的形状上下文描述子。算法的主要思

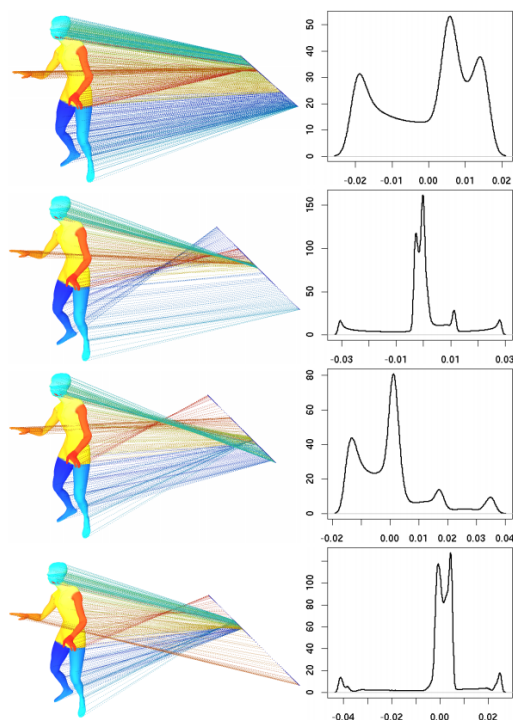


图 2.3: 图像的Laplacian矩阵的前4个最小的非零特征值对应的特征向量对应的直方图[19]。

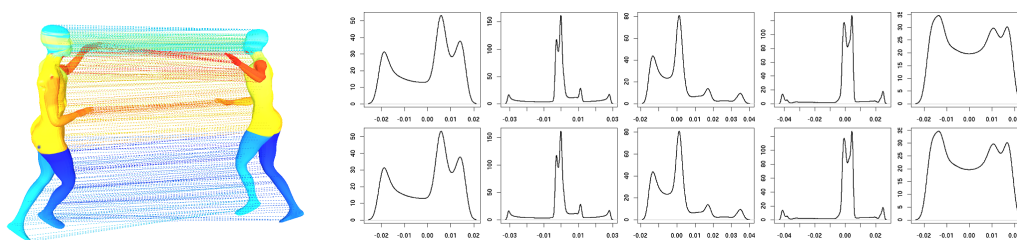


图 2.4: 左侧为两幅比较的图像，右侧为对应两图的特征向量直方图，从图中可以看出，他们非常的相似[19]。

想是[6-8]: 通过提取目标轮廓点作为特征点, 并创建两两连接的图来表示原图像, 通过点集匹配计算图像之间的相似性。

形状上下文分析是将形状的边缘问题转化为包含 N 个特征点的集合问题。但是, 这些点可以是外部轮廓, 也可以是内部轮廓, 并且不一定对应特征点, 例如: 最大曲率点、角点等。通常, 选取相同间隔的点作为形状的特征点。例如, 图中, 图(a)的抽样点为图(c)。考虑从图(c)中的点到形状上所有其他点的向量集合, 这些向量为整个形状相对参照点的表示。显然, $N - 1$ 个向量集合包含了形状的大量信息, 并且, 随着 N 的增长, 形状描述越准确。

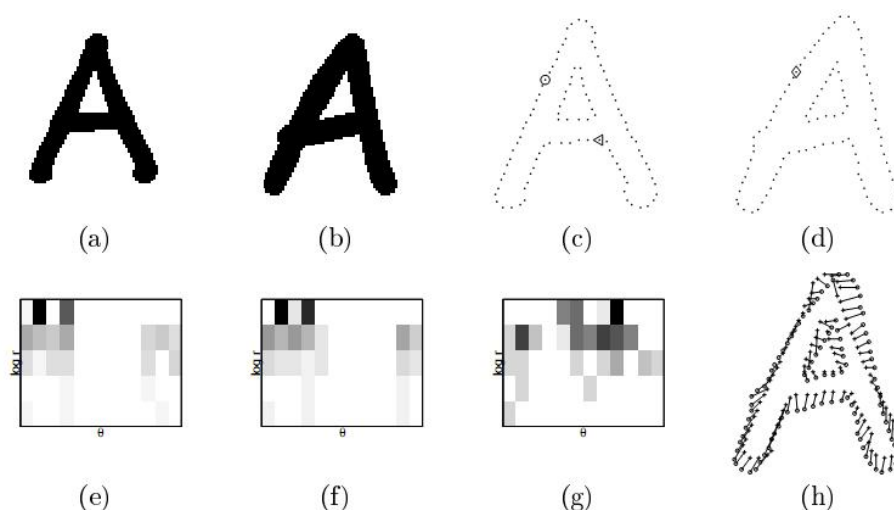


图 2.5: 形状上下文描述子的计算和匹配[6]。其中(a,b)为原始图像, (c,d)为边缘抽样点, (e,f,g)分别为图(c,d)中抽样点 \circ , \diamond , \triangleleft 的形状上下文描述子。每个描述子都是其他点相对参照点的坐标的 $\log - polar$ 直方图。(f)为两幅图像抽样点的匹配结果。

用所有向量的集合作为形状描述子显然是不恰当的, 因为每个形状和其抽样点表示都是不同的。相反, 这里采用所有其他点相对参照点的分布作为参照点的特征表示, 其不仅具有很强的鲁棒性, 并且具有很好的区分度。对于图形上的任意一点 P , 计算其相对其余 $N - 1$ 个的相对位置的分布直方图, 称为点 P 的形状上下文描述子, 其中参照坐标系的方向可以是绝对方向, 也可以取相对方向。为了让描述子对邻近像素有更好的区分度, 作者选取了 \log - $polar$ 坐标系, 如图2.6, 整平面, 按照角度平分为12份, 同时按照 \log - $polar$ 分成5等份。

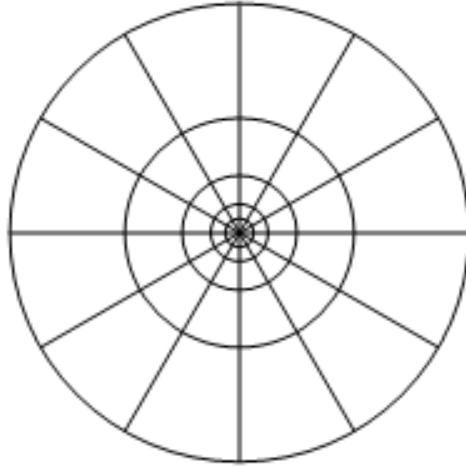


图 2.6: 计算形状上下文描述子的log-polar直方图的图表[7]。

从形状上下文描述子的计算过程，我们可以看出，其具有以下优点：

- 1 .描述子是相对于目标内的其他点计算得到的，因此具有平移不变性；
- 2 .通过选取所有 N^2 个点对之间距离的中值，归一化所有点对之间的距离，实现了描述子的尺度不变性；
- 3 .通过点集之间的迭代匹配，保证了描述子对显著性旋转的鲁棒性；

在进行形状匹配的过程中需要遵循下面两条原则：

- 1 .对应点的描述子应该非常相似；
- 2 .点与点的对应关系应该是一一对应；

考虑第一个形状上的点 i 与第二个形状上的点 j ，计算在点 i 与 j 的形状上下文描述子之间的距离 C_{ij} 来匹配两个点。假设在点 i 的 K 维直方图为 $g(k)$ ，在点 j 的为 $h(k)$ ，则两者之间的距离 C_{ij} 可以通过下式的 χ^2 统计方程计算：

$$C_{ij} = \frac{1}{2} \sum_{k=1}^K \frac{[g(k) - h(k)]^2}{g(k) + h(k)} \quad (2.11)$$

给定了第一个形状上的所有点 i 与第二个形状上的所有点 j 之间的匹配距离矩阵 $C = (C_{ij})_{i=1,\dots,n_1,j=1,\dots,n_2}$ ，为了实现点对之间的最佳匹配，按照一一对应的原则最小化点集之间的匹配距离。点集匹配问题就转化为二分权重图匹配问题，可以通过匈牙利算法在 $O(N^3)$ 复杂度内求解。当两个形状上的抽样点数不同时，可以通过增加虚拟点，设置其到其他点的距离为默认值 ϵ_d ，使最终的距离矩阵依然为平方矩阵。

上述基于形状上下文描述子的形状匹配算法，非常的简单易懂，对于存在内部轮廓的形状，也能够得到很好的应用；并且对于形状的尺度、平移和旋转比较稳定，在实验中取得了非常好的结果。但是，这种描述子必须要在轮廓已知的前提下才能应用，对边缘噪声，以及形状中各个组成部分的分布比较敏感。

2.2.4 基于内距的形状分类

通常我们选取欧式距离作为图像内两点之间距离，但是，如图2.7，在(a)和(c)中的点的空间分布相同，但却不属于同一类；另一方面，在(b)和(c)中的两点的空间分布不同，但却属于同一类；从上面的例子可以看出欧式距离本身具有自身的缺陷。H.B.Ling等人[22–24]提出了内距的概念从一定程度上解决了上述问题。假设两个点 $x, y \in \mathcal{O}$ ，其中 \mathcal{O} 定义为 \mathcal{R}^2 内连通、闭合的形状。 x 与 y 之间的内距定义为 \mathcal{O} 内连接 x 和 y 的最短路径的长度，记为 $d(x, y; \mathcal{O})$ 。当 \mathcal{O} 为凸区域时，内距等价于欧式距离；但是，对于非凸的形状，上述结论并不总是成立，如图2.7。

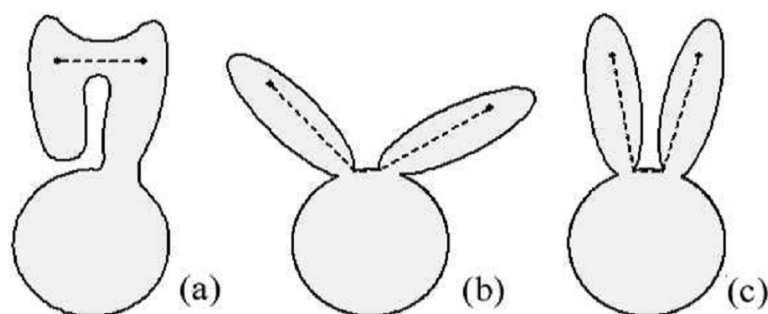


图 2.7: 3个物体，虚线表示形状内部连接两个点的最短路径[22]。

当形状 \mathcal{O} 存在连接部件时, 具有下列属性:

1. \mathcal{O} 可以被分解成一系列分区域: $\mathcal{O}_1, \mathcal{O}_2, \dots, \mathcal{O}_n$;
2. 分区域间的连接处相对于连接的分区域而言非常小;
3. \mathcal{O} 上的连接处当被限制与任何分区域 \mathcal{O}_i 作为转化是非常严格的, 但是在连接处则不严格;
4. 从 \mathcal{O} 的连接中分个出来的形状 \mathcal{O}' 能够作为连接目标重新连接回 \mathcal{O} ;

基于上述信息, 定义任意一个由 n 个分区域通过连接{构成的目标 $\mathcal{O} \in \mathcal{R}^2$:
 $\mathcal{O} = \{\bigcup_{i=1}^n \mathcal{O}_i\} \cup \{\bigcup_{i \neq j} \mathcal{J}_{ij}\}$, 其中:

1. $\forall i, 1 \leq i \leq n$, 分区域 $\mathcal{O} \subset \mathcal{R}^2$ 为连通闭合区域, 且 $\mathcal{O}_i \cap \mathcal{O}_j = \emptyset, \forall i \neq j, 1 \leq i, j \leq n$;
2. $\forall i \neq j, 1 \leq i, j \leq n$, $\mathcal{J}_{ij} \subset \mathcal{R}^2$ 为 \mathcal{O}_i 与 \mathcal{O}_j 的节点。如果 \mathcal{O}_i 与 \mathcal{O}_j 之间没有节点, 则 $\mathcal{J}_{ij} = \emptyset$; 否则, $\mathcal{J}_{ij} \cap \mathcal{O}_i \neq \emptyset$ 且 $\mathcal{J}_{ij} \cap \mathcal{O}_j \neq \emptyset$;
3. $diam(\mathcal{J}_{ij}) \leq \epsilon$, 其中 $diam(P)$ 定义为点集 P 内点对之间的最大距离, 即 $diam(p) = \max_{x, y \in P} \{d(x, y; P)\}$ 。 $\epsilon \geq 0$ 相对于任意分区域的尺度是非常小的。

目标 \mathcal{O} 的节点为从 \mathcal{O} 到 $\mathcal{O}' = f(\mathcal{O}) \subset \mathcal{R}^2$ 的一一映射 f , 具有下列属性:

1. 为连接目标 \mathcal{O} , 有如下分割: $\mathcal{O}' = \{\bigcup_{i=1}^n \mathcal{O}'_i\} \cup \{\bigcup_{i \neq j} \mathcal{J}'_{ij}\}$, 并且 $\mathcal{O}'_i = f(\mathcal{O}_i), \forall i, 1 \leq i \leq n$ 为 \mathcal{O}' 的分区域, $\mathcal{J}'_{ij} = f(\mathcal{J}_{ij}), \forall i \neq j, 1 \leq i, j \leq n$ 为 \mathcal{O}' 内的节点。
2. $\forall i, 1 \leq i \leq n$, 在 \mathcal{O}_i 上的 f 是严格的, 即任何一个分区域内的内距都是不变的。

从上面的定义可知 f^{-1} 为将 \mathcal{O}' 内的分割区域和节点映射到 \mathcal{O} 的函数。图2.8给出了一些形状的例子。

从图2.7、2.9给出的例子, 可以看出, 当欧式距离不能区分形状的不同时, 内距则表现出了很好的效果。除了区分图像结构的优越性外, 图2.10给出了一

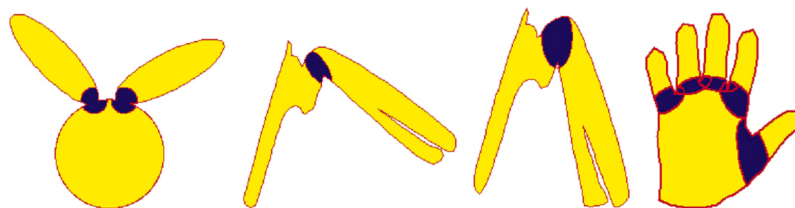


图 2.8: 图为人造形状给出的例子[22], 其中黄色分割为图形的组成部分, 蓝色为两个或者多个分割的连接处。

些例子, 即使形状不能进一步分割, 内距也能够更好的抓取形状信息。在充分了解了内距的定义及计算后, 可以将其用于目前应用比较广泛的形状上下文特征[6]中去, 从而更好的实现形状匹配和识别。

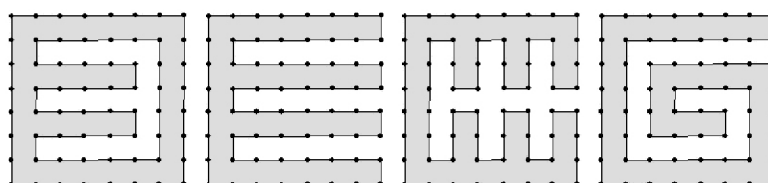


图 2.9: 图中给出的四个不同的形状, 提取同样的样本点, 任意两点之间的欧式距离分布完全相同, 但内距的分布却体现出了各自的不同[22]。

上述基于内距的形状描述子, 对于形状连接不敏感, 并且尤其对于由很多部分组合而成的图形效果更好。但是, 此方法必须基于形状轮廓已知的情况下, 能够对图形进行分割; 另外, 当图形包含很小的组成部分或者严重变形时, 内距的度量会存在严重偏差, 从而影响最终的形状识别与匹配。

2.3 基于局部轮廓特征

2.3.1 k邻接分割

通常目标检测都会受到周围环境、遮挡等影响, 且由于颜色、纹理等因素, 检测到的边缘轮廓信息通常都是不连续的。2006年V.Ferrari等人提出的基于k邻接分割的算法[11]对上述干扰具有较高的稳定下。通过观察发现: 在一定

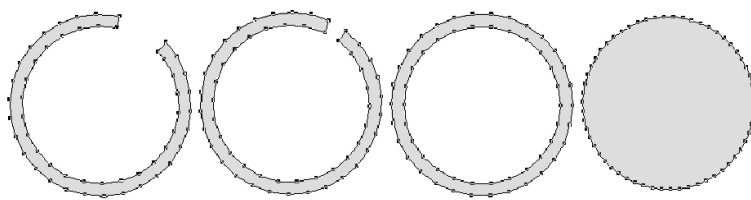


图 2.10: 从图中的4个形状中抽取同样数量的样本点, 他们的欧式距离分布基本都不同, 但是, 内距分布除了前两个相似外, 其他的都明显不同[22]。

的精度范围内, 任何一条曲线都可以表示成直线段的组合, 图像的轮廓曲线可以表示成若干条相邻直线段的组合, 并且在某种度量和一定的相似性范围内, 一定数量的直线段能够表示有限种形状组合。例如, 一条直线段只能表示成横线、竖线、斜线 (4个不同方向); 两条直线段则可以表示成更多的形状, 如图2.11, 简称为 $k-AS$ (k -Adjacent Segments)。线段越多, 构成的形状特征越复杂, 检测和匹配就越困难。

为了使用上述2-AS特征袋, 首先检测目标的边缘, 但是, 通常得到的边缘都会被分割成一段段的边缘, 根据边缘的切线方向, 利用下述方法链接相邻边缘: 如果边缘链 c_2 的任意边缘段位于边缘链 c_1 终点的搜索区域内, 那么连接边缘链 c_1 与 c_2 , 如图2.12。其中这里的搜索区域为等边梯形, 其上底与 c_1 的重心重合, 并且垂直于曲线的切线方向, 从而保证了 c_1 延续位置的不确定性。依据上述标准, 边缘链 c_1 的链接边缘位于其终点。但是, 这里 c_1 既可以链接到 c_2 的终点, 也可以链接到其内部节点, 后者允许T形链接, 即曲线可以在两个方向连续, 如图2.12b; 同时这里的链接允许切线方向不连续, 如图2.12c。

在将所有边缘进行链接后, 通过将所有的边缘链分成直线分割, 进行轮廓分割, 从而建立目标的表达网络, 如图2.12d为瓶子边缘链的分割; 同时为边缘链间切线连续的链接创建分割桥, 从而保留在分割中丢失的链接信息。

在完成上述准备工作后, 将原目标表示为轮廓分割网络。首先, 定义每个分割都是有向的, 即存在起点和终点之分, 一个分割的起点与另一个分割边缘链的终点相连; 另外, 定义边缘链也是有向的, 梯形搜索区域在边缘链 c_1 上的位置为终点, 对应边缘链 c_2 的起点。在给定上述定义之后, 按照下述规则, 通过连接每个分割的起点到其他分割集合, 创建图像目标的轮廓分割网络, 如图2.13:

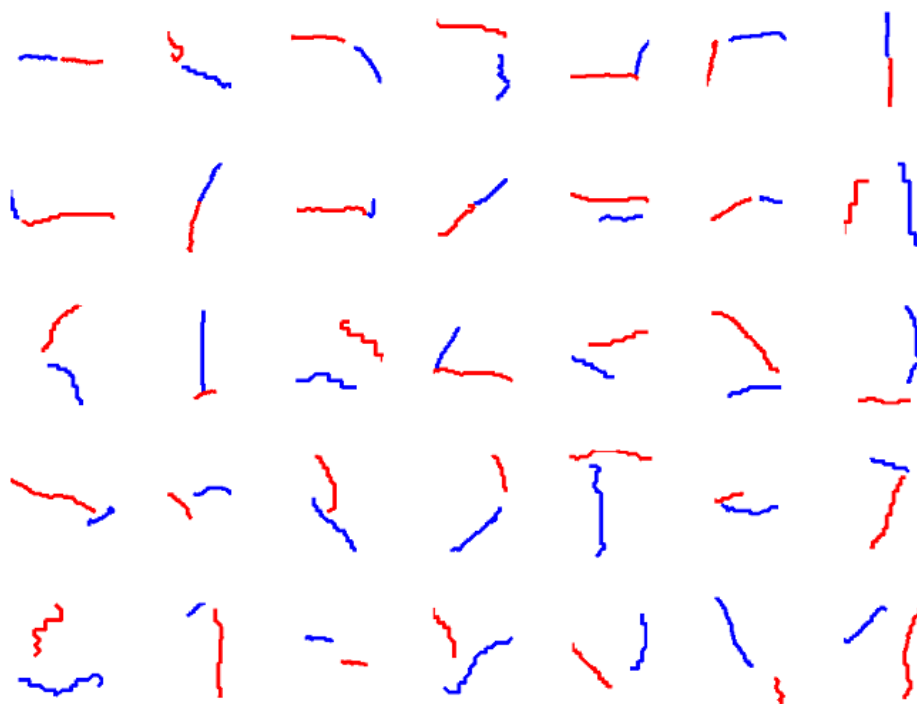


图 2.11: 实验中, 特征袋中出现频率最高的35个2AS特征[11]。

1. 同一边缘链上的两个分割, 一个分割的起点连接另一个分割的终点;
2. 当链接两个边缘链 c_1 , c_2 的终点时, 连接 c_1 链接前的分割与 c_2 链接后的分割;
3. 在T型链接中, c_1 的分割与 c_2 的两个分割中距离最近的锻炼相连, 轮廓在两个方向上继续;
4. 令 s 为连接边缘链 c_1 与 c_2 的分割桥, 则 s 连接 c_2 起点的分割与 c_1 终点的分割;
5. 两个链接桥连接同一个边缘链上的两个端点;
6. 考虑一个链接桥 s , 由于它完全覆盖了边缘链 c_2 , 因此, 并没有前向链接。当 c_2 链接到另外一个边缘链 c_3 时, 连接 s 与 c_3 起点分割;

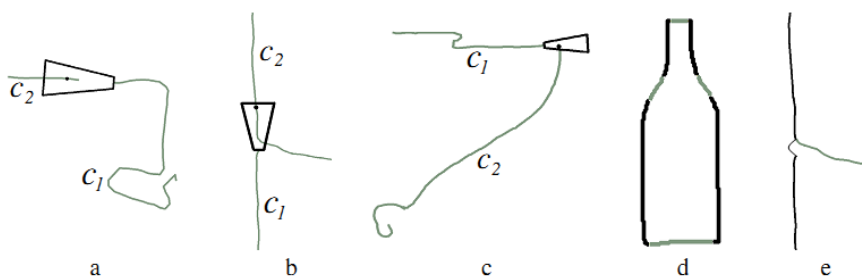


图 2.12: 边缘链之间的链接[11], 如图(a-c)。其中(a)为终点与终点连接, (b)为T形连接, (c)为切线不连续连接, (d)瓶子形状的边缘链的8个分割, (e)为连接b的分割桥, 标记为弧形。

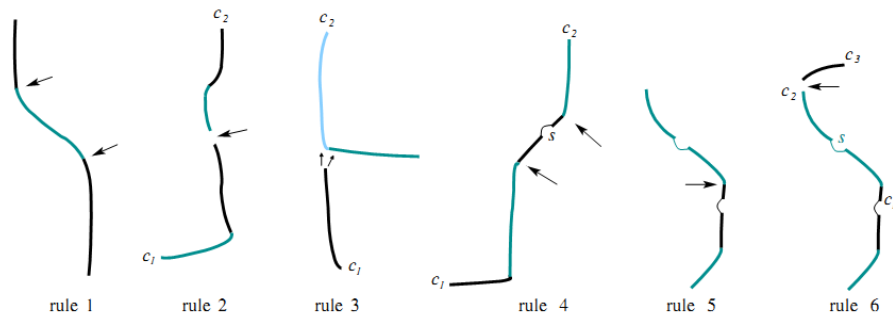


图 2.13: 创建轮廓分割的6个规则[11], 通过这些规则连接分割与分割桥 (标记为弧形), 其中规则2-6连接分割与边缘链 c_i .

k 个分割构成的边缘链称为 kAS , 当且仅当他们能够排序, 并且当 $i \in \{1, k-1\}$ 时, 在CSN (轮廓分割网络) 中的第 i -th个分割连接到第 $(i+1)$ -th个, 其中路径长度 k 称为度。例如: $k = 1$, 单独分割; $k = 2$, L型和T型分割; $k = 3$, C、F、Z型分割、3个分割的T型、三角形分割, 如图2.11, 2.14。随着 k 的增长, kAS 描述子的维度也在增长。

为了比较不同的 kAS 特征描述, 首先对 kAS 分割 $\{s_i\}_{i=1, \dots, k}$ 按照下面的规则进行排序, 使相似的 kAS 有相同的排序, 如图2.15:

1. 设所有分割的中点集合 $\{m_i = (x_i, y_i)\}_{i=1, \dots, k}$ 的中心为 \bar{m} , 选取距离 \bar{m} 最近的中心所在的分割作为第一个分割;
2. 剩余的分割按照其中点位置从左到有进行排序, 分别为 $2 - k$;

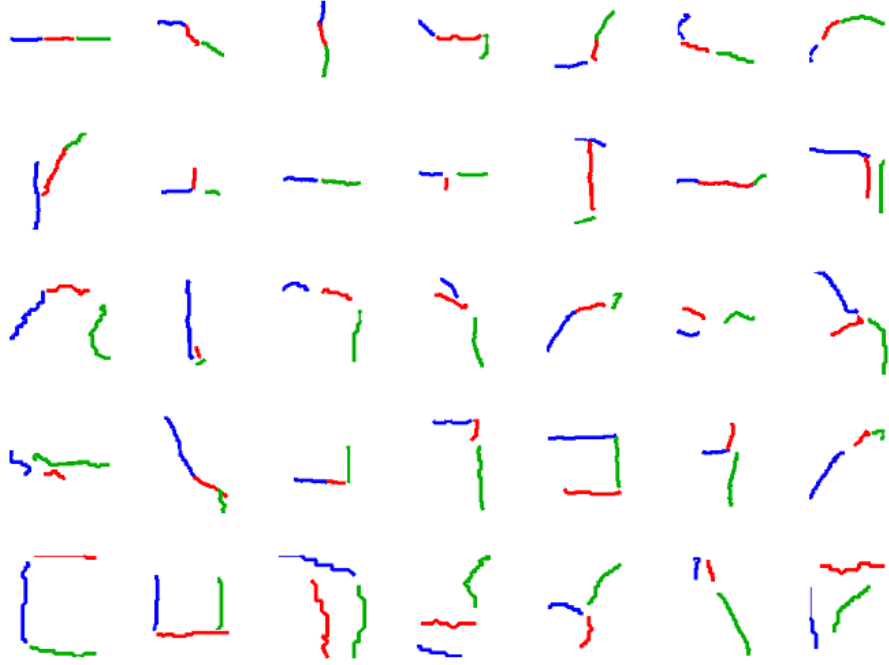


图 2.14: 特征袋中出现频率最高的35个3AS特征[11]。

3. 如果两个分割 s_i, s_j 的 x 坐标相似, 例如, $|x_i - x_j| \leq 0.2\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$, 则按照从上到下进行排序;

一旦排好序, 每个 k AS可以表示为一个分割列表 $P = (s_1, s_2, \dots, s_k)$ 。令向量 $\vec{r}_i = (r_i^x, r_i^y)$ 表示 s_1 的中点到 s_i 中点的向量; θ_i 和 $l_i = \|s_i\|$ 分别表示 s_i 的方向和长度, 则 P 的描述子可以表示成 $4k - 2$ 维向量:

$$\left(\frac{r_2^x}{N_d}, \frac{r_2^y}{N_d}, \dots, \frac{r_k^x}{N_d}, \frac{r_k^y}{N_d}, \theta_1, \dots, \theta_k, \frac{l_1}{N_d}, \dots, \frac{l_k}{N_d} \right) \quad (2.12)$$

其中 N_d 为两个中点之间的最长距离, 作为归一化因子, 使得描述子对尺度稳定。两个 k AS: P^a 和 P^b 之间的距离 $D(a, b)$ 为:

$$D(a, b) = \omega_r \sum_{i=2}^k \|\vec{r}_i^a - \vec{r}_i^b\| + \omega_\theta \sum_{i=1}^k D_\theta(\theta_i^a, \theta_i^b) + \sum_{i=1}^k \left| \log\left(\frac{l_i^a}{l_i^b}\right) \right| \quad (2.13)$$

其中第一项为分割相对位置的差别, $D_\theta \in [0, \pi/2]$ 为分割方向之间的距离; 最后一项比较分割长度的不同。

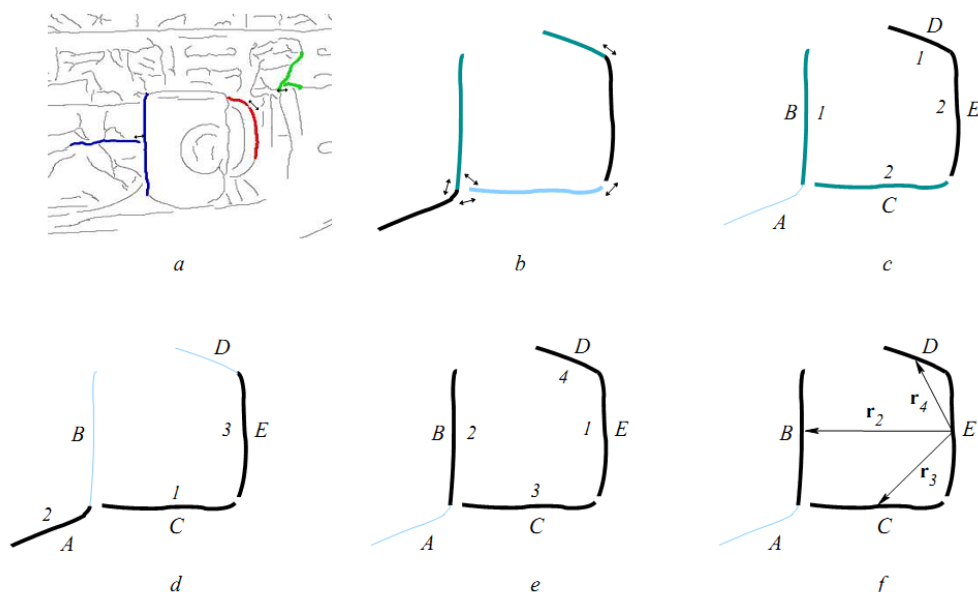


图 2.15: a), 示例图像, 3个2AS和CSN连接; b), 由5个分割和他们中间的连接组成的3个边缘链; c), 2个2AS (B,C) 和 (D, E), 每个分割的序号标记在其旁边; d), 3AS (C,A,E); e), 4AF (E,B,C,D); f), 在d)中的AS描述所涉及到的向量 \vec{r}_i [11]。

通过聚类将检测到的所有 k AS特征进行分类, 假设分为 M 类, 则创建包含 M 个元素的特征袋。在进行目标检测过程中, 对滑动窗口中的所有轮廓边缘计算特征袋中元素直方图来表示目标, 通过SVM进行训练得到期望的分类器。

2.3.2 基于轮廓框架的形状识别

在人类的视觉系统中, 目标轮廓边缘各部分的空间分布为目标的识别提供了大量信息。通常, 目标的某些特殊的组成部分就暗示了目标的类别, 并且目标各部分轮廓的空间分布具有很强的稳定性。依据上述理论依据, J.Shotton于2008年提出通过物体的框架组成来描述、识别目标[36, 37]。

如图2.16给出了基于轮廓组件的目标模型。其中, 在边界框(绿色)内, 轮廓框架 \bar{T} (黑色轮廓线)分布在目标中心(绿色十字)周围; 蓝色箭头表示偏离目标中心的向量 \bar{x}_f , 红色圆为空间不确定性, 半径为 σ ; 边界框记为:

$b = (b_{tl}, b_{br})$, 中心: $\mathbf{x} = \frac{1}{2}(\mathbf{b}_{tl} + \mathbf{b}_{br})$, 目标尺度: $s = \sqrt{area(b)}$ 。为了保持算法对目标尺度的稳定性, 将目标尺度归一化为 $s = 1$, 每一框架部分尺度归一化后, 记为: $F = (\bar{T}, \bar{x}_f, \sigma)$ 。

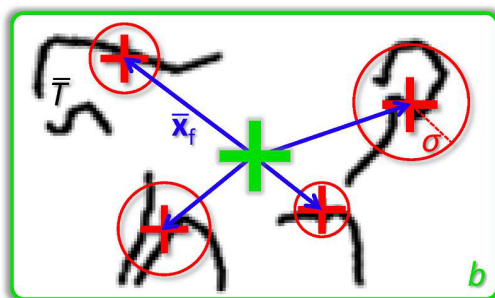


图 2.16: 在边界框 b 内 (绿色), 轮廓框架 \bar{T} (黑色轮廓线) 分布在目标中心 (绿色十字) 周围[36]。

由H.G.Barrow提出的倒角距离[4], 在度量两个轮廓之间的相似性时, 对于位置、尺度和旋转非常的稳定, 非常适合此处的轮廓框架之间匹配计算。在通常情况下, 倒角距离取两个边缘点集合作为输入: 模板 T 与边缘 E , 计算2D相对偏移的距离:

$$d_{cham}^{(T,E)}(x) = \frac{1}{|T|} \sum_{x_t \in T} \min_{x_e \in E} \|(x_t + x) - x_e\| \quad (2.14)$$

其中 $|T|$ 表示模板 T 中边缘点的个数, $\|\cdot\|$ 表示欧式距离。上述倒角距离公式给出了 T 上的点到 E 上最近的对应点的平均距离。式2.14中的距离通过距离变换给出 E 中最近邻点的距离:

$$DT_E(x) = \min_{x_e \in E} \|x - x_e\| \quad (2.15)$$

因此, 式2.14的最小值运算可以简化为下式:

$$d_{cham}^{(T,E)}(x) = \frac{1}{|T|} \sum_{x_t \in T} DT_E(x_t + x) \quad (2.16)$$

图像 I 的边缘 E 通过Canny边缘算子来提取, 为了简化由 E 中的丢失边缘而引起的问题, 设置距离阈值 τ 来增强距离函数的鲁棒性:

$$d_{cham_\tau}^{(T,E)}(x) = \frac{1}{|T|} \sum_{x_t \in T} \min(DT_E(x_t + x), \tau) \quad (2.17)$$

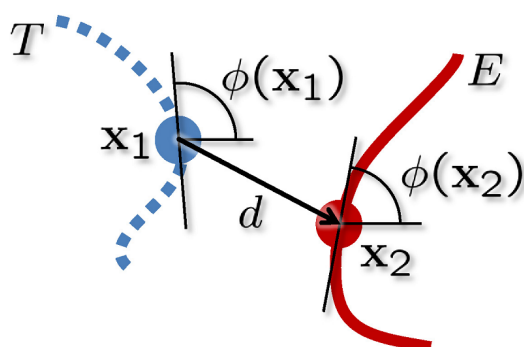


图 2.17: 方向倒角距离[36]: 模板 T 上的边缘点 x_1 的方向倒角距离为从 x_1 到边缘图 E 上的最近邻边缘点 x_2 的距离 d , 他们之间的边缘梯度差值为 $|\phi(x_1) - \phi(x_2)|$ 。

为了更进一步提高效果, 作者添加边缘的方向信息—边缘点梯度, 从而引进了方向倒角距离。方向倒角距离通过计算模板边缘点与他们最近邻边缘点之间方向的平均距离, 来衡量方向匹配的距离:

$$d_{orient}^{(T,E)}(x) = \frac{1}{|T|} \sum_{x_t \in T} |\phi(x_t) - \phi(ADT_E(x_t + x))| \quad (2.18)$$

其中 $\phi(\cdot)$ 为边缘点的方向, $|\phi_1 - \phi_2|$ 为角度的最小距离。辅角距离变换: $ADT_E(q) = \arg \min_{e \in E} \|q - e\|_2$ 。图2.17给出了详细的说明。

综上, 最终的距离度量公式为距离项与方向项的权重和:

$$d^{(T,E,\lambda)}(x) = d_{cham_r}^{(T,E)}(x) + \lambda_{orient}^{(T,E)}(x) \quad (2.19)$$

其中 λ 为方向参数。通过对每个特征学习各自的参数 λ , 可以控制通过边缘分割所不能实现的, 方向敏感性对最终结果的影响, 而这些是通过边缘分割所不能实现的。

通过将搜索框内的目标轮廓与模板按照方向倒角距离2.19进行匹配, 进行目标检测。这种方法对于目标的尺度、旋转和平移具有很强的鲁棒性, 但是, 边缘轮廓的噪声, 以及各组成部分的分布差异, 对于距离计算和目标的识别效果存在较大的影响。

2.4 小结

本章主要是对形状识别与匹配的基本概念做了描述, 并介绍了这一领域中

近年来较流行的几种算法的描述子，分析了每种算法的优势及不足。在基于密度的特征描述子中，SIFT算子具有非常强的适应能力，通常作为标准描述子；然而，在基于几何形状的特征描述子中，迄今为止我们尚未发现任何一种描述子得到了广大学者的认可，能够应用到绝大多数的场景中，必须依据不同的图像数据库和目标图像设计相应的描述子和算法。在本文接下来的章节中，将会介绍基于我们研究的图像数据，在前人的研究成果上发展的形状识别与匹配的算法。

第三章 基于谱图理论的形状识别

3.1 引言

谱图理论分析的基础是图的邻接矩阵，它是1973年由Fiedler[12]提出来，目前已经在图像分割领域得到广泛应用。Shi和Malik[35]将像素作为顶点，根据像素的亮度和空间位置确定链接像素点同边的权值，利用2-wayNcut的谱聚类方法迭代地进行图的分割，并且得到了满意的效果。近年来，越来越多的人开始关注，研究谱聚类算法在图像分割中的应用，并且针对其中的参数给出了自适应的算法[45]。各种实验都表明了谱聚类算法在图像分割中的有效性，同时也表明Laplacian矩阵的特征向量包含了大量的图像信息。

谱图理论在图像分割领域已经得到了广泛的应用，并且取得了非常好的效果[12, 35, 45]。从上述工作以及邻接矩阵本身的属性，我们可以发现邻接矩阵包含了图中每个点相对其他点的位置信息，充分表达了图中点的分布。并且根据矩阵的最大熵理论，矩阵最大的特征值对应的特征向量包含了最多的矩阵信息，在此充足的理论基础上，在一定的精度方位内，任意矩阵都可以用 K 个特征向量来表示原来的 N 维矩阵。从而所有的矩阵都可以通过上述方法进行降维，减少计算量。另外邻接矩阵的特征向量直方图对于图像的平移和旋转具有不变性，保证方法的鲁棒性。但是，应用谱图理论到图像识别中来，也存在着很多需要解决的问题：(1)，如何计算图中各个顶点之间的距离，从而保证整个矩阵对于图像尺度的稳定性；(2)，如何确定降维后特征空间的维度 K ；(3)，若特征降维取决于特征间隙，那么如何确定特征间隙；(4)，如何确定矩阵特征值的几何重数。针对上述问题，我们提出了自己的一些改进措施确定特征值的几何重数、特征间隙从而保证谱图理论在图像识别中的应用。

本章的内容组织如下：4.1节引言，介绍方法的主要思想，列出章节组织；4.2节简单回顾了一下谱图理论的基本知识[26]；4.3节着重介绍了我们提出的基于特征向量直方图的图像匹配算法；4.4节中我们给出实验结果及其分析；4.6节总结，讨论了本章方法的优点和不足，提出下一步的改进措施。

3.2 谱图理论回顾

3.2.1 图划分

给定一系列的样本点： x_1, \dots, x_n ，以及任意样本点 x_i 与 x_j 之间的相似性 s_{ij} ，聚类的初始目的就是希望将这些样本点分成几组，使得同组内的样本点相似，而不同组之间的样本点不相似。根据样本点之间的相似性，创建相似图 $G = (V, E)$ 来表示样本集，其中顶点集合 $V = \{v_1, \dots, v_n\}$ ，其中，顶点 v_i 表示样本点 x_i 。当顶点 v_i 与 v_j 对应的样本点 x_i 与 x_j 之间的相似性 s_{ij} 大于0，或者某个阈值时，则连接两个顶点，对应边的权重为 s_{ij} 。原始的聚类问题就转化为图的分割，使得不同组间的边的权重尽量小，而同组内的边的权重尽量大。在这一小节，我们将介绍一些图的基本概念，并讨论几种不同的图[26]。

给定一个无向图 $G = (V, E)$ ，其中顶点集合： $V = \{v_1, \dots, v_n\}$ ，边集： $E = \{e_{ij}\}_{i,j=1,\dots,n}$ 。假设图 G 为权重图，即：任意两个顶点 v_i 和 v_j 之间的边 e_{ij} 的权重 $w_{ij} \geq 0$ ，则图 G 的邻接权重矩阵为 $W = (w_{ij})_{i,j=1,\dots,n}$ ，其中，如果 $w_{ij} = 0$ 则顶点 v_i 与 v_j 不相连。因为图 G 为无向图，故其邻接权重矩阵为对称矩阵，即 $w_{ij} = w_{ji}$ 。顶点 $v_i \in V$ 的度定义为：

$$d_i = \sum_{j=1}^n w_{ij} \quad (3.1)$$

则图 G 的度量矩阵为：

$$D = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix} \quad (3.2)$$

给定顶点集的子集 $A \subset V$ ，则其补集为 $\bar{A} = V - A$ 。定义指示向量 $\bar{1}_A = (f_1, \dots, f_n)' \in \mathbb{R}^n$ ，其中，如果 $v_i \in A$ ，则 $f_i = 1$ ；否则 $f_i = 0$ 。则子集 $A \subset V$ 的大小为：

$$\begin{aligned} |A| &:= \# \text{ of vertices in } A \\ \text{vol}(A) &:= \sum_{i \in A} d_i \end{aligned} \quad (3.3)$$

其中 $|A|$ 通过顶点数量计算集合 A 的大小，而 $\text{vol}(A)$ 通过边的权重来衡量集合 A 的大小。

我们称图的子集 $A \subset V$ 是联通的，如果集合 A 内的任意两个顶点都能通过内部路径连接；而当子集 A 联通，且与补集 \bar{A} 之间任意两点都不相连，

则称 A 为 V 的联通分支。集合 A_1, \dots, A_k 构成图的分割, 当且仅当 $A_i \cap A_j = \emptyset$ 且 $A_1 \cup \dots \cup A_k = V$ 。

给定一个样本点集合 x_1, \dots, x_n , 及其点对之间的相似性 s_{ij} 或者距离 d_{ij} , 则可以将其转化为以下三种不同的图:

1. ε -邻域图, 连接所有距离小于 ε 的点对;
2. K -近邻图, 连接顶点 v_i 与 v_j , 其中 v_j 为与 v_i 距离最近的 K 个顶点;
3. 全连通图, 连接所有相似度 s_{ij} 大于0的点对, 并且每条边的权重为 s_{ij} ;

3.2.2 图的Laplacian矩阵及其属性

假设一个无向加权图 $G = (V, E)$, 其对称权重矩阵: $W = (w_{ij})$, 其中 $w_{ij} = w_{ji} \geq 0$ 表示连接顶点 v_i 与 v_j 的边 e_{ij} 的权重。那么该图的非归一化Laplacian矩阵定义为:

$$L = D - W \quad (3.4)$$

命题3.2.1 (L 的属性).

矩阵 L 满足下列属性:

1. 对于任意向量 $f \in \mathbb{R}^n$ 满足下式:

$$f'Lf = \frac{1}{2} \sum_{i,j=1}^n w_{ij}(f_i - f_j)^2 \quad (3.5)$$

2. L 为对称的半正定矩阵;
3. L 最小的特征值为0, 对应的特征向量为常数向量: $\bar{1}$;
4. L 有 n 个非负的实数特征值: $0 = \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$;

为归一化的图的Laplacian矩阵及其特征值、特征向量包含了图的很多属性。下述命题给出了在谱聚类中非常重要的一个应用:

命题3.2.2 (连同分支数).

假设图 G 为无向非负加权图, 则矩阵 L 的特征值为0的几何重数等于图 G 中连通分支 A_1, \dots, A_k 的数量, 并且这些分支的指示向量 $\bar{1}_{A_1}, \dots, \bar{1}_{A_k}$ 的生成集为特征值0的特征空间。

根据不同的特点，有两种归一化的Laplacian矩阵，他们彼此之间非常相似，其定义如下：

$$\begin{aligned} L_{sym} &:= D^{-\frac{1}{2}}LD^{-\frac{1}{2}} \\ L_{rw} &:= D^{-1}L = I - D^{-1}W \end{aligned} \quad (3.6)$$

这两个归一化矩阵具有下述属性：

命题3.2.3 (L_{sym} 和 L_{rw} 的属性).

归一化的Laplacian矩阵满足下列属性：

1. 对于任意的向量 $f \in \mathbb{R}^n$ 满足下式：

$$f' L_{sym} f = \frac{1}{2} \sum_{i,j=1}^n w_{ij} \left(\frac{f_i}{\sqrt{d_i}} - \frac{f_j}{\sqrt{d_j}} \right)^2 \quad (3.7)$$

2. λ 为 L_{rw} 的特征向量 v 对应的特征值，当且仅当， λ 为 L_{sym} 的特征向量 $v' = D^{\frac{1}{2}}v$ 对应的特征值；
3. λ 为 L_{rw} 的特征向量 v 对应的特征值，当且仅当， λ 和 v 是方程 $Lv = \lambda Dv$ 的解；
4. L_{rw} 和 L_{sym} 是半正定矩阵，并且有 n 个非负的实数特征值： $0 = \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$.

与未归一化的Laplacian矩阵相似， L_{rw} 和 L_{sym} 的特征值0的几何重数与矩阵的连通分支数相等。

3.3 基于特征向量直方图的形状匹配算法

在第二章的技术综述中，我们介绍了D.Knossow等人的基于Laplacian特征向量的图的模糊匹配算法，他将每个三维图用对应特征点构成的图的Laplacian矩阵的特征向量直方图来表示，并且通过直方图集合之间的直方图匹配来进行图像间的匹配。上述方法通过图像的全局特征来描述，但是，由于直接应用匈牙利算法进行直方图匹配，而不考虑由于图像尺度造成的特征值的几何重数，从而使得整个匹配过程有失欠缺。在下面介绍的我们提出的改进措施，从一定程度上解决了由于图像尺度造成的特征向量数目不同、特征几何重数不同等问题。

假设给定两幅图像 I_1 与 I_2 ，及其特征点集合： $V_1 = \{f_1^1, \dots, f_N^1\}$ 和 $V_2 = \{f_1^2, \dots, f_M^2\}$ ，由此可以得到图像对应的图 G_1 和 G_2 ，分别有 N 和 M 个顶点。根据方程 (3.1) - (3.4) 计算图 G_1 和 G_2 的Laplacian矩阵 L_1 和 L_2 ，由算法 (??) 得到Laplacian矩阵降维的特征空间： $U_1 = (\lambda_i^1, \mu_i^1)_{i \in \{0, \dots, n_1\}}$ 和 $U_2 = (\lambda_i^2, \mu_i^2)_{i \in \{0, \dots, n_2\}}$ ，其中 n_1, n_2 分别为降维后特征空间的维度。

通过比较特征值，进行特征向量的匹配，计算 U_1 与 U_2 的相似性，即 I_1 与 I_2 之间的相似性。但是，这种方法却忽略了下面三个问题：

1. 特征值可能存在几何重数，特征值并不能被有效的排序；
2. 特征向量并不是唯一定义的，例如： $\lambda v = Av$ 且 $\lambda(-v) = A(-v)$ ， v 与 $-v$ 均为 λ 对应的特征向量；
3. 当图的维度不同时，其对应的特征值和特征向量的数目，及其排序都将受到影响；

为了解决上述问题，我们提出了一些改进措施。假设图 G 对应的Laplacian矩阵的特征值为： $0 \leq \lambda_0 \leq \lambda_1 \leq \dots \leq \lambda_n$ 。为了确定每个特征值的几何重数，我们设定了经验值 ε ，当满足不等式 (3.8) 时，则称两个特征值相同；否则，不同。

$$|\lambda_{i+1} - \lambda_i| < \varepsilon \quad (3.8)$$

降维特征空间的维度取决于特征间隙，但是，特征间隙的选取依然是谱图理论中的一个开放性问题。在本文中，我们设定了另外一个经验值 η 来确定特征间隙，即：按照从大到小的顺序查找特征值中满足条件 $\lambda_{k+1} - \lambda_k > \eta$ 的 k 值， k 即为降维后的特征空间维度，原图 G 表示为特征空间 $U = \{v_1, \dots, v_k\}$ 。

考虑两幅不同尺度的图像对应的图 $G_1 \in \mathbb{R}^{N_1}$ 与 $G_2 \in \mathbb{R}^{N_2}$ 分别降维到特征空间 $U_1 \in \mathbb{R}^{N_1 \times n_1}$ 和 $U_2 \in \mathbb{R}^{N_2 \times n_2}$ ，其中 $N_1 \neq N_2, n_1 \neq n_2$ 。由于图像对应的特征空间维度不同，并且特征向量的维度一般都比较高，无法直接计算两个特征向量之间的距离，因此，我们将每个特征向量归一化为同维度（ D 维）的直方图。则原特征空间可以分别表示为 $H_1 = \{h_1^1, \dots, h_{n_1}^1\}$ ， $H_2 = \{h_1^2, \dots, h_{n_2}^2\}$ ，图像之间的距离即为对应直方图集合之间的距离： $dis(I_1, I_2) = dis(H_1, H_2)$ 。

为了计算两个特征向量直方图之间的距离，我们首先计算两个直方图集合之间的距离矩阵 $R = \{r_{ij}\}_{i=1, \dots, n_1, j=1, \dots, n_2}$ ，其中

$$r_{ij} = \min\{r(h_i^1, h_j^2), r(h_i^1, -h_j^2)\}, \forall i = 1, \dots, n_1, j = 1, \dots, n_2 \quad (3.9)$$

其中 $r(\cdot)$ 为直方图的距离函数，例如内积、欧式距离等，消除了特征向量符号的影响。一旦得到距离矩阵 R ，下一步最重要的就是特征向量匹配，计算图像之间的距离，下面我们将重点介绍如何进行直方图匹配和计算 $dis(H_1, H_2)$ 。

事实上，我们并不需要计算距离矩阵 R 中的每个元素，而只需要计算矩阵对角线子矩阵上的元素。对于不同尺度的图，特征值的几何重数也各不相同。因此，我们根据方程(3.8)，将每个图的特征向量集合分成若干组，且每组对应的特征值为组内特征值的平均值。

下面计算图之间对应特征向量组内直方图间的距离 r_{ij} ，并且在计算距离矩阵 R 中，我们遵循下述条件：

1. 特征向量直方图匹配的数量要小于等于图的最小维度， $\#ofmatch \leq \min(n_1, n_2)$ ，即：我们将低维图的特征向量直方图全部匹配；
2. 特征向量直方图之间进行一一对应；
3. 尽量多的特征向量直方图能够进行匹配；

每个特征向量组允许和其临近范围内的对应组进行匹配，例如图3.1中，第一组特征向量可以和邻域为2的特征向量组进行匹配，选取最佳的匹配。实验中，我们通过匈牙利算法对组间的特征向量进行匹配，取组间距离 $score_B$ ：最小的作为最佳匹配组，其中：

$$score_{B_k} = -\frac{M_k}{N_k} + \frac{d_k \times N_k}{M_k} \quad (3.10)$$

其中： M_k 为组内匹配的特征向量直方图数量； N_k 对应两个组的最小维度，即： $\min\{B_{k_i}, B_{k'_i}\}$ ； d_k 为组内特征向量直方图匹配的距离和。则，两幅图像之间的距离为所有小组匹配距离的和：

$$score = \sum_k score_{B_k} \quad (3.11)$$

当图像相似时，对应图像之间的分值越小；否则，分值越大。

3.4 实验结果及分析

这一节将对我们提出的算法进行验证及分析，首先介绍实验中算法参数设置，然后将展示实验结果并对其进行分析。我们从数据库MPEG-7 CE-Shape-1中随机选取图像进行试验，其中的图像均为二值图像。

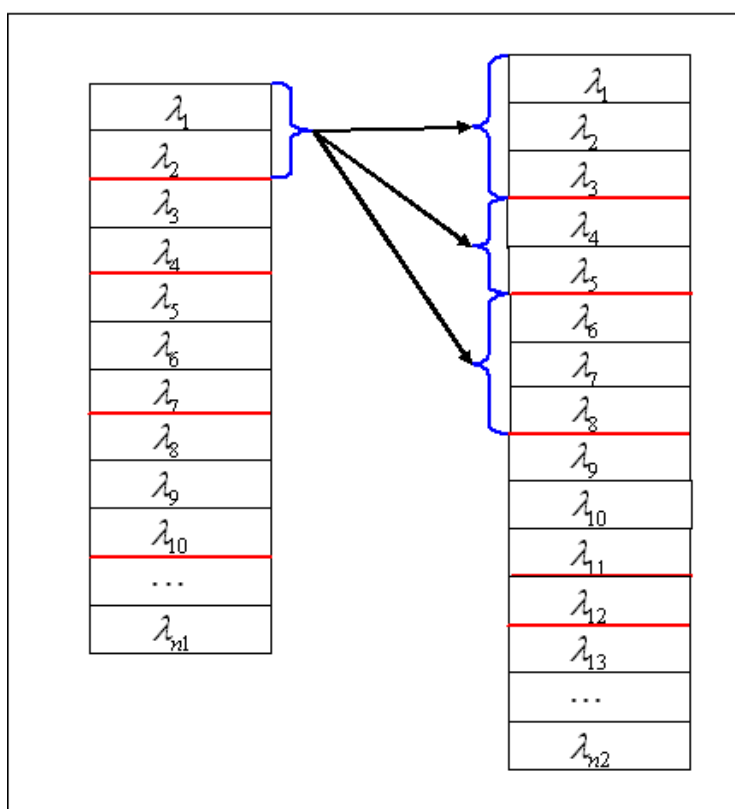


图 3.1: 特征向量的匹配: 低维图的特征向量组与邻域内的高维图的特征相量组进行无序匹配, 查找最佳匹配

在实验中, 为了保证得到单像素的边缘轮廓信息, 我们采用Canny算子[9]检测图像的边缘; 并且将每个边缘点作为顶点, 两两连接, 创建图 $G = (V, E)$ 来表示原始图像, 从而得到图的权重矩阵 W , 其中每条边的权重用下面的方程计算:

$$\omega_{ij} = \exp\left(-\frac{d_{ij}^2}{\sigma^2}\right) \quad (3.12)$$

其中 d_{ij} 为两点之间的欧式距离, σ 为平滑参数, 设为30。根据方程 (3.1-3.4) 可以得到图的Laplacian矩阵 L , 其特征值的几何重数阈值 $\epsilon = 0.01$, 特征间隙阈值 $\eta = 0.01$ 。

图3.2给出两幅形状略有不同的人首马身图, 图3.3为我们用上述方法得到的对应图的Laplacian矩阵的特征值曲线, 从图中可以看出, 其中大部分特征值为1。

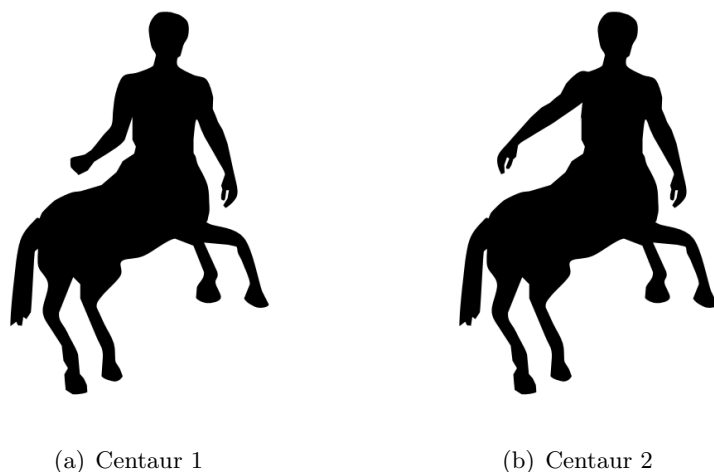


图 3.2: 不同形状的两幅人首马身原始图。

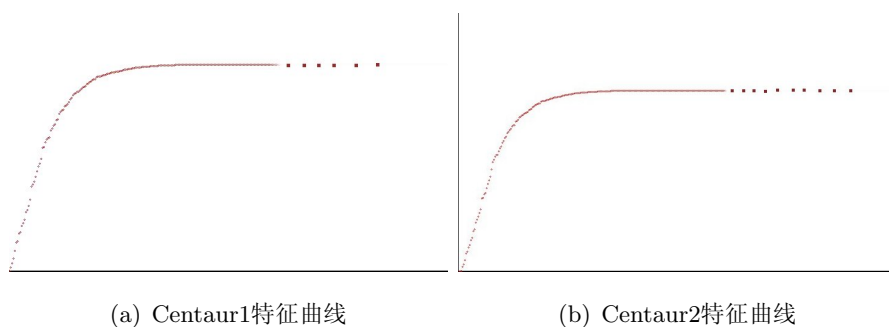


图 3.3: 图3.2对应的Laplacian矩阵的特征值曲线，为了图像的简洁，其中大部分特征值为1曲线被省略掉了。

通过特征检测，从图3.5(a)中选取58个特征向量，而图3.2(b)选取48个特征向量。在形状匹配过程中，共有23对特征向量直方图可以匹配。由于空间所限，这里我们只列出两幅图像的前9个最小特征值对应的特征向量直方图，其匹配结果如下： $(Centaur1 : Centaur2) : (1 : 1), (2 : 2), (3 : 3), (4 : 4), (5 : 5), (6 : 6), (7 : 7), (9 : 8)$ 。由方程 (3.10), (3.11) 可知，在通过特征向量直方图排序匹配后，两幅图像之间的距离为：0.116；相对而言，未经过排序匹配的距离为0.327。实验表明，经过特征向量排序匹配后，同类图像之间的距离更小，更具有区分性。

图3.5给出了两个不同目标—蜥蜴、人首马身—的原始图像，通过特征空隙

检测，从图3.5(a)中选取了48个特征向量直方图，从图3.5(b)中选取了17个。但是，通过直方图匹配，只有一对特征向量直方图匹配成功：(4, 1)，图像间距离为0.512；而没有排序匹配的图像距离为0.562。实验表明，即使经过特征向量排序匹配后，不同类的图像间依然保持了较大的距离。

上述两个例子，说明我们的方法有效的增强了图像的区分度，即：同类图像之间的距离变小，与此同时，保持不同图像之间拥有较大的距离。

3.5 总结

在这一章中，我们首先介绍了谱图理论的相关概念以及谱聚类算法，然后，重点介绍了基于谱图理论的形状匹配算法。这种算法的核心和主要思想是利用谱图理论中的最大熵理论，在保留尽量多的信息的同时，降低图像维度，通过特征匹配实现图像识别的目标。通过实验可以看出这种算法有效增强了图像之间的区分度，不但，降低了同类图像之间的距离，同时有效的保持了不同类图像之间的高距离。但是，我们也可以看到，由于谱聚类算法本身的特点，我们的方法也带有一些缺点，例如：求解矩阵特征的复杂性随着特征点数量的增加而大幅度增长；在计算图的权重矩阵时，平滑参数 σ 对图像尺度不够稳定，从一定程度上会影响后面的直方图匹配，从而影响图像识别。在以后的工作中，我们可以通过预处理对图像尺度进行归一化。

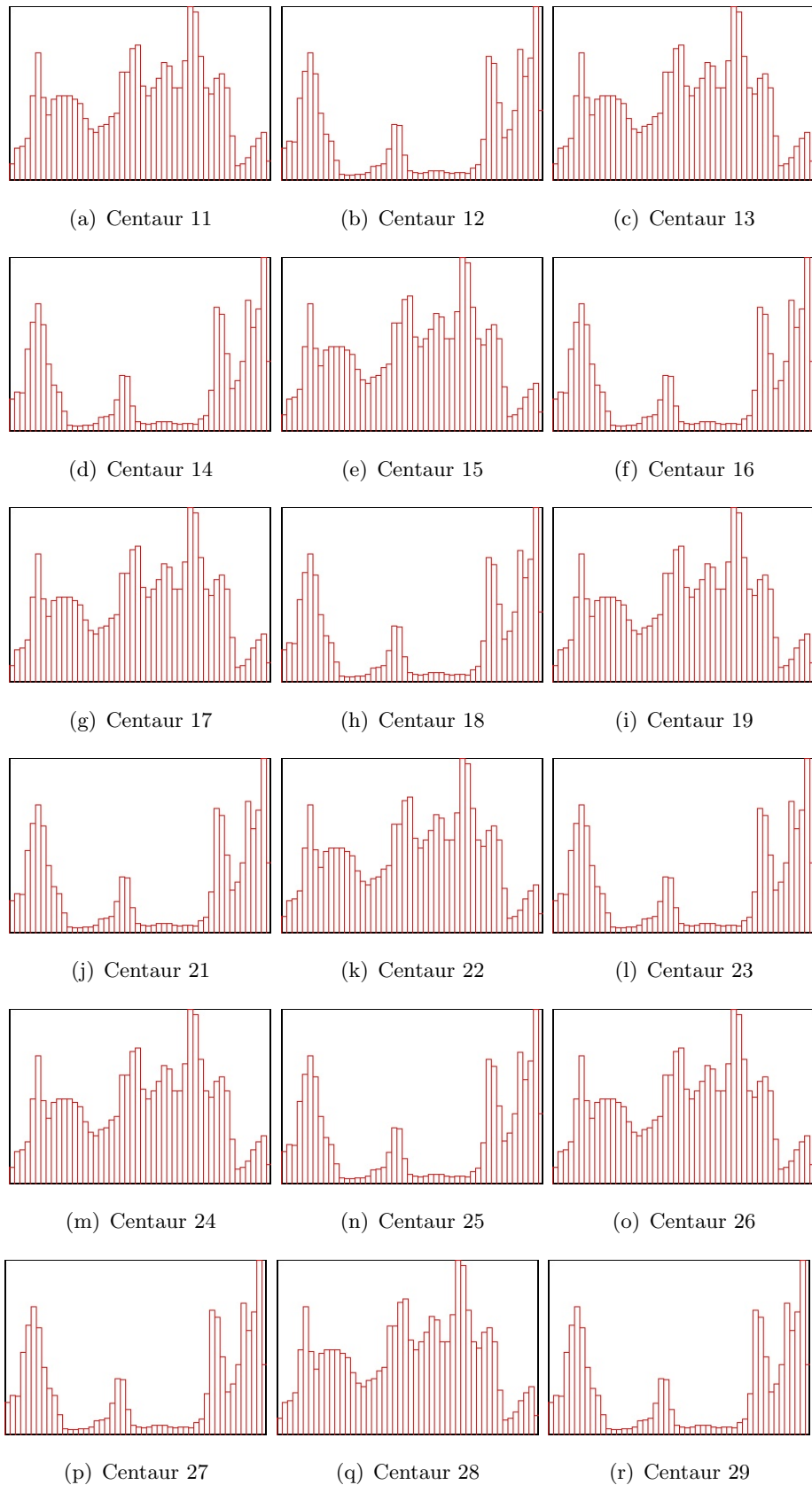


图 3.4: 图3.2对应的前9个特征向量直方图, 通过直方图匹配, 两个特征相量集合的匹配结果为: (11 : 21), (12 : 22), (13 : 23), (14 : 24), (15 : 25), (16 : 26), (17 : 27), (19 : 28), 从而可得到下述结论: 两幅图像为同类图像。

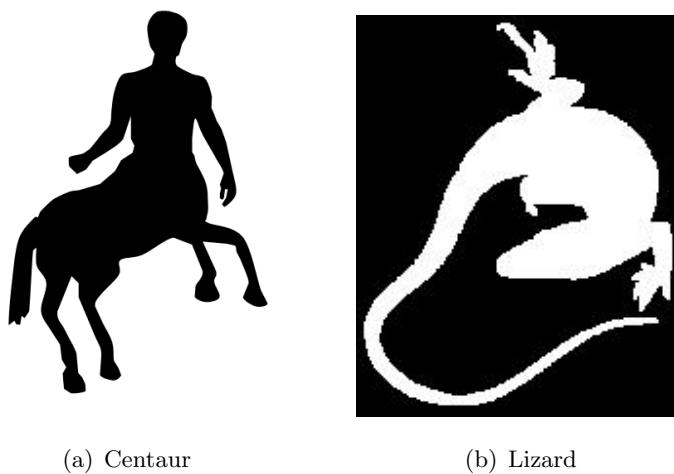


图 3.5: 不同目标的两幅原始图像: 蜥蜴与人首马身。

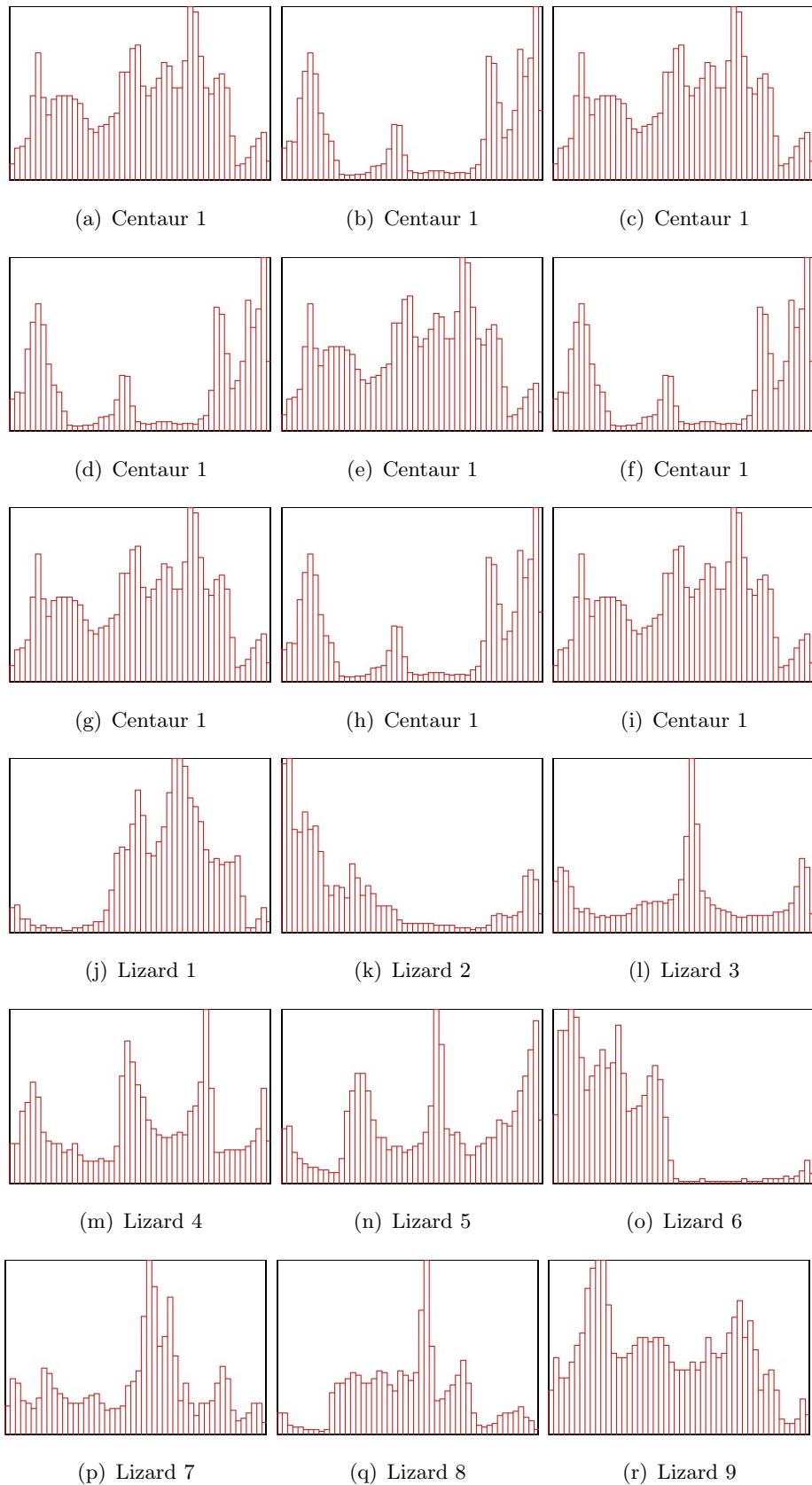


图 3.6: 图3.5对应的前9个特征向量直方图, 通过直方图匹配, 只有一对特征向量直方图能够匹配(14 : 21), 从而可得到下述结论: 两幅图像不是同类图像。

第四章 基于夹角直方图的形状匹配

4.1 引言

通过将文本处理中的上下文特征信息统计的方法引入到形状识别中来，形成了现在在形状识别与匹配中被广泛应用的形状上下文描述子。本章的算法是对原始的形状上下文描述子的一个改进。S.Belongie[6-8]等人的形状上下文描述子通过提取大量的边缘点来描述原始的目标，虽然，通过选取足够的点，能够更加精确的表达图像的原有信息，但是，这也导致了在后续的图像匹配中大量的计算，影响了算法的效率。为此，我们在其工作的基础之上做了一些改进，通过选取局部曲率最大值点作为特征点从而降低图像匹配时，从而消除大量点匹配而引起的高级算量。同时，我们选取点对之间连线向量的夹角直方图来表示点，保证描述子对于图像旋转的稳定性。

本章的内容组织如下：3.1节引言，列出章节组织；3.2节详细介绍了我们改进后的基于夹角直方图描述子的图像匹配算法，同时给出具体的算法过程；3.3节给出了实验的结果，并讨论分析了实验结果；3.4节总结，讨论了本章提出的方法的优点以及不足，并以此为基础分析了下一步工作如果展开。

4.2 基于夹角直方图的形状上下文的形状匹配

目前在图像检索、目标识别与匹配的研究中，基于形状的算法受到越来越多的重视，同时越来越多的形状描述子被提出，例如：Hough变换[28]，矩表示[13]，傅立叶描述[5]等。但是，大部分形状匹配算法都严格限制在形状轮廓或者闭合曲线上，对于存在内轮廓或者内部边缘的形状，图像匹配的效果并不十分理想，或者由于条件限制不能被应用。形状上下文算法可以理解为点集匹配算法，它对于尺度和平移具有不变性，并且对于旋转和扭曲具有很强的鲁棒性。在2.4节我们已经详细介绍了形状上下文描述子，及其匹配算法[6-8]。在这一节，我们将重点介绍基于夹角直方图的形状上下文特征的图像匹配算法。

目前，上形状上下文描述子已经被广大学者接受，并且与其他方法结合扩展出了很多有效的形状匹配算法。但是，每个点的形状上下文特征都是形状的

全局统计特性，需要建立在大量样本点的基础上，而图像的相似性是通过点集匹配得到的，因此，后续的形状匹配的计算量很大，算法的时间复杂度过高。

根据曲线逼近理论，在有限的精度范围内，任意曲线可以用有限条直线段来逼近。依据上述理论，我们选取轮廓上的局部曲率最大值点作为特征点，通过连接相邻点来近似原始形状轮廓。如图4.1，绿色的圆点为我们检测到的局部曲率最大值点，红色直线连接相邻的两个局部曲率最大值点。从图中可以看出，从某种程度上来说，局部曲率最大值点的连线能够表示原来图像，因此，我们选取局部最大曲率角点作为特征点。

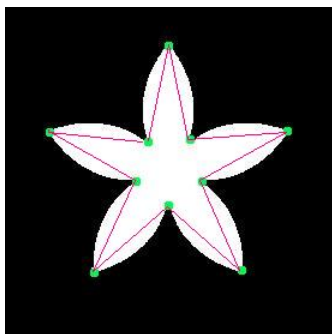


图 4.1: 图形的局部曲率最大值点（标记为绿色），以及轮廓拟合直线（标记为红色）。

如图4.2中，两幅锤子和苹果图像的特征点（绿色圆点）数量和位置不同，但是，同类目标的特征点数基本相同。由此，很容易想到，某一形状的特征点数量必然是在一定范围内。例如，图4.2(c)中包含5个特征点，可以猜想其他同类图像的特征点数量应该非常接近5；如果一幅图像有100个特征点，那么它不可能与图4.2(c)是同类图像。因此，我们首先通过图像的特征点数目对图像库中的带匹配图像进行筛选，得到目标图像的初步候选集。

显然，仅仅依靠图像提取的特征点数量来进行形状识别与匹配是远远不够的，因此，我们创建了下面的特征描述子。假设图像中检测到 N 个边缘点 (x_{ei}, y_{ei}) ，为了保证描述子对目标平移的稳定性，我们按照式4.1计算轮廓的

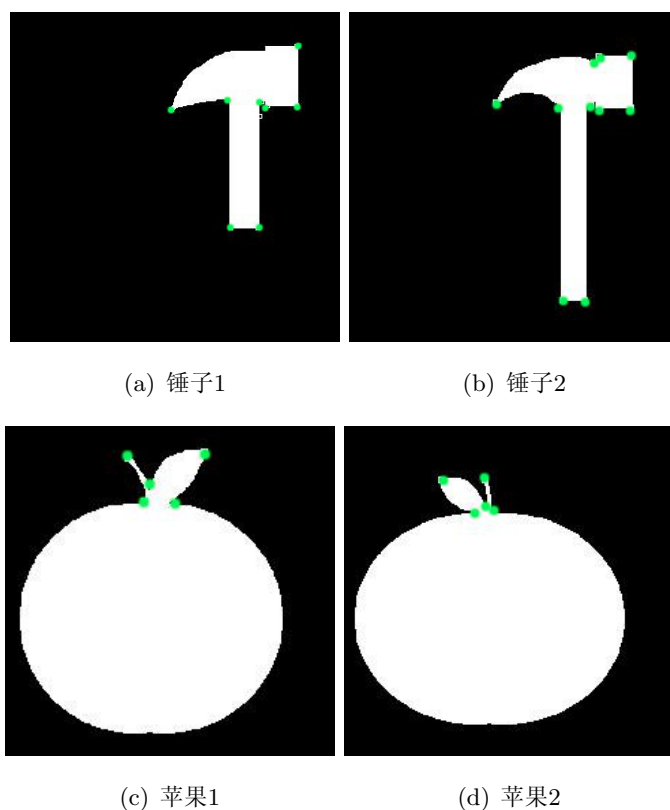


图 4.2: 锤子和苹果图像在不同的位置, 具有不同数量的曲率特征点。

重心 (\bar{x}, \bar{y}) , 作为参照点:

$$\begin{aligned}\bar{x} &= \frac{1}{N} \sum_{ei=1}^N x_{ei} \\ \bar{y} &= \frac{1}{N} \sum_{ei=1}^N y_{ei}\end{aligned}\tag{4.1}$$

通过计算边缘点的梯度, 假设检测到 K 个局部曲率最大值点 $\{(x_i, y_i)\}_{i=1, \dots, K}$, 连接轮廓的重心与特征点, 得到向量 $\vec{v}_i = (x_i - \bar{x}, y_i - \bar{y})$, 如图4.3, 其中每个特征点对应一个向量。则所有特征点对应向量集合: $V = \{\vec{v}_1, \dots, \vec{v}_K\}$ 。为了更好的描述每个特征点, 及其相对其他特征点的位置信息和对目标旋转的稳定性, 计算特征点 P_i 的向量相对其他特征点向量之间的夹角来表示点 P_i 。首先计算向量集合 V 中任意两个向量之间的夹角:

$$\theta_{ij} = \arccos\left(\frac{\langle \vec{v}_i, \vec{v}_j \rangle}{\|\vec{v}_i\| \|\vec{v}_j\|}\right)\tag{4.2}$$

其中 $\langle \cdot \rangle$ 表示两个向量的内积, $i, j = 1, \dots, K$, 得到图像的夹角矩阵 $\Theta = (\theta_{ij})_{i,j=1,\dots,K}$, 其中每行表示一个特征点。由形状上下文特征匹配算法可知, 图像之间的匹配可以转化为两幅图像中的特征点集合之间的匹配。由于每幅图像中的特征点数量不同, 因此, 我们将每个特征点对应的行向量表示为统计直方图, 从而每幅图像表示为 (h_1, \dots, h_K) 。

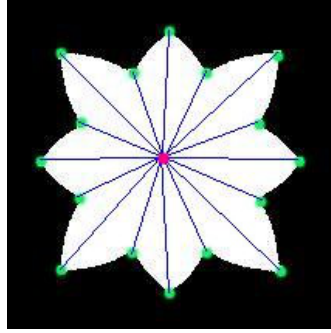


图 4.3: 图中绿色点为角点, 中心的红点为轮廓重心, 重心与角点之间的蓝色线为对应角点的表示向量。

基于S.Belongie等人的形状上下文算子的形状匹配算法[6-8], 通过两个形状的特征点集合之间的最佳匹配, 计算图像之间的相似性。假设两个形状 S_1 和 S_2 , 分别有 K_1 和 K_2 个角点, 表示为 $S_1 = \{h_1^1, \dots, h_{K_1}^1\}$ 和 $S_2 = \{h_1^2, \dots, h_{K_2}^2\}$, 其中 h_i^j 为形状 j 对应的第 i 个特征点的角度直方图。假设 S_1 上的角点 p_i 与 S_2 的角点 q_j 对应, 则两点之间的距离可以通过 χ^2 统计计算:

$$C_{ij} = C(p_i, q_j) = \frac{1}{2} \sum_{k=1}^M \frac{[h_i^1(k) - h_j^2(k)]^2}{h_i^1(k) + h_j^2(k)} \quad (4.3)$$

其中 M 为角度直方图的维度。角点匹配的最优解不仅使得形状 S_1 的角点能够在 S_2 的角点找到最佳匹配, 同时 S_2 的角点也能够能够在 S_1 的角点中找到最佳匹配。因此, 我们引入下面的距离公式:

$$D(S_1, S_2) = \frac{1}{K_1} \sum_{p \in S_1} \arg \min_{q \in S_2} C(p, q) + \frac{1}{K_2} \sum_{q \in S_2} \arg \min_{p \in S_1} C(p, q) \quad (4.4)$$

其中 p, q 分别代表 S_1 和 S_2 得角点, K_1, K_2 为其对应角点的数量。式4.4的第一项表示形状 S_1 的角点的最佳匹配距离, 第二项表示形状 S_2 的角点最佳匹配距

离。通过这两项多对一的点匹配，两幅图像之间的距离为两幅图像各自最佳匹配的距离的平均值之和。

综上所述我们给出整个算法如下算法1：

Algorithm 1 基于夹角直方图的形状上下文的形状匹配算法.

Input: 输入目标图像 I_o 和测试图像 $I_t = \{I_i\}_{i=1,\dots,N}$

- 1: 检测图像的外部轮廓点 (x_{ei}, y_{ei}) ，及其重心 (\bar{x}, \bar{y}) ，并计算目标形状的面积和周长比： $Ratio_1$ ；
- 2: 设定 $Ratio_1$ 的邻域，筛选测试图像，得到初步的图像候选集： I_t^1 ；
- 3: 计算轮廓点的曲率，选取局部曲率最大值点作为特征点： $(x_i, y_i)_{i=1,\dots,K}$ ；
- 4: 用连接特征点与轮廓重心的向量 $\vec{v}_i_{i=1,\dots,K}$ 表示特征点，并计算向量之间的夹角矩阵： $\Theta = (\theta)_{ij}$ ，其中 $\theta_{ij} = \arccos(\frac{\langle \vec{v}_i, \vec{v}_j \rangle}{\|\vec{v}_i\| \|\vec{v}_j\|})$ ；
- 5: 计算夹角矩阵 Θ 每行的夹角直方图 $\{h_i\}_{i=1,\dots,K}$ ；
- 6: 通过图像之间的点集的夹角直方图匹配计算图像之间的相似性： $\vec{S} = \{s_j\}$ ；

Output:

目标图像的前10个最佳匹配图像。

4.3 实验结果及分析

这一节将对我们上面提出的改进的算法进行验证和分析，实验中我们选取众所周知的KIMIA99[34]图像库和MPEG-7CE-Shape-1图像库，其中KIMIA99图像库包含99幅图像，分为9组，每组11幅图像；MPEG-7CE-Shape-1图像库被广泛应用在检测形状描述子及其匹配算法的有效性，包含1400张图片，共70类，每类20幅。

为了更加准确的表示图像，我们首先对图像进行预处理，如图4.4中的两幅图片，他们属于同一类，但是，图4.4(a)的轮廓内部包含大量细节化的边缘信息，相对于图4.4(b)，边缘信息更加简单。如果直接利用这些边缘的形状上下文描述子来表达图像，将影响最终的匹配效果。因此，我们首先去除轮廓内部的噪声边缘信息，得到如图4.5的填充图。同时计算每个图像中目标的面积和轮廓周长，计算其面积与周长比。由于同一类目标图像，虽然会出现旋转、平移和尺度变化，但是，其相对的面积与周长比是稳定不变的，并且在一定范围内

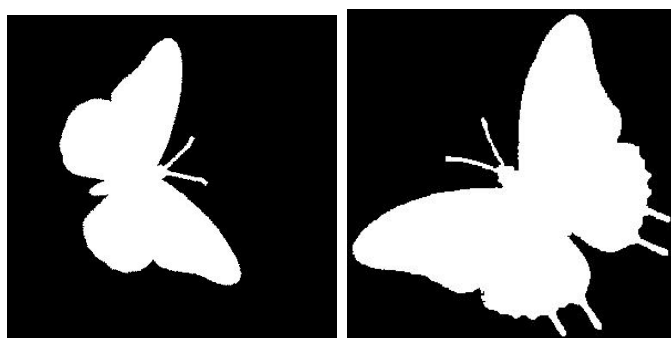
浮动，通过设定阈值，对图像库中的图像进行初步筛选，缩小后续的精确定匹配候选集。



(a) 图一

(b) 图二

图 4.4: 同类图像的两张图片，相对与图二，图一轮廓内包含大量的噪声—内部边缘。



(a) 图一

(b) 图二

图 4.5: 图4.4填充后的对应图像。

实验中，我们设置角点数量邻域为10来进行预处理选取候选集，当候选图像的角点数量与目标图像的角点数量的差值超出范围 $[-10, 10]$ ，则从候选集中删除。任意两个角点向量之间的夹角属于 $[0, \pi]$ ，并且量化为12份，如图4.6。图4.7给出了本章算法在MPEG-7CE-Shape-1图像库上的匹配结果，其中第一列为目标图像，后面的10列为我们的方法在图像库中检索到最佳匹配图像。实验中，我们采用下面的公式4.5计算识别的召回率来衡量算法的有效性。将数据库中的每一幅图像作为模板，通过和数据库中的所有图像进行匹配，选取前10幅

最佳匹配结果作为最终结果。实验结果显示，该方法的匹配召回率为38%：

$$ratio = \frac{\sum_{k=1}^N N_k}{10 \times N} \quad (4.5)$$

其中 N 为图像数据库中的图片数量， N_k 为每幅图像的检索结果中，前10幅最佳匹配中正确匹配的图像数目。从图中可以看出，并不是每一类都能够得到很好的匹配，例如第四行与第五行中的结果，虽然图像之间有一定能够的相似性。仔细分析可以发现，我们的方法忽略的圆的存在，造成误检，这是因为，我们选择局部曲率最大值点作为特征点，而这种特征的缺点就是不能检测到圆上的点，从而在整体描述中丢失了圆弧信息。

图4.8给出了在KIMIA99上的实验结果，其中第一列为目标图像，后面给出了在图像库中前11个最佳匹配结果。

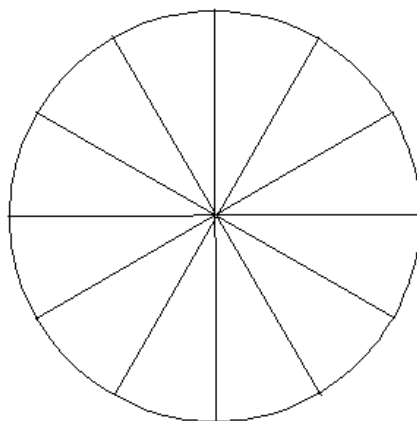


图 4.6: 计算夹角直方图形状上下文描述子时，角度量化图表。

上述实验基于Ubuntu9.10，320G硬盘和4G的显卡，在MPEG-7 CE-Shape-1图像库的运行时间大约14分钟，与Felzenszwalb 136小时[10]、Scott 122小时[33]，相比速度更快。上述实验结果表明我们提出的方法基本上能够实时的完成图像之间的正确匹配，证实了本章算法的有效性。

4.4 总结

在这一章中，我们重点介绍了我们改进后的基于夹角直方图的形状上下文

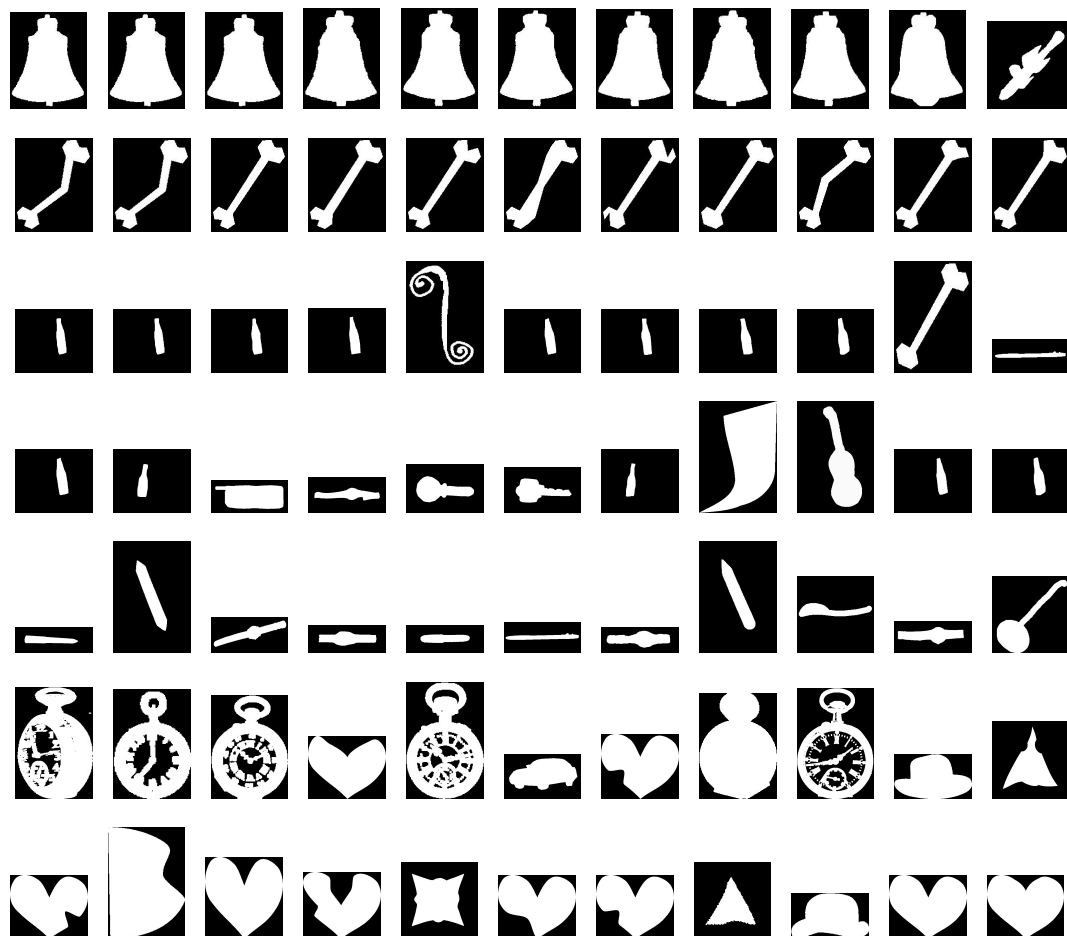


图 4.7: 在MPEG-7 CE-Shape-1图像库上的匹配结果。第一列为目标图像，后面的10个为匹配到的结果图像。

特征匹配算法。其核心思想是计算每个点相对其他点的分布信息来描述对应点，通过图像之间点的最佳匹配来计算图像之间的相似性。为了减少点集匹配所消耗的时间，我们提取图像边缘上的局部曲率最大值点来替代原来的随机点，进而用相邻两点之间的连线构成的形状近似描述图像。实验证明这种表示方法的能够实现图像之间的实时匹配，由于我们采用曲率点作为特征点，且用角度信息来描述每个点，因此，整个描述子对于图像的旋转比较稳定；通过选取轮廓点的重心作为目标参照点，确定每个图像的坐标系，保证了图像表示对于目标平移的稳定性。但是，通过对实验结果的分析，上述方法依然存在一些需要解决的问题：

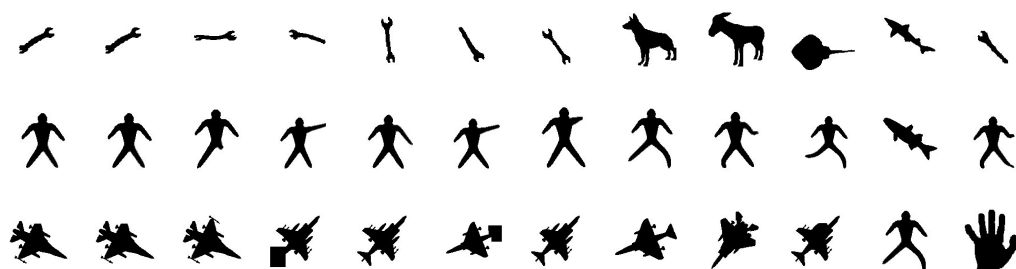


图 4.8: 在图像库KIMIA[18]的匹配结果。第一列为目标图像, 后面为其前11个最佳匹配结果。

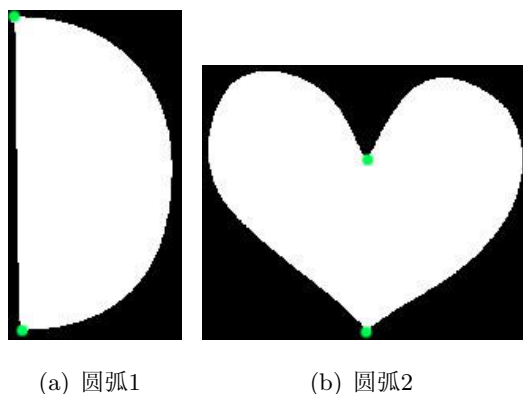


图 4.9: 由于包含大量的圆弧和直线, 他们检测到同样数量角点, 并且相对位置相同, 特征描述完全相同, 但是他们完全是不同的两个形状。

1. 上述描述子仅仅包含了特征点的相对角度分布信息, 却没有边缘信息, 这就限制了其应用到包含大量的平滑弧的形状上, 如图4.9中, 两幅图像中包含大量的弧形边缘, 选取局部曲率最大值点作为特征点, 其数目和角度分布完全相同, 由于丢失了边缘的形状信息, 他们被误判为同一目标; 在极端情况下, 当图像目标为圆形时, 其边缘将不存在局部曲率最大值点, 整幅图像将丢失所有信息。因此在以后的工作中, 我们将添加相邻角点之间的边缘信息, 例如, 长度, 平均曲率等, 用尽量少的特征点表达尽量多的信息。
2. 上述描述子对边缘随机噪声非常的敏感, 如图4.10, 虽然, 两幅图像属于同一目标, 但是, 由于出现过多的边缘随机噪声, 使得我们最终检测到的角点数量远远超出了目标图像; 在未来的工作中, 我们将尝试使用平

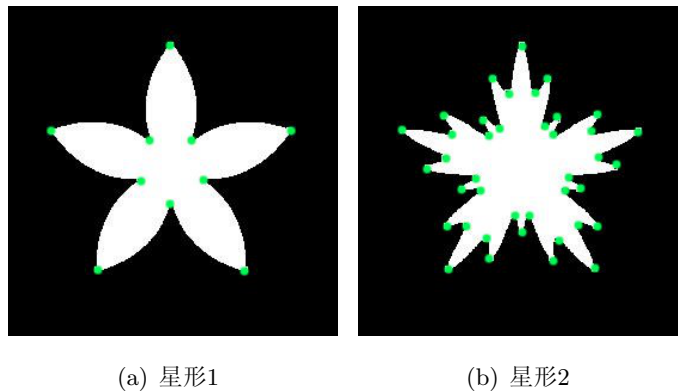


图 4.10: 属于同一目标的两个形状，但是，由于毛刺的存在，他们的角点数量相差很大，在预处理中被删除。

滑算子，或者膨胀、腐蚀等算法消除噪声对目标描述的影响。

3. 由于我们选取的每幅图像的特征点的数目不确定，我们采用直方图这一统计描述子来表示特征点。但是，事实上，当目标轮廓包含大量平滑的弧形时，特征点的数目将大量减少。在极端情况下，如图4.9，圆、半圆等图形上，特征点只有1 - 2个，甚至没有，采用直方图这种方法来描述并不恰当。

在未来工作中我们将添加更多的边缘信息，用高维特征来描述每个特征点，从而更精确的表达图像。

第五章 结束语

本章我们将对本文的主要工作及其主要贡献进行总结，并对下一步的工作进行展望。

5.1 本文主要工作

本文在对基于形状的图像匹配的各种方法的分析和总结的基础上，结合不同方法的优点，以二值形状数据库MPEG-7CE-Shape-1中的图像作为实验数据，围绕快速的形状识别与匹配这一主题进行了研究。本文的主要工作及贡献如下：

1. 发展了基于特征向量直方图的形状匹配算法

本文提出了一种基于谱聚类的理论基础上，并集合图匹配理论进行形状匹配的算法。在这种算法中，我们用目标形状的轮廓上的点建立图来表示原图像，并根据谱图理论计算其拉普拉斯矩阵来表示目标轮廓；根据最大熵理论，矩阵的最大特征值对应特征向量包含最多的信息量，我们选取矩阵的前 K 个最大的特征值对应的特征向量来表示目标；由于每个图的维度不同，将每个特征向量映射为同样维度的直方图，用直方图集合来表示原图，最终通过直方图匹配实现图像直接的匹配。其中，我们改进了直方图匹配原则，通过排序后特征向量按照特征值在一定的邻域内进行小组的无序匹配，而不再按照特征大小进行匹配，这种方法的优点是对于图的维度具有很强的鲁棒性，并且使得类内图像间的距离更小，同时保持了类间图像间的最大距离。与此同时，此方法对于图像的尺度、平移和旋转具有很强的鲁棒性。

2. 提出了基于夹角直方图的形状匹配算法

本文提出了一种基于S.Belongie等人的形状上下文特征，提出了基于特征点的夹角直方图的形状匹配算法。在这种算法中，我们计算图像边缘上任何一点的曲率，通过选取局部最大曲率点作为我们的特征点。在一定的相似度范围内，任何一条曲线都可以表示为有线条直线段的组合。基

于上述理论，我们连接相邻的特征点来近似表示原图像；并且利用形状上下文特征来表示每个点，通过图之间的点匹配计算图像间的相似性，其中我们用对应点的夹角直方图表示每个点。这种方法的优点是对图像的尺度具有很强的鲁棒性，相对于原始的选取大量的边缘点计算形状上下文特征，进行形状识别与匹配的运算速度更快。

5.2 未来工作展望

本文提出的两种形状匹配算法，都展示出了很好的性能或者巨大的潜力，也为今后的工作打下了良好的基础。对于我们未来的工作可以从如下方面展开：

1. 基于特征向量直方图的形状匹配算法改进

由于算法主要基于谱图理论，虽然我们已经对下面几个参数的设定做了一些改进，但是，效果并不是特别的理想：选取的特征向量数目、特征间隙的设定。今后的工作中，我们可以选取更加恰当的核函数来计算特征矩阵、寻更合理的特征度量来计算点对之间的距离，而不是坐标的欧式距离、尝试使用半监督监督的方法来设定相应的参数。

基于谱图理论的特征向量直方图在形状匹配中的应用，可以拓展到场景识别和分类中。由于一般的场景图片中SIFT特征点的数量是非常多的，从而可以通过提取SIFT点来建立场景图像的对应图，再利用我们的方法，对场景进行分类和识别。

2. 基于夹角直方图的形状匹配算法改进

由于我们的工作中采用轮廓点，从某种程度上依赖对轮廓点的提取，因此，在今后的工作中，我们可以寻找某种方法对图像进行预处理，例如平滑算子或者腐蚀、膨胀等算法，去除毛刺等噪声、形变的影响，保证同类目标的轮廓相似性。我们选取局部曲率最大点作为特征点来近似目标，且采用夹角直方图来表示点，丢失了相邻点的边缘信息，尤其是相邻两点之间存在较大的圆弧时，对图像近似的影响更大，因此，在今后的工作中，我们将寻找某种方式将相邻点的信息加入到特征中去，用高维信息来描述点。

夹角直方图算法在经过改进，添加边缘信息之后，可以拓展在手写体数字的识别中。跟踪手写体数字的轮廓，利用边缘的曲率信息提取出手写体数字的特征，如角点、圈、局部曲率最大值点、端点、交叉点等。再利用这些特征对手写体数字进行分类。

参考文献

- [1] Alta Vista Photo Finder, <http://image.altavista.com/cgi-bin/avn.cgi>.
- [2] X.Bai, B.Wang, X.G.Wang, W.Y.Liu, Z.W.Tu. *Co-Transduction for Shape Retrieval* In Pro. of the 11th European Conference on Computer Vision, pp. 328-341, Sep.2010
- [3] D.H.Ballard. *Generalized Hough Transform to Detect Arbitrary Patterns*. IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.13, pp.111-122, 1981.
- [4] H.G.Barrow, J.M.Tenenbaum, R.C.Bolles, H.C.Wolf. *Parametric correspondence and chamfer matching: Two new techniques for image matching*. In Proc. 5th Int. Joint Conf. Artificial Intelligence, pp.659-663, 1977.
- [5] S.O.Belkasim, M.Shridhar, M.Ahmadi. *Pattern Recognition with Moment Invariant: A Descriptor in 2D Shape Classification*. IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.2, pp.201-207, 1994.
- [6] S.Belongie, J.Malik, J.Puzicha. *Shape Contexts: A New Descriptor for Shape Matching and Object Recognition*. In Pro. 2000 Conf. in Neural Information Processing System, pp.831-837, 2001.
- [7] S.Belongie, J.Malik, J.Puzicha. *Matching with shape contexts*. IEEE Workshop on Content-based access of Image and Video-Libraries, pp.20-26, 2000.
- [8] S.Belongie, J.Malik, J.Puzicha. *Shape Matching and Object Recognition Using Shape Contexts*. IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 24, pp.705-522, 2002.
- [9] J.Canny. *A Computational Approach To Edge Detection*. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 6, pp.679-698, 1986.

- [10] P.F.Felzenszwalb, J.D.Schwartz. *Hierarchical Matching of Deformable Shapes*. Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 2007
- [11] V.Ferrari, L.Fevrier, F.Jurie, C.Schmid. *Groups of Adjacent Contour Segments for Object Detection*. IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 30, no. 1, pp. 36-51, Jan. 2008.
- [12] I.Dhillon, Y.Guan, R.Kulis. *A Unified View of Kernel k means, Spectral Clustering and Graph Cuts* In Proc. 2nd International Conference of Pattern Recognition, pp.289-292, 2002
- [13] H.Freeman. *On the Encoding of Arbitrary Geometric Configurations*. IRE Transactions on Electronic Computers EC-10, pp.260-266, 1961.
- [14] S.Gold. *Matching and Learning Structural and Spatial Representations with Neural Networks*. PhD thesis, Yale University, 1995.
- [15] K.Grauman, T.Darrell. *The Pyramid Match Kernel: Discriminative Classification with Sets of Image Features*. Tenth IEEE International Conference on Computer Vision, vol.2, pp.1458-1465, 2005
- [16] D.Huttenlocher, S.Ullman. *Object Recognition Using Alignment*. International Conference on Computer Vision, London, pp.102-111, 1987.
- [17] C.Jacobs, A.Finkelstein, D.Salesin. *Fast Multiresolution Image Querying* Computer Graphics Proceeding SIGGRAPH, pp.277-286, 1995.
- [18] T.B.Sebastian, P.N.Klein, B.B.Kimia. *Recognition of Shapes by Editing Their Shock Graphs*. IEEE Trans. Pattern Anal. Mach. Intell, vol.26, no.5, pp.550-571, 2004.
- [19] D.Knosow, A.Sharma, D.Mateus, R.Horaud. *Inexact Matching of Large and Sparse Graphs Using Laplacian Eigenvectors*. In Proceedings 7th Workshop on Graph-based Representations in Pattern Recognition, 2009.

- [20] D.Mateus, R.Horaud, D.Knossov, F.Cuzzolin, E.Boyer. Articulated Shape Matching Using Laplacian Eigenfunctions and Unsupervised Point Registration. IEEE Conference on Computer Vision and Pattern Recognition, pp. 1-8, 2008.
- [21] S.Lazebnik, C.Schmid, J.Ponce. *Beyond Bags of Features: Spatial Pyramid Matching for Recognizing Natural Scene Categories*. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 2169-2178, 2006.
- [22] H.B.Ling, D.W.Jacobs. *Using the Inner-Distance for Classification of Articulated Shapes*. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 719-726, 2005.
- [23] H.B.Ling, X.W.Yang, L.J.Latecki. *Balancing Deformability and Discriminability for Shape Matching*.
- [24] H.B.Ling, D.W.Jacobs. *Shape Classification Using the Inner-Distance*. IEEE Trans. on Pattern Anal. and Mach. Intell, vol.29, no.2, pp.286-299, 2007
- [25] S.Loncaric. *A Survey of Shape Analysis Techniques*. Pattern Recognition, vol.31, pp.983-1001, 1998.
- [26] U.V.Luxburg. *A Tutorial on Spectral Clustering*. Statistics and Computing, vol.17, no.4, pp.395-416, 2007.
- [27] F.Mokhtarian, S.Abbasi, J.Kittler. *Efficient and Robust Retrieval by Shape Content Through Curvature Scale Space*. Image Databases and Multimedia Search, First International Workshop IDB-MMS'96, Amsterdam, The Netherlands, pp.35-42, 1996.
- [28] *Hough P V C. Methods and means for recognizing complex patterns*. USA. United States Patent, 3069644, 1962.
- [29] QBIC project. <http://www.qbic.almaden.ibm.com/>.

- [30] P.L.Rosin. *Techniques for Assessing Polygonal Approximations of Curves*. IEEE Trans. Pattern Recognition and Machine Intelligence, vol.19, pp.659-666, 1997.
- [31] Schomaker, L, de Leau, E, Vuurppijl, L. *Using Pen-Based Outlines for Object-Based Annotation and Image-Based Queries*. Visual Information and Information Systems-Third International Conference VISUAL'99, Amsterdam, The Netherlands, Lecture Notes in Computer Science 1614, pp.585-592, June, 1999.
- [32] S.Sclaroff, A.P.Pentland. *Modal Matching for Correspondence and Recognition*. IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.17, pp.545-561, 1995.
- [33] C.Scott, R.Nowak. *Robust contour Matching Via the Order-Preserving Assignment Problem*. Proc. IEEE Trans. on Image Processing, vol.15, no.7, pp.1831-1838, 2006.
- [34] T.B.Sebastian, P.N.Klein, B.B.Kimia. *Recognition of Shapes by Editing Their Shock Graphs*. IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.26, no.5, pp.550-571, 2004.
- [35] J.B.Shi, J.Malik. *New Spectral Methods for Ratio Cut Partitioning and Clustering*. IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.22, no.8, pp. 888-905, 2000.
- [36] J.Shotton, M.Johnson, R.Clipolla. *Multi-Scale Categorical Object Recognition Using Contour Fragments*. IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.30, no.7, pp.1270-1281, 2008.
- [37] J.Shotton, A.Blake, R.Cipolla. *Contour-Based Learning for Object Detection*. 10th IEEE International Conference on Computer Vision, vol.1, pp.503-510, 2005.
- [38] C.G.Small. *The Statistical Theory of Shapes*. Springer Series in Statistics, Springer, 1996.

- [39] G.Stockman. *Object Recognition and Localization via Pose Clustering*. Computer Vision Graphics Image Process, vol:40, pp.361-387, 1987.
- [40] S.Umeyama. *An eigen decomposition approach to weighted graph matching problems*. IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.10, pp.695-703, 1988.
- [41] S.Umeyama. *Parameterized Point Pattern Matching and its Application to Recognition of Object Families*. IEEE Trans. Pattern Analysis and Machine Intelligence, vol.15, pp.136-144, 1993.
- [42] D. Vernon. *Machine Vision : Automated Visual Inspection and Robot Vision*. Prentice Hall, 1991, ISBN 0-13-543398-3.
- [43] H.J.Wolfson, I.Rigoutsos. *Geometric Hashing: An Overview*. IEEE Computer Science Eigene, pp.10-21, Oct.-Dec, 1997.
- [44] I.Yahiaoui, N.Hervé, N.Boujemaa. *Shape-based Image Retrieval Inbotanical Collections*. In PCM, vol.4261, pp.357-364, 2006.
- [45] L.Zelnik-Manor, P.Perona. *Self-Tuning spectral Clustering*. Advances in Neural Information Processing Systems 17, vol.2, pp.1601-1608, 2004.

发表文章目录

- [1] Zhang Cui, Veronique Prinnet, "Shape Matching Using the Included Angle Histogram of Vectors", Chinese Conference of Pattern Reconition, (C-CPR2010), ChongQing, China.
- [2] Zhang Cui, Veronique Prinnet, "Spectral Analysis for Shape Matching", Chinese Conference of Pattern Reconition, (CCPR2010), ChongQing, China.

简 历

基本情况

张翠，女，河北省枣强县人，1984年6月出生，未婚，中国科学院自动化研究所模式识别国家重点实验室，中法信息、自动化与应用数学联合实验室在读硕士研究生。

教育情况

2004年9月至2008年7月，武汉大学数学与统计学院，本科，专业：信息与计算科学。

2008年9月至2011年7月，中国科学院自动化研究所，模式识别国家重点实验室，中法信息、自动化与应用数学联合实验室在读硕士研究生，专业：模式识别与智能系统。

工作经历

无。

研究兴趣

形状识别，计算机视觉，模式识别

联系方式

通讯地址：北京市2728信箱，中国科学院自动化研究所，模式识别国家重点实验室，中法信息、自动化与应用数学联合实验室

邮编：100190

E-mail:czhang@nlpr.ia.ac.cn

致 谢

我在攻读硕士期间从事的学习和科研工作，是在导师和实验室各位老师、同学的指导和帮助下完成的。在完成论文之际，我谨表最诚挚的谢意。

本文是在Veronique Prinèt老师和程健老师的亲切关怀和悉心指导下完成的。他们严肃的科学态度，严谨的治学精神，精益求精的工作作风，深深的感染和激励着我。他们不仅在学业上给我精心指导，同时还在思想、生活上给予我以无微不至的关怀，在此谨向所有关心和指导我的老师致以诚挚的谢意和崇高的敬意。

特别感谢马东民、何亮亮、袁飞、龚星同学，他们在项目和论文撰写中给予我极大的帮助。正式由于这些同学热心的帮助和支持，我才能克服一个个困难和疑惑，直至本文的顺利完成。另外非常感谢课题组的彭婷、Regis、Paul、Cyril、Antoine、Pierre，在学习、工作和生活中你们给予了我很多的指导和帮助，丰富和充实了我的生活，同时也让我更加充满自信的面对未来。同时感谢中科院自动化所各位老师和同学在学习和生活上的无私帮助，感谢你们给予的良好学习和研究环境。最后，我还要感谢培养我长大含辛茹苦的父母，以及一路默默支持和关心我的兄长、家人和朋友，谢谢你们！