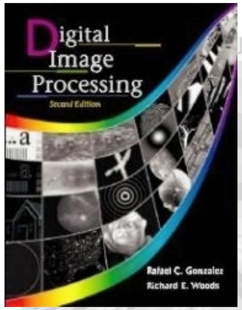
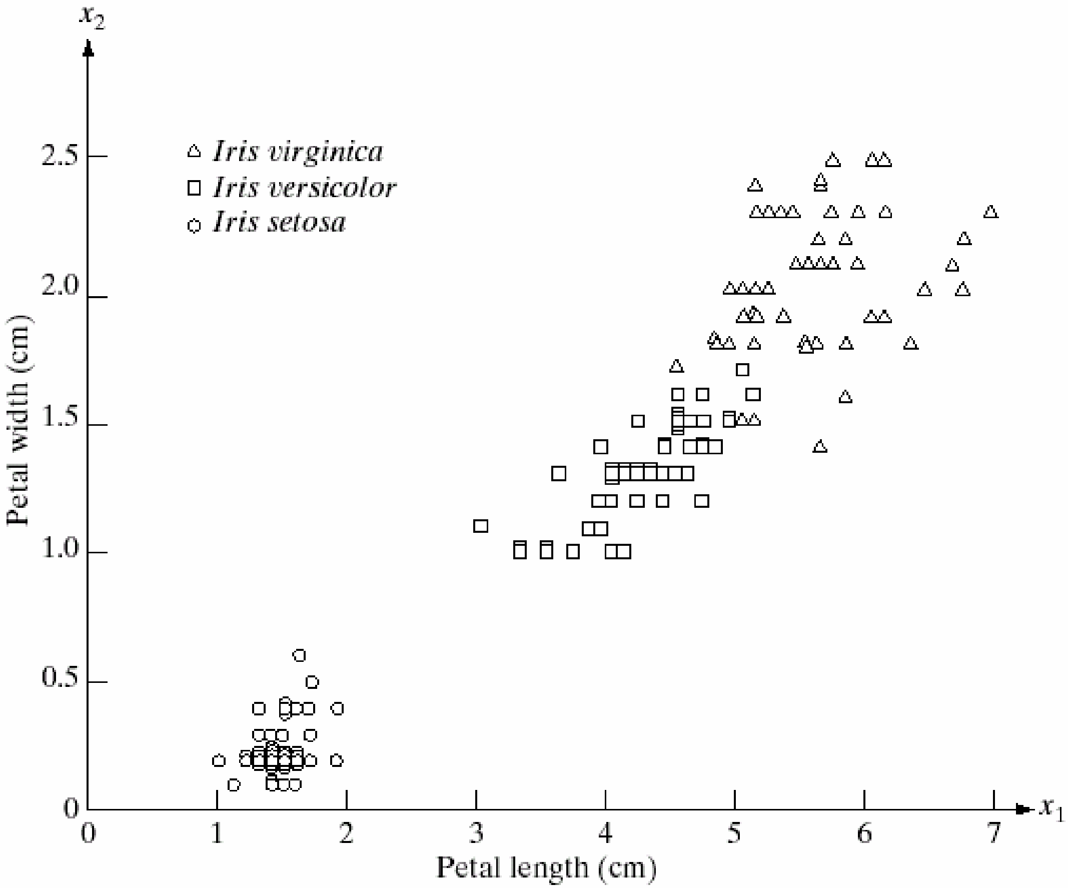
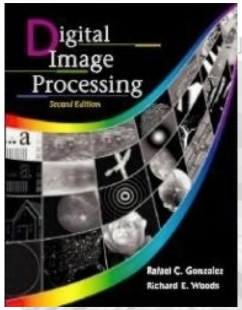


Chapter 12
Object Recognition

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





Patterns and Pattern Classes

- A *pattern* is an *arrangement of descriptors*, such as those discussed in Chapter 11.
- The name feature is used often in the pattern recognition literature to denote a descriptor.
- A pattern class is a family of patterns that share some common properties.

Chapter 12
Object Recognition

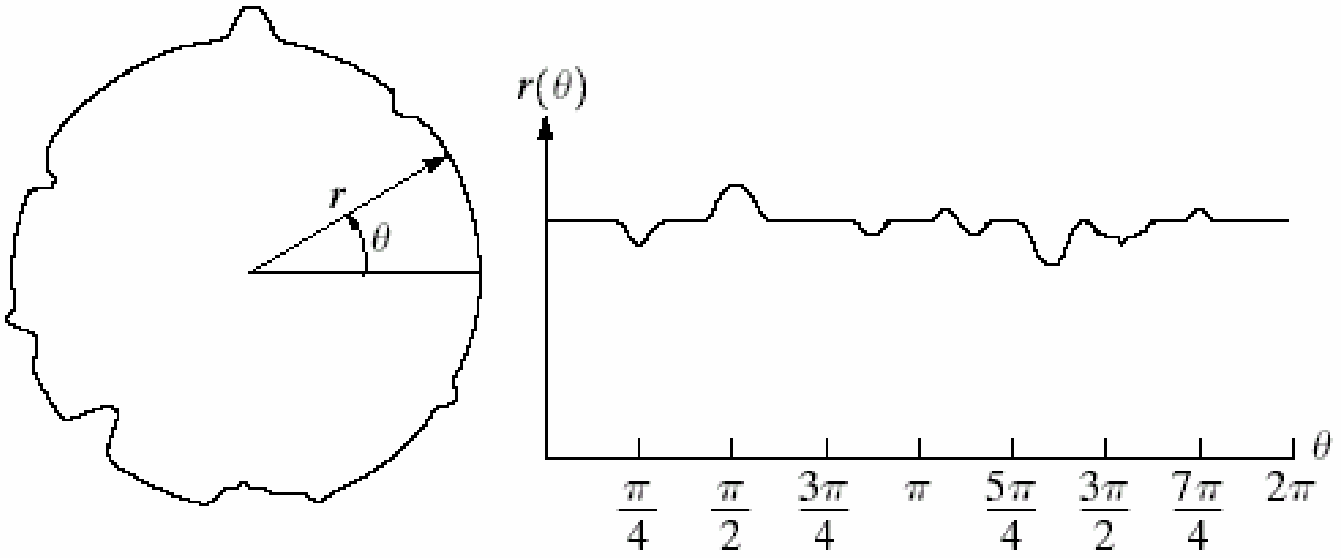


FIGURE 12.2 A noisy object and its corresponding signature.

Chapter 12
Object Recognition

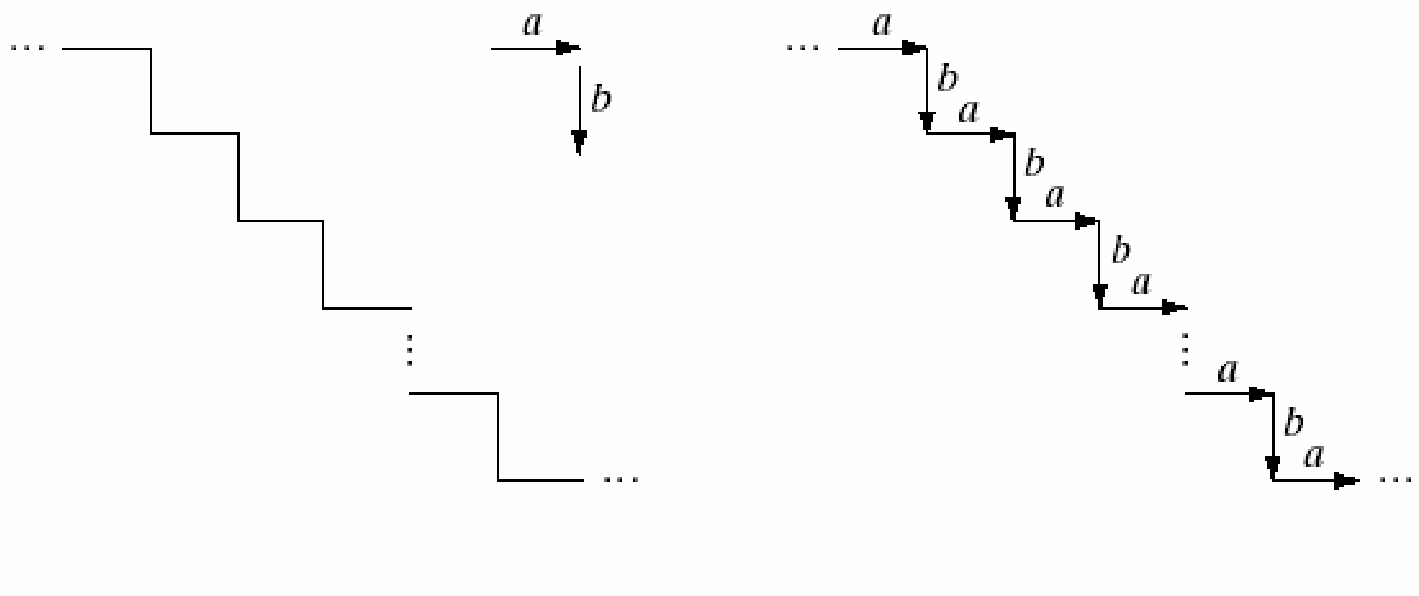


FIGURE 12.3 (a) Staircase structure. (b) Structure coded in terms of the primitives *a* and *b* to yield the string description . . . *ababab* . . . .

Chapter 12
Object Recognition

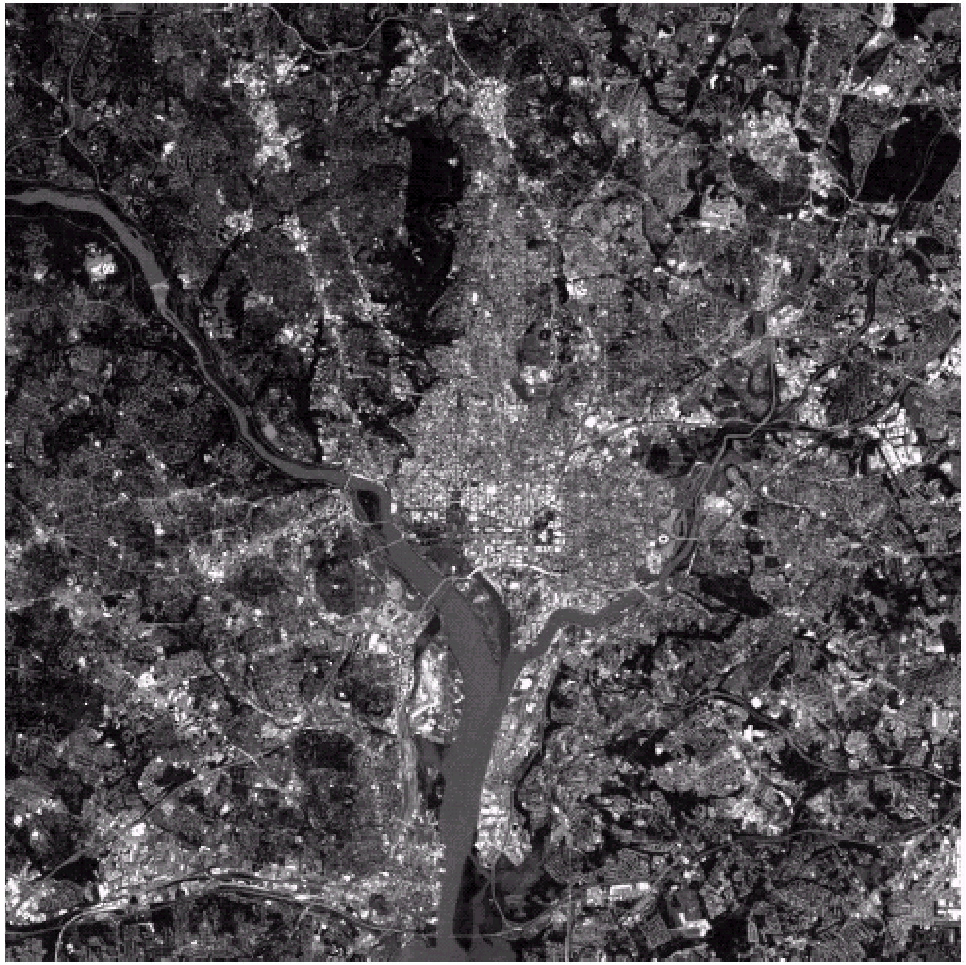


FIGURE 12.4 Satellite image of a heavily built downtown area (Washington, D.C.) and surrounding residential areas. (Courtesy of NASA.)

Chapter 12
Object Recognition

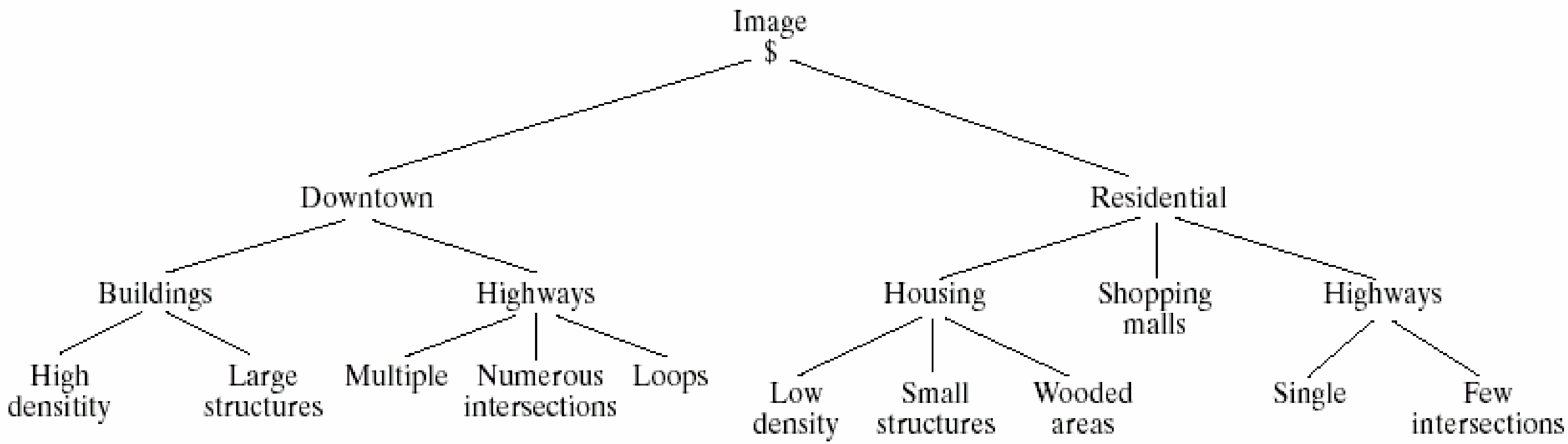


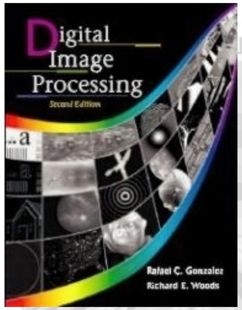
FIGURE 12.5 A tree description of the image in Fig. 12.4.

Recognition Based on Decision-Theoretic Methods

Let $x = (x_1, x_2, \ldots, x_n)^T$ for W pattern classes $\omega_1, \omega_2, \ldots, \omega_W$

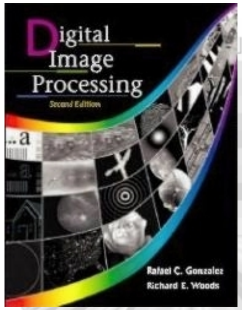
$$d_i(x) > d_j(x) \quad j = 1, 2, \ldots, W;\ j \neq i$$

- In other words, an unknown pattern **x** is said to belong to the *i*th pattern class if, upon substitution of **x** into all decision functions, $d_i(x)$ yields the largest numerical value.



Matching
Minimum distance classifier

- Suppose that we define the prototype of each pattern class to be the mean vector of the patterns of that class: $m_j = \frac{1}{N_j} \sum_{x \in \omega_j} x_j \quad j = 1, 2, \ldots, W$
- We then assign **x** to class ω_i if $D_i(\mathbf{x})$ is the smallest distance. $D_j(x) = \left\| x - m_j \right\|$



Minimum distance classifier

- It is not difficult to show (Problem 12.2) that selecting the smallest distance is equivalent to evaluating the functions $d_j(x) = x^T m_j - \frac{1}{2} m_j^T m_j \quad j = 1,2,\ldots,W$
- assign **x** to class ω_i if $d_i(\mathbf{x})$ is the largest numerical value.
- This formulation agrees with the concept of a decision function, as defined in Eq. (12.2-1).

Chapter 12
Object Recognition

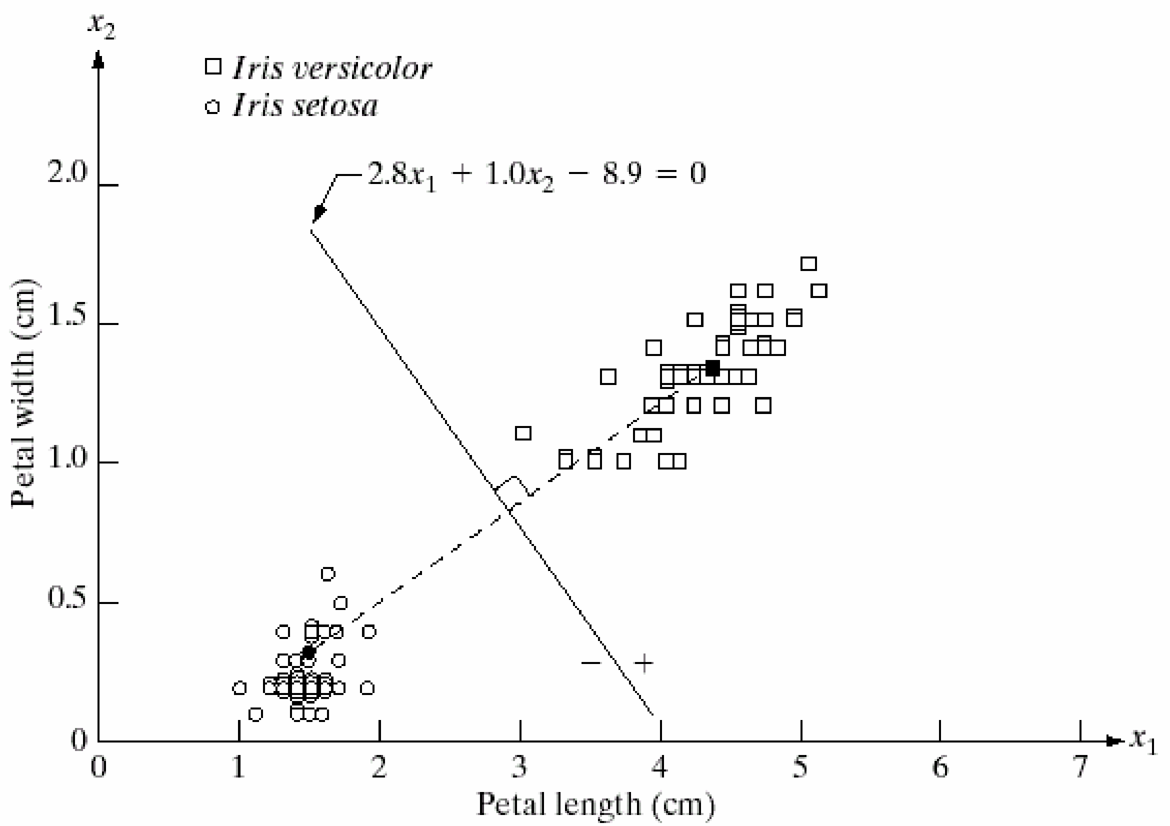
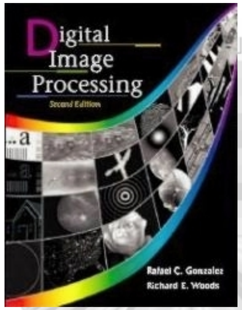


FIGURE 12.6 Decision boundary of minimum distance classifier for the classes of ***Iris versicolor*** and ***Iris setosa***. The dark dot and square are the means.



Chapter 12
Object Recognition

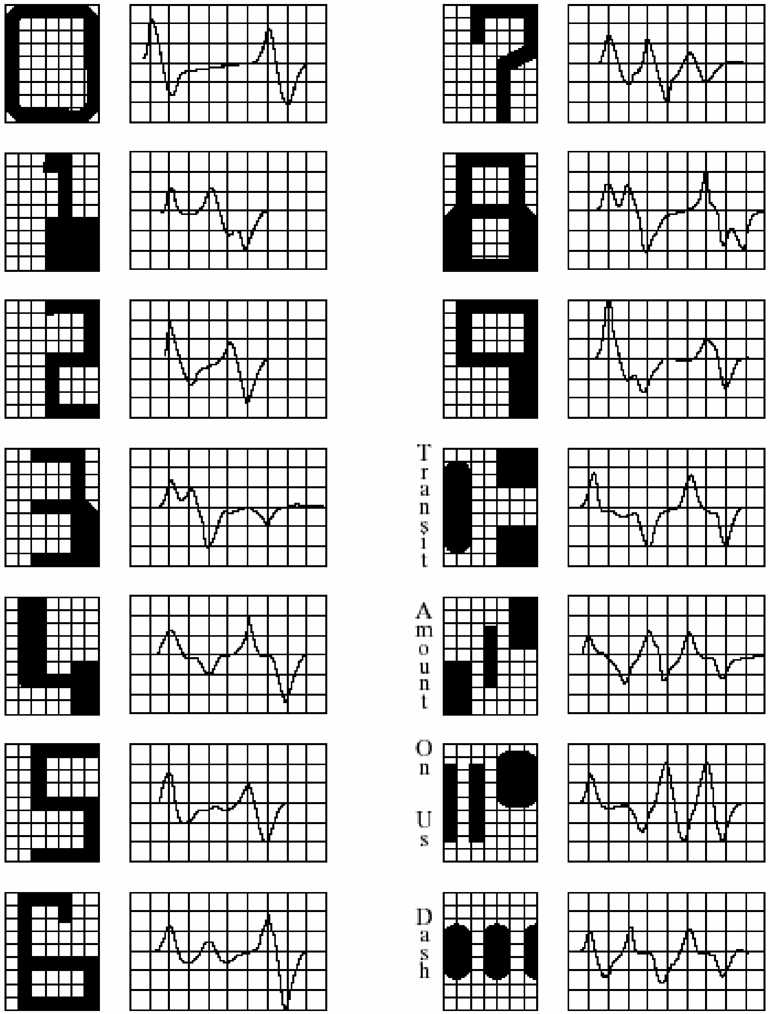
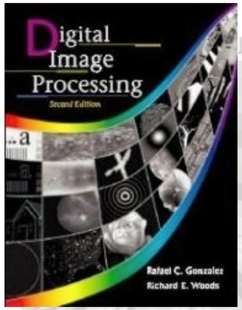


FIGURE 12.7 American Bankers Association E-13B font character set and corresponding waveforms.



Matching by correlation

$$c(x,y)=\sum_{s}\sum_{t} f(s,t)\,w(x+s,y+t)$$

- *correlation coefficient*, which is defined as

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

Chapter 12
Object Recognition

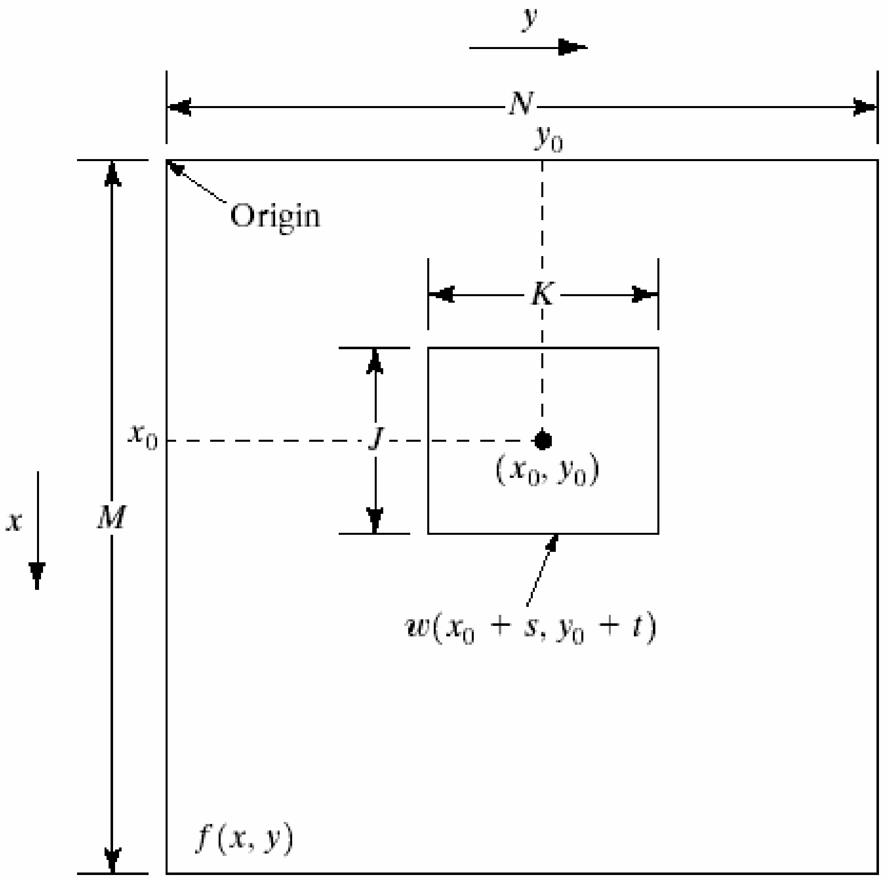
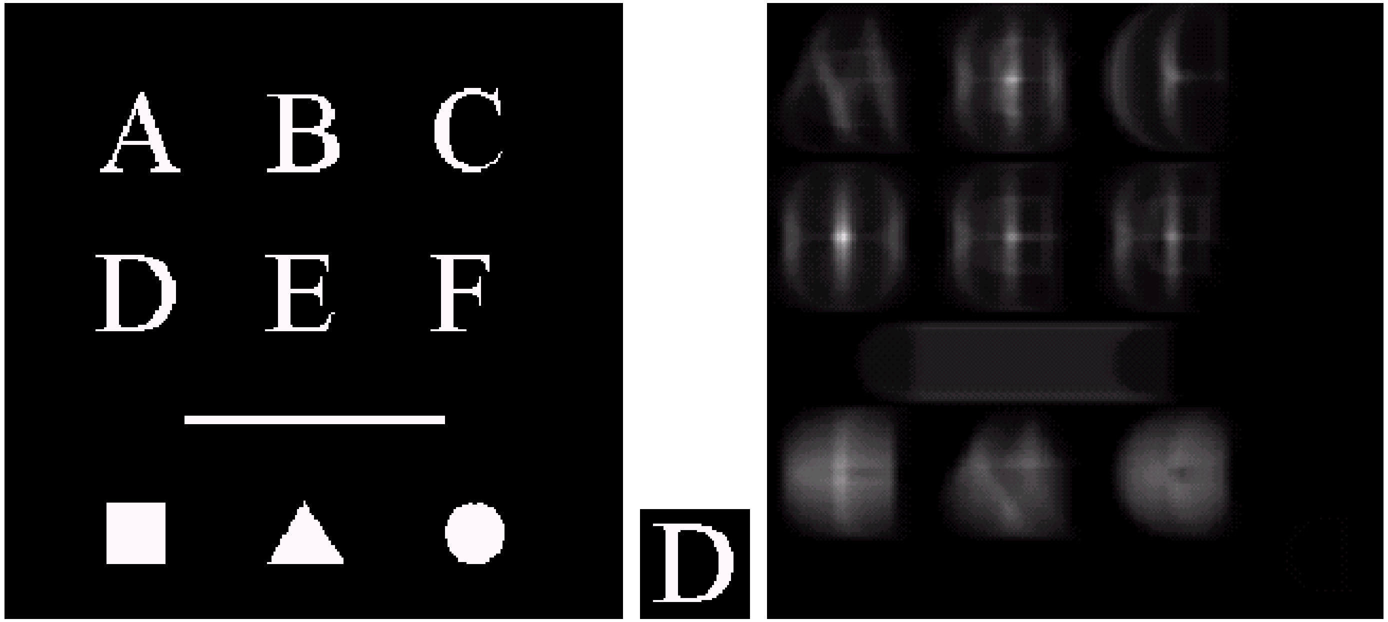


FIGURE 12.8 Arrangement for obtaining the correlation of f and w at point (x_0, y_0).

Chapter 12
Object Recognition



a b c

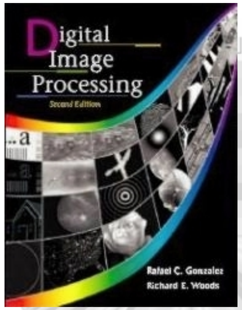
FIGURE 12.9 (a) Image. (b) Subimage. (c) Correlation coefficient of (a) and (b). Note that the highest (brighter) point in (c) occurs when subimage (b) is coincident with the letter "D" in (a).

Optimum Statistical Classifiers

- $r_j(x) = \sum_{k=1}^{W} L_{kj}\, p(\omega_k / x)$ the equation often called the *conditional average risk* or *loss* in decision-theory terminology.
- The classifier that minimizes the total average loss is called the *Bayes classifier*.
- Thus the Bayes classifier assigns an unknown pattern x to class ω_i if $r_i(x) < r_j(x) \quad for \; j = 1, 2, \ldots, W;\; j \neq i$

in other words, **x** is assigned to class ω_i if

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



Bayes classifier for Gaussian pattern classes

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

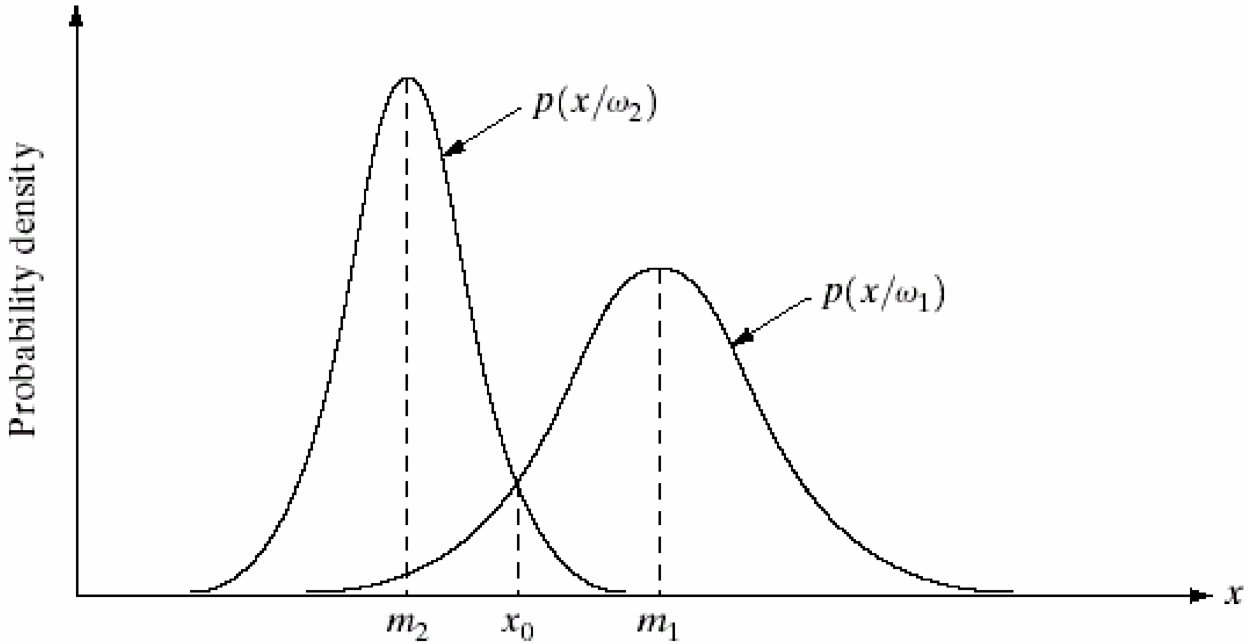
$$p(x/\omega_j) = \frac{1}{(2\pi)^{n/2}\left|C_j\right|^{n/2}} e^{-\frac{1}{2}(x-m_j)^T C_j^{-1}(x-m_j)} \qquad m_j = E_j\{x\}$$

$$C_j = E_j\left\{(x-m_j)(x-m_j)^T\right\} \qquad C_j = \frac{1}{N_j}\sum_{x\in\omega_j} xx^T - m_j m_j^T$$

Bayes decision function for class ω_j is $d_j(x) = p(x/\omega_j)P(\omega_j)$

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

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.



Chapter 12
Object Recognition

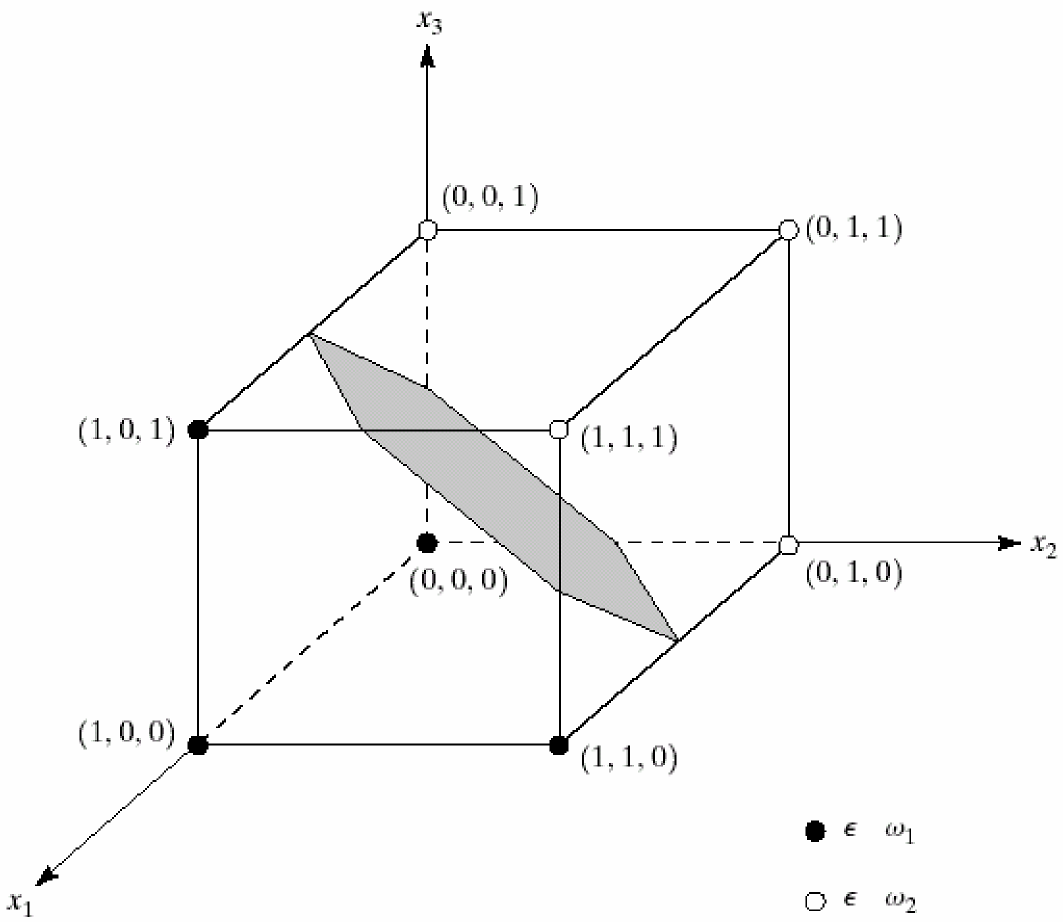
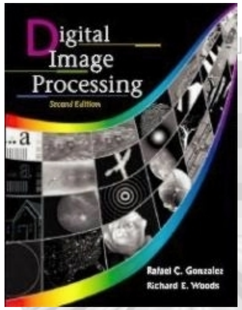
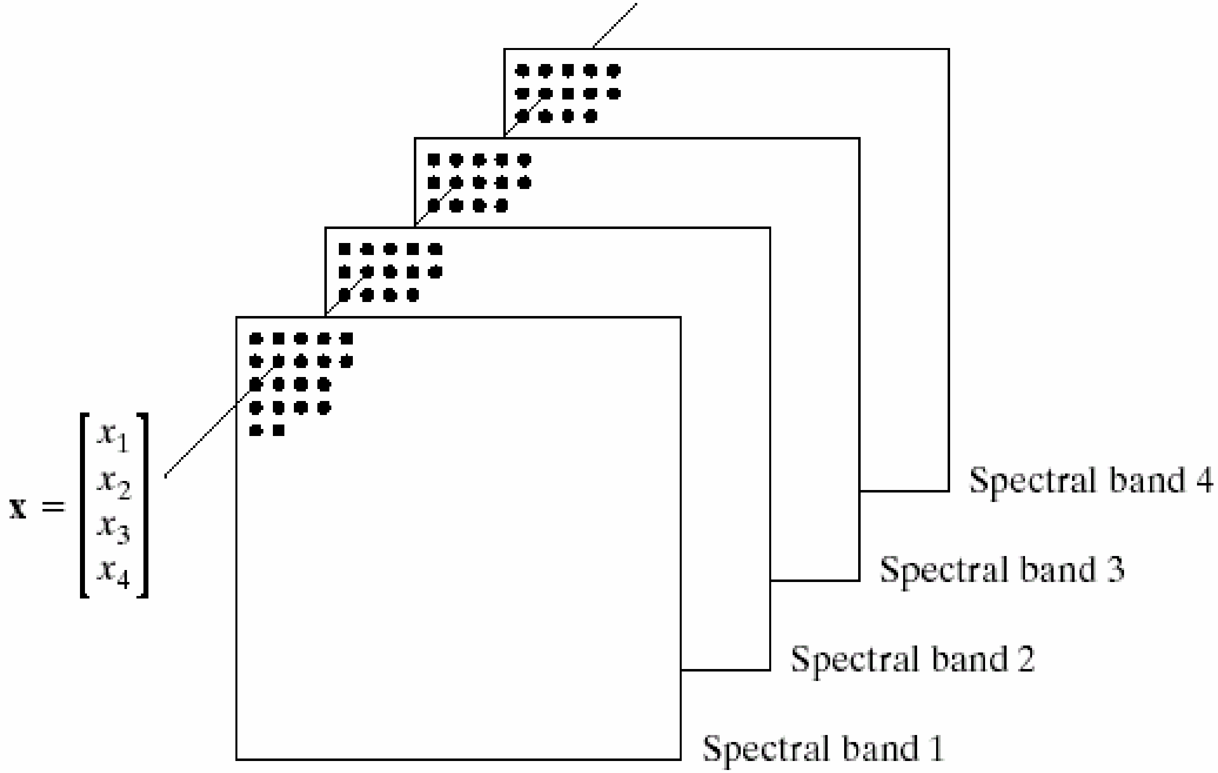


FIGURE 12.11 Two simple pattern classes and their Bayes decision boundary (shown shaded).

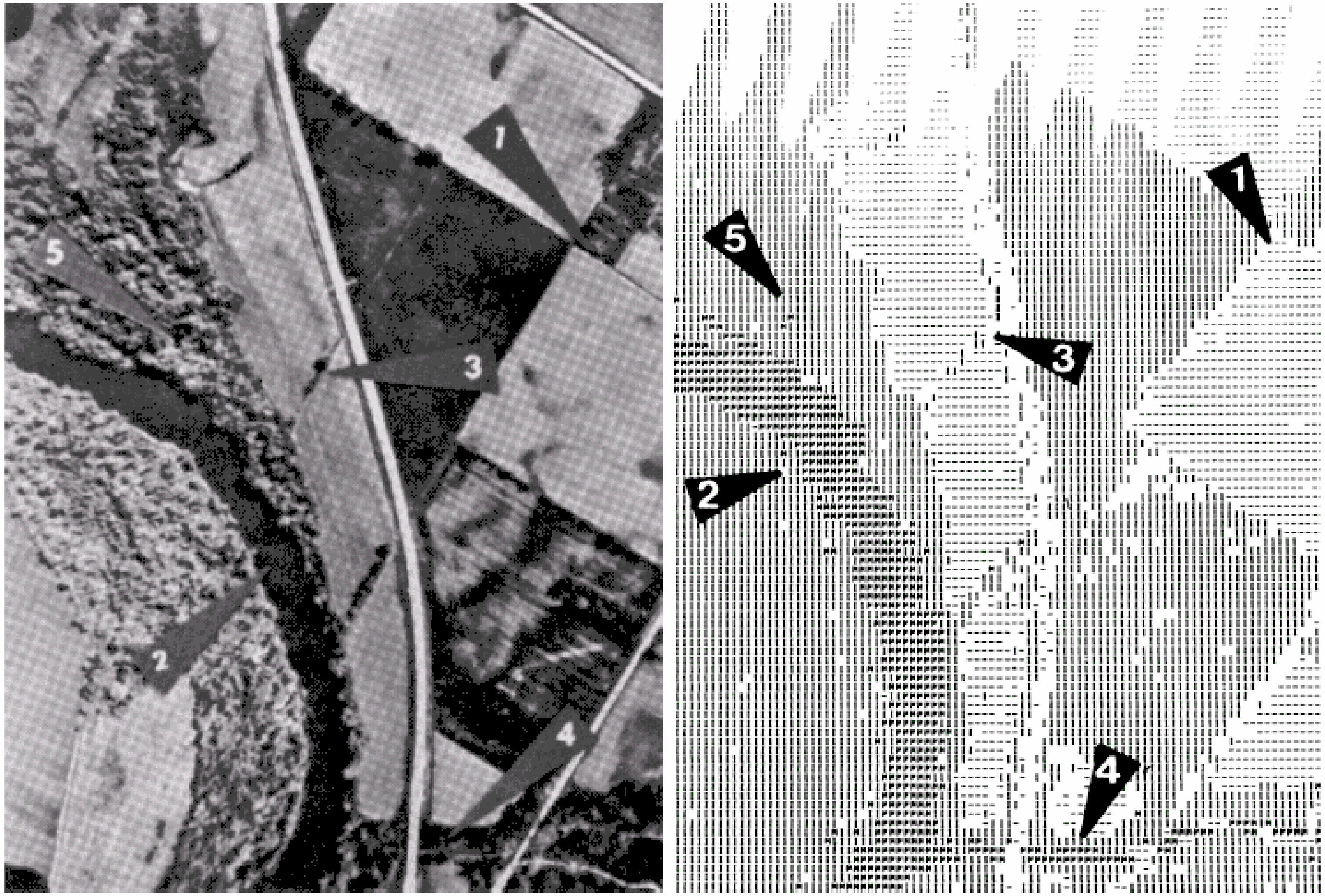


Chapter 12
Object Recognition

FIGURE 12.12
Formation of a pattern vector from registered pixels of four digital images generated by a multispectral scanner.



Chapter 12
Object Recognition



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.)

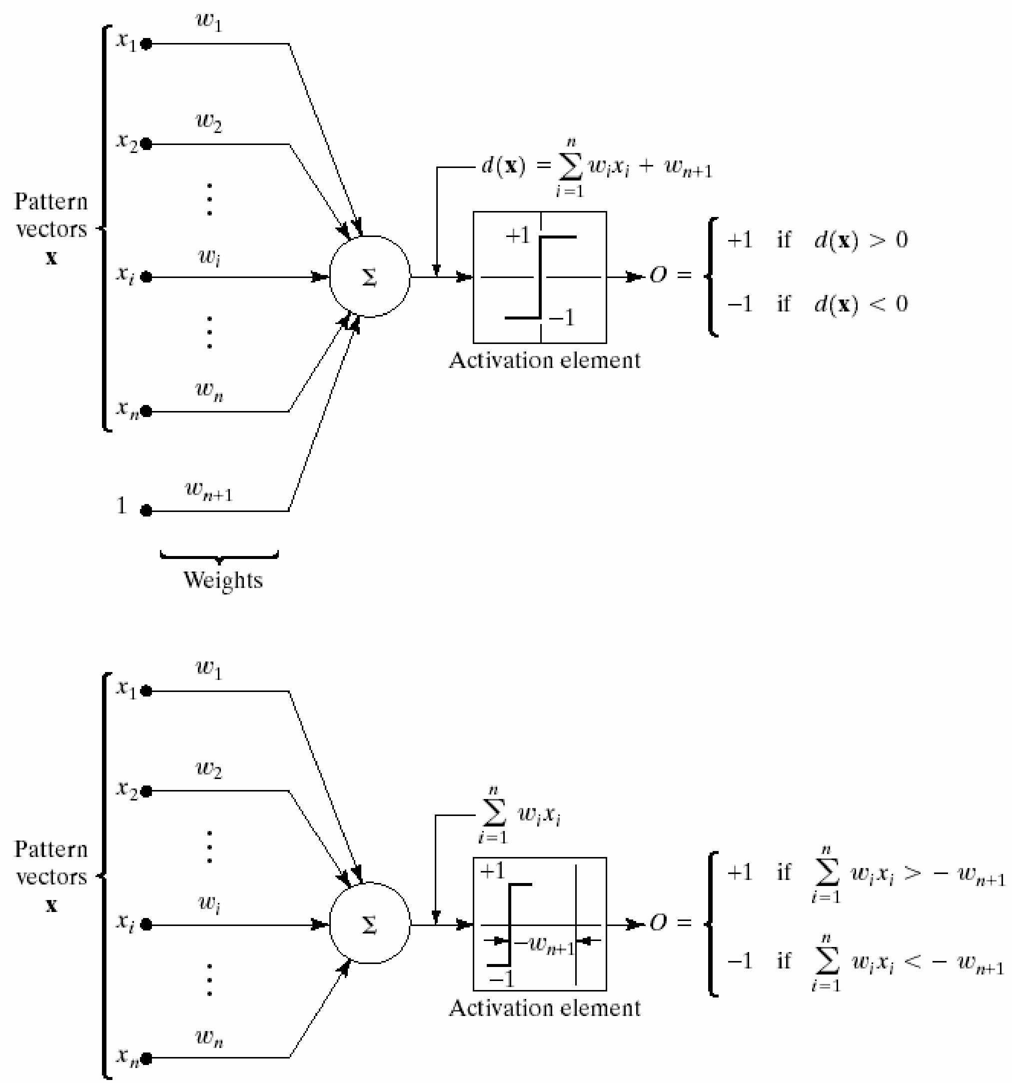
Neural Network

$$w(k+1) = w(k) - \alpha \left[\frac{\partial J(w)}{\partial w} \right]_{w=w(k)}$$

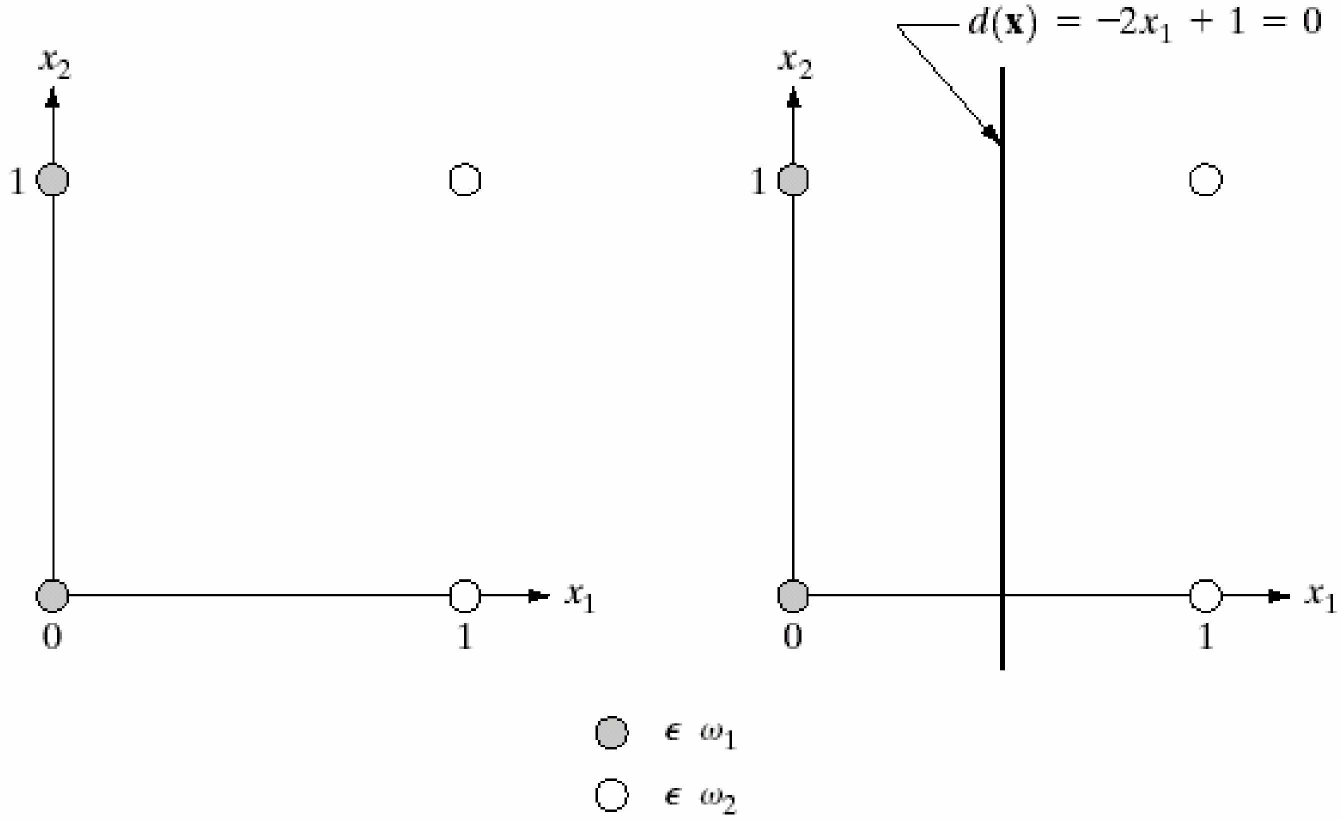
Chapter 12
Object Recognition

a
b

FIGURE 12.14 Two equivalent representations of the perceptron model for two pattern classes.

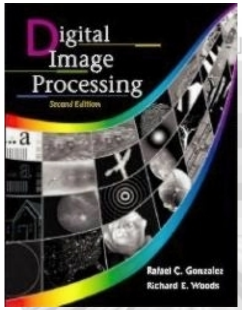


Chapter 12
Object Recognition



a b

FIGURE 12.15
(a) Patterns belonging to two classes.
(b) Decision boundary determined by training.



Chapter 12
Object Recognition

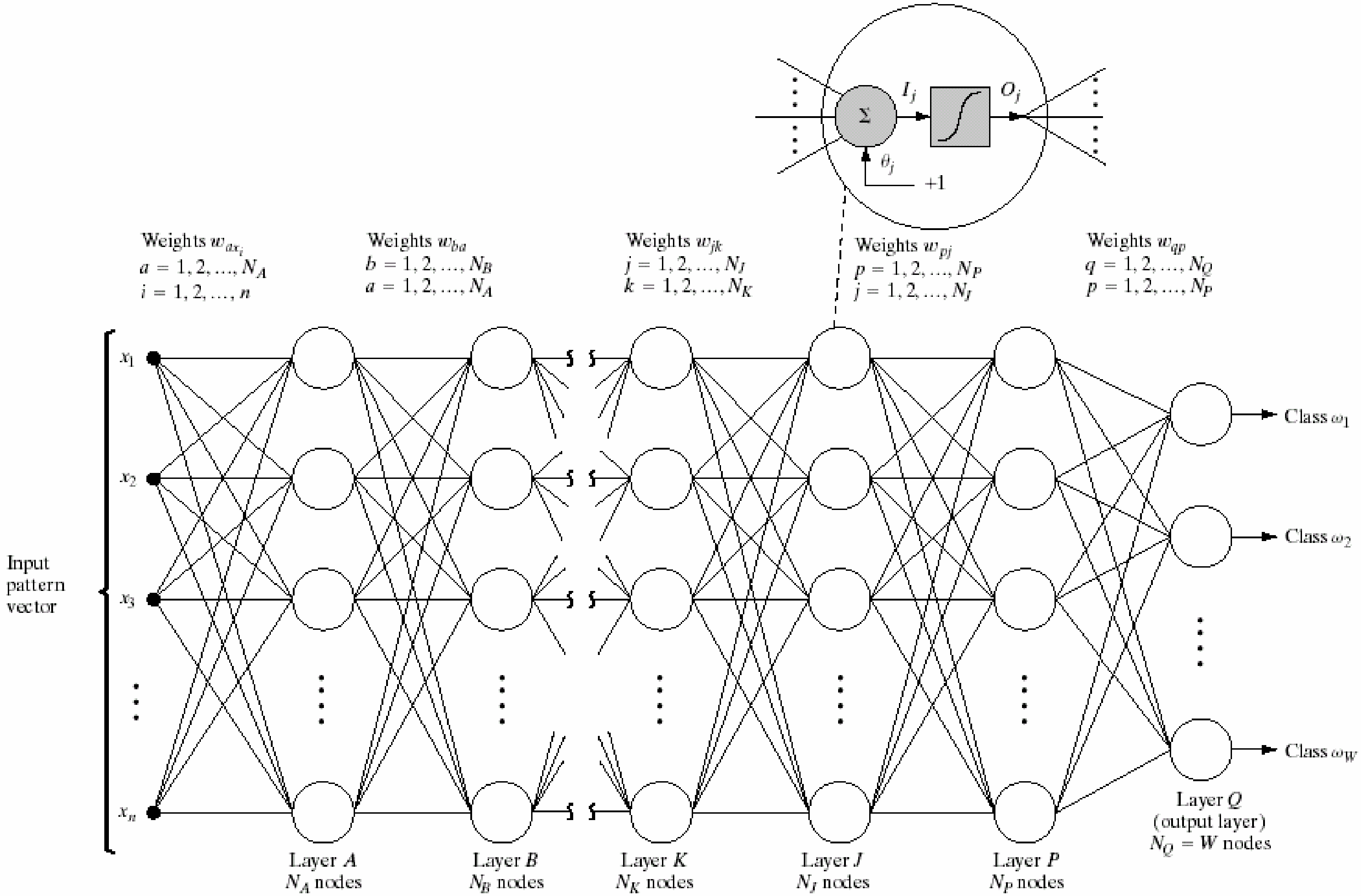


FIGURE 12.16 Multilayer feedforward neural network model. The blowup shows the basic structure of each neuron element throughout the network. The offset, θ_j, is treated as just another weight.

Chapter 12
Object Recognition

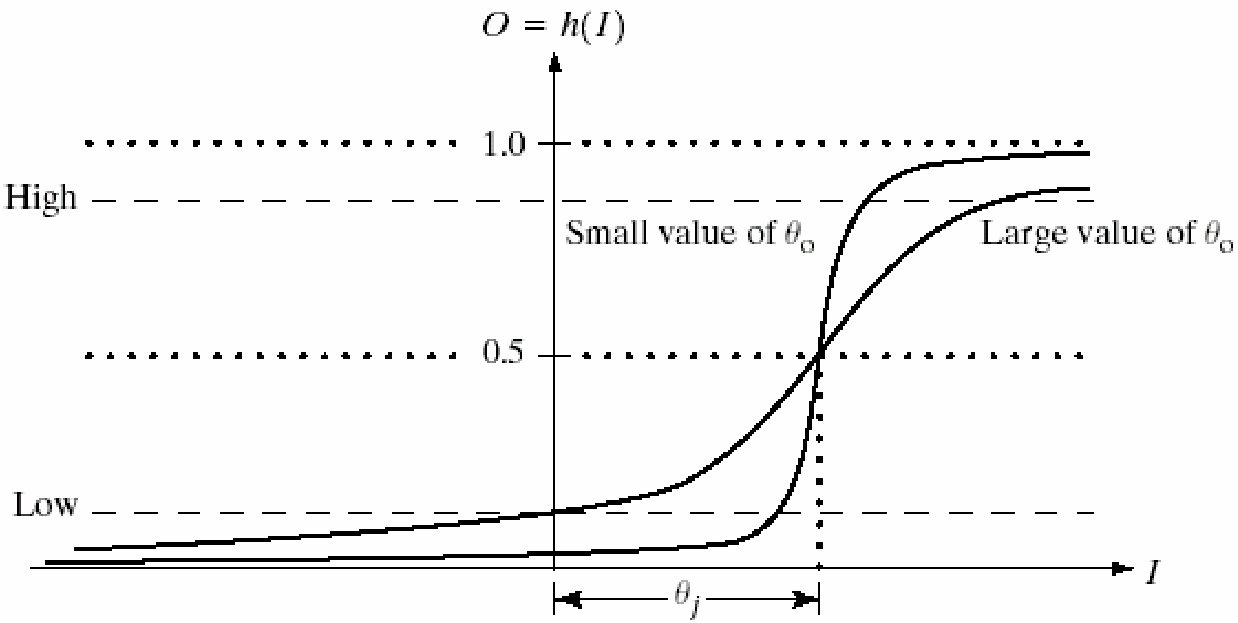
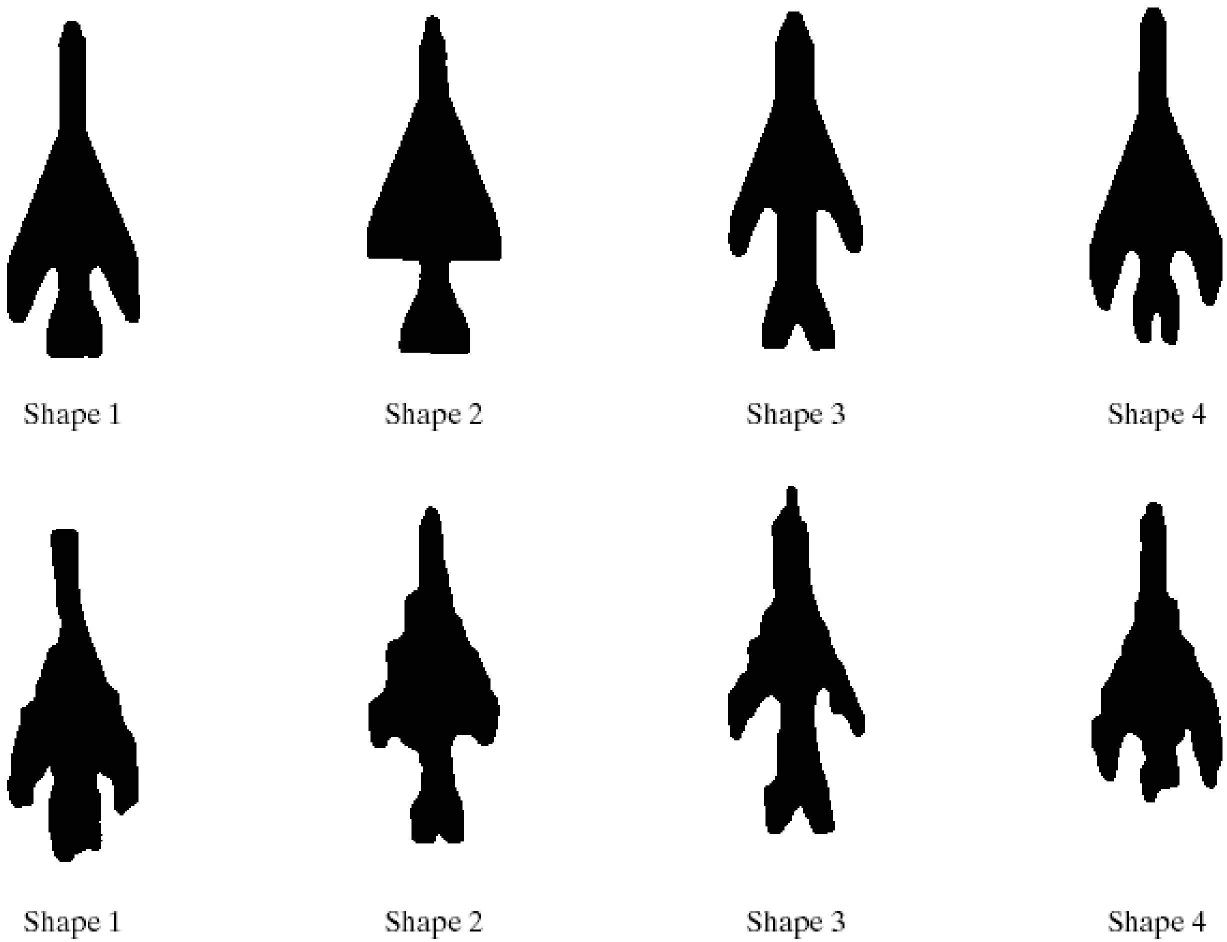


FIGURE 12.17 The sigmoidal activation function of Eq. (12.2-47).

Chapter 12
Object Recognition



a
b

FIGURE 12.18 (a) Reference shapes and (b) typical noisy shapes used in training the neural network of Fig. 12.19. (Courtesy of Dr. Lalit Gupta, ECE Department, Southern Illinois University.)

Chapter 12
Object Recognition

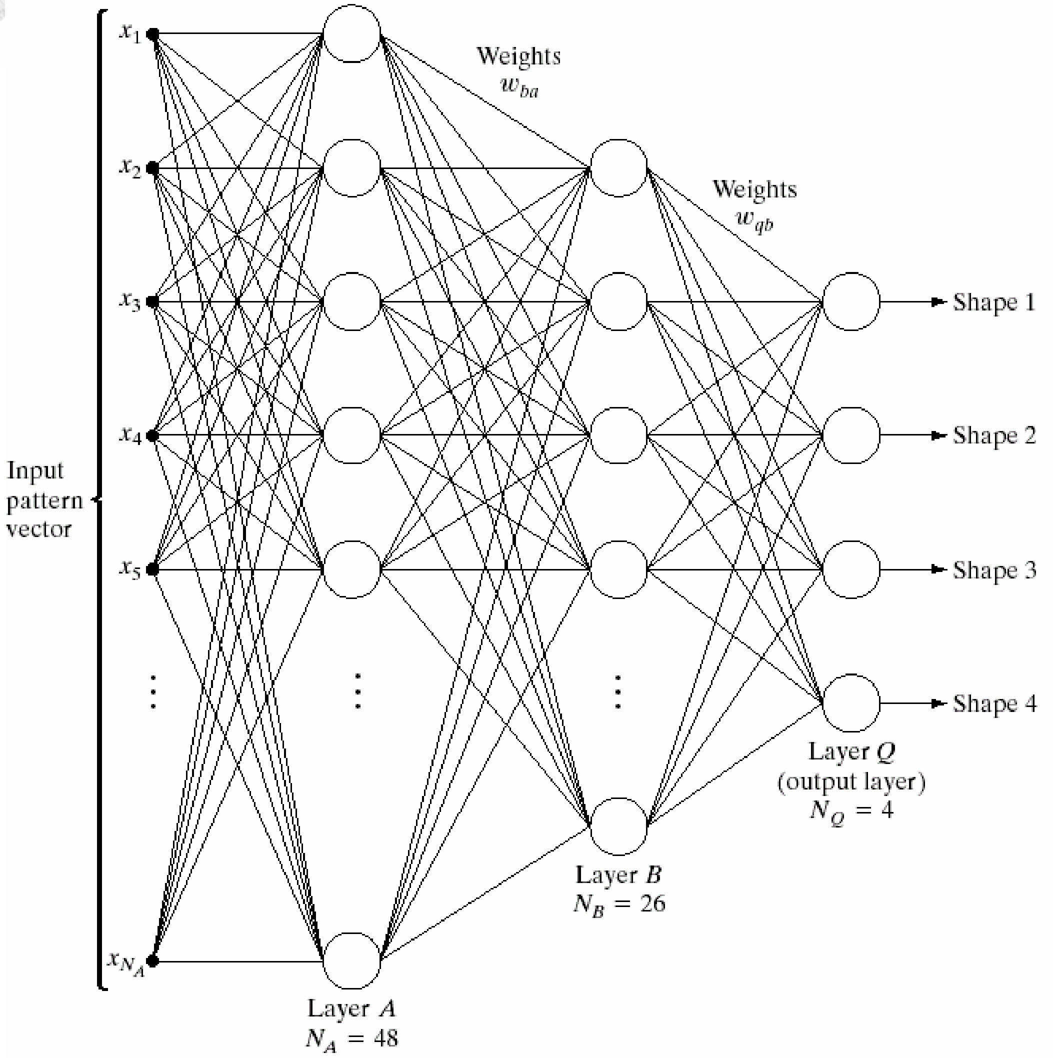
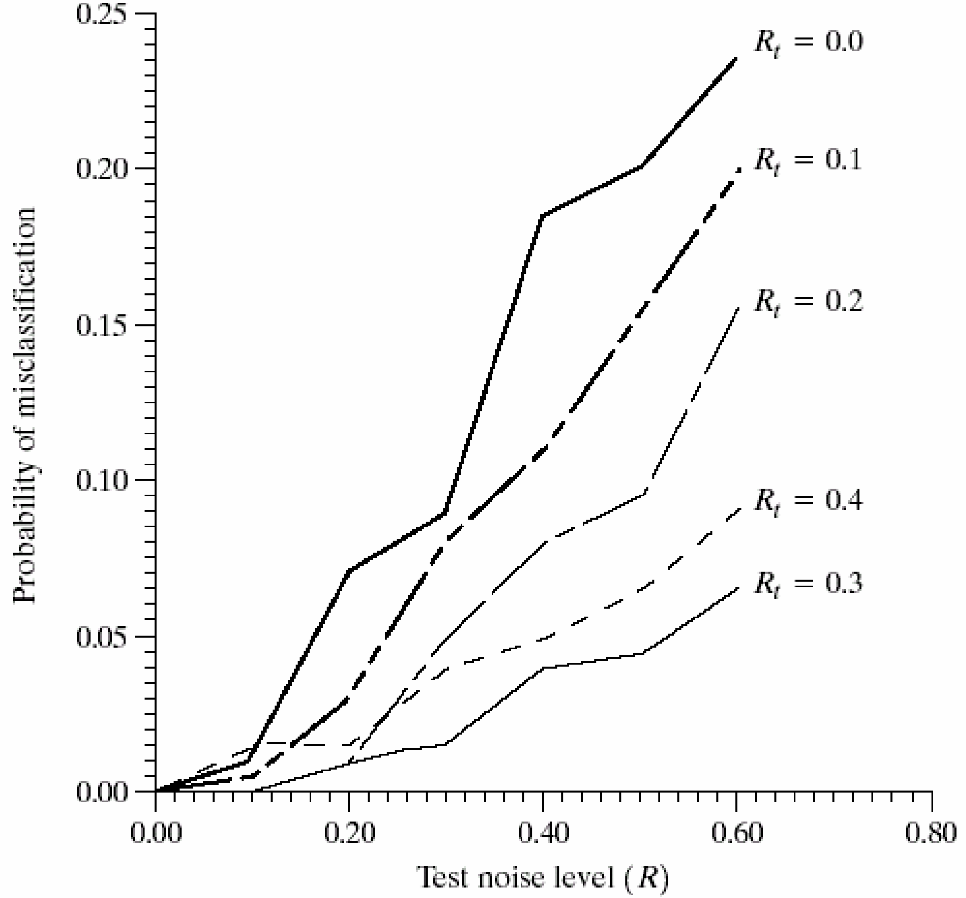


FIGURE 12.19 Three-layer neural network used to recognize the shapes in Fig. 12.18. (Courtesy of Dr. Lalit Gupta, ECE Department, Southern Illinois University.)

Chapter 12
Object Recognition

FIGURE 12.20 Performance of the neural network as a function of noise level. (Courtesy of Dr. Lalit Gupta, ECE Department, Southern Illinois University.)



Chapter 12
Object Recognition

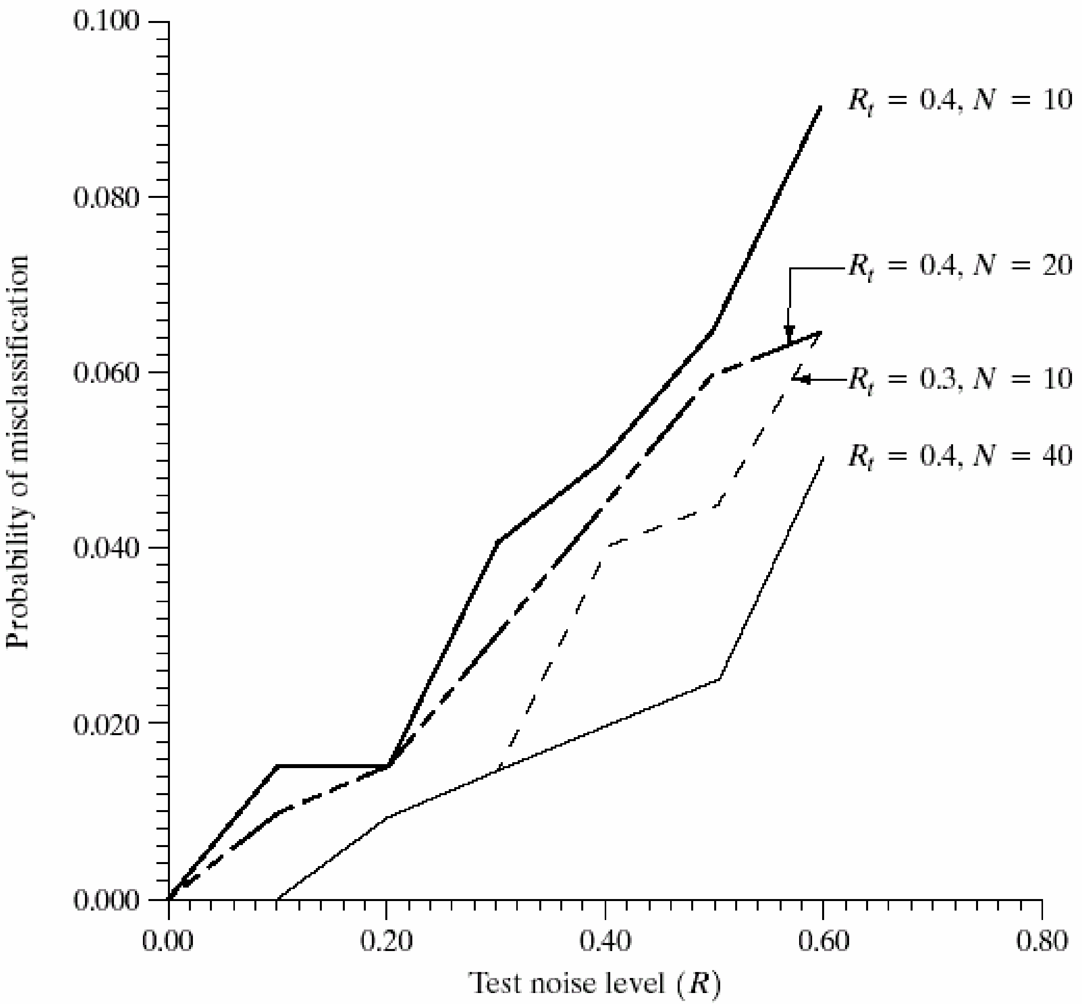


FIGURE 12.21 Improvement in performance for $R_t = 0.4$ by increasing the number of training patterns (the curve for $R_t = 0.3$ is shown for reference). (Courtesy of Dr. Lalit Gupta, ECE Department, Southern Illinois University.)

Chapter 12
Object Recognition

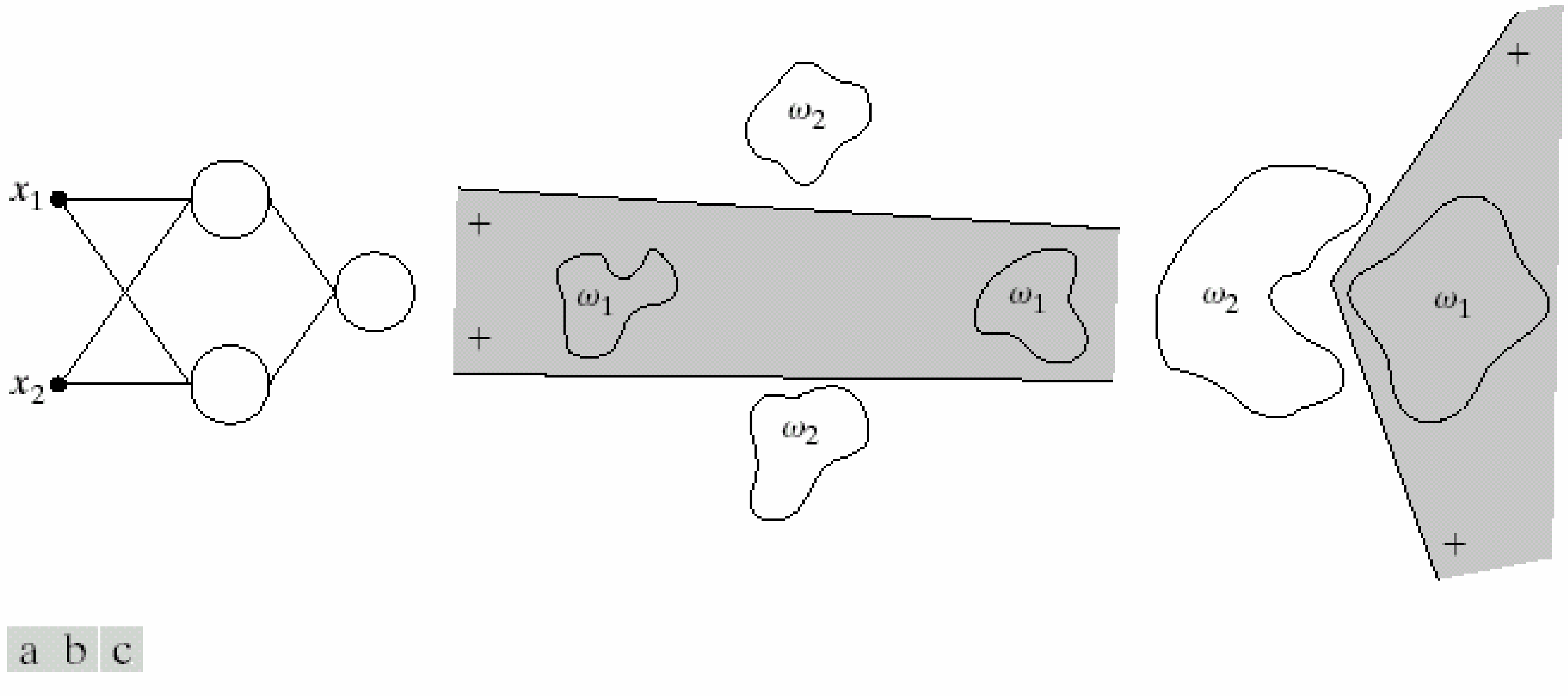
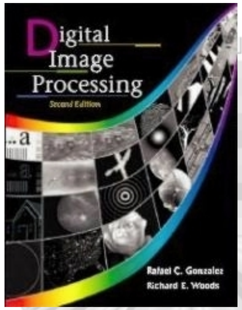


FIGURE 12.22 (a) A two-input, two-layer, feedforward neural network. (b) and (c) Examples of decision boundaries that can be implemented with this network.



Chapter 12
Object Recognition

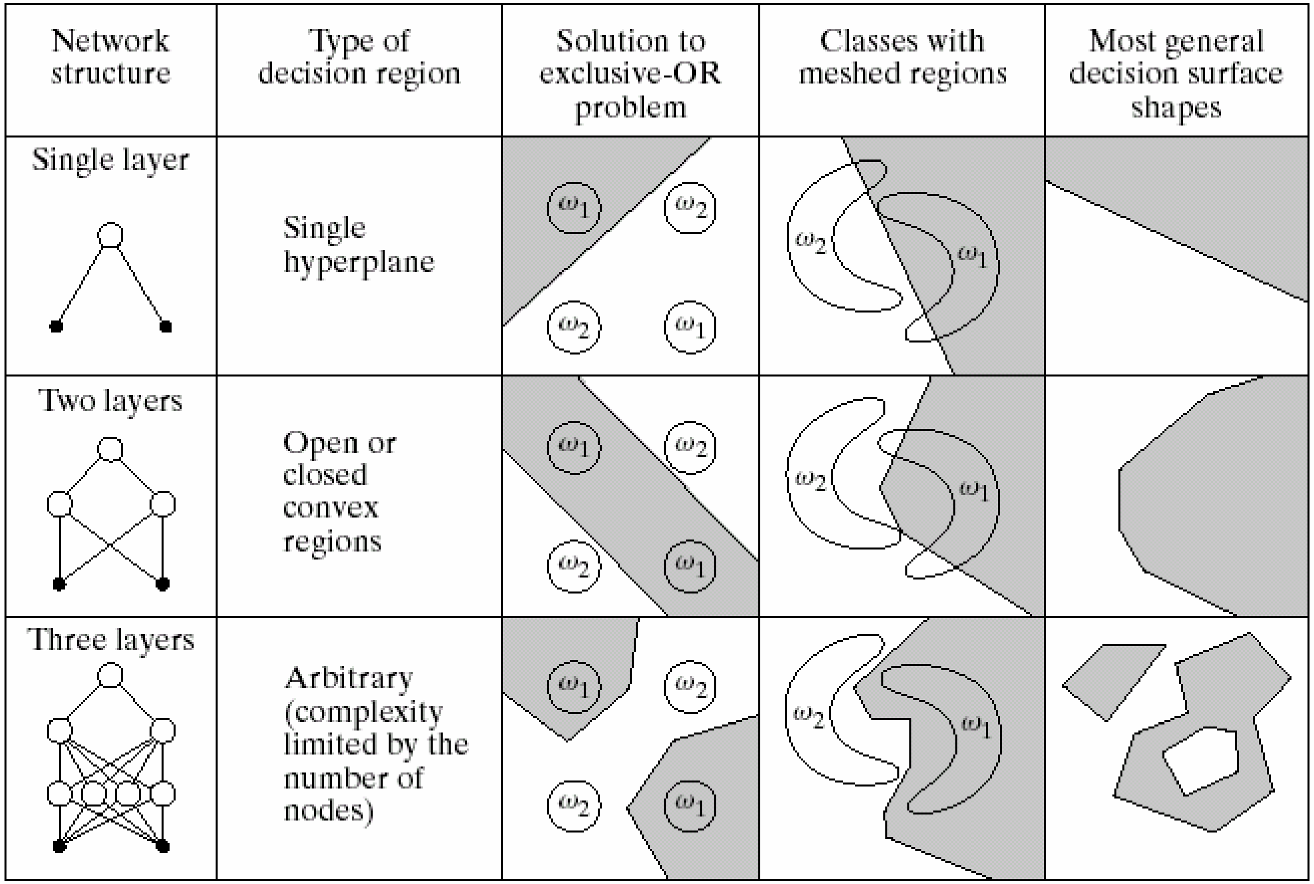
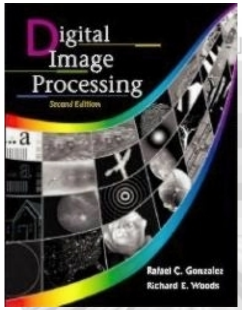
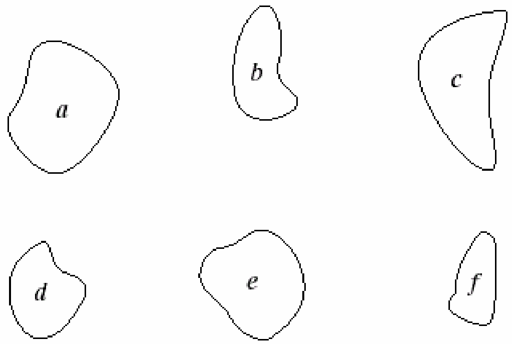


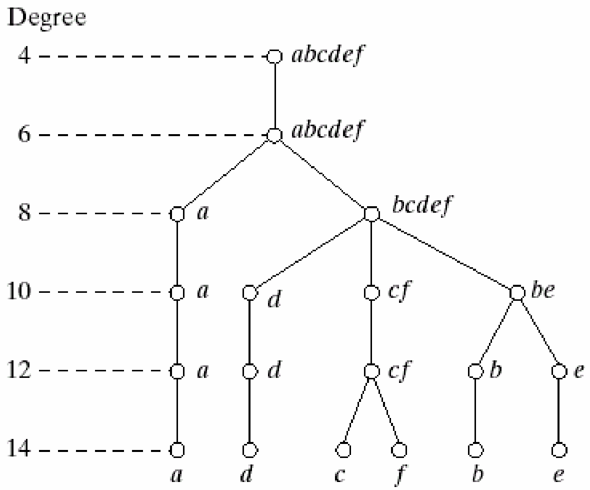
FIGURE 12.23 Types of decision regions that can be formed by single- and multilayer feed-forward networks with one and two layers of hidden units and two inputs. (Lippman)

Structural Method



Chapter 12
Object Recognition



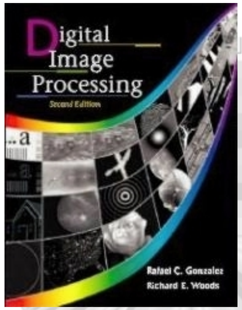


	a	b	c	d	e	f
a	∞	6	6	6	6	6
b		∞	8	8	10	8
c			∞	8	8	12
d				∞	8	8
e					∞	8
f						∞

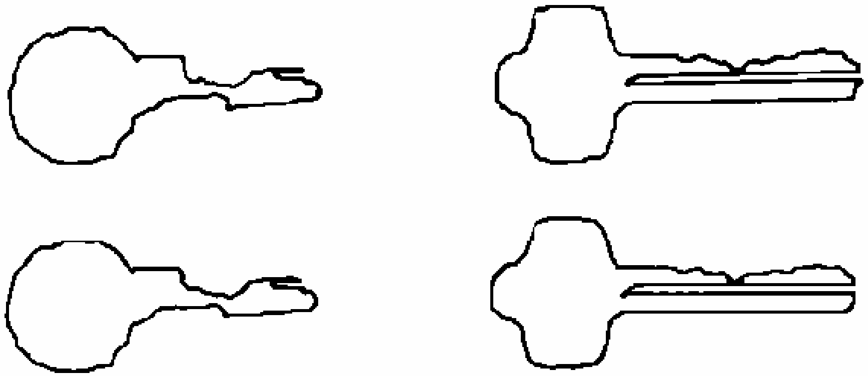
a
b c

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

String Matching



Chapter 12
Object Recognition



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	∞

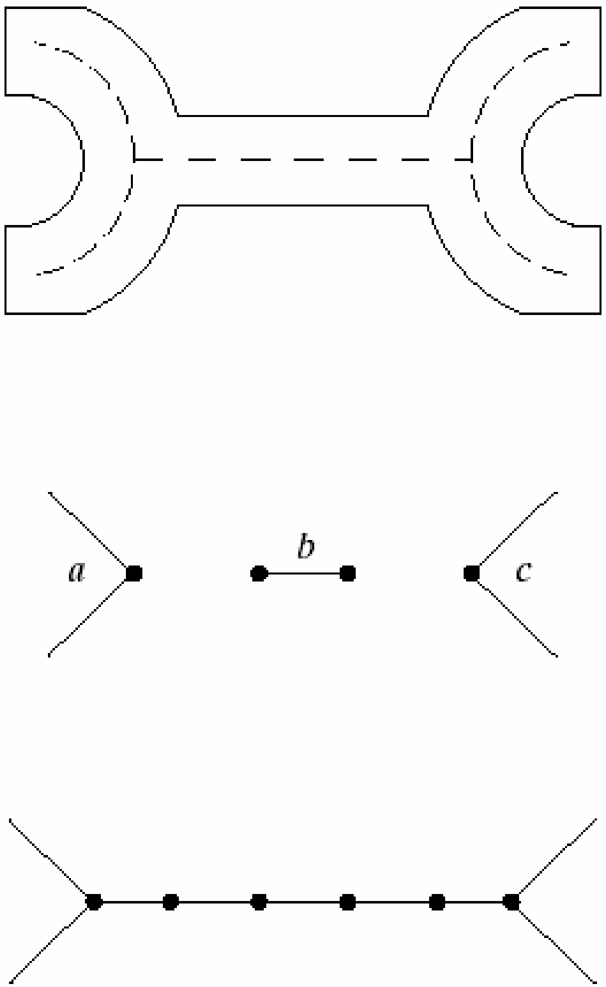
R	2.a	2.b	2.c	2.d	2.e	2.f
2.a	∞					
2.b	33.5	∞				
2.c	4.8	5.8	∞			
2.d	3.6	4.2	19.3	∞		
2.e	2.8	3.3	9.2	18.3	∞	
2.f	2.6	3.0	7.7	13.5	27.0	∞

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.)

Chapter 12
Object Recognition



a
b
c

FIGURE 12.26 (a) Object represented by its (pruned) skeleton. (b) Primitives. (c) Structure generated by using a regular string grammar.

TABLE 12.1 Example of semantic information attached to production rules.

Production	Semantic Information
$S \rightarrow aA$	Connections to a are made only at the dot. The direction of a, denoted θ, is given by the direction of the perpendicular bisector of the line joining the end points of the two undotted segments. The line segments are 3 cm each.
$A \rightarrow bA$	Connections to b are made only at the dots. No multiple connections are allowed. The direction of b must be the same as the direction of a. The length of b is 0.25 cm. This production cannot be applied more than 10 times.
$A \rightarrow bB$	The direction of a and b must be the same. Connections must be simple and made only at the dots.
$B \rightarrow c$	The direction of c and a must be the same. Connections must be simple and made only at the dots.

Automatic as string recognizers

Chapter 12
Object Recognition

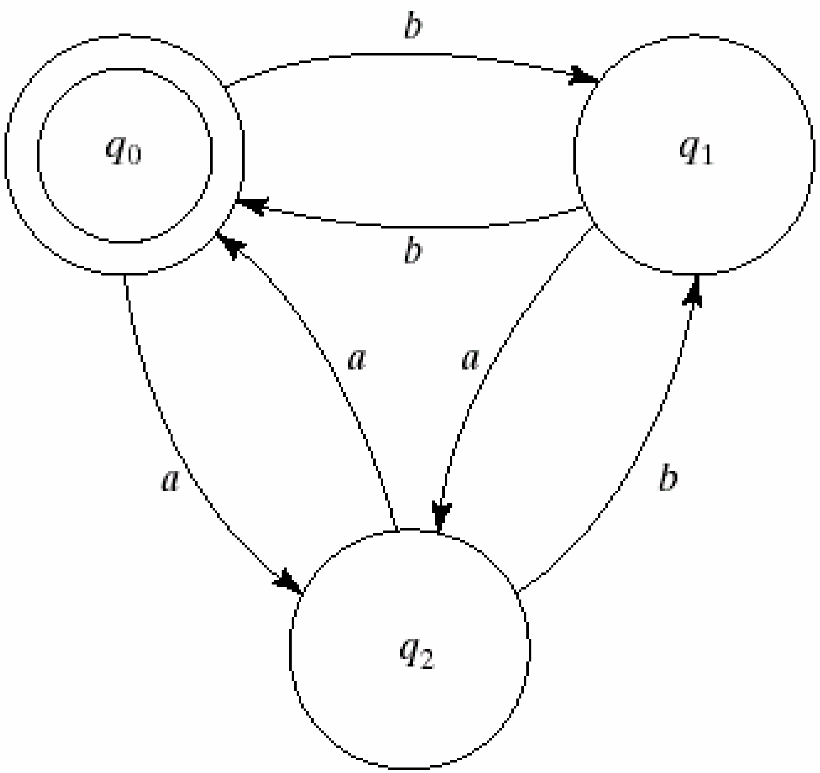
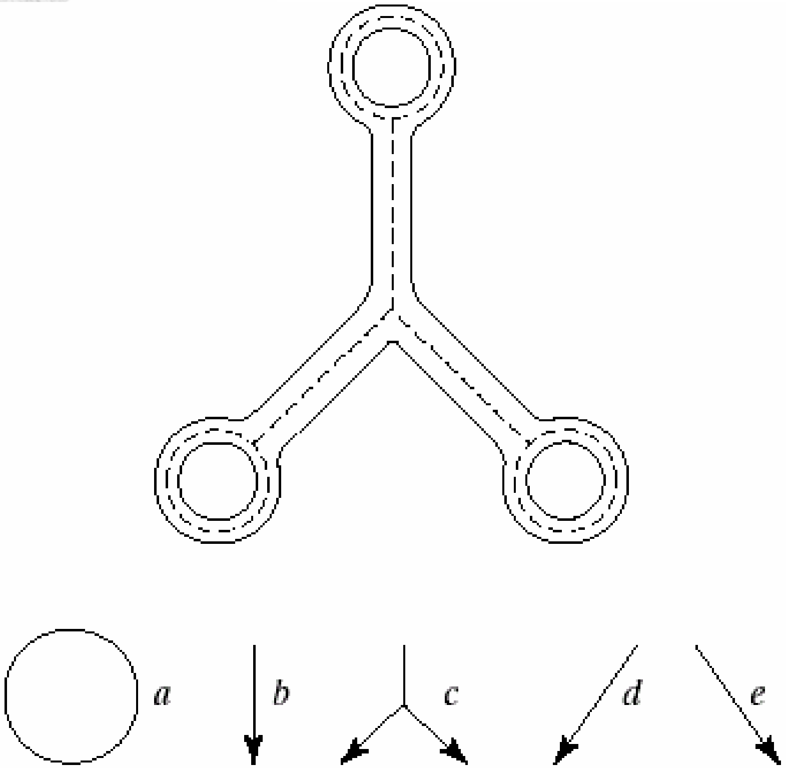


FIGURE 12.27 A finite automaton.

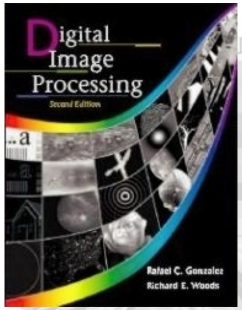
Syntactic Recognition of Trees

Chapter 12
Object Recognition

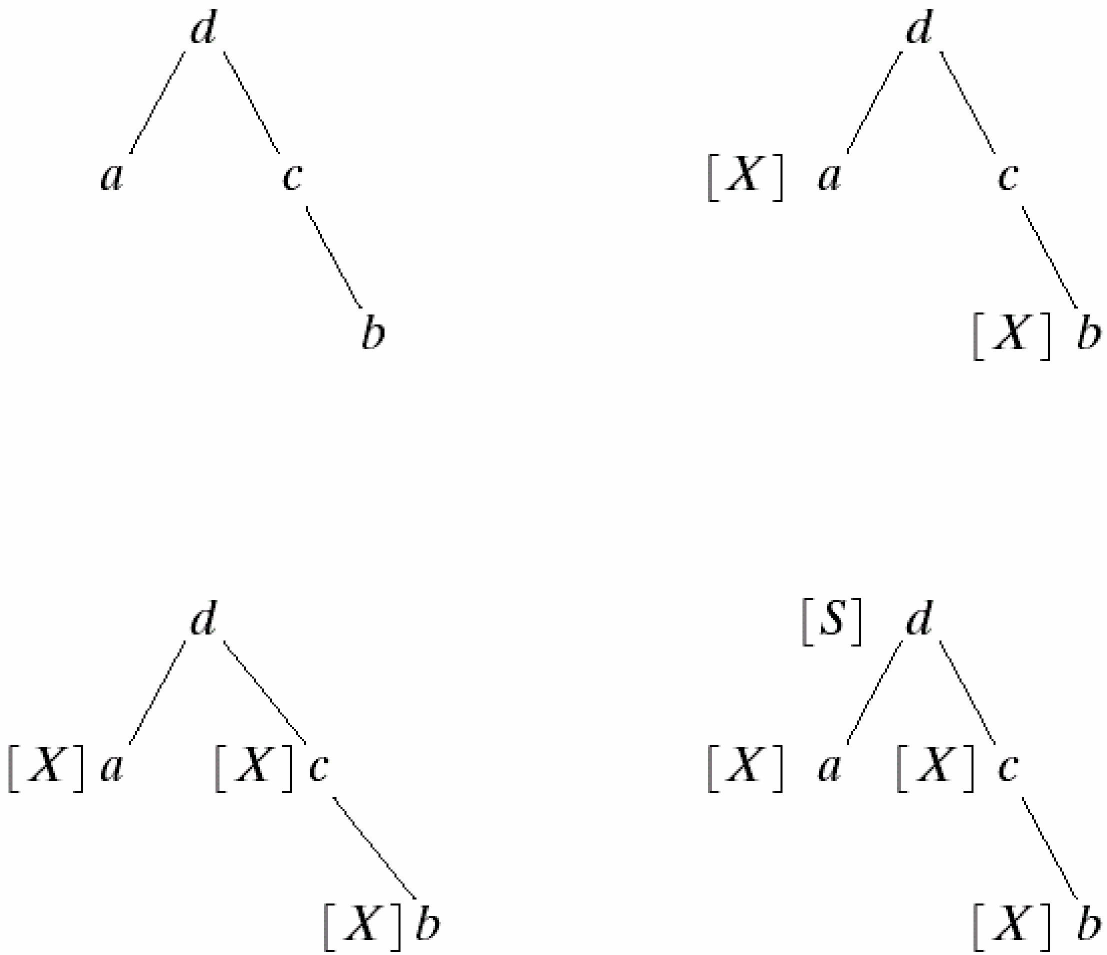


a
b

FIGURE 12.28 (a) An object and (b) primitives used for representing the skeleton by means of a tree grammar.



Chapter 12
Object Recognition

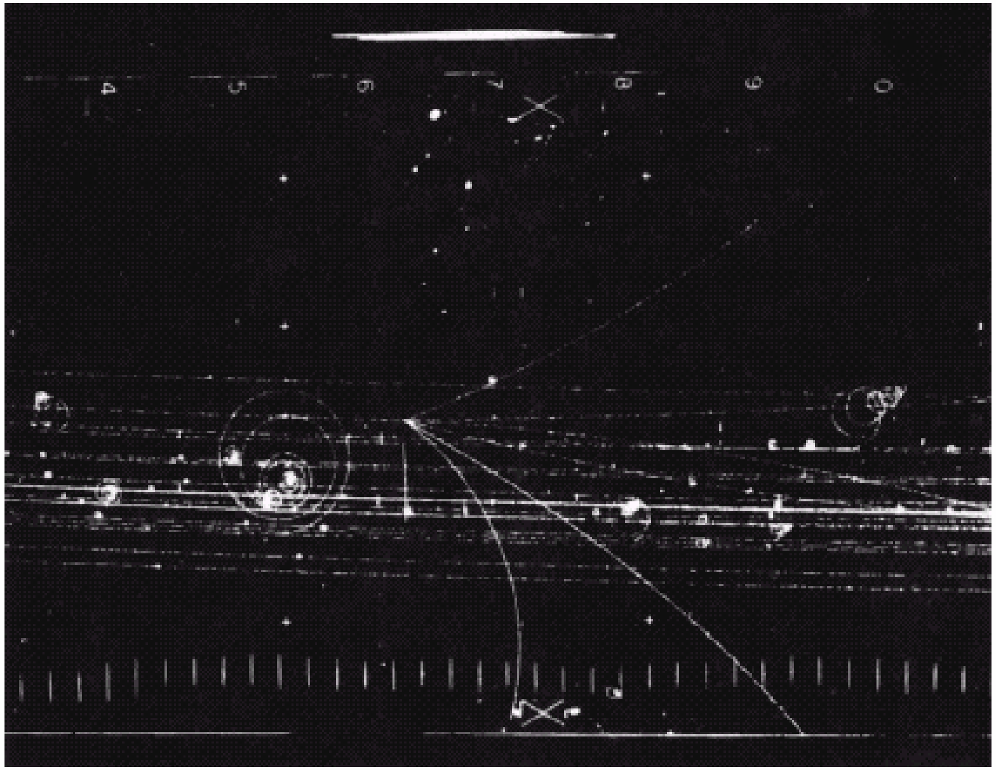


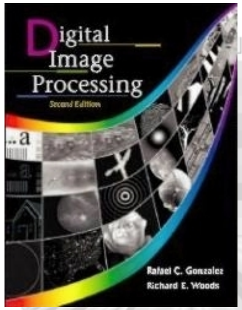
a b
c d

FIGURE 12.29 Processing stages of a frontier-to-root tree automaton: (a) Input tree. (b) State assignment to frontier nodes. (c) State assignment to intermediate nodes. (d) State assignment to root node.

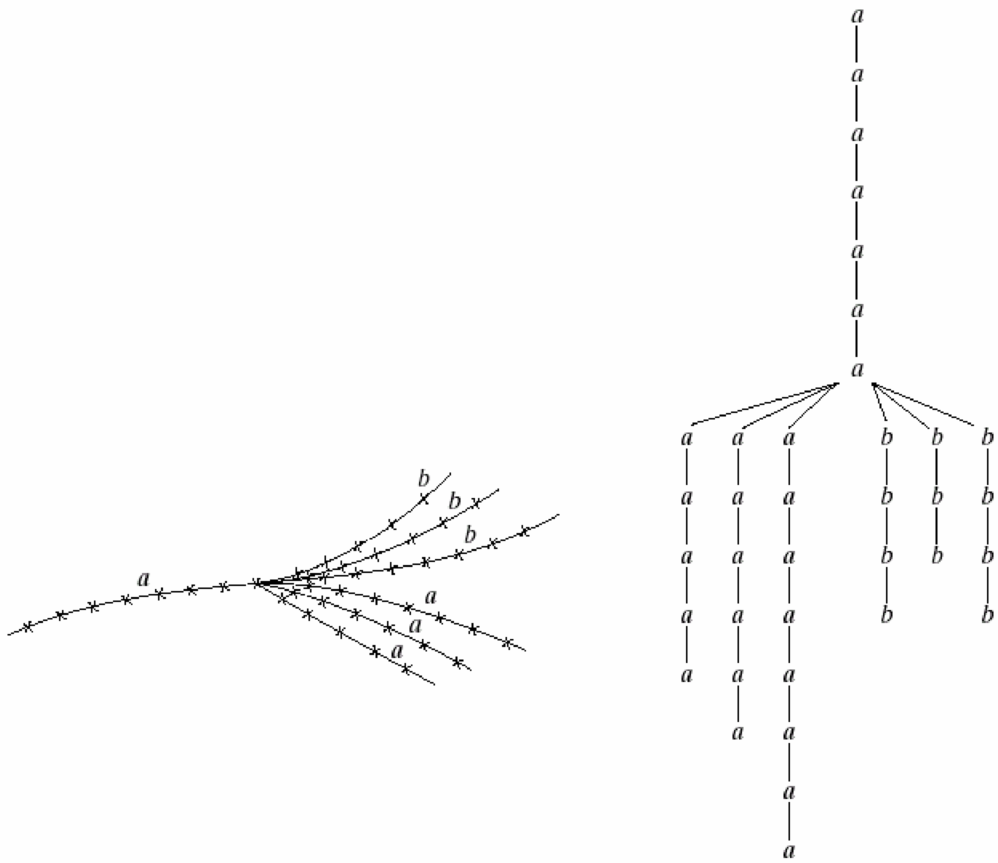
Chapter 12
Object Recognition

FIGURE 12.30 A bubble chamber photograph. (Fu and Bhargava.)





Chapter 12
Object Recognition



a b

FIGURE 12.31 (a) Coded event from Fig. 12.30. (b) Corresponding tree representation. (Fu and Bhargava.)

Chapter 12
Object Recognition

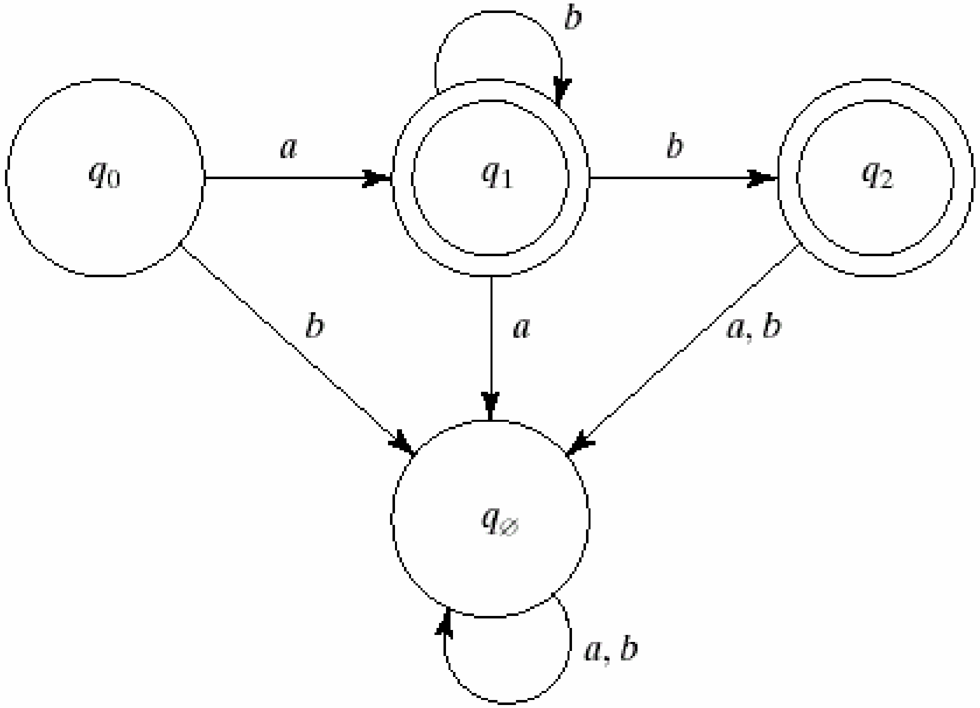
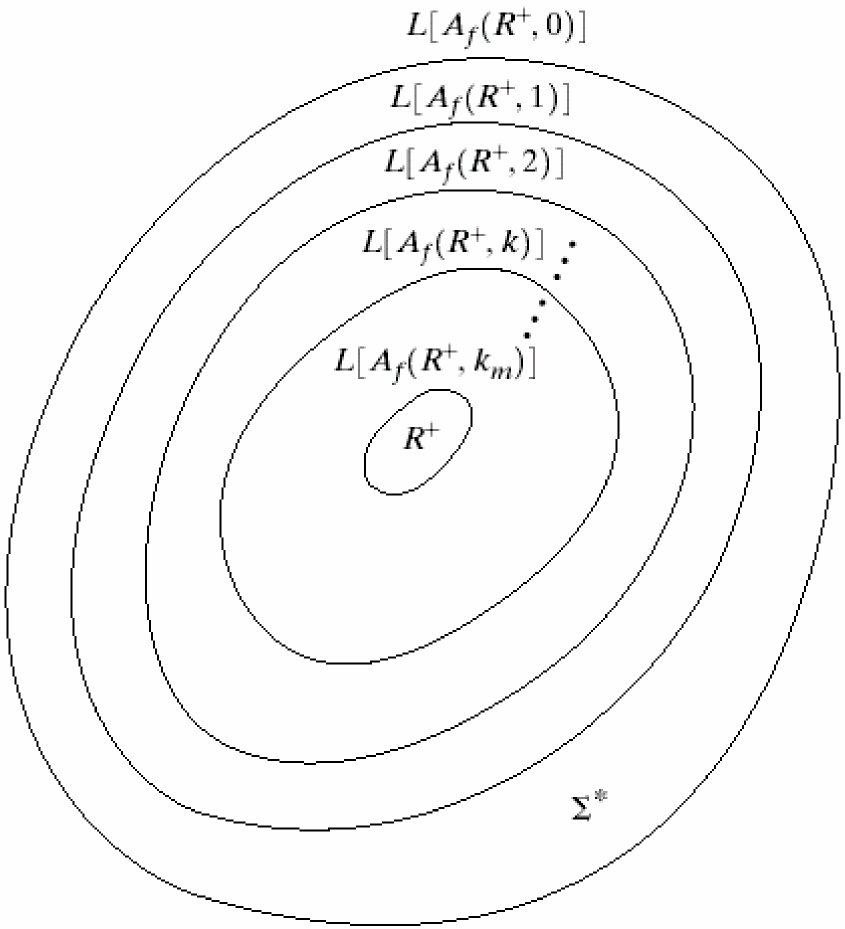


FIGURE 12.32 State diagram for the finite automaton inferred from the sample set $R^+ = \{a, ab, abb\}$.

Chapter 12
Object Recognition

FIGURE 12.33 Relationship between $L[A_f(R^+, k)]$ and k. The value of k_m is such that $k_m \geq$ (length of the longest string in R^+).



Chapter 12
Object Recognition

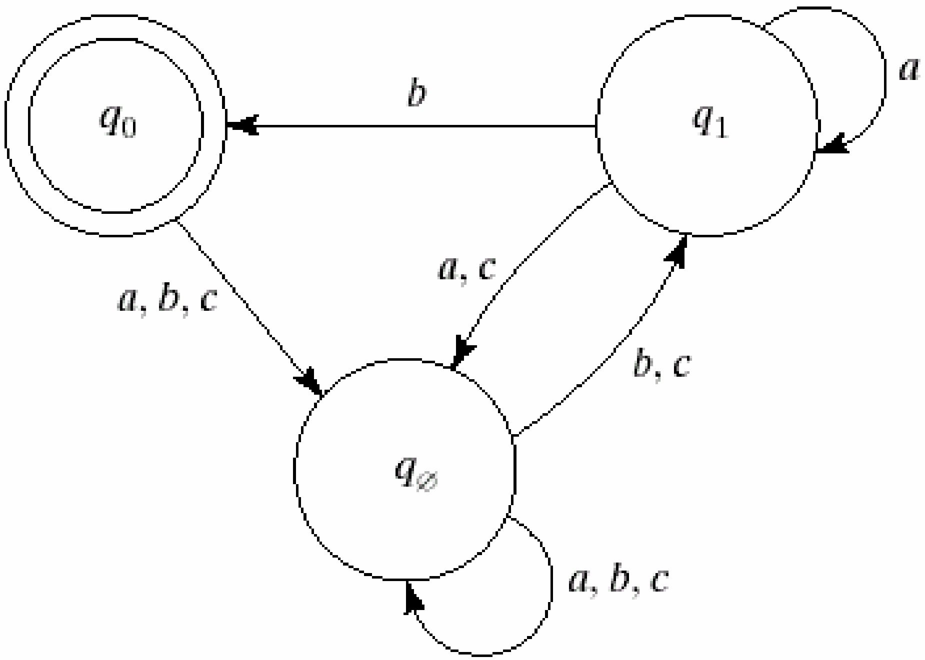


FIGURE 12.34 State diagram for the automaton $A_f(R^+, 1)$ inferred from the sample set $R^+ = \{caaab, bbaab, caab, bbab, cab, bbb, cb\}$.