



# 数字图象处理

---

中国科学技术大学  
电子工程与信息科学系

主讲教师：李厚强 (lihq@ustc.edu.cn)  
周文罡 (zhwg@ustc.edu.cn)

助教：王宁 (wn6149@mail.ustc.edu.cn)



# 第12章 目标识别

- 12.1 模式和模式类
- 12.2 基于决策理论方法的识别
  - 匹配
  - 最佳统计分类器
  - 神经网络
- 12.3 结构方法
  - 匹配形状数
  - 串匹配



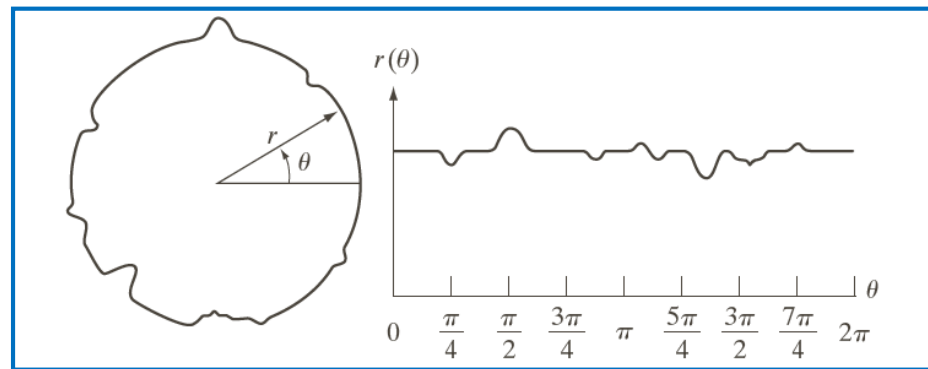
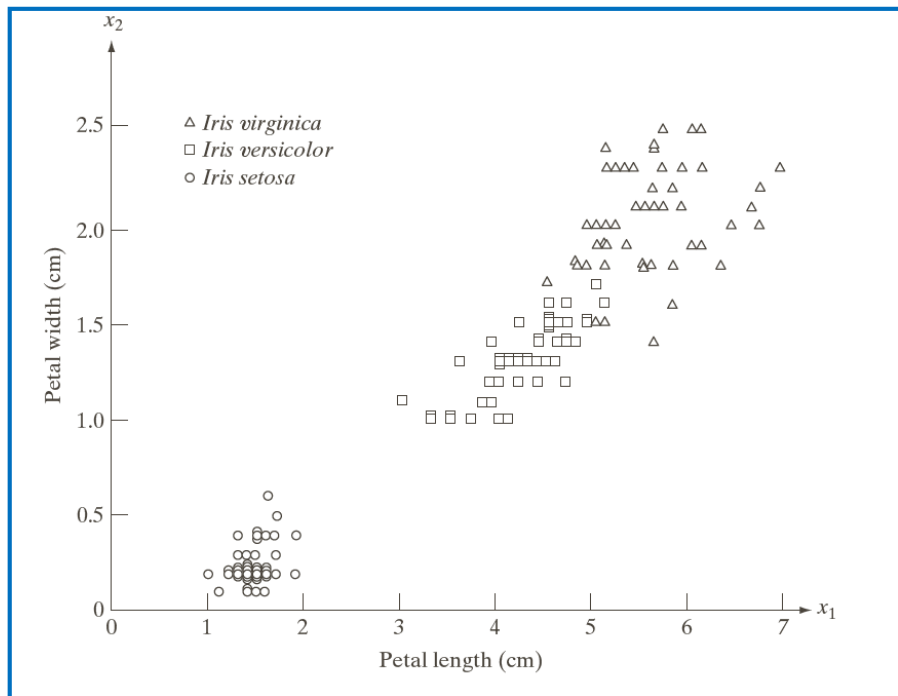
# 第12章 目标识别

- 12.1 模式和模式类
- 12.2 基于决策理论方法的识别
  - 匹配
  - 最佳统计分类器
  - 神经网络
- 12.3 结构方法
  - 匹配形状数
  - 串匹配

# 12.1 模式和模式类

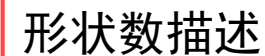
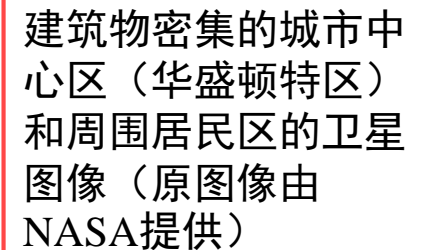
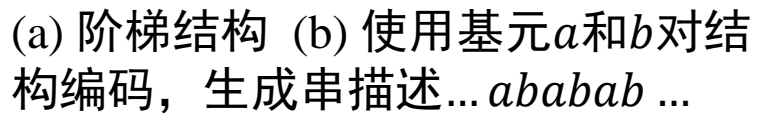
## □ 模式向量

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



一个带有噪声的目标及其相应的信号

由两个测度描述的三种鸢尾花





# 第12章 目标识别

- 12.1 模式和模式类
- 12.2 基于决策理论方法的识别
  - 匹配
  - 最佳统计分类器
  - 神经网络
- 12.3 结构方法
  - 匹配形状数
  - 串匹配



## 12.2 基于决策理论方法的识别

- $W$ 个模式类 $\omega_1, \omega_2, \dots, \omega_W$ , 决策函数 $d_1(x), d_2(x), \dots, d_W(x)$   
如果模式 $x$ 属于类 $\omega_i$ , 则
$$d_i(x) > d_j \quad j = 1, 2, \dots, W; j \neq i$$

- 决策边界

$$d_i(x) = d_j(x)$$

## 12.2.1 匹配

### □ 最小距离分类器

模式类的原型

$$\mathbf{m}_j = \frac{1}{N_j} \sum_{\mathbf{x} \in \omega_j} \mathbf{x}_j \quad j = 1, 2, \dots, W$$

距离测度

$$D_j(\mathbf{x}) = \|\mathbf{x} - \mathbf{m}_j\| \quad j = 1, 2, \dots, W$$

选择最小距离等价于

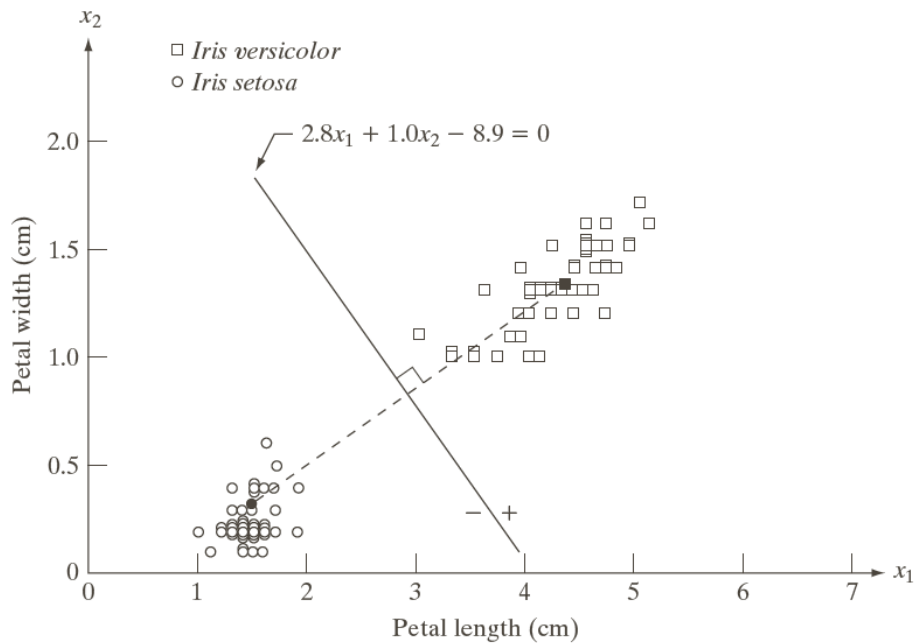
$$d_j(\mathbf{x}) = \mathbf{x}^T \mathbf{m}_j - \frac{1}{2} \mathbf{m}_j^T \mathbf{m}_j \quad j = 1, 2, \dots, W$$

决策边界

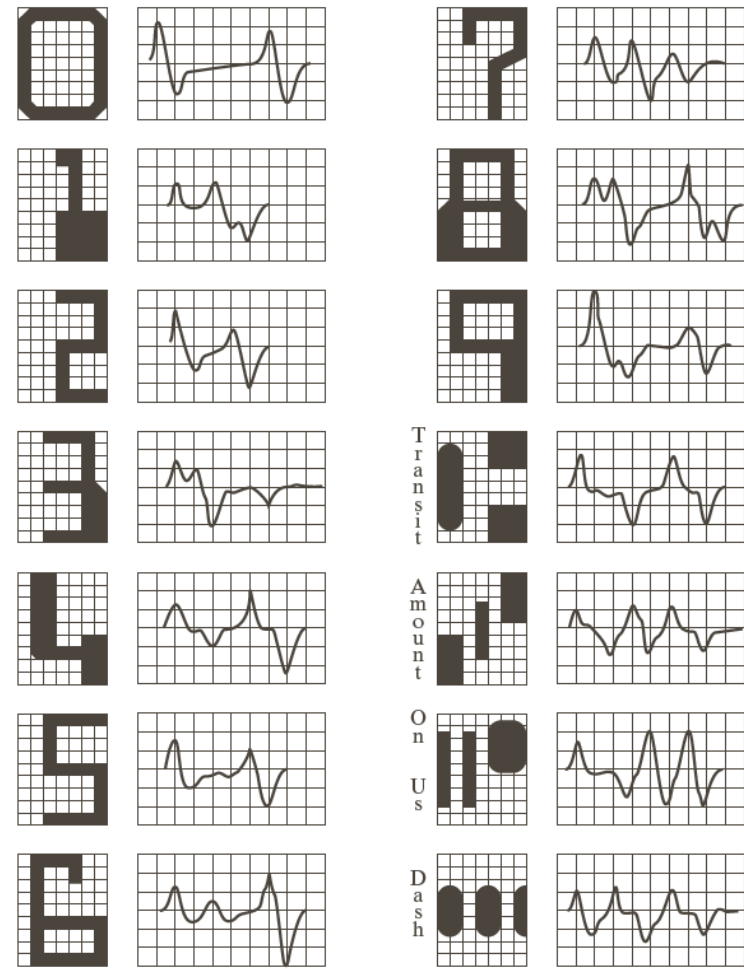
$$\begin{aligned} d_{ij}(\mathbf{x}) &= d_i(\mathbf{x}) - d_j(\mathbf{x}) \\ &= \mathbf{x}^T (\mathbf{m}_i - \mathbf{m}_j) - \frac{1}{2} (\mathbf{m}_i - \mathbf{m}_j)^T (\mathbf{m}_i + \mathbf{m}_j) = 0 \end{aligned}$$



# 12.2.1 匹配



类 *Iris versicolor* 和类 *Iris setosa* 的最小距离分类器的决策边界



美国银行家协会E-138  
字符集和对应波形

# 12.2.1 匹配

## □ 相关匹配

相关表示

$$c(x, y) = \sum_s \sum_t w(s, t) f(x + s, y + t)$$

模板      图像

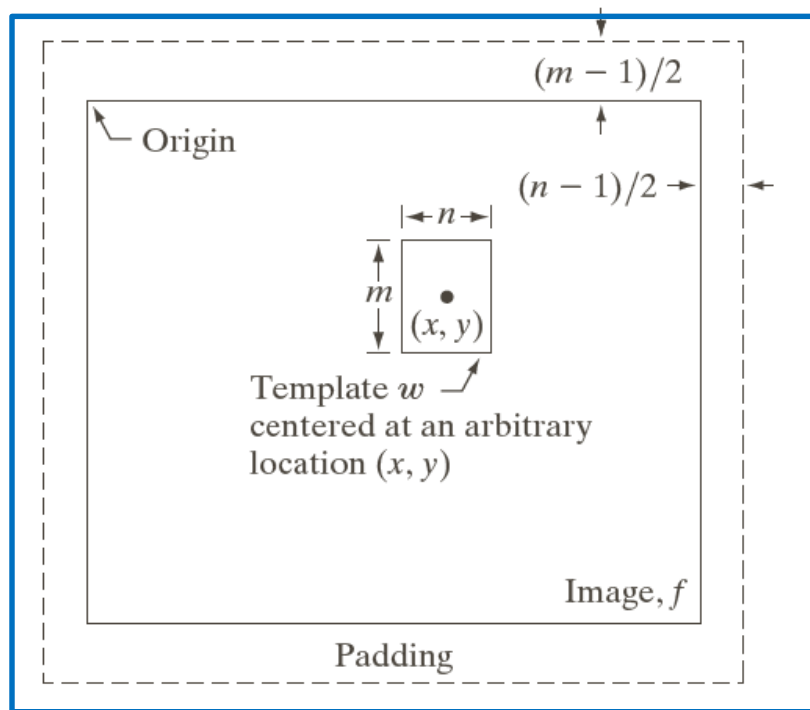
空间相关

$$f(x, y) \star w(x, y) \Leftrightarrow F^*(u, v) W(u, v)$$

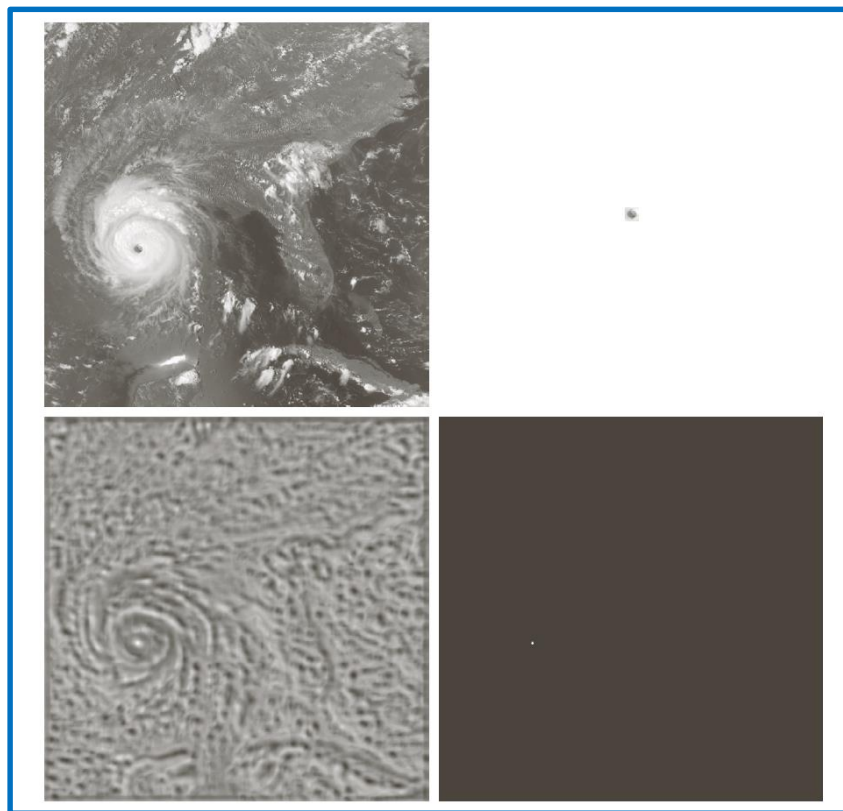
$$\gamma(x, y) = \frac{\sum_s \sum_t [w(s, t) - \bar{w}] [f(x + s, y + t) - \overline{f_{xy}}]}{\left\{ \sum_s \sum_t [w(s, t) - \bar{w}]^2 \sum_s \sum_t [f(x + s, y + t) - \overline{f_{xy}}]^2 \right\}^{\frac{1}{2}}}$$

归一化相关系数

# 12.2.1 匹配



模板匹配的机理



(a) 飓风“安德鲁”的卫星图像，摄于1992年8月24日；(b) 飓风眼的模板；(c) 显示为图像的相关系数（注意最亮点）；(d) 匹配最好的位置。该点是单个像素，但其已被放大，以便于查看（原图像由NOAA提供）

## 12.2.2 最佳统计分类器

### □ 基础知识

平均损失

$$\begin{aligned} r_j(\mathbf{x}) &= \sum_{k=1}^W L_{kj} p(\omega_k | \mathbf{x}) \\ &= \frac{1}{p}(\mathbf{x}) \sum_{k=1}^W L_{kj} p(\mathbf{x} | \omega_k) P(\omega_k) \end{aligned}$$

0-1损失

$$L_{ij} = 1 - \delta_{ij}$$

损失函数

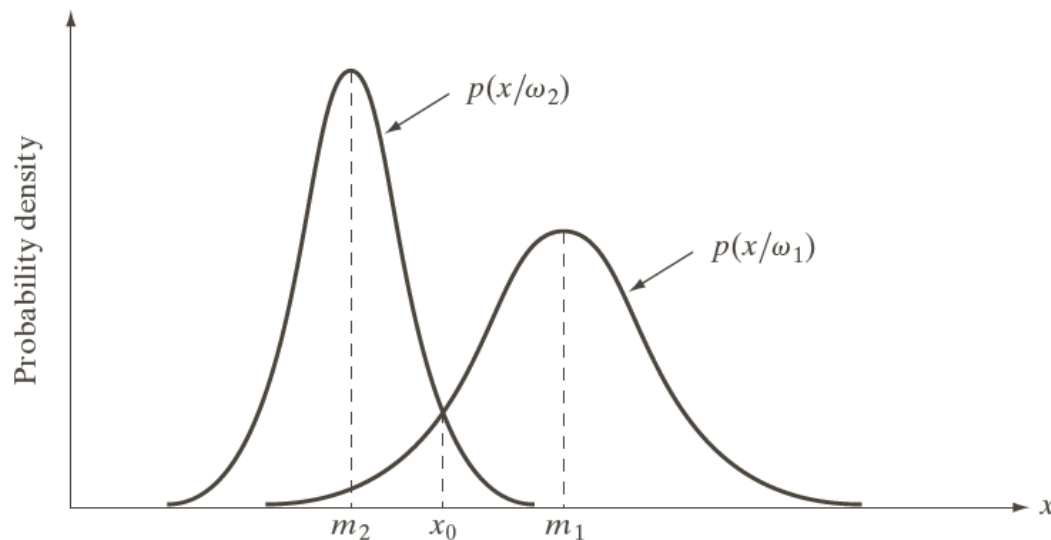
$$\begin{aligned} r_j(\mathbf{x}) &= \sum_{k=1}^W (1 - \delta_{ij}) p(\omega_k | \mathbf{x}) \\ &= P(\mathbf{x}) - p(\mathbf{x} | \omega_j) P(\omega_j) \end{aligned}$$

## 12.2.2 最佳统计分类器

### □ 高斯模式类的贝叶斯分类器

贝叶斯决策函数

$$\begin{aligned} d_j(x) &= p(x|\omega_j)P(\omega_j) \\ &= \frac{1}{\sqrt{2\pi}\sigma_j} e^{-\frac{(x-m_j)^2}{2\sigma_j^2}} P(\omega_j) \quad j = 1, 2 \end{aligned}$$



两个一维模式类的概率密度函数。如果两个类出现概率相等，那么所示的点 $x_0$ 就是决策边界



## 12.2.2 最佳统计分类器

### □ 高斯模式类的贝叶斯分类器

$n$ 维情形下,  $j$ 个  
模式类中的向量  
高斯密度

$$p(\mathbf{x}|\omega_j) = \frac{1}{(2\pi)^{n/2} |\mathbf{C}_j|^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\mathbf{m}_j)^T(\mathbf{x}-\mathbf{m}_j)}$$

$$\mathbf{m}_j = \frac{1}{N_j} \sum_{\mathbf{x} \in \omega_j} \mathbf{x}$$

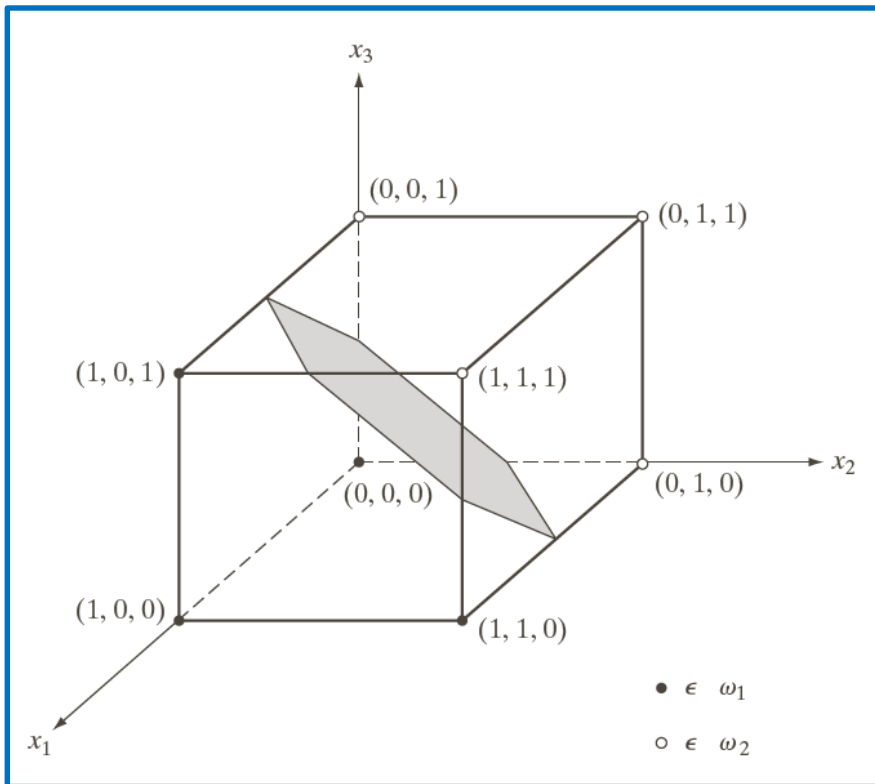
$$\mathbf{C}_j = \frac{1}{N_j} \sum_{\mathbf{x} \in \omega_j} \mathbf{x}\mathbf{x}^T - \mathbf{m}_j\mathbf{m}_j^T$$

线性决策函数  
(超平面)

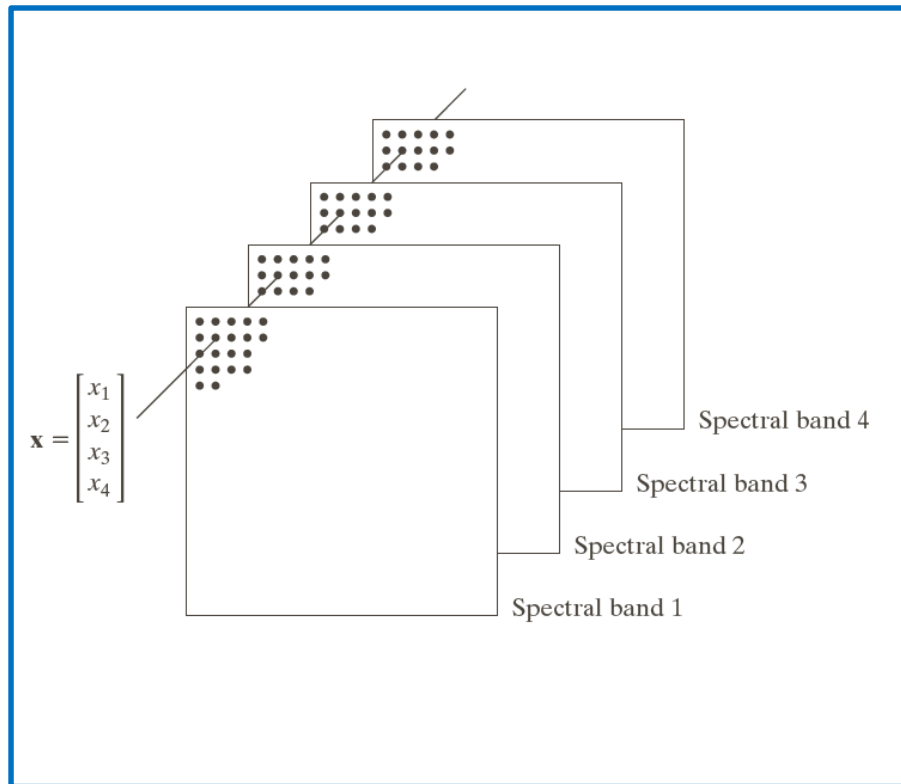
$$d_j(\mathbf{x}) = \ln P(\omega_j) + \mathbf{x}^T \mathbf{C}^{-1} \mathbf{m}_j - \frac{1}{2} \mathbf{m}_j^T \mathbf{C}^{-1} \mathbf{m}_j$$

# 12.2.2 最佳统计分类器

## □ 高斯模式类的贝叶斯分类器



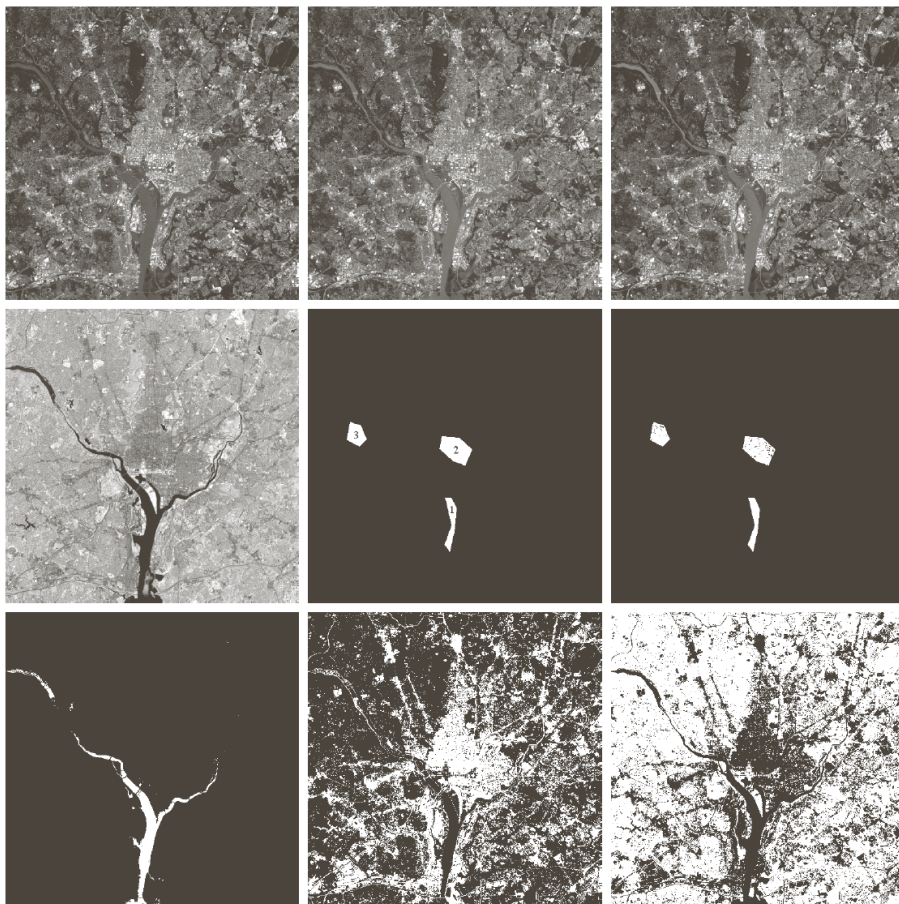
两个简单的模式类  
及其贝叶斯决策边  
界（阴影所示）



由多光谱扫描器生成的4  
幅数字图像经像素配准后，  
所形成的的模式向量

## 12.2.2 最佳统计分类器

### □ 高斯模式类的贝叶斯分类器



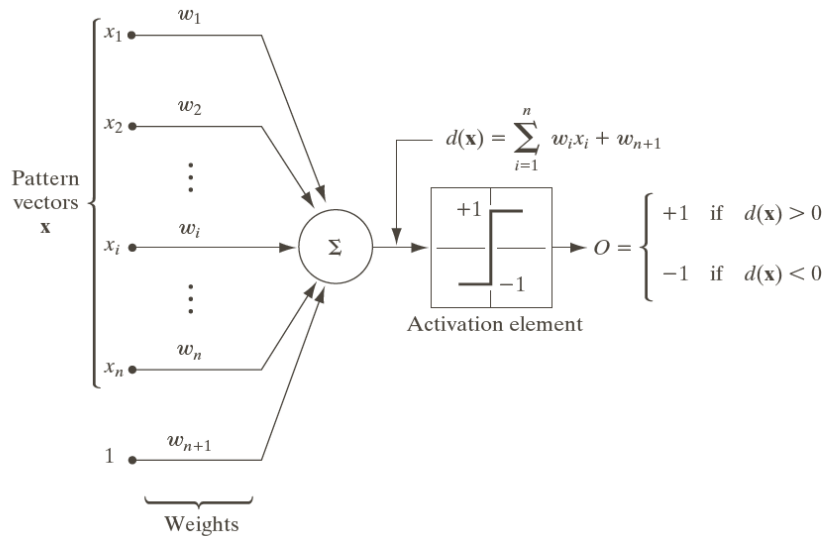
a	b	c
d	e	f
g	h	i

多光谱数据的贝叶斯分类：  
(a)~(d) 可见蓝光、可见绿光、可见红光和古今红外波长图像；  
(e) 显示(1)水体、(2)市区和(3)植被的样本区域的模板；  
(f) 分类结果。黑点表示为正确分类的点，其他（白）点是正确分类的点；  
(g) 分类为水体的所有图像像素（白色）；  
(h) 分类为市区的所有图像像素（白色）；  
(i) 分类为植被的所有图像像素（白色）

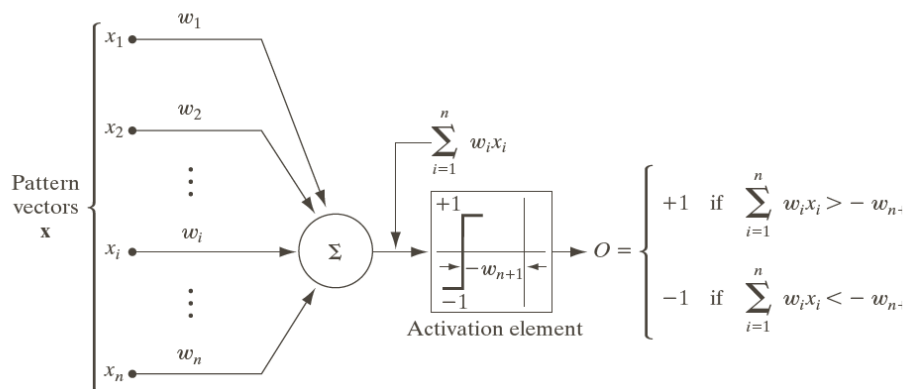


# 12.2.3 神经网络

## □ 两个模式类的感知机



$$d(x) = \sum_{i=1}^n w_i x_i + w_{n+1}$$



$$O = \begin{cases} +1, & \sum_{i=1}^n w_i x_i > -w_{n+1} \\ -1, & \sum_{i=1}^n w_i x_i < -w_{n+1} \end{cases}$$



## 12.2.3 神经网络

### □ 训练算法

#### ■ 线性可分的类:

如果  $\mathbf{y}(k) \in \omega_1$  且  $\mathbf{w}^T \mathbf{y}(k) \leq 0$ , 则

$$\mathbf{w}(k+1) = \mathbf{w}(k) + c\mathbf{y}(k)$$

相反, 如果  $\mathbf{y}(k) \in \omega_2$  且  $\mathbf{w}^T \mathbf{y}(k) \geq 0$ , 则

$$\mathbf{w}(k+1) = \mathbf{w}(k) - c\mathbf{y}(k)$$

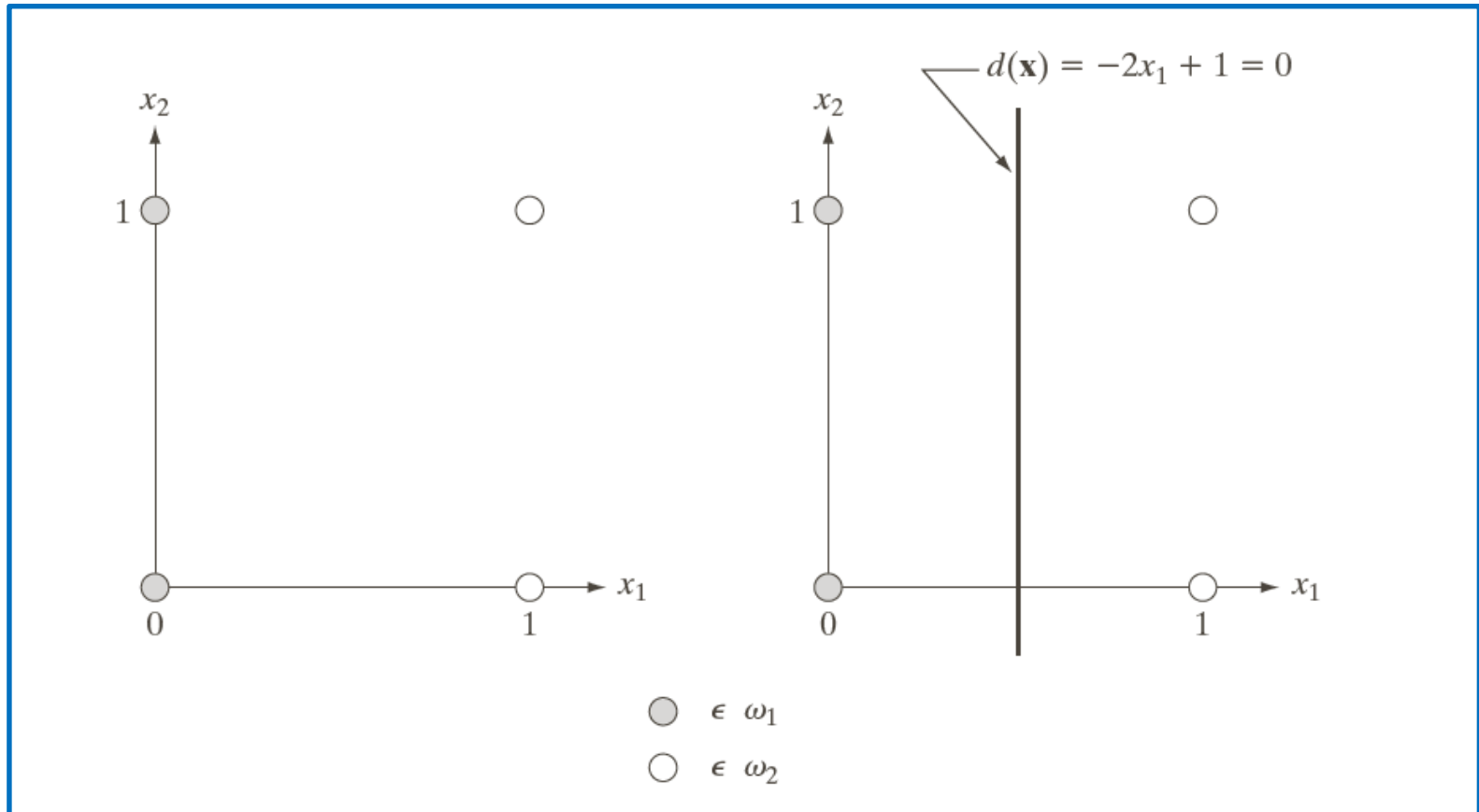
否则,  $\mathbf{w}(k)$  保持不变

#### ■ 线性不可分的类:

$$\text{准则函数 } J(\mathbf{w}) = \frac{1}{2} (r - \mathbf{w}^T \mathbf{y})^2$$

$$\text{梯度下降算法 } \mathbf{w}(k+1) = \mathbf{w}(k) - \alpha \left[ \frac{\partial J(\mathbf{w})}{\partial \mathbf{w}} \right]_{\mathbf{w}=\mathbf{w}(k)}$$

## 12.2.3 神经网络

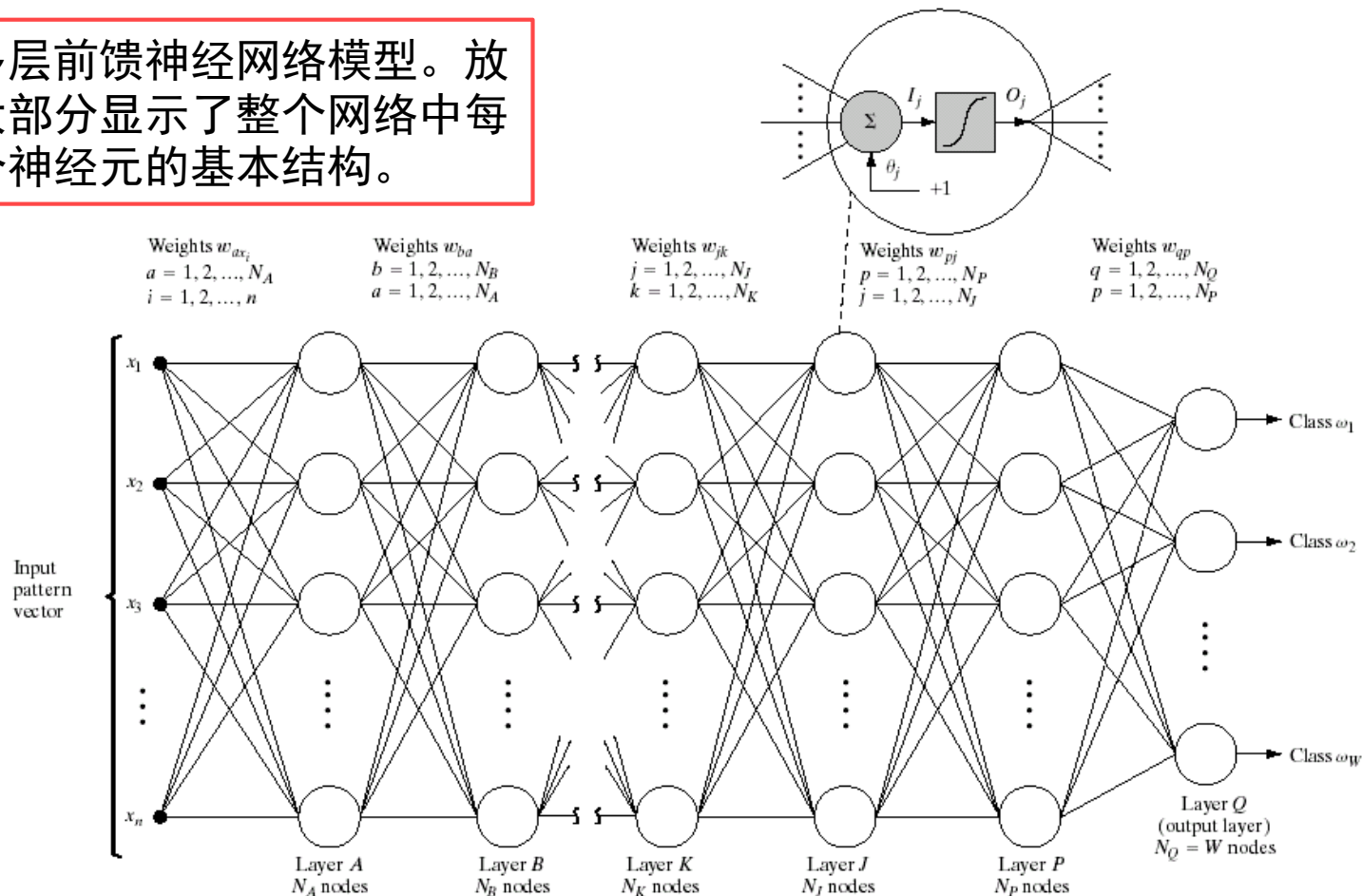


- (a) 属于两个类的模式;
- (b) 由训练确定的决策边界

# 12.2.3 神经网络

## □ 多层前馈神经网络

多层前馈神经网络模型。放大部分显示了整个网络中每个神经元的基本结构。



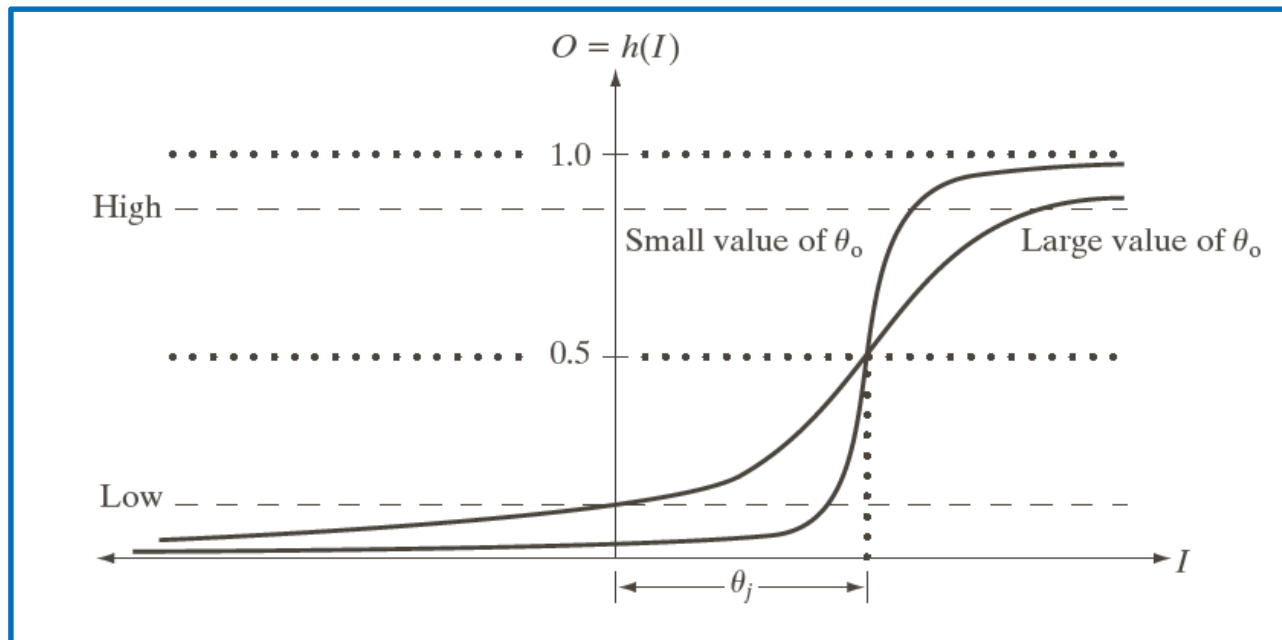
**FIGURE 12.16** Multilayer feedforward neural network model. The blowup shows the basic structure of each neuron element throughout the network. The offset,  $\theta_j$ , is treated as just another weight.

## 12.2.3 神经网络

### □ 多层前馈神经网络

激活函数

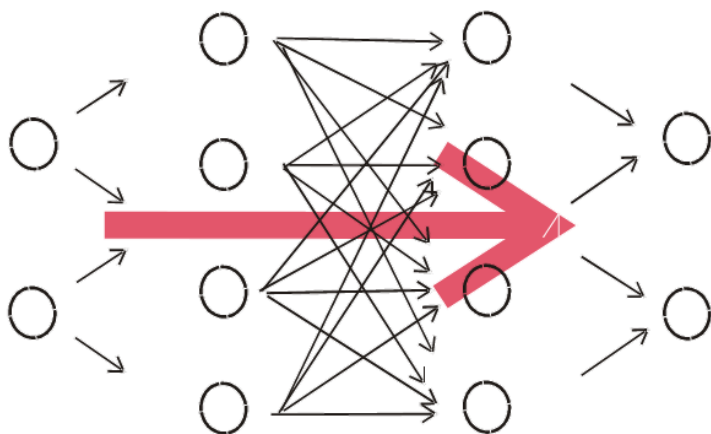
$$h_j(I_j) = \frac{1}{1 + e^{-(I_j + \theta_j)/\theta_o}}$$



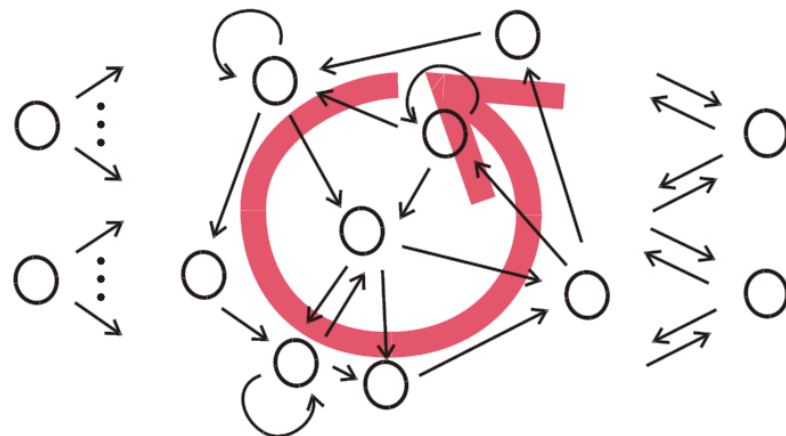
## 12.2.3 神经网络

### □ 按连接方式划分

- Forward Neural Network (前馈神经网络)
  - ✓ 主要是函数映射, 可用于模式识别和模拟非线性映射 (函数逼近)
- Feedback Neural Network (反馈神经网络)
  - ✓ 会对前面的信息进行记忆并应用于当前的输出计算中
  - ✓ 非线性动态系统, 可用来实现联想记忆和求解优化等问题。



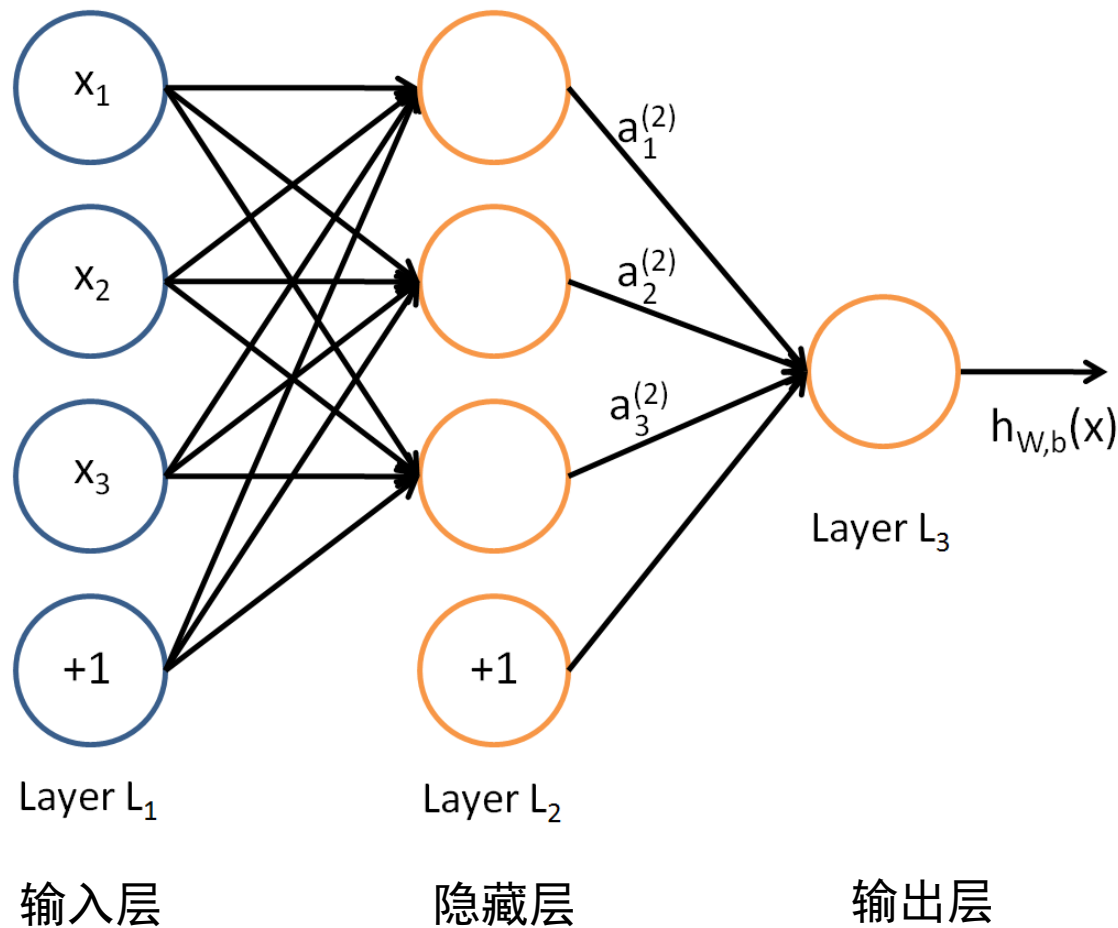
前馈网络 - DAG



反馈网络 - Digraph

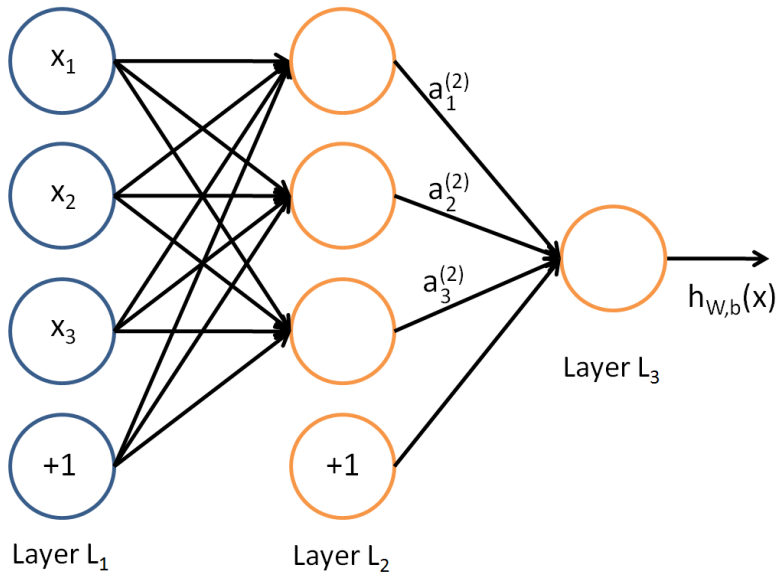
# Forward Network (前馈网络)

## □ 神经网络模型



# Forward Network (前馈网络)

## □ 神经网络模型



$a_i^{(l)}$ : 第 $l$ 层第 $i$ 单元的输出值

$z_i^{(l)}$ : 第 $l$ 层第 $i$ 单元的输入加权和

$$\begin{aligned} a_1^{(2)} &= f(z_1^{(2)}) \\ &= f(W_{11}^{(1)}x_1 + W_{12}^{(1)}x_2 + W_{13}^{(1)}x_3 + b_1^{(1)}) \end{aligned}$$

$$\begin{aligned} a_2^{(2)} &= f(z_2^{(2)}) \\ &= f(W_{21}^{(1)}x_1 + W_{22}^{(1)}x_2 + W_{23}^{(1)}x_3 + b_2^{(1)}) \end{aligned}$$

$$\begin{aligned} a_3^{(2)} &= f(z_3^{(2)}) \\ &= f(W_{31}^{(1)}x_1 + W_{32}^{(1)}x_2 + W_{33}^{(1)}x_3 + b_3^{(1)}) \end{aligned}$$

$$\begin{aligned} h_{W,b}(\mathbf{x}) &= a_1^{(3)} = f(z_1^{(3)}) \\ &= f(W_{11}^{(2)}a_1^{(2)} + W_{12}^{(2)}a_2^{(2)} + W_{13}^{(2)}a_3^{(2)} + b_1^{(2)}) \end{aligned}$$



# 训练前馈网络

## □ 误差后向传播算法（BP）

- 训练样本集  $\{(\mathbf{x}^{(1)}, y^{(1)}), \dots, (\mathbf{x}^{(m)}, y^{(m)})\}$
- 代价函数

对于单个样例  $(\mathbf{x}, y)$  的代价函数  $J(\mathbf{W}, b; \mathbf{x}, y) = \frac{1}{2} \|\mathbf{h}_{\mathbf{W}, b}(\mathbf{x}) - y\|^2$

对于训练集的整体代价函数

$$\begin{aligned} J(\mathbf{W}, b) &= \left[ \frac{1}{m} \sum_{i=1}^m J(\mathbf{W}, b; \mathbf{x}^{(i)}, y^{(i)}) \right] + \frac{\lambda}{2} \sum_{l=1}^{n_l-1} \sum_{i=1}^{s_l} \sum_{j=1}^{s_{l+1}} (W_{ji}^{(l)})^2 \\ &= \left[ \frac{1}{m} \sum_{i=1}^m \left( \frac{1}{2} \|\mathbf{h}_{\mathbf{W}, b}(\mathbf{x}^{(i)}) - y^{(i)}\|^2 \right) \right] + \frac{\lambda}{2} \sum_{l=1}^{n_l-1} \sum_{i=1}^{s_l} \sum_{j=1}^{s_{l+1}} (W_{ji}^{(l)})^2 \end{aligned}$$

以上公式中第一项是一个均方差项，第二项是一个规则化项（也叫权重衰减weight decay项），其目的是减小权重的幅度，防止过度拟合。

# 训练前馈网络

## □ 后向传播算法

- 参数随机初始化
- 进行前馈传导计算，利用前向传导公式，得到 $L_2, L_3, \dots, L_{n_l}$ 的输出值；
- 对于第 $n_l$ 层（输出层）的每个输出单元 $i$ ，我们根据以下公式计算残差：

$$\begin{aligned}\delta_i^{(n_l)} &= \frac{\partial}{\partial z_i^{(n_l)}} J(\mathbf{W}, b; \mathbf{x}, y) = \frac{\partial}{\partial z_i^{(n_l)}} \frac{1}{2} \|y - h_{\mathbf{W}, b}(\mathbf{x})\|^2 \\ &= -(y - a_i^{(n_l)}) \cdot f'(z_i^{(n_l)})\end{aligned}$$

# 训练前馈网络

## □ 后向传播算法

- 对  $l = n_l - 1, n_l - 2, n_l - 3, \dots, 2$  的各个层，第  $l$  层的第  $i$  个节点的残差计算方法如下：

$$\delta_i^{(n_l-1)} = \frac{\partial}{\partial z_i^{(n_l-1)}} J(\mathbf{W}, b; \mathbf{x}, y) = \frac{\partial}{\partial z_i^{(n_l)}} J(\mathbf{W}, b; \mathbf{x}, y) \cdot \frac{\partial z_i^{(n_l)}}{\partial z_i^{(n_l-1)}}$$

$$= \delta_i^{(n_l)} \cdot \frac{\partial z_i^{(n_l)}}{\partial z_i^{(n_l-1)}} = \delta_i^{(n_l)} \cdot \frac{\partial}{\partial z_i^{(n_l-1)}} \sum_{j=1}^{s_{n_l}} W_{ji}^{(n_l-1)} f(z_i^{(n_l-1)})$$

$$= (\sum_{j=1}^{s_{n_l}} W_{ji}^{(n_l-1)} \delta_i^{(n_l)}) f'(z_i^{(n_l-1)})$$

$$\delta_i^{(l)} = (\sum_{j=1}^{s_{l+1}} W_{ji}^{(l)} \delta_i^{(l+1)}) f'(z_i^{(l)})$$

# 训练前馈网络

## □ 后向传播算法（BP）

- 计算我们需要的偏导数，计算方法如下：

$$\frac{\partial}{\partial W_{ij}^{(l)}} J(\mathbf{W}, b; \mathbf{x}, y) = a_j^{(l)} \delta_i^{(l+1)}$$

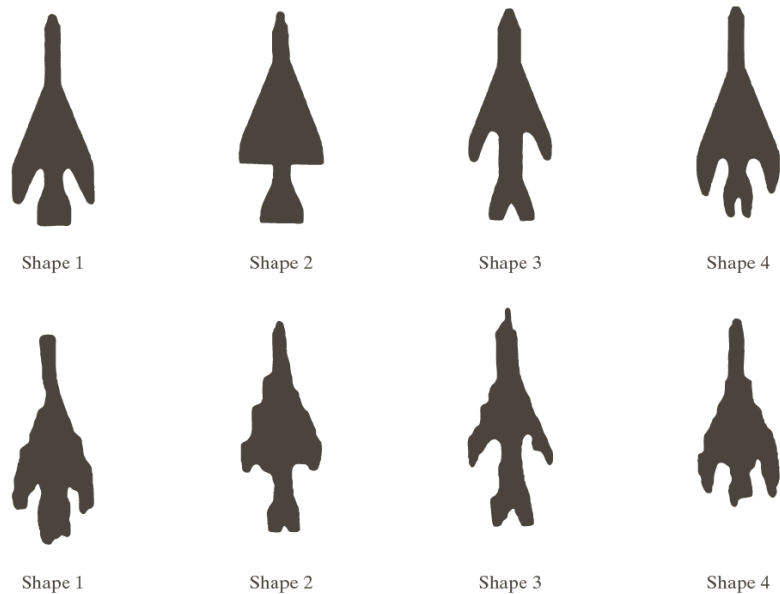
$$\frac{\partial}{\partial b_i^{(l)}} J(\mathbf{W}, b; \mathbf{x}, y) = \delta_i^{(l+1)}$$

- 更新参数：

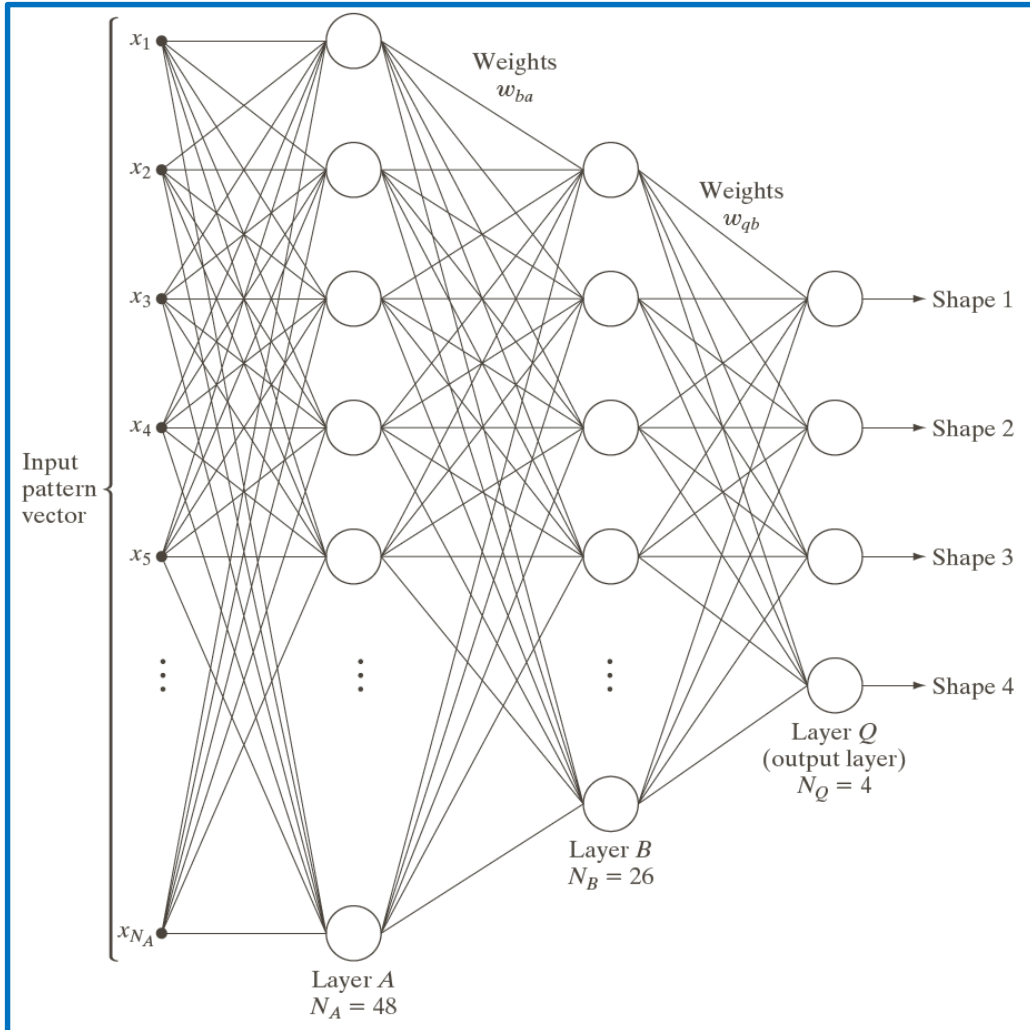
$$W_{ij}^{(l)} = W_{ij}^{(l)} - \alpha \left[ \frac{1}{m} a_j^{(l)} \delta_i^{(l+1)} + \lambda W_{ij}^{(l)} \right]$$

$$b_i^{(l)} = b_i^{(l)} - \frac{\alpha}{m} \delta_i^{(l+1)}$$

# 12.2.3 神经网络

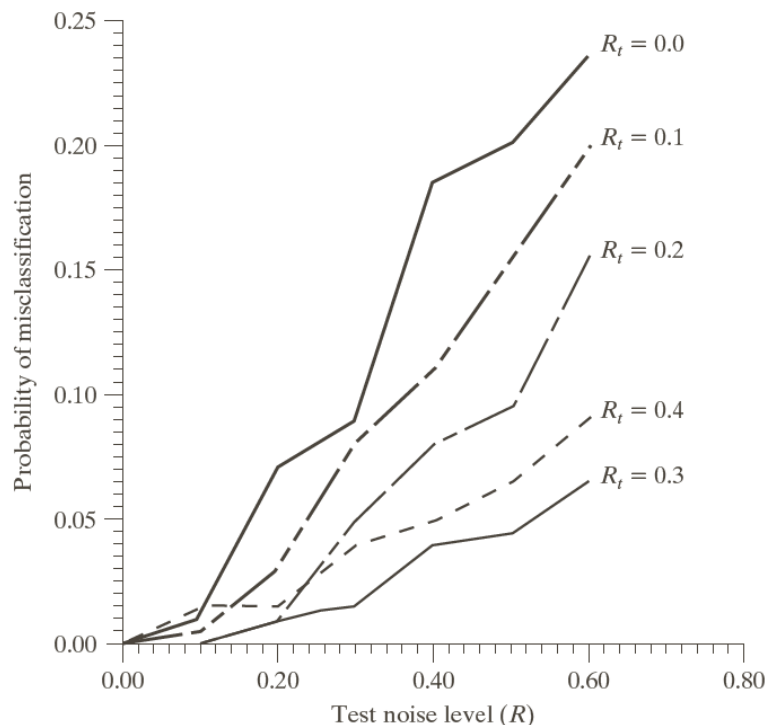


训练右图中的神经网络时所使用的  
(a) 参考图形和 (b) 典型的带噪声图形（图像由南伊利诺伊大学ECE系的Lalit Gupta博士提供）

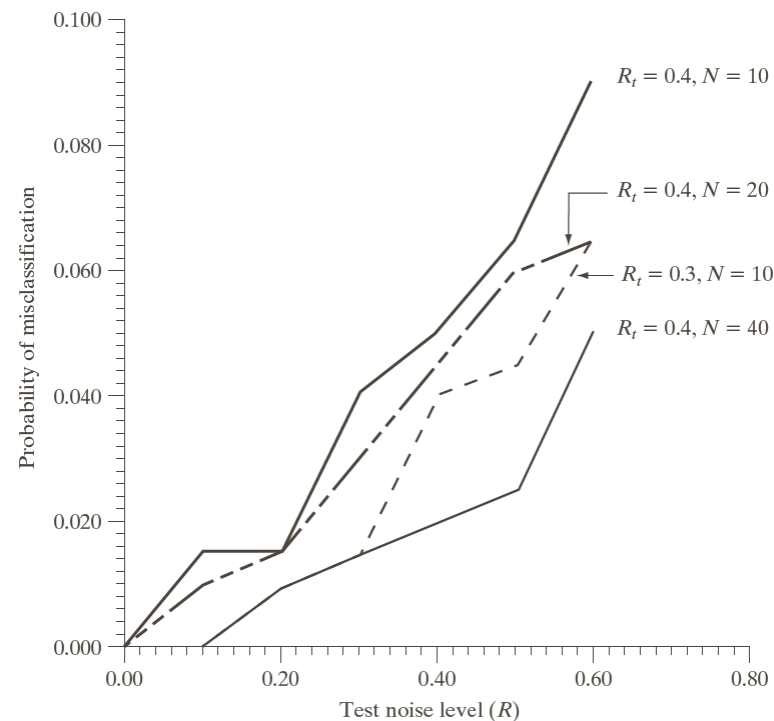


用于识别左图所示形状 of 三层神经网络

## 12.2.3 神经网络



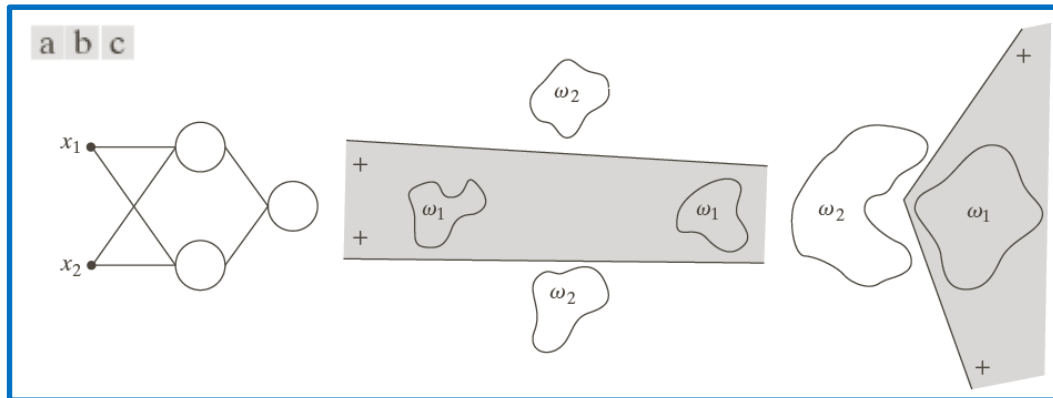
神经网络的性能是噪声水平的函数（原图由南伊利诺伊大学ECE系Lalit Gupta博士提供）




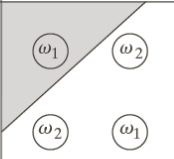
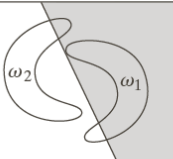
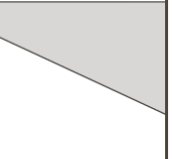

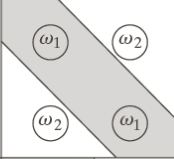
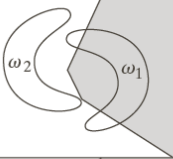
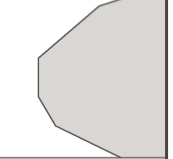

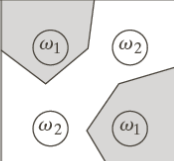
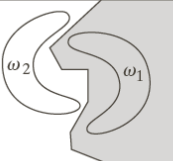
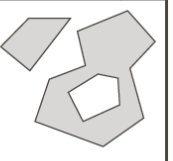
增加训练模式的数量时,  $R_t = 0.4$  时的性能改进 ( $R_t = 0.3$  时的曲线仅作为参考) (原图由南伊利诺伊大学ECE系Lalit Gupta博士提供)

# 12.2.3 神经网络

## □ 决策面的复杂性



(a) 一个双输入双层前馈神经网络；  
(b) 和 (c) 可使用该网络实现的决策边界示例

Network structure	Type of decision region	Solution to exclusive-OR problem	Classes with meshed regions	Most general decision surface shapes
Single layer 	Single hyperplane			
Two layers 	Open or closed convex regions			
Three layers 	Arbitrary (complexity limited by the number of nodes)			

可以由单层和多层前馈网络带有一层或两层隐藏单元与两个输入形成的决策区域的类型



# 第12章 目标识别

- 12.1 模式和模式类
- 12.2 基于决策理论方法的识别
  - 匹配
  - 最佳统计分类器
  - 神经网络
- 12.3 结构方法
  - 匹配形状数
  - 串匹配





## 12.3 结构方法

### □ 匹配形状数

$$\begin{aligned} s_j(a) &= s_j(b), & i &= 4, 6, 8, \dots, k \\ s_j(a) &\neq s_j(b), & j &= k + 2, k + 4, \dots \end{aligned}$$

距离

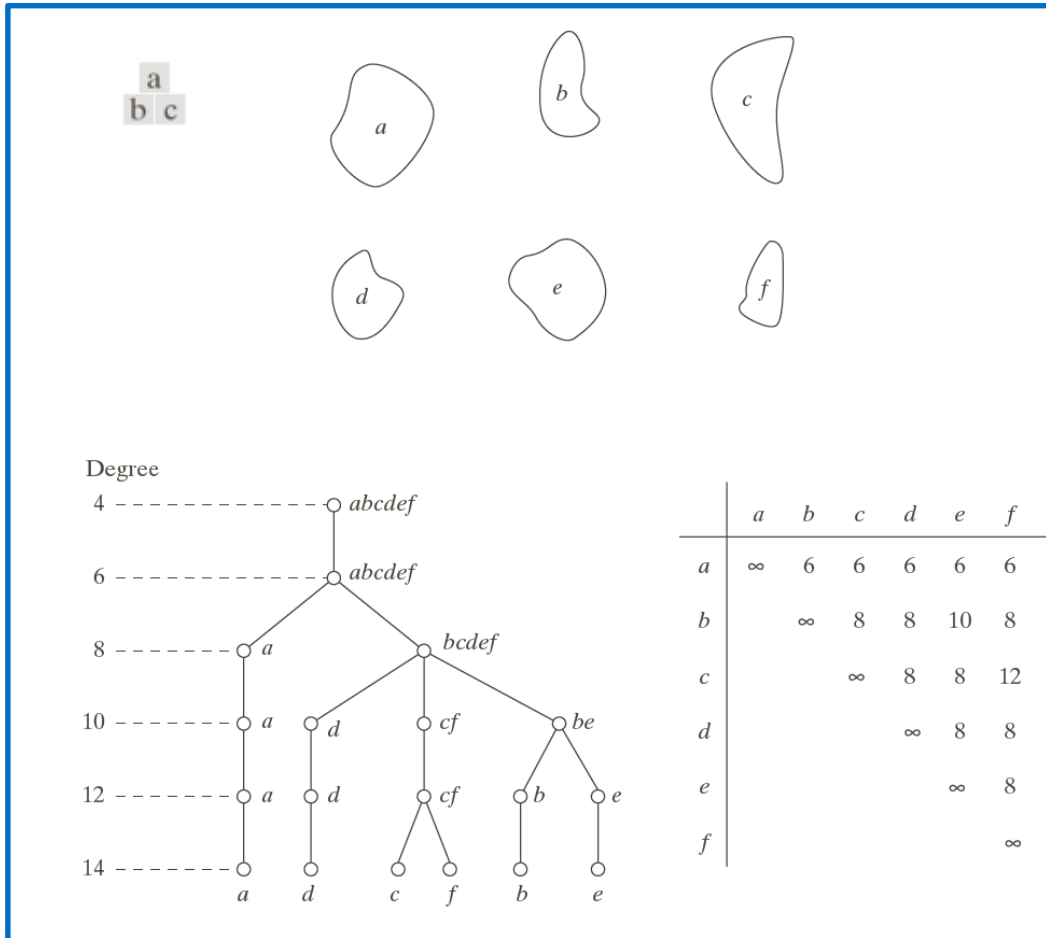
$$D(a, b) = \frac{1}{k}$$

性质

$$\begin{aligned} D(a, b) &\geq 0 \\ D(a, b) &= 0, & \text{if } a &= b \\ D(a, c) &\leq \max[(D(a, b), D(b, c))] \end{aligned}$$

# 12.3 结构方法

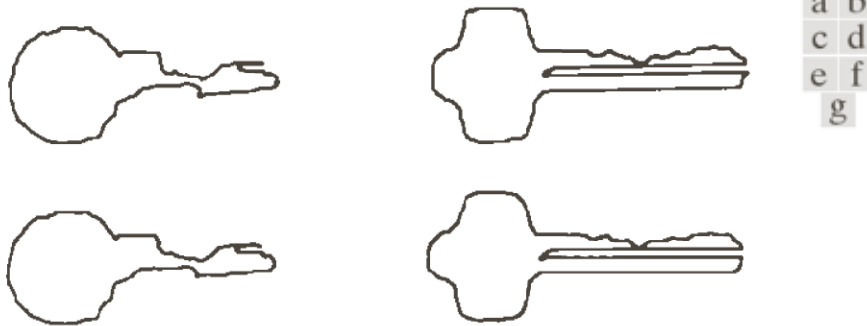
## 匹配形状数



- (a) 各种形状;
- (b) 假想的相似树;
- (c) 相似性矩阵

# 12.3 结构方法

## 串匹配



R	1.a	1.b	1.c	1.d	1.e	1.f
1.a	∞					
1.b	16.0	∞				
1.c	9.6	26.3	∞			
1.d	5.1	8.1	10.3	∞		
1.e	4.7	7.2	10.3	14.2	∞	
1.f	4.7	7.2	10.3	8.4	23.7	∞

R	2.a	2.b	2.c	2.d	2.e	2.f
2.a	∞					
2.b	33.5	∞				
2.c	4.8	5.8	∞			
2.d	3.6	4.2	19.3	∞		
2.e	2.8	3.3	9.2	18.3	∞	
2.f	2.6	3.0	7.7	13.5	27.0	∞

R	1.a	1.b	1.c	1.d	1.e	1.f
2.a	1.24	1.50	1.32	1.47	1.55	1.48
2.b	1.18	1.43	1.32	1.47	1.55	1.48
2.c	1.02	1.18	1.19	1.32	1.39	1.48
2.d	1.02	1.18	1.19	1.32	1.29	1.40
2.e	0.93	1.07	1.08	1.19	1.24	1.25
2.f	0.89	1.02	1.02	1.24	1.22	1.18

不匹配符号数

$$\beta = \max(|a|, |b|) - \alpha$$

相似性测度

$$R = \frac{\alpha}{\beta} = \frac{\alpha}{\max(|a|, |b|) - \alpha}$$

(a)和(b)两个不同目标类的  
样本边界;  
(c)和(d)对应的多边形近似;  
(e)~(g)R值表



# 第12章 目标识别

---

Thanks