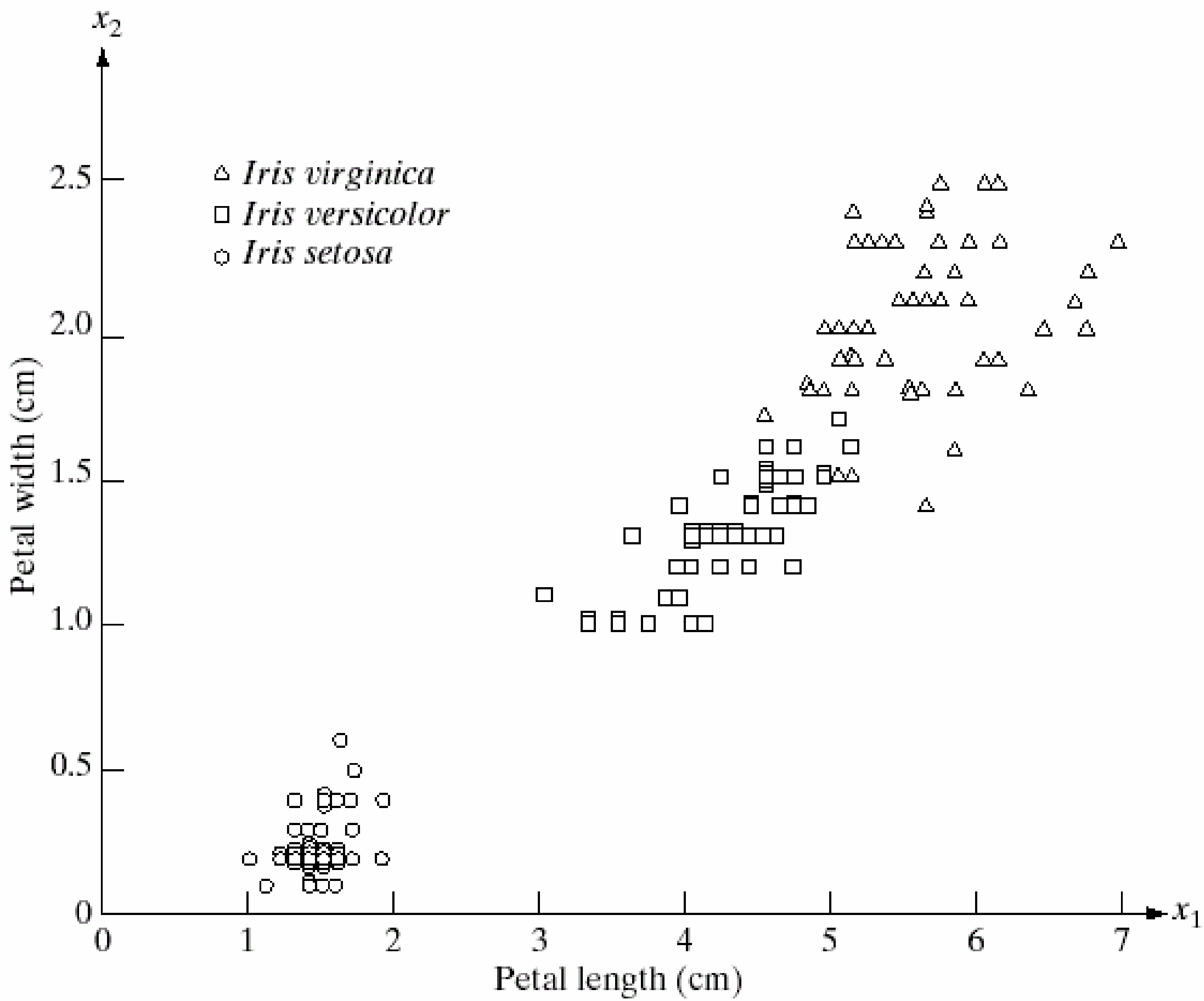
Chapter 12
Object Recognition

Yinghua He

- Patterns and Pattern Classes
- Recognition Based on Decision-Theoretic Methods
- Structural Methods

FIGURE 12.1
Three types of iris flowers described by two measurements.



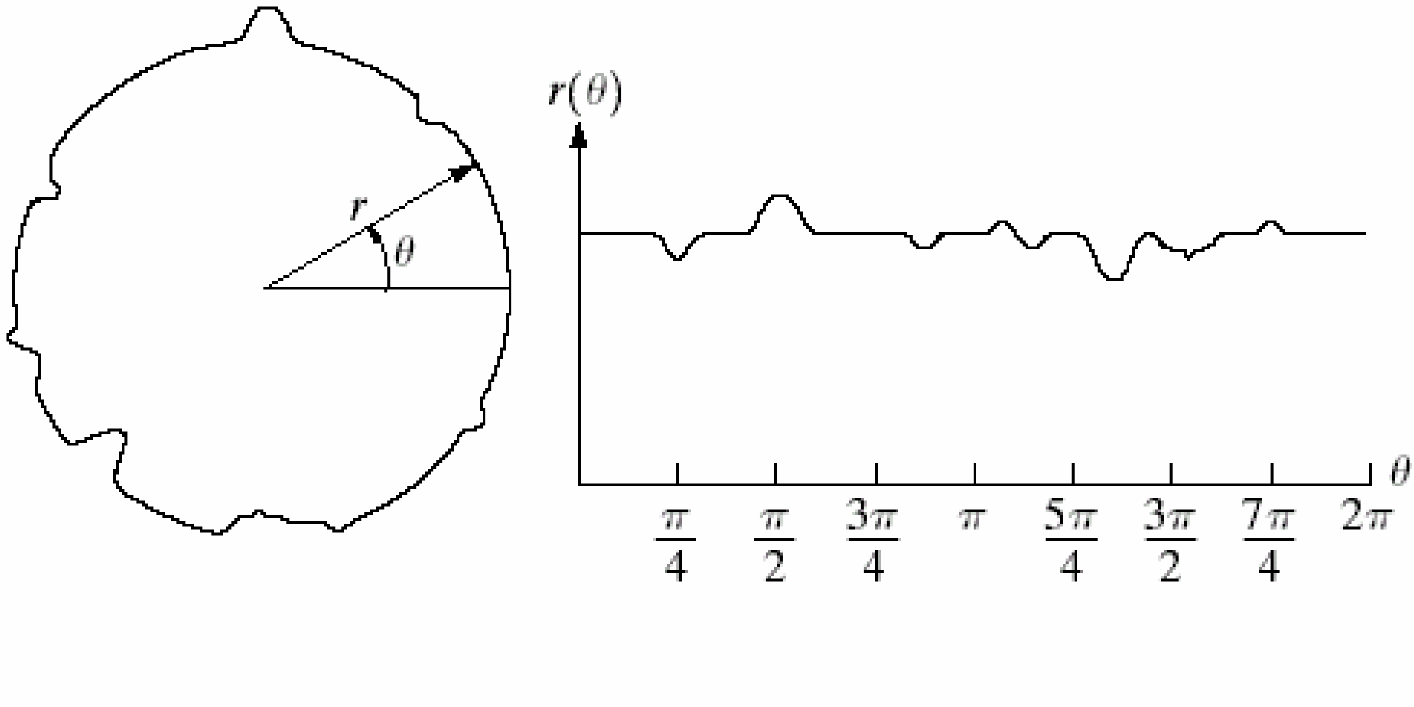
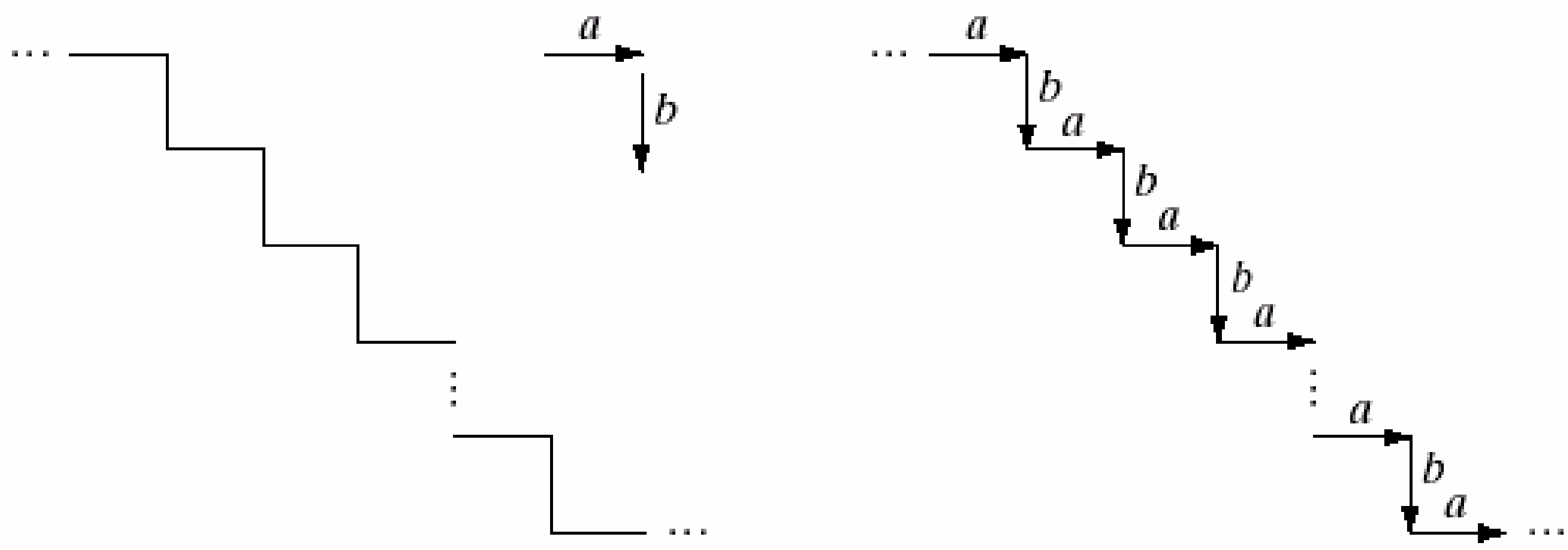


FIGURE 12.2 A noisy object and its corresponding signature.



a b

FIGURE 12.3 (a) Staircase structure. (b) Structure coded in terms of the primitives *a* and *b* to yield the string description . . . *ababab* . . . .

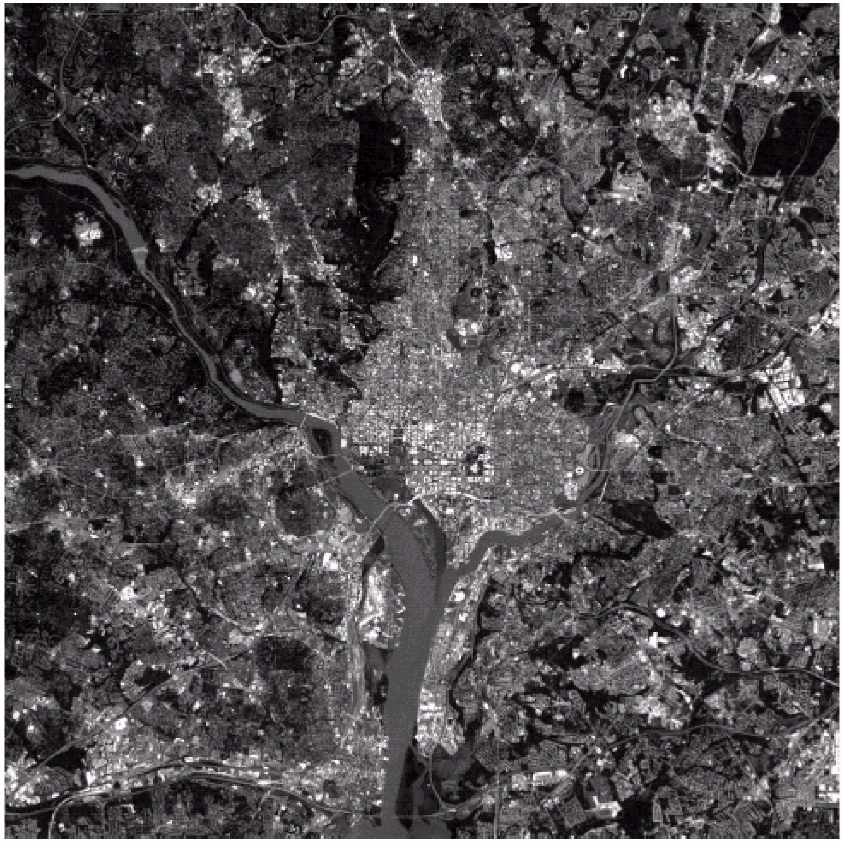


FIGURE 12.4 Satellite image of a heavily built downtown area (Washington, D.C.) and surrounding residential areas. (Courtesy of NASA.)

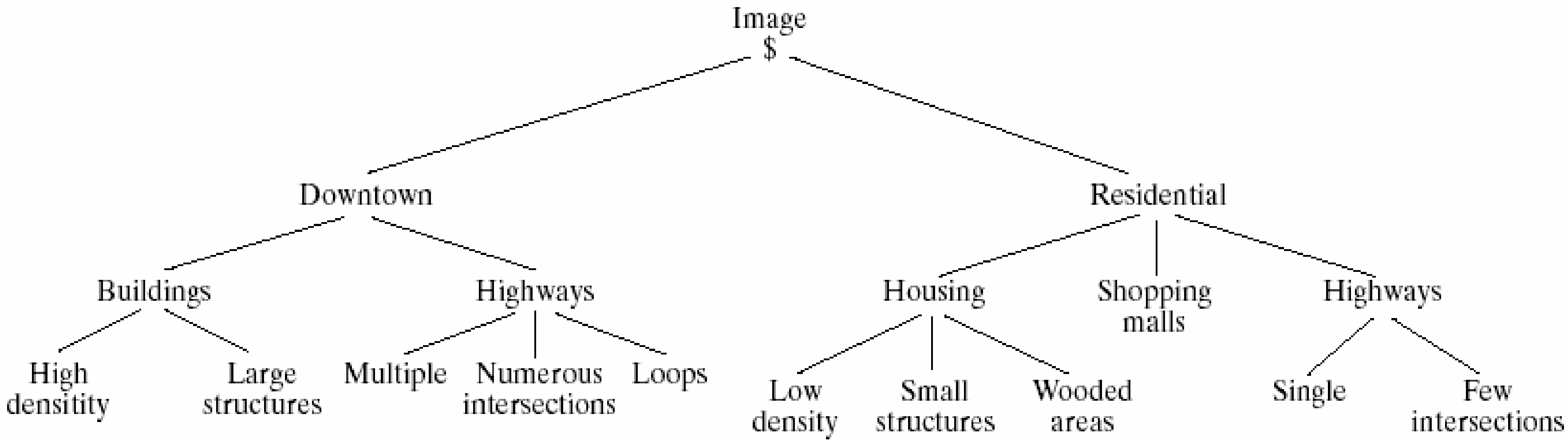


FIGURE 12.5 A tree description of the image in Fig. 12.4.

- Patterns and Pattern Classes
- Recognition Based on Decision-Theoretic Methods
- Structural Methods

- Matching
- Optimum Statistical Classifiers

$$m = \frac{1}{N_j} \sum_{x \in \omega_j} x_j \qquad j = 1,2,\ldots,W$$

$$D_j(x) = \left\| x - m_j \right\| \qquad j = 1,2,\ldots,W$$

$$\begin{cases} x \in \omega_i, & d_i(x) \geq d_j(x) \\ x \in \omega_j, & d_i(x) < d_j(x) \end{cases}$$

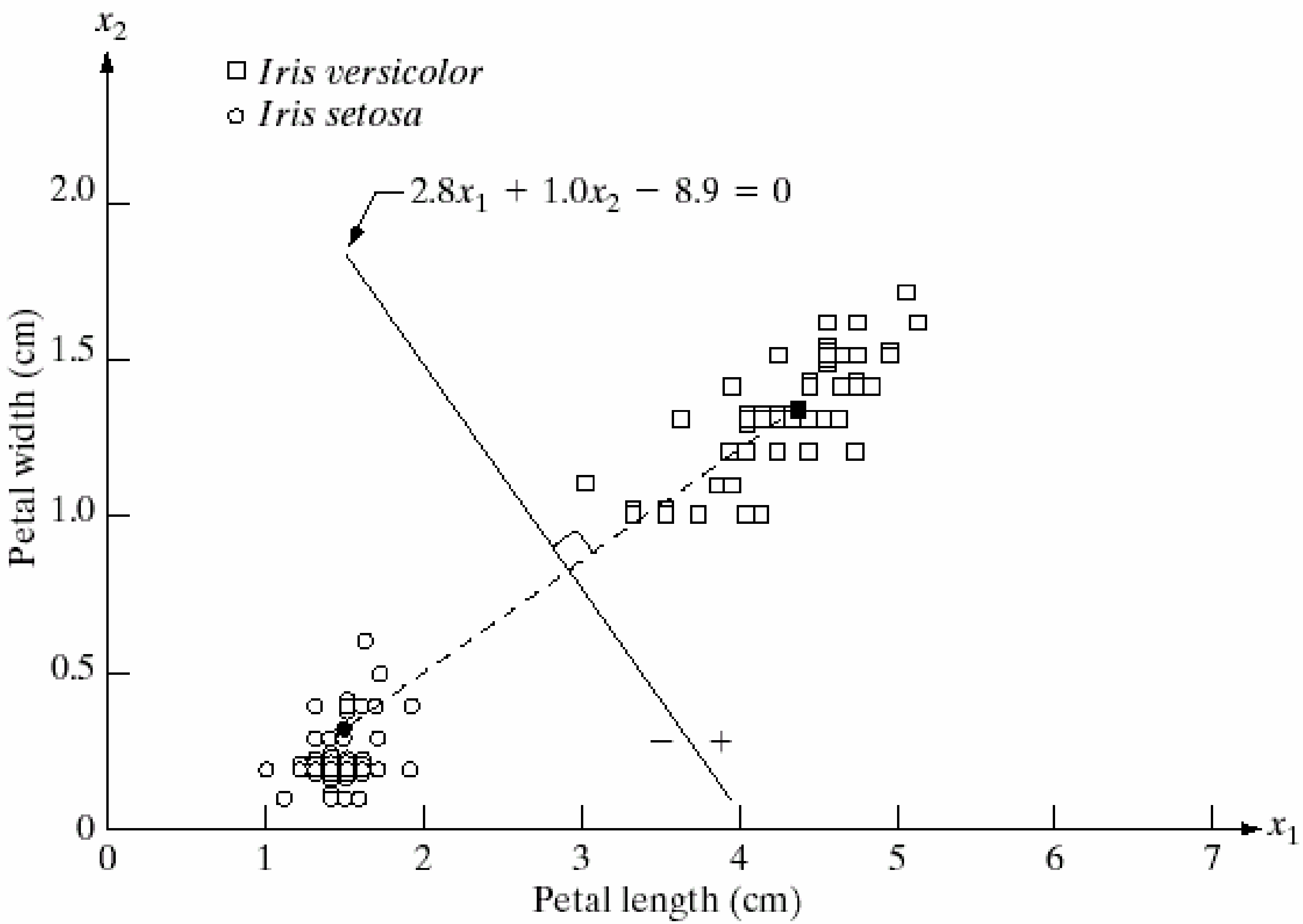


FIGURE 12.6 Decision boundary of minimum distance classifier for the classes of ***Iris versicolor*** and ***Iris setosa***. The dark dot and square are the means.

$$d_1(x) = x^T m_1 - \frac{1}{2} m_1^T m_1$$
$$= 4.3x_1 + 1.3x_2 - 10.1$$

$$d_2(x) = x^T m_2 - \frac{1}{2} m_2^T m_2$$
$$= 1.5x_1 + 0.3x_2 - 1.17$$

$$d_{12}(x) = d_1(x) - d_2(x)$$
$$= 2.8x_1 + 1.0x_2 - 8.9$$

- Matching
- Optimum Statistical Classifiers

- The conditional average risk or loss in decision theory terminology.

$$r_j(x) = \sum_{k=1}^{W} L_{kj}\, p(\omega_k / x)$$

$$r_j(x) = \frac{1}{p(x)} \sum_{k=1}^{W} L_{kj}\, p(\omega_k / x) P(\omega_k)$$

$$r_j(x) = \sum_{k=1}^{W} L_{kj}\, p(x / \omega_k) P(\omega_k)$$

$$\sum_{k=1}^{W} L_{ki}\, p(x/\omega_k) P(\omega_k) < \sum_{k=1}^{W} L_{qj}\, p(x/\omega_q) P(\omega_q)$$

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

Where $\delta_{ij} = 1$ if i=j and $\delta_{ij} = 0$ if $i \neq j$.

$$r_j(\mathbf{x}) = \sum_{k=1}^{W} (1 - \delta_{kj}) p(\mathbf{x}/\omega_k) P(\omega_k)$$
$$= p(\mathbf{x}) - p(\mathbf{x}/\omega_j) P(\omega_j).$$

The Bayes classifier then assigns a pattern $\mathbf{x}$ to class ω_i if, for all $j \neq i$,

$$p(\mathbf{x}) - p(\mathbf{x}/\omega_i)P(\omega_i) < p(\mathbf{x}) - p(\mathbf{x}/\omega_j)P(\omega_j) \qquad (12.2\text{-}15)$$

or, equivalently, if

$$p(\mathbf{x}/\omega_i)P(\omega_i) > p(\mathbf{x}/\omega_j)P(\omega_j) \qquad j = 1, 2, \dots, W;\ j \neq i. \qquad (12.2\text{-}16)$$

With reference to the discussion leading to Eq. (12.2-1), we see that the Bayes classifier for a 0-1 loss function is nothing more than computation of decision functions of the form

$$d_j(\mathbf{x}) = p(\mathbf{x}/\omega_j)P(\omega_j) \qquad j = 1, 2, \dots, W \qquad (12.2\text{-}17)$$

$$d_j(x) = p(x/\omega_j)P(\omega_j)$$

$$= \frac{1}{\sqrt{2\pi}\sigma_j} e^{-\frac{(x-m_j)^2}{2\sigma_j^2}} P(\omega_j) \qquad j = 1, 2$$

In the n-dimensional case, the Gaussian density of the vectors in the jth pattern class has the form

$$p(\mathbf{x}/\omega_j) = \frac{1}{(2\pi)^{n/2}|\mathbf{C}_j|^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\mathbf{m}_j)^T \mathbf{C}_j^{-1}(\mathbf{x}-\mathbf{m}_j)} \qquad (12.2\text{-}19)$$

$$\mathbf{m}_j = E_j\{\mathbf{x}\}$$

and

$$\mathbf{C}_j = E_j\{(\mathbf{x} - \mathbf{m}_j)(\mathbf{x} - \mathbf{m}_j)^T\}$$

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

and

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

$$\begin{aligned} d_j(\mathbf{x}) &= \ln\left[p(\mathbf{x}/\omega_j) P(\omega_j) \right] \\ &= \ln p(\mathbf{x}/\omega_j) + \ln P(\omega_j). \end{aligned}$$

$$d_j(\mathbf{x}) = \ln P(\omega_j) - \frac{n}{2} \ln 2\pi - \frac{1}{2} \ln|\mathbf{C}_j| - \frac{1}{2}\left[(\mathbf{x} - \mathbf{m}_j)^T \mathbf{C}_j^{-1} (\mathbf{x} - \mathbf{m}_j) \right]. \quad (12.2\text{-}25)$$

The term $(n/2) \ln 2\pi$ is the same for all classes, so it can be eliminated from Eq. (12.2-25), which then becomes

$$d_j(\mathbf{x}) = \ln P(\omega_j) - \frac{1}{2} \ln|\mathbf{C}_j| - \frac{1}{2}\left[(\mathbf{x} - \mathbf{m}_j)^T \mathbf{C}_j^{-1} (\mathbf{x} - \mathbf{m}_j) \right] \quad (12.2\text{-}26)$$

If all covariance matrices are equal, then $\mathbf{C}_j = \mathbf{C}$, for $j = 1, 2, \dots, W$. By expanding Eq. (12.2-26) and dropping all terms independent of j, we obtain

$$d_j(\mathbf{x}) = \ln P(\omega_j) + \mathbf{x}^T\mathbf{C}^{-1}\mathbf{m}_j - \frac{1}{2}\mathbf{m}_j^T\mathbf{C}^{-1}\mathbf{m}_j, \qquad (12.2\text{-}27)$$

which are linear decision functions (*hyperplanes*) for $j = 1, 2, \dots, W$.

If, in addition, $\mathbf{C} = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix, and also $P(\omega_j) = 1/W$, for $j = 1, 2, \dots, W$, then

$$d_j(\mathbf{x}) = \mathbf{x}^T\mathbf{m}_j - \frac{1}{2}\mathbf{m}_j^T\mathbf{m}_j \qquad j = 1, 2, \dots, W. \qquad (12.2\text{-}28)$$

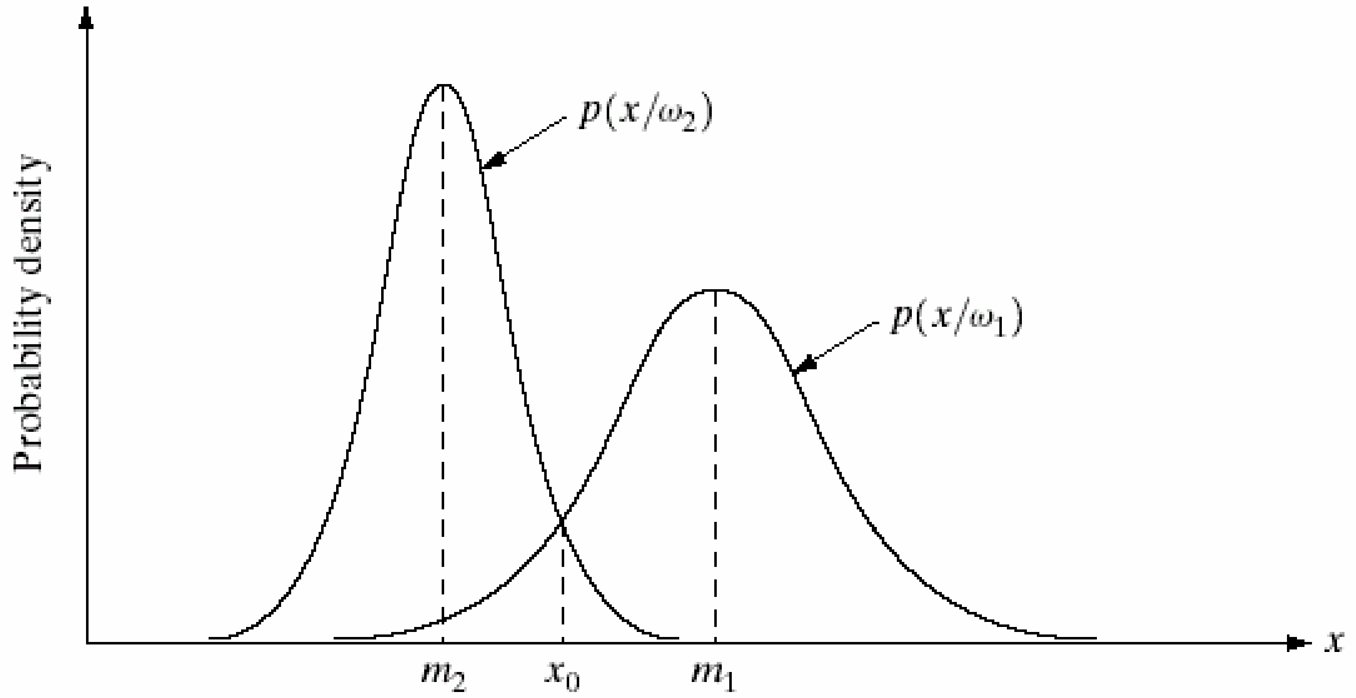
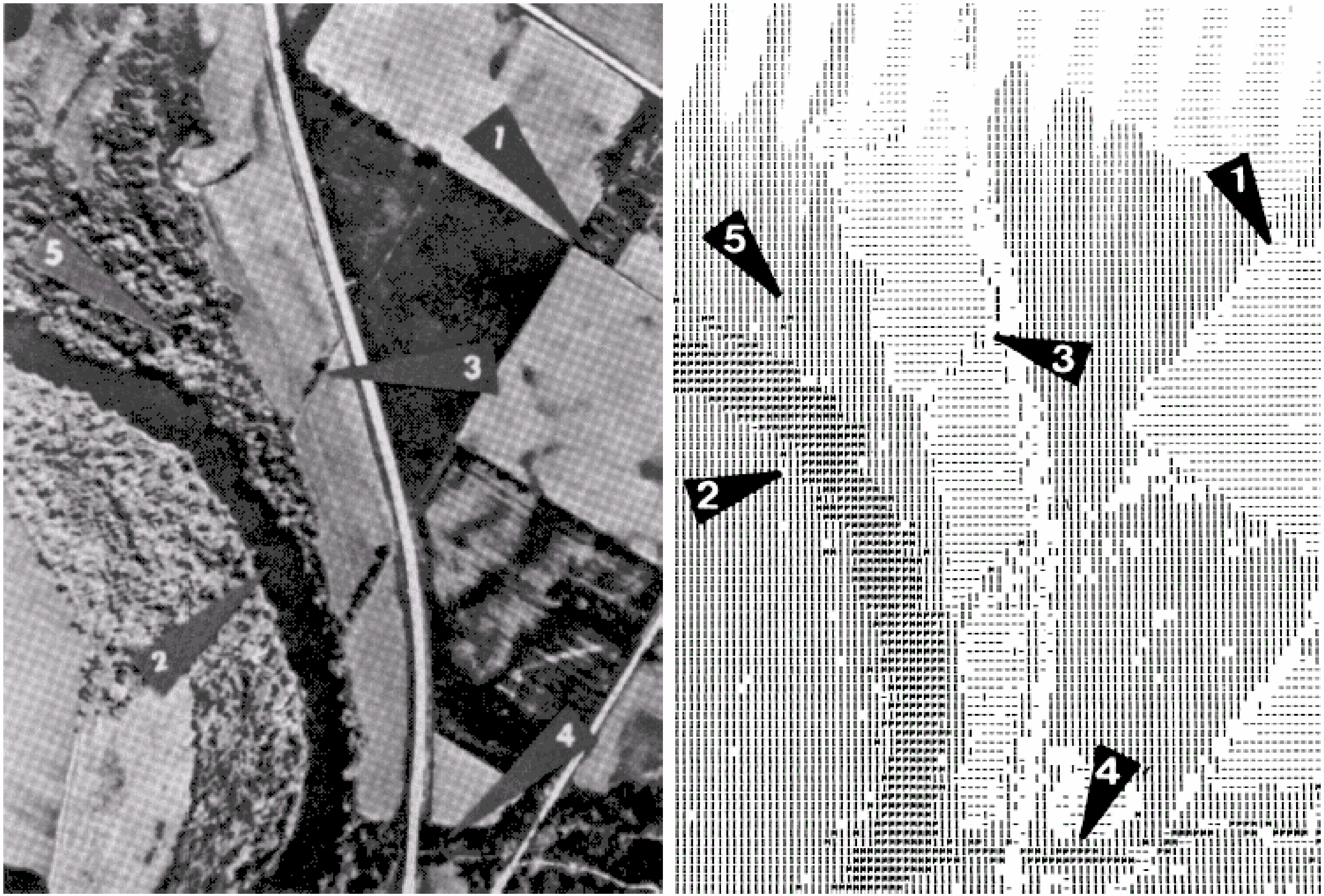


FIGURE 12.10 Probability density functions for two 1-D pattern classes. The point x_0 shown is the decision boundary if the two classes are equally likely to occur.



a b

FIGURE 12.13 (a) Multispectral image. (b) Printout of machine classification results using a Bayes classifier. (Courtesy of the Laboratory for Applications of Remote Sensing, Purdue University.)

- Patterns and Pattern Classes
- Recognition Based on Decision-Theoretic Methods
- Structural Methods

- Matching Shape Numbers
- String Matching

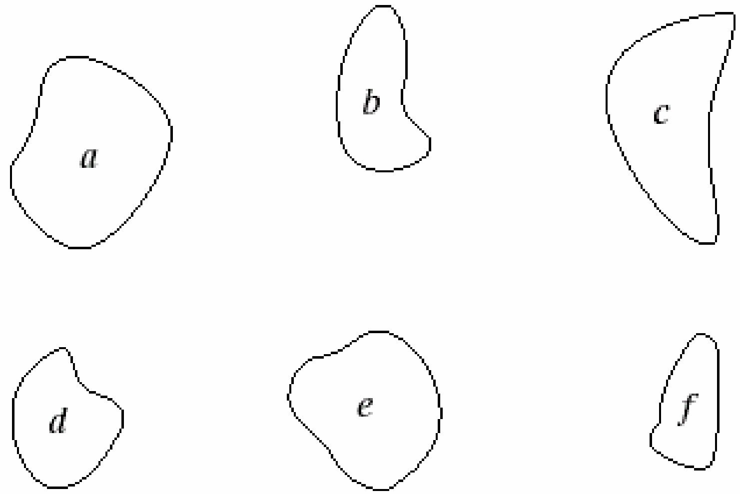
$$s_j(a) = s_j(b) \quad \text{for } j = 4, 6, 8, \dots, k$$
$$s_j(a) \neq s_j(b) \quad \text{for } j = k + 2, k + 4, \dots$$

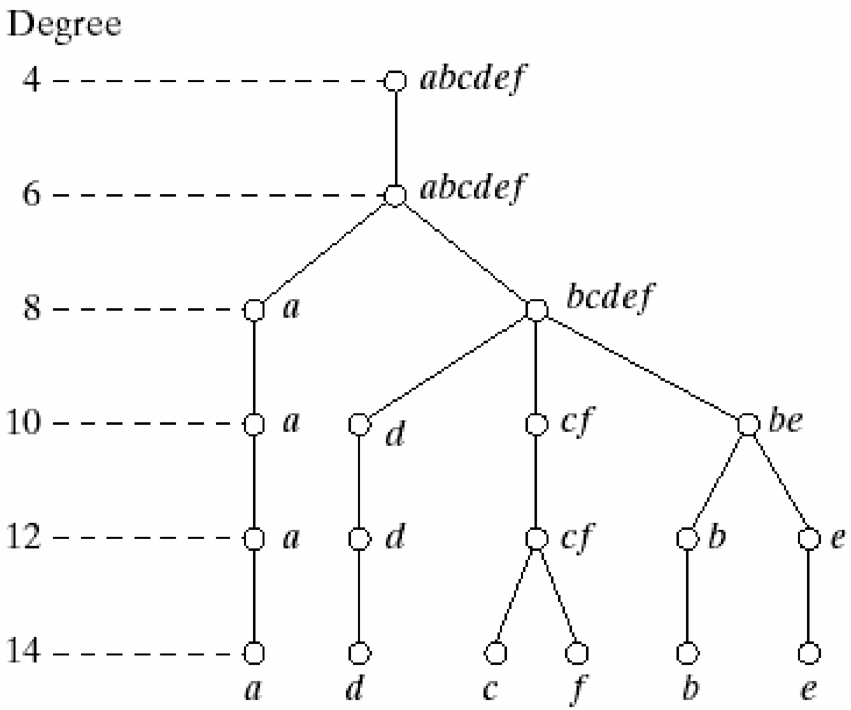
where s indicates shape number and the subscript indicates order. The *distance* between two shapes a and b is defined as the inverse of their degree of similarity:

$$D(a, b) = \frac{1}{k}. \qquad (12.3\text{-}2)$$

This distance satisfies the following properties:

$$D(a,b) \geq 0$$
$$D(a,b) = 0 \quad \text{iff } a = b$$
$$D(a,c) \leq \max\left[D(a,b), D(b,c)\right].$$





	a	*b*	*c*	*d*	*e*	*f*
a	∞	6	6	6	6	6
b		∞	8	8	10	8
c			∞	8	8	12
d				∞	8	8
e					∞	8
f						∞

a
b c

FIGURE 12.24 (a) Shapes. (b) Hypothetical similarity tree. (c) Similarity matrix. (Bribiesca and Guzman.)

- Matching Shape Numbers
- String Matching

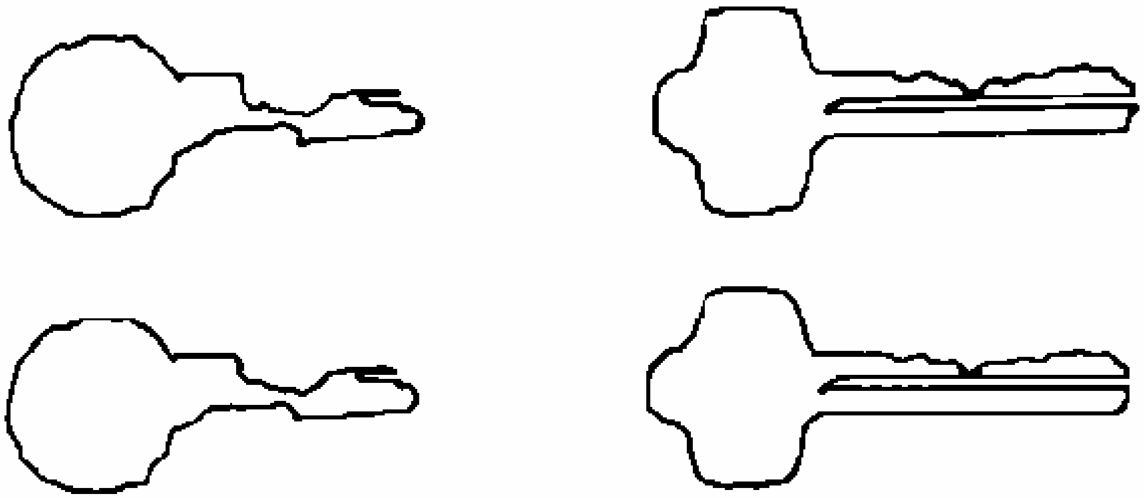
Suppose that two region boundaries, a and b, are coded into strings (see Section 11.5) denoted $a_1 a_2 \ldots a_n$ and $b_1 b_2 \ldots b_m$, respectively. Let α represent the number of matches between the two strings, where a match occurs in the kth position if $a_k = b_k$. The number of symbols that do not match is

$$\beta = \max(|a|, |b|) - \alpha \tag{12.3-4}$$

where |arg| is the length (number of symbols) in the string representation of the argument. It can be shown that $\beta = 0$ if and only if a and b are identical (see Problem 12.21).

A simple measure of similarity between a and b is the ratio

$$R = \frac{\alpha}{\beta} = \frac{\alpha}{\max(|a|, |b|) - \alpha}. \tag{12.3-5}$$



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

a b
c d
e f
g

FIGURE 12.25 (a) and (b) Sample boundaries of two different object classes; (c) and (d) their corresponding polygonal approximations; (e)–(g) tabulations of R. (Sze and Yang.)

The end