CHAPTER 4

PROPOSED METHOD OF IRIS FEATURE EXTRACTION

4.1 INTRODUCTION

The primary objective of iris coding is to obtain a good interclass separation between the authentic user and imposter. The recognition rate of such systems depends on the discrimination power of classifier and invariant features of the image (robust feature extraction technique). This thesis emphasizes both because the decision taken by the classifier depends on the invariant features. This chapter covers the proposed method of feature extraction for images captured in uncontrolled environment. Further, this thesis proposes an iris subimage method that is based on the division of the normalized iris image into eight regions. The independent feature extraction and comparison of each subimage region, avoiding localized noisy iris regions corrupting the whole biometric signature. A multiresolution scheme called undecimated pyramidal decomposition method is developed, which produces an undecimated output (input and output are of same size). A directional filter bank is employed to extract the directional features from an undecimated output. A novel coding scheme is introduced to generate rotation invariant features and a fusion algorithm to combine the similarity measure of each subimage. Decision is taken based on the final fusion score. The experimental results in the next chapter demonstrate that the proposed algorithm performs better than other algorithms.
Figure 4.1 Proposed feature extraction algorithm

The proposed approach of feature extraction for nonideal iris images is shown in Figure 4.1, in which the normalized image is divided into eight subimages. Each subimage is decomposed by undecimated pyramid before the DFB decomposition. Rotation invariant features are obtained from the DFB subbands using optimal projection analysis. Classifiers of each subimage compare the invariant features of the test image with the template
image. A fusion rule combines the scores from individual classifiers for final decision.

4.2 IMAGE PREPROCESSING

An iris image not only contains the region of interest, but also some useless parts such as eyelid, eyelash and pupil. Size of the same iris varies as the camera to eye distance changes. Practical conditions such as different illumination, noise levels also affect the recognition rate. It is difficult to acquire a good image, but all the subsequent steps produce an image suitable for further processing. Various steps involved in iris recognition are localization, normalization, feature extraction and classification.

Localization

In the pattern recognition domain, segmentation is the assignment of each pixel to an image region, which can be regarded as a typical classification problem. Regarding the iris biometrics compass, the segmentation stage receives a close-up eye image and localizes the pupillary and scleric iris borders in the image as shown in Figure 4.2.

(a) eye image    (b) Edge detected image   (c) Localized iris image

Figure 4.2 Segmented iris image with the inner (pupillary) and outer (scleric) borders signalled by circles
In most of the iris recognition algorithms Li Ma et al (2004) method is used for localization and normalization of the iris image. Since it accurately localizes the iris regions I also follow the same algorithm for my method. The iris localization is to detect the iris area between pupil and sclera from an eye image. The boundaries of the iris and the pupil are detected by considering the pupil as the darkest area in the image. The localization algorithm is based on Li Ma (2004) method.

Estimate of its center $(C_x, C_y)$ is performed using the following formula:

\[
C_x = \arg \min_x \left( \sum_y I(x, y) \right) \quad (4.1)
\]

\[
C_y = \arg \min_y \left( \sum_x I(x, y) \right) \quad (4.2)
\]

where $I(x, y)$ is the iris image intensity at point $(x, y)$. To find the exact centre of the pupil, a part of image is binarized using an adaptive threshold obtained by the histogram of a square window and centered at the estimated center $(C_x, C_y)$ and then the radius of the pupil is obtained by tracing from the centre of the pupil to the boundary between the iris and the pupil on different directions in the binary image and averaging them. In order to detect the boundary between the iris and the sclera, the image is convolved with a blurring function, which is a 2-D Gaussian operator with centre at $(x_0, y_0)$.

\[
G(x, y) = \frac{1}{2\pi \delta^2} e^{-\frac{(x-x_0)^2 + (y-y_0)^2}{2\delta^2}} \quad (4.3)
\]
where $\delta$ is standard deviation, that smoothes the image and then apply Canny operator with the threshold values 0.005 and 0.1 as the low and the upper limits. Now finding the radius of the iris is similar to that for the pupil. These two radii localize the iris as shown in Figure 4.2

**Normalization**

Assuming that each iris has a unique pattern that does not change over time, it is desired to capture these patterns every time an image is taken from the eye. However, an iris texture is highly sensitive to illumination and the texture deforms to change the pupil diameter and control the amount of light entering the eye. The deformation of the texture also results in the deformation of the iris patterns. In other words, the iris patterns change due to different illumination conditions. In order to obtain the unique patterns of the iris, it is desired to be able to track the changes of the iris to recover the unique patterns. It should be mentioned that the way an iris responses to illumination also varies from eye to eye due to different distributions of the muscles controlling the pupil.

Locating iris in the image delineates the circular iris zone of analysis by its own inner and outer boundaries. The system normalizes the iris region to overcome the problem of a change in camera-to-eye distance and pupil’s size variation derived from illumination. To compensate the stretching of the iris texture as the pupil changes in size, and have a new model of iris which removes the non-concentricity of the iris and the pupil the Cartesian to polar reference transform suggested by Daugman(2004) authorizes equivalent rectangular representation of the zone of interest. In this way I compensate the stretching of the iris texture as the pupil changes in size, and I unfold the frequency information contained in the circular texture in order to facilitate next features extraction. Moreover this new representation of the iris breaks the noneccentricity of the iris and the pupil.
The angle $\theta \in (0 - 2\pi)$ and radius $p \in (0 - 1)$ describe the polar coordinate system (Daughman 2004) $I(x(p,\theta), y(p,\theta)) = I(p,\theta)$

\[
\begin{align*}
x(p,\theta) &= (1 - p) * x_p(\theta) + p * x_i(\theta) \\
y(p,\theta) &= (1 - p) * y_p(\theta) + p * y_i(\theta)
\end{align*}
\] (4.4)

with

\[
\begin{align*}
x_p(\theta) &= x_{p0}(\theta) + r_p * \cos(\theta) \\
y_p(\theta) &= y_{p0}(\theta) + r_p * \sin(\theta)
\end{align*}
\] (4.5)

\[
\begin{align*}
x_i(\theta) &= x_{i0}(\theta) + r_i * \cos(\theta) \\
y_i(\theta) &= y_{i0}(\theta) + r_i * \sin(\theta)
\end{align*}
\] (4.6)

where $r_p, r_i$ are the radius of the pupil and the iris respectively, while $(x_p(\theta), y_p(\theta))$ and $(x_i(\theta), y_i(\theta))$ are the coordinates of the papillary and iris boundaries in the direction $\theta$. The normalized image is shown in Figure 4.3. The normalization not only reduces exactly the distortion of the iris caused by pupil movement, but also simplifies subsequent processing.

![Figure 4.3 Normalized iris image](image1)

![Figure 4.4 Enhanced iris image](image2)
Image Enhancement

On account of imaging conditions and situations of light sources, the normalized iris image does not have an appropriate quality. These factors may affect the performance of feature extraction and matching processes. Hence for getting a uniform distributed illumination and better contrast, the unwrapped flat iris is enhanced by eliminating background and applying histogram equalization (Li Ma et al 2004).

In order to obtain a well-distributed texture image, I first approximate the intensity variations across the whole image. The mean of each 16x16 small block constitutes a coarse estimate of the background illumination. This estimate is further expanded to the same size as the normalized image by bicubic interpolation. The estimated background illumination is subtracted from the normalized image to compensate for a variety of lighting conditions. Then, I enhance the lighting corrected image by means of histogram equalization in each 32x32 region. Such processing compensates for the nonuniform illumination, as well as improves the contrast of the image. Figure 4.4 shows the preprocessing result of an iris image, from which I can see that finer texture characteristics of the iris become clearer than those in Figure 4.3.

4.3 NOISY IRIS RECOGNITION

The capture of iris images at large distances, under less controlled lighting conditions, and without active participation of the subjects results in low quality images. This increases the probability of capturing heterogeneous images regarding focus, contrast, several noises (ie) iris obstructions and reflections. This will directly affect the performance of the system.
Considering the above said problems, the normalized iris image is divided into separate regions so that some of these regions are noise-free. Feature extraction and comparison are done for each region separately. This will avoid the noise localized in some of the iris subparts corrupting the whole iris features and reduces the false rejection rate. Matching scores from individual classifiers are fused together using fusion rule as explained in section 4.6.

4.3.1 Noise Factors in Normalized Iris Images

The captured irises resultant of non-cooperative imaging environments contain several other types of information considered as noise. The image region that corresponds to any other types of information apart from the iris and is localized within the region delimited by the pupillary and scleric iris borders is considered as noise. In the iris recognition literature it is common to consider all the information that obstructs portions of the iris as noise. The pupil's size and the area of the regions corresponding to the iris will vary due to variations in the image capturing distances and due to lighting conditions. In order to compensate for this variation, common iris recognition proposals translate the segmented iris image into a double dimensionless pseudo-polar coordinate system, in a process known as the iris normalization stage.

![Figure 4.5 Noisy iris image](image)
Figure 4.6 Normalized iris image with noise

It is expectable that unconstrained image acquisition processes decrease the quality of the captured data and increase its heterogeneity. In iris recognition literature, different noise factors are classified into one of two major categories (Wai-Kin Kong and David Zhang 2003): local or global, as they affect exclusive image regions or the complete image. The local category comprises iris obstructions, reflections, off-angle and partial images, while the global comprises poorly focused, motion blurred, rotated, improperly illuminated and out-of-iris images. The most common noise regions that result from non-cooperative imaging processes is shown in the captured iris image of Figure 4.5, and the normalized iris image with noise is shown in Figure 4.6. The normalized image has low contrast and non-uniform brightness. Moreover the portion of the pupil incorrectly located on the normalized image, is considered as noise. Other noises encountered are eyelid and eyelash occlusion, lighting reflections from artificial light sources near to the subject and specular reflections corresponding to reflected information from the environment where the user is located. Figure 4.7 shows the various types of noisy images. Several moving parts that interact in the iris image capturing, results in motion blurred image and the eyelids movement has a significant contribution to this type of noise.
Eyelid and eyelash obstructions usually occur in the lower part of the normalized images, corresponding to the upper and lower iris extremes which are naturally obstructed by eyelid movement. Oppositely, reflections are determined by heterogeneous lighting conditions, both specular and lighting reflections are highly disseminated across the irises. Finally, the noise region due to the inaccurate pupil segmentation results in the wrong classification of a portion of the pupil as belonging to the iris.

Eyelashes can occlude portions of the iris in two distinct forms (Yingzi Du et al 2005) (i) isolated or (ii) grouped. If an eyelash is isolated it appears as a very thin and dark line in the iris region. The existence of grouped or multiple eyelashes in the iris region generates a uniform dark region. Isolated or separable eyelashes can be distinguished against the texture of the iris, whereas multiple eyelashes present a larger region of occlusion.
4.3.2 Subimage Formation

Iris images captured in an uncontrolled environment produce nonideal iris images with varying image quality. If the eyes are not properly opened, certain regions of the iris cannot be captured due to occlusion, which further affects the recognition performance. Images may also suffer from motion blurr, presence of eyelids and eyelashes, head rotation, reflections, contrast, luminosity, and problems due to contraction and dilation. New challenges in human authentication systems such as noncooperative behavior have attracted researcher’s attention in recent years. Some of the features of non-cooperative iris recognition systems are

- **Security**: Since there is no need for user cooperation, it is possible to recognize the individual’s identity when and where the user is not aware of identification process. This achievement is suitable to identify subjects in terrorist attacks.

- **User commodity**: The less cooperative behavior leads to more user commodity in image capture process and reduces the time which is necessary to capture an appropriate image.

- **Functioning Radius**: Noncooperative recognition enables to perform identification process in larger distance than that of cooperative systems (usually less than 1 m).

It is observed that, in most cases, the noise is localized in some subpart of the iris. The proposed method is based on the division of the segmented iris into eight regions, followed by the independent feature extraction on each one. Further, through the feature comparison between features extracted from corresponding regions, eight dissimilarities are obtained and fused according to a classification fusion rule. The hope is that
most of the iris regions are noise-free and that accurate recognition can be achieved, even in highly noisy images. The following sub-section describes the proposed noisy recognition method.

Eyelashes and eyelids are the two occluded noise portions of the upper and lower part of the iris. Under natural and artificial lightning conditions, reflections are mostly seen in outer and inner iris region. When an error on the pupil segmentation occurs, a portion of the pupil will be considered as belonging to the iris. Moreover, in the normalized iris image, the wrongly identified portion of the pupil is located on the upper band of the image. Previous methods (Li Ma et al 2004, Tisse et al 2002, Daugman 2004) performed recognition based on small portions of the iris, where noise is less probable. However, they are static and do not take into account the dynamics of the imaging environments, which determine the iris regions corrupted by noise.

When iris is divided into different regions, some of these regions may be noise free. When I compare the noise free part of the iris with the enrolled region, it produces accurate recognition. This decreases the error rate in the recognition of noisy iris images in which the independent feature extraction and comparison for each of these regions effectively avoid the noise, localized in some of the iris subparts corrupting the whole iris features.

![Figure 4.8. Division of iris image into subregions](image)
As shown in Figure 4.8 the iris pattern is divided into 8 parts of equal size. Normalized eye image of size 32x256 is divided into eight segments of size 32x32. The noisy information is localized in some of the iris subparts. Most of the types of noise such as iris occlusions and reflections are predominantly localized in the upper/lower and left/right portions of the iris. Also, reflections that are resultants of natural and artificial lighting environments are in the outer and inner iris regions. Hence various types of noise do not affect all the subregions.

This research work considers “noise” of occlusion and pupil dilation. Specular highlights are not a serious problem on this set of images as compared with eyelid occlusion. Due to inaccurate segmentation, portion of the pupil occupies the lower portion of the normalized image, which is shown in Figure 4.8. By dividing the image into different regions, the effect of this noise is present only in some (second, third and sixth) subimages.

4.4 DIRECTIONAL FILTERING AND QUALITY MEASURE

The images captured in an uncontrolled environment produce heterogeneous images with varying image quality. Poor quality images reduce the accuracy of iris-recognition systems. The quality measure is a combination of the distinctiveness of the iris region from the amount of iris region available. In this thesis quality of the image is measured by the energy estimate of the directional subbands using directional filter bank.

Directional information from an image can be extracted in many ways. Directional filter banks are efficient directional filtering structures that decompose the image into different wedge like directional subbands. The iris image consists of features like freckles, coronas and stripes which are more
directional. Therefore directional filter banks are used to get the directional features from an iris image. The normalized iris image is divided into eight subimage of equal size. Each subimage is further decomposed into eight directional subbands. Feature vectors of each subimage are formed independently. Matching scores of each subimage are obtained by comparing the feature vector of each subimage with their respective template feature vector. Finally a fusion algorithm combines the matching scores from each subimage depends on the quality of the subimage and a decision is made.

4.4.1 Directional Analysis Techniques

Directional decomposition of images is a process, which consists of decomposing an image into several subbands, each representing directional information or energy along one direction. A directional decomposition can be used in many fields in image processing (Rosiles and Smith 2001, Rosik 2004, Chul-Hyun Park et al 2004b), for example, common automatic target recognition, texture analysis, segmentation, classification, (Mohammed at al 2004) image denoising and enhancement.

Filter banks and wavelets are able to generate features from images by decomposing them based on different frequency regions. Directional information in a given image can be extracted by various methods such as Discrete wavelet transform (Jafar Ali and Aboul Hassanien 2003) and 2D Gabor filter bank (Li Ma et al 2002a). Gabor filter bank is one of the most well known methods used for this purpose and many algorithms have been proposed. However, as described in (Park et al 1999), the use of a Gabor filter
bank inherently results in some overlapping and missing subband regions. In
Gabor filter bank, each Gabor filter represents a separate frequency channel
thereby mimicking the human visual system (HVS). Gabor representations are
significantly influenced by many parameters in the feature extraction process
(spatial position, orientation, center frequencies, and size parameters of 2D
Gabor filter), which may vary, depending on the environmental factors of iris
image acquisition. For a set of test iris images, extensive parameter
optimization is required to achieve high recognition rate.

The discrete wavelet transform (DWT) (Boles and Boashash 1998)
has enjoyed success in the field of iris image classification. The 2-D DWT has
limited angular selectivity along horizontal, vertical and diagonal orientations.

The Directional Filter Bank (DFB) (Bamberger and Smith 1992,
Smith and Eddins 1990), on the other hand, is a contiguous subband
representation that preserves all image information. Accordingly, a DFB can
represent linear patterns, as found in iris patterns, more effectively than a
Gabor filter bank. DFB is implemented using a novel class of linear phase
filters. These filters enable the DFB to achieve additional reduction in
complexity and allow the image boundaries to be handled by symmetric
extension. The positive effect of extracting features using a specific
directional and frequency subband that emphasizes the dominant frequency
information and suppresses noise components is much offset by the
suppression of useful information existing outside the specified frequency
range.

In the proposed method, directional filter bank (DFB) is used to get
the directional information. The directional filter bank has been proven to be
effective in feature extraction due to the attractive property that they can
decompose an image into directional subbands, each corresponding to a
unique angular orientation and reconstruct that image perfectly by using a
dual synthesis filter bank. The DFB divides the two-dimensional spectrum of
an image into wedge-like directional subbands. The DFB provides a strong
case for multichannel approaches to iris recognition, and in particular, results
obtained using the DFB indicate that there is not necessarily a large trade-off
between computational efficiency and high classification accuracy (Park 1999).

4.4.2 Directional Filtering using Directional Filter Bank

Iris patterns include many linear features that can be considered as
a combination of directional linear pattern components. As such, the unique
c characteristics of an iris pattern can be effectively represented by a feature
vector constructed by extracting the linear components of an iris according to
directionality. A DFB is suitable for extracting iris features, since it can
accurately decompose an image into directional subband outputs. The method
of multichannel analysis extracts information by passing the normalized iris
image through a bank of filters. The output of each filter provides information
about different features of the iris. For feature extraction, the image is
decomposed into eight directional subband outputs using the DFB.
Directional filter bank extracts iris discriminating features because iris has
extraordinary structure and provides many interlacing minute characteristics
such as freckles, coronas, stripes, etc which are more directional. The
proposed method incorporates directionality as a prominent feature
component and represents the iris in terms of directional subband.

Filter Bank Structure

The DFB was originally proposed by Bamberger and Smith (1990)
and later refined by Park et al (1999). This employs fan filters, sampling and
resampling matrices to extract directional information. DFB decomposition
uses a tree structure of two-band splits. The tree-structure allows an arbitrary number of directional subbands. Each level of the structure can be implemented using separable poly phase filters (Lim 1990). The DFB decomposition shares two important properties with the traditional discrete wavelet transform (DWT) namely maximally decimated and perfectly reconstructed (PR).

The DFB divides the two-dimensional (2-D) spectrum of an image into wedge-like directional subbands (Minh 2001). Figure 4.9 shows an example of the directional subband images decomposed by the eight-band DFB, where each directional component is captured in its own subband image. Directional filter bank can be implemented by 1-D separable filtering and extract 2-D directional orientation information, into $2^n$ subbands.

![Figure 4.9 Eight directional subbands](image)

The DFB basically consists of lowpass filters $H_o(\omega)$, quincunx down samplers $Q$, diamond conversion matrices $R$, modulators $e^{-j\omega_1 \pi}$ and postsampling matrices $B$ as shown in Figure 4.10. The modulator varies the spectrum of the image so that the modulated image can be divided into two directional subband images by a lowpass filter with a diamond-shaped passband. One of the attractive features of the DFB is the fact that it can be implemented by only one filter prototype (Vaidyanathan 1993), which is a
strong motivation to use unimodular matrices in a DFB. By using carefully
designed unimodular matrices, the filter design process can be reduced to
require only one prototype filter $H_0(\omega)$ as shown in Figure 4.11. Therefore, if
the unimodular matrices that change the frequency components from $R'_0(\omega)$
to $H_0(\omega)$, for $I = 1, 2, 3$ and $4$, respectively, are found, then the systems in
Figure 4.11 (a) and (b) are identical and only one filter prototype $H_0(\omega)$ is
needed. Consequently, $H_0(\omega)$ has replaced the four remaining filters
$R'_0(\omega)$ by using unimodular matrices.

Figure 4.10 Two paths in an eight-band DFB with the backsampling matrices

(a) 

(b)
A 2-D directional filter bank (DFB) be maximally decimated while achieving perfect reconstruction. The DFB is efficiently implemented via a \( l \)-level tree-structured decomposition. This leads to \( 2^l \) subbands with wedge-shaped frequency partition. The basic building block of the original DFB is the two-band filter bank using filters with diamond-shaped passbands and stopbands.

The Two band DFB: In the first stage the input image is modulated, filtered via the diamond filters (Lowpass filter \( h_0(z) \) and highpass filter \( h_1(z) \)) and downsampled resulting in two subband regions.

The Four-band DFB: In the second stage, the two subband images are further divided into four subband images using a procedure similar to that used in the first stage. A four-band DFB is composed of two-band DFBs arranged in a tree like structure. After the modulator, the constituent frequency components are shifted, resulting in a diamond-like shape. Then, via the diamond filters, \( h_0(z) \) and \( h_1(z) \) each of the four frequency regions are filtered then downsampled by a quincunx down sampler \( Q \). By cascading
another set of two-band DFBs at the ends of the first two-band DFBs, a four-band directional decomposition is achieved.

The $2^n$-band DFB: The two-band and four-band DFB’s lead to $2^n$ band extensions. To expand to eight bands, one can apply a third stage in a cascade fashion. In the third stage, resampling matrices are required to transform the parallelogram-shaped passband into one with a diamond shape. Resampling operations rearrange the samples; it does not change the data rate. Skewing (rearrange) the image in the spatial domain also affects the frequency domain in a similar manner. The spectrum is skewed by the frequency domain equivalent of the resampling matrix. The following four basic resampling matrices are used in the DFB in order to provide the equivalence of the rotation operations (Park et al 2004, Smith et al 1990).

\[
R_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad R_2 = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \\
R_3 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad R_4 = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}
\] (4.7)

These resampling matrices enable the DFB to be implemented by only a one-dimensional (1-D) filter prototype using ladder structure proposed by Ansari et al (1999). The ladder filter $\beta(z)$ is designed with the Parks-McClellan transform with filter length $L$ varies between 2 to 12. Frequency response of half band fan filter using ladder structure is shown in Figure 4.12.
Postsampling matrices $B$ are then appended to the end of the filter bank to remove the phenomenon of frequency scrambling, resulting from the frequency shift due to the nondiagonality of the overall down sampling matrix. Filter bank stages can be implemented in a separable polyphase form to achieve highly efficient realizations. All polyphase components of prototype filters used are assumed to have the same length that reduces the computation time.

**Quincunx Downsampling**

Quincunx downsampling uses quincunx resampling matrices, which are matrices whose entries are ±1 and thus having a determinant equal to 2. In the 2-D case the desired passband supports of the filters depend not only on the lattice but also on the choice of dilation matrix $M$. There are several quincunx downsampling matrices (Bamberger and Smith 1990) that generate the sampling lattice shown in Figure 4.13, but the most commonly used is (Park 2004)
\[ Q_0 = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \quad Q_1 = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \]  \tag{4.8}

For the 2-D sequence \( x(n) \), the \( M \) fold down sampled sequence \( x_d(n) \) is defined as

\[ x_d(n) = x(Mn) \]  \tag{4.9}

where \( M \) is a 2x2 nonsingular matrix of integers.

![Figure 4.13 Quincunx sampling lattice](image)

The quincunx sampling matrix \( Q \) generates the same sub-lattice but the downsampling operation rotates the input image by \(-45^\circ\) and \(45^\circ\) respectively. A quincunx downsampling corresponds to a rotated downsampling. Figure 4.14 shows the original cameraman image and its corresponding downsampled image by the quincunx matrix \( Q_0 \) and \( Q_1 \). The quincunx subsampling process has a diagonal cut-off and keeps more important features.
Figure 4.14 Original cameraman image down sampled by $Q_0$ and $Q_1$

The ladder structure is used in the DFB implementation (Vaidyanathan 1993). The ladder structure has the addition of structurally enforced perfect reconstruction. Polyphase form and the ladder structure allow the diamond filtering to be performed in the decimated domain. Directional filter bank is a decimated structure and is implemented with single prototype filter. The prototype filter is a linear phase filter. The linear phase filters in the ladder implementation also provide higher computational efficiency because they are symmetric filters and can halve the number of multipliers required.

**Directional filtering of iris image**

Directional images are obtained by applying directional filter bank constructed as described above. As given in chapter 5, experiments show that better results are achieved when applying a three level DFB design and hence eight directional subband images are obtained at the end of the DFB as shown in Figure 4.15. With these subband coefficients, other features are extracted from further analysis and can form a powerful discriminating feature vector. These directional images can be regarded as a decomposition of the original
image in eight directions. By creating directional images, noise of the original image was divided into eight different directions, thus reducing noise energy by a factor of eight (Mohammad Khan et al 2004).

Directional images contain features associated with global directions rather than local directions. For an N x N image, the size of the first half of the subband outputs is $\frac{N}{2^{n-1}} \times \frac{N}{2}$ while the size of the other half is $\frac{N}{2} \times \frac{N}{2^{n-1}}$. In the case of an 8 band DFB, for an N x N image, the size of the first half of the eight subband outputs is $\frac{N}{4} \times \frac{N}{2}$, while the size of the other half is $\frac{N}{2} \times \frac{N}{4}$. Hence, the subbands corresponding to an $m \times n$ block have size of $\frac{m}{4} \times \frac{n}{2}$ for directions 0 to 3, and size of $\frac{m}{2} \times \frac{n}{4}$ for directions 4 to 7. Directionally decomposed images contain features associated with global directions rather than local direction. On application of the above procedure, for an image of size 32 x 32, DFB decomposition yields subband images of size 8 x 16 for directions 0 to 3 and size 16 x 8 for directions 4 to 7.

![Figure 4.15 Directional subband images](image_url)
The quincunx-downsampled image has half the number of pixels that are considered in the original input image. But, due to the geometry of quincunx sampling, the actual dimension of the array needed to contain the downsampled image can be as large as the input. This results in redundant memory utilization even if the DFB is a maximally decimated filter bank. Therefore tiling method (Bamberger and Smith 1992) is utilized to prevent this problem. The tiling method is a generalized two-dimensional version of symmetric extension. The tiling method literally assumes that a given image is periodically repeated. Therefore, the quincunx-downsampled version does not have any empty entries, repeated pixels around the four corners. The quincunx downsampled image is periodic either only horizontally or only vertically.

4.4.3 Quality Measure

Current iris matching proposals (Daugman 2007, Thornton et al 2007) (feature extraction and comparison) are independent of the imaging environments and do not take into account this information in the recognition task. In the iris recognition systems, the image quality (Kalka et al 2006) is considered as a critical issue, which directly affects the overall performance of identification process.

In this section I propose a method that measures the quality of each subimage and takes into account this information in the computation of the similarity between iris images. This quality measure is based on the proportion of noise present in each subimage and its main purpose is to minimize their effect in similarity measure by assigning them a small comparison weight. Experiments led me to conclude that this method
significantly decreases the error rates in the recognition of noisy iris images, which makes it adequate for the uncontrolled environment.

The quality of iris image is the key point to affect the accuracy of iris recognition system. Poor quality images can significantly affect the accuracy of iris-recognition systems because they do not have enough feature information. The quality of iris is a combination of the distinctiveness of the iris region and the amount of iris region available.

The iris texture is so localized that the quality varies from region to region. For example, the upper iris regions are more often occluded than lower regions, and the inner regions often provide finer texture compared to the outer regions. Further, I combine the local quality into a single image quality measure and demonstrate its capability of predicting the matching performance.

Let $I(x, y)$ be the normalized iris image with coordinates of an image pixel $x$ and $y$. Let $A = \{a_1, ..., a_n\}$ be the set of subimages extracted from the image $I$. Every subimage block is decomposed into eight sub bands using directional filter bank (DFB). The energy $E_{mn}$ of the $n$th subband ($n \in \{1, 2 \ldots 8\}$) associated with the $m$th subimage block is defined as

$$E_{mn} = \frac{1}{N} \sum_{x,y} |I_{mn}(x, y) - I_{mn}|$$

($4.10$)

$I_{mn}(x, y)$ denotes the image coefficient at position $(x, y)$ of subband $n$ corresponding to subimage block $m$. $N$ is the total number of coefficients in the $n$th subband. $I_{mn}$ is the mean of pixel values of $I_{mn}(x, y)$. The energy, $E_{mn}$ is a good indicator of the distinctiveness of the iris features, and hence, a
reliable measure of local quality; high values of $E_{mn}$ indicate good quality and vice versa. I define the function $q(B_m)$ that gives the quality measure of the subimage as

$$q(B_m) = \frac{1}{M} \sum_{n=1}^{M} \log E_{mn}$$

(4.11)

This function $q(B_m)$ gives the proportion of information considered in the subimage. Every subimage has a correspondent quality measure, which is inverse proportion of noisy information involved in it. In the similarity measure the comparison is allowed for all subimages independently of their quality measure and the result is weighted by the average quality of the operands subimages. All the subimages are considered for comparison and weighted according to their respective quality value. The proposed quality measure is effective in predicting the matching performance. Matching score from individual subimage are combined together based on this quality measure.

![Figure 4.16 Regions of iris image with varying quality measures](image)

**Figure 4.16 Regions of iris image with varying quality measures**

According to the process described in section 4.3.2, I subdivide the image into eight regions as in Figure 4.16 and I compute quality measure (4.11) of each region. As shown in the Figure 4.16 the quality measure varies
depending on the noise in the image. The regions 1, 4, 5, 8 have large quality measures and the other regions have smaller measures.

This thesis proposes an energy based approach of iris image quality for iris recognition. The quality score can be used to calculate the confidence level of the recognition result.

4.5 UNDECIMATED PYRAMIDAL DECOMPOSITION

Discriminative information usually resides in high frequency regions. Although the DFB provides good directional resolution, it does not provide frequency resolution. Each subband covers the whole frequency spectrum. Since directional information resides mostly in the mid-frequency to high-frequency range the simplest approach to deal with this issue is to use the “lowpass-highpass” decomposition using a undecimated pyramidal structure before the DFB. Undecimated pyramid is obtained by manipulating the decimated Laplacian pyramid, such that low computational complexity is preserved, while achieving shift invariance.

4.5.1 Lowpass-Highpass Decomposition

The distinctive spatial characteristics of the human iris are manifest at a variety of scales. For example, distinguishing structures of the iris ranges from the overall shape of the iris to the distribution of tiny crypts and detailed texture. To capture this range of spatial detail, it is advantageous to make use of a multiscale representation. The iris-recognition systems under discussion make use of bandpass image decompositions to avail them of multiscale information.
When the image is directly decomposed with the DFB, the wedge-shaped passbands of the filters would introduce low frequency energy in the subbands. Since directional information is related to mid and high frequency information (like edges and textures), lowpass information can hinder the ability to capture these features. Moreover, the DFB is not capable of distinguishing between structures and detail at different scales. Hence, this limits the DFB ability to perform multiscale/multiresolution analysis. To deal with these issues, DFB is combined with multiscale pyramids.

The simplest approach to deal with the above mentioned issue is to use lowpass-highpass decomposition using Laplacian pyramid (Burt and Adelson 1983) as in Figure 4.17. The image is first filtered with a lowpass filter. To lower the computational complexity, filtering can be computed with a row-column separable filter. The high frequency component is obtained by a simple difference with the original image. Next, the directional analysis is performed only on the high frequency information.

Let \( f(x,y) \) be the original image and \( f_0(x,y) \) be the result of applying an appropriate low-pass filter \( H_0(\omega) \) to \( f(x,y) \) and down sampling. The prediction error \( E_1(x,y) \) is the subtraction of \( f_{11}(x,y) \) (upsampled, filtered version using \( G_0(\omega) \) of \( f_0(x,y) \) from \( f(x,y) \). The prediction error in the first level is given by

\[
E_1(x,y) = f(x,y) - f_{11}(x,y). \tag{4.12}
\]

The prediction error provides the detailed information.

The reduced image \( f_1(x,y) \) is itself low-pass filtered; down sampled to yield \( f_2(x,y) \) and a second error image \( E_2(x,y) \) is given by
\[ E_2(x, y) = f_1(x, y) - f_{22}(x, y) \] (4.13)

where \( f_{22}(x, y) \) is the upsampled lowpass version of \( f_2(x, y) \). By repeating these steps in the next level, I obtain a sequence of two-dimensional error \( E_1, E_2, E_3 \) resulting in discrete approximations of an image at the resolutions 1, 1/2, 1/4 and 1/8. The number of pixels decreases by a factor of two at each scale. It is well known that texture pattern concentrates mostly in mid-frequency band. Thus accurate feature analysis and characterization can be obtained by using a fine directional component in the mid-frequency range.

Figure 4.17 Multiresolution pyramid
4.5.2 Undecimated Pyramid

Pyramid extracts information across different resolutions. For image analysis and processing, scale information has been used in conjunction with directional information in applications where it might be necessary to distinguish objects or features of different sizes. Information across different resolutions and orientations is obtained by combining a radial pyramid and an angular filter bank. Radial decomposition of Laplacian pyramid (Burt and Adelson 1983) provides the necessary scale information as explained in the previous section. Each detail level $D_j$ of the pyramid is one quarter the size of the previous level $D_{j-1}$. In the case of the N-band DFB, the size of directional subbands reduces with number of levels of DFB so that the size of the subband is not significant. To overcome this drawback, the proposed method uses the structures in which the radial components are not downsampled or undecimated. If the same number of directional bands is used for all $J$ pyramid levels, all the subbands corresponding to the same orientation will have the same size.

An undecimated pyramid has no downsampling, and the modified system is shown in Figure 4.18. In order to eliminate the downsamplers in the pyramidal structure multirate identities are applied. One single lowpass prototype filter $L_{\frac{\pi}{\pi}}(z_0, z_1)$ that satisfies the perfect reconstruction condition is used which is modified for each resolution level as $L_{\frac{\pi}{\pi}}(z_0^{2j}, z_1^{2j})$ where $j = 0,1,...,J-1$. In the spatial domain this modification corresponds to inserting zeros in the impulse response of $L_{\frac{\pi}{\pi}}(z_0, z_1)$. For each level, all filters are upsampled by 2 in both dimensions. Hence, the undecimated pyramid eliminates the parent-child ambiguities for inter-subband processing. The filters for subsequent stages are obtained by upsampling the filters of the first
stage. This gives the multi-scale property without the need for additional filter design.

Undecimated pyramid does not decimate signals. Thus it produces more precise information for further processing. It replaces the sub-sampling of filtered image by up-sampling the low-pass filters. This up-sampling is done by inserting zeros between the filter coefficients at each level. The detail coefficients are computed as difference between low-pass images from two consecutive levels. The inverse transform is computed by adding the detail coefficients from all levels to the final low-resolution image. For the first level the detail image is obtained as the difference of the lowpass filtered image with the input image. It is a multiresolution (multiscale) decomposition of the $L^2(R^2)$ space in to a series of increasing resolution (Minh 2001).

$$L^2(R^2) = V_{j_0} \oplus \sum_{j=j_0}^{w} (W_j)$$

(4.14)

where $V_{j_0}$ is the approximation at level $J$ and multi resolution $W_j$ contains the added detail to the finer level $J-1$. By using $J$ appropriate low pass filters, $J$ low pass approximations of the image are created. The result is a multiresolution pyramid with $J+1$ equal size levels; one coarse image approximation and $J$ bandpass images. In this, multi-scale property is a shift-invariant filtering structure that achieves a subband decomposition similar to that of the Laplacian pyramid. Due to the absence of downsamplers in pyramid decomposition, the lowpass subband has no frequency aliasing. This improves the Recognition rate of the system.
Using the iris subdivision method the normalized image is divided into eight subimages, of size 32x32. Each subimage of size 32x32 is decomposed into J pyramid levels. The outputs of all the levels have the same size as 32x32. Except the lowpass channel, each of the J bandpass channels are decomposed by the maximally decimated DFB. Figure 4.19 shows the outputs of pyramidal structure at various levels along with the original image.
4.6 ROTATION INVARIANT FEATURES

One of the key performance limitations in iris recognition is the low image quality including rotated iris images. In this thesis proposes a rotation invariant feature extraction method using optimal projection analysis. Experiments have been done on the CASIA, UBIRIS image database to validate this algorithm and the results are discussed in chapter 5.

The need for robust iris recognition systems due to the variation of image size, position, and orientation still persists. Changes in position and size may be readily normalized in the pre-processing stage as they depend mainly on optical magnification and distance of the camera from the eye. One of the problems associated with iris recognition is the task of identifying a person at different orientations. Two different iris codes from the same eye are not necessarily the same orientation when acquired from the raw images.

Iris orientation depends upon a large number of internal and external factors including torsional eye rotation and head tilt. Optical systems may introduce image rotation depending on eye position, camera position, and mirror angles. Feature information in the normalized image varies in position due to image rotation. The feature information in Figure 4.21 is not same as in Figure 4.20 with respect to spatial position.
4.6.1 Rotation Compensation Techniques

For non-wall-mounted cameras, accidental misalignment during capture may lead to rotated images even when the subject is looking straight into the lens. Rotation invariance is achieved by rotating the feature vector before matching (Daughman 2004), by registering the input image with the model before feature extraction (Wildes 1997) or unwrapping the iris image to different initial angles (Li Ma et al 2002a).

Daugman’s technique (2007) and other iris classification techniques (Mayank Vatsa et al 2008) based on various intensity and frequency transforms have assumed a restricted range of rotation due to head tilt and ocular torsion. Daugman computes the iris code in a single canonical orientation and compares it with several orientations by cyclic scrolling (Daughman 1995). The use of multiple comparisons in many systems leads to higher storage requirements and increased time to enroll and verify.

In order to make the iris recognition system rotation invariant, Monro and Zhang (2007) codes every registered iris image from several initial positions around the circumference. They made seven ‘slip’ templates
from one registered iris image. An iris to be recognized will be matched to each slip template of the registered iris image and the minimum distance taken as the matching distance. Monro and Rakshit (2007) studied the effect of iris rotation through circular shifts and seen to have minimal effects on match/nonmatch scores. Du et al (2006) utilizes Global iris histograms method to extract rotation invariant features in the spatial domain, thus providing low overall computational complexity. However, FAR and FRR are worse compared to state of the art techniques. A recent approach proposed by (Ives et al 2004) uses rotation invariant 1D signatures with radial locality extracted from the spatial domain. Performance of their algorithm seems to be poor.

Vatsa (2005) proposes rotation invariant algorithm by the circular shift-based matching. The method for solving rotation invariance is similar to the cyclic scrolling method proposed by Daugman (2007). Vijaya Kumar et al (2002) investigate the performance of advanced correlation filters for such applications. (Koichi et al 2005) used phase-based image matching to achieve good results in fingerprint and iris recognition using correlation filters. Apart from a dramatic increase in storage requirements, two dimensional operations are computationally more intensive and slow down the verification/identification process significantly.

Zhenan Sun et al (2005) obtains approximate rotation invariance by unwrapping the iris ring at different initial angles. The Wildes (1997) et al system uses an image-registration technique for rotation compensation. His approach geometrically warps a newly acquired image into alignment with a selected data base image according to a mapping function. Comparing all this algorithms Daugman method is more efficient than other methods.
However the above mentioned methods may increase the complexity of the classifier or increase the computational complexity by storing codes from multiple orientations of the same image. As a consequence, rotation invariant iris features are highly desired to avoid these additional computations. In this thesis I explore rotation invariant iris features using optimal projection analysis.

4.6.2 Optimal Projection Analysis

Optimal projection analysis is used to obtain the rotation invariant iris features. Principle component analysis (PCA) (Kirby and Sirovich 1990, Cui et al 2004) is one of the popular feature extraction and data representation techniques widely used in the area of computer vision and pattern recognition. This is a straightforward image projection technique, which is developed for image feature extraction. As opposed to conventional PCA, Optimal projection analysis is based on 2D matrices rather than 1D vectors (Jain and Chengjun 2007, Liwei et al 2006, Yanwei Pang et al 2008, Jain and David 2004).

Let $X$ denote an $n$ dimensional unitary column vector. Projection of image $A$ (an $m \times n$ random matrix) onto $X$ is

$$Y = AX$$

(4.15)

where $Y$ is an $m$-dimensional projected vector, which is called the projected feature vector of image $A$. The total scatter of the projected samples is characterized by the trace of the covariance matrix of the projected feature vectors.

$$J(X) = tr(S_Y)$$

(4.16)
where $S_x$ denotes the covariance matrix of the projected feature vectors of the training samples and $tr(S_x)$ denotes the trace of $S_x$. The covariance matrix $S_x$ is given by

$$S_x = E[(A - E(A))X][(A - E(A))X]^T$$  \hspace{1cm} (4.17)$$

where

$$tr(S_x) = X^T E[(A - E(A))^T (A - E(A))] X$$  \hspace{1cm} (4.18)$$

Let me define the following Matrix

$$G_i = E[(A - E(A))^T (A - E(A))]$$  \hspace{1cm} (4.19)$$

The matrix $G_i$ is called the image covariance matrix. Suppose there are $M$ samples in total, the $j$th image is denoted by an $m \times n$ matrix $A_j$ ($j=1,2, \ldots, m$) and the average image of all is $\overline{A}$. Then $G_i$ can be written as

$$G_i = \frac{1}{M} \sum_{j=1}^{M} (A_j - \overline{A})(A_j - \overline{A})$$  \hspace{1cm} (4.20)$$

The criterion $J(x)$ can be expressed by

$$J(x) = X^T G_i X$$  \hspace{1cm} (4.21)$$

where $X$ is a unitary column vector. This criterion is called the generalized total scatter criterion. The optimal projection axis $X_{opt}$ is the unitary vector that maximizes $J(X)$ i.e the eigenvector of $G_i$ corresponding to the largest eigenvalue
\[ \{X_1, ..., X_d \} = \text{arg max } J(X) \]
\[ X_i^T X_j = 0, i \neq j, i, j = 1, ..., d \] (4.22)

where \( X_1, ..., X_d \) are the orthonormal eigenvectors of \( G \) corresponding to the first \( d \) largest eigenvalues. Using the optimal projection vectors \( X_1, ..., X_d \), the features are extracted for the iris images. After the eigen value decomposition of \( G \), the optimal projection vectors \( X_1, ..., X_d \) of \( G \) corresponding to the first \( d \) largest eigenvalues construct optimal projection axis. For a given image sample \( A \), its projection on \( X_k \) is

\[ Y_k = AX_k, \ k = 1, 2, ..., d \] (4.23)

Thus I obtain a family of projected vectors \( B = [Y_1, ..., Y_d] \) of size \( m \times d \) are called principle component vectors of \( A \).

### 4.6.3 Optimal Projection Analysis on Rows and Columns

The matrix \( B \) obtained above is based on optimal projection analysis performed on the row vectors of all the available images; the correlation information among the column vectors of the images is lost. I consider another projection scheme which can effectively remove the redundancies among both rows and columns of the images and thus lower down the number of coefficients used to represent an image (Dinh et al 2005). Also, the correlation information in both rows and columns of the images are considered and this will benefit the subsequent classification performed in the obtained subspaces.

Let \( B_k = [(B_k^{(1)})^T (B_k^{(2)})^T .... (B_k^{(m)})^T]^T \) (4.24)
\[
\overline{B} = [(\overline{B}^{(1)})^T (\overline{B}^{(2)})^T \ldots (\overline{B}^{(m)})^T]^T
\]  
(4.25)

where \(B_k^{(i)}\) and \(\overline{B}^{(i)}\) denote \(i\)-th row vectors of \(B_k\) and \(\overline{B}\). Then

\[
G = \frac{1}{M} \sum_{k=1}^{M} \sum_{i=1}^{m} (B_k^{(i)} - \overline{B}^{(i)})^T (B_k^{(i)} - \overline{B}^{(i)})
\]  
(4.26)

Similarly, the optimal projection matrix \(Z\) can be obtained by computing the eigenvectors \(Z = [Z_1, Z_2, \ldots, Z_d]\) corresponding to the \(d\) largest eigenvalues. The eigenvalues are sorted in descending order and each of them representing the variance along the principle axis determined by the corresponding eigenvectors. A further advantage of this feature vector is that its size can be easily reduced by keeping only a few largest eigenvalues, i.e. principle components. This approach can effectively reduce the computation at both training and testing phases. The optimal projection vectors \(Z = [Z_1, Z_2, \ldots, Z_d]\) are used for final feature extraction for the row correlation projected image \(B = [Y_1, \ldots, Y_d]\) of size \(m \times d\), its projection on \(Z_k\) is

\[
W_k = B^T Z_k, \quad k = 1, 2, \ldots, d
\]  
(4.27)

I obtain a family of projected vectors \(W = [W_1, W_2, \ldots, W_d]\) of size \(d \times d\) which are called the principle component vectors of \(B\).

4.6.4 Rotation Invariant Feature Extraction using Optimal Projection Analysis

I take one coefficient from each resulting DFB subband of the subimage at the same location to form a subband vector. When all coefficients
within each subband are scanned, a sequence of subband vector is generated. Let $A$ be the matrix obtained from the DFB subband vectors. Suppose there are $M$ samples in total, the $j$th image is denoted by an $m \times n$ matrix $A_j (j = 1, 2, \ldots, M)$ and the average image of all is $\overline{A}$. The covariance matrix $G_t$ can be written as

$$G_t = \frac{1}{M} \sum_{j=1}^{M} (A_j - \overline{A})(A_j - \overline{A})$$

The covariance matrix not only describes the distributions for individual subbands, but also indicates the correlation among the distributions of different subbands. Therefore the covariance matrices of different images belonging to the same class will cluster in a high dimensional space. When the original image is rotated, I expect that all the coefficients inside each subband are collectively rotated by the same angle, i.e. the subband domain image will have the same orientation as the rotated original image (Park et al 2004). However the magnitude of coefficients inside each specific subband may change. For example, if an image has a strong directional component its energy in directional subband domain will be mostly concentrated in one or two subbands corresponding to that direction. After a rotation, this energy concentration still exists, but will be in different subbands. This represents the special energy compaction property of the DFB.

The observation matrix $x^\theta$ from an image rotated at angle $\theta$ can be calculated through a linear transform of the corresponding non-rotated observation matrix $x$.

$$x^\theta = R^\theta x$$

(4.29)
where $R^\theta$ is the rotation matrix with columns of interpolation vectors. 2-D rotation operator $R^\theta$ performs a rigid rotation by an angle $\theta$. The covariance matrix $C^\theta$ from the rotated image and the covariance matrix $C$ from the non-rotated image are related as:

$$C^\theta = R^\theta C R^{\theta T} \quad (4.30)$$

If the covariance matrix with rotated image $C^\theta$ and non-rotated images $C$ are diagonalized, they will share the same eigenvalues. Therefore the eigenvalues of the covariance matrices can be used as rotation invariant features. A further advantage of this feature matrix is that its size can be easily reduced by keeping only a few largest eigen values, i.e. principle components. This approach can effectively reduce the computation at both training and testing phases.

### 4.6.5 Selection of Largest Projection Values

The selection of the number of projection values $d$ is very important in achieving good recognition rates. A large $d$ will result in a small compression rate while a small $d$ will lose some important information for classification. A further advantage of this feature matrix $W=[W_1, W_2, ..., W_d]$ is that its size can be easily reduced by keeping only a few largest eigen values, i.e. principle components. This approach can effectively reduce the computation at both training and testing phases. There are $d$ projection values obtained from the covariance matrix. The first $d$ principal components represent the salient features of the image that can be used to differentiate from other images. Therefore, it is sufficient to use these component vectors to represent the image for recognition purposes.
Figure 4.22 Magnitude of eigen values versus number of eigen values

As the number of eigen values increases, the corresponding magnitude of the eigen value decreases. Figure 4.22 shows the magnitude of eigen values \(d\) against the number of eigen values. As observed the first few eigen values contain most of the energy of the original image. As the value of \(d\) increases, the information (the energy of image) contained becomes gradually weaker as shown in Figure 4.22. Single optimal projected vector is not enough in classification of images; so I need to find a set of optimal projected vectors. As Figure 4.22 shows the magnitude of the eigenvalues quickly converges to zero and I can conclude that the energy of an image is concentrated on its first small number of component vectors.

Suppose that the \(n\) eigen-values of \(A\) are \(\lambda_1, \lambda_2, \ldots, \lambda_N\) and that they are ordered in such a way that \(|\lambda_1| > |\lambda_2| > \ldots > |\lambda_N|\). I choose the number of eigen values \(d\) to maximize the recognition rate. Figure 4.23 shows the recognition rate of the proposed system against number of eigen values. The
recognition rate is poor for less number of eigen values. Small $d$ will lose some important information for recognition. A large $d$ will result in a small compression rate. As the number of eigen values is increased beyond 8 the recognition rate remains the same because only the first few eigen values contain the maximum energy content necessary to represent an image. For eigen values greater than 8 the energy content is very small and they donot give any additional information for recognition. Therefore the recognition rate remains constant after first 8 eigen values. The value of $d$ is chosen as 8 after which there is no improvement in performance. The proposed method performs well and achieves a classification rate of 99.82%.

![Figure 4.23 Performance of the system against number of eigen values](image)

**Figure 4.23 Performance of the system against number of eigen values**

### 4.7 FUSION OF CLASSIFIERS

Iris images captured in an uncontrolled environment produce nonideal iris images with varying image quality. In most cases, the noise is localized in some subpart of the iris. The proposed method is based on the division of the normalized iris image into eight regions, followed by the
independent feature extraction on each one. The feature comparison between features extracted from correspondent regions produces eight matching scores. A new matching score is obtained by using the quality measure of the corresponding subimages. The matching scores from each classifier are fused according to a classification fusion rule based on the quality of each subimage. Since most of the iris regions are noise-free accurate recognition can be achieved, even in highly noisy images. As my experiments confirm, the proposed classification method decreases the false rejection.

4.7.1 Fusion Methods

Information from multiple sources to compensate for the limitations in performance of each individual source can be combined together using some fusion rules (Brunelli and Falavigna 1995, Kuncheva 2004, Kuncheva 2002, Husken et al 2005). The matter of how to fuse two or more sources of information is important regarding the performance of the system. The performance improvement is dependent on the degree of correlation among individual decisions. Fusion of decisions with low mutual correlation can improve the performance. The fusion can be done at the feature level, matching score level, or decision level.

Fusion at the Feature Extraction Level

The information extracted from the different sources is combined into a joint feature vector (Lam and Suen 1995, Xu et al 1992). This is then compared to an enrollment template and assigned a matching score as in a single biometric system. Enrollment template is also a joint feature vector stored in a database. Drawback of feature level fusion is that the feature vectors to be joined might be incompatible due to numerical problems.
Fusion at the Matching Score Level

It is possible to fuse the information present in the classifier of different sources. Feature vectors are created independently for each sample and then compared to the enrollment templates, which are stored separately for each biometric. Based on the feature vector and template, each subsystem computes its own matching score. These individual scores are finally combined into a total score (kittler et al 1998, Kittler and Alkoot 2003, Kuncheva 2002), which is given to the decision module. The process inside the subsystem is the same as in a single system, thus allowing the use of same algorithms for feature extraction and matching.

Fusion at the Decision Level

In decision level fusion method, a separate recognition decision is made for each classifier (Ho et al 1994). These decisions are then combined into a final decision. In decision level fusion, each subsystem performs likes a single biometric system. Fusion at the feature extraction level is for immediate data integration at the beginning of the processing, while fusion at the decision level represents late integration at the end of the process.

This thesis deals with the fusion at the matching score level. Matching score level fusion is generally preferred because it provides the best trade-off between information content and ease of fusion (Xu et al 1992). There are different ways of combining different matching scores to achieve the best decision such as, by majority vote, sum rule, multiplication rule, median rule and minimum rule.
4.7.2 Existing Fusion Algorithms in Iris Recognition

The performance of the recognition system is improved by fusion rules. Many fusion rules exist in iris recognition. Mayank Vatsa proposes 2ν-SVM (Mayank Vatsa et al 2008) to fuse the information obtained by matching the textural and topological features of the iris image. The global textural feature is extracted using the 1-D log polar Gabor transform, and the local topological feature is extracted using Euler numbers. An intelligent fusion algorithm combines the textural and topological features.

Hyun-Ae Park et al (2007) propose a new iris recognition method based on score level fusion, using two Gabor wavelet filters and SVM (support vector machine). For score level fusion, they use the typical HD (Hamming distance) produced by a Gabor filter, which can easily be applied to conventional iris recognition systems. Jong Hyun Park et al (Jong Hyun Park et al 2005) present an iris recognition system considering counterfeit attacks. The proposed system takes multi-spectral images instead of one infrared iris image. The energy of the multi-spectral images is checked and the authentication is failed if the amount of the energy is not in the proper range. Then the images are normalized and merged into a grayscale image by using a gradient-based image fusion algorithm. In the fusion process, the images considered to be from a counterfeited iris are merged into a poor-quality image which successively generates poor matching score.

An iris recognition method based on multialgorithmic fusion is proposed (Fenghua Wang et al 2007). The method combines two kinds of iris recognition algorithms: one is based on phase information and the other is based on zero-crossing representation. Two algorithms are fused at the matching score level and a novel fusion strategy based on Minimax Probability Machine (MPM) is applied to generate a fused score which is used
to make the final decision. Mehrotra et al (2007) propose an efficient iris recognition algorithm, obtained through the fusion of Haar Wavelet and Circular Mellin operator. The features for the iris pattern are extracted using Haar Wavelet and Circular Mellin operator. The features are compared using Hamming Distance method and the fusion is done at decision level using Conjunction rule.

4.7.3 Fusion Algorithm

Due to the easy accessing and processing of match scores, fusion at the matching score level is the most commonly used approach compared to the feature level or decision level fusion. The iris images captured in the uncontrolled environment produce nonideal images containing several types of noise. Division of the normalized images into several regions reduces the number of regions affected by noise. The quality of each region varies depending on the amount of noise present in that region. The Independent feature extraction and matching of each region produce separate matching scores. New matching scores are obtained based on the quality of each region. These matching scores are combined by fusion rules.

Independent feature extraction and comparison of the query subimages with template subimages form the matching scores of individual subimage. Matching score level fusion is computed to achieve the final matching score.

The query input image is divided into eight subimages after segmentation and normalization. The optimal projection image for every subimage pair is obtained after scale decomposition and directional filtering. Then the matching scores of individual pairs are computed using nearest center classifier (Li Ma et al 2003). Matching core from individual classifier
can be combined together using fusion rules. This technique leads to better discrimination capability, even for highly degraded images.

The normalized training iris image \( f(n_1, n_2) \) and testing sample \( g(n_1, n_2) \) are divided into multiple subimages (eight subimages). The size of each subimage block is \( q_1 \times q_2 \) pixels (32x32 pixels in my experiments) and have eight nonoverlapping blocks. Let \( j_i(n_1, n_2) \) and \( t_i(n_1, n_2) \) be the \( i \)th block extracted from the image, \( f(n_1, n_2) \) and \( g(n_1, n_2) \) respectively, where \( i=1,2,..,8 \). Let \( J_{ij}(n_1, n_2) \) and \( T_{ij}(n_1, n_2) \) be the outputs of undecimated pyramid where \( j=1,2,3 \). Three levels of pyramidal decomposition is considered. The detail component in the first level of the pyramid is very low and it contains low frequency components, which in turn reduce the recognition rate of the system. The higher levels of pyramid contain more detail components and it improves the recognition rate. Therefore three levels (second, third, fourth) of the pyramid are considered for the proposed method. The size of each level of the pyramid is 32x32. The output of the pyramid is decomposed by eight band DFB (three level DFB structure produce eight subbands) which produce subbands of size 16x8 and 8x16 for 1 to 4 and 5 to 8 directions respectively. Total number of subbands for each subimage is 24(3 levels x 8 subbands). The optimal projection analysis produce the rotation invariant feature image \( W_i=\left[W_{1i},W_{2i},...,W_{di}\right] \) and \( W_j=\left[W_{1j},W_{2j},...,W_{dj}\right] \) as explained in section 4.6. The size of the feature image is 8x8.

The distance between two arbitrary feature matrices \( W_i=\left[W_{1i},W_{2i},...,W_{di}\right] \) and \( W_j=\left[W_{1j},W_{2j},...,W_{dj}\right] \) using nearest center classifier is

\[
g(W_i, W_j) = \sum_{q=1}^{d} \left[ 1 - \frac{W^T_q W^j_q}{\|W^q_i\| \|W^q_j\|} \right]
\]  
(4.31)
where \( d_{i,j} = \frac{W_{i}^j W_{j}^i}{\|W_{i}\| \|W_{j}\|} \) denotes the distance between two principal components \( W_{i}^j \) and \( W_{j}^i \) of each subimage.

It has been suggested that the fusion (Kittler et al 1998) of match scores from two or more classifiers provides better performance compared to a single classifier. In general, match score fusion is performed using some statistical rules (Kuncheva 2002). Match score fusion is performed using the sum rule, the product rule, or other statistical rules.

In this I have introduced a new matching score by combining the quality measures of different subimages with their respective matching scores. The quality measure differs depending on the amount of noise present in the subimage. During fusion more weightage is given to the noise free regions and less weightage to the noisy regions. After fusion the overall matching score represents the necessary information for recognition. This improves the recognition rate of the system.

Suppose \([C_1, C_2, ..., C_8]\) are the match scores from the subimages. The fused score for sample \( V \) is denoted as \( f(V) \). The match score fusion using the sum rule is

\[
\text{Weighted-Sum (WS): } f(V) = \sum_{i=1}^{8} q(B_i) C_i
\]

\( \text{(4.32)} \)

where \( q(B_i) \) is the weight assigned to the \( i^{th} \) matcher.
The quality measure of individual subimages is used as the corresponding weight for the matching score. The quality measure of the \( m \)th subimage is given by

\[
q(B_m) = \frac{1}{M} \sum_{n=1}^{M} E_{mn}
\]  

(4.33)

where \( E_{mn} \) is the energy estimate of the subbands corresponding to the subimage as defined in section 4.4.3. In general product rule, minimum and median rules are applied as fusion rules (Kittler and Alkoot 2003).

The match score fusion algorithm using the product rule is

Weighted-product (WP): \( f(V) = \prod_{i=1}^{8} q(B_i)C_i \)  

(4.34)

The match score fusion algorithm using the minimum rule is

Weighted-minimum (WM): \( f(V) = \min(q(B_i)C_i) \)  

(4.35)

Similarly the match score fusion algorithm using the median rule is

Weighted-median (WMe): \( f(V) = \text{median}(q(B_i)C_i) \)  

(4.36)

In the verification mode, the final match score after fusion is compared with the threshold. If it is greater than the threshold, the corresponding subject is accepted otherwise rejected. The proposed threshold selections are computed via intra and inter class information gathered from the training dataset (Daugman 1994). The intra-class is a set where the distances between the images of the same individual are calculated as
The inter-class is a set where the distances between the images of an individual are measured against the images of other individuals in the training dataset as

\[ d_{ik} = \| h_{ik} - h_{ik'} \|^2 \] where \( i \in I, k^j \in k \) and \( k \neq k^j \)

where \( h_{ik} \) is the feature vector obtained with \( k \) th image of the \( i \)th individual and \( h_{jl} \) is the feature vector obtained with \( l \)th image of the \( j \)th individual

The intra class distance \( D \) and inter class distance \( P \) are calculated as

\[ D = \{ d_{ik}^{jk} / k^j, k, k^j \in K, i \in I \} \]
\[ P = \{ p_{ik}^{jk} / j \in I, j \neq i, k, l \in K \} \]

\( D_{\text{max}} \) is a measure of generalization among images for all individuals \( P_{\text{min}} \) is a measure of differences between one individual against others.

\[ D_{\text{max}} = \max(d_{ik}^{jk} / k \neq k^j, k, k^j \in K, i \in I) \]
\[ P_{\text{min}} = \min(p_{ik}^{jk} / j \in I, j \neq i, k, l \in K) \]

The classification threshold \( \theta \) is defined

\[ \theta = \frac{D_{\text{max}} + P_{\text{min}}}{2} \] (4.37)
Depending on the threshold value the subject is accepted or rejected.

4.8 SUMMARY

The captured irises resultant of non-cooperative imaging environments contains several types of noise. In this chapter I have discussed the proposed method of feature extraction for the recognition of iris images captured in an uncontrolled environment. The iris subimage method reduces the number of localized noise regions corrupting the whole iris features. The directional features of the iris image have been extracted using DFB which is a more reliable structure compared to other directional filter banks like Gabor filter and wavelet filter bank. The proposed undecimated pyramid provides the bandpass information to the DFB without decimation. Rotation invariant feature vectors have been obtained using optimal projection analysis. It avoids multiple comparisons in the matching level by cyclic scrolling and reduces the memory requirement and system complexity. A fusion algorithm combines the matching score from individual subimages depending on the quality of each subimage, which in turn reduces the effect of noise and reduces the false rejection rate. The proposed method improves the recognition rate with less number of training samples. These are validated experimentally in the next chapter.