图像处理与分析

第十八讲：图像表示与描述

胡晰远

中国科学院自动化研究所

xiyuan.hu@ia.ac.cn

内容大纲

- 基本概念
- 表示方法
- 图像描述
  - 边界描述子
  - 区域描述子
  - 关系描述子
  - 特征描述子
- 图像目标检测

基本概念

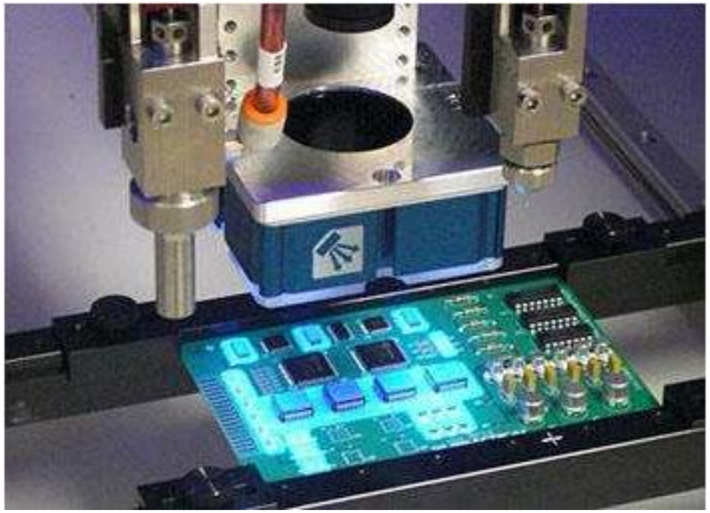
- 表示：直接具体地表示目标
  - 好的表示方法应具有节省存储空间、易于特征计算等优点
- 描述：描述是较抽象地表示目标
  - 好的描述应在尽可能区别不同目标的基础上对目标的尺度、平移、旋转等不敏感

基本概念

- 表示和描述是密切联系的
  - 表示的方法对描述很重要，因为它限定了描述的精确性；
  - 而通过对目标的描述，各种表示方法才有实际意义；
- 表示和描述又有区别
  - 表示侧重于数据结构，而描述侧重于区域特性以及不同区域间的联系和差别

基本概念

- 表示和描述在具体应用中非常重要
  - 在工业视觉检测中，好的表示方法才能实现高速高效；
  - 在基于内容的检索中，选用什么特征来描述目标，如何精确地测量这些目标；
  - ……



内容大纲

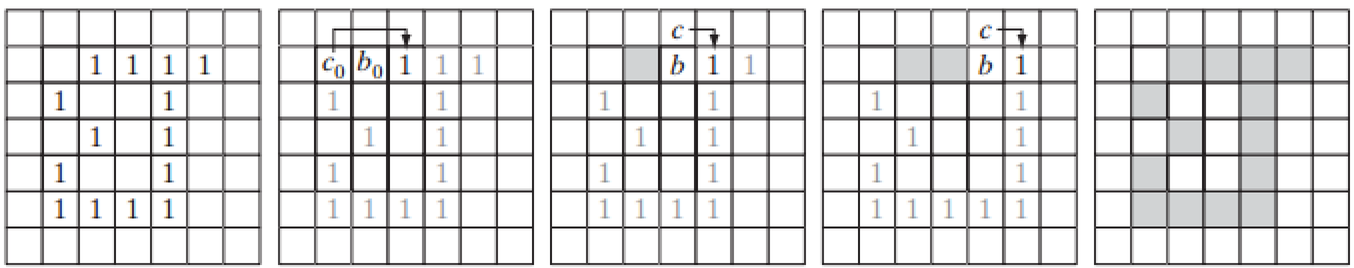
- 基本概念
- 表示方法
- 图像描述
  - 边界描述子
  - 区域描述子
  - 关系描述子
  - 特征描述子
- 图像目标检测

表示方法

- 链码
- 多边形近似
- 外形特征
- 边界分段
- 区域骨架

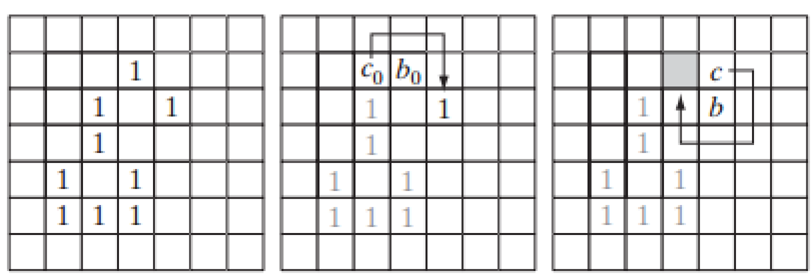
表示方法

- 边界追踪



a b c d e

FIGURE 11.1 Illustration of the first few steps in the boundary-following algorithm. The point to be processed next is labeled in black, the points yet to be processed are gray, and the points found by the algorithm are labeled as gray squares.



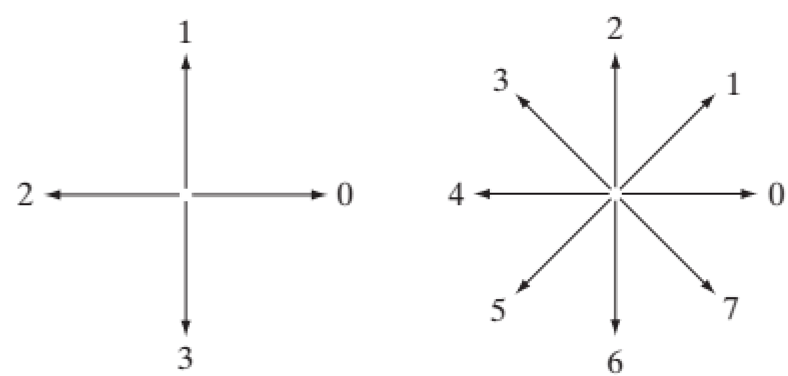
Error case

表示方法

- 链码
  - 链码用于表示由顺序连接的具有指定长度和方向的直线段组成的边界线
  - 这种表示方法基于线段的4或8连接
  - 每一段的方向使用数字编号方法进行编码

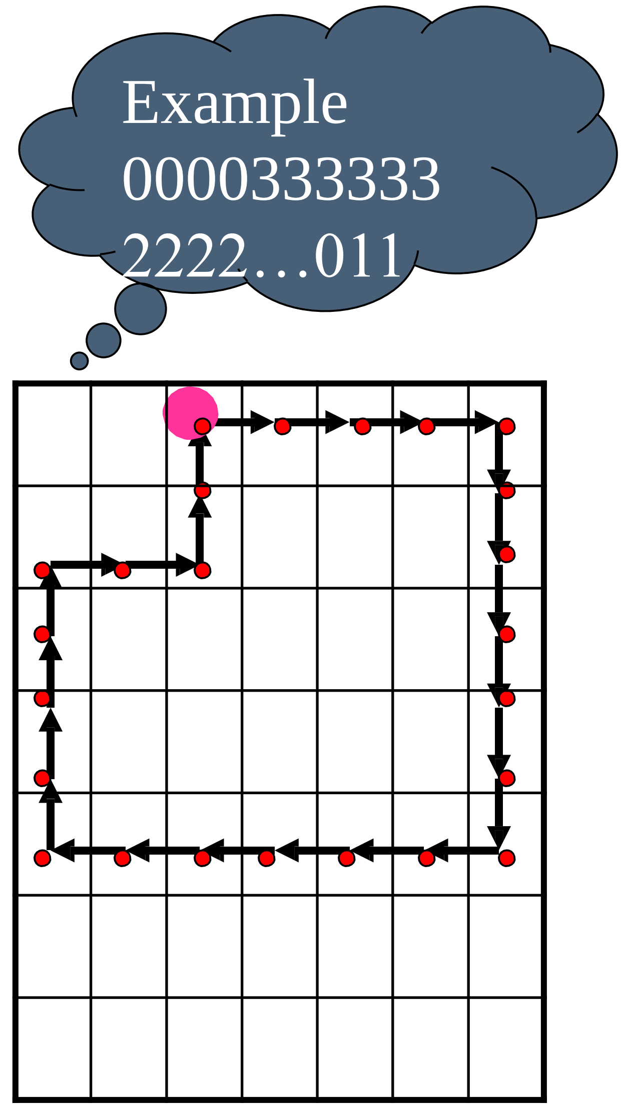
a b

FIGURE 11.3 Direction numbers for (a) 4-directional chain code, and (b) 8-directional chain code.



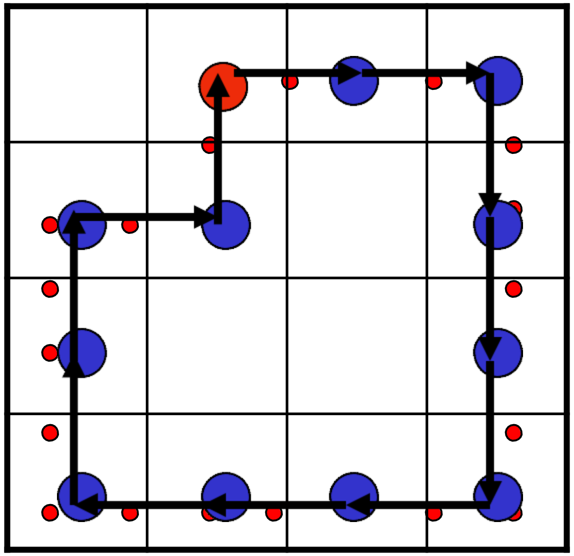
表示方法

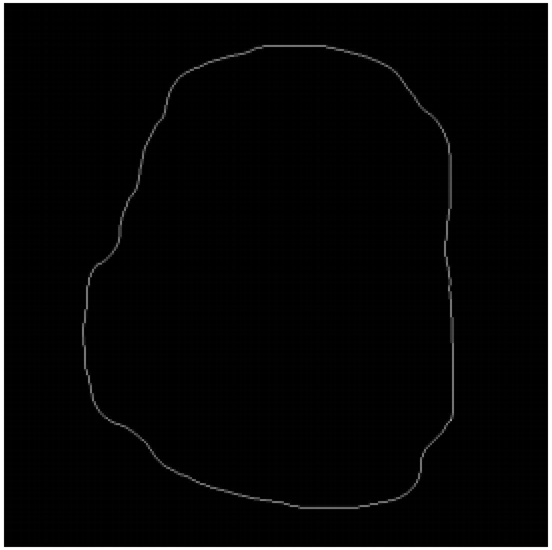
- 链码算法
  - 给每一个线段边界一个方向编码；
  - 有4链码和8链码两种编码方法；
  - 从起点开始，沿边界编码，至起点被重新碰到，结束一个对象的编码。

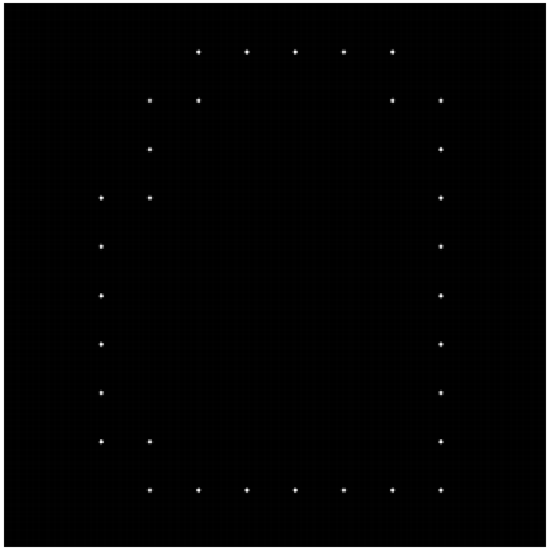


表示方法——链码

- 问题1：会不会受到噪声的干扰?
  - 加大采样网格空间
  - 依据原始边界与结果的接近程度，确定新点的位置





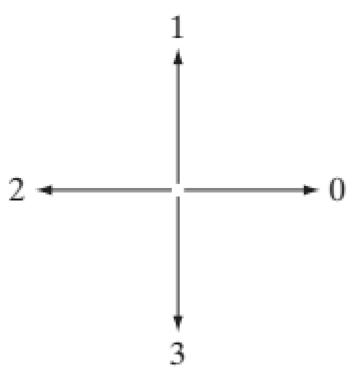


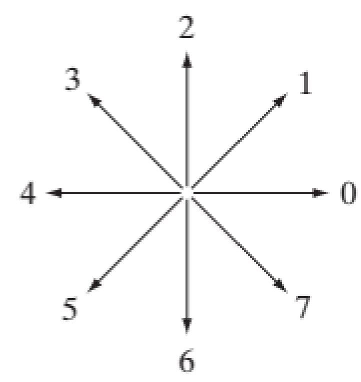
4链码：003332221101

表示方法——链码

- 问题2：由于起点和角度等不同，造成编码不同?
  - 从固定位置作为起点(最左最上)开始编码
  - 通过使用链码的差分代替码字本身的方式

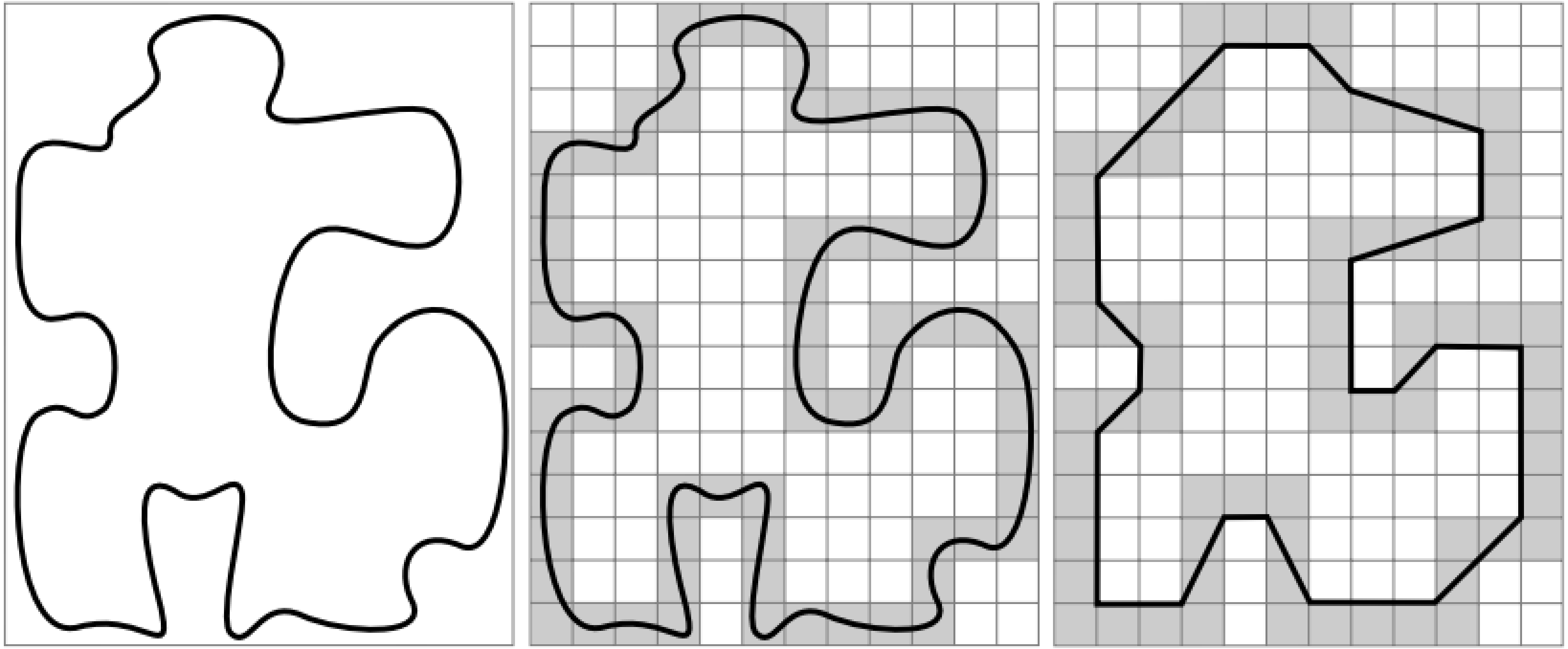
Example：
4-direction chain code：
10103322
Circular first difference：
33133030





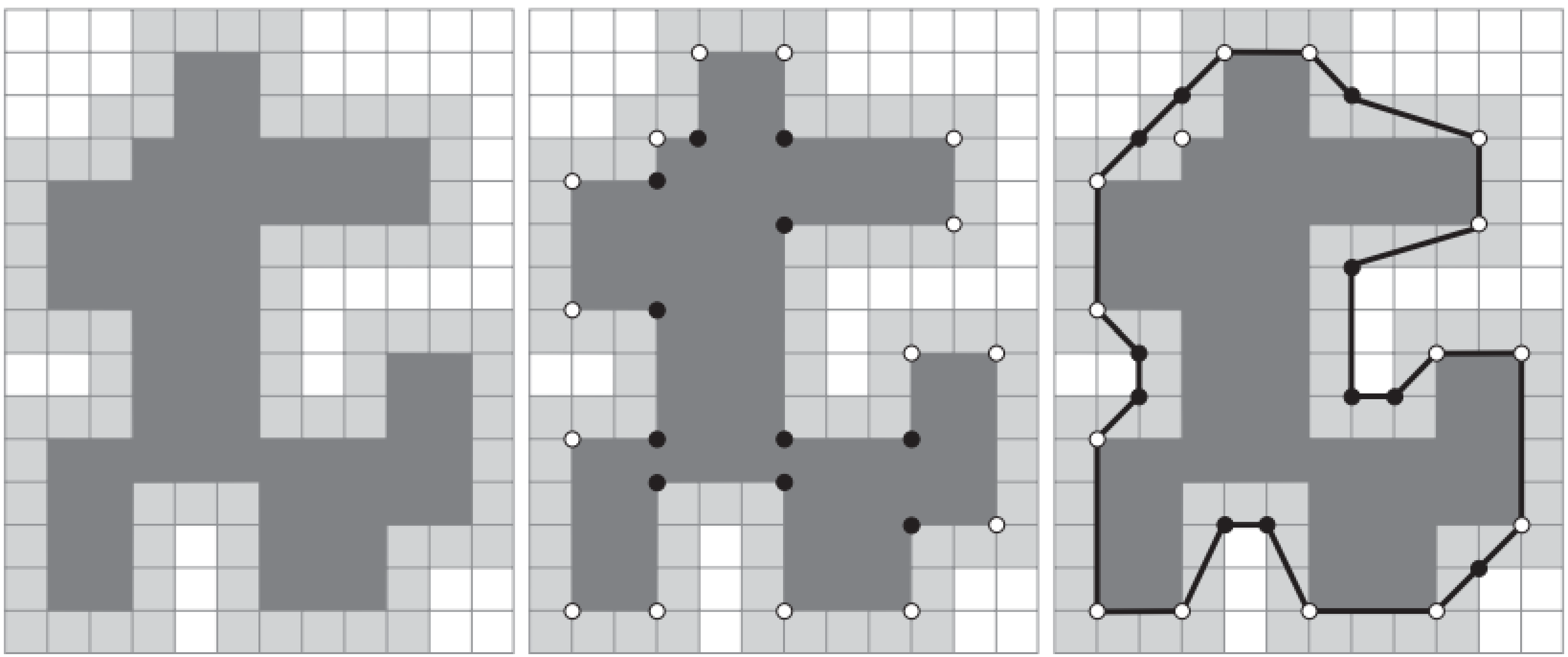
表示方法

- 最小周长多边形近似（MPP算法，Minimum-Perimeter Polygons）
  - 用最少的多边形线段，获取边界形状的本质



表示方法

- 最小周长多边形近似（MPP算法）
  - 定义凸顶点（W）
  - 定义镜像凹顶点（B）



表示方法

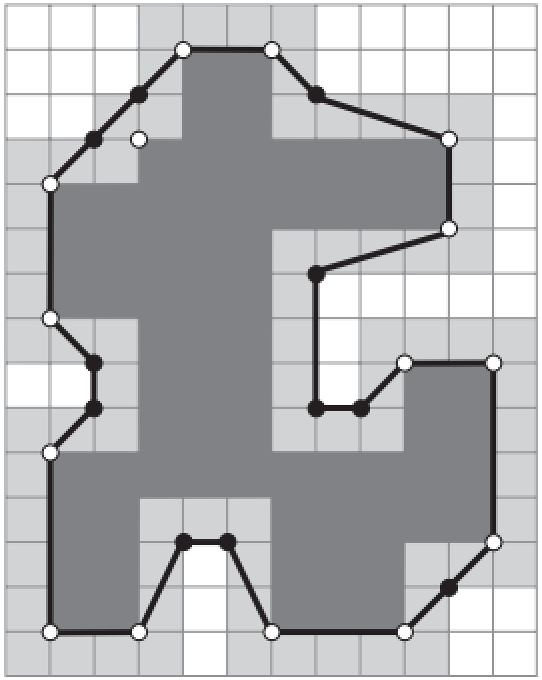
- 最小周长多边形近似（MPP算法）
  - 如何判断平面上三点的关系？
  - $\frac{y-y_1}{y_2-y_1} = \frac{x-x_1}{x_2-x_1}$
  - $|A| = \left|\begin{bmatrix} y-y_1 & x-x_1 \\ y_2-y_1 & x_2-x_1 \end{bmatrix}\right| = \left\|\begin{bmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{bmatrix}\right\|$
  - $|A| = \begin{cases} > 0 & L(x,y) > 0,\text{ 逆时针、上侧} \\ = 0 & L(x,y) = 0,\text{ 共线} \\ < 0 & L(x,y) < 0,\text{ 顺时针、下侧} \end{cases}$

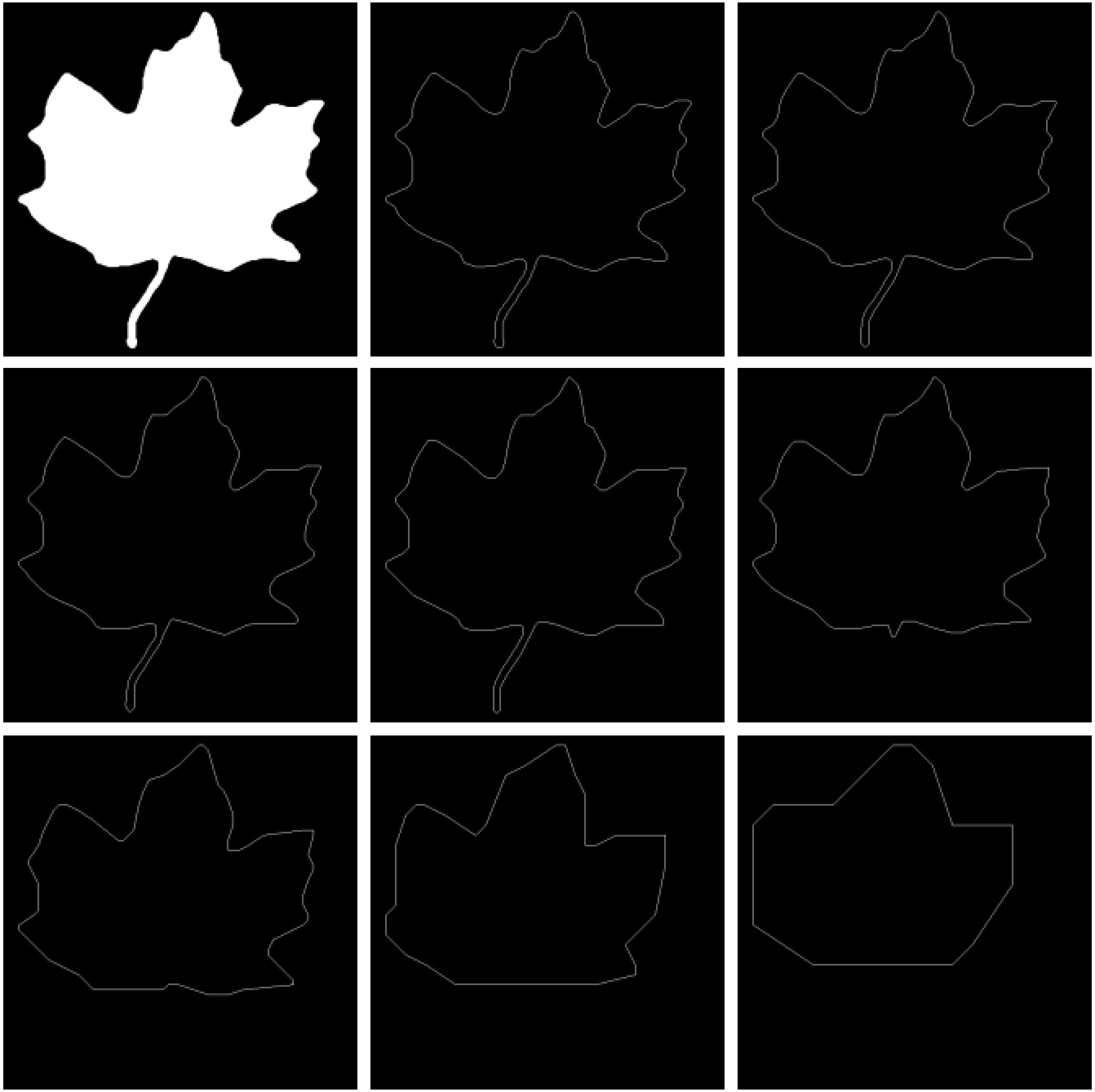
表示方法

- 最小周长多边形近似（MPP算法）
  - 按顺序遍历$W_c \to W, B_c \to B$
  - $W_c = B_c = V_0 = V_L$

1. $Sgn(V_L, W_c, V_k) > 0, V_L = W_c, reinit$
2. $\begin{cases} Sgn(V_L, W_c, V_k) \le 0 \\ Sgn(V_L, B_c, V_k) \ge 0 \end{cases}$, $\begin{cases} if\ V_k \in W, W_c = V_k \\ otherwise\ B_c = V_k \end{cases}$
3. $Sgn(V_L, B_c, V_k) < 0, V_L = B_c, reinit$

V_0	$(1, 4)$	W
V_1	$(2, 3)$	B
V_2	$(3, 3)$	W
V_3	$(3, 2)$	B
V_4	$(4, 1)$	W
V_5	$(7, 1)$	W
V_6	$(8, 2)$	B
V_7	$(9, 2)$	B



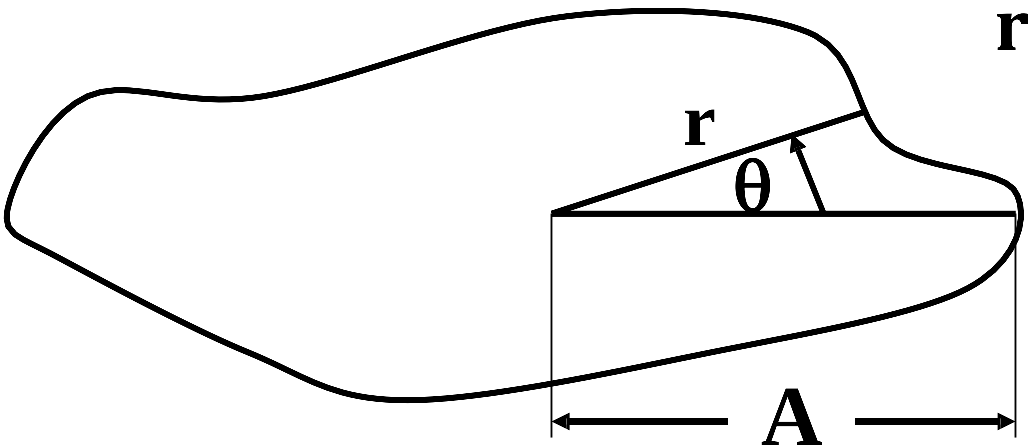


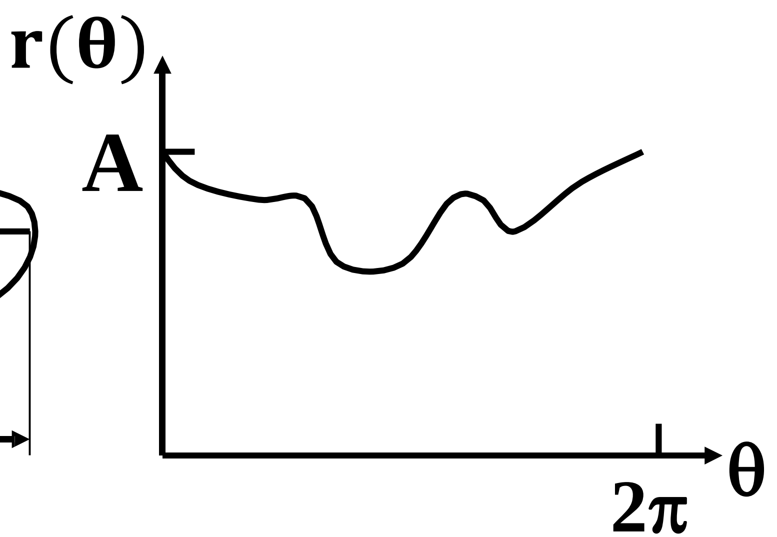
a b c
d e f
g h i

FIGURE 11.8 (a) 566 × 566 binary image. (b) 8-connected boundary. (c) through (i), MMPs obtained using square cells of sizes 2, 3, 4, 6, 8, 16, and 32, respectively (the vertices were joined by straight lines for display). The number of boundary points in (b) is 1900. The numbers of vertices in (c) through (i) are 206, 160, 127, 92, 66, 32, and 13, respectively.

表示方法

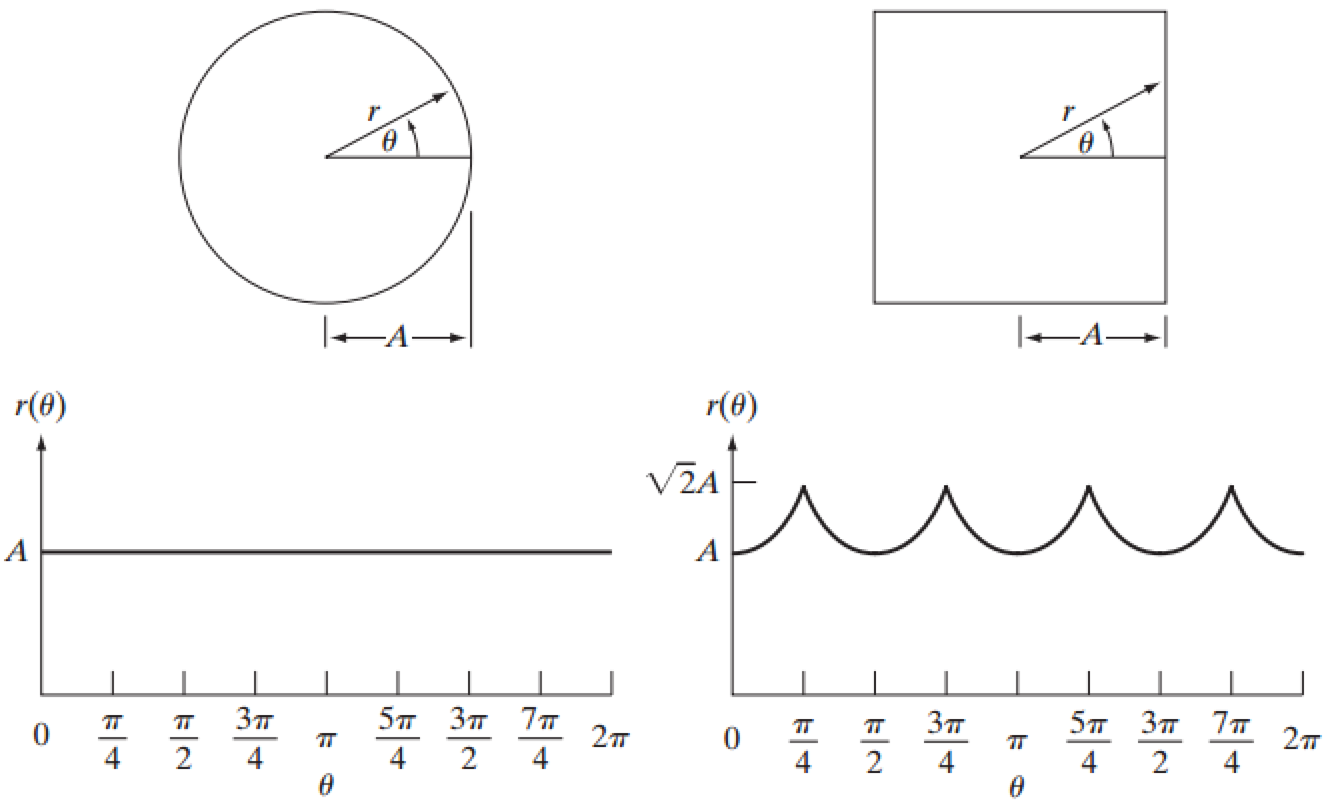
- 外形特征（标记图）
  - 基本思想：用一维函数表达边界，降低表示维度
  - 函数定义：质心角函数$r(\theta)$，夹角及距离





表示方法

- 外形特征（标记图）



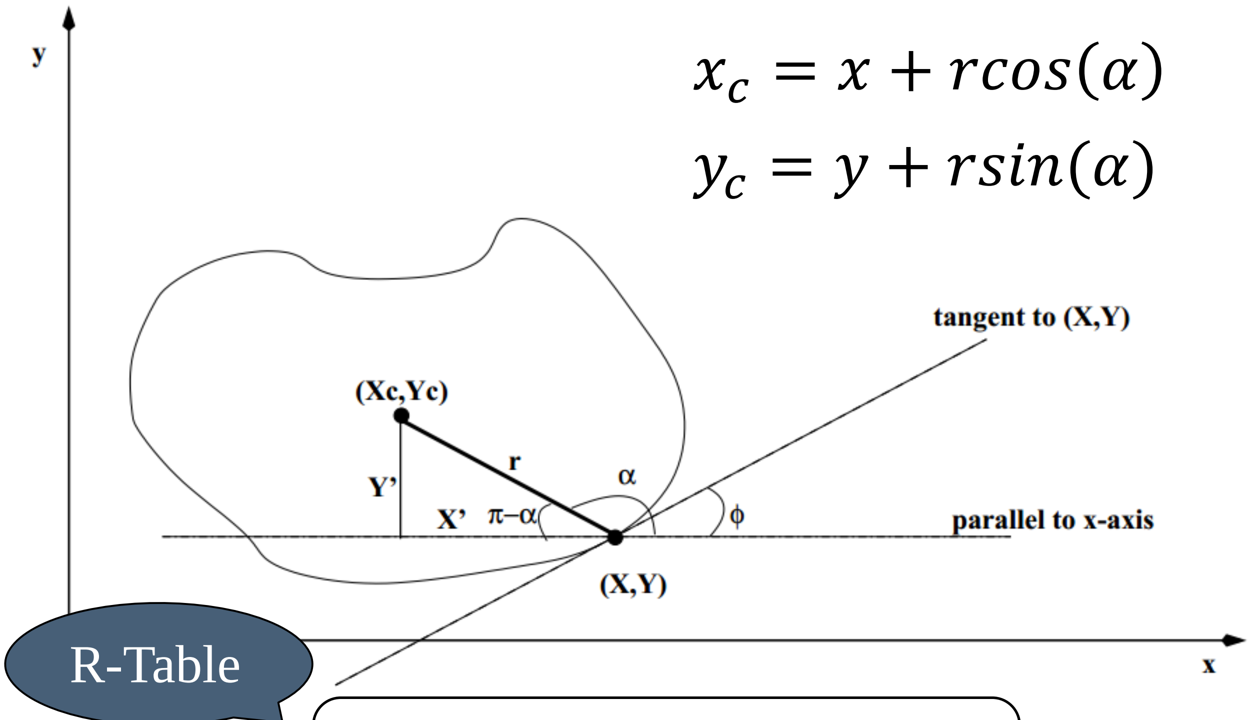
a b

FIGURE 11.10 Distance-versus-angle signatures. In (a) $r(\theta)$ is constant. In (b), the signature consists of repetitions of the pattern $r(\theta) = A \sec \theta$ for $0 \leq \theta \leq \pi/4$ and $r(\theta) = A \csc \theta$ for $\pi/4 < \theta \leq \pi/2$.

表示方法

- 外形特征（标记图）
  - 具有平移不变性；
  - 如何解决旋转问题：选择合适的起点（循环平移）
  - 如何解决缩放问题：函数值归一化
- 典型应用：广义霍夫变换（GHT）

广义霍夫变换（GHT）



$$x_c = x + r\cos(\alpha)$$
$$y_c = y + r\sin(\alpha)$$

$$\phi_1: (r_1^1, \alpha_1^1), (r_2^1, \alpha_2^1), \cdots$$
$$\phi_2: (r_1^2, \alpha_1^2), (r_2^2, \alpha_2^2), \cdots$$
$$\vdots$$
$$\phi_N: (r_1^N, \alpha_1^N), (r_2^N, \alpha_2^N), \cdots$$

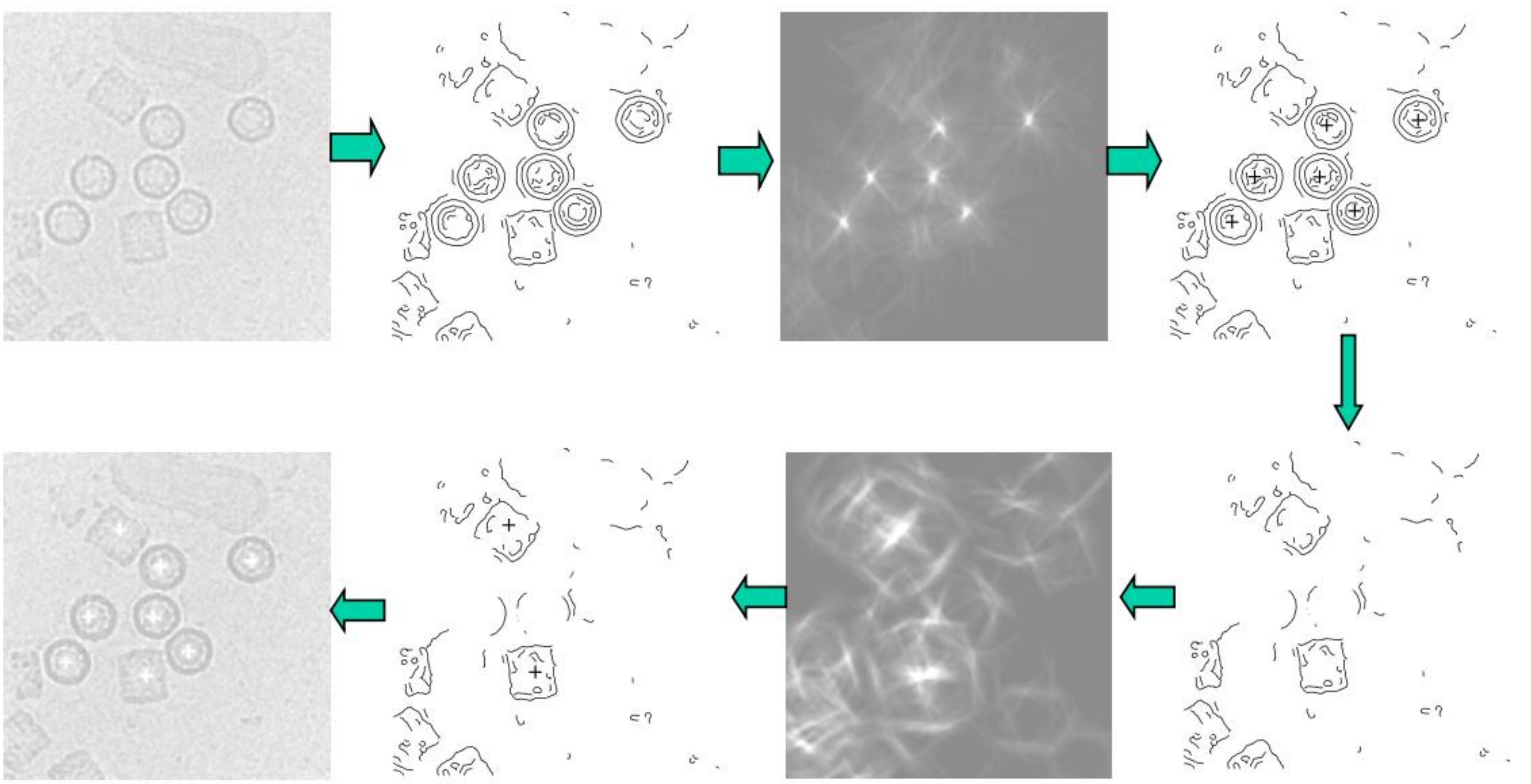
广义霍夫变换（GHT）

- 算法介绍

1. 参数空间量化：$P\left[x_{c_{min}} \cdots x_{c_{max}}\right]\left[y_{c_{min}} \cdots y_{c_{max}}\right]$
2. 遍历每个边界点(x, y)
   2.1 计算ϕ, 在R-Table中检索其对应的所有(r, α)
   2.2 对每个(r, α)，计算(x_c', y_c')
   2.3 在参数空间中投票：$++(P[x_c'][y_c'])$
3. 在投票空间中找出最大匹配点

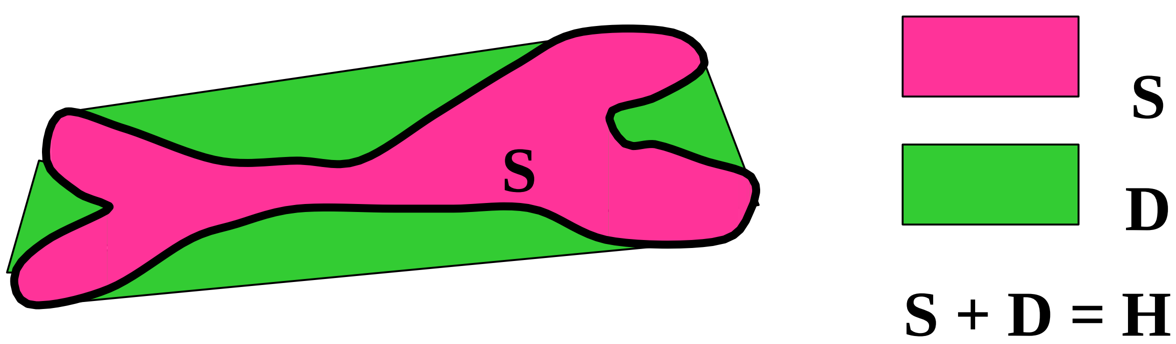
广义霍夫变换（GHT）

• 广义霍夫变换匹配示例



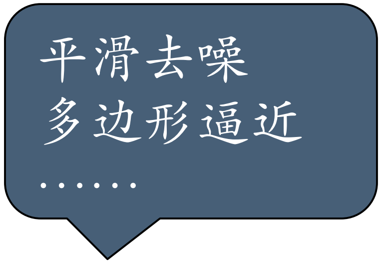
表示方法

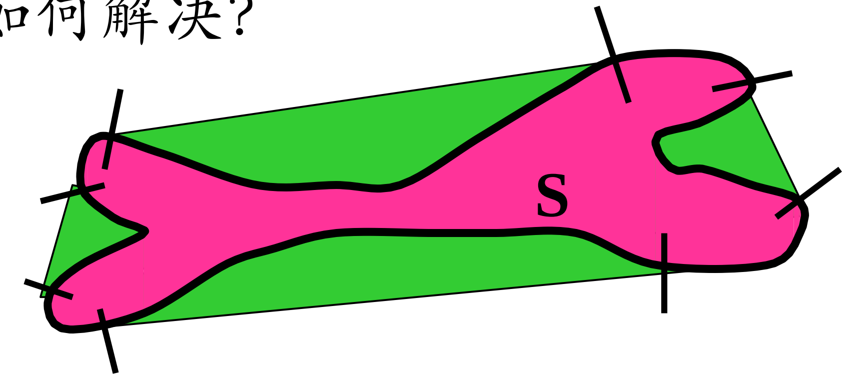
- 边界线段
  - 基本概念：
  - 一个任意集合S（区域）的凸起外缘H是：包含S的最小凸起的集合
  - $H-S$的差的集合被称为集合S的凸起补集D



表示方法

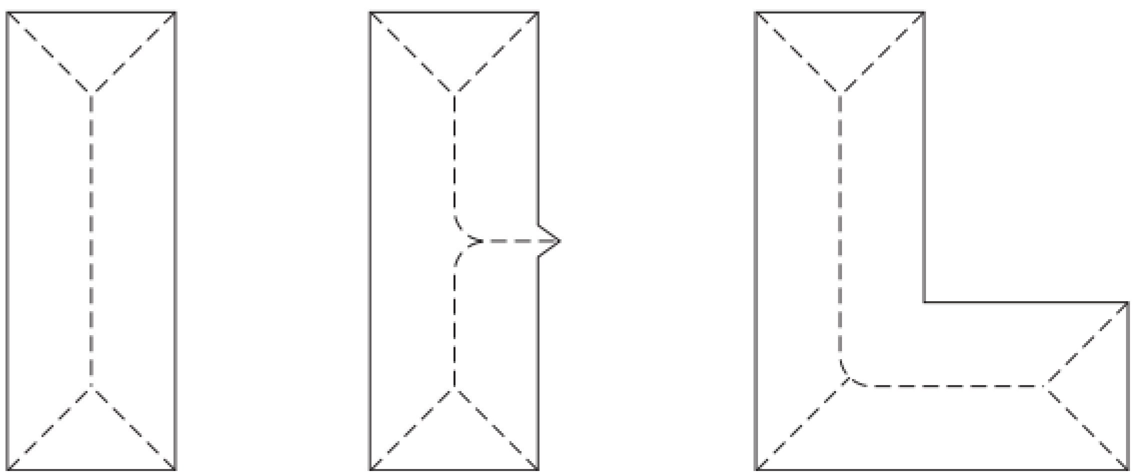
- 边界线段
  - 分段算法：
  - 给进入和离开凸起补集D的变换点打标记来划分边界段；
  - 优点：不依赖于方向和比例的变化；
  - 缺点：噪声影响零碎的边界划分，
    如何解决？





表示方法

- 区域骨架
  - 基本思想：把一个平面区域结构形状削减成图形，这种削减可以通过细化（也称为骨架提取）算法，获取区域的骨架来实现。
  - Blum的中轴变换方法（MAT）

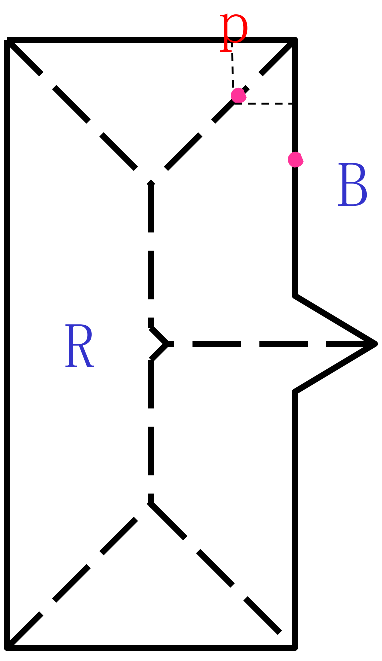


a b c

FIGURE 11.13 Medial axes (dashed) of three simple regions.

表示方法

- Blum中轴变换方法
  - 设：R是一个区域，B为R的边界点，对于R中的点p，找p在B上“最近”的邻居。如果p有多于一个的邻居，称它属于R的中轴（骨架）
  - 缺点：计算量大，需要计算区域的每个内部点到其边界点的距离。



表示方法

- 改进算法
  - 不断删除区域边界点，但保证删除满足:

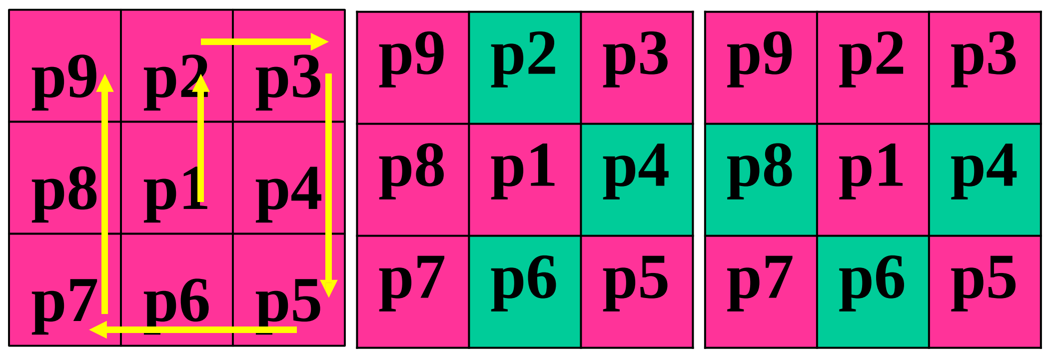
1. 不删除端点
2. 不破坏联通性
3. 不造成过度腐蚀

  - 改进算法举例：假设区域内点值为1，背景为0，对边界点进行两个基本操作（边界以8连通为例）

表示方法

- 改进算法
  - 基本操作1
  - 对于满足以下四个条件的边界点打标记准备删除：

1. $2 \le N(p_1) \le 6$
2. $S(p_1) = 1$
3. $p_2 \cdot p_4 \cdot p_6 = 0$
4. $p_4 \cdot p_6 \cdot p_8 = 0$



表示方法

- 改进算法
  - 所有条件都满足，才打删除标记。
  - 删除并不立即进行，而是等到对所有边界点都打完标记后，再把作了标记的点一起删除

举例

$N(p_1) = 4$

$T(p_1) = 3$

(3)(4)满足

0	0	1
1	p1	0
1	0	1

p9	p2	p3
p8	p1	p4
p7	p6	p5

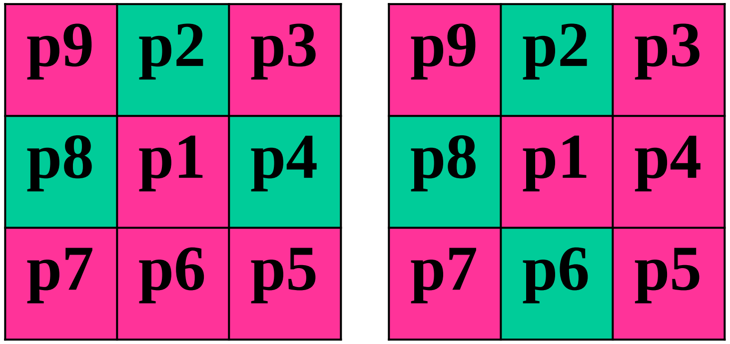
p9	p2	p3
p8	p1	p4
p7	p6	p5

表示方法

- 改进算法
  - 基本操作2
  - 条件（1）、（2）与操作1相同，条件（3）、（4）改为：

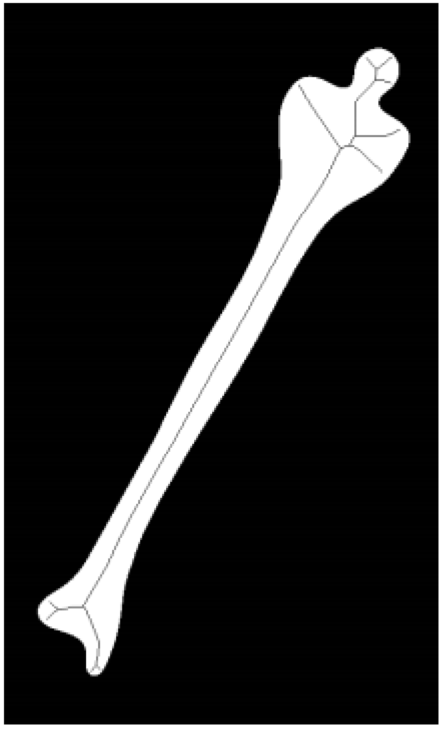
3. $p_2 \cdot p_4 \cdot p_8 = 0$

4. $p_2 \cdot p_6 \cdot p_8 = 0$



表示方法

- 改进算法总结
  - 按基本操作1，给边界点打标记——删除点
  - 按基本操作2，给边界点打标记——删除点
  - 这个基本过程反复进行，直至没有点可以删除为止。
  - 算法终止。

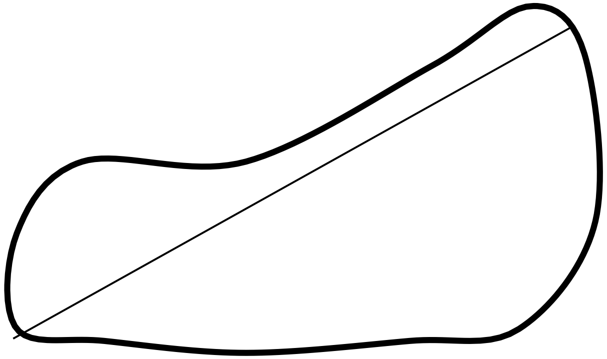


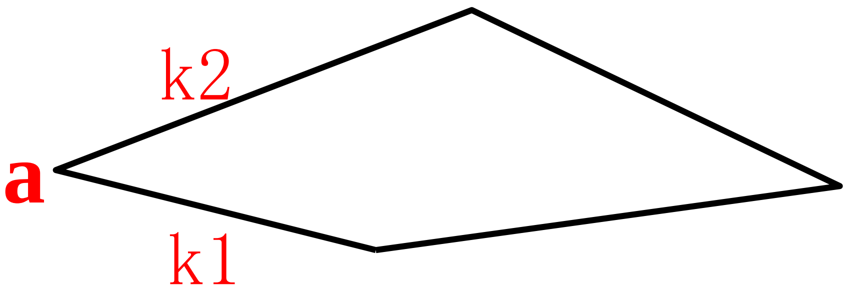
内容大纲

- 基本概念
- 表示方法
- 图像描述
  - 边界描述子
  - 区域描述子
  - 关系描述子
  - 特征描述子
- 图像目标检测

边界描述子

- 简单描述子
  - 周长
  - 直径：$Diam(B) = \max_{i,j}\left[D(p_i, p_j)\right]$
  - 曲率：$d_k = k_1 - k_2$





边界描述子

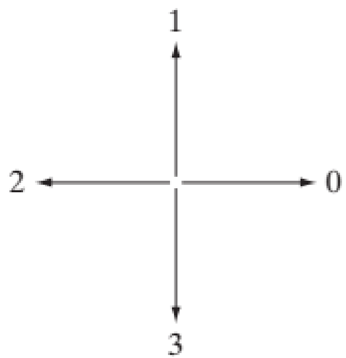
- 形状数——链码的实用化
  - 定义：最小循环首差链码
  - 示例：

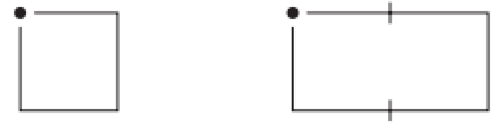
4-链码	10103322
循环首差	33133 030
形状数	03033133

  - 形状数的阶：形状数表达形式中的位数
  - 上例中形状数的阶是8;
  - 对于封闭边界序号一定是偶数，如4、6、8。

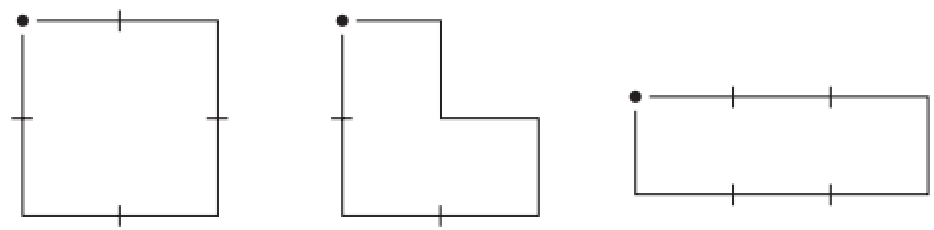
边界描述子

- 形状数示例





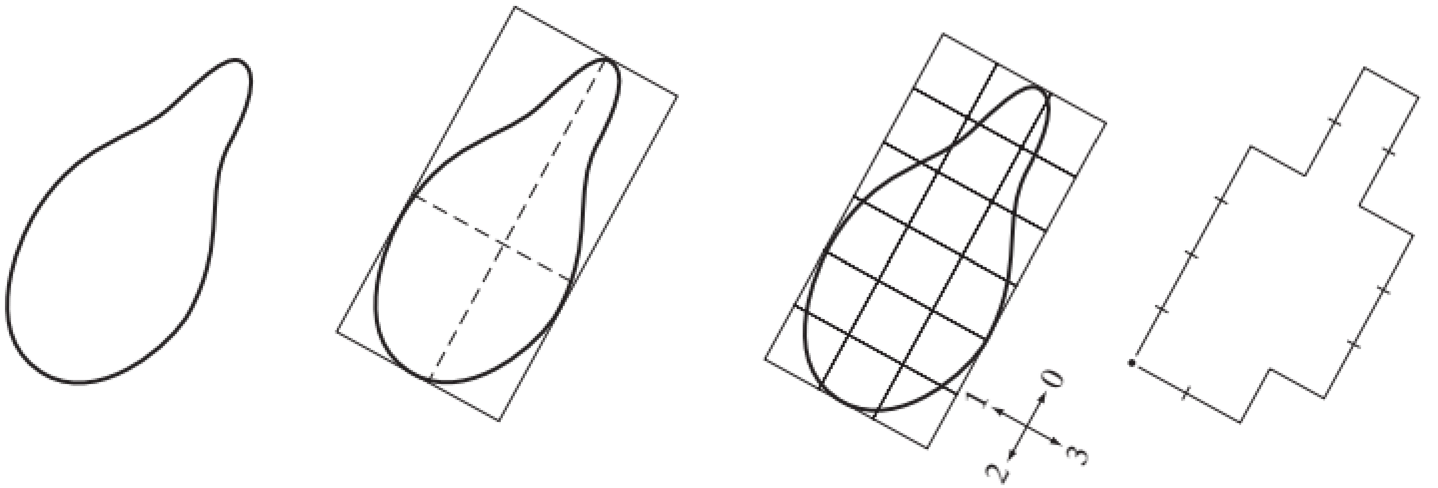
	Order 4	Order 6
Chain code:	0 3 2 1	0 0 3 2 2 1
Difference:	3 3 3 3	3 0 3 3 0 3
Shape no.:	3 3 3 3	0 3 3 0 3 3



Chain code:	0 0 3 3 2 2 1 1	0 3 0 3 2 2 1 1	0 0 0 3 2 2 2 1
Difference:	3 0 3 0 3 0 3 0	3 3 1 3 3 0 3 0	3 0 0 3 3 0 0 3
Shape no.:	0 3 0 3 0 3 0 3	0 3 0 3 3 1 3 3	0 0 3 3 0 0 3 3

边界描述子

- 形状数示例：生成步骤



Chain code: 0 0 0 0 3 0 0 3 2 2 3 2 2 2 1 2 1 1

Difference: 3 0 0 0 3 1 0 3 3 0 1 3 0 0 3 1 3 0

Shape no.: 0 0 0 3 1 0 3 3 0 1 3 0 0 3 1 3 0 3

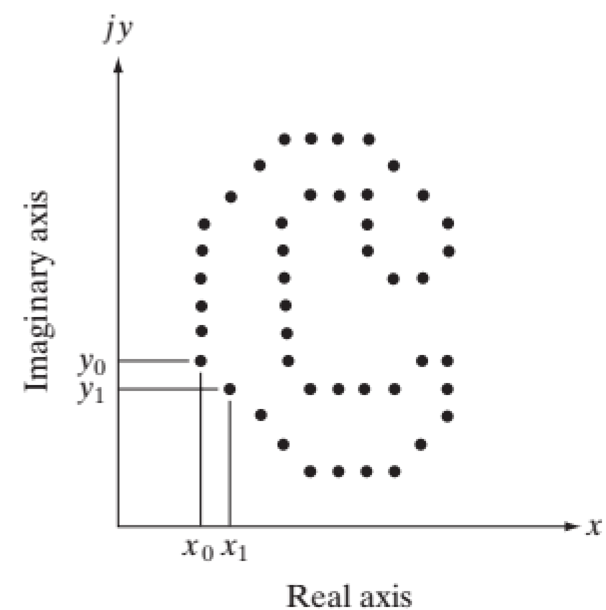
边界描述子

- 傅里叶描述子
  - 基本思想：将二维问题转化为一维问题
  - 边界点用复数表示：$s(k) = x(k) + jy(k)$
  - 进行离散傅里叶变换

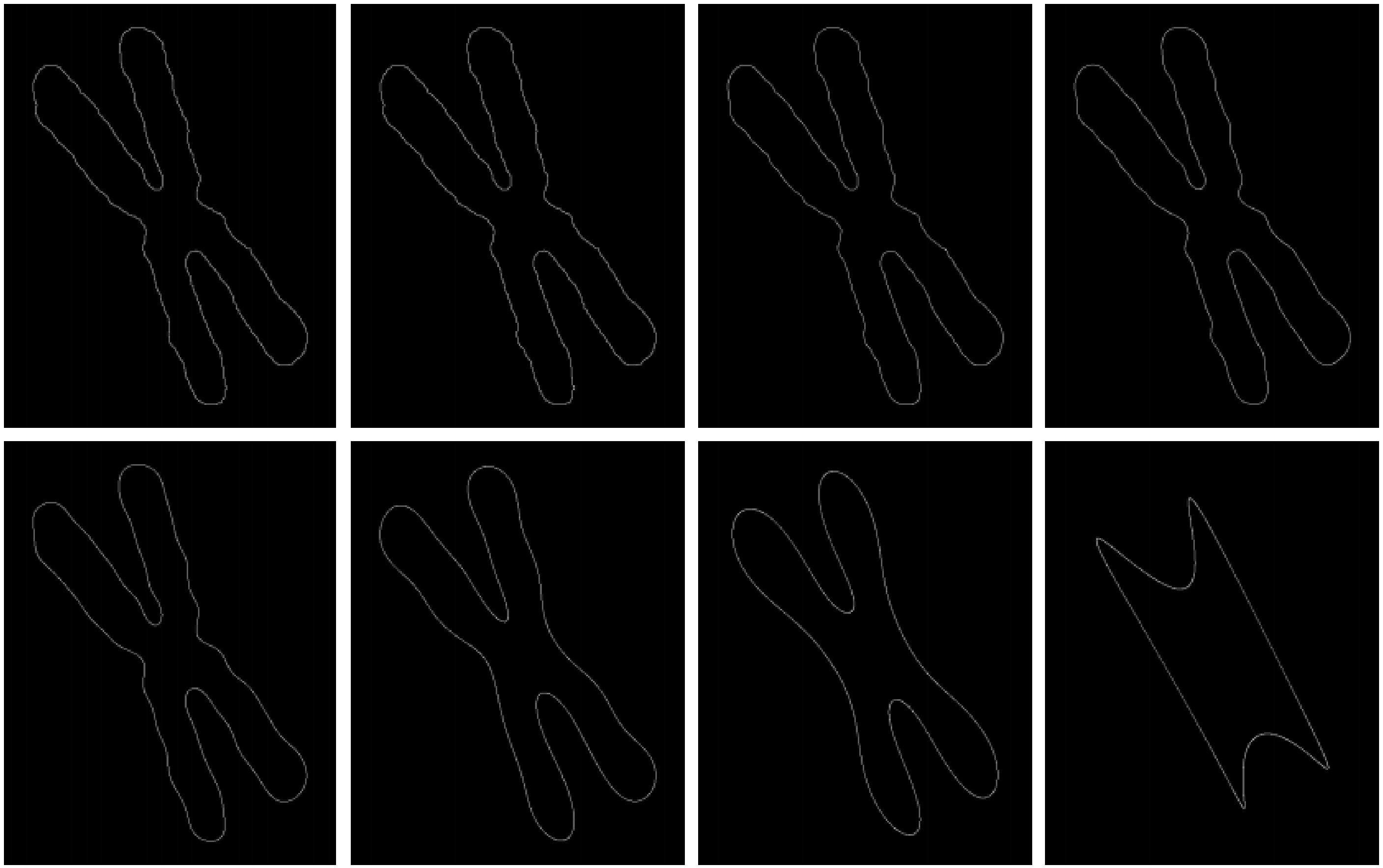
$$a(u) = \sum_{k=0}^{K-1} s(k) e^{-j2\pi uk/K}$$

  - $a(u)$被称为边界的傅里叶描述子
  - 边界逼近

$$\hat{s}(k) = \frac{1}{K} \sum_{u=0}^{P-1} a(u) e^{j2\pi uk/P}$$



傅里叶描述子示例



a b c d
e f g h

FIGURE 11.20 (a) Boundary of human chromosome (2868 points). (b)–(h) Boundaries reconstructed using 1434, 286, 144, 72, 36, 18, and 8 Fourier descriptors, respectively. These numbers are approximately 50%, 10%, 5%, 2.5%, 1.25%, 0.63%, and 0.28% of 2868, respectively.

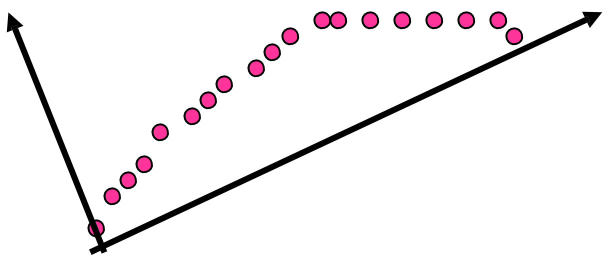
边界描述子

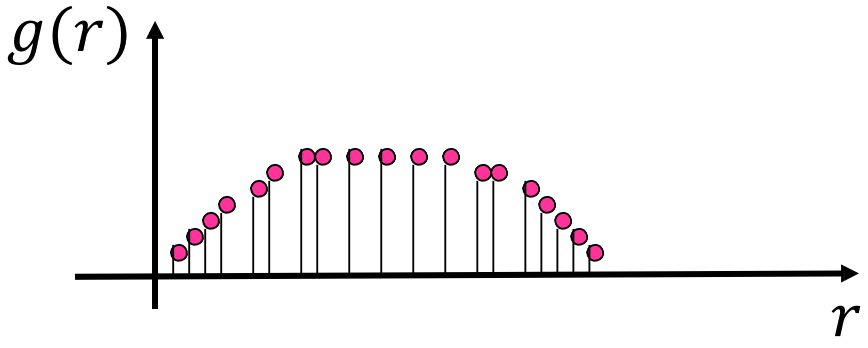
- 傅里叶描述子的性质
  - 对于旋转、平移、缩放等操作和起始点的选取不十分敏感；
  - 几何变换的描述子可通过对函数作简单变换来获得

Transformation	Boundary	Fourier Descriptor
Identity	$s(k)$	$a(u)$
Rotation	$s_r(k) = s(k)e^{j\theta}$	$a_r(u) = a(u)e^{j\theta}$
Translation	$s_t(k) = s(k) + \Delta_{xy}$	$a_t(u) = a(u) + \Delta_{xy}\delta(u)$
Scaling	$s_s(k) = \alpha s(k)$	$a_s(u) = \alpha a(u)$
Starting point	$s_p(k) = s(k - k_0)$	$a_p(u) = a(u)e^{-j2\pi k_0 u/K}$

边界描述子

- 统计矩
  - 将描述形状的任务简化为一维函数的矩统计量，例如均值、方差、高阶矩等；
  - 将边界当作直方图函数来处理：$g(r)$



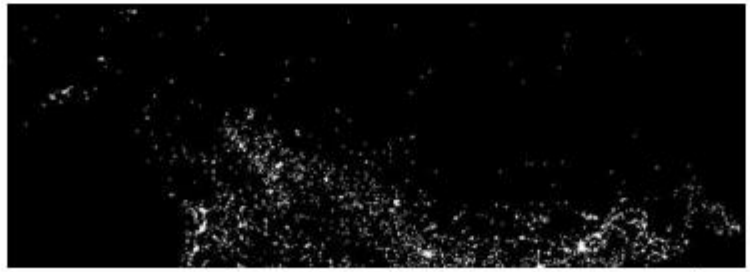


边界描述子

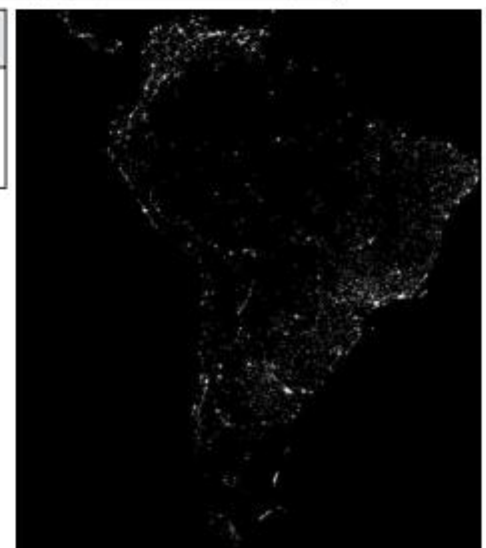
- 统计矩定义
  - $\mu_n(r)=\sum_{i=1}^{L}(r_i-m)^n g(r_i)$
  - 其中，$m=\sum_{i=1}^{L} r_i g(r_i)$
- 统计矩的优点
  - 实现简单
  - 提供一种关于边界形状的“物理”解释
  - 对旋转不敏感
  - 为了对缩放不敏感，可以通过伸缩r的范围来实现

区域描述子

- 致密性
  P/A
- 圆度率
  $R_c = 4\pi A/P^2$

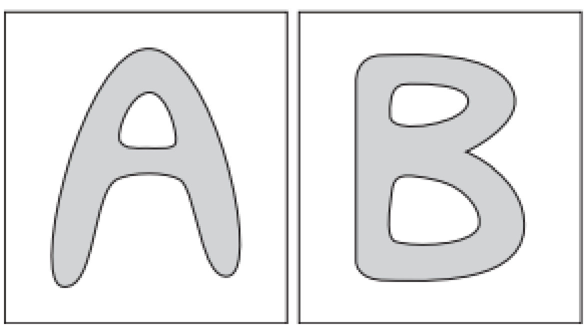


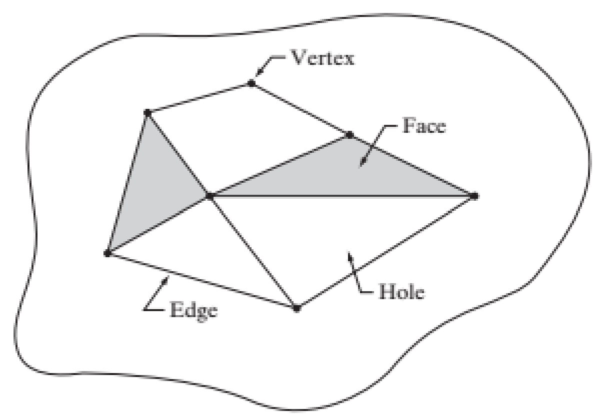
Region no. (from top)	Ratio of lights per region to total lights
1	0.204
2	0.640
3	0.049
4	0.107



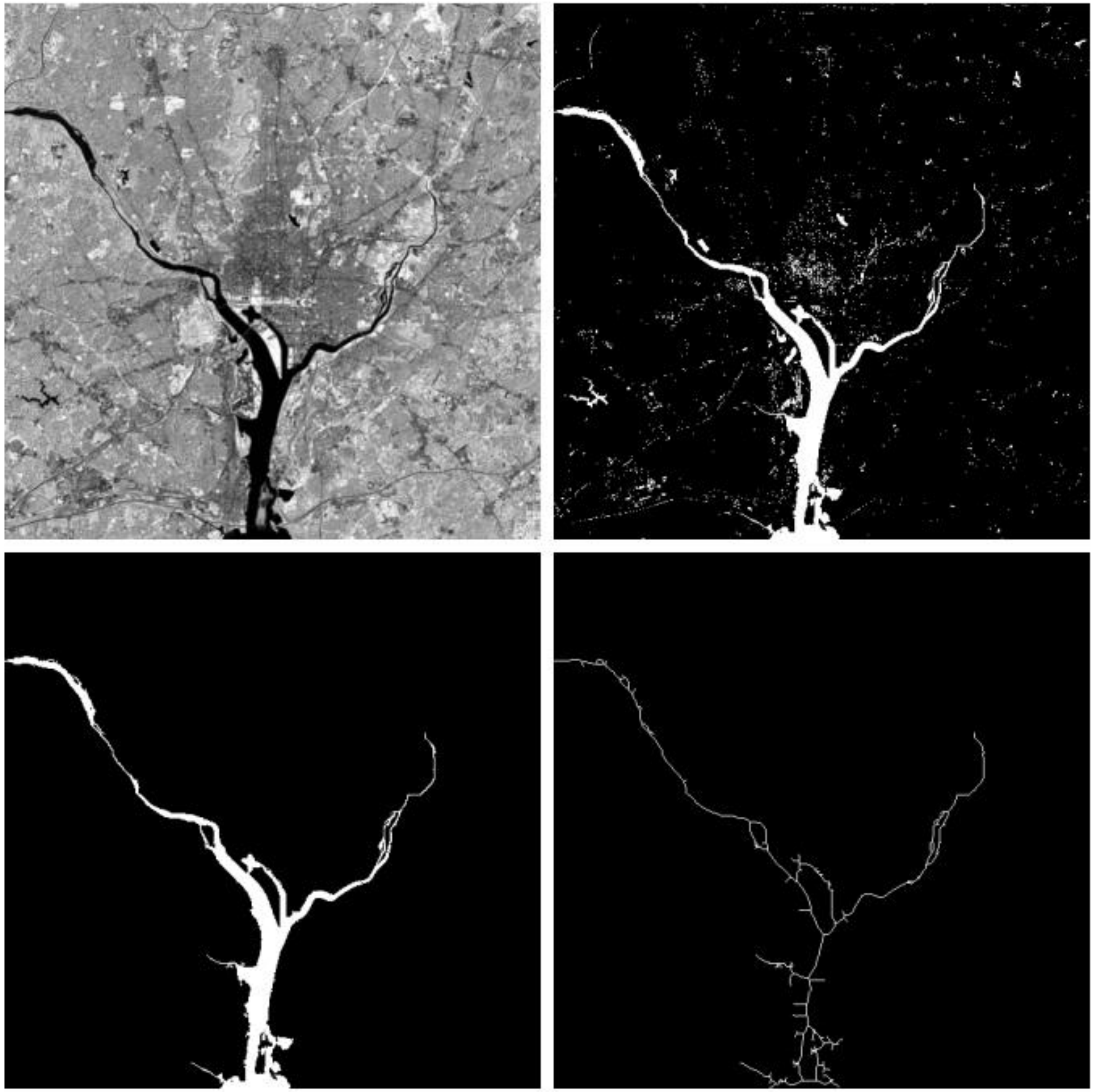
区域描述子

- 拓扑描述子
  - 欧拉数：$E = C - H$
  - 欧拉公式：$V - Q + F = E = C - H$
  - 拓扑特性与距离或基于距离测度概念的任何特性都无关；
  - 可以非常简单地解释多边形网络；





利用拓扑描述子的示例



a b
c d

FIGURE 11.27 (a) Infrared image of the Washington, D.C. area. (b) Thresholded image. (c) The largest connected component of (b). Skeleton of (c).

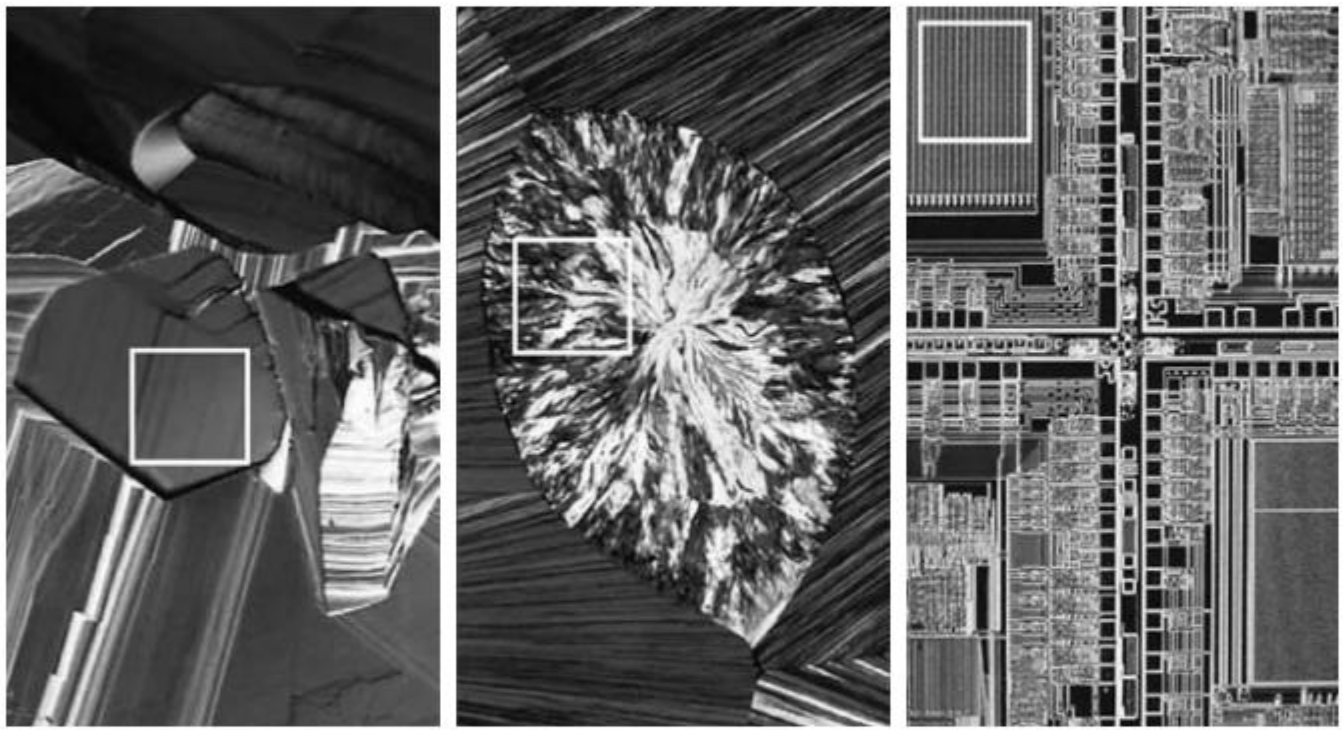
区域描述子

- 纹理描述子
  - 统计法：描述图像内容的平滑程度
  - 结构法：描述图像像元的排列规律
  - 频谱法：识别图像中的全局周期性

纹理描述子

- 统计法
  - 针对]

Texture	Mean	Standard deviation	R (normalized)	Third moment	Uniformity	Entropy
Smooth	82.64	11.79	0.002	−0.105	0.026	5.434
Coarse	143.56	74.63	0.079	−0.151	0.005	7.783
Regular	99.72	33.73	0.017	0.750	0.013	6.674



a b c

FIGURE 11.28 The white squares mark, from left to right, smooth, coarse, and regular textures. These are optical microscope images of a superconductor, human cholesterol, and a microprocessor. (Courtesy of Dr. Michael W. Davidson, Florida State University.)

纹理描述子

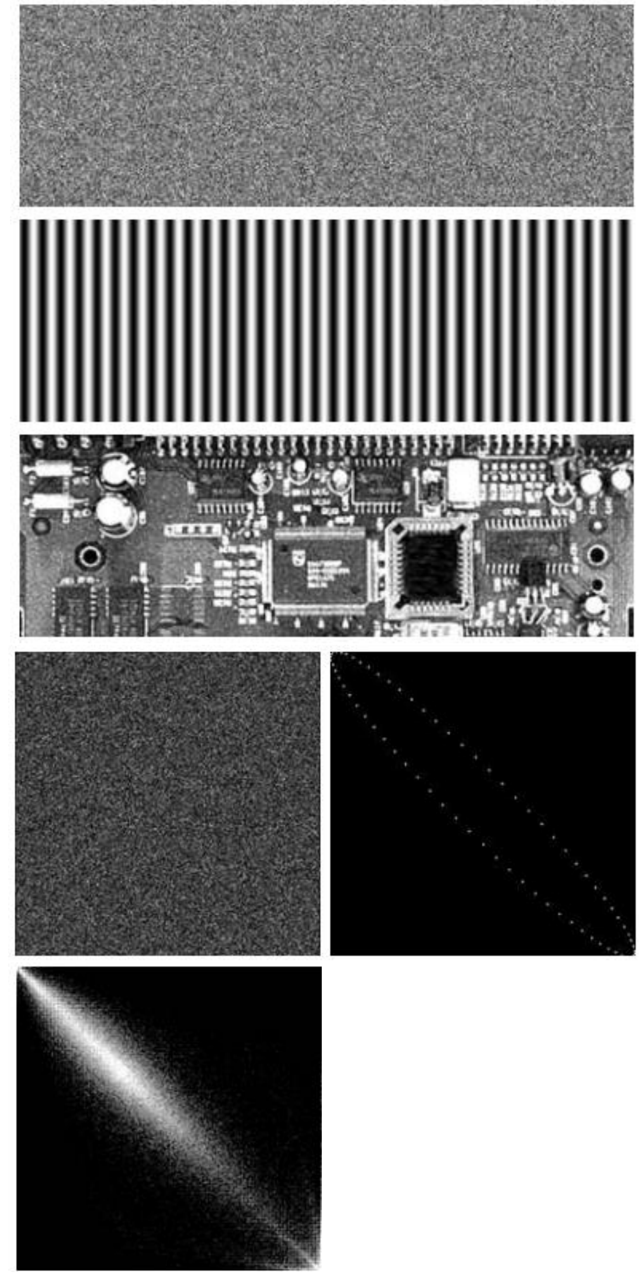
- 统计法
  - 直方图不包含空间位置信息
  - 针对共生矩阵的统计量

1	1	7	5	3	2
5	1	6	1	2	5
8	8	6	8	1	2
4	3	4	5	5	1
8	7	8	7	6	2
7	8	6	2	6	2

Image *f*

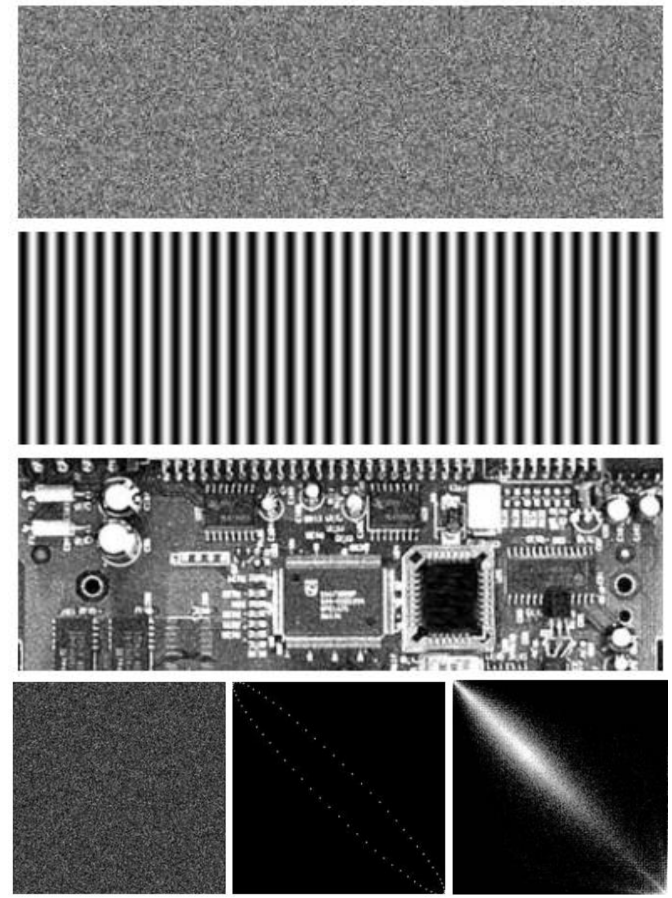
	1	2	3	4	5	6	7	8
1	1	2	0	0	0	1	1	0
2	0	0	0	0	1	1	0	0
3	0	1	0	1	0	0	0	0
4	0	0	1	0	1	0	0	0
5	2	0	1	0	1	0	0	0
6	1	3	0	0	0	0	0	1
7	0	0	0	0	1	1	0	2
8	1	0	0	0	0	2	2	1

Co-occurrence matrix **G**



纹理描述子——统计法

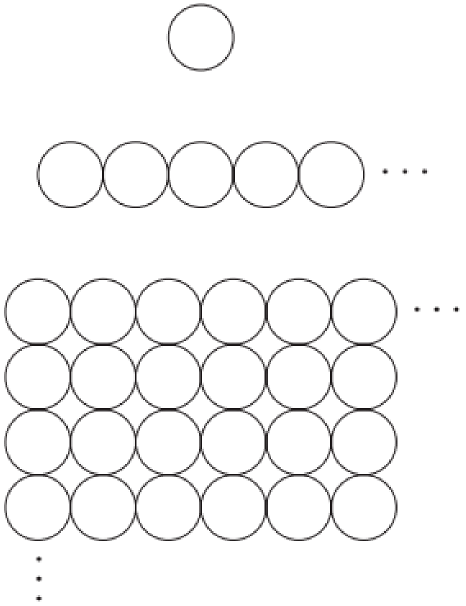
Descriptor	Explanation	Formula
Maximum probability	Measures the strongest response of **G**. The range of values is [0, 1].	$\max_{i,j}(p_{ij})$
Correlation	A measure of how correlated a pixel is to its neighbor over the entire image. Range of values is 1 to −1, corresponding to perfect positive and perfect negative correlations. This measure is not defined if either standard deviation is zero.	$\sum_{i=1}^{K}\sum_{j=1}^{K}\frac{(i-m_r)(j-m_c)p_{ij}}{\sigma_r\sigma_c}$ $\sigma_r \neq 0;\ \sigma_c \neq 0$
Contrast	A measure of intensity contrast between a pixel and its neighbor over the entire image. The range of values is 0 (when **G** is constant) to $(K-1)^2$.	$\sum_{i=1}^{K}\sum_{j=1}^{K}(i-j)^2 p_{ij}$
Uniformity (also called Energy)	A measure of uniformity in the range [0, 1]. Uniformity is 1 for a constant image.	$\sum_{i=1}^{K}\sum_{j=1}^{K}p_{ij}^2$
Homogeneity	Measures the spatial closeness of the distribution of elements in **G** to the diagonal. The range of values is [0, 1], with the maximum being achieved when **G** is a diagonal matrix.	$\sum_{i=1}^{K}\sum_{j=1}^{K}\frac{p_{ij}}{1+\lvert i-j\rvert}$
Entropy	Measures the randomness of the elements of **G**. The entropy is 0 when all p_{ij}'s are 0 and is maximum when all p_{ij}'s are equal. The maximum value is $2\log_2 K$. (See Eq. (11.3-9) regarding entropy).	$-\sum_{i=1}^{K}\sum_{j=1}^{K}p_{ij}\log_2 p_{ij}$



Normalized Co-occurrence Matrix	Descriptor					
	Max Probability	Correlation	Contrast	Uniformity	Homogeneity	Entropy
$\mathbf{G}_1/n_1$	0.00006	−0.0005	10838	0.00002	0.0366	15.75
$\mathbf{G}_2/n_2$	0.01500	0.9650	570	0.01230	0.0824	6.43
$\mathbf{G}_3/n_3$	0.06860	0.8798	1356	0.00480	0.2048	13.58

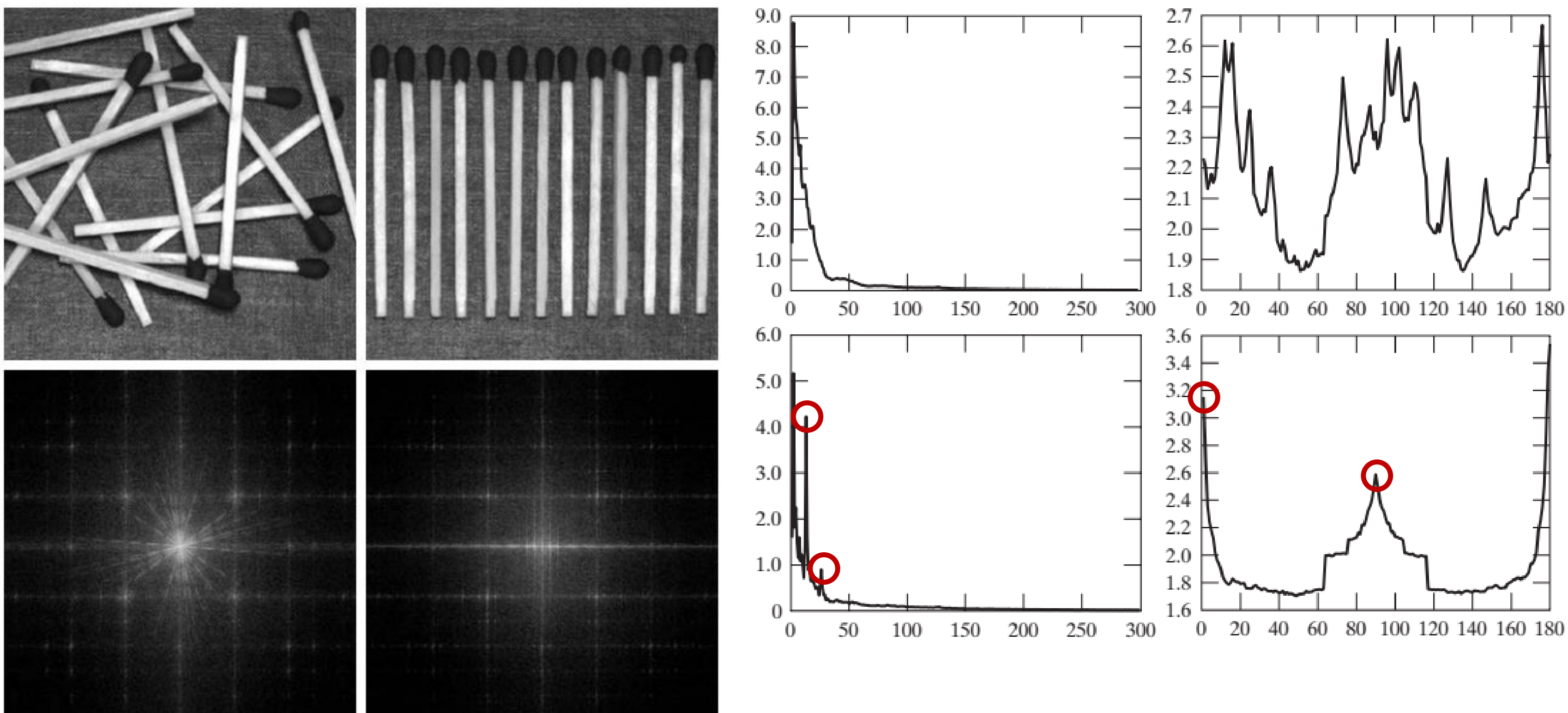
纹理描述子

- 结构法
  - 基于规则的纹理描述
  - 借助“纹理基元”（texture primitive）



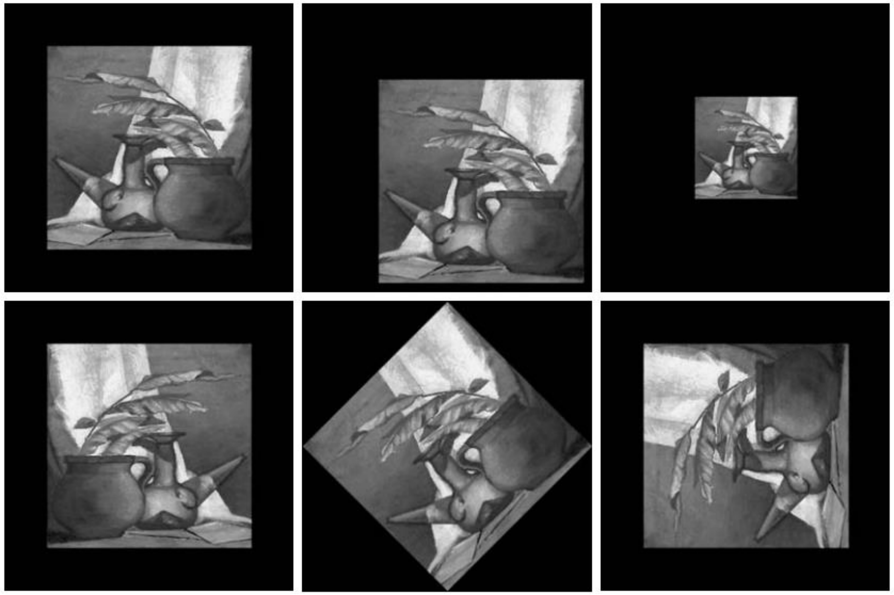
纹理描述子

- 频谱法：按极坐标积分
  - $S(r)=\sum_{\theta=0}^{\pi}S_\theta(r),S(\theta)=\sum_{r=0}^{R_0}S_r(\theta)$



区域描述子

- 不变矩
  - 对于图像的一些变换，例如平移、旋转，缩放等，是不变的；
  - 在传统的图像检索方法中非常有用



Moment Invariant	Original Image	Translated	Half Size	Mirrored	Rotated 45°	Rotated 90°
ϕ_1	2.8662	2.8662	2.8664	2.8662	2.8661	2.8662
ϕ_2	7.1265	7.1265	7.1257	7.1265	7.1266	7.1265
ϕ_3	10.4109	10.4109	10.4047	10.4109	10.4115	10.4109
ϕ_4	10.3742	10.3742	10.3719	10.3742	10.3742	10.3742
ϕ_5	21.3674	21.3674	21.3924	21.3674	21.3663	21.3674
ϕ_6	13.9417	13.9417	13.9383	13.9417	13.9417	13.9417
ϕ_7	−20.7809	−20.7809	−20.7724	20.7809	−20.7813	−20.7809

关系描述子

- 基本思想
- 阶梯关系编码
- 骨架关系编码
- 方向关系编码
- 树结构关系编码
- ……

关系描述子

- 基本思想
  - 通过挖掘各个成分之间的结构关系来描述边界
  - 图像中各个部分间的结构关系是二维的，而串是一维的，期望找到一种方法把二维关系转化为一维的串
  - 主导思想是考虑物体各个部分的连接线段

关系描述子

- 阶梯关系编码
  - $S \rightarrow aA$: 起始符S可以被图元a和变元A代替
  - $A \rightarrow bS$: 变元A可以被b和S代替
  - $A \rightarrow b$: 变元A可以只被b代替

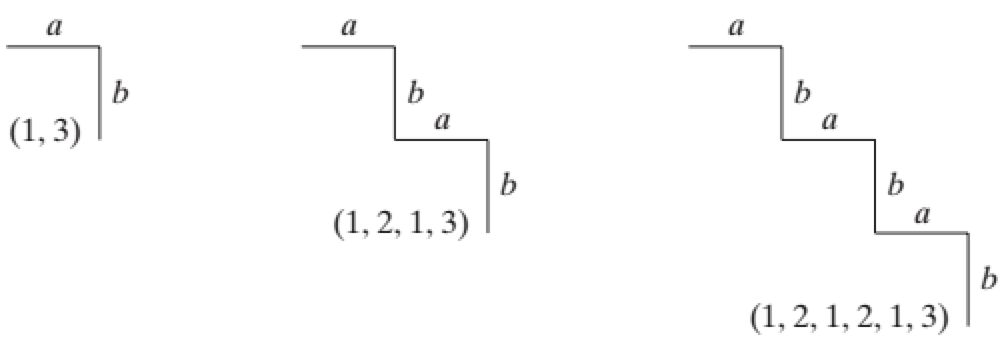
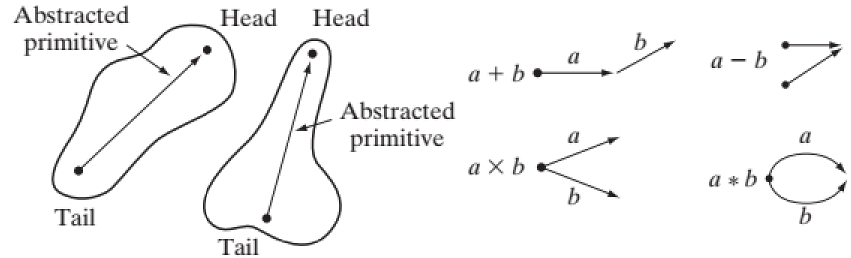
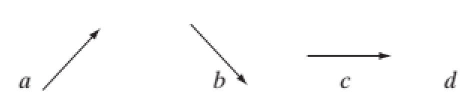


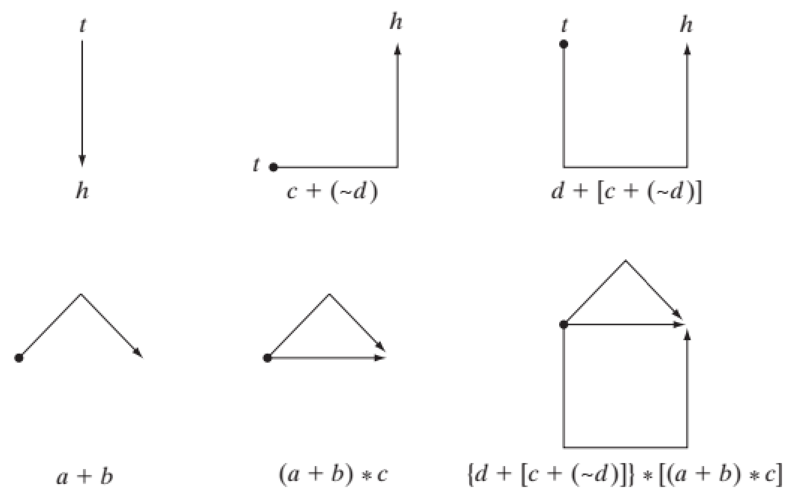
FIGURE 11.46 Sample derivations for the rules $S \rightarrow aA$, $A \rightarrow bS$, and $A \rightarrow b$.

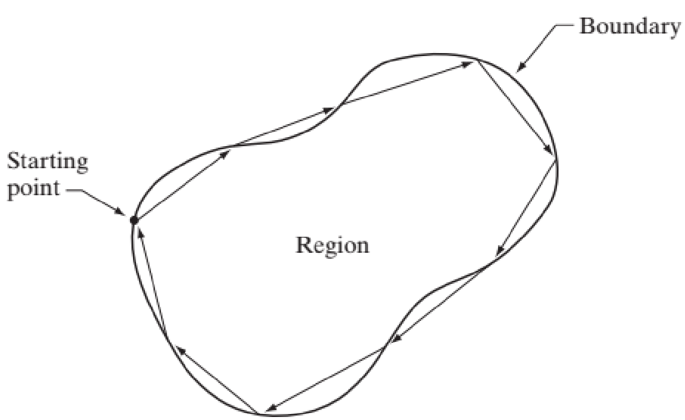
关系描述子

- 骨架关系编码
  - 用有向线段来描述一个
  - 线段之间的不同运算代
  - 适合于图像的连通性可的方式描述的时候。



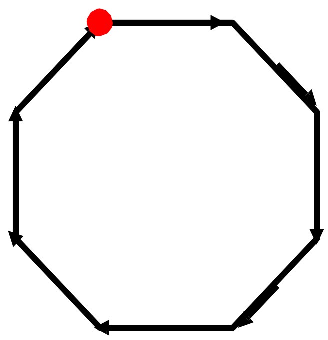




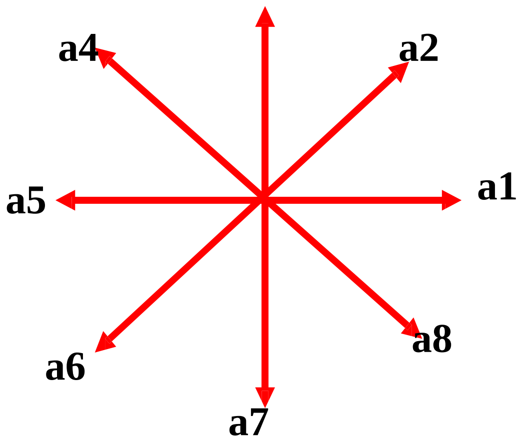


关系描述子

- 方向关系编码
  - 跟踪对象的边界，将跟踪得到的线段按照方向或长度来编码

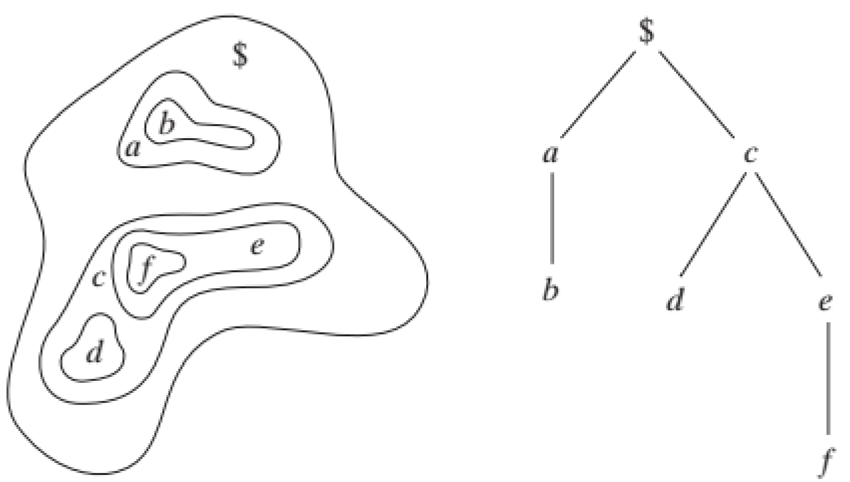


a1a8a7a6a5a4a3a2



关系描述子

- 树结构关系编码
  - 仅有一个表示根的节点\$
  - 余下节点被分成不相交的集合，$T_1, T_2, \cdots, T_m$，每个集合都是一棵树，称为T的子树；
  - 树结构中每个结点的意义和它们之间的关系最为重要



a b

FIGURE 11.50
(a) A simple composite region. (b) Tree representation obtained by using the relationship "inside of."

Any Questions?