CHAPTER 3

FACE RECOGNITION TECHNIQUES

3.1 INTRODUCTION

This chapter deals with the techniques for face recognition, used in the present study. The concepts and of the face recognition techniques is explained in detail through the block diagram shown in Figure 3.1

Figure 3.1 Proposed Block Diagram for Face Recognition
The proposed approach for face recognition using neural networks is explained in the block diagram shown in Figure 3.1.

### 3.1.1 Feature Extraction

The intensity of an image is the only source from a camera used for object recognition. However, a lot of variations, such as colour and shape of the object, lighting, etc., are all encoded the intensity. To eliminate the extrinsic factors, various feature extraction and selection methods are widely used. For almost three decades the use of features based on Gabor filters has been promoted for their useful properties in image processing. The most important properties are related to invariance to illumination, rotation, scale, and translation. These properties are based on the fact that they are all parameters of Gabor filters themselves. This is especially useful in feature extraction, where Gabor filters have succeeded in many applications, from texture analysis to iris and face recognition.

The Gabor wavelet was first introduced by David Gabor in 1946. The Gabor wavelet is a sinusoidal plane wave with a particular frequency and orientation, modulated by a Gaussian envelope. It can characterize the spatial frequency structure in the image while preserving information of spatial relations and, thus, is suitable for extracting the orientation-dependent frequency contents of patterns.

Also, the use of Gabor filters in extracting textured image features is motivated by various factors. The Gabor representation has been shown to be optimal in the sense of minimizing the joint two-dimensional uncertainty in space and frequency. These filters can be considered as orientation and scale tunable edge and line (bar) detectors, and the statistics of these micro features in a given region are often used to characterize the underlying texture information. Gabor features have been used in several image analysis.
applications including texture classification and segmentation, image recognition, image registration, and motion tracking.

Lades (1993) pioneered the use of the Gabor wavelet for face recognition using the dynamic link architecture framework. Wiskott (1997) subsequently developed a Gabor wavelet-based elastic bunch graph matching (EBGM) method to label and recognize human faces. In the EBGM method, the face is represented as a graph, each node of which contains a group of coefficients, known as a jet. It can also measure the geometry of the face by using the labeled distance vector, which is the edge part of the graph. Liu and Wechsler (2002) showed that the face representation based on the magnitude part of Gabor feature had been a promising way towards achieving high accuracy face recognition. Shan (2005) provided an AdaBoost-based strategy to select the discriminative features from the magnitude part of the Gabor feature, and then trained a Fisher classifier to make a final classification. As a powerful descriptor, the Gabor wavelet is also used in many applications, such as data compression optical character recognition (OCR), texture analysis, fingerprint recognition, and so on. Most of the above applications are based on the magnitude part of Gabor feature. In fact, the Gabor phase is a very discriminative information source, and has been successfully used in iris and palm print identification.

3.1.2 Gabor Functions and Wavelets

The Gabor space is very useful in image processing applications such as iris recognition and fingerprint recognition. Relations between activations for a specific spatial location are very distinctive between objects in an image. Furthermore, important activations can be extracted from the Gabor space in order to create a sparse object representation.
3.1.2.1 Gaber Wavelets and Analysis

One difficulty with the Fourier transform is that Fourier coefficients depend on the entire image, the value of the Fourier transform for particular \((u,v)\) is computed using every image pixel. This is an inconvenient way to think of images, because we have lost all spatial information. For example, the stripes of the Figure 3.2 get wider as one moves across the images. If we think in terms of spatial frequency only locally defined, then we can think of this phenomenon in terms of spatial frequency content of the image changing as we move across it. In some window around a point, the narrow stripes look like high spatial frequency terms and the wide stripes look like low spatial frequency terms.

Gabor filters achieve this. The Kernels look like Fourier basis elements that are multiplied by Gaussian, meaning that a Gabor filter responds strongly at points in an image where there are components that locally have a particular spatial frequency and orientation. Gabor filters come in pairs, often referred to as quadrature pairs, one of the pair recovers symmetric components in a particular direction and the other recovers antisymmetric components.

Gabor filters are directly related to Gabor wavelets, since they can be designed for number of dilations and rotations. However, in general, expansion is not applied for Gabor wavelets, since this requires computation of biorthogonal wavelets, which may be very time-consuming. Therefore, usually, a filter bank consisting of Gabor filters with various scales and rotations is created. The filters are convolved with the signal, resulting in a so-called Gabor space. This process is closely related to processes in the primary visual cortex.
A **Gabor filter** (Appendix 1) is a linear filter whose impulse response is defined by a harmonic function multiplied by a Gaussian function. Because of the multiplication-convolution property (Convolution theorem), the Fourier transform of a Gabor filter's impulse response is the convolution of the Fourier transform of the harmonic function and the Fourier transform of the Gaussian function.

\[
g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(\frac{x^2 + \gamma^2 y^2}{2\sigma^2}\right) \cos\left(2\pi \frac{x'}{\lambda} + \psi\right)
\]  

(3.1)

Where

\[
x' = x \cos \theta + y \sin \theta
\]

(3.2)

And

\[
y' = -x \sin \theta + y \cos \theta
\]

(3.3)

In this equation, \( \lambda \) represents the wavelength of the cosine factor \( \cos\left(2\pi \frac{x'}{\lambda} + \psi\right) \), \( \theta \) represents the orientation of the normal to the parallel stripes of a Gabor function, \( \psi \) is the phase offset, \( \sigma \) is the sigma of the Gaussian envelope and \( \gamma \) is the spatial aspect ratio, and specifies the ellipticity of the support of the Gabor function.
A two dimensional Gabor function and its Fourier transform

\[ G(u, v) \] can be written as:

\[
g(x, y) = \frac{1}{2\pi \sigma_x \sigma_y} \exp \left[ -\frac{1}{2} \left( \frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) + 2\pi j W_x \right]
\] (3.4)

\[
G(u, v) = \exp \left\{ -\frac{1}{2} \left[ \frac{(u - W)^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2} \right] \right\}
\] (3.5)

where \( \sigma_u = 1/2\pi \sigma_x \) and \( \sigma_v = 1/2\pi \sigma_y \). Gabor functions form a complete but nonorthogonal basis set. Expanding a signal using this basis provides a localized frequency description. A class of self-similar functions, referred to as Gabor wavelets in the following discussion, is now considered.

Let \( g(x, y) \) be the mother Gabor wavelet, then this self-similar filter dictionary can be obtained by appropriate dilations and rotations of \( g(x, y) \) through the generating function:
\[ g_{mn}(x, y) = a^{-m} G(x', y'), \quad a > 1, \quad (3.6) \]

Where, \( m, n = \text{integer} \)

\[ x' = a^{-m} (x \cos \Theta + y \sin \Theta), \quad (3.7) \]

And

\[ y' = a^{-m} (-x \sin \Theta + y \cos \Theta), \quad (3.8) \]

Where \( \Theta = n \pi / K \) and \( K \) is the total number of orientations. The scale factor \( a^{-m} \) is meant to ensure that the energy is independent of \( m \). The cosine factor is defined as \( \cos \left( \frac{2\pi x'}{\lambda - v} \right) \)

### 3.1.3 Affine Transformation

An affine transformation is any transformation that preserves colinearity (i.e., all points lying on a line initially still lie on a line after transformation) and ratios of distances (e.g., the midpoint of a line segment remains the midpoint after transformation). In this sense, affine indicates a special class of projective transformations that do not move any objects from the affine space to the plane at infinity or conversely. An affine transformation is also called an affinity.

Geometric contraction, expansion, dilation, reflection, rotation, shear, similarity transformations, spiral similarities, and translation are all affine transformations, as are their combinations. In general, an affine transformation is a composition of rotations, translations, dilations, and shears.

While an affine transformation preserves proportions on lines, it does not necessarily preserve angles or lengths. Any triangle can be
transformed into any other by an affine transformation, so all triangles are
affine and, in this sense, affine is a generalization of congruent and similar.

A particular example combining rotation and expansion is the
rotation-enlargement transformation

\[
\begin{bmatrix}
  x' \\
  y'
\end{bmatrix}
= s
\begin{bmatrix}
  \cos \alpha & \sin \alpha \\
  -\sin \alpha & \cos \alpha
\end{bmatrix}
\begin{bmatrix}
  x - x_o \\
  y - y_o
\end{bmatrix}
\]  \hspace{1cm} (3.9)

\[
\begin{bmatrix}
  x' \\
  y'
\end{bmatrix}
= s
\begin{bmatrix}
  \cos \alpha (x - x_o) + \sin \alpha (y - y_o) \\
  -\sin \alpha (x - x_o) + \cos \alpha (y - y_o)
\end{bmatrix}
\]  \hspace{1cm} (3.10)

Separating the equations,

\[
x' = (s \cos \alpha) x + s (\sin \alpha) y - s (x_o \cos \alpha + y_o \sin \alpha)
\]  \hspace{1cm} (3.11)

\[
y' = (-s \sin \alpha) x + (s \cos \alpha) y + s (x_o \sin \alpha - y_o \cos \alpha)
\]  \hspace{1cm} (3.12)

This can be also written as

\[
x' = ax - by + c
\]  \hspace{1cm} (3.13)

\[
y' = bx + ay + d
\]  \hspace{1cm} (3.14)

where

\[
a = s \cos \alpha
\]  \hspace{1cm} (3.15)

\[
b = -s \sin \alpha
\]  \hspace{1cm} (3.16)

The scale factor is then defined by

\[
s = \sqrt{a^2 + b^2}
\]  \hspace{1cm} (3.17)
and the rotation angle by

\[
\alpha = \tan^{-1}\left(-\frac{b}{a}\right) \tag{3.18}
\]

### 3.1.4 Histogram

An image histogram is type of histogram which acts as a graphical representation of the tonal distribution in a digital image. It plots the number of pixels for each tonal value. By looking at the histogram for a specific image a viewer will be able to judge the entire tonal distribution at a glance. Image histograms are present on many modern digital cameras and can be used as an aid to show the photographer whether he or she has captured and adequate amount of tones or whether image detail has been lost to blown-out highlights or blacked-out shadows. The horizontal axis of the graph represents the tonal variations, while the vertical axis represents the amount of pixels in that particular tone. The left side of the horizontal axis represents the black and dark areas, the middle represents medium grey and the right hand side represents light and pure white areas. The vertical axis, on the other hand, represents the size of the area which is captured in each one of these zones.

The histogram method has been widely used to represent, analyze, and recognize images because it can be calculated easily and efficiently, and robust to the noise and local image transformations. One of initial applications of histograms was the work of Swain and Ballard (1991) for the identification of 3-D objects. Subsequently, various recognition systems proposed by Schmid (1997), Hadjidemetriou (2004) based on histograms were developed. However, the histogram is not adequate for many applications as it suffers from losing the structure information of the object. A compact and effective object descriptor – histogram of Gabor phase pattern is proposed for robust face recognition. The Gabor phase pattern is having large number of values.
Hence histogram of each matrix is compared. The approach is based on the combination of the spatial histogram and the Gabor phase information encoding scheme.

3.1.5 KLT for facial expression recognition

Karhunen–Loeve Transform based dimensionality reduction for face images was first proposed by Kirby and Sirovich (1990). The KLT is a linear transform, where the basis functions are taken from the statistics of the signal and can thus be adaptive. Mathematically, the eigenface method tries to represent a face image as a linear combination of orthonormal vectors, called eigenfaces. These eigenfaces are obtained by finding the eigenvectors of the covariance matrix of the training face image set. Let $I_1, I_2, I_3, \ldots, I_k$ be a set of $k$ face images, each ordered lexicographically.

The eigenvectors of the matrix

$$C = \sum_{i=1}^{k} I_i I_i^T$$

(3.19)

that corresponds to the largest eigenvalues span a linear subspace that can reconstruct the face images with minimum reconstruction error in the least squares sense. This $L$-dimensional subspace is called the face space. Assuming is a lexicographically ordered face image and is the matrix that contains the eigenfaces as its columns, we can write

$$x = \phi a + e_x$$

(3.20)

Where ‘$a$’ is the feature vector that represents the face, and $e_x$ is the subspace representation error for the face image. As a larger training data set is used and the dimensionality of the face space is increased, the representation error gets smaller. Letting
be the feature vector, and

\[ \phi = [ \phi_1, \phi_2, \ldots, \phi_L ] \] (3.22)

be the matrix where are the eigenface vectors, is computed as follows:

\[ a_i = \phi_i^T x, \] (3.23)

The computed eigen values are compared for similarity and the expression of the face is recognized.

3.2 SUMMARY

In this chapter, the different techniques for face recognition are clearly explained. The important concepts, artificial neural networks and optical neural networks are explained in the subsequent chapters.