



北京航空航天大学

# 数字图像处理

## 图像的识别与解释

---

计算机学院



## ■ 内容回顾

## ■ 图像的识别与解释



## ✓ 图像的表示与描述

- 表示与描述的基本概念
- 边界表示法设计
- 边界描述子
- 边界关系描述子



- 基本概念

- 图像分割后，为了进一步处理，分割后的图像一般要进行形式化的表达和描述
- 表达侧重于数据结构，而描述侧重于区域特性以及不同区域之间的联系和差别
- 解决形式化表达问题一般有两种选择：
  - 1) 根据区域的外部特征来进行形式化表示
  - 2) 根据区域的内部特征（比较区域内部的像素值）来进行形式化表示



- 基本概念

- 选择表达方式，要本着使数据变得更有利于下一步的计算工作（**节省存储空间、易于特征计算**）。下一步工作是基于所选的表达方式描述这个区域，一般情况下：

- 1) 如果关注的焦点是**形状特性**，选择外部表示方式
- 2) 如果关注的焦点是**反射率特性**，如颜色、纹理时，选择内部表示方式
- 3) 所选表示方式，应该对尺寸、变换、旋转等变量尽可能的不敏感

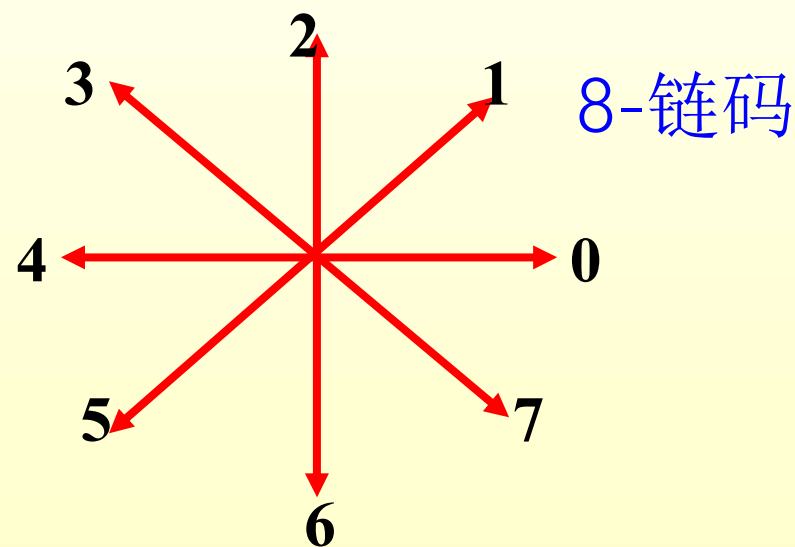
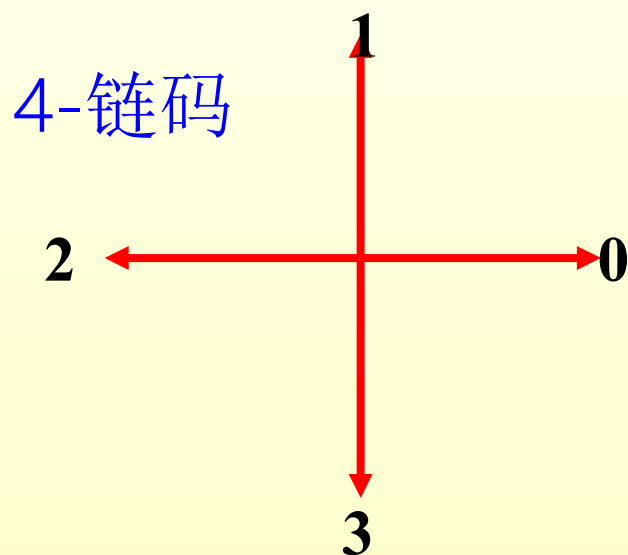


- 表示法设计
  - 链码
  - 多边形逼近
  - 外形特征
  - 边界分段
  - 区域骨架



### • 链码

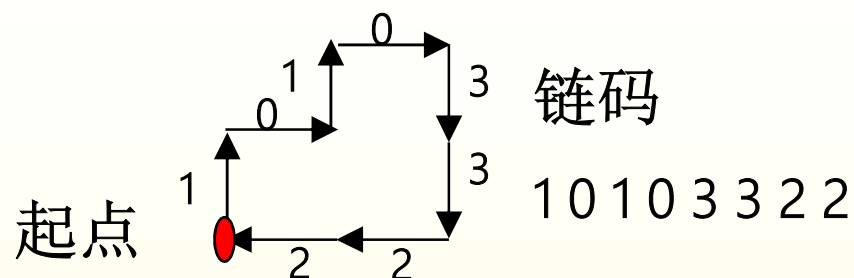
- 定义：1) 链码是一种边界的编码表示法。  
2) 用边界的方向作为编码依据。



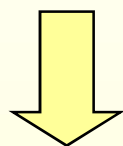


- 边界的起点需用（绝对）坐标表示，其余点都可只用连续方向来代表偏移量。
- 由于表示1个方向数比表示1个坐标值所需比特数少，而且对每1个点又只需1个方向数就可以代替2个坐标值，所以链码表达可以大大减少边界表示所需的数据量。
  - 如果目标平移，链码不会发生变化
  - 主要问题：如果目标旋转，链码会发生变化
  - 改进：通过使用链码的首差代替码子本身的方式

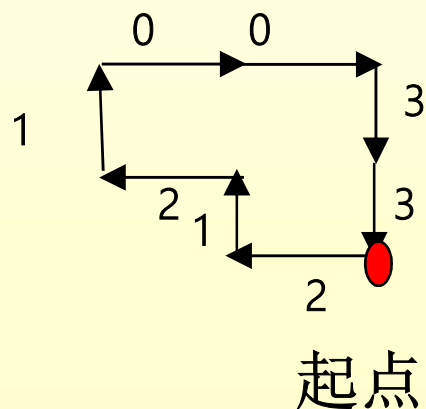




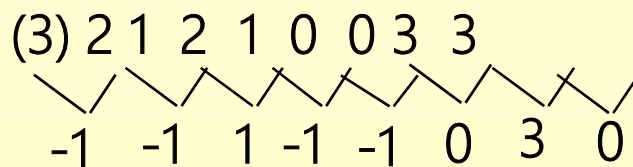
循环首差为：33133030



向左旋转90度



循环首差为：33133030





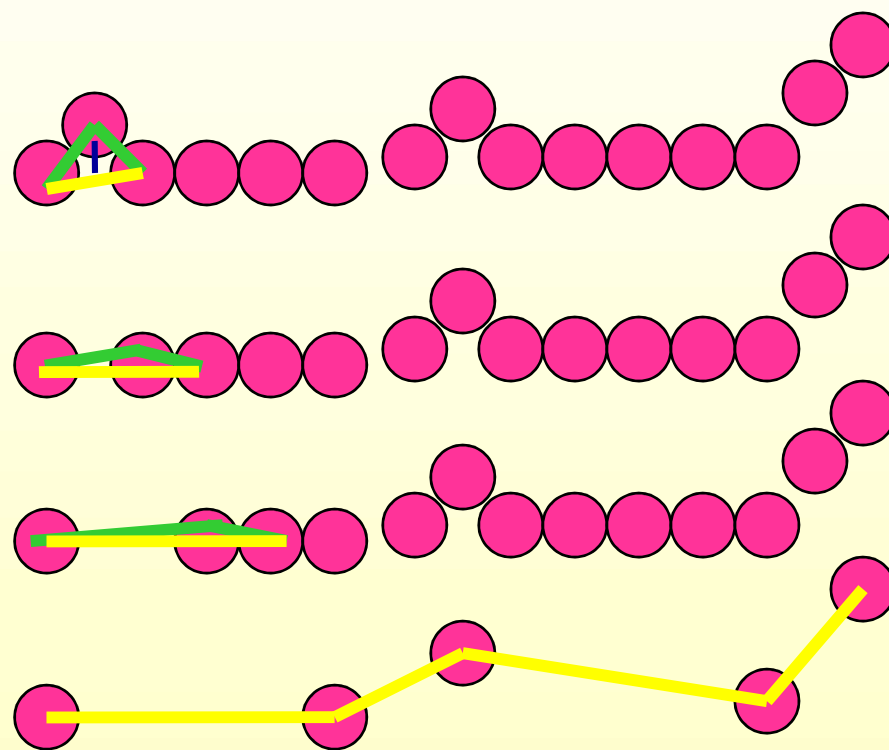
- 多边形逼近

- 在数字图像中，如果多边形的线段数与边界上的点数相等，则多边形可以完全准确地表达边界。
- 基本思想：用最少的多边形线段，获取边界形状的本质。
- 寻找最小基本多边形的方法一般有两种：  
点合成法和边分裂法



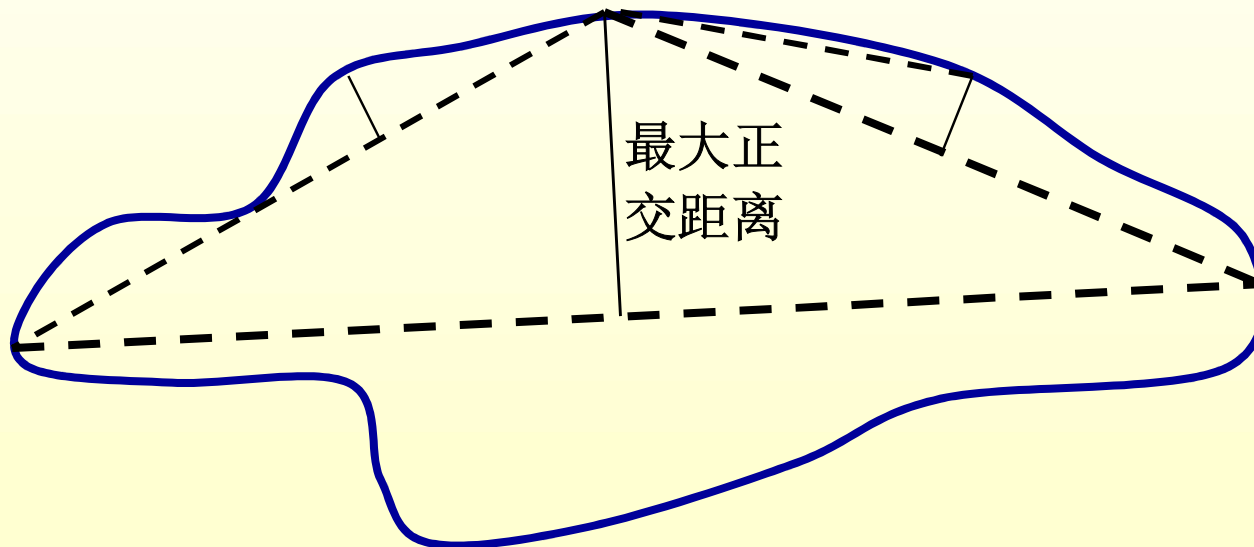
- 多边形逼近

— 点合成算法思想举例：





- 多边形逼近
  - 边分裂算法思想举例：



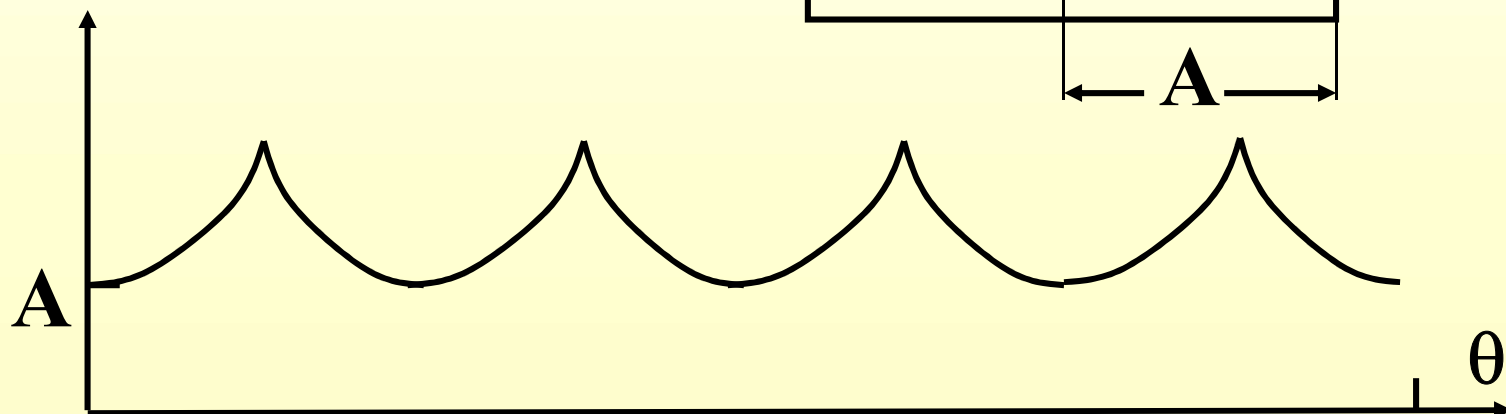


## • 外形特征

### — 基本思想：

外形特征是一种用一维函数表达边界的方法。基本思想是把边界的表示降到一维函数。

$$r(\theta) = A * \sec(\theta)$$





- 外形特征

- 问题：函数过分依赖于旋转和比例的变化

- 改进：

- 对于旋转——两种改进：

- a. 选择离质心最远的点作为起点

- b. 选择从质心到主轴最远的点作为起点

唯一的起点

- 对于比例变换：

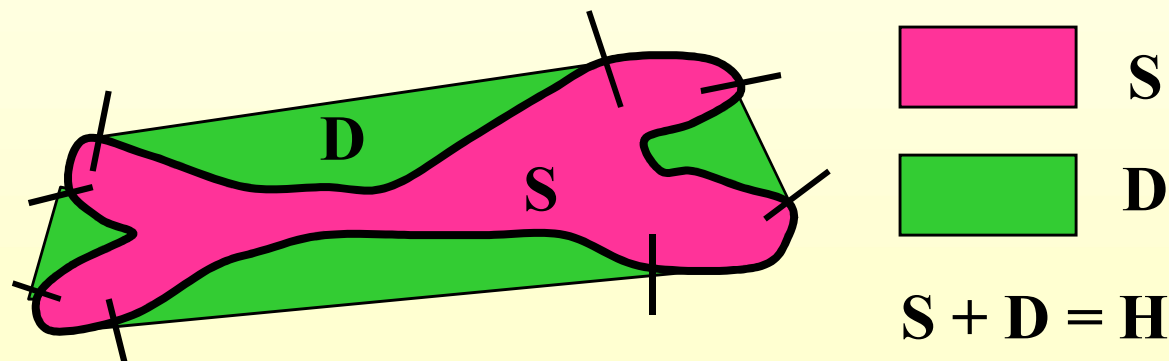
- 对函数进行正则化，使函数值总是分布在相同的值域里，比如说 $[0, 1]$



- 边界分段

- 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

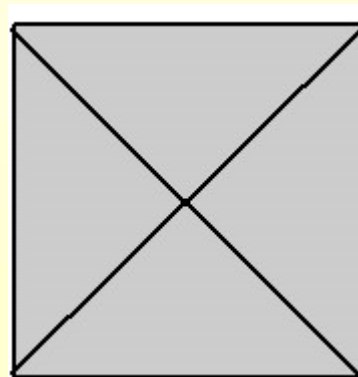
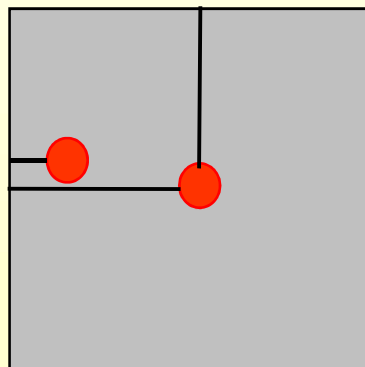




# 北京航空航天大学 表示法设计 - 区域骨架

## • 区域骨架

- 针对边界B的区域R，中轴变换方法(MAT)如下：
- 对每个R中的点P，在R中搜寻与它最近的点，如果对P能找到多于一个这样的点（即有两个或两个以上的B中的点）与P同时最近，就可以认为P属于R的骨架

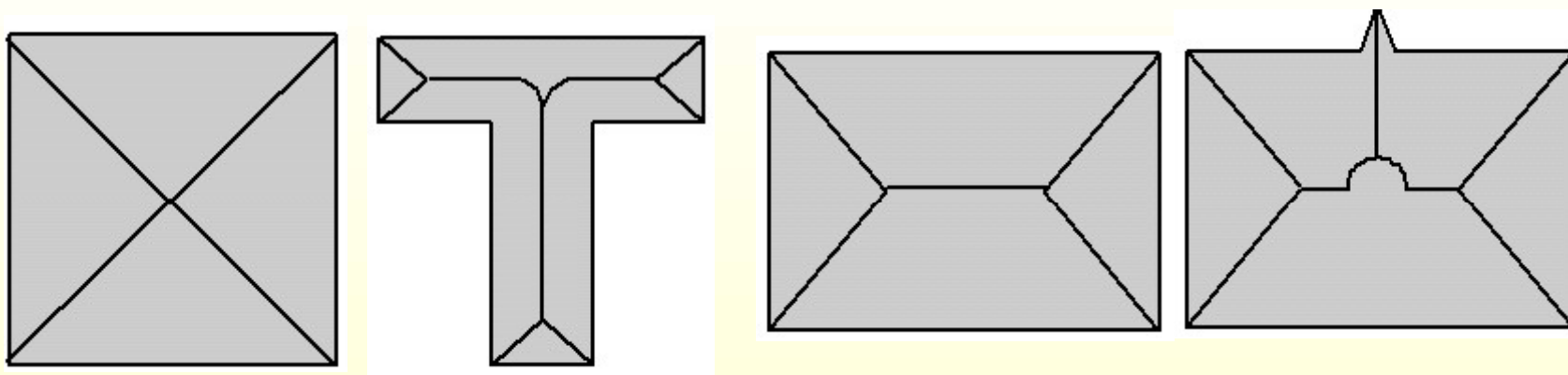






## • 区域骨架

### • 中轴变换方法(MAT)



骨架上的点P能在边界上找到至少2个最近点和P距离相等！

迭代细化算法，改进算法的效率：两个基本操作

(1) 不移去端点； (2) 不破坏连通性； (3) 不引起区域的过度腐蚀



- 边界描述子
  - 简单描述子
  - 形状数
  - 傅立叶描述子
  - 矩量



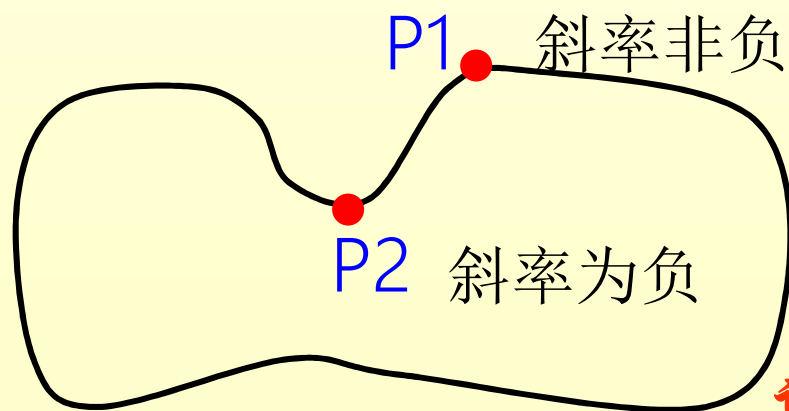
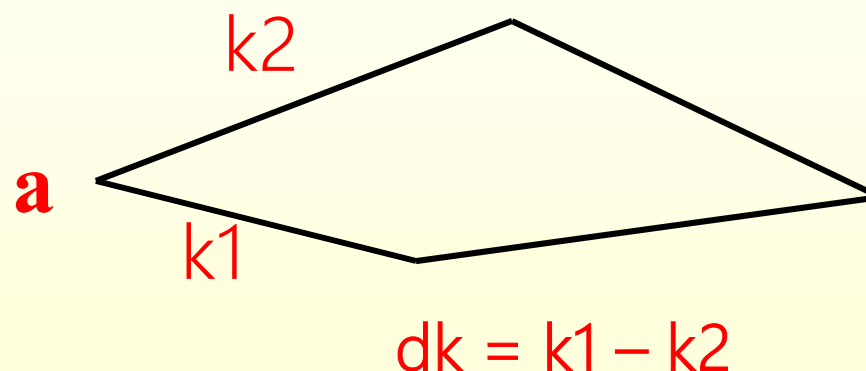
# 北京航空航天大学——边界描述子——简单描述子

- 简单描述子

- 边界的周长
- 边界的直径
- 边界的曲率
- 边界的凸线段点P1
- 边界的凹线段点P2


$$\text{Diam}(B) = \max[D(p_i, p_j)]$$

沿轮廓线计算像素的个数，给出一个长度的近似估计



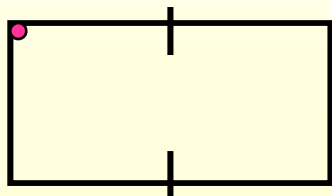


# 北京航空航天大学3—边界描述子—形状数

- 形状数

- 最小循环首差链码

阶数6

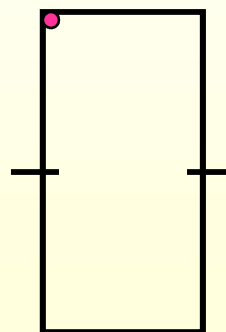


链码：003221

首差：303303

形状数：033033

阶数6



链码：033211

首差：330330

形状数：033033



## • 傅立叶描述子

- (1) 对于XY平面上的每个边界点，将其坐标用复数表示
- (2) 进行离散傅立叶变换

$$a(u) = \frac{1}{N} \sum_{k=0}^{N-1} s(k) \exp(-j2\pi uk/N) \quad u=0,1,\dots,N-1$$

$a(u)$ 被称为边界的傅立叶描述子

- (3) 选取整数  $M \leq N-1$ ，进行逆傅立叶变换（重构）

$$s'(k) = \sum_{u=0}^{M-1} a(u) \exp(j2\pi uk/N) \quad k=0,1,\dots,N-1$$



- 傅立叶描述符

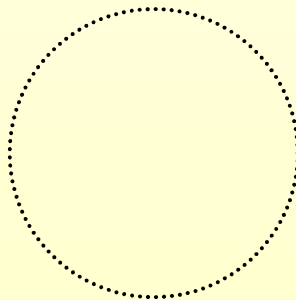
- $M$  的选取与描述符的关系

在上述方法中，相当于对于  $u > M-1$  的部分舍去不予计算。由于傅立叶变换中高频部分对应于图像的细节描述，因此  $M$  取得越小，细节部分丢失得越多。

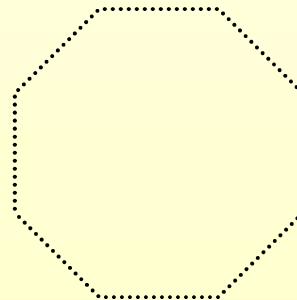
只用一些低频分量的傅立叶系数来近似描绘边界形状。



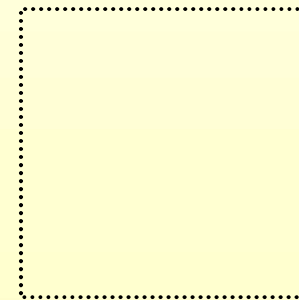
$N=64$



$M=2$



$M=32$



$M=62$



- 矩和边缘矩
- 对旋转不敏感

f(r)的均值 
$$m = \sum_{i=1}^L r_i f(r_i)$$

f(r)对均值的n阶矩 
$$\mu_n(r) = \sum_{i=1}^L (r_i - m)^n f(r_i)$$



### • 关系描述子

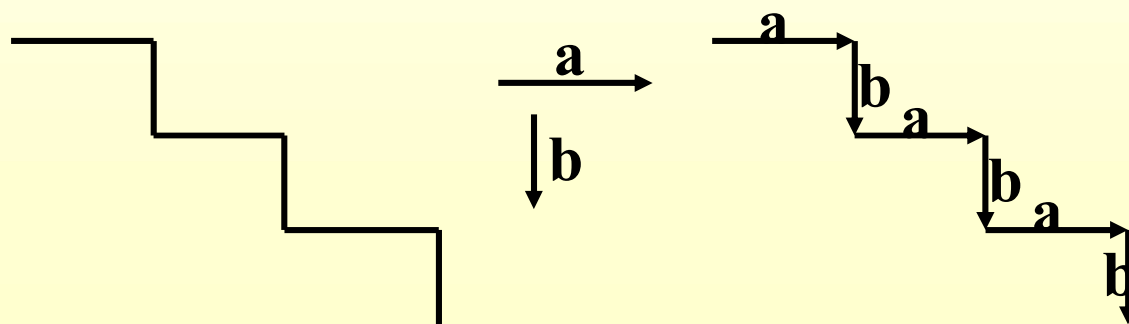
— 基本思想：通过挖掘各个成分之间的结构关系来描述边界。图像中各个部分间的结构关系是二维的，而串是一维的，期望找到一种方法把二维关系转化为一维的串。主导思想是考虑各个部分的连接线段。

— 阶梯关系编码：阶梯形边界

— 骨架关系编码

— 方向关系编码

— 树结构关系编码

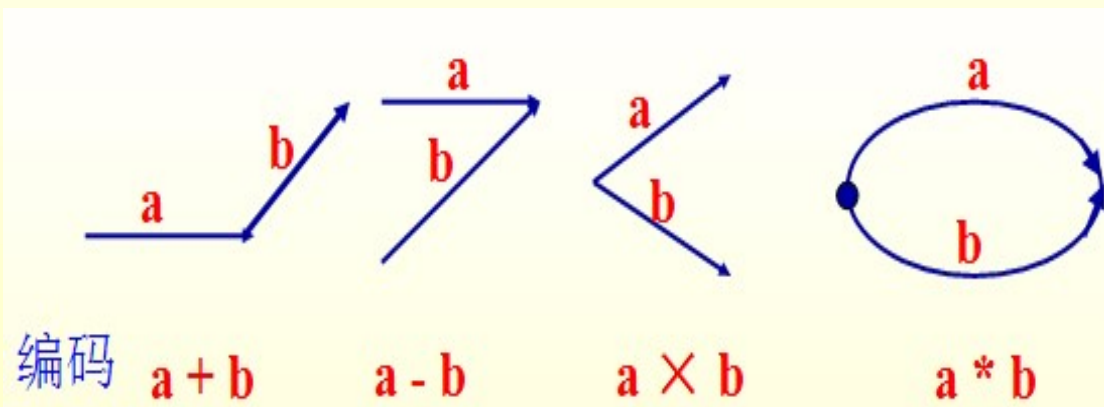






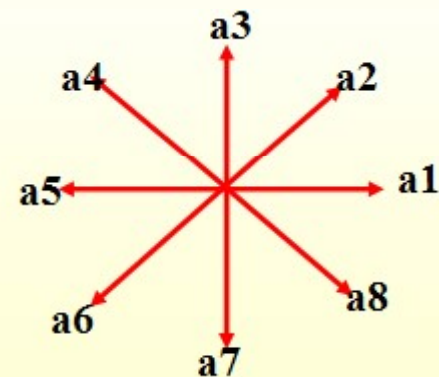
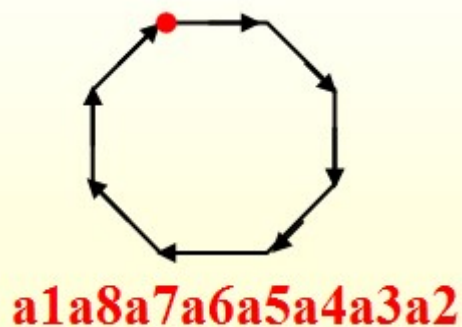
- 关系描述子

- 基本思想
- 阶梯关系编码
- 骨架关系编码：用有向线段来描述一个图像的各个部分，这个线段是通过头尾连接等方法得到的。线段之间的不同运算代表了区域的不同组合。
- 方向关系编码
- 树结构关系编码



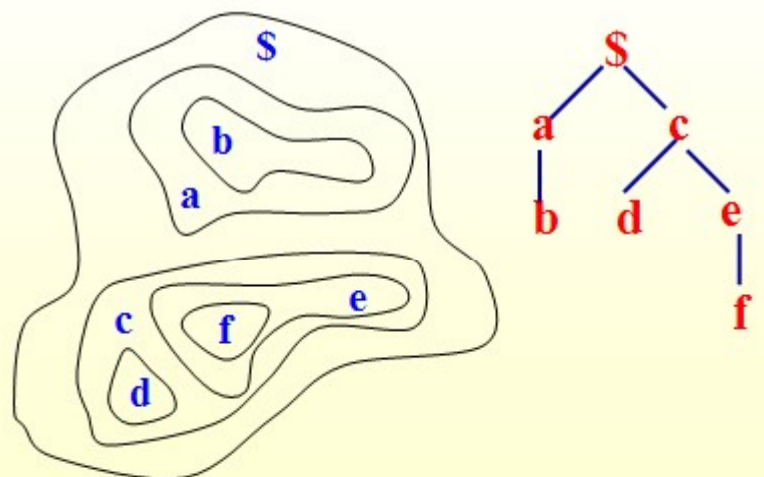


- 关系描述子
  - 基本思想
  - 阶梯关系编码
  - 骨架关系编码
  - 方向关系编码
  - 树结构关系编码





- 关系描述子
  - 基本思想
  - 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





北京航空航天大学

# 数字图像处理

## 图像的识别与解释

---

计算机学院



- 图像分析理解引言
- 图像工程的组成
- 模式与模式类
- 决策论法
- 结构法



- 图像分析理解引言
  - 图像分析理解的定义
  - 图像分析理解的目标
  - 自动图像分析理解系统的概念化分类
  - 自动图像分析理解系统的现状



- 图像分析的定义

- 图像分析是一个发现、辨认和理解模式的过程，这些模式都与执行和图像相关的任务有关。

- 图像理解的定义

- 图像理解是在图像分析的基础上，结合人工智能和认知理论，进一步研究图像中各目标的性质和它们之间的相互联系，并理解图像内容的含义以及解释原来的客观场景，从而指导和规划行为。



- 图像分析理解的目标

- 计算机图像分析理解的主要目的之一是，赋予某些具有感觉能力的机器，以类似人的大脑的能力。例如OCR（Optical Character Recognition光学字符识别，是指电子设备（例如扫描仪或数码相机）检查纸上打印的字符，通过检测暗、亮的模式确定其形状，然后用字符识别方法将形状翻译成计算机文字的过程）





- 自动图像分析理解系统的概念化分类
  - 获取、发现信息：从背景中提取有关信息。
  - 学习、应用知识：抽象、归纳信息特征的学习过程，并应用到新的对象中。
  - 构造、推理知识：从不完整的信息中构造推论出新的知识，并加以应用。

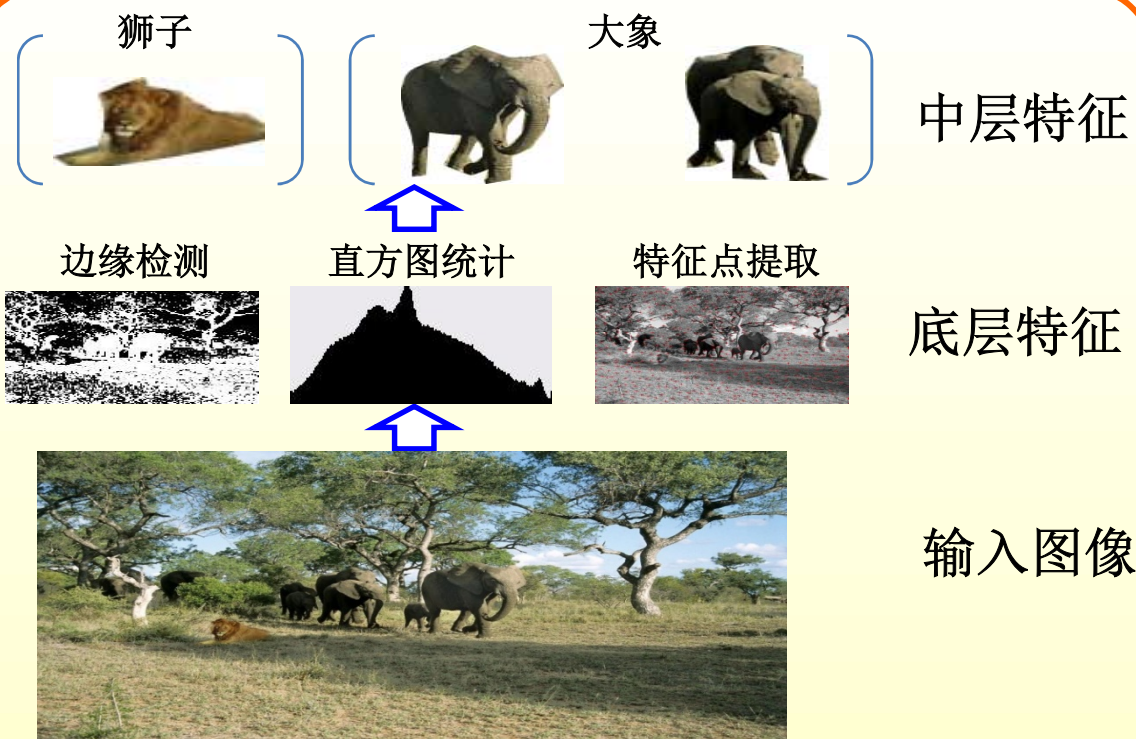


北京航空航天大学

# 图像分析理解引言

学习、应用知识

获取、发现信息



高层语义：

非洲场景  
动物世界

构造、推理知识



- 自动图像分析理解系统的现状
  - 我们可以设计出初级的这种系统，但仍然缺乏理论依据。有待人类视觉认识理论的进一步研究。
  - 我们可以做出在某一应用上超过人的能力的系统，但缺乏扩展性。过分依赖应用。



您搜索的图片可能是：  
**唐装旗袍**    **短袖**  
图片尺寸：234X505

百度图搜图

服饰搜索



¥439.00

京东商城

Sc...季新款  
短...



¥119.00

京东商城

花图+夏装新款女装优雅  
气质唐装旗袍...



¥940.00

京东商城

舒朗+2015夏季新款女  
装通勤气质...



¥398.00

京东商城

曼薇琪2015夏装新款+  
修身短袖图...



- 图像工程分类的三种基本范畴

- 低级处理：

- 图像获取、预处理，不需要智能（后期研究结合智能）

- 中级处理：

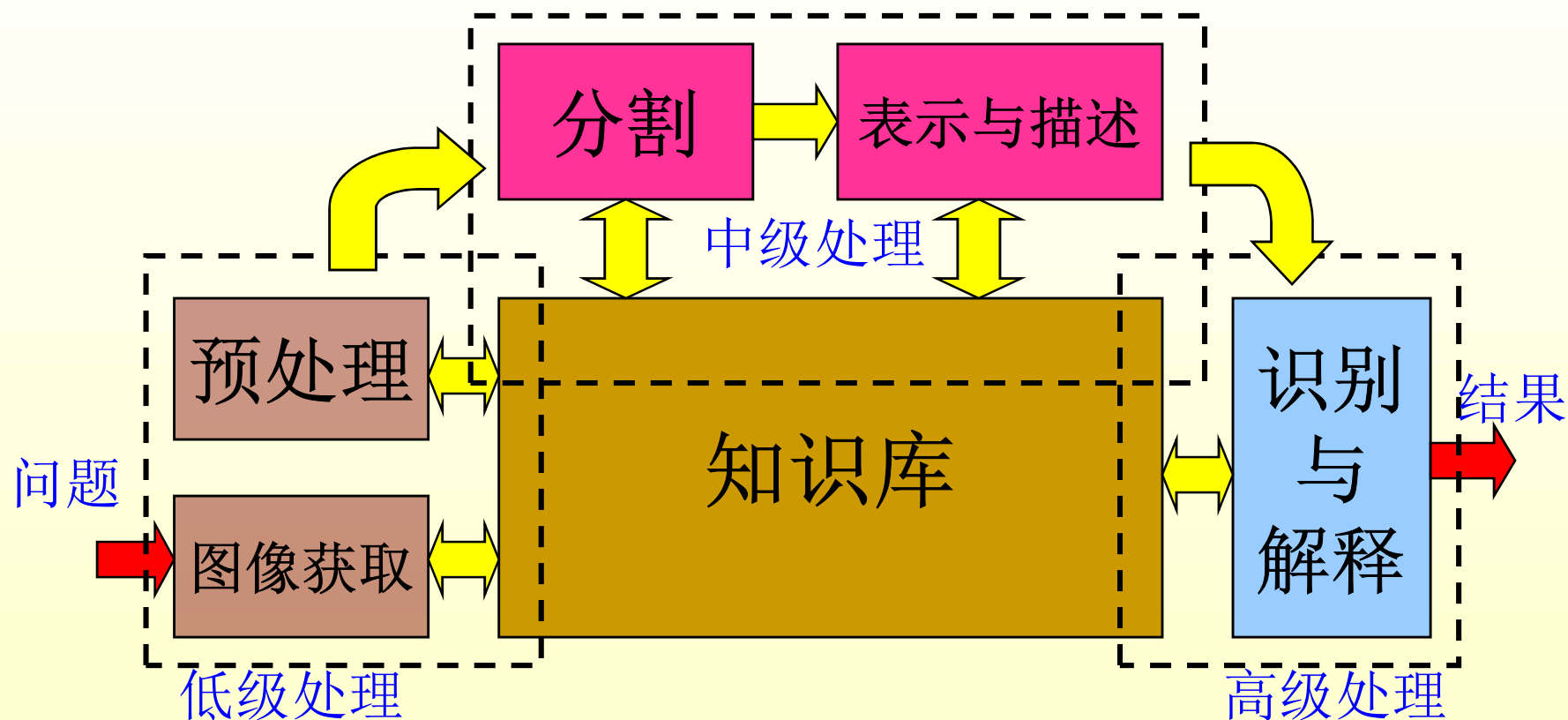
- 图像分割、表示与描述，需要智能

- 高级处理：

- 图像识别、解释，缺少理论，为降低难度，设计得更专用



- 图像工程分类的三种基本范畴





- 图像识别与解释的基本方法

- 识别的统计分类方法：

- 用向量形式表达模式，分派模式向量到不同的模式类。

- 识别的结构方法：

- 用符号匹配，模式被表示为符号形式（如形状数、串和树）。

- 图像解释的方法：

- 图像解释技术是基于谓词逻辑、语义网络和特定产品的系统。



- 模式与模式类
  - 模式的定义
  - 模式类的定义
  - 模式识别的定义
  - 常用的模式序列
    - 模式向量
    - 模式串
    - 模式树



- 模式的定义

- 模式包括的范围很广，图像就是模式的一种。
- 本课程中模式是图像中的一个对象或某些感兴趣内容的数量或结构的描述。
- 模式是由一个或多个描述子来组成，换句话说，模式是一个描述子的序列（名词“特征”经常被用来代指描述子）。
- 模式是一组特征或一组描述子。





- 模式识别

- 识别是指从客观事实中自动建立符号描述或进行逻辑推理的数学和技术，因而人们定义模式识别是对客观世界中物体和过程进行分类、描述的学科。
- 目前，对图像模式的识别主要集中在对图像中感兴趣的内容（目标）的分类、分析和描述。



- 模式类的定义
  - 模式类是具有某些公共特征的模式系列
  - 模式类用 $w_1, w_2, \dots, w_M$ 表示， $M$ 是类的个数
- 模式识别的定义
  - 根据图像中对象的特征组成的模式，确定对象是属于哪一个模式类，即为模式识别
- 模式与模式类举例
  - 已知汽车的长、宽、高（ $x_1, x_2, x_3$ ），希望识别出，大客车、小轿车、卡车（ $w_1, w_2, w_3$ ）



- 常用的模式序列

定量描述

结构描述

- 模式向量、模式串、模式树

- 模式向量

- 模式向量用粗体的小写字母表示，如 $x, y$ 形式如下：

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

其中每一个 $x_i$ 代表第 $i$ 个描述子， $n$ 是这种描述子的数量。模式向量被表示为一列或表示成

$$x = (x_1, x_2, \dots, x_n)^T,$$



### —模式向量举例

假设我们想描述三种蝴蝶花（多毛的、维吉尼亚、多色的），通过测量它们花瓣的宽度和长度。这里涉及一个两维的模式向量：

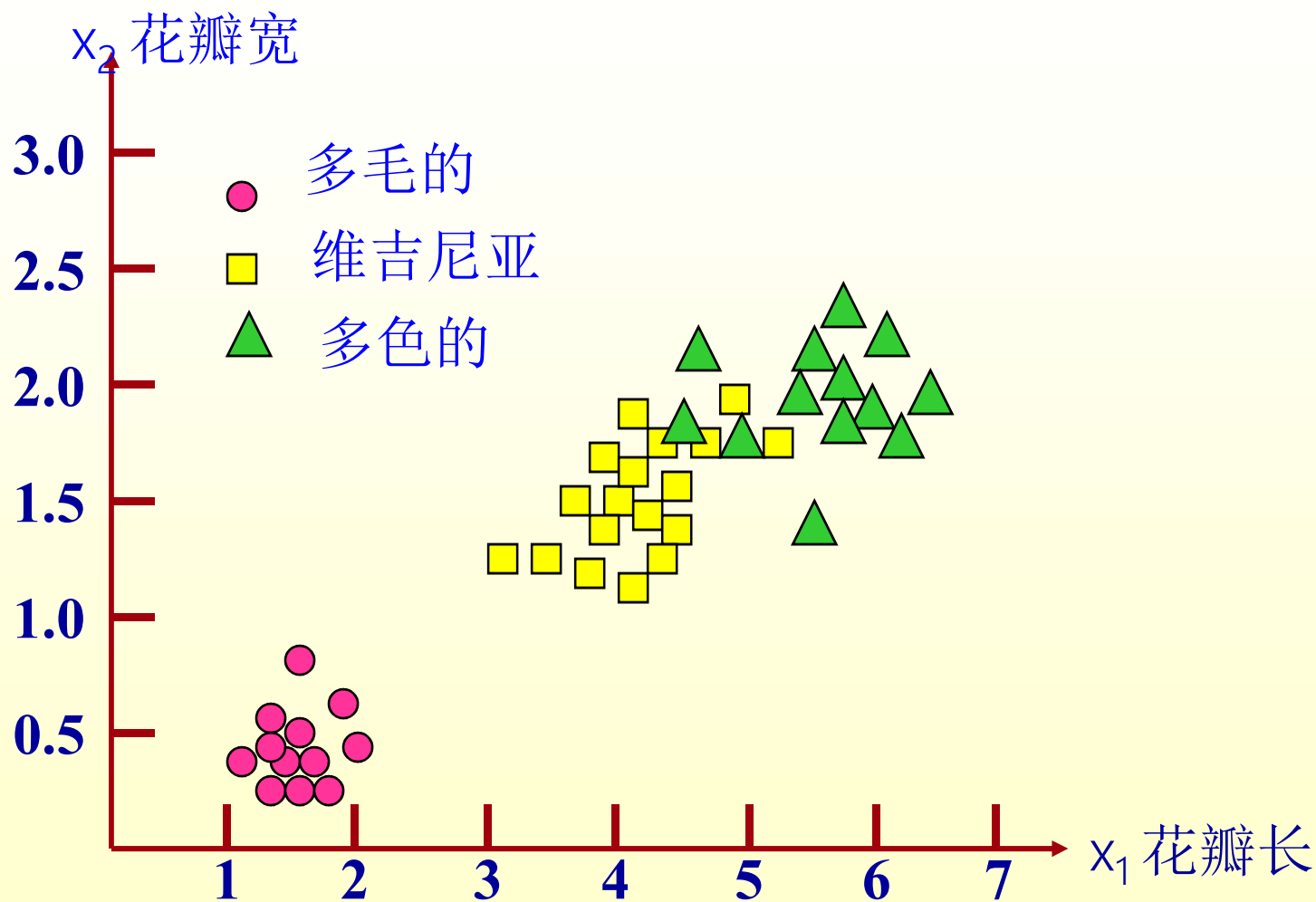
$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

其中 $x_1$ 、 $x_2$ 分别对应花瓣的长和宽  
三种模式类用 $w_1$ 、 $w_2$ 、 $w_3$ 表示



由于所有的蝴蝶花瓣在宽和长上都有某种程度的变化，所以描述这些花瓣的模式向量也将有变化，不仅在不同的类之间，而且也在类的内部。

在这种情况下每一种蝴蝶花瓣变成二维欧几里德空间的一个点。





## 一对结果的分析

对花瓣长宽的测量，成功地将多毛的蝴蝶花与其它两种分离，但对于分离维吉尼亚和多色的是失败的。

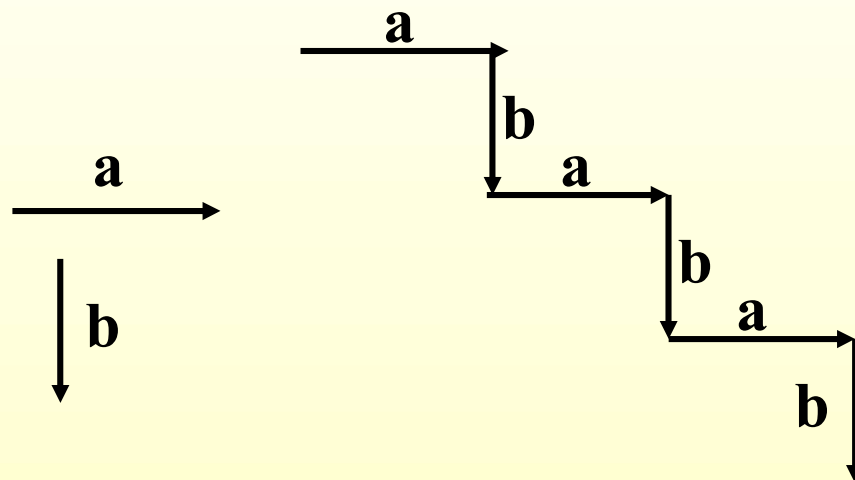
这个结论说明了分类的特性选择问题，在这个问题中，类的可分离性的程度，完全依赖于对模式尺寸测量的选择。



### —模式串

用于以对象特征的**结构或空间关系**作为模式的识别。

### —模式串举例：梯状的模式



规定每个元素都头尾相连，而且允许它们交替出现，可把握这类模式的基本结构。这种结构化的构造法可适用于任何长度的阶梯状模式，其他一些形式的结构也可依靠对a和b的不同组合得到。

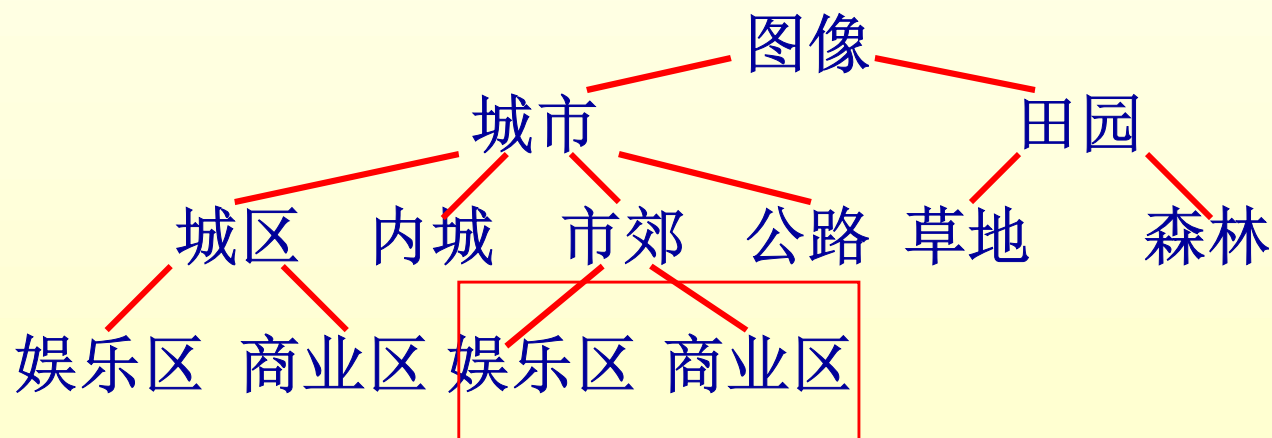




### 一模式树

以分层目录结构排序的模式类，一般多采用树结构。

模式树举例





对模式的分类主要基于决策理论

- 决策论法
  - 决策论法的基本概念
  - 匹配
    - 最小距离分类器
    - 相关匹配



- 决策论法的基本概念

- 决策论识别法的定义

设：模式向量： $x = (x_1, x_2, \dots, x_n)^T$ ，

对于：M个模式类  $w_1, w_2, \dots, w_M$ ，

寻找M个决策函数  $d_1(x), d_2(x), \dots, d_M(x)$ ，具有这样的特性：如果模式  $x$  属于模式类  $w_i$ ，那么：

$$d_i(x) > d_j(x) \quad j = 1, 2, \dots, M; j \neq i$$

换句话说，如果一个未知模式  $x$  属于第  $i$  个模式类，把  $x$  代入所有的决策函数， $d_i(x)$  的数值最大。



## • 决策论法的基本概念

### — 决策论识别法的定义

设：模式向量： $x = (x_1, x_2, \dots, x_n)^T$ ，

对于：M个模式类  $w_1, w_2, \dots, w_M$ ，

M个决策函数  $d_1(x), d_2(x), \dots, d_M(x)$ ：

$d_i(x) > d_j(x) \quad j = 1, 2, \dots, M; j \neq i$



$x = (R:120, G:19, B:7)$

$w_1 = (255, 0, 0), w_2 = (0, 255, 0), w_3 = (0, 0, 255)$

$d_1 = 1/(|255-120| + |0-19| + |0-7|) = 0.0062$

$d_2 = 1/(|0-120| + |255-19| + |0-7|) = 0.0028$

$d_3 = 1/(|0-120| + |0-19| + |255-7|) = 0.0026$





- 决策论法的基本概念

- 决策边界的定义

- 对于模式 $x$ ，如果决策函数值有：

$$d_i(x) - d_j(x) = 0$$

此 $x$ 值，被称为 $w_i$ 与 $w_j$ 的决策边界。

- 通常用一个单一的函数标识两个类之间的决策边界，定义为： $d_{ij}(x) = d_i(x) - d_j(x) = 0$

如果  $d_{ij}(x) > 0$   $x$  属于类 $w_i$

如果  $d_{ij}(x) < 0$   $x$  属于类 $w_j$



- 基于决策函数可设计多种分类器。
- 统计模式识别中根据统计参数来分类，一般将用来估计统计参数的（已知其类别的）模式称为训练模式，将一组这样的模式称为训练集，将用一个训练集去获取决策函数的过程称为学习或训练。



- 匹配

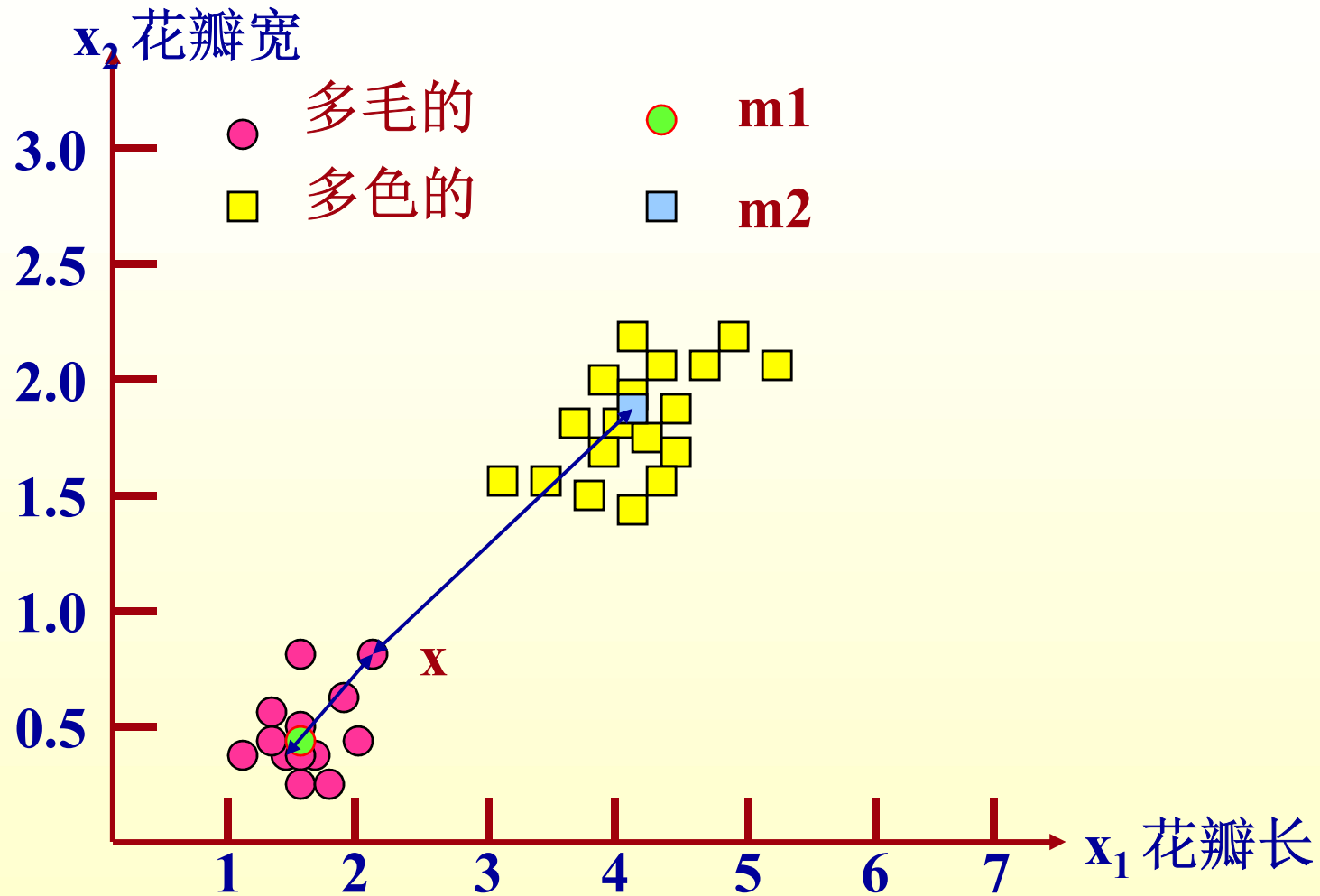
- 最小距离分类器

- 基于对模式的采样来估计各类模式的统计参数，完全由各类的均值和方差确定。

以蝴蝶花的例子为例：

〈1〉为多色 ( $w_1$ ) 和多毛 ( $w_2$ ) 的两种蝴蝶花，确定两个原形（或称模板） $m_1$  和  $m_2$ 。

〈2〉对于一个未知模式向量  $x$ ，判断  $x$  与  $m_1$  和  $m_2$  的距离，如果与  $m_1$  的距离小于与  $m_2$  的距离，则  $x$  属于  $w_1$ ，否则属于  $w_2$ 。







- 匹配

- 最小距离分类器

- (1) 算法思想：

- 对于M个模式类  $w_j$   $j = 1, 2, \dots, M$   
为每一个模式类确定一个原形模式向量  $m_j$
      - 对于一个未知模式向量  $x$ ，如果  $x$  与  $m_i$  的距离最小，就称  $x$  属于  $w_i$ 。



- 匹配

- 最小距离分类器

(2) 最小距离分类器定义：

Step1—计算模式类 $w_j$ 的原形向量：

$$m_j = 1/N_j \sum_{x \in w_j} x \quad j = 1, 2, \dots, M$$

其中 $N_j$ 是属于模式类 $w_j$ 的模式向量的个数。

通过计算已知属于 $w_j$ 的模式向量的各分量的均值得到原形模式向量 $m_j$



- 匹配

- 最小距离分类器

Step2——计算 $x$  与  $m_j$ 的距离

$$d_j(x) = \|x - m_j\| \quad j = 1, 2, \dots, M$$

Step3——决策

如果 ,  $d_i(x) = \min(d_j(x)) \quad j = 1, 2, \dots, M$

就说 :  $x$  属于  $w_i$



$$d_j(x) = \|x - m_j\|, \quad j = 1, 2, \dots, M$$

其中 $\|a\| = (a^T a)^{1/2}$ 是欧氏距离。可以证明这等价于计算：

$$d_j(x) = x^T m_j - 1/2 m_j^T m_j, \quad j = 1, 2, \dots, M$$

在 $d_j(x)$ 给出最大值时将 $x$ 赋给模式类 $w_i$



- 匹配

- 最小距离分类器

改写成求最大的标准形式，决策函数为：

$$d_j(x) = x^T m_j - 1/2 m_j^T m_j \quad j = 1, 2, \dots, M$$

如果， $d_i(x) = \max(d_j(x)) \quad j = 1, 2, \dots, M$

就说：x 属于  $w_i$

Step4——类  $w_i$  和  $w_j$  之间的决策边界是：

$$\begin{aligned} d_{ij}(x) &= d_i(x) - d_j(x) \\ &= x^T (m_i - m_j) - 1/2 (m_i - m_j)^T (m_i + m_j) = 0 \end{aligned}$$

上式实际上给出一个连接  $m_i$  和  $m_j$  线段的垂直二分界。

线 -> 平面 -> 超平面

$n=2, n=3, n>3$



(3) 举例：

多色的和多毛的蝴蝶花，用 $w_1$ 和 $w_2$ 分别表示，  
有简单的原形向量

$$m_1 = (4.3, 1.3)^T \quad m_2 = (1.5, 0.3)^T$$

决策函数是：

$$d_1(x) = x^T m_1 - 1/2 m_1^T m_1 = 4.3x_1 + 1.3x_2 - 10.09$$

$$d_2(x) = x^T m_2 - 1/2 m_2^T m_2 = 1.5x_1 + 0.3x_2 - 1.17$$

决策边界的等式：

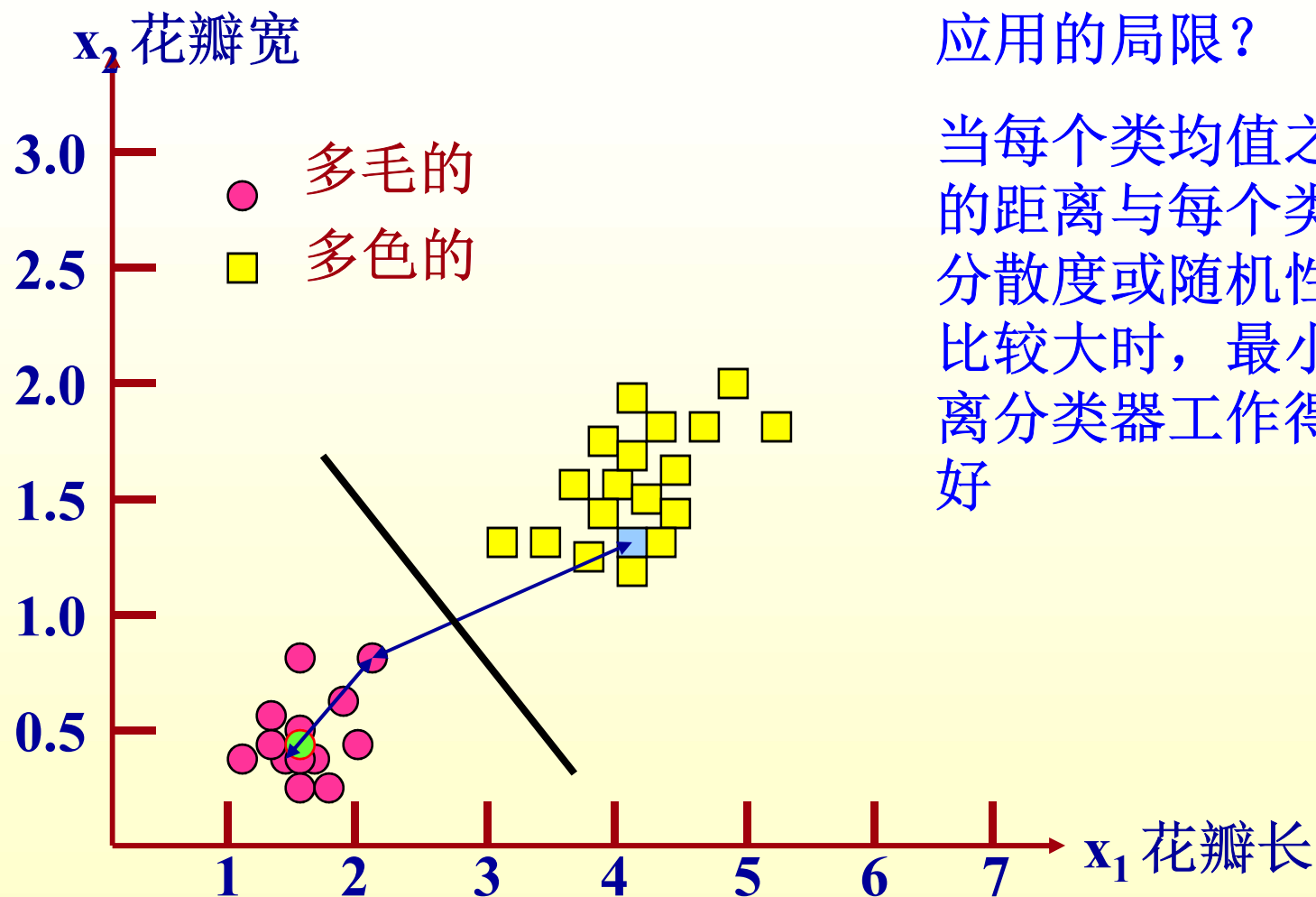
$$d_{12}(x) = d_1(x) - d_2(x) = 2.8x_1 + 1.0x_2 - 8.92 = 0$$



- 边界方程为

$$d_{12}(x) = 2.8x_1 + 1.0x_2 - 8.92 = 0$$

- 将属于模式类w1中的任一个模式代入将使  $d_{12}(x) < 0$ 。反之，将属于模式类w2中的任一个模式代入将使  $d_{12}(x) > 0$ 。



应用的局限？

当每个类均值之间的距离与每个类的分散度或随机性相比较大时，最小距离分类器工作得很好

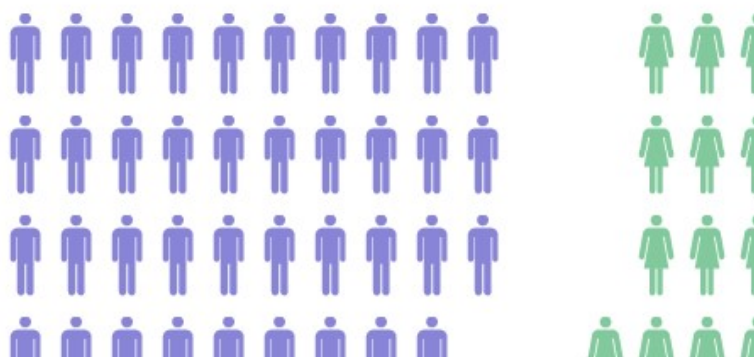




□ 问题的引出：市场分析员如何从海量、杂乱数据中发现不同的客户群？

阿里指数

性别占比



年龄阶段占比



排名	搜索词	搜索指数
1	连衣裙	29,189
2	爱奇艺会员	23,766
3	小白鞋	20,686
4	手机	19,843
5	外套女春	19,281



北京航空航天大学

# 聚类算法



objects in cluster 1



objects in cluster 2



计算机学院



- **聚类**：根据客观属性对一系列**未分类的个体进行类别的识别**，把一组个体按照相似性归纳成若干类。聚类属于**无监督学习**。
- **聚类问题的数学描述**：给定数据集合 $X$ ，根据数据对象间的相似程度将数据集合分组，并满足：

$$\{C_j | j = 1, 2, \dots, k\}$$

$$C_j \subseteq X$$

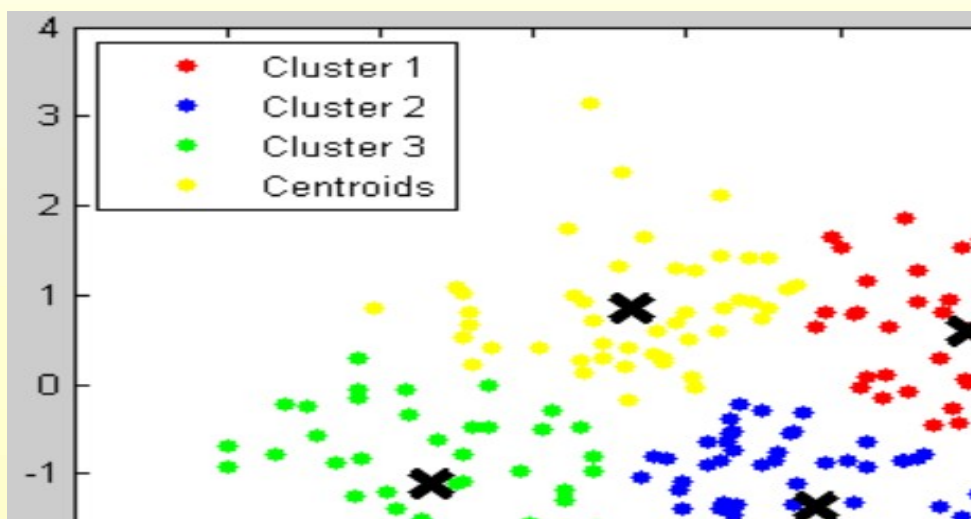
$$C_i \cap C_j = \emptyset$$

$$\bigcup_{i=1}^k C_i = X$$

则该过程称为聚类， $C_i$  称为簇。



- **聚类分析算法**：划分法、层次法、基于密度的方法、基于网格的方法、基于模型的方法
- **K-means 算法**—划分法(最著名和常用的划分方法，易于学生理解聚类的思想)
- **K-means 算法概述**： **牧师-居民问题**





## □ K-means 算法概述

- 基于误差平方和准则进行动态聚类，以K为参数，把N个样本分为K个簇，以使簇内具有较高的相似度，而簇间的相似度较低。
- K: 聚类算法中类的个数
- Means: 均值算法（样本中心）
- K-means: 利用均值算法将样本分为k个类



## □ 算法设计思路

1. 随机抽取  $k$  个样本点作为初始聚类的中心，由各中心代表各聚类

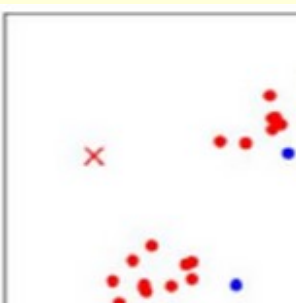
2. 计算所有点到这  $k$  个中心的聚类，并将点归到离其最近的聚类

3. 调整聚类中心，即将聚类中心移到聚类的几何中心（即平均值）

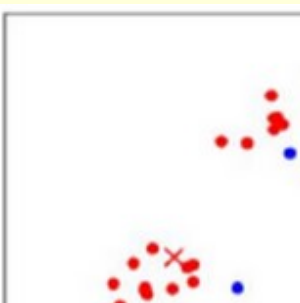
重复第2, 3步，直到聚类中心不再移动，此时算法收敛



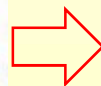
初始选择  
中心点



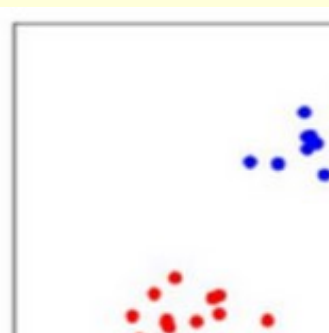
各点划分进  
最近聚类



重新计算  
聚类中心



迭代计算中心点



收敛



### □ 算法流程

从数据集  $X$  中任意选取  $K$  个赋给初始的聚类中心  $C_1, C_2, \dots, C_K$

对各样本点，计算其与各聚类中心的欧式距离并获取其类别标号  $\text{Label}(i) = \arg \min_j \|x_i - c_j\|^2$

重新计算  $K$  个聚类中心  $c_j = \sum_{s: \text{Label}(s)=j} X_s / N_j, j = 1, 2, \dots, K$

$$I_k = \sum_{j=1}^k \sum_{i=1}^{n_j} \|x_i - c_j\|^2$$

$$c_j = \frac{1}{n_j} \sum_{i=1}^{n_j} x_i, j = 1, 2, \dots, K$$

K均值聚类算法在初始划分的基础上，使用迭代算法逐步优化聚类结果，使准则函数  $I_k$  达到极小值，获得  $K$  个类型。



### 影响聚类结果的几个因素

#### 初始中心点

- 随机选点
- 经验选有代表性的点
- 基于取样的方法确定
- 基于密度的选择方法

#### K的取值

- 凭检验直观选择
- 按密度大小选代表点确定k
- 使距离度量方法值最小的k
- 最大最小距离法确定

#### 距离度量

- 曼哈顿距离
- 欧式距离
- 明科夫斯基距离





- K均值聚类算法在初始划分的基础上，使用迭代算法逐步优化聚类结果，使准则函数  $J_k$  达到极小值，获得  $k$  个类型。
- 当结果簇是密集的，而簇与簇之间区别明显时，K均值聚类算法的效果较好。
- 速度快，时间复杂度接近线性。对处理大数据集，该算法是相对可伸缩的和高效的，复杂度为  $O(nkt)$ ，其中， $n$  是样本个数， $k$  是簇的数目， $t$  是迭代的次数，通常， $k \ll n$  且  $t \ll n$ 。



- 但是，该算法为了降低运算时间复杂度可能以局部最优结束。
- 此外，K均值法要求用户事先给出簇的数目  $k$ ，不适合于发现非凸面形状的簇或者大小差别很大的簇。
- 而且，它对于“噪声”和孤立点数据敏感，少量的该类数据能够对均值产生极大的影响。



解  
决  
实  
际  
问  
题

H&E Image-Copyright(c) 2009  
Microsoft Corporation



原始图



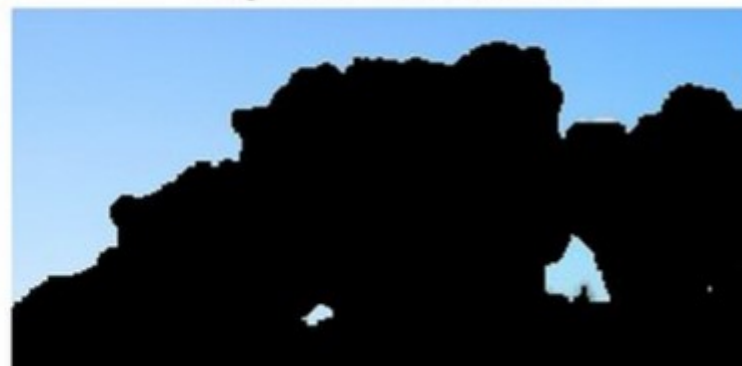
聚类成3类的结果

objects in cluster 1



类型1对应的目标

objects in cluster 2



类型2对应的目标

类型3对应的目标

计算机学院



## • 匹配

- 匹配可在（不同）抽象层次上进行。对每个具体的匹配，它都可以看成是对两个（已有的）表达找其对应性。
- 如果两个表达是相似的，匹配就是在相似的意义上进行。例如当两个表达都是图像结构时，可称为**图像匹配**；如果两个表达都代表图像中的目标时，可称为**目标匹配**；如果两个表达都是关系结构，则是**关系匹配**；如果两个表达一个是图像结构，一个是关系结构，这时也可以在扩展的意义上进行匹配，或称“**拟合**”。



- 对图像的匹配可以是整幅图像和整幅图像间的匹配，但常见的是一个较小的图像与一幅较大图像中一部分（子图像）的匹配。
- 匹配的目的：
  - 确定在大图像中是否存在小图像
  - 确定小图像在大图像中的位置



- 匹配

- 相关匹配（模板匹配）

- (1) 相关匹配的基本思想:

- Step1. 用模板子图像直接作为模式（不是用描述子）

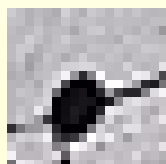
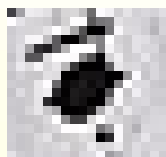
- Step2. 通过子图像与原图像直接进行相关计算，把相关计算作为决策函数。

- Step3. 相关计算获得最大值的位置，就被认为匹配成功。



## — 相关匹配

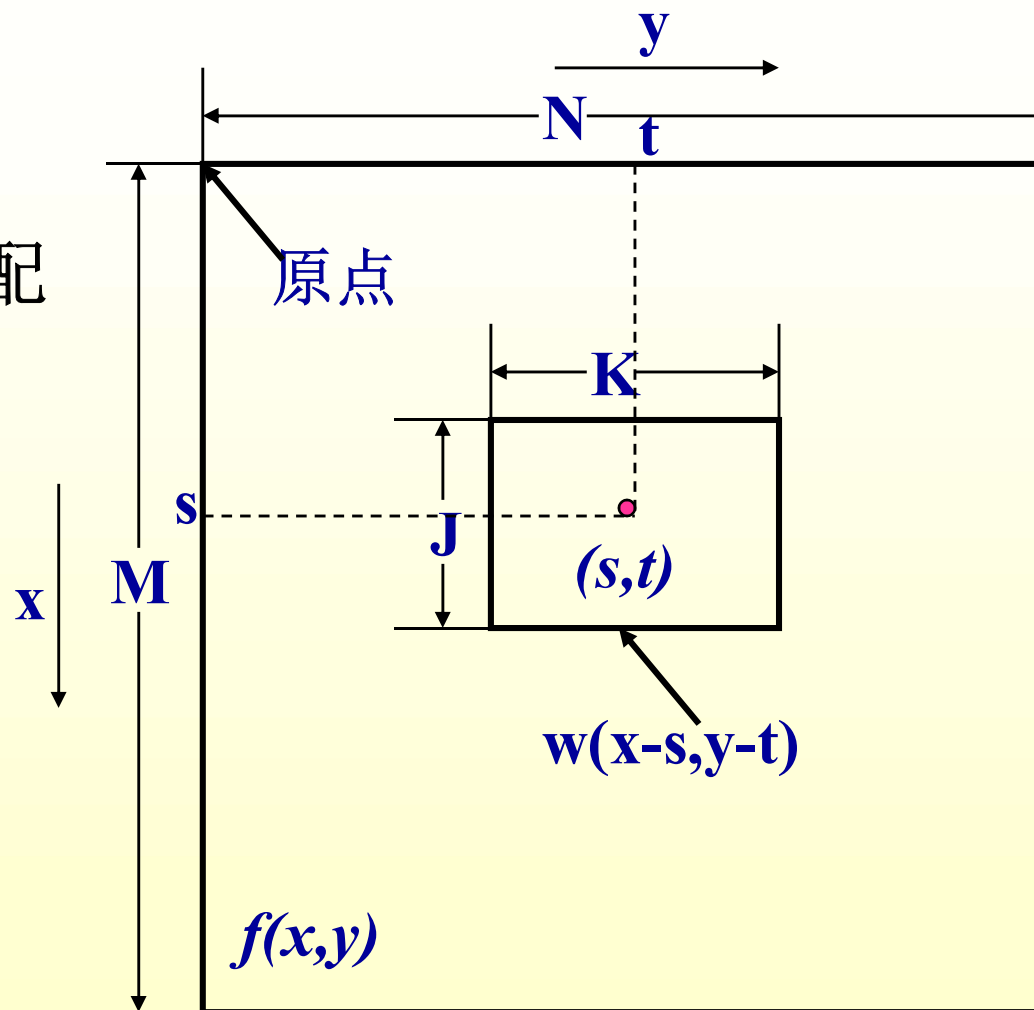
### 相关匹配基本思想







- 匹配
  - 相关性匹配







- 匹配

- 相关匹配

- (2) 算法描述

- 决策函数是相关函数

- $$c(s,t) = \sum_x \sum_y f(x,y)w(x-s,y-t)$$

- 对图像的每一个点进行相关计算，只计算重叠部分。

- 问题：（1）在边界处将失去准确性。（2）其误差与子图像的尺寸成正比,相关函数对振幅的变化太敏感。用相关系数函数代替相关函数。



$$\gamma(s, t) = \frac{\sum_x \sum_y [f(x, y) - \bar{f}(x, y)][w(x - s, y - t) - \bar{w}]}{\left\{ \sum_x \sum_y [f(x, y) - \bar{f}(x, y)]^2 \sum_x \sum_y [w(x - s, y - t) - \bar{w}]^2 \right\}^{1/2}}$$

$\bar{w}$  是w的均值

$\bar{f}(x, y)$  是f(x,y)中与w当前位置相对应区域的均值

因为相关系数已尺度变换到区间[-1,1]，因此其值与f(x,y)和w(x,y)的幅度变化无关。



## — 相关性匹配

### (3) 对旋转和比例变化的分析

- 当被匹配图像中，对象的尺寸和角度与模式不一致，此方法将失效。
- 尺寸的正则化，解决空间比例的问题。正则化模板与原图。
- 如果知道原图像的旋转角度，我们可以通过旋转原图像，对齐模式解决。
- 如果被匹配的对象的角度任意，有结论:此方法不能用于这种问题。



- 快速模板匹配策略
  - 运动预测缩小搜索范围

目标运动的连续性必然体现为时间和空间上的相关性。统计结果显示，相邻帧同一目标运动矢量相同的比例高达97.9%。根据目标在一段时间的运动轨迹，预测出目标在下一时刻的位置，优先对这一位置附近进行搜索可以减少计算量。当估计得到的目标运动速度为零时，以当前目标位置为中心点，由中心点向周围辐射匹配的方式效果最理想。



### — 粗细两级匹配

由于匹配点附近的匹配误差迅速下降，明显区别于其他位置，因此采用粗细匹配结合的方法能够迅速锁定匹配点大致所在的区域，大大降低整体匹配次数。



## —序贯相似性检测(SSDA)及时终止匹配

- ✓ 序贯相似性检测的原理认为累积误差增长越慢的点越可能是匹配点。
- ✓ 假设当前最佳匹配位置及最小SAD已求得，在计算下一候选匹配位置SAD值的过程中，一旦比已求最小SAD值都大了，就立即停止，没有必要继续计算下去。这一候选匹配位置的SAD计算及时终止,该候选匹配位置及时被淘汰，以此减少计算量。
- ✓ 如果匹配完成后该候选位置的SAD值仍小于最小SAD值，则用当前误差SAD替换最小SAD值，该位置作为最佳匹配位置记录下来。



- 先验知识
- 在相邻匹配位置上，利用模板覆盖范围有相当大重合的特点减少重新计算相关值的数量。



决策法，通过量化的方法处理模式，最大限度地忽略了模式形状的内在结构关系。

结构法，则力求通过准确地抓住这些不同模式类的内在结构关系来进行模式识别。

- 结构法
  - 匹配形状数
  - 匹配串





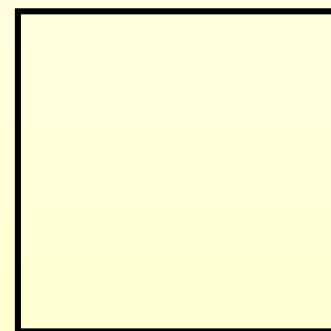
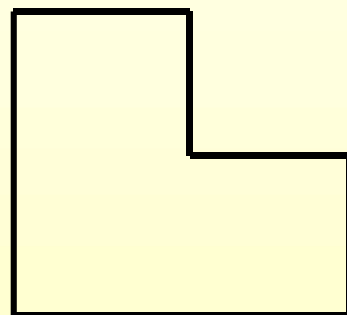
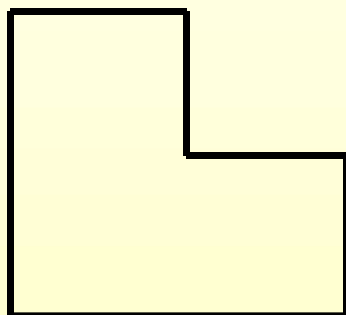
- 匹配形状数

- (1) 匹配形状数的基本思想

通过比较两个对象边界的形状数的相似程度，来匹配对象。例如：

未知模式

原形模式类





- 匹配形状数

- (2) 基本概念-a

- a. 两个区域边界相似级别k的定义：

- 相同形状数的最大序号,即—

- 当考虑用4向链码表示封闭区域边界的形状数时，A和B具有相似级别k，如果满足 $s_4(A) = s_4(B)$ ,  $s_6(A) = s_6(B)$ ,  $s_8(A) = s_8(B)$ , ...,  $s_k(A) = s_k(B)$ ,  $s_{k+2}(A) \neq s_{k+2}(B)$ ,  $s_{k+4}(A) \neq s_{k+4}(B)$ , ... , 这里s表示形状数, 下标表示阶数。



- 匹配形状数

- (2) 基本概念-b

- b.两个区域边界A和B形状数的距离 $D(A,B)$ 定义为相似级别的倒数 :  $D(A,B) = 1 / k$

- 距离满足如下性质 :

- $D(A,B) \geq 0$

- $D(A,B) = 0$  if  $A=B$

- $D(A,C) \leq \max[D(A,B), D(B,C)]$

对同一个边界，它的自身相似度为无穷！



- 匹配形状数

- (3) 算法思想

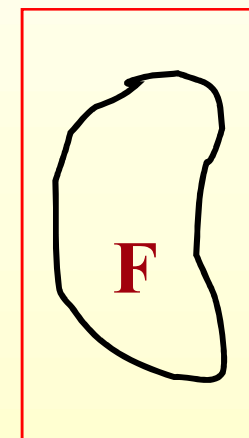
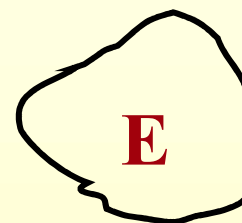
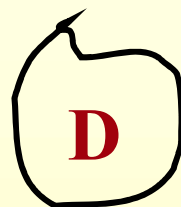
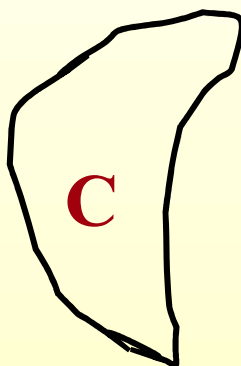
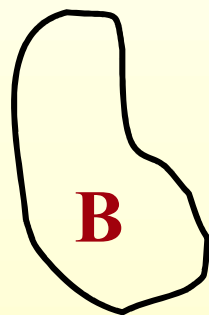
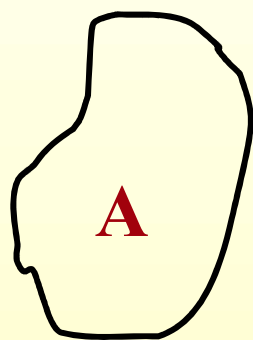
- a. 用不同密度的网格划分边界区域，获得不同序数的形状数。
    - b. 如果使用相似级别 $k$ ， $k$ 越大说明越相似。
    - c. 如果使用相似距离 $D$ ， $D$ 越小说明越相似
    - d. 可以利用相似树来进行判别



- 匹配形状数

- (4) 举例

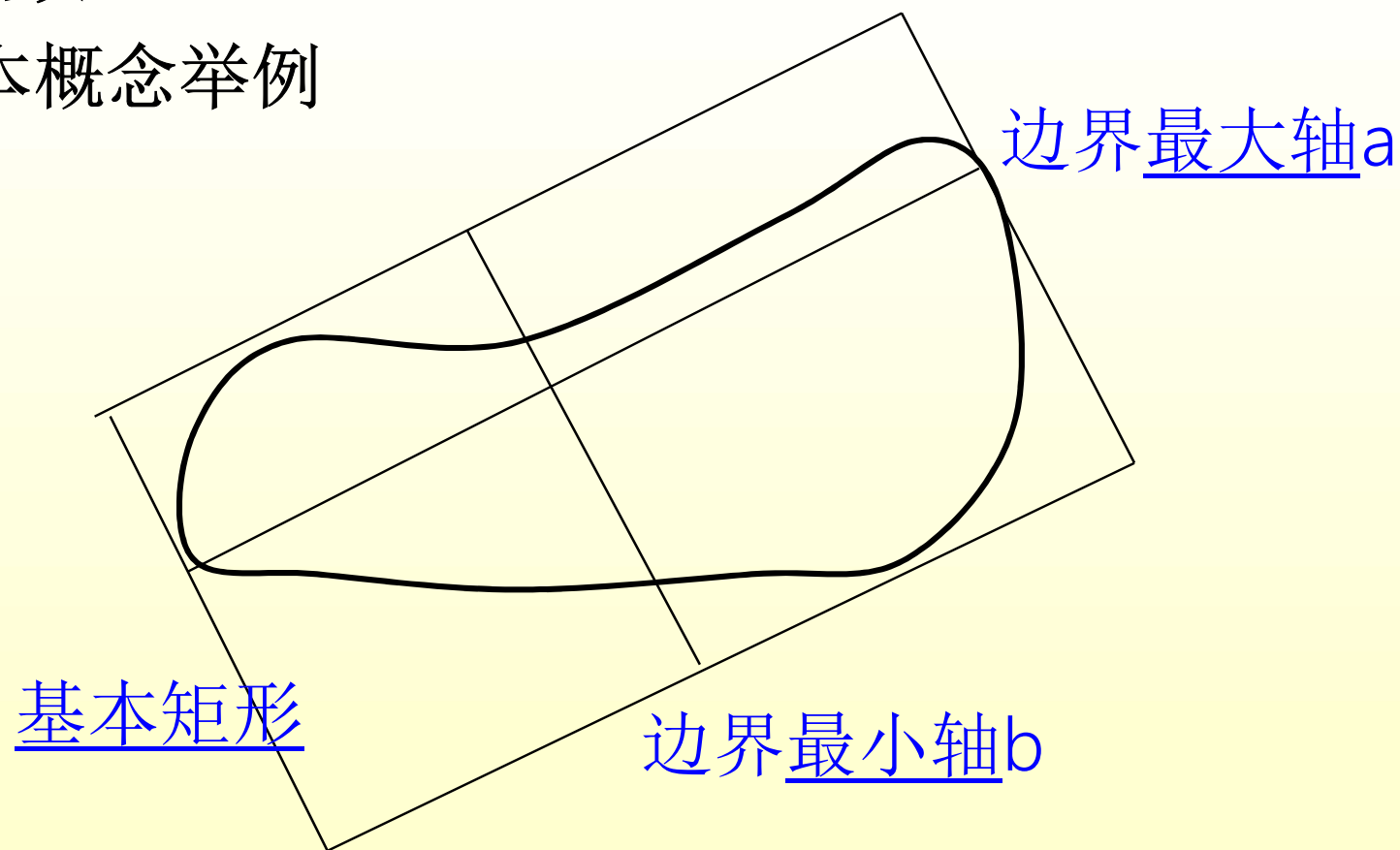
假设我们有一个形状F，想在另5个形状（A，B，C，D，E）中找到与其最相似的形状





- 形状数

- 基本概念举例





- 形状数

- 规整化网格方向算法的思想：

- 大多数情况下，将链码网格与基本矩形对齐，即可得到一个唯一的形状数。

- 对一个给定的形状序号，处理步骤如下：

- (1) 从所有满足给定阶数要求的矩形中选取一个矩形，它的离心率最接近于给定形状的基本矩形的离心率。



- 形状数

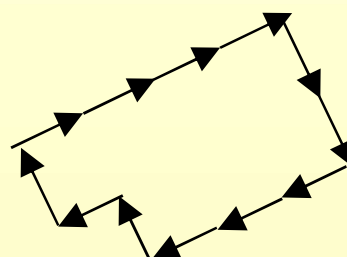
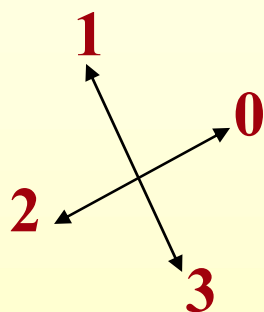
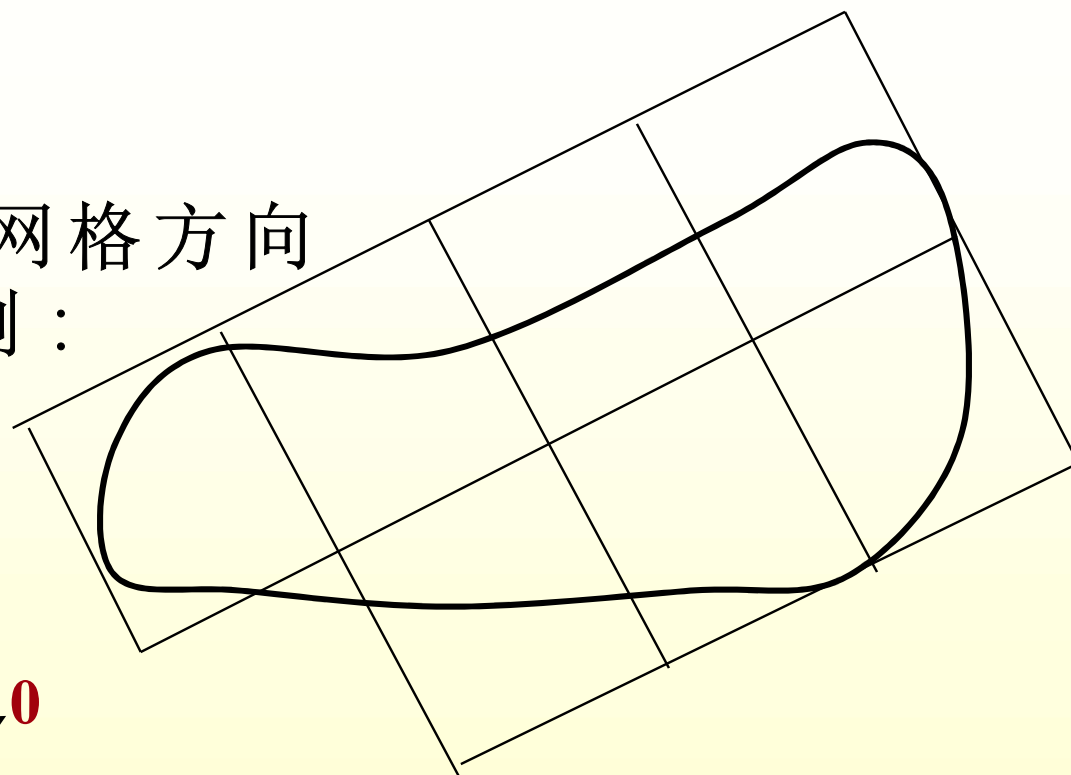
- (2) 然后再用这个矩形构造网格。
- (3) 求出与边界最吻合的多边形，例如将面积的50%以上包含在边界内的正方形划入内容。
- (4) 再得到链码。
- (5) 最后，再得到循环首差。
- (6) 首差中的最小循环数即为形状数。





- 形状数

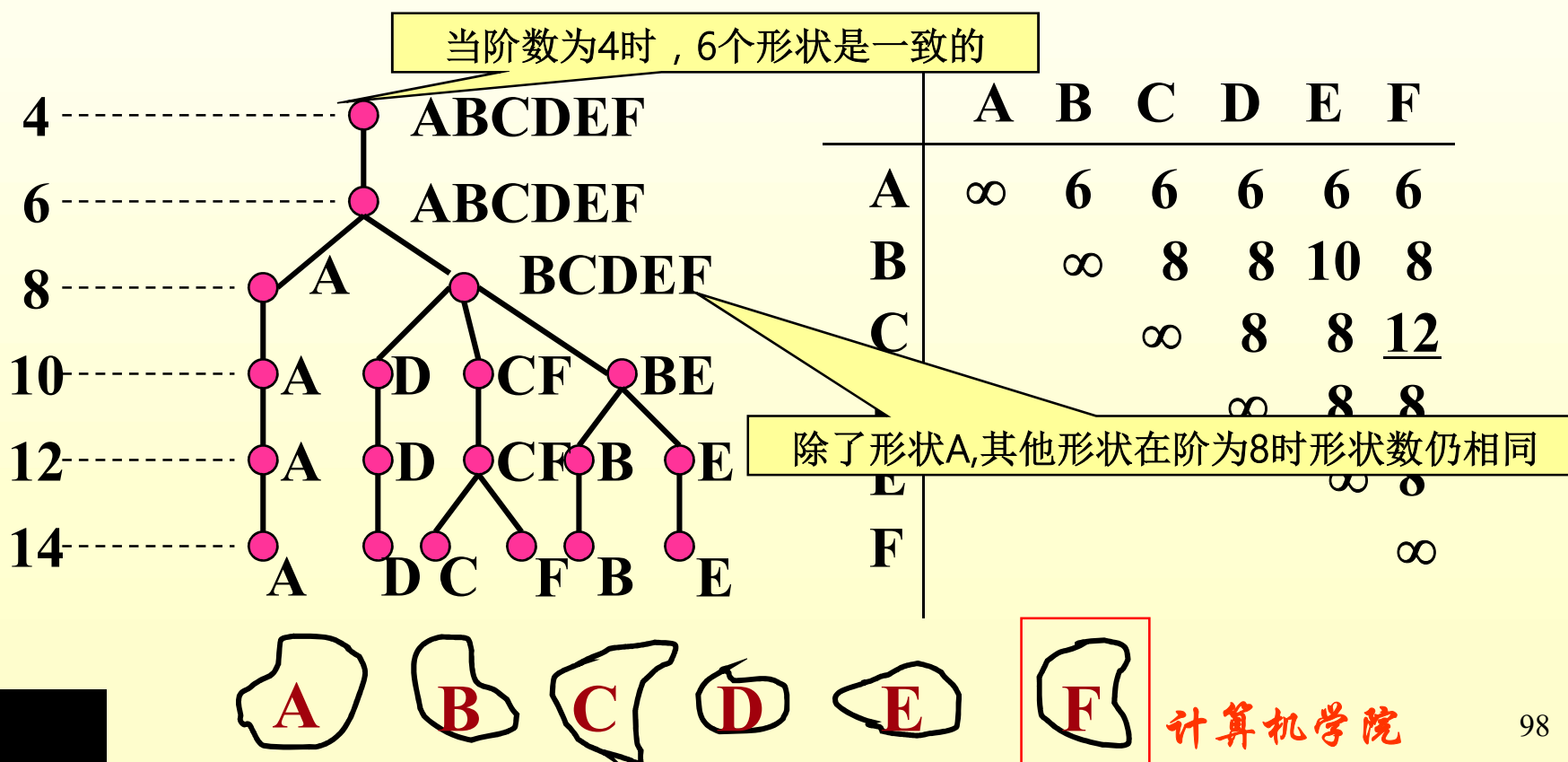
- 规整化网格方向  
算法举例：



链码：000033222121  
首差：300030300313  
形状：000303003133



这个问题类似于有五个原型形状，想找出一个给定的尚不确定的形状的最佳匹配的问题。这个问题可以利用相似树和矩阵来可视化





- 串匹配

- (1) 串匹配的基本思想

- 比较两个边界的串编码的相似程度，来进行匹配

- (2) 三个基本概念

- 设: 两个区域边界A和B已分别被编码为串 $a_1a_2\dots a_n$ 和 $b_1b_2\dots b_m$ 。



- 串匹配

- a. 两个串的匹配数M:

当  $a_k = b_k$  时我们说发生了一个匹配。令M代表A、B中匹配的总数。

- b. 不匹配的符号数量Q :

$$Q = \max ( |A| , |B| ) - M$$

这里|arg|是字符串的长度。当且仅当A和B完全相同时， $Q = 0$ 。



- 串匹配

- c. A和B相似度的简便衡量R:

$$R = M/Q = M / [\max(|A|, |B|) - M]$$

因此， 当A和B完全匹配时，  $R = \infty$ ；

当A和B中任何字符都不匹配时，

$$M = 0, R = 0。$$



- 串匹配

- (3) 算法思想

- a. 由于匹配是逐字符进行的，
    - b. 选择一个好的开始点，可以大大减少计算量。
    - c. 如果从任意一点开始计算，然后每次移动一个符号的位置再计算，整个计算非常耗时（正比于  $|A| \times |B|$ ）。
    - d. 任何将两个串规则化为相同字符开头的方法都是有效的，只要这种方法不是穷举起点。
    - e. 最大的R给出了最好的匹配。