CHAPTER 3
DATABASES, ROI EXTRACTION, ENERGY COMPACTION METHOD, SIMILARITY MEASURES AND PERFORMANCE EVALUATION PARAMETERS

Hand-based biometric identification has been performed on the three most reliable hand biometrics viz. fingerprint, palmprint and finger-knuckle print. The experiments for identification have been performed on the biometric images downloaded from authentic websites. The databases of fingerprint, palmprint and finger-knuckle print are taken from well known publicly available database websites. Each of the three biometric databases has been explained as follows:

3.1. Databases

3.1.1. Fingerprint: Fingerprint database consists of images of size 256x256 pixels and 500 dpi resolution in the bmp image format. These are low resolution images and can be captured using any of the existing scanners. These fingerprints belong to 21 persons and 8 samples of each person are taken giving a total of 168 fingerprint images. These images have a full 2-D view of the surface of the fingerprint and contain very less background. Hence there is no requirement of extraction of region of interest as most of the full image part is foreground and contains information. The fingerprint consists of ridges and valleys taking different widths, orientation and spacing as shown in Figure 3.1 a). Also the ridges may end abruptly...
or bifurcate at a point making the fingerprint unique. The texture which fingerprints offers is also unique and hence can be analyzed and the unique features extracted.

Figure 3.1. Sample images of a) Fingerprint b) Palmprint c) Finger-knuckle print

3.1.2. Palmprint: The palmprint database has been taken from Hong Kong Polytechnic University database which is publicly available database for biometric research. Hong Kong PolyU Palmprint Database [87] contains 1800 grayscale single-spectral images corresponding to 90 different palms in BMP image format. Twenty samples from each of these palms were collected in two sessions, where around 10 samples were captured in the first session and the second session, respectively. The average interval between the first and the second collection was two months. The size of the original palmprint image is 384x284 pixels. A sample of the palmprint images is shown in the Figure 3.1 b) 0 and it can be seen that the there is a considerable portion of background part in the image. Hence there is a need of extraction of region of interest (ROI) which has been explained in section 3.2.1. The palmprint ROI consists of set of unique lines at low and high resolutions. These lines form good texture which helps in extracting features of the palmprint.
3.1.3. Finger-knuckle print: The finger-knuckle print image database has also been downloaded from Hong Kong Polytechnic University website [87] which is publicly available database for biometric research. These images have been collected from 165 people (12 samples each), including 125 males and 40 females. Among them, 143 subjects were between 20 and 30 years old and the others were between 30 and 50 years old. The samples have been collected in two separate sessions. In each session, the subject was asked to provide 6 images for each of the left index finger, the left middle finger, the right index finger, and the right middle finger. Therefore, 48 images from 4 fingers were collected from each subject. The average time interval between the first and the second sessions was about 25 days. The sizes of the original finger-knuckle print images are 384x288 pixels and since these images have substantial amount of background portion which contains negligible information, we need to extract the region of interest from these images.

3.2. Extraction Of ROI

Though there exist many algorithms for ROI extraction of palmprints [88-91] and finger-knuckle prints [92, 93], here unique techniques have been developed for extraction of region of interest of palmprint [94] and finger-knuckle prints explained as follows. The extraction of ROI for each is done in a different way.
3.2.1. ROI of Palmprint

3.2.1.1. Palm-Centroid Calculation

The palmprint is binarized using dynamic thresholding of histogram and its centroid is calculated. Histogram of an image $f(x,y)$ is the probability of existence of grey-levels in the image. A gray-level histogram of an image, $f(x,y)$, which is composed of dark objects in a light background or vice-versa, is given in such a way that object and background pixels have gray levels grouped into two dominant regions. We can extract the object from the background by selecting a threshold ‘T’ that separates these regions. An object pixel $(x,y)$ is defined for which $f(x,y) > T$, otherwise, the pixel corresponds to background. Figure 3.2 shows a palmprint and its corresponding histogram.

![Palmprint and its Histogram](image)

The left dominant region of the histogram corresponds to the background pixels and the right dominant region to the Palm area. The minimum value between the peaks of the two regions is chosen
as ‘T’. This value may vary with respect to size and illumination and hence binarization of the palmprints is done using dynamic thresholding. Dynamic thresholding helps in minimizing the effects of noise, shadows and severe changes in illumination. Figure 3.3 shows the binary image.

![Figure 3.3. Binarized Palmprint](image1)

![Figure 3.4. Centroid of Palmprint](image2)

The centroid co-ordinates of the binarized palmprint \( (X_{\text{centroid}}, Y_{\text{centroid}}) \) is calculated by the equation 3.1.

\[
X_{\text{centroid}} = \frac{m_{01}}{m_{11}}, \quad Y_{\text{centroid}} = \frac{m_{10}}{m_{11}}
\]

(3.1)

where the moments \( m_{pq} \) are computed by equation (3.2), where \( p \) and \( q \) can take values of 0 and 1.

\[
m_{pq} = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} i^p j^q * f(i,j)
\]

(3.2)

The centroid of the palmprint thus obtained is shown in Figure 3.4.
3.2.1.2. Boundary Extraction

The boundary $B(A)$ of the palmprint $A$ is given as subtraction of the morphological eroded binary palmprint from the original binary palmprint as defined in equation (3.3). The erosion [95] of an image $A$ by $S$ is denoted as $A ⊕ S$.

$$B(A) = A - (A ⊕ S) \tag{3.3}$$

where $S$ is the structuring element used to erode the palmprint image given by

$$S = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

The morphological erosion technique helps in getting rid of certain stray pixels from the binary image. This gives better connected boundary as compared to the differential technique, where the boundary of the palmprint is obtained by differentiating the binary image. The extracted boundary of the palmprint is shown in Figure 3.5.

![Extracted boundary of palmprint](image)

Figure 3.5. Extracted boundary of palmprint
3.2.1.3. Detection of Valley Points

The co-ordinates of the boundary pixels are obtained by tracing the exterior boundaries of objects using eight-point connectivity. The binary image contains certain stray pixels, some of which are removed by morphological erosion. The remaining stray pixels are eliminated by selecting the largest object which is the actual boundary of the palmprint. The boundary pixels of largest object are traced from the left-most pixel. Figure 3.6 below shows the boundary of the palm extracted with the left-most pixel (Start Co-ordinate).

The Euclidian Distances (ED) between the centroid \((X_{Centroid}, Y_{Centroid})\) and the boundary pixels is computed using equation (3.4).

\[
ED(x, y) = \sqrt{(x - X_{Centroid})^2 - (y - Y_{Centroid})^2}
\]  

(3.4)

The plot of the Euclidian distance between the centroid and the boundary pixels from the start co-ordinates is shown in Figure 3.6.
The peaks (maximas) and the valleys (minimas) in Figure 3.6 are the finger-tips and the valley-points respectively in the palmprint. The valley points are closer to the centroid, and hence their ED will be minimum. To extract these points we first differentiate the ED function and note the zero-crossing points. The differential-graph of a function will reach zero value at both the minima and maxima points of the function. To differentiate between the two, we choose those zero-crossing points, where the differential graph goes from negative value to positive value at point of inflexion.

The following figure shows the differential-graph.

Figure 3.7. (a) Differential-graph of Euclidean distance function (b) Zoomed graph of some part of (a)

Here we can see that there are many zero-crossing and points of inflexions which change their sign from negative to positive. This problem arises due to the high frequency components present in ED function. These high frequency components are generated due to the uneven contour of palm. To remove these high frequencies, we degenerate the ED function into its constituent frequencies using Fourier transform and discard the high frequency components and
regenerate the same using only low frequencies. The Discrete Fourier transform of the N point Euclidean function is given by,

\[ F(k) = \sum_{x=0}^{N-1} f(x) \cdot \left( e^{-2\pi j \frac{kx}{N}} \right) \] (3.5)

where \( f(x) \) is the ED function with respect to distance \( x \) from starting point and \( F(k) \) is its Fourier transform.

The graph clearly depicts that the energy of ED function is concentrated in the lower frequency range. Thus if the ED function is reconstructed considering only low frequency components, it will retain its basic characteristics and thus the position of the valley points will not change. Hence all the high frequency components whose cumulative energy is less than some percentage threshold energy are removed. The regenerated ED function, its slope and sign of second derivative of ED function is plotted in Figure 3.9. It can be

---

Figure 3.8. Magnitude plot of Frequency distribution of Euclidean function
clearly seen that at the minimas (valleys points) of ED function there is a zero-crossing of the first derivative and the sign of its second derivative is positive. Hence the first, second and third minimas are obtained as valley points.

![Figure 3.9. Reconstructed ED function, its slope function and sign of its second derivative](image)

The valley point between the index finger and middle finger is considered as a reference point to extract the ROI. Figure 3.11 shows a palmprint and its corresponding extracted ROI.

![Figure 3.10. Detected valley Points of a palm](image)
This ROI is a centre-most square region of size 128x128 and contains good texture and line features which can be extracted for identification. Various algorithms have been applied on this ROI extracted palmprint discussed in the later part of the thesis.

3.2.2. ROI of Finger Knuckle Print

3.2.2.1. Region Growing
Region of the finger-knuckle is selected using region growing technique [95], where seed point is selected within the finger-knuckle. From the seed point growing outward a region which has less variance around the point is selected as shown in Figure 3.12. Thus most of the finger-knuckle area is selected.
3.2.2.2. Boundary Extraction

The boundary of the region selected as shown in Figure 3.13, is extracted using the boundary extraction algorithm using erosion and the bottom most point of the boundary is selected as the lowermost point for ROI.

3.2.2.3. Canny Edge Detection

In order to find the centre of x-axis, the centre region of the finger-knuckle has to be extracted which has least amount of curvature.
These curvatures can be detected using canny edge detection [96] and the convex operator can be applied to it to get the region where the lines are minimum.

![Canny edge detection on finger-knuckle print](image)

Figure 3.14. Canny edge detection on finger-knuckle print

A window is moved through the length of finger-knuckle, within which the convex lines are summed. The least value of the sum gives the desired window and the centre of that window as shown in Figure 3.15, is the center of the ROI horizontally.

![Window of width 50 pixels is moved from one end to the other](image)

Figure 3.15. Window of width 50 pixels is moved from one end to the other

Figure 3.16 shows the original finger knuckle and the extracted ROI of size 112x224. This extracted ROI contains the maximum and unique texture portion.
3.3. Energy Compaction method for Feature Extraction

The basic technique that has been applied to all transformed image in order to extract the feature vector is a unique energy compaction method. All transforms and wavelets exhibit the property of energy compaction wherein more energy is held in the prominent frequency components [67]. In most of the hand-based biometrics the lines and creases form different frequency components and since these are unique the magnitude of frequency components representing them are high i.e. they contain more energy and can be used to represent the feature vector. Those frequency components which contain very low energy are either noise or not so significant features and can be discarded to reduce the size of feature vector. The following procedure making use of energy compaction extracts the feature vector of the biometrics:

1. The transform ‘\(I\)’ of the biometric image is generated using one of the transforms or wavelets with equation (3.6), where \(T_1\) and \(T_2\) are row and column transforms of image ‘\(i\)’.

\[
I = [T_1] * [i] * [T_2]^t
\]  (3.6)
2. Energy coefficients, ‘E’ of the transformed image is generated by

\[ [E] = [I^2] \]  

(3.7)

The figure shows the energy coefficient matrices of all the transformed biometric images in the database.

![Energy matrices of the transformed database images](image)

Figure 3.17. Energy matrices of the transformed database images

3. The mean energy matrix, ‘E_M’ of all the N database images is calculated by

\[ [E_M] = \frac{1}{N} \left[ \sum_{i=1}^{N} E_i \right] \]  

(3.8)

![A mean energy matrix which is divided into 4x4 blocks](image)

Figure 3.18. A mean energy matrix which is divided into 4x4 blocks
The mean energy matrix is divided into 4x4 energy blocks and the average of the energy coefficients of each block, called mean block energy is calculated. Each block represents similar frequency components. Those blocks whose cumulative mean energy is more than a threshold energy value are selected and the rest are discarded. Generally the threshold energy is high and even with around 99% threshold energy almost 80-90% components are discarded hence giving a small feature vector size. The selected blocks form the feature vector map. Figure 3.19 shows a sample feature vector map of the selected blocks as white and the discarded ones as black.

![Figure 3.19. The selected (white) blocks of the mean block energy matrix](image)

The standard deviation of the transform coefficients of the selected blocks corresponding to the feature vector map is the feature vector of each database image. The standard deviation gives the deviation of the frequencies in a block as similar frequency components are grouped in a block.

\[ FV = \{SD_s\} \]  

(3.9)
For matching the same blocks of the energy matrix of the test image are selected and the same parameter (standard deviation) of each block represents the feature vector of the test image. These are then compared with the feature vector of each of the database images using a similarity measure as discussed in the next section.

### 3.4. Similarity Measures

Similarity or distance measure is a function which determines how closely a biometric data sample resembles the template data. A good similarity measure takes into account statistical characteristics of the data distribution. The distance score should be very small for two feature vectors belonging to the same individual. At the same time it needs to be large enough for feature vectors coming from different individuals. The Minkowski distance is a generalized distance parameter. Given two vectors $P$ and $Q$, both of same lengths

$$P = (x_1, x_2, \ldots \ldots, x_n) \text{ and } Q = (y_1, y_2, \ldots \ldots, y_n)$$

this distance of order $p$ between them is given by

$$d_{pq} = \left(\sum_{j=1}^{n}|x_j - y_j|^p\right)^{1/p} \quad (3.10)$$

The most common distance measures which are all based on Minkowski distance with different orders are:

1. Euclidean Distance
2. Manhattan Distance
3.4.1. **Euclidean distance**: One of the most popular similarity distance functions is the Euclidean distance. It is given by the sum of squared distances of two vectors. For the same vectors P and Q, Euclidean distance, $d_E$ is given by

$$d_E = \sqrt{\left(\sum_{j=1}^{n}(x_j - y_j)^2\right)}$$

(3.11)

3.4.2. **Absolute distance**: The absolute distance between two points in space with fixed Cartesian co-ordinate system is the sum of the lengths of projections of the line segment between the points onto the co-ordinate axes. In other words, absolute distance is the sum of the absolute differences of the two vector $p$ and $q$ with values $x_j$ and $y_j$ respectively, is given by,

$$d_A = \sum_{j=1}^{n}|x_j - y_j|$$

(3.12)

There exist many other similarity measures like Mahalanobis and weighted Euclidean distance etc, but in the experiments performed for hand-based identification in this thesis the distance measure used is absolute distance. **All the experiments were also performed using Euclidean distance as similarity measure also, but the results obtained with absolute distance were far better than with those using Euclidean distance, hence they have not been included.**
3.5. Decision for Matching

The decision whether the query image matches with any of the enrolled image in the database depends on the value of similarity measure. Consider the vector \( P \) as the feature vector generated by query image after a feature extraction algorithm has been applied on it and vector \( Q_i \) as the feature vector of the \( i^{th} \) image in the database generated using the same feature extraction algorithm. If

\[
P = (x_1, x_2, \ldots, x_n) \quad \text{and} \quad Q_i = (y_{i_1}, y_{i_2}, \ldots, y_{i_n})
\]

Then, \( d_{Ai} \), the absolute distance between query image and the \( i^{th} \) database image is given by,

\[
d_{Ai} = \sum_{j=1}^{n} |x_i - y_{ij}|
\]  
(3.13)

There are two ways in which identification can be performed, closed set and open set. In closed set identification, only the enrolled users try to identify themselves. These are called genuine users and the match is simply determined on the basis of the first match obtained. If the first match belongs to the same person, decision is that of match or else non-match.

So the decision in closed set is simply that person whose sample has generated least distance with respect to the query image, i.e. if \( \min(d_{Ai}) \) is obtained with \( i^{th} \) sample which belongs to the same person who has given query image, then it is a match, else a non-match.

In open set, persons other than those enrolled with feature vectors are present in the database also try to seek identification. These users are called imposters and here the decision has to be taken based on some threshold value, \( \theta \). Hence, if
\[ \min(d_{Ai}) < \theta, \text{accept} \]
\[ > \theta, \text{reject} \quad (3.14) \]

For a query image, if the value of similarity measure i.e. minimum absolute distance between the features of the query and a database image is less than a threshold value then there is a match and it indicates that the features of the query image are similar to the matched image in the database. The value of threshold can be varied; at low values of threshold there would be lesser matches and at higher value more matches.

### 3.6. Performance Evaluation Parameters

The performance of any biometric system has to be evaluated and is done with the help of certain parameters [3]. Any biometric system should be able to identify the genuine users and reject the imposters. The parameters for closed set system are different from those of open set and are:

**3.6.1. For Closed Set:** For closed set the efficiency \( \eta \), of the system is evaluated which is given by the percentage of number of correct matches with respect to the total number of query images. The higher the value of efficiency the better is the system.

\[ \eta = \frac{\text{number of matches}}{\text{number of query images}} \times 100 \quad (3.15) \]
3.6.2. For Open Set: There are two types of errors which can be generated in open sets: a) when the genuine person gets rejected b) when the imposter is accepted. In the first case, False Rejection Rate (FRR) is evaluated, which is given by the percentage of genuine users which are falsely rejected. Another parameter which is related to FRR is Genuine Acceptance Rate (GAR).

\[
FRR = \frac{\text{number of genuine users rejected}}{\text{number of genuine users query images}} \times 100 \quad (3.16)
\]

and GAR is given by,

\[
GAR = 100 - FRR \quad (3.17)
\]

In the second case, a parameter which evaluates the number of imposters which are falsely accepted is specified. This is called False Acceptance Rate (FAR) and is defined as the percentage of imposters accepted for a given number of imposter query images. Thus,

\[
FAR = \frac{\text{number of imposters accepted}}{\text{number of imposters/ query images}} \times 100 \quad (3.18)
\]

and Genuine Rejection Rate (GRR), which indicates the numbers of imposters which are rightly rejected, is given by,

\[
GRR = 100 - FAR \quad (3.19)
\]

Both FAR and FRR are errors and should be as low as possible. When these two errors are evaluated by varying threshold from a small value to a large value, a receiver operating curve (ROC) can be obtained by plotting the two against threshold. Figure 3.20 shows a sample ROC.
ROC can also be obtained by plotting FRR vs. FAR, each for various values of threshold.

As can be seen from the Figure 3.20 the point where FRR curve meets FAR curve, is called Equal Error Rate (EER). It is the value of FAR or FRR at that point where both the errors are equal. Though EER of the system should be as low as possible, it does not suggest the operating point of a biometric system.

![Figure 3.20. ROC of a typical biometric system indicating FAR and FRR vs. Threshold, EER and SPI](image)

Another parameter can be defined for open set system, which is newly used in this thesis is Security Parameter Index (SPI). This is defined by the ratio of threshold values at two points $T_1$ and $T_2$ in the figure. $T_1$ is the maximum value of threshold with FAR as zero and $T_2$ is the threshold value at EER. Hence SPI is given by,
\[ SPI = \frac{\text{Maximum value of threshold where FAR is zero}}{\text{Value of threshold at EER}} \times 100 \] (3.20)

If there are no errors, EER will be zero and SPI will be 100%. It would symbolize an ideal system.