CHAPTER 5

INTEGRATION OF FACE AND FINGER KNUCKLE PRINT RECOGNITION SYSTEM

5.1 GENERAL

This chapter discusses the novel way of integrating face and finger knuckle print as Multimodal Biometric Authentication System (MMBAS) using Scale Invariant Feature Transforms (SIFT). A biometric system which is based only on a single biometric identifier in making a personal identification is often not able to meet the desired performance requirements. Multimodal biometrics is an emerging field of biometric technology, where more than one biometric trait to improve the combined performance. The proposed method designs a multimodal biometric system which integrates Face and FKP. The objective of applying SIFT on MMBAS is to reduce the computational complexity, to overcome the inference of those samples that may cause error in recognition and to increase the accuracy of recognition. Also it explains the research contribution with parameter analysis and comparison on experimental results in MMBAS using SIFT.

5.2 MOTIVATION TO APPLY MMBAS USING SIFT

The local feature extraction methods like Gabor Filter Based Competitive Code (Compcode), Combined Orientation and Magnitude Information (Imcompcode and Magcode) (Morales et al. 2011), Speeded Up Robust Feature (SURF) (Herbert Bay et al. 2008) and Scale Invariant Feature Transform (SIFT) are providing meticulous information from specific local region and is suitable for better representation (Lin Zhang et al. 2011 and David G. Lowe, 2004). Among these, SIFT is a more popular method which can result in an innovative way to authorize a person’s identity in a Multimodal Biometric Authentication System (MMBAS) for a specific problem in addition to the following advantages,

- SIFT transforms image data into scale-invariant coordinates relative to local features and perform better matching
- It corrects the non-uniform brightness and improves the texture of the image

- SIFT is extremely acquiescent to hybridize the biometrics to authorize a person’s uniqueness

In the proposed MMBAS the combination of face and finger knuckle recognition is the basic idea to use SIFT local feature extraction to decrease the size of data by choosing an image template from the database. In last decade, many researches show that local features are more effective to describe the detailed and stable information of an image. Among these local features, Scale Invariant Feature Transform (SIFT) (David G. Lowe 2004) became popular in face recognition. SIFT is capable to capture the main grey level features of an object’s view by means of local patterns (Manuele Bicego et al. 2006) extracted from a scale-space decomposition of an image. Luo et al. (2007) showed the good representation ability of SIFT features by combining the person specific SIFT features and simple matching strategy. So far, researchers have made a lot of research on SIFT, including some improvements. This motivates us to combine the facial features along with finger knuckle to improve the security and decrease the size of data at a low computational cost, by choosing some candidates as image templates and combine the Euclidian distance (Yu-Yao Wang et al. 2013) with the local feature (SIFT) for face and finger knuckle recognition as Multimodal Biometric Authentication System (MMBAS).

### 5.3 SYSTEM ARCHITECTURE OF MMBAS

Multimodal biometrics refers to the use of a combination of two or more biometric modalities in a verification or identification system. They address the problem of non-universality, since multiple traits ensure sufficient population coverage (Ross and Jain 2003). Multimodal biometrics also addresses the problem of spoofing. As it concerns with multiple traits or modalities, it would be very difficult for an imposter to spoof or attack multiple traits of genuine user simultaneously. The general recognition process was shown in the Figure.5.1, which consists of two phase, i.e., enrolment phase and verification phase.
5.3.1 Enrolment Phase

In enrolment phase, biometric traits of a user are captured and these are stored in the system database as a template for that user and which is further used for authentication phase.

5.3.2 Authentication Phase

In authentication phase, once again a trait of a user is captured and system uses this to either identify or verify a person. Identification is one to many matching which involves comparing captured data with templates corresponding to all users in database while verification is one to one matching which involves comparing captured data with template of claimed identity only (Golfarelli et al. 1997).

Figure 5.1 Illustration of general biometric recognition process

Unimodal biometric systems have to contend with a variety of problems such as noisy data, intra-class variations, restricted degrees of freedom, non-universality, spoof attacks, and unacceptable error rates. Some of these limitations can be addressed by deploying multimodal biometric systems that integrate the evidence presented by multiple sources of information (Ross and Jain 2004).

The research on multimodal started, and different multimodal biometrics has been developed with combination of various traits, that is, face and finger print, face and iris, iris and finger print etc. The most commonly used biometrics is face, that is, either as a single trait or combined with other trait as multimodal biometrics.
5.4 INTEGRATION OF FACE AND FINGER KNUCKLE PRINT IN MMBAS

The proposed method discusses the combination of face and FKP by extracting features from face and FKP using Scale Invariant Feature Transform (SIFT). The key points are derived from face and FKP. The combinations of face, finger print, palm print, iris recognition are readily available. But there is no possibility of having the combination of FKP and Face. This method combines the facial features along with FKP to improve the security as shown in the Figure 5.2. Results are performed on the Indian face (IITK) and Poly-U database to check the proposed face and FKP recognition method respectively.

![Figure 5.2 Illustration of face and finger knuckle print images](image)

Face recognition is a nonintrusive method, and facial images are probably the most common biometric characteristic used by humans to make a personal recognition. David G. Lowe presented a method for extracting distinctive invariant features from images that can be used to perform reliable matching between different views of an object or scene. For face authentication to determine the real potential and applicability of the Scale invariant feature transform method, different matching schemes are proposed and tested using the Indian face database (IITK) and protocol, showing better result. FKP verification was done by applying Scale invariant feature transform which shows better results than minutiae points and ridge patterns of finger print (Adithya Nigam 2011).
Due to the wide application of biometric application such as face detection and fingerprint recognition in authentication process, still many unresolved problems are left in the system as the system deals with rigid and complex data. The accuracy of the result plays a vital role to characterize the authentication system as it vary based on the algorithm deployed in the system. The most intrinsic features from the image extracted using SIFT and compared with the nearest score credentials to get the exact match. The close encounter of the distance measures evaluated with the key point descriptor of face and finger knuckle image recognition.

5.4.1 Generation of class ID in MMBAS using Scale Invariant Feature Transform (SIFT)

The proposed method split the authentication process into two phases starting with face detection using distance comparison measures followed by finger knuckle recognition phase. The data base chosen for this work uses the combination of face and figure knuckle of a person with the data size of about 100 users. The SIFT features makes the similarity analysis of face and finger knuckle combined together to authenticate a person. The key point detector of SIFT combines the detector of face and finger knuckle and make fixed detector points in the training image data sets. The different Gaussian filters are applied on the various key point detectors with difference in scales.

The following are the steps involved in proposed algorithm for evaluating the MMBAS in proposed method. Before starting the recognition process the following are ensured.

Step 1: Select the valid class id of face recognition process and knuckle image from the data sets to implement the SIFT process

Step 2: Once the process is done it is stored in the valid class id

Step 3: Check the SIFT scores that match with the valid class id
Step 4: If score matches store it in the class id or else try another set unless better match is found

Step 5: The process is done until the SIFT matches with the face and finger knuckle recognition

The likelihood multimodal biometric human recognition system was developed specifically for the verification scenario where the goal is to decide whether an input sample belongs to the genuine or impostor class.

The local image gradients are measured at the selected scale in the region around each keypoint. These are transformed into a representation that allows for significant levels of local shape distortion and change in illumination. This approach has been named the Scale Invariant Feature Transform (SIFT), as it transforms image data into scale invariant coordinates relative to local features. An important aspect of this approach is that it generates large numbers of features that densely cover the image over the full range of scales and locations.

Figure 5.3 Illustration of angle based image posture of face and finger knuckle

The SIFT features represent a compact representation of the local grey level structure, invariant to image scaling, translation, and rotation, and partially invariant to illumination changes and affine or 3D projections. SIFT has emerged as a very powerful image descriptor and its employment for face analysis and recognition was
systematically investigated in (Manuele Bicego et al. 2006) where the matching was performed using three techniques: (a) minimum pair distance, (b) matching eyes and mouth, and (c) matching on a regular grid. The present system considers spatial, orientation and keypoint descriptor information of each extracted SIFT point. Thus the input to the present system is the face image and the output is the set of extracted SIFT features \( s = (s_1, s_2, \ldots, s_m) \) where each feature point \( s_i = (x, y, \theta, k) \) consist of the \((x, y)\) spatial location, the local orientation \(\theta\) and \(k\) is the key descriptor of size \(1 \times 128\).

The angle \(\theta\) defines the SIFT boundary and by drawing the tangent line that cuts the portion into two halves by splitting the recognition boundary. The boundary set within the curves is the targeted zone and generally it assists for recognition of the image where the recognition zone falls within the space followed by the SIFT zone that is compared along with the class is based on the spilled location. It sets the SIFT region for face and finger knuckle recognition. The observation made after getting the sample results, states that the related class id falls on the same SIFT comparison zone on both occasions when dealing with the results in terms of face and finger knuckle recognition.

The output of face recognition along with the valid class id is stored in the database before starting the process of finger knuckle recognition. Once after the completion of finger knuckle process the class id is stored by comparing with the existing class id. The value of class id is created based on SIFT index that provide the exact match between two face images in terms of face recognition and followed by finger knuckle comparison. The proposed method makes use of Euclidian distance to choose a certain number of candidates, and apply SIFT features extracted from those candidates to classify the test sample. Large number of features can be extracted from typical images and increase the storage space with our proposed algorithm. Hence development of an efficient authentication system based on hybrid features becomes great demand in most of the real time authentication based applications.
5.4.2 Normalization

The fusion framework proposed in the presented multimodal biometric human recognition system was developed specifically for the verification scenario where the goal is to decide whether an input sample belongs to the genuine or impostor class. In verification, the biometric query is compared only to the template of the claimed identity, resulting in a single match score for each matcher. However, in an identification system, the biometric query is compared with all templates in the database resulting in N match scores for each matcher, where, N is the number of persons enrolled in the database. Scores generated from individual biometric traits are combined at matching score level using sum rule. \( MS_{\text{face}} \) and \( MS_{\text{fkp}} \) are the matching scores generated by face and FKP respectively. Since, the matching scores output by the two traits are heterogeneous because they are not on the same numerical range. The normalization of these scores is done.

\[
N_{\text{face}} = \frac{MS_{\text{face}} - \min_{\text{face}}}{\max_{\text{face}} - \min_{\text{face}}} \quad (5.1)
\]

\[
N_{\text{fkp}} = \frac{MS_{\text{fkp}} - \min_{\text{fkp}}}{\max_{\text{fkp}} - \min_{\text{fkp}}} \quad (5.2)
\]

The proposed method uses min – max normalization technique that can be time consuming. Selecting inappropriate normalization technique can lead to very poor recognition performance.

Min-Max normalization transforms all the scores into a common range \([0, 1]\) (Ross, A. et. al. 2005). This normalization is sensitive to outliers, but maintains the original distribution. \( N_{\text{face}} \) and \( N_{\text{fkp}} \) are the normalized matching scores of face and finger knuckle print respectively. Prior to combining the normalized scores, it is necessary that the two normalized scores are transformed as either similarity or dissimilarity measure. In the proposed method, the normalized scores of face and fkp are converted to similarity measure by subtracting them from as given below:
Finally, the two normalized similarity scores $N_{face}'$ and $N_{fkp}'$ are fused using weighted sum rule to generate final matching score as follows:

\[ MS_{final} = X \times N_{face}' + Y \times N_{fkp}' \]  

(5.5)

Here $X$ and $Y$ are the weightage assigned for face and finger knuckle print images. Weightage of images are varying from each other.

### 5.5 RESULTS AND DISCUSSION

This section discuss about the performance of proposed method with existing method. Proposed method implemented in MATLAB, face and FKP recognition was performed using the set of face and FKP images available in IITK and PolyU Database respectively. This is briefly discussed in Annexure - I.

Feature detection and matching are important steps in the design of a biometric based authentication system. The process output of face recognition along with the valid class id is stored in the data base before starting the process of finger knuckle recognition. Once after the completion of finger knuckle process, the class id is stored by comparing the existing class id. The value of class id is created based on SIFT index that provide the exact match between two face images in terms of face recognition and followed by finger knuckle comparison.

#### 5.5.1 Face and Finger Knuckle Mapping

While selecting an image from finger database the resultant would be the class id of face recognition that was already stored in the face image data base. Figure 5.4 shows the resultant nearest class id with a distance from face space. Figure 5.5 shows the feature extraction of finger knuckle and recognition process of MMBAS.
Figure 5.4 Illustration of resultant nearest class id with a distance from face space

Figure 5.5 Illustration of feature extraction of finger knuckle and recognition process of MMBAS

From the illustrated example, when finger knuckle image of class id 6 is selected it will provide the exact face recognition image whose class id is 1. The
The proposed method consists of various parameters namely translation, rotation, scaling and other image parameters, but using SIFT algorithm, only key points information is taken out from image. For image matching and recognition, SIFT features are first extracted from a set of reference images and stored in a database. A new image is matched by individually comparing each feature from the new image to this previous database and finding candidate matching features based on their feature vectors. SIFT is a proven candidate algorithm for machine vision related applications in which, generally the detected feature vectors of one scene will be compared with some set of trained feature vectors to recognize different objects in that scene. The comparison of result was done by combining the key description of face and finger knuckle, and further subjected to SIFT recognition process of face and finger knuckle recognition.

Class id comprises of both face and finger knuckle images. The SIFT library includes a simple image matching module which reads in two sets of keypoint descriptors from the scene image and the object image and measures the similarity analysis of two features by computing the Euclidian distance between two feature vectors. This stage attempts to eliminate these unstable keypoints from the final list of keypoints by finding those that have low contrast on an edge. This may be achieved by calculating the Laplacian value for each keypoint found in stage one (David G. Lowe 2004). The location of extremum, z, where D and its derivatives are evaluated at the sample point and x is the offset from this point. The location of the extremum can be determined by taking the derivative of partial differential equation \( \partial \) with respect to x and setting it to zero is given by equation (5.6).

\[
z = - \frac{\partial D^{-1}}{\partial x^2} \frac{\partial D}{\partial x}
\]  

(5.6)

To set the location range the key point and its corresponding neighbour pixel value is calculated using complete data set in terms of location, scale and proportion of primary curvatures. The low contrast points are discarded that have low contrasts (sensitive to noise) which are poorly localized along an edge.
The adjusted sample is the best retrieved image of close neighbour using SIFT score justifying best match. The SIFT value is compared with the nearest image SIFT score for making the comparison process until the match found. The proposed method is used to assign a consistent orientation to the keypoints based on local image properties. The keypoint descriptor can then be represented relative to this orientation, achieving invariance to rotation. The image is simply chosen if it is bigger compared with its neighbours of smaller size. This test occupies less constraints by the following first few checks using the equation (5.2) and (5.3). The gradient magnitude, \( m \), and orientation, \( \mu \), of \((x, y)\) are given in equations (5.7) and (5.8) below.

\[
m(x,y) = \left( L(x + 1, y) - L(x - 1, y) \right) \uparrow 2 + \left( L(x, y + 1) - L(x, y - 1) \right) \uparrow 2 \tag{5.7}
\]

\[
\mu(x,y) = \tan^{-1} \left( \frac{(L(x, y + 1) - L(x, y - 1))/(L(x + 1, y) - L(x - 1, y))} \right) \tag{5.8}
\]

The invariance of the image accessed using input point descriptor relates to its orientation based on image. The invariance achieved to produce the results that are different as change in image direction (rotational). The downside of this approach will get rid of the images, which are outfit using invariance consistency. The input key points are also balanced with location gradient of the particular images (David G. Lowe 2004). The operations on the image data is done with direction, degree and its position based on its attribute with invariance transformations.

The previous operation has assigned an image position, size, and direction to every key point. These parameters require resident image invariance of two dimensional coordinates that describes an image area followed by resident image area with its description. The area that is extremely distinctive yet is as invariant as possible to continue variations, such as transform in illumination or 3D viewpoint. The resident image grades are measured using scaling features that are around the region of every point focused towards it. The illustrated Figure 5.6 shows the gradient measures of all level.
The image gradient and its orientation measures are created using key point descriptor (Zhang et al. 2010). The local image gradients are measured at the selected scale in the region around each keypoint. These are transformed into a representation that allows for significant levels of local shape distortion and change in illumination.

The comparison of face and finger knuckle are independent to start with and concluded the result by combining both processes. Figure 5.7 and 5.8 show the feature, direction and neighbourhood comparison using SIFT.
The local gradient data, used above, is also used to create keypoint descriptors. The feature descriptor is computed as a set of orientation histograms on 4 x 4 pixel neighbourhoods. The orientation histograms are relative to the keypoint orientation and the orientation data comes from the Gaussian image closest in scale to the key point’s scale. Histograms contain 8 bins each, and each descriptor contains a 4x4 array of 16 histograms around the keypoint. This leads to a SIFT feature vector with (4x 4x 8 = 128 elements). This vector is normalized to enhance invariance to changes in illumination.

Once match is found it will be stored in the data base with a valid class id as mentioned in the beginning of this chapter. While applying SIFT the condition is checked for nearest or the exact match and the steps to be taken during the course of the result. To be able to recognize a person by biometric characteristics and the derived biometric features, first a learning phase must take place. The procedure is called enrolment and comprehends the creation of an enrolment data record of the biometric data subject (the person to be enrolled) and to store it in a biometric enrolment database. The enrolment data record comprises one or multiple biometric references and arbitrary non-biometric data such as a name or a personnel number. For example, if the system takes 165 enrolled images in 6 different postures then total number of images can be stored in database is 165*6 = 990 images. From 990 images the proposed SIFT system can extract 45 features. Along with SIFT, another algorithm called Speeded Up Robust Features (SURF) is reliable feature extraction technique that can be used nowadays in biometric authentication system.

The proposed MMBAS is compared with other existing SIFT and SURF multimodal systems which are discussed in the following subsection 5.6. and it shows the feature extraction performance of SIFT and SURF with respect to different number of images. This operation is usually done during enrolment of the users into the database.
5.6 COMPARATIVE ANALYSIS OF PROPOSED METHOD WITH EXISTING METHOD

The average time is reduced as the image data sets grow in terms of number and by using SIFT features average time is reduced as stated in the data sets.

The different data sets have been used in enrolment process and it was compared to estimate the processing cost. Table 5.1 shows that the proposed system provides high level of feature extraction performance using SIFT compared to other existing methods which uses both SIFT and SURF. Though the proposed system consumes time it reduces the cost in enrolment process.

Table 5.1 Comparison of proposed and existing method for feature extraction with respect to different number of images based on SIFT and SURF

<table>
<thead>
<tr>
<th>System</th>
<th>Method</th>
<th>Database Used</th>
<th>Number of images</th>
<th>Extracted Features (in nos.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hassan Soliman et al. (2009)</td>
<td>Feature Level Fusion of Palm Veins and Signature</td>
<td>Local database</td>
<td>Signature 37<em>10(370) Palm Vein 37</em>5(185)</td>
<td>92 33</td>
</tr>
<tr>
<td>Muhammad Imran Razzak et al. (2011)</td>
<td>Fusion of low resolution face and finger vein</td>
<td>Local database</td>
<td>35*6(210)</td>
<td>45 61</td>
</tr>
<tr>
<td>Proposed Method</td>
<td>MMBAS face and finger knuckle</td>
<td>Face database (IIT-Kanpur) and FKP of Hong Kong Polytechnic University</td>
<td>Face 11<em>40 (440) FKP 165</em>48 (7920)</td>
<td>364 122</td>
</tr>
</tbody>
</table>
After enrolment process based on Table 5.1 the enrolled information is taken for feature matching to identify the person in authentication process. In authentication mode, the system will verify a personal identity by comparing a submitted biometric sample with the biometric reference template of a user whose identity is being claimed. In this system, the user needs to claim an identity and the system will conduct a one-to-one comparison to determine whether the claimed identity matches with the user template. In the following Table 5.2, proposed method presents the authentication / feature matching performance of SIFT feature descriptors with respect to different number of images / person.

Table 5.2 Comparison of time taken for matching in authentication process

<table>
<thead>
<tr>
<th>Image Template Per Person</th>
<th>Total Training Images</th>
<th>Total Test Images</th>
<th>Time Taken For Matching (in seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>SIFT</td>
</tr>
<tr>
<td>2</td>
<td>370</td>
<td>185</td>
<td>76.6</td>
</tr>
<tr>
<td>4</td>
<td>740</td>
<td>185</td>
<td>152.5</td>
</tr>
<tr>
<td>6</td>
<td>1110</td>
<td>185</td>
<td>230.2</td>
</tr>
<tr>
<td>8</td>
<td>1480</td>
<td>185</td>
<td>307.7</td>
</tr>
</tbody>
</table>

If the proposed method uses 8 images per person for enrolment, then for searching database of 185 entries to find a match of input image it is taking around 1.6 second (307.7/185) for SIFT based MMBAS. It means, for authenticating one person by matching his original face and FKP image features with a template of 185x8 size feature database, approximately 1.6 seconds are needed. So obviously, this 1.6 second will grow rapidly if we increase the enrolment in the database. For example, if the proposed system needs to enroll 1000 sets of original face and FKP feature descriptors (1000 x 8) of persons, then it may take around 10 seconds to authenticate a person using that, huge template set. The processing cost for authentication process is compared using different number of training images and it is shown in Figure 5.9.
In general, feature detection and matching are important steps in the design of a MMBAS or any biometric based authentication system. Scale Invariant Feature Transform (SIFT) is the most reliable feature extraction technique that was used in recent works on MMBAS. The feature descriptors detected by SIFT claimed to be capable of distinguish each and every image in the dataset from one another. But certainly, there are some costs involved in it. The storage and computational cost involved in any biometrics based identification system will mostly depend on the feature detection, description and matching techniques used in the system. In SIFT based MMBAS, the storage and computational cost will directly depend on the size of the feature descriptors used. Because, each and every time the matching process will directly match these feature descriptors to find an exact match and the descriptors so that these descriptors were directly stored in storage media as templates. So we need to store all the feature descriptors of the enrolled images for future use. Even, the size of these feature descriptors data will be greater than the original image dataset and the performance of the system will rapidly decrease with respect to the increase in enrolment in the database.

**Figure 5.9 Time taken for Feature extraction matching in authentication process**
Table 5.3 Storage space consumed by the SIFT and SURF based feature descriptor

<table>
<thead>
<tr>
<th>Number of Images</th>
<th>Database Storage Size on Disk (Mega Bytes)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SIFT</td>
</tr>
<tr>
<td>850</td>
<td>10.9</td>
</tr>
<tr>
<td>1780</td>
<td>22</td>
</tr>
<tr>
<td>3560</td>
<td>43.8</td>
</tr>
<tr>
<td>7320</td>
<td>87.5</td>
</tr>
</tbody>
</table>

The above table shows the storage space consumed by the SIFT and SURF based descriptor feature data sets. In some previous works claimed that SURF was faster than SIFT. The column SIFT +SURF is just the addition of SIFT and SURF values to show how the system will perform if we combine the two features.

Even though feature vector and the number of detected points of SIFT and SURF are approximately same, SURF consumed much storage space than SIFT. This is because, SURF is using ‘signed single precision number’ to represent fractional values in its feature vector but SIFT only uses ‘8bit unsigned integers’. Since the basic data type used in SIFT is smaller than the data type used in SURF. So, the data type used to represent the feature vector values much effects the storage size in the case of SURF. For example, for storing the feature vectors of 7320 images, it almost consumed around 160MB. This size is very high; even the original images itself consumed around 36 Mb (5kb * 7320). In fact, both SIFT and SURF features consumed much space than the original image size and SURF consumed double the space of SIFT. The comparison of storage space in SIFT feature descriptors with respect to different number of images and images/person stored in the database is shown in Figure 5.10.
Figure 5.10 Comparison of storage space for various image sets based on SIFT and SURF

The processing speed of the proposed system calculated by the execution time of each process taken by the system for feature extraction and matching is given in Table 5.4.

Table 5.4 Execution time for enrolment and authentication

<table>
<thead>
<tr>
<th>Process</th>
<th>Time (in ms.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SIFT</td>
</tr>
<tr>
<td>Feature Extraction</td>
<td>56.073</td>
</tr>
<tr>
<td>Matching</td>
<td>4.321</td>
</tr>
<tr>
<td>Total</td>
<td>60.394</td>
</tr>
</tbody>
</table>

The time taken by the proposed system for feature extraction and matching has been implemented on a Intel-Core i3-2330M processor (2.20 GHz) with 2 GB RAM. The feature extraction time is the time taken to extract SIFT and SURF
features on entire database and dividing by the total number of images. It is found that the system takes 56.073ms and 18.701ms for extracting SIFT and SURF features respectively. For average matching time, all possible matches that include both genuine and imposter cases are considered. It is observed that the system takes 4.321ms and 0.075ms for matching SIFT and SURF feature vectors respectively. Total time taken by the system is 60.394ms, out of which matching takes 4.321ms only and is significantly fast.

The proposed design of a MMBAS, using SIFT is better than SURF. From the above Figure 5.14 it is understood that the SURF consumed much storage space than SIFT. From these two efficient algorithms it can be concluded that the SURF based system consumed much storage space than SIFT. So, SIFT is one of the best feature extraction and matching algorithm that can be implemented in any biometric based authentication system. The proposed method is implemented using SIFT, and feature vector of SIFT values are computed from the local image region around the keypoint as Keypoint descriptor.

For example, if we use 7320 images in enrolment / authentication process, the storage space needed on disk in a database is 87.5 megabytes using SIFT, but using SURF the storage space needed on disk in a database is 163.7 megabytes, which is shown in Table 5.3. Using SIFT feature comparison while increasing per person score the time taken is 413.5 unit for 16 person which is half the score achieved when using either face or finger knuckle image data sets. The proposed system tries to reduce the amount of storage space compared to other existing algorithms. So far, there is no other algorithm that has tried to reduce the storage space in any other multimodal biometric based authentication system.

The result projects the comparison of true vs. false positive rate, at various threshold settings. Accuracy can also be measured in terms of Area Under ROC Curve (AUC). The proposed system considered Equal Error Rate (EER) as the main quality metric for measuring the utility of the system because, often, most of the previous reference works used this metric.
Table 5.5 presents the comparison of proposed MMBAS and existing multimodal system which uses Two and Three biometric Traits (Shekhar Karanwal 2013), Speech and Signature (Dapinder Kaur et al. 2013), Face and Finger Vein (Muhammad Imran Razzak et al. 2011), Finger Print and FKP (Muthukumar and Kannan 2013), Iris and Face (Cortland Tompkins 2011), Palm veins and Signature (Hassan Soliman et al. 2009)

Table 5.5 Comparison of proposed MMBAS with existing multimodal biometric systems

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Method</th>
<th>EER (in%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Face, Finger and Palm print (Shekhar Karanwal 2013)</td>
<td>Wavelet decomposition and SIFT</td>
<td>3.3</td>
</tr>
<tr>
<td>Speech and Signature (Dapinder Kaur et al. 2013)</td>
<td>Mel Frequency Cepstral Coefficient (MFCC), Feature level Fusion, and SIFT algorithm</td>
<td>0.91</td>
</tr>
<tr>
<td>Face and Finger Vein (Muhammad Imran Razzak et al. 2011)</td>
<td>Fusion of low resolution face and finger vein</td>
<td>0.69</td>
</tr>
<tr>
<td>Finger print and FKP (Muthukumar and Kannan 2013)</td>
<td>K-means clustering using SIFT</td>
<td>0.6</td>
</tr>
<tr>
<td>Iris and Face (Cortland Tompkins 2011)</td>
<td>Opportunistic feature selection with SIFT and SURF</td>
<td>2.13</td>
</tr>
<tr>
<td>Palm Veins and Signature (Hassan Soliman et al. 2009)</td>
<td>Feature Level Fusion</td>
<td>0.9</td>
</tr>
<tr>
<td>Proposed MMBAS Face and FKP</td>
<td>SIFT</td>
<td>0.63</td>
</tr>
</tbody>
</table>

The following Figure 5.11 shows the comparison of EER with the existing system and the proposed MMBAS based on Indian face (IITK) and Poly-U FKP database.
Figure 5.11 Comparison of EER with existing systems and proposed MMBAS

Based on Table 5.5 dataset values, the above chart Figure 5.11 has been prepared to show the comparison of EER between existing and proposed MMBAS. Training: One image per person from each modality i.e., face and FKP is used enrolment in database and which are further used for feature extraction and feature level fusion. Fused feature vector is then encoded and is saved as feature vector and is used for identification and verification.

Testing: Pair of face and FKP images is used for testing. Imposter matching scores are generated by validating and testing the first client against itself and also against the remaining subjects. Fused feature vector is generated from a pair of face and FKP images and is compared with the feature vectors of images which are stored database.

The proposed method compares with the SURF method to check the accuracy. The accuracy measured in terms of Area Under Curve (AUC), and the SIFT provided 99.46% in MMBAS. In the proposed system implementation of SIFT, feature vector of 128 values is computed from the local image region around the key-point as key-point descriptor and only feature vector of 64 values is
computed around key-point in the case of SURF. And then to equally match the two algorithms, parameter Hessian threshold of SURF was also adjusted to detect almost equal number of feature points like that of SIFT. Figure 5.12 shows the graphical representation of AUC and EER of proposed MMBAS along with the ROC curve of proposed system.

![Figure 5.12 ROC Curve for both SIFT and SURF](image)

From the above results it is determined that,

- SIFT provided EER of 0.63 and SURF provided EER of 2.13.

Accuracy measured in terms of AUC (area under curve),

- SIFT provided 99.46% and SURF provided 98.50%.

The SURF’s performance in terms of EER was much lower than the results of previous works. The reason may be the fact that we did not use any enhancement to the original face and FKP images during this experiment and set the parameters of the SURF to produce approximately equal number of feature like in the case of SIFT for the purpose of this comparison. As far as the results are compared and tested, SURF provided poor EER than SIFT.
The proposed method have successfully implemented MMBAS using SIFT feature descriptors and evaluated their performance using the standard database. The performance of SIFT in term of EER, was almost equal to that of existing works. But, the performance of SIFT during our evaluation was somewhat different from earlier results.

5.7 SOME IMPORTANT OBSERVATIONS

- For searching database of 165 to find a match of input image,

- If Enrolment = 8 images / person, then SIFT taking around 1.6 second (307.7/185) and SURF is taking around 1.5 seconds (286.9/185) in Multimodal Biometric Authentication System (MMBAS).

- It means, for authenticating one person by matching his features with a template of 185x8 bit size feature database, approximately 1.6 seconds needed. So obviously, this 1.8 second will grow rapidly if we increase the enrolment in the database.

- For example, if we need to enrol 1000 sets of feature descriptors (1000 x 8) of persons, then it may take around 10 seconds to authenticate a person using that huge template set. In SURF Feature vector = 64 values is computed around key-point. In SIFT Feature vector = 128 values is computed around key-point. But the SURF storage cost is higher than SIFT.

- This difference is mainly due to the data type used in both cases. Even though SURF is claimed as faster method than SIFT, SURF was three times faster than SIFT only during the enrolment process and only provided equal speed during the authentication / matching stage. This may be the reason that the proposed system used almost equal number of feature points while comparing them.

- The EER of SURF was much poor than EER of SIFT.
5.8 DATASET AND SYSTEM REQUIREMENT

There are many face databases are used in biometric authentication system like FERET, YALE, AR, IITK face database etc. Compared to other, IITK face database contains only Indian images which perform better results in Indian voting system through online. FERET database consists of face images of persons from different nationality in varying poses and times. It contains 261 images in $+15^\circ$ pose, among which 100 are used for learning the transformation matrix and 161 are used for testing. Contains 423 images in $\pm 15^\circ$ pose, among which 50 are used for learning the transformation and 373 are used for testing. Unlike IITK database which contains only Indian face images, it contains face images of people from different nationality including blacks, whites, asians, caucasians, etc. Hence, results not as good as IITK Database. So, this research work uses Indian face database (IITK) for face recognition process.

FKP images were collected from 165 volunteers, including 125 males and 40 females. Among them, 143 subjects were 20–30 years old and the others were 30–50 years old. We collected samples 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. In total, the database contains 7,920 images from 660 different fingers. The average time interval between the first and the second sessions was about 25 days. The maximum and minimum intervals were 96 days and 14 days, respectively. Detail information about database is discussed in Annexure – I.

In recognition process, the estimated average images of 140 persons along with their different angled images are considered. The total training images that were split into two equal portions of face and finger knuckle images. Total test image for the per person template of 2 has the score 370 which means nearly 185 image of face images and the remaining 185 image of finger knuckle images are used to get the matching score 99.46% (on an average per seconds).
The implementation of result image shows the histogram samples that are equal in length and magnitude. The feature vector is modified to reduce the effects of illumination transforms. The resultant set overrules the difference of modification that affects vector normalization. The pixel difference was calculated by adding a constant value to every pixel and such component will distinguish from noise and will never affect the pixel value as stated by (David G Lowe 2004). Hence the resultant set achieves the experiment result of 128 sets sample featuring and its vector for each key point. Therefore, the descriptor is invariant to affine changes in illumination. Even though the cost of extracting the image feature measure was done by cascade filter approach, by which only the initial set pixel value is deputed for the valuable operation that pass the initial test. It is obvious that the above mentioned operation will consume considerable time since the operation is repeated in several scales of different image frames.

For implementing SIFT based authentication system, the proposed methods used VLFeat open source library for computer vision algorithms and Matlab version of VLFeat-0.9.16. It is a cross-platform open source collection of vision algorithms with a special focus on visual features and clustering. For better performance, binary versions of these libraries were used in both the designs.

5.8 SUMMARY

Biometric systems can effectively and efficiently operate in ultra large-scale applications, i.e. those capable of supporting hundreds of millions of registered users, have a number of potential opportunities. Such systems will be able to support National ID programs or improve homeland security, e-commerce, e-voting and more effective implementation of social welfare programs in countries with large population (e.g., India, China and United States). The expectations from the proposed method for such large-scale applications can be summarized as follows:

- High accuracy and throughput under varying operating conditions and user Composition
- Sensor interoperability
Rapid collection of biometric data in harsh operating environments with virtually no failure to enrol rate

High levels of privacy and template protection and

Secure supporting information/operating systems.

This chapter discussed about an efficient Multimodal Biometric Authentication System (MMBAS). The proposed method is successfully implemented and developed an efficient Multimodal Biometric Authentication System (MMBAS) using Scale Invariant Feature Transform (SIFT) based on feature descriptors and evaluated their performance using the standard database. The accuracy of proposed MMBAS is high compared with existing MMBAS. Thus the proposed multimodal biometric recognition is efficient.