

Object Recognition

The automatic recognition of *objects* or *patterns* is one of the important image analysis tasks. The approaches to pattern recognition are divided into two principal areas:

- Decision-theoretic methods: deal with patterns described using quantitative descriptors, such as length, area, and texture.
- Structural methods: deal with patterns best described by qualitative descriptors (symbolic information), such as the relational descriptors.

Patterns and Pattern Classes

- A pattern is an arrangement of descriptors (or features).
- A pattern class is a family of patterns that share some common properties. Pattern classes are denoted w_1, w_2, \dots, w_N where N is the number of classes.
- Pattern recognition by machine involves techniques for assigning patterns to their respective classes—automatically and with as little human intervention as possible.

The object or pattern recognition task consists of two steps:

- feature selection (extraction)
- matching (classification)

There are three common pattern arrangements used in practice:

- Numeric vectors (for quantitative descriptions)

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

- Strings and trees (for structural descriptions)

$x = \text{abababa}\dots$

Recognition Based on Decision-Theoretic Methods

These methods are based on the use of decision functions. Let $x = (x_1, x_2, \dots, x_n)^T$ represent an n -dimensional pattern vector. For N known pattern classes w_1, w_2, \dots, w_N the idea here is to find N decision functions $d_1(x), d_2(x), \dots, d_N(x)$ with the property that, if 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

Recognition techniques based on matching represent each class by a prototype pattern vector. Set of patterns of known classes is called the training set. Set of patterns of unknown classes is called the testing set. An unknown pattern is assigned to the class to which it is closest in terms of a predefined metric. The simplest approach is the minimum-distance classifier, which, as its name implies, computes the (Euclidean) distance between the unknown and each of the prototype vectors. Then, it chooses the smallest distance to make a decision.

Wavelet-based Face Recognition Application

In the enrolment stage, each face image in the training set is transformed to the wavelet domain to extract its pattern vector (i.e. subband). The choice of an appropriate subband varies depending on the operational circumstances of the face recognition application. The decomposition

level is predetermined based on the efficiency and accuracy requirements and the size of the face image. In the recognition stage, a minimum-distance classification method was used to classify the unknown face images. Figure 15.1 illustrates the stages of this approach.

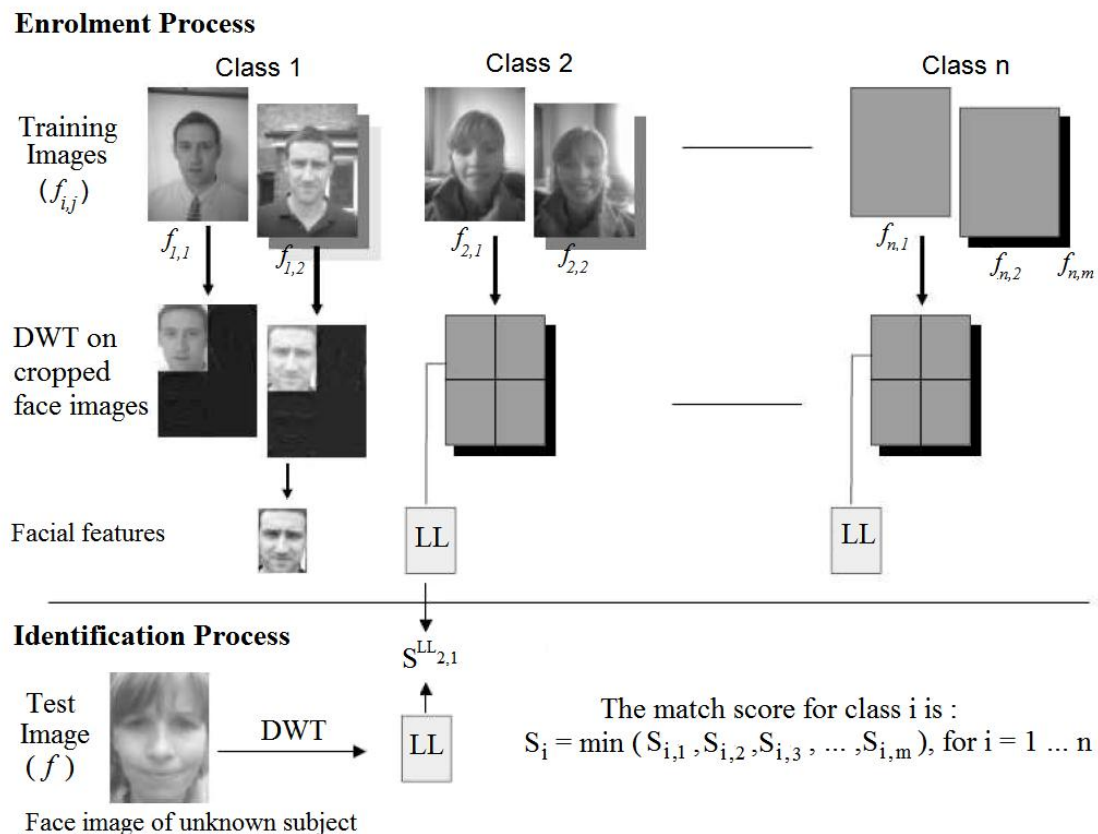


Figure 16.1 Wavelet-based face recognition

Let the set $F = \{ f_{i,1}, f_{i,2}, f_{i,3}, \dots, f_{i,m} \}$ be a training set of face images of n subjects, where each subject i has m images. In the enrolment stage, wavelet transform is applied on each training image so that a set $W_k(F)$ of multi-resolution decomposed images result. A new set $LL_k(F)$ of all k -level LL -subbands will be obtained from the transformed face images in the set $W_k(F)$. The new set $LL_k(F)$ forms the set of features for the training images. Thus, the training face image l of subject i ($f_{i,l}$) is

expressed by its feature vector $LL_{k,i,1}$. The collection of feature vectors $LL_{k,i,1}, LL_{k,i,2}, \dots, LL_{k,i,m}$ represents the stored template of subject i . In a similar manner, 3 new sets $HL_k(F), LH_k(F), HH_k(F)$ can be created at the same decomposition level k from the k th-level HL -, LH -, HH -subbands respectively.

In the recognition phase, a minimum-distance classifier is used to classify the input face image. When a probe face image is introduced to the system, it is decomposed by wavelet transform, and a certain subband (e.g. LL_k) is chosen to represent the feature vector of the probe image. A match score $S_{i,j}$ can now be computed between the probe feature vector and each of the feature vectors j of the subject i in the feature set $LL_k(F)$. Then, the identity (i.e. class) of the training image which gives the minimum score is assigned to the probe image:

$$S_i = \min (S_{i,j}) \quad (j = 1, \dots, m)$$

Many similarity measures can be used for the minimum-distance classifier, for example CityBlock or Euclidean distance functions.

Structural methods

Structural recognition techniques are based on representing objects as strings, trees or graphs and then defining descriptors and recognition rules based on those representations.

The key difference between decision-theoretic and structural methods is that the former uses quantitative descriptors expressed in the form of numeric vectors, while the structural techniques deal with symbolic information.