Chapter 7: Introduction to Object (Pattern) Recognition

7. Pattern Recognition

- A **pattern** (**object**) is an entity, vaguely defined, that could be given a name, e.g.,
  - fingerprint image,
  - handwritten word,
  - human face,
  - speech signal,
  - DNA sequence,

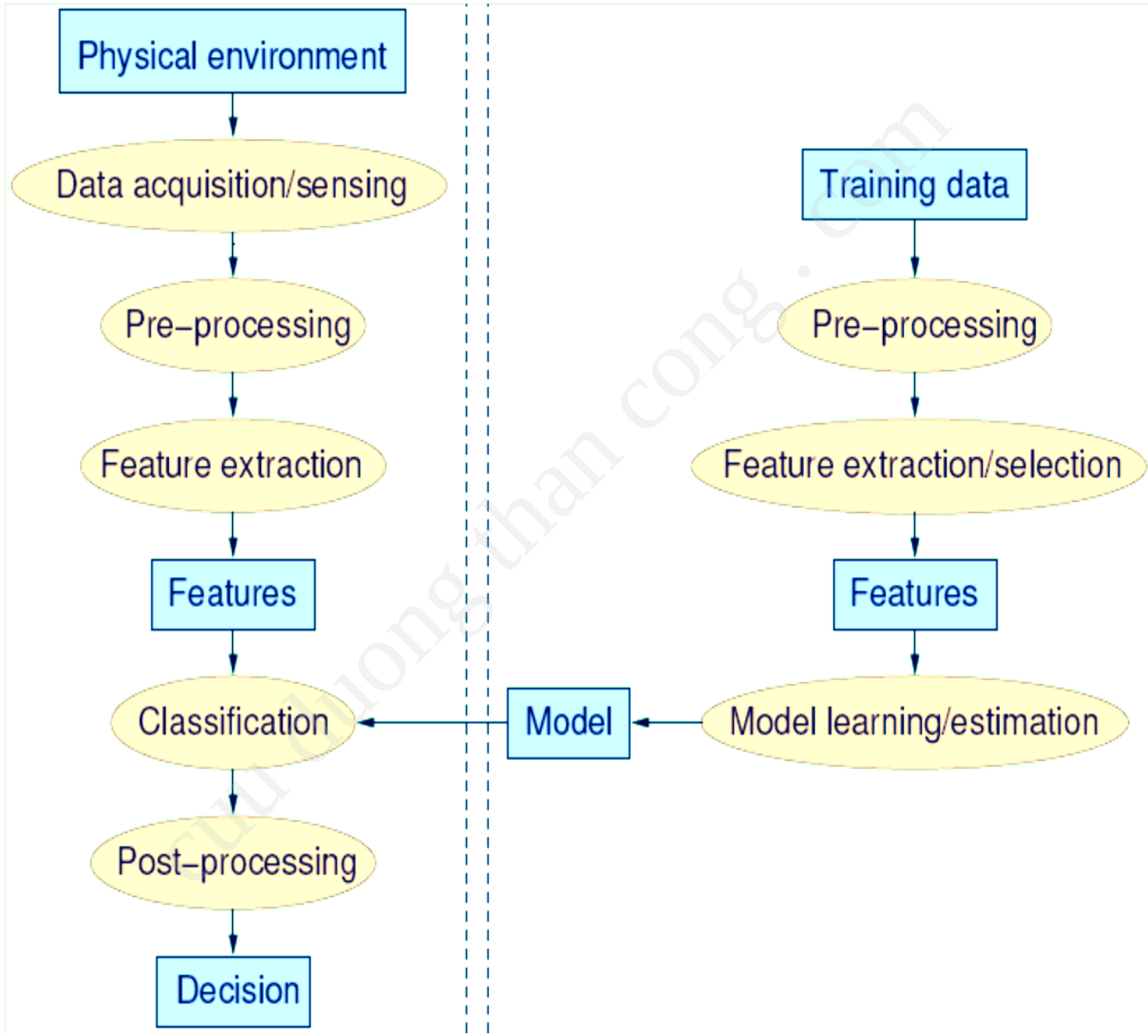
    . . .
- **Pattern recognition** is the study of how machines can
  - observe the environment,
  - learn to distinguish patterns of interest,
  - make sound and reasonable decisions about the categories of the patterns.

7. Pattern Recognition Applications

Example pattern recognition applications:

Problem Domain	Application	Input Pattern	Pattern Classes
Document image analysis	Optical character recognition	Document image	Characters, words
Document classification	Internet search	Text document	Semantic categories
Document classification	Junk mail filtering	Email	Junk/non-junk
Multimedia database retrieval	Internet search	Video clip	Video genres
Speech recognition	Telephone directory assistance	Speech waveform	Spoken words
Natural language processing	Information extraction	Sentences	Parts of speech
Biometric recognition	Personal identification	Face, iris, fingerprint	Authorized users for access control
Medical	Computer aided diagnosis	Microscopic image	Cancerous/healthy cell
Military	Automatic target recognition	Optical or infrared image	Target type
Industrial automation	Printed circuit board inspection	Intensity or range image	Defective/non-defective product
Industrial automation	Fruit sorting	Images taken on a conveyor belt	Grade of quality
Remote sensing	Forecasting crop yield	Multispectral image	Land use categories
Bioinformatics	Sequence analysis	DNA sequence	Known types of genes
Data mining	Searching for meaningful patterns	Points in multidimensional space	Compact and well-separated clusters

7. Process of a Pattern Recognition System (1)



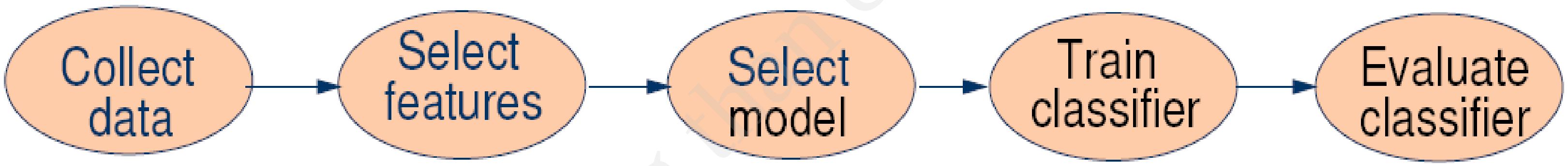
7. Process of a Pattern Recognition System (2)

- Data acquisition and sensing:
  - Measurements of physical variables.
  - Important issues: bandwidth, resolution, sensitivity, distortion, SNR, latency, etc.
- Pre-processing:
  - Removal of noise in data.
  - Isolation of patterns of interest from the background.
- Feature extraction:
  - Finding a new representation in terms of features.
- Model learning and estimation:
  - Learning a mapping between features and pattern groups and categories.

7. Process of a Pattern Recognition System (3)

- Classification:
  - Using features and learned models to assign a pattern to a category.
- Post-processing:
  - Evaluation of confidence in decisions.
  - Exploitation of context to improve performance.
  - Combination of experts.

7. Design Cycle (1)



7. Design Cycle (2)

- Data collection:
  - Collecting training and testing data.
  - How can we know when we have adequately large and representative set of samples?
- Feature selection:
  - Domain dependence and prior information.
  - Computational cost and feasibility.
  - Discriminative features.
    - Similar values for similar patterns.
    - Different values for different patterns.
  - Invariant features with respect to translation, rotation and scale.
  - Robust features with respect to occlusion, distortion, deformation, and variations in environment.

7. Design Cycle (3)

- Model selection:
  - Domain dependence and prior information.
  - Definition of design criteria.
  - Parametric vs. non-parametric models.
  - Handling of missing features.
  - Computational complexity.
  - Types of models: templates, decision-theoretic or statistical, syntactic or structural, neural, and hybrid.
  - How can we know how close we are to the true model underlying the patterns?

7. Design Cycle (4)

- Training:
  - How can we learn the rule from data?
  - Supervised learning: a teacher provides a category label or cost for each pattern in the training set.
  - Unsupervised learning: the system forms clusters or natural groupings of the input patterns.
  - Reinforcement learning: no desired category is given but the teacher provides feedback to the system such as the decision is right or wrong.
- Evaluation:
  - How can we estimate the performance with training samples?
  - How can we predict the performance with future data?
  - Problems of overfitting and generalization.

7. Pattern Recognition Techniques (1)

- **Pattern** is an **arrangement of descriptors (features)**.
- **Pattern class** is a family of patterns that share some common properties.
- The approaches to pattern recognition developed are divided into two principal areas: **decision-theoretic** and **structural**
  - The first category deals with patterns described using quantitative descriptors, such as length, area, and texture.
  - The second category deals with patterns best described by qualitative descriptors, such as the relational descriptors.

7. Pattern Recognition Techniques (2)

Example: 3 pattern classes (for 3 types of iris flowers) arranged in vectors with 2 descriptors (2 measurements: width and length of their petals) for each pattern class.

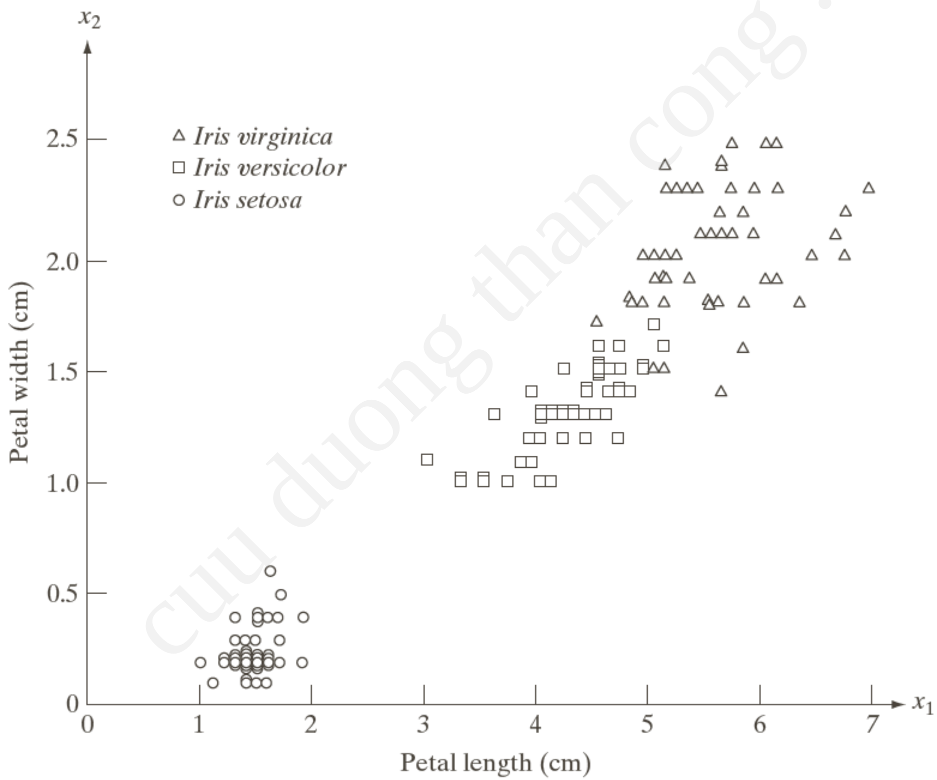


FIGURE 12.1 Three types of iris flowers described by two measurements.

7. Recognition Based on Decision-Theoretic Methods (1)

- Let $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T$ for W pattern classes $\omega_1, \ldots, \omega_W$, $d_i(\mathbf{x}) > d_j(\mathbf{x}), j = 1, \ldots, W, i \neq j$.
- In other words, an unknown pattern **x** is said to belong to the ith pattern class if, upon substitution of **x** into all decision functions, $d_i(\mathbf{x})$ yields the largest numerical value.
- Suppose that we define the prototype of each pattern class to be the mean vector of the patterns of that class:

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

- Using the Euclidean distance to determine closeness reduces the problem to computing the distance measures:

$$D_j(\mathbf{x}) = \left\| \mathbf{x} - \mathbf{m}_j \right\|$$

7. Recognition Based on Decision-Theoretic Methods (2)

Matching with Minimum Distance Classifier

- The smallest distance is equivalent to evaluating the functions:

$$d_j(\mathbf{x}) = \mathbf{x}^T \mathbf{m}_j - \frac{1}{2}\mathbf{m}_j^T \mathbf{m}_j$$

- The decision boundary between classes and for a **minimum distance classifier** is

$$d_{ij}(\mathbf{x}) = d_i(\mathbf{x}) - d_j(\mathbf{x})$$

$$= \mathbf{x}^T(\mathbf{m}_i - \mathbf{m}_j) - \frac{1}{2}(\mathbf{m}_i - \mathbf{m}_j)^T(\mathbf{m}_i + \mathbf{m}_j) = 0$$

7. Recognition Based on Decision-Theoretic Methods (3)

- Decision boundary of minimum distance classifier:

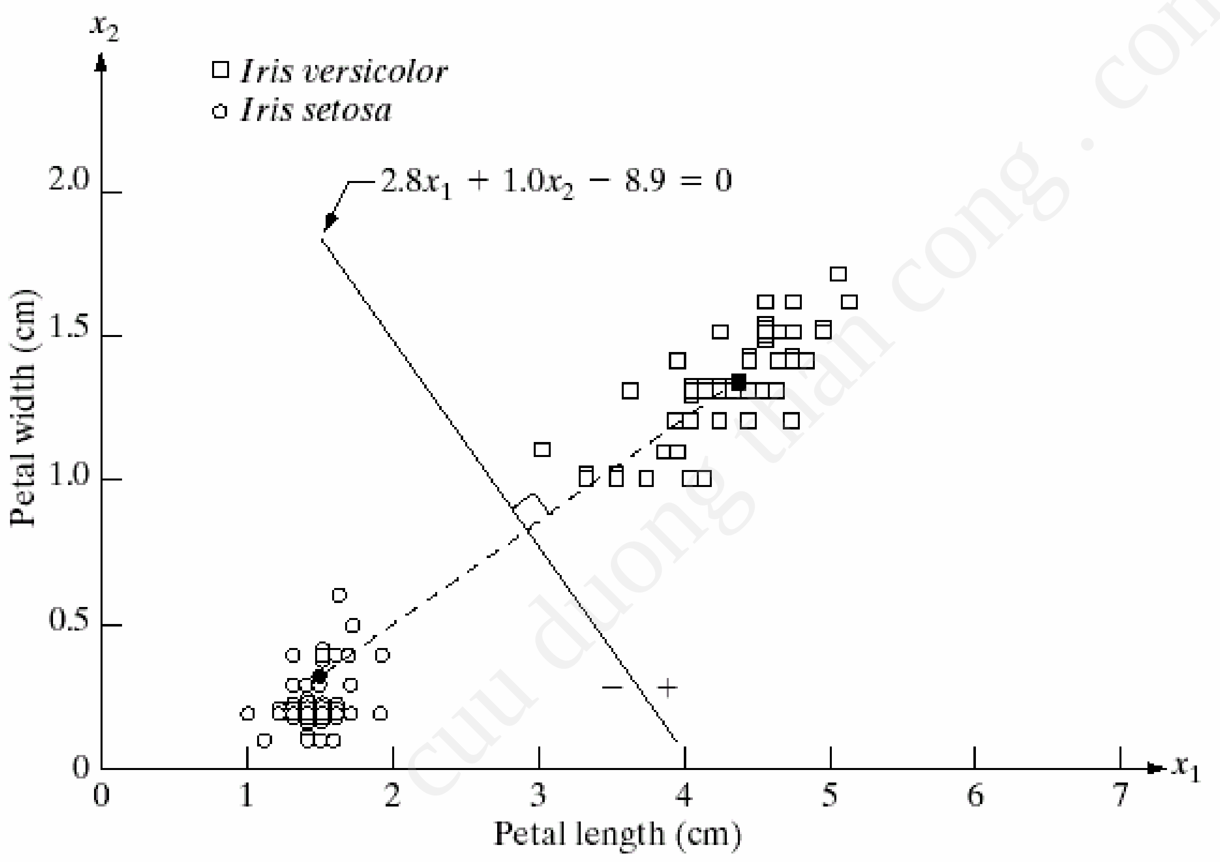


FIGURE 12.6 Decision boundary of minimum distance classifier for the classes of ***Iris versicolor*** and ***Iris setosa***. The dark dot and square are the means.

7. Recognition Based on Decision-Theoretic Methods (4)

❑ Matching by correlation

We consider it as the basis for finding matches of a sub-image of size $J \times K$ within $f(x,y)$ an image of $M \times N$ size , where we assume that $J \leq M$ and $K \leq N$

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

for x = 0,1, 2,…, M-1, y = 0,1, 2,…, N-1

7. Recognition Based on Decision-Theoretic Methods (5)

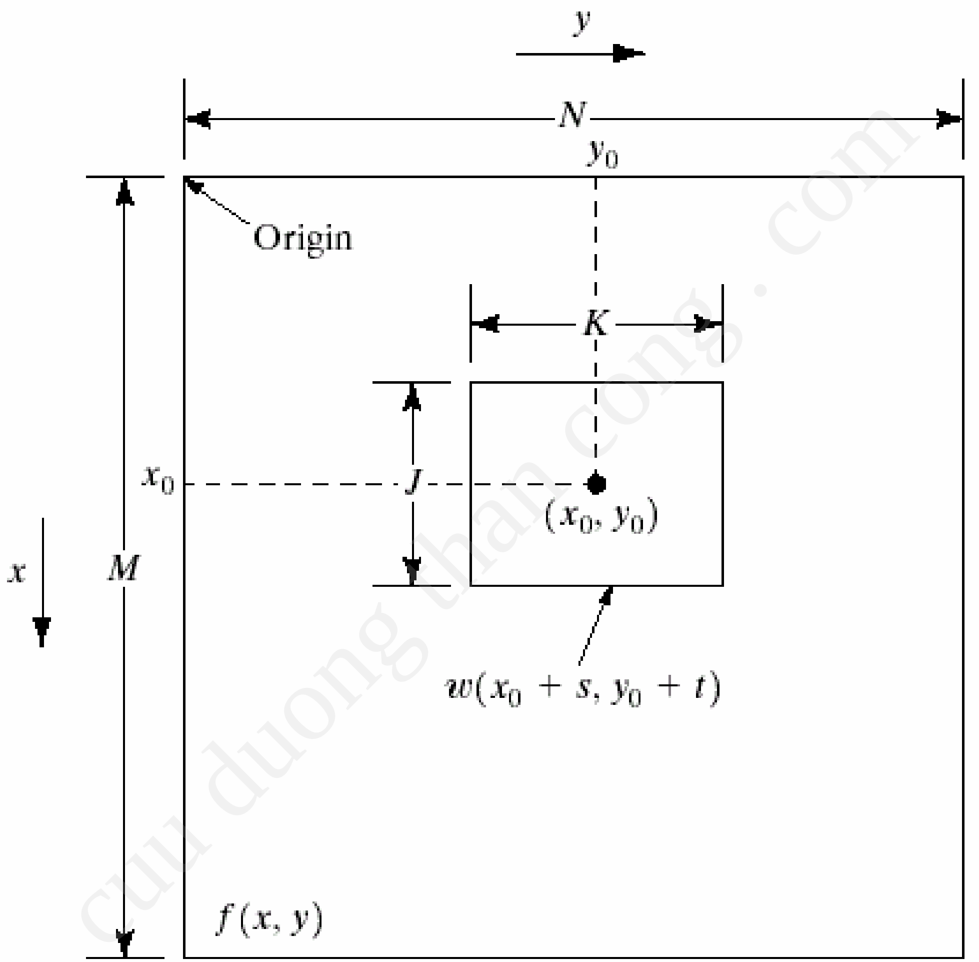


FIGURE 12.8 Arrangement for obtaining the correlation of f and w at point (x_0, y_0).

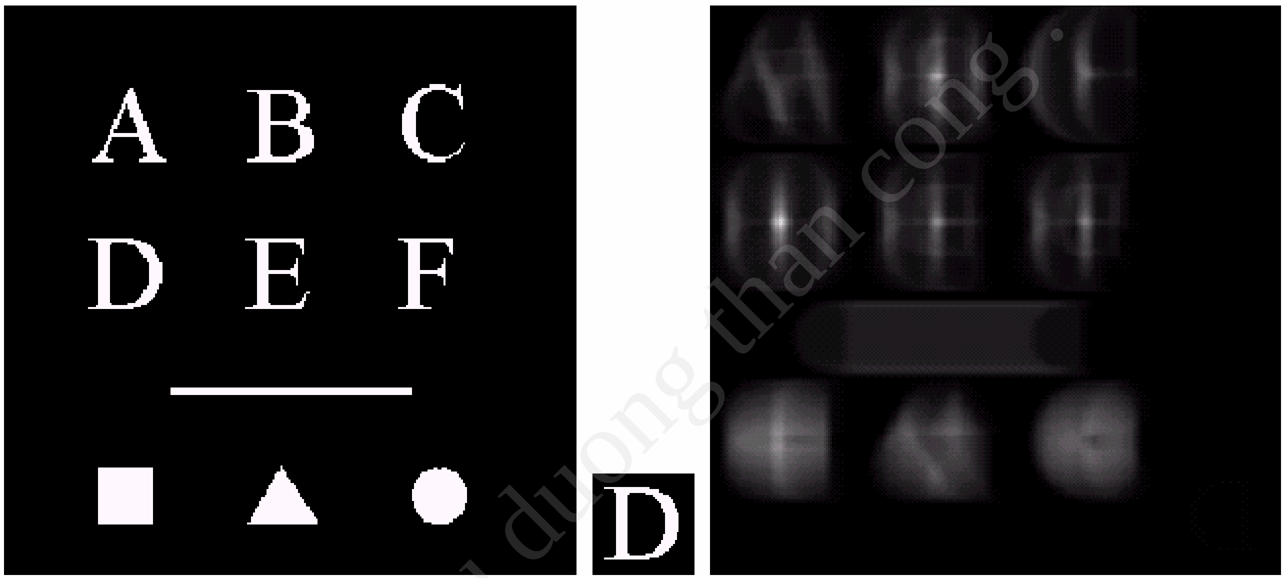
7. Recognition Based on Decision-Theoretic Methods (6)

- The correlation function has the disadvantage of being sensitive to changes in the amplitude of f and w.
- For example, doubling all values of f doubles the value of $c(x, y)$.
- An approach frequently used to overcome this difficulty is to perform matching via the **correlation coefficient**

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

- The correlation coefficient is scaled in the range -1 to 1, independent of scale changes in the amplitude of f and w.

7. Recognition Based on Decision-Theoretic Methods (7)



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).

7. Recognition Based on Decision-Theoretic Methods (8)

❑ Optimum Statistical Classifiers

- The probability that a particular pattern **x** comes from class ω_i is denoted $p(\omega_i/\mathbf{x})$
- If the pattern classifier decides that **x** came from ω_j when it actually came from ω_i, it incurs a **loss**, denoted L_{ij}

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

- From basic probability theory, we know that

$$p(A/B) = \left[p(A)\,p(B/A)\right]/p(B)$$

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

7. Recognition Based on Decision-Theoretic Methods (9)

- Since $1/p(\mathbf{x})$ is positive and common to allthe $r_j(\mathbf{x})$, $j = 1, 2, \ldots, W$, it can be dropped without affecting the relative order of these functions from the smallest to the largest value.

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

- Thus the **Bayes classifier** assigns an unknown pattern $\mathbf{x}$ to class ω_i if

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

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

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

$$= p(\mathbf{x}) - p(\mathbf{x}/\omega_j)\, P(\omega_j)$$

7. Recognition Based on Decision-Theoretic Methods (10)

- The Bayes classifier then assigns a pattern **x** to class ω_i if,

$$p(\mathbf{x}) - p(\mathbf{x}/\omega_i)P(\omega_i) < p(\mathbf{x}) - p(\mathbf{x}/\omega_j)P(\omega_j)$$

or, equivalently, if

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

or in the form of decision functions:

$$d_j(\mathbf{x}) = p(\mathbf{x}/\omega_j)P(\omega_j), \qquad j = 1, 2, ..., W$$

7. Recognition Based on Decision-Theoretic Methods (11)

- **Bayes classifier for Gaussian pattern classes**

  Let us consider a 1-D problem (n = 1) involving two pattern classes (W = 2) governed by Gaussian densities

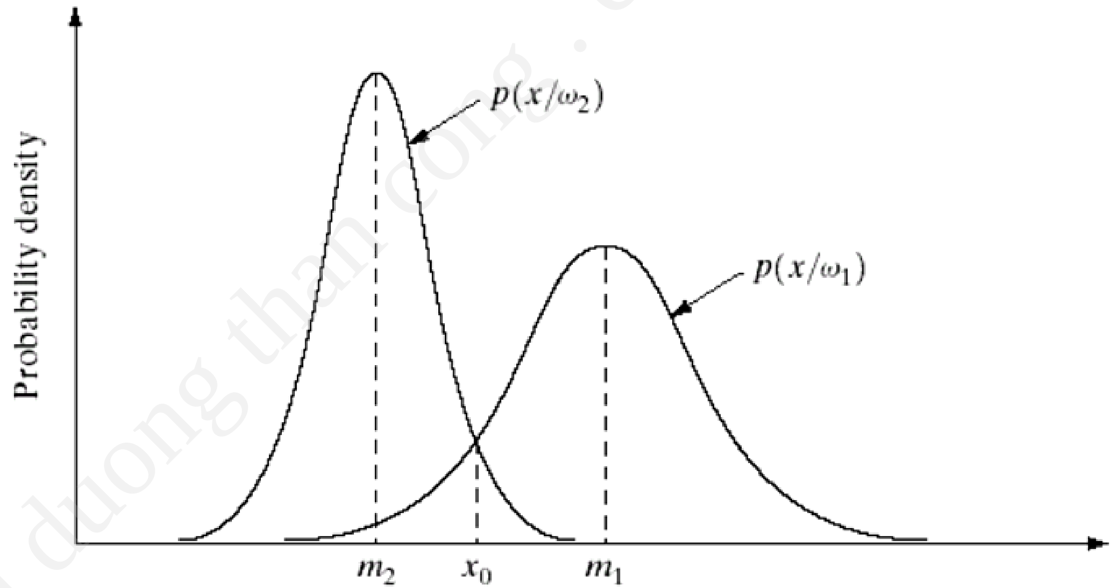
$$d_j(x) = p(x/\omega_j)\,P(\omega_j)$$

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

$$j = 1, 2$$

7. Recognition Based on Decision-Theoretic Methods (12)

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.



7. Recognition Based on Decision-Theoretic Methods (13)

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

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

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

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

Approximating:

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

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

7. Recognition Based on Decision-Theoretic Methods (14)

- Bayes decision function for class ω_j is

$$d_j(\mathbf{x}) = \ln[p(\mathbf{x}/\omega_j)P(\omega_j)]$$

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

7. Recognition Based on Decision-Theoretic Methods (15)

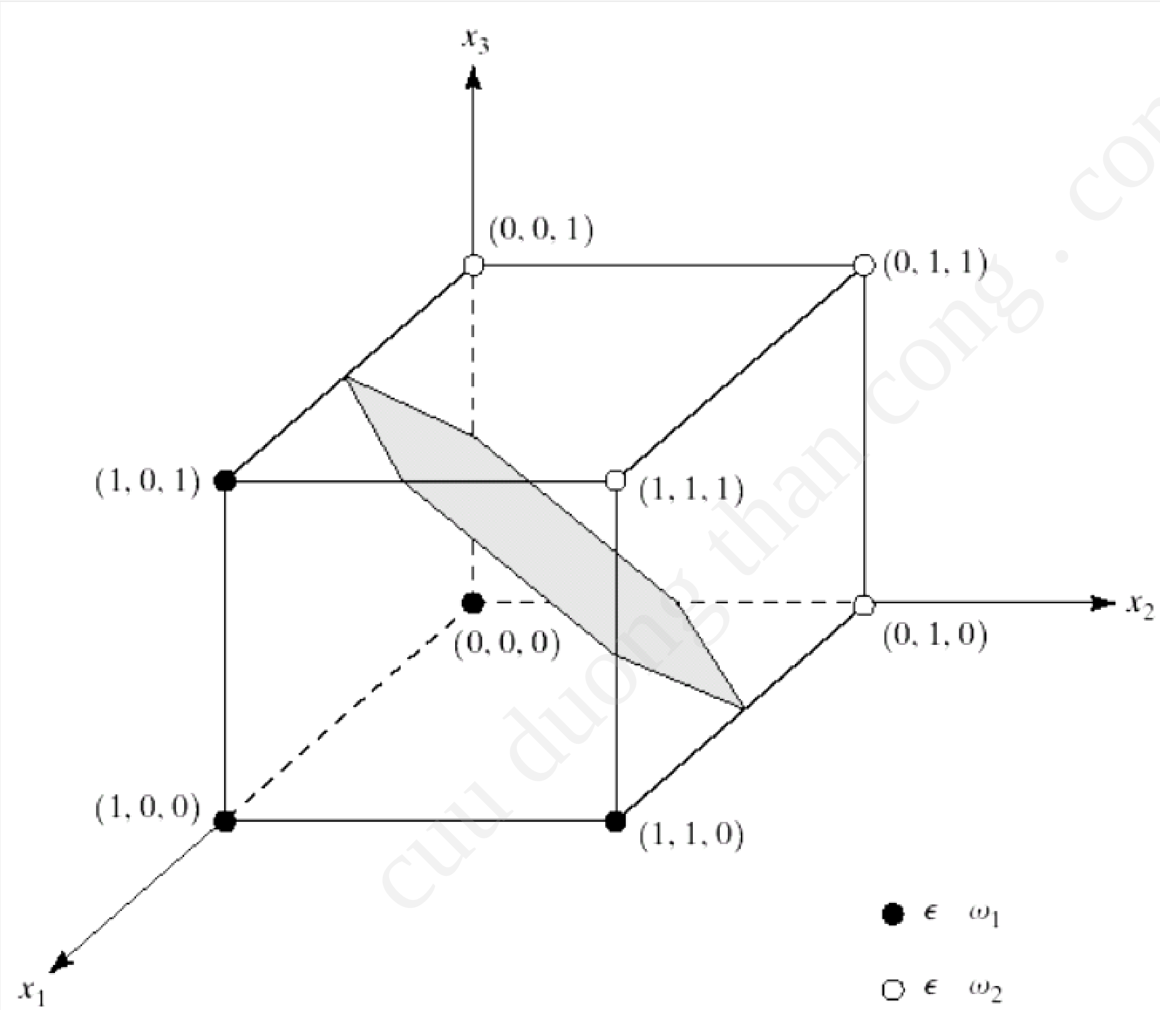


FIGURE 12.11 Two simple pattern classes and their Bayes decision boundary (shown shaded).

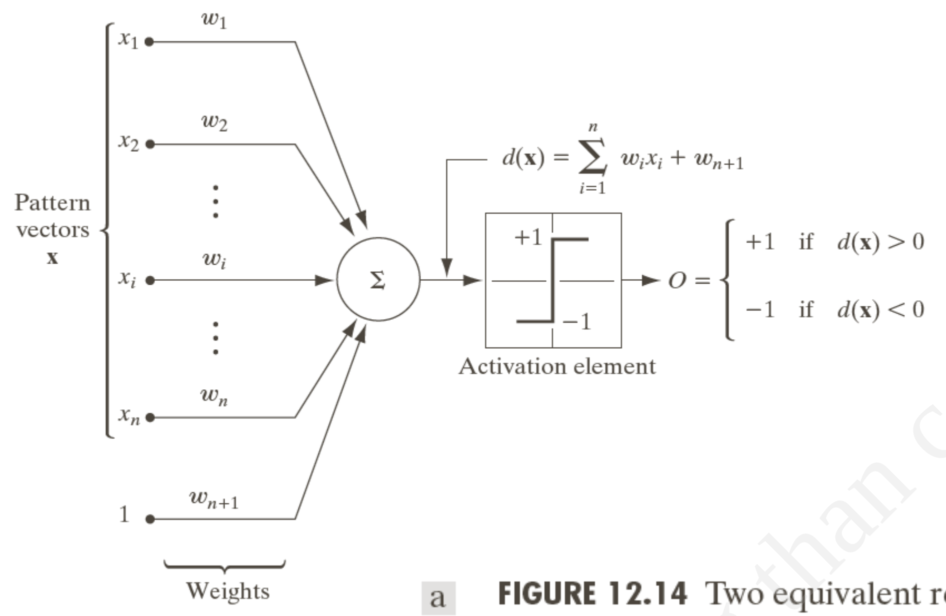
7. Recognition Based on Decision-Theoretic Methods (16)

❑ <u>Neuron Networks</u>

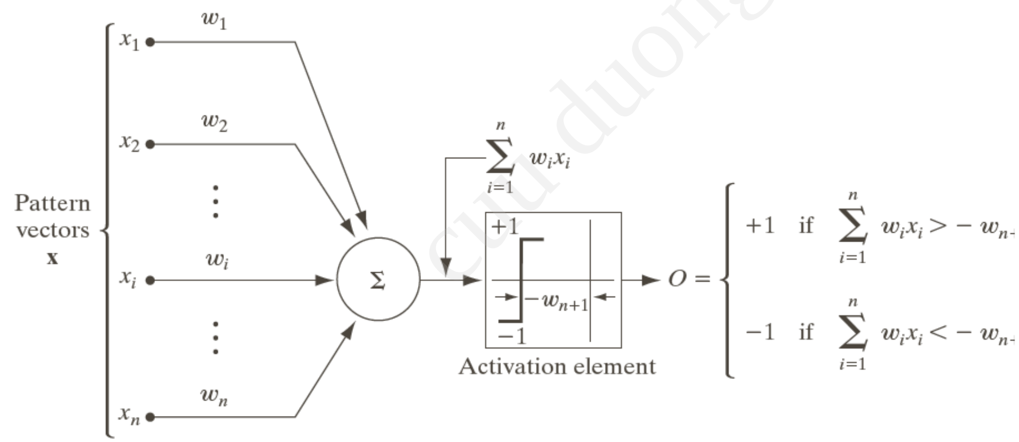
– Ideas stem from the operation of human neural networks.

– Networks of interconnected nonlinear computing elements called neurons.

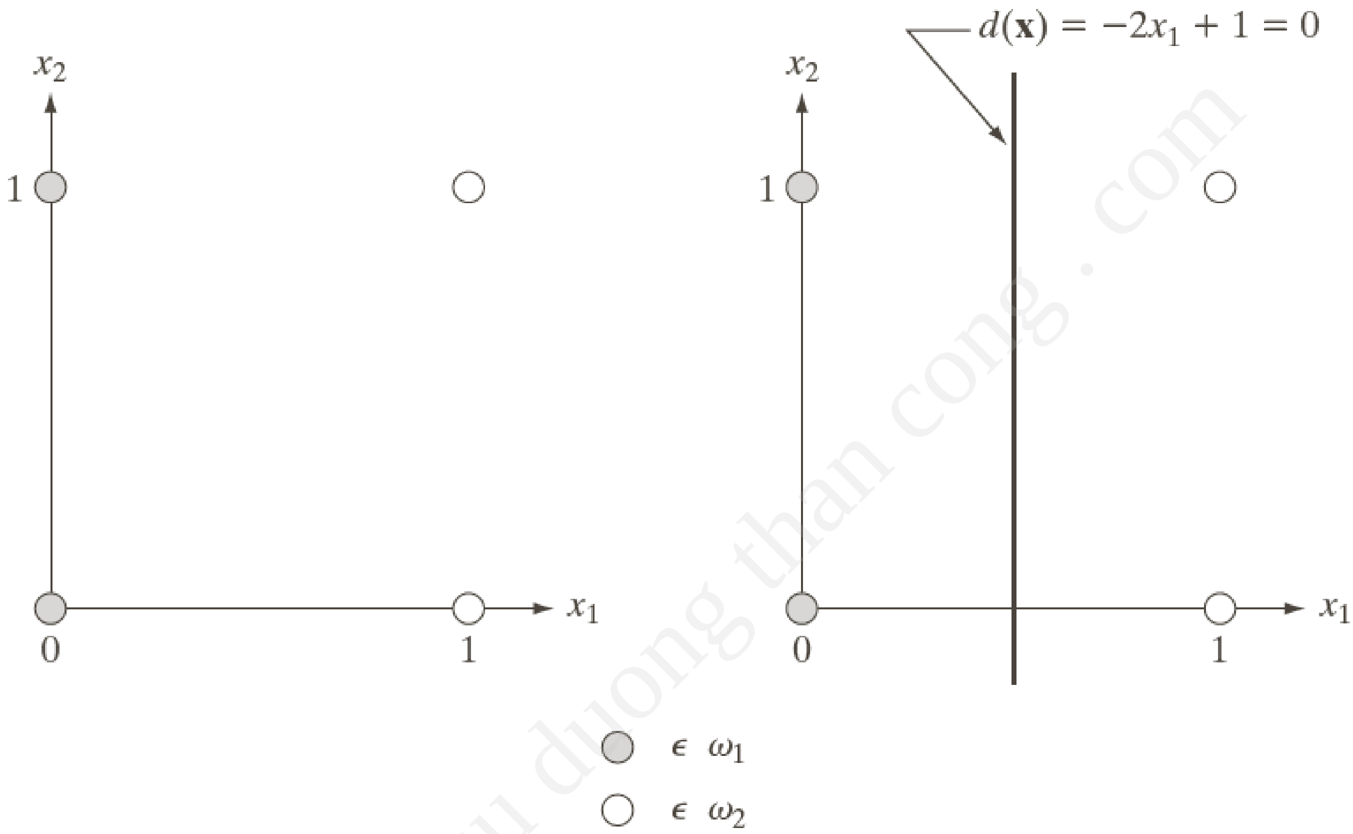
7. Recognition Based on Decision-Theoretic Methods (17)



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



7. Recognition Based on Decision-Theoretic Methods (18)



a b

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

7. Recognition Based on Decision-Theoretic Methods (19)

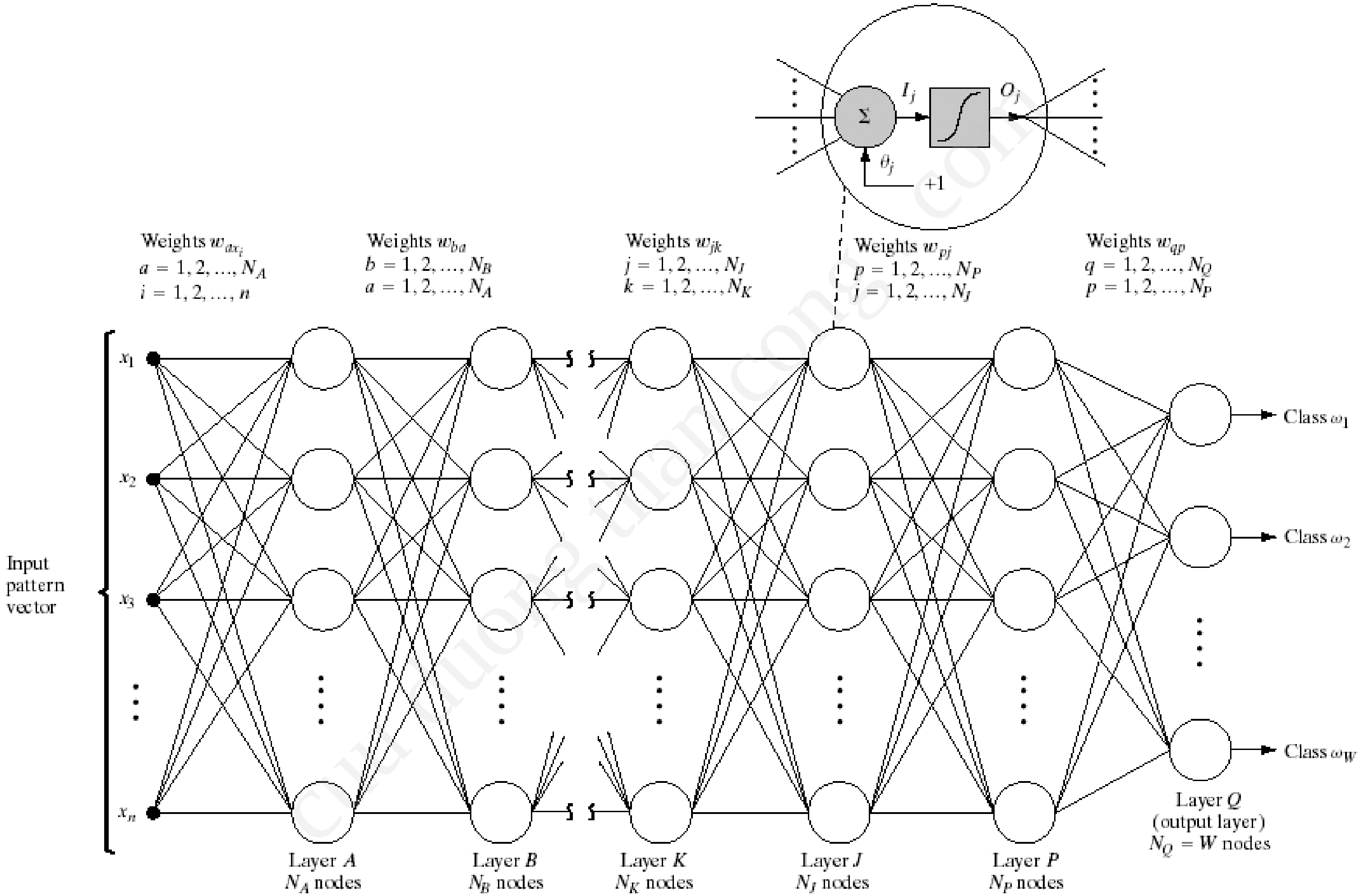


FIGURE 12.16 Multilayer feedforward neural network model. The blowup shows the basic structure of each neuron element throughout the network. The offset, θ_j, is treated as just another weight.

7. Recognition Based on Decision-Theoretic Methods (20)

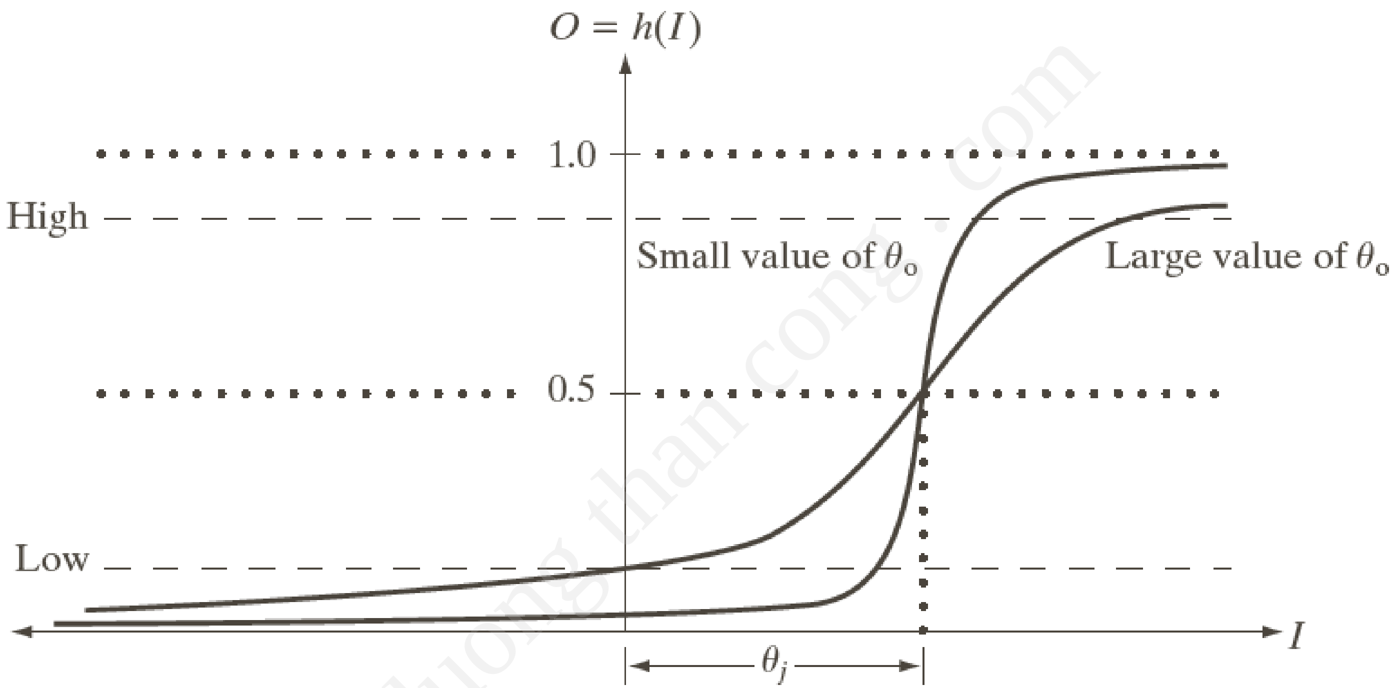
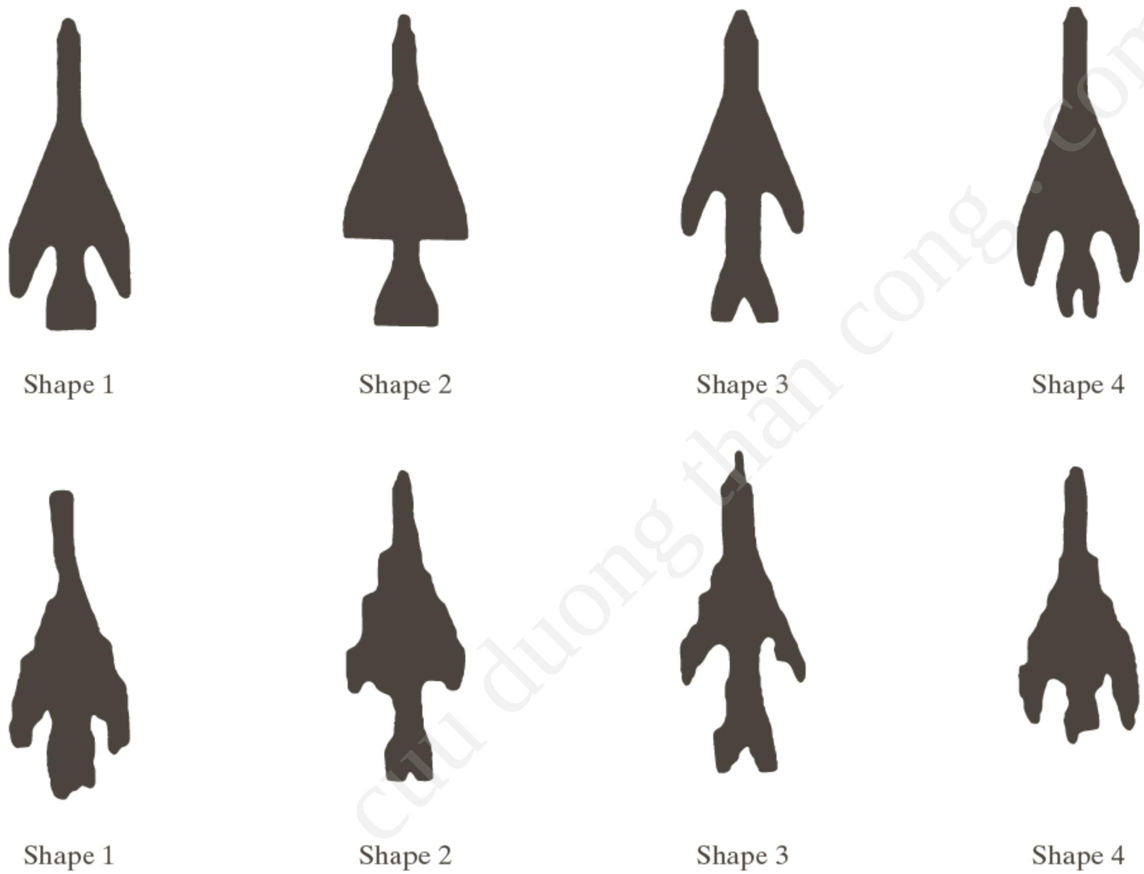


FIGURE 12.17 The sigmoidal activation function of Eq. (12.2-47).

7. Recognition Based on Decision-Theoretic Methods (21)



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.)

7. Recognition Based on Decision-Theoretic Methods (22)

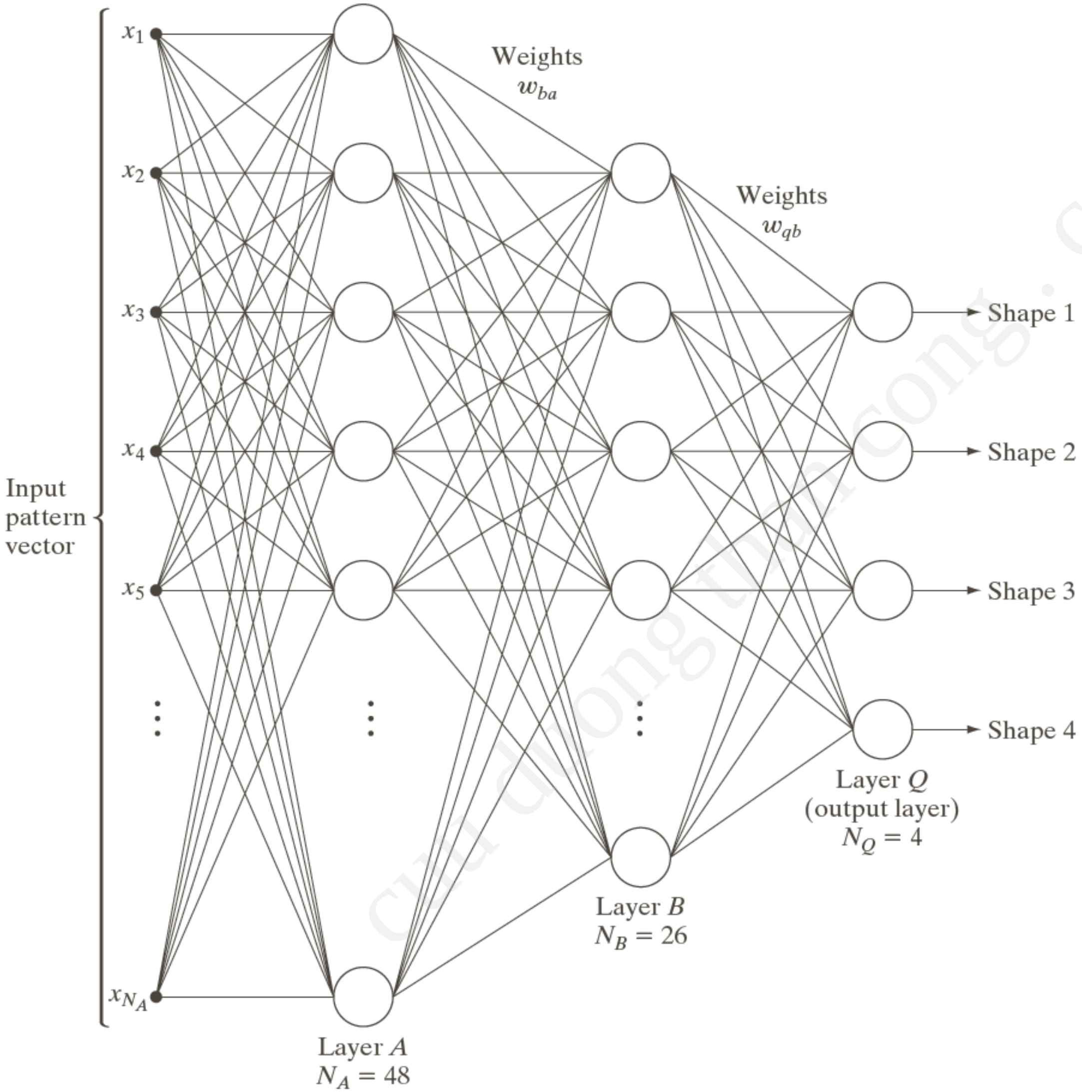


FIGURE 12.19 Three-layer neural network used to recognize the shapes in Fig. 12.18. (Courtesy of Dr. Lalit Gupta, ECE Department, Southern Illinois University.)

7. Recognition Based on Decision-Theoretic Methods (23)

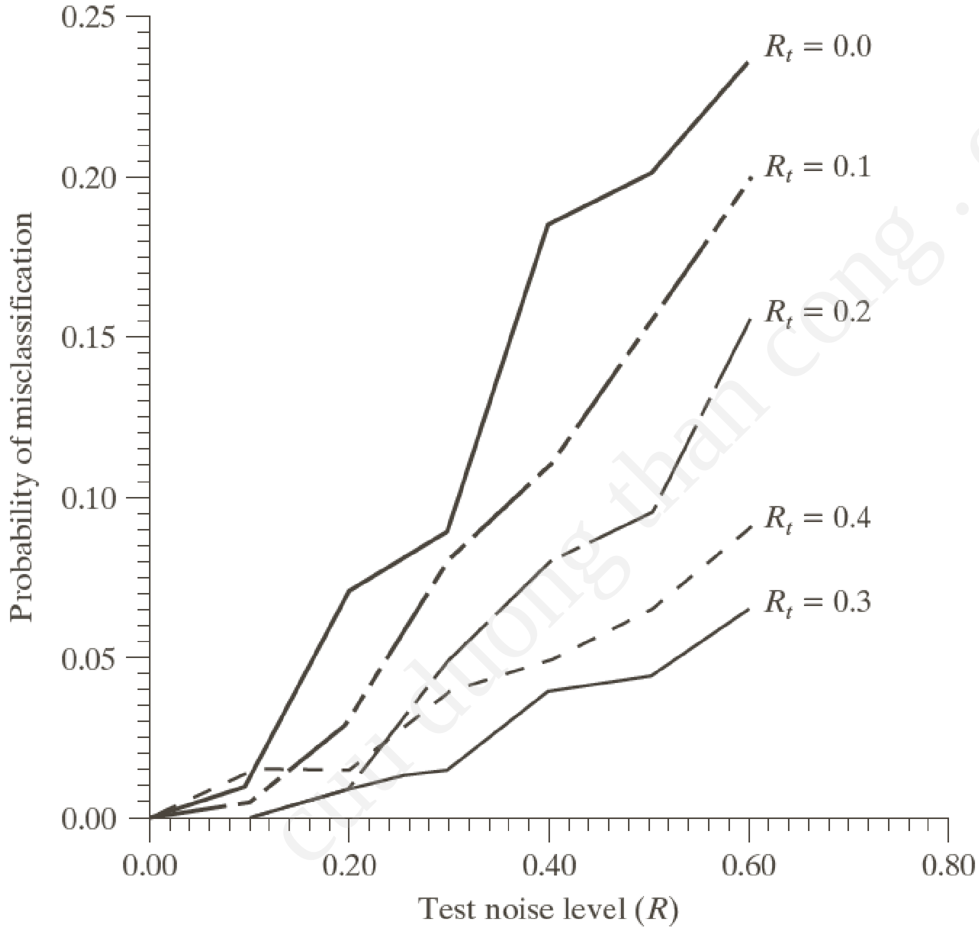


FIGURE 12.20 Performance of the neural network as a function of noise level. (Courtesy of Dr. Lalit Gupta, ECE Department, Southern Illinois University.)

7. Recognition Based on Decision-Theoretic Methods (24)

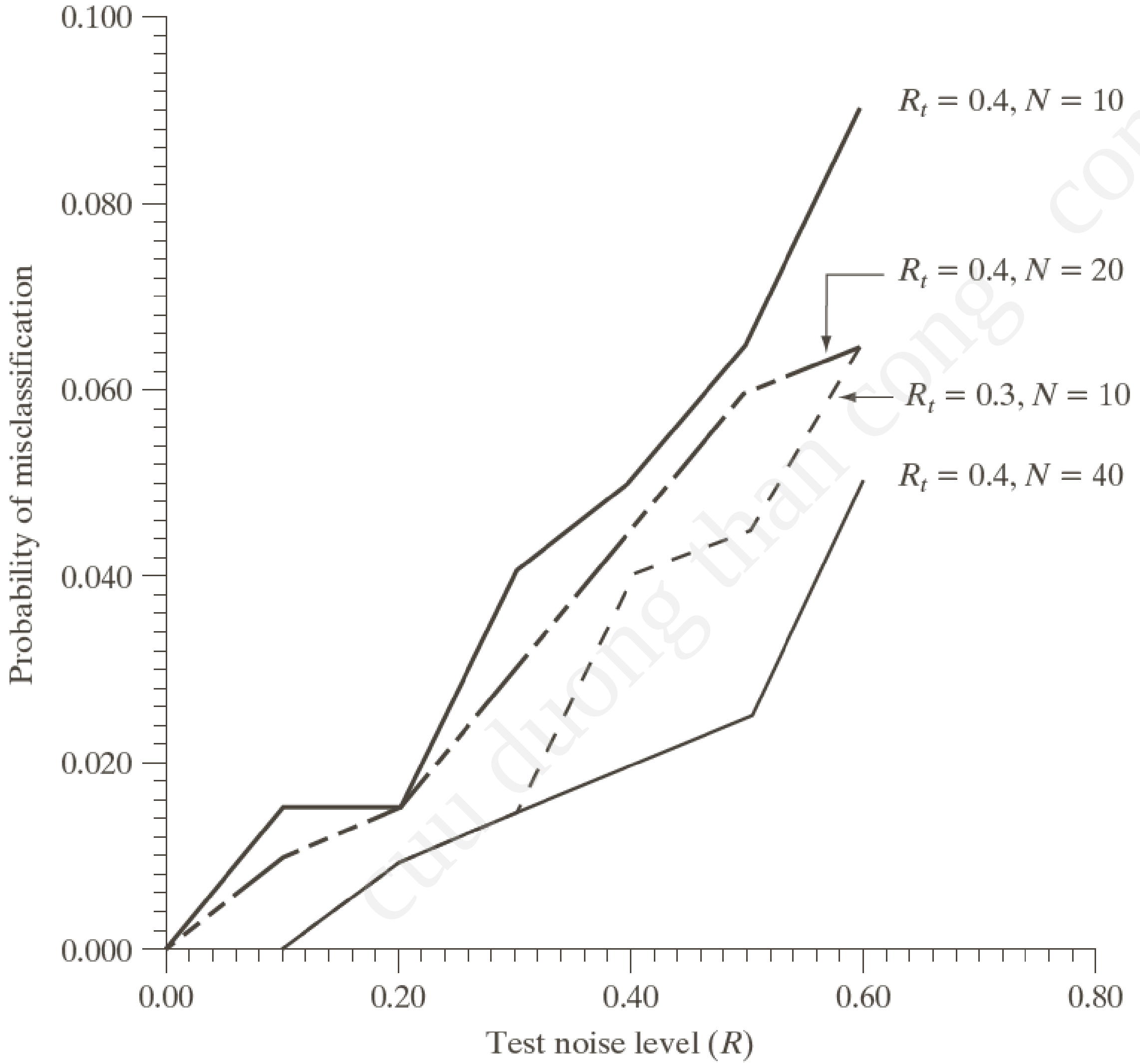
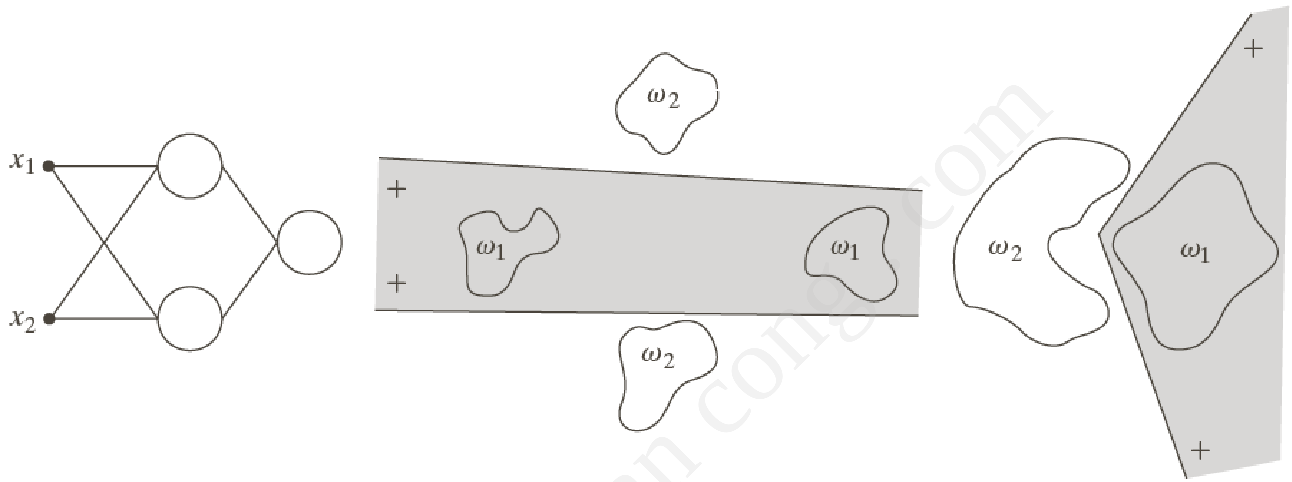


FIGURE 12.21 Improvement in performance for $R_t = 0.4$ by increasing the number of training patterns (the curve for $R_t = 0.3$ is shown for reference). (Courtesy of Dr. Lalit Gupta, ECE Department, Southern Illinois University.)

7. Recognition Based on Decision-Theoretic Methods (25)



a b c

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.

7. Recognition Based on Decision-Theoretic Methods (26)

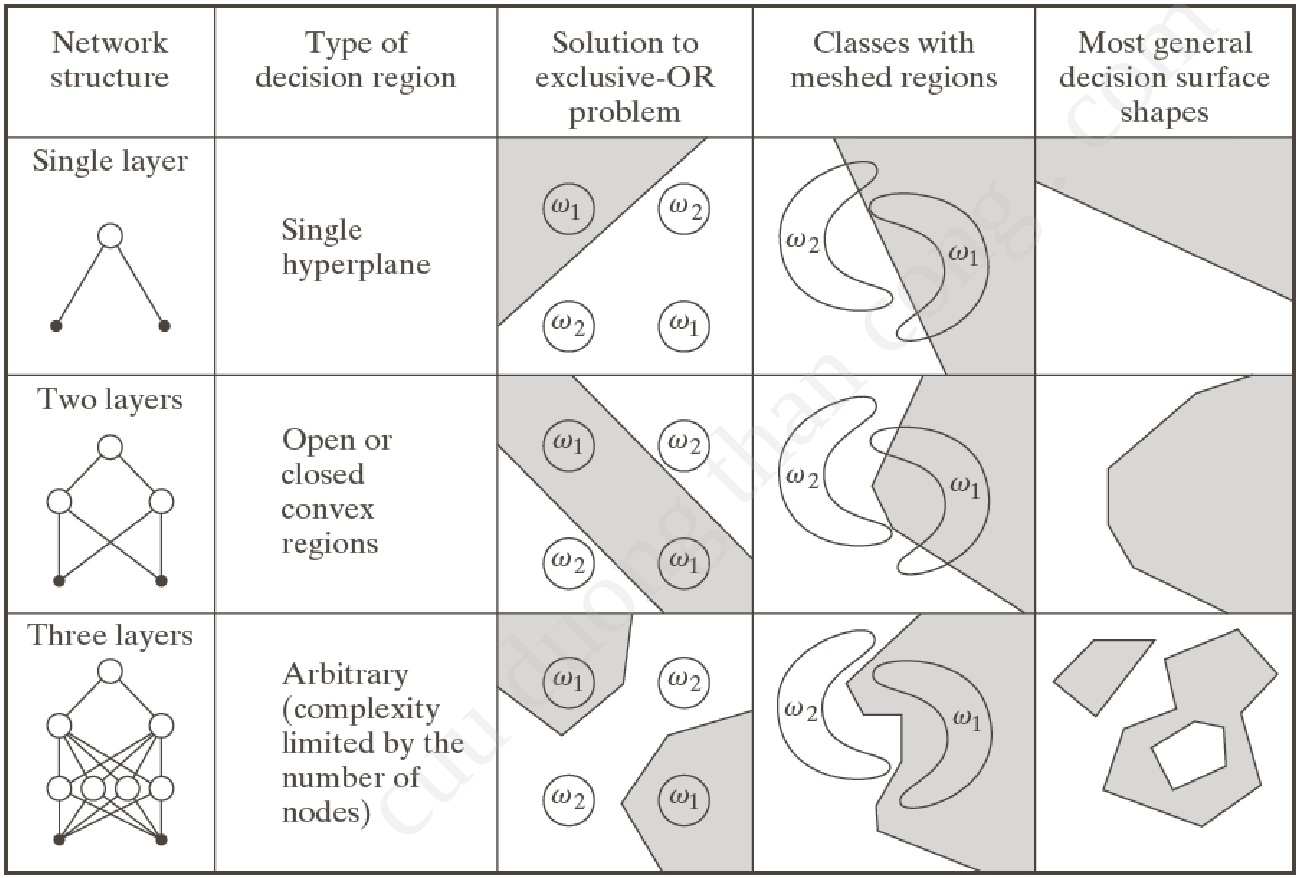


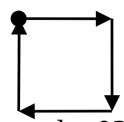
FIGURE 12.23 Types of decision regions that can be formed by single- and multilayer feed-forward networks with one and two layers of hidden units and two inputs. (Lippman.)

7. Recognition Based on Structural Methods (1)

❑ Matching Shape Numbers

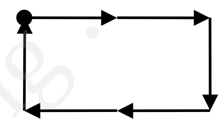
- All shapes of order 4, 6, and 8

Order4



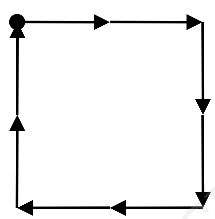
Chain code: **0321**
Difference : **3333**
Shape no. : **3333**

Order6

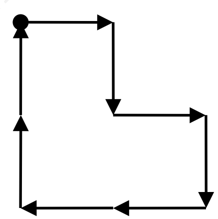


Chain code: **003221**
Difference : **303303**
Shape no. : **033033**

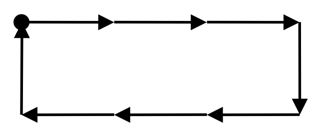
Order8



Chain code: **00332211**
Difference : **30303030**
Shape no. : **03030303**



Chain code:**03032211**
Difference :**33133030**
Shape no. :**03033133**



Chain code: **00032221**
Difference : **30033003**
Shape no. : **00330033**

7. Recognition Based on Structural Methods (2)

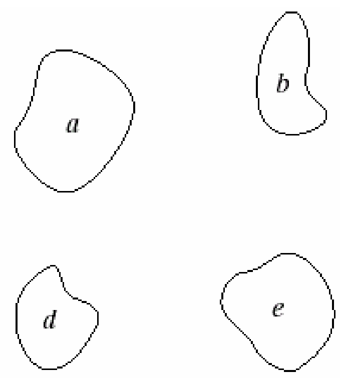
- Advantages:

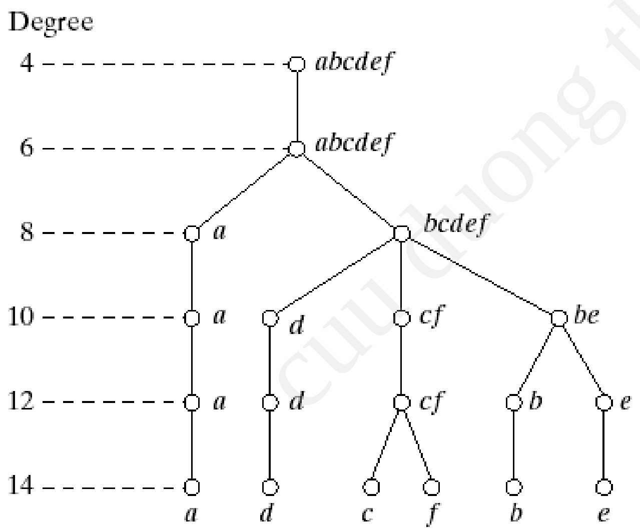
  1. Matching Shape Numbers suits the processing structure simple graph, specially becomes by the line combination.
  2. Can solve rotation the question.
  3. Matching Shape Numbers most emphatically to the graph outline, shape similarity also may completely overcome.
  4. The Displacement question definitely may overcome, because of this method emphatically to the relative position but is not to the position.

7. Recognition Based on Structural Methods (3)

- Disadvantages:
  1. It can not uses for a hollow structure.
  2. Scaling is a shortcoming which needs to change, perhaps coordinates the alternative means .
  3. Intensity.
  4. Mirror problem.
  5. The color is unable to recognize.

7. Recognition Based on Structural Methods (4)





	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.)

7. Recognition Based on Structural Methods (5)

String Matching

- Suppose that two region boundaries, a and b, are coded into strings 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 k th position if $a_k = b_k$

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

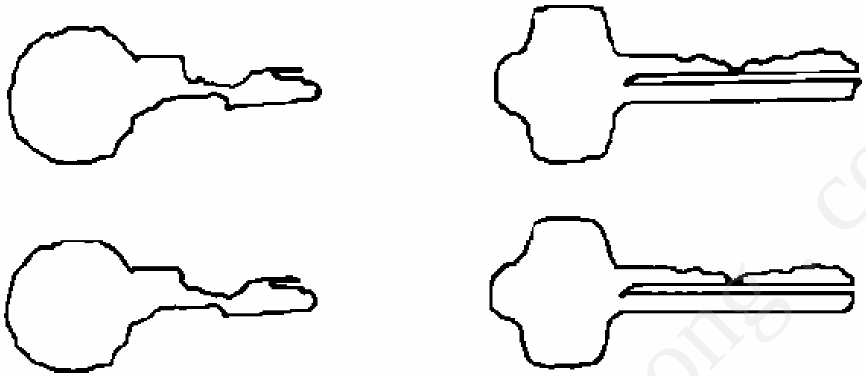
7. Recognition Based on Structural Methods (6)

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

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

Hence R is infinite for a perfect match and 0 when none of the corresponding symbols in and match (in this case)

7. Recognition Based on Structural Methods (7)



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.)