CHAPTER 1

INTRODUCTION

1.1 Face recognition

Face recognition is a biometric technique for automatic identification. It makes use of face, the most usual human identifier and unique facial characteristics. Although research in automated face recognition has been conducted since the 1960's, it has only recently caught the attention of the scientific community. It has emerged as the most successful application of image processing and pattern analysis in computer vision. This because face recognition, in additional to having numerous practical applications such as bankcard identification, access control, Mug shots searching, security monitoring, and surveillance system, is a fundamental human behaviour that is essential for effective communications and interactions among people. A formal method of classifying faces was first proposed in [1]. The author proposed collecting facial profiles as curves, finding their norm, and then classifying other profiles by their deviations from the norm. This classification is multi-modal, i.e. resulting in a vector of independent measures that could be compared with other vectors in a database. Progress has advanced to the point that face recognition systems are being demonstrated in real-world settings [2]. The rapid development of face recognition is due to a combination of factors such as active development of algorithms, the availability of a large databases of facial images, and a method for evaluating the performance of face recognition algorithms. Face recognition is a biometric approach that employs automated methods to verify or recognize the identity of aliving person based on his/her physiological characteristics. In general, a biometric identification system makes use of either physiological characteristics (such as a fingerprint, iris pattern, or face) or behaviour patterns (such as hand-writing, voice, or key-stroke pattern) to identify a person. Face recognition has the benefit of being a passive, non intrusive system to verify personal identity in a "natural" and friendly way. Though technologies in face recognition have been evolving through years, but its importance has recently grown in a significant manner. This is due to certain increase in civilian and commercial research projects, the need for surveillance in trafficking and increased terrorist activities. Enhanced real time computation and the relative ease of obtaining face biometric samples from a distance make the face recognition system very desirable.

In general, biometric devices can be explained with a three step procedure:

(1) A sensor takes an observation. The type of sensor and its observation depend on the type of biometric devices used. This observation gives us a “Biometric Signature” of the individual.

(2) A computer algorithm “normalizes” the biometric signature so that it is in the same format (size, resolution, view, etc.) as the signatures on the system’s database. The normalization of the biometric signature gives us a “Normalized Signature” of the individual.

(3) A matcher compares the normalized signature with the set (or sub-set) of normalized signatures on the system's database and provides a “similarity score” that compares the Individual’s normalized signature with each signature in the database set (or sub-set). What is then done with the similarity scores depends on the biometric system’s application?

The application of face recognition technique can be categorized into two main parts: law enforcement application and commercial application. Face recognition technology is primarily used in law enforcement applications, especially Mug shot albums (static matching) and video surveillance (real-time matching by video image sequences). The commercial applications range from static matching of photographs on credit cards, ATM cards, passports, driver’s licenses, and photo ID to real-time matching with still images or video image sequences for access control. Each application presents different constraints in terms of processing.

However, the reliability of face recognition schemes still poses a great challenge to the scientific community. It is a challenging task for a machine to recognize human faces accurately in real-time, especially under variable circumstances such as variations in illumination, pose, facial expression, makeup etc. The similarity of human faces and the unpredictable variations are the greatest obstacles in face recognition. Despite the challenges involved, face recognition systems are being widely used in commercial systems to perform real-time face detection, image registration and image matching.

Though technologies in face recognition have been evolving through years, but its importance has recently grown in a significant manner. This is due to certain increase in civilian and commercial research projects, the need for surveillance in trafficking and increased terrorist activities. Enhanced real time computation and the relative ease of obtaining face biometric samples from a distance make the face recognition system very desirable.

1.2 Role of Face Recognition

Within today's environment of increased importance of security and organisation, identification and authentication methods have developed into a key technology in various areas: entrance control in building; access control for computers in general or for automatic teller machines in particular; day to day affairs like withdrawing money from a bank account or dealing with the post office; or in the prominent field of criminal investigation. Such requirement for reliable personal identification in computerized access control has resulted in an increased interest in biometrics.

Biometric identification is the technique of automatically identifying or verifying an individual by a physical characteristics or personal trait. The term "automatically" means the biometric identification system must identify or verify a human characteristic or trait quickly with little or no intervention from the user. Biometric technology was developed for use in high-level security system and law enforcement markets. The key element of biometric technology is its ability to identify a human being and enforce security.

Biometric characteristic and traits are divided into behavioural or physical categories. Behavioural biometrics encompasses such behaviours as signature and typing rhythms. Physical biometric system use the eye, finger, hand, voice and face for identification.

Another well-known biometric measure is that of fingerprints. Various institutions around the world have carried out research in the field. Fingerprints system are unobtrusive and relatively cheap to buy. They are used in banks and to control entrance to restricted access areas. Fowler [3] has produced a short summary of the available system.

Fingerprints are unique to each human being. It has been observed that the iris of the eye, like fingerprints, displays patterns and textures unique to each human and that it remains stable over decades of life as detail by Siedlarz [4]. Daugman [5] designed a robust pattern recognition method on 2-D Gabor transforms to classify human irises. Speech recognition also offers one of the most natural and less obstructive biometric measures, where a user is identified through his or her spoken words. AT&T have produced a prototype that stores a person's voice on a memory card, details of which are described by Mandelbaun [6].

While appropriate for bank transaction and entry into secure areas, such technologies have the disadvantage that they are intrusive both physically and socially. They require the user to position their body relative to the sensor, then pause for a second to declare himself or herself. this pause and declare interaction is unlikely to change because of the fine-grain spatial sensing required. Moreover, since people cannot recognize people using this sort of data, these types of identification do not have a place in normal interactions and social structures.

While the 'pause and present' interaction perception is useful in high security applications, they are exactly the opposite of what is required when building a store that aims to recognize its best customers, or an information that remembers people, or a house that knows the people who live there.

A face recognition system would allow user to be identified by simply walking past a surveillance camera. Human beings often recognize one another by unique facial characteristics. Automatic facial recognition, is based on this phenomenon. Facial recognition technology is being used to improve human efficiency when recognizing faces and is one of the fastest growing fields in biometric industry. Interest in facial recognition is being fuelled by the availability and low cost of video hardware, the ever- increasing number of video cameras being placed in the workspace, and the non-invasive aspect of facial recognition system.

Although facial recognition is still in the research and development phase, several commercial systems are currently available and research organizations are working on the development of more accurate and reliable systems

1.3 Challenges for the Technology

Even though face recognition has been studied for a long time, the researchers are still trying to develop a robust automated face recognition system. This is due to the challenges that can fail even the more recent technologies. Hence, the techniques are facing certain issues attributed as: [7, 8]

1. **Facial expression**: Changes in the facial expression can directly affect the shape, size and structure of facial components and thus the face appearance. Various Facial Expression is shown in Fig. 1.1

Fig. 1.1: Various Facial Expression.

2. **Head-pose**: Most of the face detectors use the features of the facial components like the relative size and position of the components. Extracting these kinds of features may not be easy when the head is rotated. The situation becomes worse or all of the facial components are hidden due to the changes in the head-pose as shown in Fig. 1.2

Fig. 1.2: Various Head Pose.

3. **Presence or absence of structural elements**: Facial features such as beards, moustaches, hair and glasses may or may not be present and there is a wide range of changes among these components including shape, colour, and size as shown in Fig. 1.3

Fig. 1.3: Presence of various Facial Features.

4. **Occlusion**: Faces may be partially occluded by other objects. In an image with a group of people, some faces may partially occlude other faces occlude other faces as shown in Fig. 1.4(a) or by some objects in the scene Fig. 1.4(b).

Fig. 1.4: (a) Face Occluded by dome face, (b) Face occluded by some object.

5. **Image orientation:** Face images directly vary for different rotations about the camera's optical axis as shown in Fig. 1.5.

Fig. 1.5: Variation due to Image Orientation.

6. **Imaging condition**: When the image is formed, factors such as lighting (spectra, source distribution and intensity) and camera characteristics (sensor response, lenses) affect the appearance of a face as shown in Fig. 1.6.

Fig. 1.6: Variation due to Various Imaging Condition.

The goal of face recognition is to successfully match the image of faces in spite of variations in above issues hence these issues are the challenges of each techniques being used for face recognition.

1.4 Objectives

This thesis aims to develop a robust face recognition system which is able to perform adaptive image pre-processing in case of illumination variation in face input image and increase the efficiency of feature extraction and matching. The light variation is there in database and input image to add to the ambiguity. To achieve this goal a number of sub-objective has been set.

- To perform regional adaptive histogram equalization of the input images.
- To use SIFT for feature extraction from the face images.
- To perform matching of test and reference images by an efficient classifier and perform face recognition.

1.5 Overview of the Proposed Technique in the Thesis

This thesis presents a face recognition technique comprises of an adaptive image quality based normalization algorithm. The features of the normalized face image have been extracted by SIFT. And further the test and reference images have been matched by the use of K-nearest neighbor classifier [9]. The technique has been proposed to have a high recognition rate automated face recognition system.

The flow diagram is shown below:-

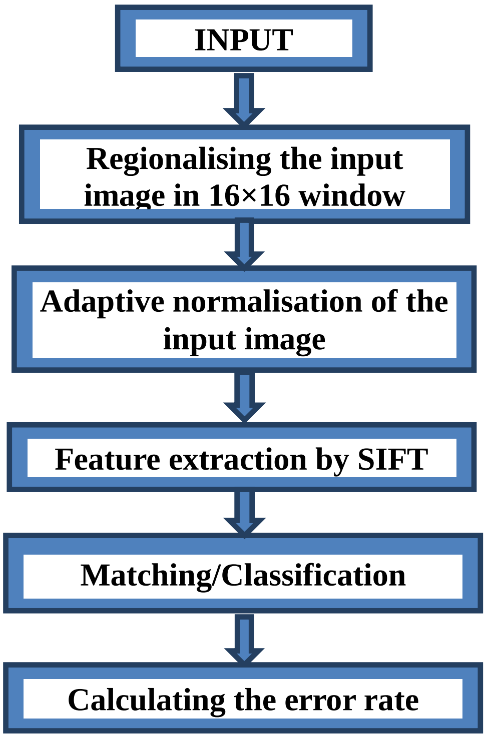

Fig. 1.7: System Flowchart

1.6 Organization of the Thesis

After the Introductory Chapter, Rest of the thesis is organized as follows:-

- **Chapter 2** conducts a literature review of face recognition. In this chapter, generic face recognition system has been explained.
- **Chapter 3** describes the proposed methodology for image quality measure based adaptive image per-processing.
- **Chapter 4** describes SIFT based feature Extraction and classifier used for matching.
- **Chapter 5** provides with the Experimental Results of the Algorithm with various Datasets. It also gives a comparative analysis of our method with other existing methods and also shows the error rate of the Algorithm.
- **Chapter 6** summarizes the Thesis with Conclusion, Advantages, Limitations and Future Work.

CHAPTER 2

FACE RECOGNITION

2.1 Literature Review

Kirby and Sirovich [10] were among the first to apply principal component analysis (PCA) to face images. It was showed that PCA is an optimal compression scheme that minimizes mean squared error between original images and their reconstructions for any given level of compression [11, 12]. Turk and Pentland popularized the use of PCA for face recognition [13]. They used PCA to compute a set of subspace basis vectors (which they called "eigenfaces" [13]) for a database of face images, and projected the images in the database into the compressed subspace. New test images were then matched to images in the database by projecting them onto the basis vectors and finding the nearest compressed image in the subspace (Eigen space).

The initial success of Eigen faces popularized the idea of matching images in compressed subspaces. Researchers began to search for other subspaces that might improve the performance. One alternative is Fisher's linear discriminant analysis (LDA, a.k.a. "fisherfaces") [14]. For any N-class classification problem, the goal of LDA is to find the N-1 basis vectors that maximize the interclass distances while minimizing the intraclass distances. At one level, PCA and LDA are very different: LDA is a supervised learning technique that relies on class labels, whereas PCA is an unsupervised technique. Nonetheless, in circumstances where class labels are available both technique can be used and LDA has been compared to PCA in several studies [15].

One characteristic of both PCA and LDA is that they produce spatially global feature vectors. In other words, the basis vectors produced by PCA and LDA are non-zero for almost all dimensions. It implies that a change to a single input pixel will alter every dimension of its subspace projection. There is also a lot of interest in techniques that create spatially localized feature vectors, in the hopes that they might be less susceptible to occlusion and would implement recognition by parts. The most common method for generating spatially localized features is to apply independent component analysis (ICA) [16] to produce basis vectors that are statistically independent (not just linearly decorrelated, as with PCA) [17]. ICA can also be used to create feature vectors that uniformly distribute data samples in subspace. Conceptually, this is a very different use of ICA producing feature vectors that are not spatially localized. Instead, it produces feature vectors that draw fine distinctions between similar images in order to spread the samples in subspace. Another well-known approach is the Fisherfaces in which the Fisher's linear discriminant (FLD) is employed after the PCA is used for dimensionality reduction [18].

Compared with the Eigenface (PCA) approach, the Fisherface approach is more insensitive to large variations in lighting direction and facial expression. More recently, some variants of FLD (LDA) have been developed for face recognition such as F-LDA [19], D-LDA [20], FD-LDA [21], and KDDA [22] etc. However, the computational requirements of these approaches are greatly related to the dimensionality of the original data and the number of training samples. Later, discrete cosine transform (DCT) has also been employed in face recognition [12, 23]. The DCT has several advantages over the PCA. First, the DCT is data independent and second, the DCT can be implemented using a fast algorithm and can be applied for dimensionality reduction. It is well-known that the problems arising from the curse of dimensionality should be considered in pattern recognition. It has been suggested that as the dimensionality increases, the sample size needs to increase exponentially in order to have an effective estimate of multivariate densities [24]. In face recognition applications, the original input data are usually of high dimension, whereas only limited training samples are available. Therefore, dimensionality reduction is a very important step which will greatly improve the performance of the face recognition system.

Neural networks have been widely applied in pattern recognition for the reason that neural-networks-based classifiers can incorporate both statistical and structural information and achieve better performance than the simple minimum distance classifiers [25]. Multi-layered networks (MLNs), usually employing the back propagation (BP) algorithm, are widely used in face recognition [26]. Recently, RBF neural networks have been applied in many engineering and scientific applications including face recognition. The sub-clustering process implemented in helps in structure determination of the radial basis function (RBF) neural networks, as the number of clusters is just the number of hidden neurons in the RBF neural networks.

Generally speaking, the technical approaches for face recognition can be grouped into three categories. First is the Feature based approach, which is based on the shapes and geometrical relationships of key facial features including eyes, mouth, nose, chin and curvature based face components [27]. These key features of an image are extracted after data acquisition and later on used for face recognition. Second is the Holistic approach (Template matching approach), which takes the input face images globally and extract important facial features based on the high-dimensional intensity values of face images automatically. Although feature-based schemes are more robust against rotation, scale, and illumination variations. They greatly rely on the accuracy of facial feature detection methods and it has been argued that existing feature-based techniques are not reliable enough for extracting individual facial features.

Holistic face recognition has attracted more attention since the well-known statistical methods, the principal component analysis (PCA) also known as Karhunen-Loeve transform (KLT) [10], Independent component analysis (ICA) were applied in face recognition and later on support vector machine [28] also followed this approach. Lots of researches have been performed to analyse the performance of PCA vs. ICA under different conditions viz., visible light, infra-red images, LWIR images etc. Third one is hybrid approach, this approach uses both the face images together with the local features for face recognition.

The incredible human intelligence can be demonstrated by its ability to recognize human faces. Over the last three decades researches have been going on to study this outstanding visual perception of human beings in machine recognition of faces. While coping up with the challenges in face recognition numerous techniques have been implemented. This paper discusses some of the primitive and latest face recognition techniques in the following subsections.

2.1.1 PCA

Principal component analysis (PCA) is a statistical dimensionality reduction method, which produces the optimal linear least-square decomposition of a training set. This subspace projection technique has found application in fields such as face recognition, pattern recognition and image compression. It is computationally efficient to compare images in subspaces with significantly reduced dimensions, PCA helps to reduce image vectors with 65,536 pixels (256×256) might be projected into a subspace with only 100 to 300 dimensions. PCA reduces the data dimensionality to reveal the most effective low dimensional structure of facial patterns by decomposing the face structure into orthogonal (uncorrelated) components known as eigenvectors and eigenvalues [29]. These eigenvectors represent a set of features which together characterize the variation between face images [27]. In PCA, a set of training images I, are used to compute basis vectors. In first step, the average image in I is computed and subtracted from the training images, creating a set of data samples, given by equation (1)

$$i_1, i_2, \ldots, i_n \in I - \bar{I} \tag{1}$$

These data samples are then arrayed in a matrix with one column per sample image, as represented by X in equation (2),

$$X = \begin{bmatrix} \begin{bmatrix} \vdots \\ i1 \\ \vdots \end{bmatrix} \ldots\ldots\ldots \begin{bmatrix} \vdots \\ i_n \\ \vdots \end{bmatrix} \end{bmatrix} \tag{2}$$

XX^T is then the sample covariance matrix for the training images. The principal components of the covariance matrix are computed by solving equation (3),

$$R^T (XX^T) R = \Lambda \tag{3}$$

Where Λ is the diagonal matrix of eigenvalues and R is the matrix of orthonormal eigenvectors.

Geometrically, R is a rotation matrix that rotates the original coordinate system onto the eigenvectors. Larger the eigenvalue of the associated eigenvectors, more is the variance. The N eigenvectors associated with the largest eigenvalues are used to define the subspace, where N is the desired subspace dimensionality. Input image is compared with training set data image by measuring distance between their respective eigenvector corresponding to each feature. For using PCA technique the size of input images should be same with frontal face representation, normalization of illumination and facial features like eyes, mouth etc. is required. PCA gives robust performance under different lighting conditions by significant correlation between images with changes in illumination.

2.1.2 Artificial Neural Network

Artificial neural network is a nonlinear mathematical approach derived from structure of biological neural network having interconnected group of neurons. These neurons of biological neural network resembles the artificial neurons or nodes of artificial neural network. It is a popular tool in pattern recognition and computes the data using a connectionist approach. Kohonen [17] was the first to demonstrate neural network as a efficient technique to recognize aligned and normalized faces. One of the first artificial neural networks (ANN) technique used for face recognition is a single layer adaptive network called WISARD containing a separate network for each stored individual [30]. A conventional artificial neural network consist of interconnected group of nodes from input to output layers via some hidden network layers. Here Fig 2.1 shows a single hidden layer between input and output layers.

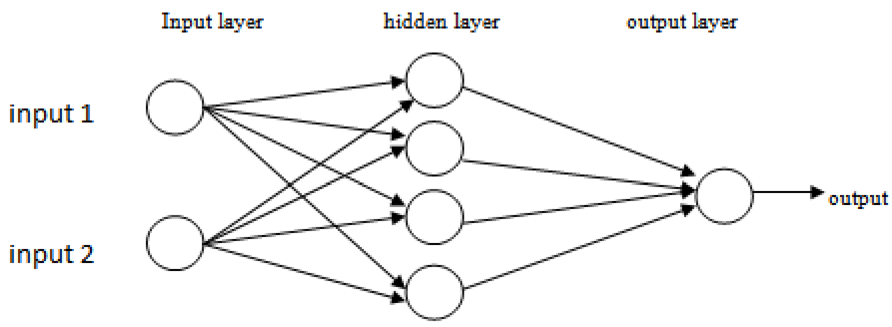

Fig. 2.1: Three Layer Artificial Neural Network.

The input test images are given at input layer and face or non-face image match as 0 or 1 value of node are given at output layer. During the process, weights of network are continuously updated until error rate in recognition is close to zero. Nonlinearity of the network helps in finding patterns in data by modelling relationship between input and output. The way of constructing a neural network structure is crucial for successful recognition and the process is application dependent. For face detection, multilayer perceptron [11, 31] and convolutional neural network [32] have been applied. Hybrid neural network [32] combines local image sampling, self-organizing map (SOM) neural network and convolutional neural network. Quantization of input image is done by SOM keeping input samples in topological space leading to dimensionality reduction and invariance to minor changes in image sample. Feature extraction in hierarchical set of layer is done by convolution network. The reported recognition rate of hybrid neural network is 96.2% on ORL database of 400 images of 40 individuals.

2.1.3 Support Vector Machines

SVM [28] are considered an effective method for general purpose pattern and face recognition as it gives a high generalization performance without the need to add extra information. SVM were introduced as learning machines but now efficiently used as linear classifiers. For the first time Osuna et al. [34] used it as a linear classifier for face detection. A support vector machine constructs a hyper-plane or set of hyper-planes in a high- or infinite-dimensional space, which can be used for classification, regression or other tasks. Intuitively, maximization of the margin between the decision hyper-plane and the data in the training set is required. So, an optimal hyper-plane is selected that should minimize the classification error of the unseen test set. According to [35], a hyper-plane is selected called as Optimal Separating Hyper-plane (OSH), which minimizes the risk of misclassifying the images in the training set and input test image data set. As shown in Fig 2.2, a solid hyper-plane (OSH) is separating training data of class 1 and 2, outliers in the training data set can be handled by means of soft margins. SVM is a more suitable approach for average size face recognition systems because normally those systems have only a small number of training image samples. It uses an additional principal called structural risk minimization. The purpose of structural risk minimization is to give an upper bound on the expected generalization error.

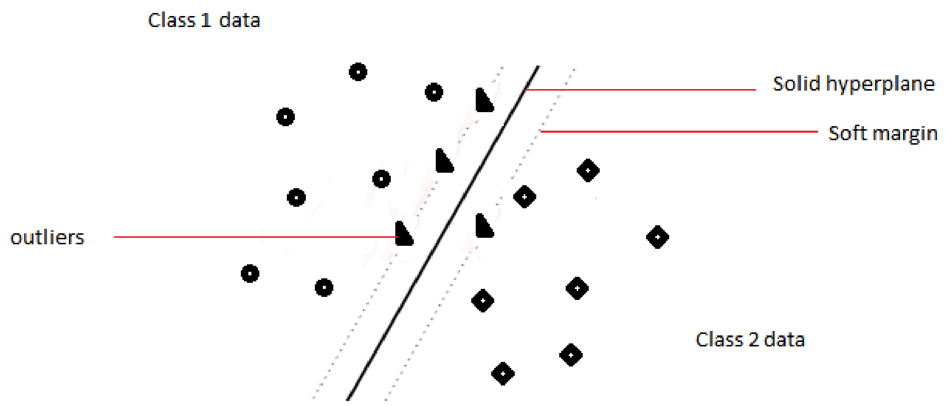

Fig. 2.2: Classification Pattern of SVM.

SVMs can work on high-dimensional feature spaces by means of a dual formulation in terms of kernels. C. Shavers et al. presented a SVM-based approach for face detection. They used the ORL database consisting of 200 images. A 20×20 window of pixels (400 dimension vectors) is extracted from each image to create a sample. The hyper-plane decision is based on a degree-one polynomial. The decision function to separate training set data is constructed equidistance between support vector archetypes to give an optimal hyper-plane decision function. Using one degree polynomial as a kernel function, SVM maps images into higher dimension transform space to construct decision function and giving a high recognition rate.

2.1.4 Hidden Markov Model

A Hidden Markov model (HMM) [28, 45] is a statistical Markov model in which the system being modelled is assumed to be a Markov process with unobserved (hidden) states. The state is directly visible to the observer in a regular Markov model, thus the state transition probabilities are the only parameters. The state is not directly visible but output dependent on the state is visible in a hidden Markov model. Each state has a probability distribution over the possible output sequence. Therefore the sequence of output generated by an HMM gives some information about the sequence of states. Hence the goal of HMM as a learning machine, is to find the best set of state transition and output probabilities for the given output sequence. Here 'hidden' refers to the state sequence through which the model passes, not to the parameters of the model. HMM require 1D observation sequence thus a 2D image is converted either to 1D temporal sequence or 1D spatial sequence. Based on a set of output sequences, HMM estimates the maximum likelihood of the parameters. Baum-Welch algorithm or Baldi-chauvin algorithm [36] can be applied for estimating maximum likelihood, an example of a forward-backward algorithm. After training HMM, output probability of an observation determines the class to which is belongs. An efficient method for getting the observation vector for face extraction is by using Karhunen-Loave transform [29] used under lightening variations. A pseudo 2D HMM [28] is reported with higher recognition rate.

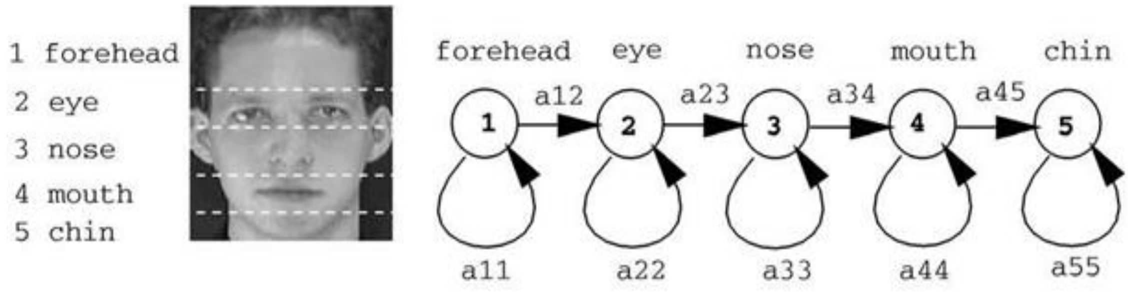

Fig. 2.3: Five States is Trained with the Sequence of Observation Vectors [27]

From Fig 2.3, it is clearly seen that for applying HMM, samples should be converted to observation vector. Observation sequence is produced by vertical scanning, so HMM is trained with five states with sequence of observation.

2.1.5 MPCA and LPP

PCA is a linear projection method in which dimensionality reduction is applied to the original image space. MPCA is an improved version of PCA employing multi-linear algebra in which each image is divided into number of sub-block image and then PCA is applied for each sub-block image. A multilinear approach is required to analyze image ensemble when a face image is embedded with multiple factors like geometry, view points and lighting. MPCA is a multilinear subspace learning method used for face extraction directly from multidimensional object.

Locality Preserving Projection (LPP) is a linear approximations of the nonlinear Laplacian Eigen map [30]. It is also well-known as a linear graph embedding method [40, 41]. The locality preserving quality of LPP gives it application in the field of information retrieval. It performs a nearest neighbour search in the low dimensional space to retrieve audio, video and text documents by locality preservation under a vector space model. LPP intends to preserve only local structure, it is probable that a nearest neighbour search in the low dimensional space will produce similar results compared to that in the high dimensional space. While LPP allows an indexing scheme to access quickly.

The combined approach of using MPCA and LPP consist of mainly 4 steps:

1. Image pre-processing comprising of face normalization and resizing of face image.

2. Appling MPCA for dimensionality reduction.

3. Applying LPP for face extraction.

4. Face recognition using L2 similarity distance measure. The L2 distance is computed between the face images present in the database and the query image for matching process.

The similarity distance measure for a pair of face images is computed. In this a threshold determines whether the face pair is classified as same or different database and the query image for matching process.

$$d\,(a, b) = \sqrt{\sum_{i=1}^{n} [a(i) - b(i)]^2} \tag{4}$$

The formula used to compute the L2 distance measure is given by equation (4).

2.1.6 Face Recognition using Texture and Depth Information

In this technique, 3D face image or depth information is used for face recognition. Texture information is more efficient than depth information for face recognition. However texture information is more sensitive to illumination and poses variation, thus recognition rate drops in environment with illumination changes. For enhancing the accuracy of face recognition, algorithms utilize both depth and texture information [42, 43]. 3D information is used for the estimation of face rotation and orientation. The rotation compensated 2D images are then used for face recognition.

The proposed face recognition algorithm [46] has the following steps:

1. Proper local features are extracted from input texture image and are compared with all local features that are extracted from texture face images in database.

2. Some predefined criteria are checked for matched local features. If the criteria are satisfied, we determine the recognized face; otherwise depending on the matched local features some face images are selected from database for recognition. The selected faces from database are called restricted face database.

3. The depth information is used to calculate and compensate for the 3D rotation matrix and translation vector between input image and each image in the restricted face database.

4. Using proper thresholding algorithm, unnecessary parts of depth images like hair, neck, dress and collar are removed from input image and the restricted face database.

5. Tip of nose is detected in depth images and used to compensate for scale change in depth images.

6. The correlation between the input image and restricted face database are calculated and the final recognized face is determined based on the correlation values.

It is a two level algorithm, first level is face recognition using texture and second is face recognition using depth. Face extraction and matching done by SIFT descriptor [44] and matching criteria are included in face recognition using texture. The face images which are not recognized using texture information are transferred to the stage of face recognition using depth.

2.2 Generic Face recognition system

Facial recognition is a visual pattern recognition task. The three-dimensional human face, which is subject to varying illumination, pose, expression etc. has to be recognized. This recognition can be performed on a variety of input data sources such as:

• A single 2D image.

• Stereo 2D images (two or more 2D images).

• 3D laser scans.

Facial recognition systems usually consist of steps, as shown in Fig 2.4; image acquisition, face detection (localization), face pre-processing (face alignment/normalization, light correction and etc.), feature extraction and feature matching. These steps are described in the following sections and are shown as flow chart below.

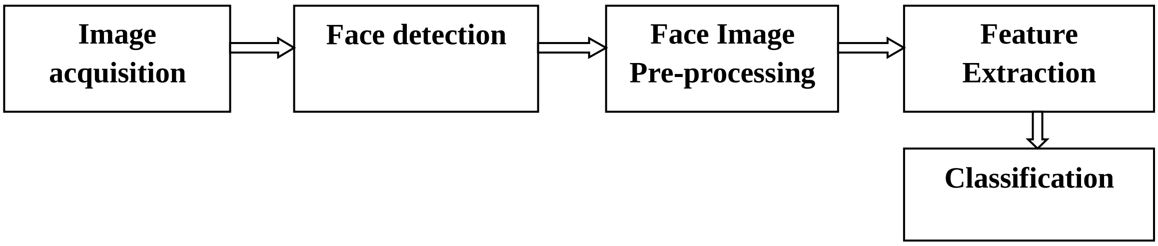

Fig. 2.4: Generic Face Recognition Model.

2.2.1 Image Acquisition

A digital image is a two-dimensional (2D) discrete signal. Mathematically, such signals can be represented as functions of two independent variables - for example, a brightness function of two spatial variables. A monochrome digital image $f(x, y)$ is a 2D array of luminance values. Each element of the array is called a pel (picture element), or more commonly a pixel. A colour digital image is typically represented by a triplet of values, one for each of the colour channels, as in the frequently used RGB colour scheme. The individual colour values are almost universally 8-bit values, resulting in a total of 3 bytes (or 24 bits) per pixel. This yields a threefold increase in the storage requirements for colour versus monochrome images. Naturally, there are a number of alternative methods of storing the image data. Most widely used are the so-called pixel-interleaved (or meshed) and colour-interleaved (or planar) formats. Row-wise or column-wise interleaving methods are less frequent. In a pixel interleaved format, every image pixel is represented by a list of three values. The imaging sensors plays important role in the image acquisition. The structure and operation of the eye is very similar to an electronic camera, which is most often used to acquire real world images. Both are based on two major components: a lens assembly, and an imaging sensor. The lens assembly captures a portion of the light emanating from an object, and focuses it onto the imaging sensor. The imaging sensor then transforms the pattern of light into a video signal, either electronic or neural. The term focus means there is a one-to one match of every point on the object with a corresponding point on the screen.

For example, consider a 1mm region on the object. In bright light, there are roughly 100 trillion photons of light striking this one square millimeter area each second. Depending on the characteristics of the surface, between 1 and 99 percent of these incident light photons will be reflected in random directions. Only a small portion of these reflected photons will pass through the lens. For example, only about one-millionth of the reflected light will pass through a one centimeter diameter lens located 3meters from the object. Refraction in the lens changes the direction of the individual photons, depending on the location and angle they strike the glass/air interface. These direction changes cause light expanding from a single point to return to a single point on the projection screen. All of the Omni-directional image processing for human detection and tracking photons that reflect from the object and pass through the lens are brought back together at the "object" in the projected image. In a similar way, a portion of the light coming from any point on the object will pass through the lens, and be focused to a corresponding point in the projected image.

2.2.2 Face detection

The second step in any automatic face recognition systems is the detection of faces in images. Face detection algorithms usually share common steps. Firstly, some data dimension reduction is done, in order to achieve an admissible response time. Some pre-processing could also be done to adapt the input image to the algorithm prerequisites. Then, some algorithms analyze the image as it is, and some others try to extract certain relevant facial regions. The next phase usually involves extracting facial features or measurements. These will then be weighted, evaluated or compared to decide if there is a face and where is it. Finally, some algorithms have a learning routine and they include new data to their models. Face detection is, therefore, a two class problem where we have to decide if there is a face or not in a picture. This approach can be seen as a simplified face recognition problem. Face recognition has to classify a given face, and there are as many classes as candidates. Consequently, many face detection methods are very similar to face recognition algorithms. Or put another way, techniques used in face detection are often used in face recognition.

Up to the mid-1990s, most work on segmentation was focused on single-face segmentation from a simple or complex background. These approaches included using a whole-face template, a deformable feature-based template, skin colour, and a neural network. Significant advances have been made in recent years in achieving automatic face detection under various conditions. Compared to feature-based methods and template-matching methods, appearance or image-based methods [Rowley et al. 1998; Sung and Poggio 1997] that train machine systems on large numbers of samples have achieved the best results. This may not be surprising since face objects are complicated, very similar to each other, and different from non-face objects. Through extensive training, computers can be quite good at detecting faces. More recently, detection of faces under rotation in depth has been studied. One approach is based on training on multipleview samples [Gu et al. 2001; Schneiderman and Kanade 2000]. Compared to invariant-feature-based methods [Wiskott et al. 1997], multiview-based methods of face detection and recognition seem to be able to achieve better results when the angle of out-of-plane rotation is large (35◦). In the psychology community, a similar debate exists on whether face recognition is viewpoint-invariant or not. Studies in both disciplines seem to support the idea that for small angles, face perception is view-independent, while for large angles, it is view-dependent.

In a detection problem, two statistics are important: true positives (also referred to as detection rate) and false positives (reported detections in non-face regions). An ideal system would have very high true positive and very low false positive rates. In practice, these two requirements are conflicting. Treating face detection as a two-class classification problem helps to reduce false positives dramatically [Rowley et al. 1998; Sung and Poggio 1997] while maintaining true positives. This is achieved by retraining systems with false positive samples that are generated by previously trained systems.

2.2.3 Image Pre-processing

The face pre-processing step aims at normalizing, i.e. reducing the variation of images obtained during the face detection step. Using AAM in the process of face detection provides a well-defined framework to retrieve the photometric information as a shape free image as well as the geometric information as a shape.

Light Correction

The unpredictable change in lighting conditions is a problem in facial recognition. Therefore, it is desirable to normalize the photometric information in terms of light correction to optimize the facial recognition. Histogram equalisation has been used for lighting correction and briefly explained. Histogram equalization (HE) can be used as a simple but very robust way to obtain light correction when applied to small regions such as faces. The aim of HE is to maximize the contrast of an input image, resulting in a histogram of the output image which is as close to a uniform histogram as possible. However, this does not remove the effect of a strong light source but maximizes the entropy of an image, thus reducing the effect of differences in illumination within the same “setup” of light sources. By doing so, HE makes facial recognition a somehow simpler task. Two examples of HE of images can be seen in Figure 8.1. The algorithm of HE is straight forward and effective, [47].

2.2.4 Feature Extraction

Face recognition’s core problem is to extract information from images. This feature extraction process can be defined as the procedure of extracting relevant information from a face image. This information must be valuable to the later step of identifying the subject with an acceptable error rate. The feature extraction process must be efficient in terms of computing time and memory usage. The output should also be optimized for the classification step. Feature extraction involves several steps - dimensionality reduction, feature extraction and feature selection. This steps may overlap, and dimensionality reduction could be seen as a consequence of the feature extraction and selection algorithms. Both algorithms could also be defined as cases of dimensionality reduction.

Dimensionality reduction is an essential task in any pattern recognition system. The performance of a classifier depends on the amount of sample images, number of features and classifier complexity. One could think that the false positive ratio of a classifier does not increase as the number of features increases. However, added features may degrade the performance of a classification algorithm. This may happen when the number of training samples is small relative to the number the features. This problem is called “curse of dimensionality” or “peaking phenomenon”. A generally accepted method of avoiding this phenomenon is to use at least ten times as many training samples per class as the number of features. This requirement should be satisfied when building a classifier. This “curse” is one of the reasons why it’s important to keep the number of features as small as possible. The other main reason is the speed. The classifier will be faster and will use less memory. Moreover, a large set of features can result in a false positive when these features are redundant. Ultimately, the number of features must be carefully chosen. Too less or redundant features can lead to a loss of accuracy of the recognition system. We can make a distinction between feature extraction and feature selection. Both terms are usually used interchangeably. Nevertheless, it is recommendable to make a distinction. A feature extraction algorithm extracts features from the data. It creates those new features based on transformations or combinations of the original data. In other words, it transforms or combines the data in order to select a proper subspace in the original feature space. On the other hand, a feature selection algorithm selects the best subset of the input feature set. It discards non-relevant features. Feature selection is often performed after feature extraction. So, features are extracted from the face images, then a optimum subset of these features is selected. The dimensionality reduction process can be embedded in some of these steps, or performed before them. Researchers in face recognition have used, modified and adapted many algorithms and methods for the purpose of feature extraction. For example, PCA was invented by Karl Pearson in 1901 [48], but proposed for pattern recognition 64 years later [49]. Finally, it was applied to face representation and recognition in the early 90's [50, 51].

2.3 Illumination effect on face recognition and literature review

Illumination variation has enormously complex effects on the image of an object. In the image of a familiar face, changing the direction of illumination leads to shifts in the location and shape of shadows, changes in highlights, and reversal of contrast gradients. The large image variations that result from changing the illumination direction have been demonstrated by Adini, Moses, and Ullman (1995). They compared images of several faces rendered with the same or different lighting direction. Several representations of these images were considered: gray-scale images, images filtered with Gabor functions, edge maps, and 1st and 2nd derivatives of gray-scale images. For all of these representations, they found that varying the illumination direction resulted in larger image differences than did varying the identity of the face. Lighting condition variations during input and identification stages significantly leads to intraclass variations of face images. Variation in illumination conditions have been typically addressed by methods categorized as follows:

1) Feature-based methods;
2) Generative methods
3) Holistic methods.

In feature-based approaches, illumination invariant features have been used to represent a face. Typically, these are geometrical measurements and relationships between local facial features such as the eyes, mouth, nose, and chin [52], [53]. Feature-based methods are robust against varying illumination condition. Accurate face and facial feature point detection are needed for the method, which is a tough and challenging task.

Generative methods [54]–[57] have been proposed to deal the problem of varying illumination based on the assumption of the Lambertian model. It has been demonstrated that the variability of images under a fixed pose, consisting of only diffuse reflection components and varying illumination conditions can be represented by a linear combination of three basis images [58], [59]. Belhumeur and Kriegman [60] addressed the problem and proposed that a set of images of an object under fixed posed, consisting of diffuse reflection components and shadows under arbitrary lighting conditions, forms a convex cone (called the illumination cone) in the image space and that this illumination cone can be approximated by a low-dimensional subspace. Experimental results show that these generative methods perform well under varying illumination conditions but a large number of training samples are required to represent extreme illumination condition.

In holistic approaches, any geometrical feature of face image is not taken into account. A face image is considered as a point in a high-dimensional image space. Thus, linearly transformation of the face image is done into a low-dimensional subspace than extraction a feature vector is done in order to avoid computational complexities and to reduce redundant data. PCA [61] is most commonly used for dimensionality reduction and is known to retain intraclass variations due to changes in illumination. Experimental results have shown that leaving out the first three eigenfaces that corresponds to the three most significant eigenvalues could reduce the effect of variations in illumination [62].

However, PCA for dimensionality reduction leads to loss of useful information. An alternative approach to PCA based linear projection is Fisher's linear discriminant, or the linear discriminant analysis (LDA), which is used to maximize the ratio of the determinant of the interclass scatter to that of interclass scatter [62], [63]. Like generative approaches, holistic approaches also require a number of training samples from different conditions to identify faces in uncontrolled environments. A more common approach to address the effects of varying lighting conditions is to preprocess face biometric samples to normalize illumination before extracting facial features for identification. Normalization techniques can be applied to an image either globally or regionally. Widely used normalization techniques include histogram equalization (HE), histogram matching, gamma intensity correction, and quotient image.

It has been accepted that illumination normalization improves recognition rate [64], [65]and regional normalization have higher accuracy than global normalization Shan et al. [64] proposed a region-based approach to illumination normalization where an image is first partitioned into four regions. The selected normalization technique (e.g., HE) is applied to each region separately (i.e., RHE). However, the improvements depend on the extent of variation in illumination present between enrolled and test [7], [66]. In a recent study, Sellahewa and Jassim [9] show that normalizing well-lit face images could lead to a decrease in identification accuracy. Thus the need for a quality-based adaptive approach to illumination normalization has been highlighted prior to feature extraction. This work focuses on wavelet-based face recognition in the presence of varying illumination

Face Recognition procedure works with the results from the previous stages of interest point detection and extraction of the descriptor. The data in the neighbourhood of each interest point becomes very important at the recognition stage, as one can claim to have a match between two images if and only if the images have sufficient similarity in the vicinity of two corresponding feature points, taking the rotation and scaling into the account. If these information is utilized effectively, then the expensive operation of the exhaustive comparison may be prevented.

CHAPTER 3

PROPOSED IMAGE PRE-PROCESSING METHODOLOGY

3.1 Introduction

The unpredictable change in lighting conditions is a problem in facial recognition. Therefore, it is desirable to normalize the photometric information in terms of light correction to optimize the facial recognition. Histogram equalisation has been used for lighting correction and briefly explained. Histogram equalization (HE) can be used as a simple but very robust way to obtain light correction when applied to small regions such as faces. The aim of HE is to maximize the contrast of an input image, resulting in a histogram of the output image which is as close to a uniform histogram as possible. For the purpose of adaptive normalisation, histogram equalization has been applied to each region of 16×16 pixel window of input image as per the thresholding of image quality measure.

3.2 Basic approach to image pre-processing

The work proposes a quality based adaptive approach for automated face recognition.

The proposed approach performs a three step process:

1. An objective measure of illumination quality of a given face image is done to decide if the image should be pre-processed to normalize its illumination.

2. The global quantity based normalization scheme is extended to a quality based approach to adaptive illumination normalization.

The adaptive regional normalised image has been used for face recognition using SIFT algorithm.

Real-time computation of a quantitative objective image quality measure is an essential tool for biometric-based identification applications. These measures act as a quality control to accept, reject, or reacquire biometric samples, as quality-based processing to select a biometric algorithm, and/or system parameters, and as confidence estimators of reliability of decision.

3.3 Image Quality Measure

The use of an image quality measure as a base for an adaptive approach to face recognition in the presence of varying illumination is investigated. Image quality measures acts as figure of merit used for the evaluation of imaging systems or of processing techniques. A good objective quality measure should reflect the distortion on the image well due to, for example, blurring, lighting noise, and compression. These measures could be instrumental in predicting the performance of vision-based algorithms such as feature extraction, image-based measurements, detection, tracking, and segmentation, etc., tasks. In the image based processing, the most frequently used measures are deviations between the reference and test images with varieties of the mean square error or signal to noise ratio being the most common measures. The reasons for their widespread popularity are their mathematical tractability and the fact that it is often straightforward to design systems that minimize the Mean Square Error. Luminance distortion factor of the given face image defines the illumination quality in comparison to a known reference image. The mathematically defined quality measure proposed by Wang and Bovik [69], i.e., the universal image quality index (Q), incorporates all the necessary required information. The Q provides meaningful comparisons across different types of image distortions by modelling any image distortion as a combination of the following three factors:

1) Loss of correlation;
2) Luminance distortion; and
3) Contrast distortion.

Here, the luminance distortion factor of Q is used to measure global or regional illumination quality of images. This will be called the luminance quality (LQ) index.

3.3.1 Universal Quality Index

Let $x = \{xi|i = 1, 2, . . .,N\}$ and $y = \{yi|i = 1, 2, . . . , N\}$ be the reference and the test images, respectively. The universal quality index in [69] is defined as

$$\mathrm{Q} = \frac{4\sigma_{xy}\bar{x}\bar{y}}{(\sigma_x^2+\sigma_y^2)[(\bar{x})^2+(\bar{y})^2]} \tag{4}$$

where

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

$$\bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i$$

$$\sigma_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2$$

$$\sigma_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{y})^2$$

$$\sigma_{xy} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})$$

Statistical features in (1) are measured individually to represent space-variant nature of image quality and then combined together to a single quality measure Q for the entire image.

A local quality index *Qj* is calculated by sliding a window of size $B \times B$ pixel by pixel from the top-left corner until the window reaches the bottom-right corner of the image. For a total of *M* steps, the overall quality index is given by

$$\mathrm{Q} = \frac{1}{M}\sum_{j=1}^{M} Q_j$$

3.2.2 Global LQ (GLQ) and Region LQ (RLQ) Indexes

The universal quality index *Q* can be written as a product of three components, i.e.,

$$\mathrm{Q} = \frac{\sigma_{xy}}{\sigma_x\sigma_y} \cdot \frac{2\bar{x}\bar{y}}{(\bar{x})^2+(\bar{y})^2} \cdot \frac{2\sigma_x\sigma_y}{\sigma_x^2+\sigma_y^2} \tag{5}$$

LQ, which is the luminance distortion factor in Q, is defined as

$$LQ = \frac{2\bar{x}\bar{y}}{(\bar{x})^2+(\bar{y})^2} \qquad (6)$$

LQ have the value range of [0, 1], LQ measures the closeness of mean luminance between x and y. LQ equals 1, if and only if $\bar{x} = \bar{y}$. The window size used in this work is the default 16 × 16 pixels. GLQ is calculated similarly as the calculation of a single Q value in (4). RLQ represents the LQ of a 16 × 16 pixel region of an image resulting. The LQ of a region of an image is calculated by partitioning the local quality.

3.4 Normalization

Normalization technique is applied to the image, which can effectively and efficiently eliminate the effect of uneven illumination. Well known contrast enhancement algorithms, such as histogram equalization, are global methods used for image normalisation for further processing. Shan *et al.* [64] investigated several illumination normalization methods and propose some novel solutions, such as Gamma Intensity Correction (GIC), Region-based strategy combining GIC with Histogram Equalization (HE) and Quotient Illumination Relighting (QIR) method. Chen *et al.* [71] employed a discrete cosine transform (DCT) to compensate for illumination variations in the logarithm domain.

3.4.1 Histogram equalization (HE)

Histogram: is the graphical representation of the frequency distribution of the elements in a digital image. It plots the number of pixels for each frequency value in the grey scale or colour. The horizontal axis of the graph represents the frequency variation while the vertical axis represents the number of pixels in the particular frequency. Complex fixed-point inputs are distributed according to their magnitude squared values; complex floating-point inputs are distributed by their normalized values.

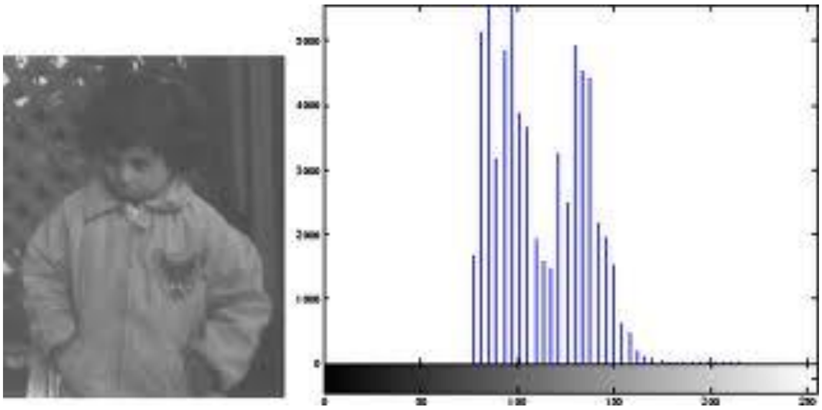

Fig. 3.1: Histogram of a Gray Scale Image.

Fig. 3.2: Histogram of Colour Image in RGB Colour Space.

Histogram equalization is a method in image processing of contrast adjustment using the histogram of the image. This method usually increases the global contrast of many images by close contrast values which leads to better distribution of intensities on histogram, as Fig. 3.3 shows a better distribution of components of histogram after applying the transform. The local area with less contrast values gains the value. This is performed by effectively spreading out the most frequent intensity values. This method of image normalisation performs best under conditions with bright or dark backgrounds and foregrounds. It gives a better view of X-ray images and other under or over illuminated images.

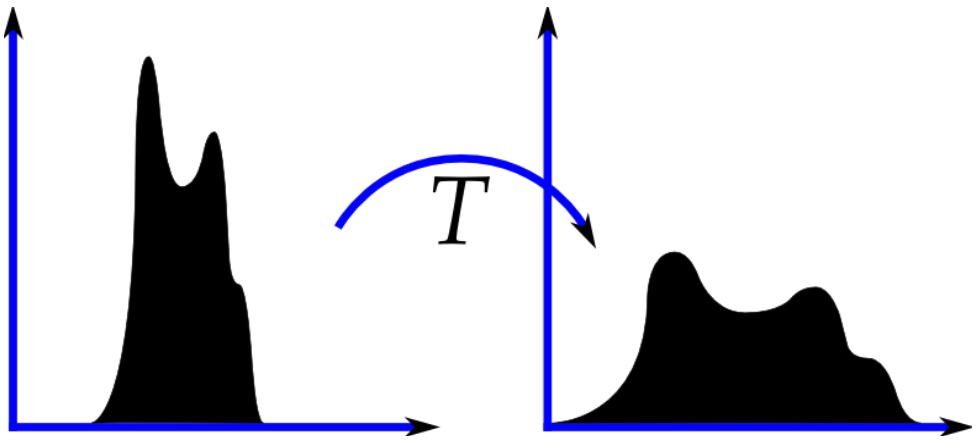

Fig. 3.3: Pictorial Representation of Histogram Equalization.

Histogram equalisation can be easily applied to grey scale images but in case of RGB colour space images it may lead to dramatic changes in colour balance of image since the relative distribution of colour channels change by applying the algorithm. So, the RGB colour image should first undergo change in colour space e.g. LAB colour space or HSV colour space. Thus the algorithm can be applied to the luminance or the value channel without resulting in change of hue or saturation.

Fig. 3.4 shows the histogram equalization of grey scale Fig. 3.4(a) and RGB colour image Fig. 3.4(b).

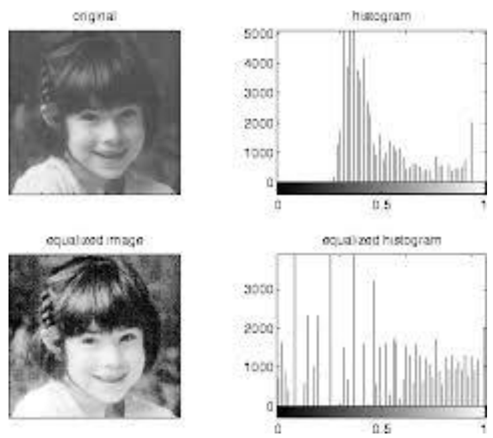

Fig. 3.4(a): Histogram Equalization of Gray Scale Image.

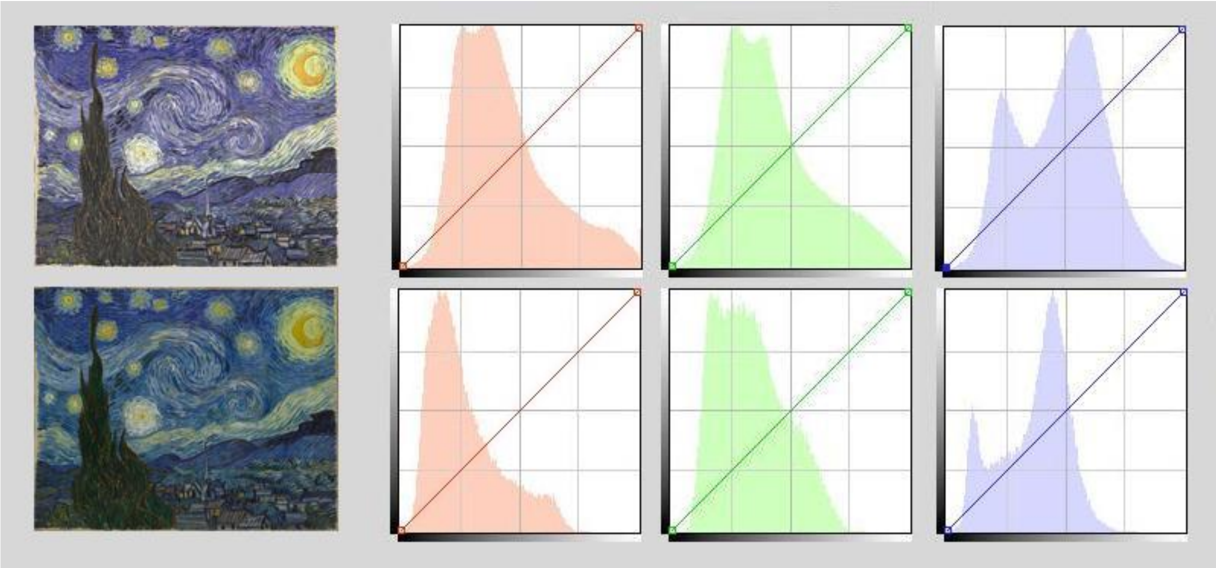

Fig. 3.4(b): Histogram Equalization of RGB Image.

3.4.2 Image-Quality-Based Adaptive Normalization

The proposed method works by first calculating the GLQ of a given image and then normalizing only if it's GLQ is less than a predefined threshold for the purpose of image-quality-based adaptive normalization. Inspired by the work of Shan *et al.* [64] and the fact that the direction of light source leads to regional variation in image quality of an image, the global quality based adaptive normalization is extended by introducing a region quality-based adaptive approach to normalization. A region of an image e.g. 16×16 pixel window, is normalized only if the RLQ score is lower than a predefined threshold. The commonly used histogram equalization (HE) is adopted here for illumination normalization. Hence, the two proposed approaches to adaptive normalization will be referred to as global quality-based HE (GQbHE) and regional quality-based HE (RQbHE). The threshold can be determined empirically, here it is determined by taking the GLQ and RLQ values of the database images under consideration.

Chapter 4

FEATURE EXTRACTION METHODOLGY

4.1 Basic approach

The proposed face recognition technique in this work comprises of image quality measure based adaptive normalisation technique which has been discussed the previous section. Further the facial features are extracted from the image and by the means of an efficient classifier used for the purpose of matching in face recognition.

A normalised image has been produced at the output of image pre-processing block by regional histogram equalization of each 16×16 pixel window region of the input image. Normalization of each region has been done by thresholding the RLQ value for the particular region as discussed earlier. This normalized image has been given as input to feature extraction block of generic face recognition system. For the purpose of getting higher efficiency of the system scale invariant feature extraction has been opted and the task is performed by SIFT.

4.2 The Scale-invariant Feature Transform(SIFT)

SIFT is an algorithm in computer vision to detect and describe local features in images. The algorithm was published by David Lowe in 1999. It find applications in object recognition, robotic mapping and navigation, image stitching, 3D modelling, gesture recognition, video tracking, individual identification of wildlife and match moving.

SIFT algorithm, which according to [71] consists of four computational stages:

(i) Scale-space extrema detection,

(ii) Removal of unreliable key-points,

(iii) Orientation assignment

(iv) Key-point descriptor calculation.

4.2.1 Scale-space extrema detection

In the first stage, interest points called key-points, are identified in the scalespace by looking for image locations that represent maxima or minima of the difference-of-Gaussian function. The scale space of an image is defined as a function $L(x, y, \sigma)$, that is produced from the convolution of a variable-scale Gaussian, $G(x, y, \sigma)$, with the input image, $I(x, y)$:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y), \qquad (7)$$

With

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \qquad (8)$$

Where σ denotes the standard deviation of the Gaussian $G(x, y, \sigma)$.

The difference-of-Gaussian function $D(x, y, \sigma)$ can be computed from the difference of Gaussians of two scales that are separated by a factor k:

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \qquad (9)$$

Local maxima and minima of $D(x, y, \sigma)$ are computed based on the comparison of the sample point and its eight neighbours in the current image as well as the nine neighbours in the scale above and below. If the pixel represents a local maximum or minimum, it is selected as a candidate key-point.

4.2.2 Removal of unreliable key-points

The final key-points are selected based on measures of their stability. During this stage low contrast points (sensitive to noise) and poorly localized points along edges (unstable) are discarded. Two criteria are used for the detection of unreliable key-points. The first criterion evaluates the value of $|D(x, y, \sigma)|$ at each candidate key-point. If the value is below some threshold, which means that the structure has low contrast, the key-point is removed. The second criterion evaluates the ratio of principal curvatures of each candidate key-point to search for poorly defined peaks in the Difference-of-Gaussian function. For key-points with high edge responses, the principal curvature across the edge will be much larger than the principal curvature along it. Hence, to remove unstable edge key-points based on the second criterion, the ratio of principal curvatures of each candidate key-point is checked. If the ratio is below some threshold, the key-point is kept, otherwise it is removed.

4.2.3 Orientation assignment

An orientation is assigned to each key-point by building a histogram of gradient orientations $\theta(x, y)$ weighted by the gradient magnitudes $m(x, y)$ from the key-points neighbourhood.

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))2^2} \quad (10)$$

$$\theta(x, y) = \tanh\ (L(x, y+1) - L(x, y-1))/(L(x+1, y) - L(x-1, y)); \quad (11)$$

Where L is a Gaussian smoothed image with a closest scale to that of a key-point. By assigning a consistent orientation to each key-point, the key-point descriptor can be represented relative to this orientation and, therefore, invariance to image rotation is achieved.

4.2.4 Key-point descriptor calculation

The key-point descriptor is created by first computing the gradient magnitude and orientation at each image point of the 16×16 key-point neighborhood (left side of Fig. 4.1). This neighbourhood is weighted by a Gaussian window and then accumulated into orientation histograms summarizing the contents over subregions of the neighborhood of size 4×4 (right side of Fig. 4.1), with the length of each arrow in Figure 4.1(right) corresponding to the sum of the gradient magnitudes near that direction within the region [4]. Each histogram contains 8 bins, therefore each key-point descriptor features 4 × 4 × 8 = 128 elements. The coordinates of the descriptor and the gradient orientations are rotated relative to the key-point orientation to achieve orientation invariance and the descriptor is normalized to enhance invariance to changes in illumination.

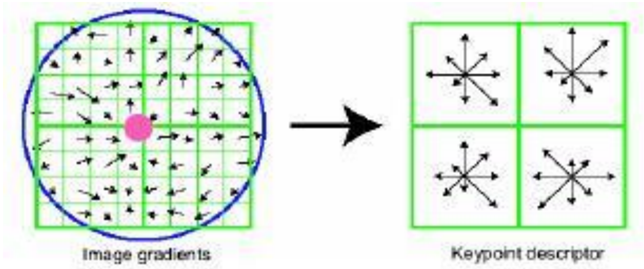

Fig. 4.1: Computation of 2×2 sub-regions from 8×8 neighbourhood

4.2.5 Matching

When using the SIFT algorithm for object recognition, each key-point descriptor extracted from the query (or test) image is matched independently to the database of descriptors extracted from all training images. The best match for each descriptor is found by identifying its nearest neighbour (closest descriptor) in the database of key-point descriptors from the training images. Generally, many features from a test image do not have any correct match in the training database, because they were either not detected in the training image or they arose from background clutter. To discard key-points whose descriptors do not have any good match in the training database, a subsequent threshold is used, which rejects matches that are too ambiguous. If the distance ratio between the closest neighbour and the second-closest neighbour, (i.e., the closest neighbour that is known to come from a different object than the first) is below some threshold, than the match is kept, otherwise the match is rejected and the key-point is removed. The object in the database with the largest number of matching points is considered the matched object, and is used for the classification of the object in the test image.

4.3 SIFT-based Face Recognition

Over the past few years there have been some studies (from the early studies, e.g., [72], [73] to more recent ones, such as [78]) assessing the feasibility of the SIFT approach for face recognition. One of the first attempts to use the SIFT algorithm for face recognition was presented in [72]. Selected key point features in an face image by SIFT have been shown below in Fig 4.2.

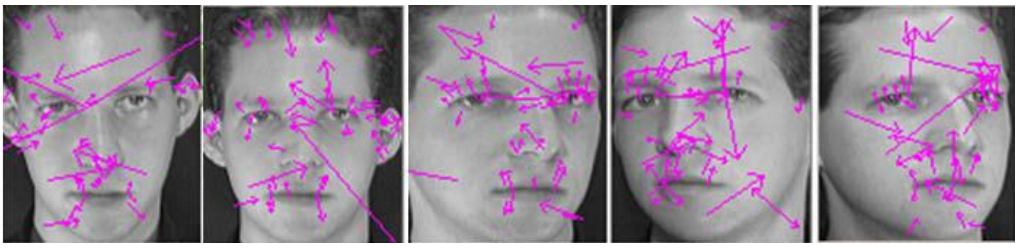

Fig. 4.2: Feature Extraction using SIFT of a human face with different pose.

The algorithm used here, differs from original SIFT algorithm in the implementation of the matching stage. Each SIFT descriptor in the test image is matched with every descriptor in each training image. Matching is done using a distance based criterion. A descriptor from the test image is said to match a descriptor from the training image, if the distance between the 2 descriptors is less than a specific fraction of the distance to the next nearest descriptor. The problem with this method is that it is very time consuming. Matching between two images has a computational complexity of $O\ (n^2)$, where n is the average number of SIFT descriptors in each image. In [73], the original SIFT algorithm is rendered more robust by following one of two strategies that aim at imposing local constraints on the matching procedure: the first matches only SIFT descriptors extracted from image-windows corresponding to the mouth and the two eyes, while the second relies on grid based matching, Local matching, i.e. within a grid or a cluster, constrains the SIFT features to match features from nearby areas only. Local matching also reduces the computational complexity linearly. The computational complexity required for matching a pair of images by a local method is $O\ (n^2/s)$, where s is the number of grids or clusters. As seen from Fig. 4.3, where the basic SIFT algorithm from [71] was used to match the SIFT descriptors, there are some key-points matched, that do not represent the same characteristic of the face. Although we would expect the distance between such key-points to be high, since they correspond to different regions of the faces, this is clearly not the case. Therefore better results are achieved, if certain subsets of SIFT key-points are used for matching and only (spatially) corresponding subsets of SIFT descriptors are matched (as is [73] and later in [74], [75], [76] and [77]).

Fig. 4.3: Matching of Test Images (bottom) with Training faces (top) using Basic SIFT.

Both local and global information for face recognition are used in [74]. Instead of using a grid based approach, the SIFT features are clustered into 5 clusters using k means clustering (2 clusters for the eyes, one for the nose, and 2 clusters at the edges of the mouth). Only the SIFT descriptors between two corresponding clusters are matched. This ensures that matching is done locally. As a global matching criterion, the total number of descriptor matches (as in [71]) is used.

In [79] SIFT features are extracted from the frontal and half left and right profiles. An augmented set of SIFT features is then formed from the fusion of features from the frontal and side profiles of an individual, after removing feature redundancy. SIFT feature sets from the database and query images are matched using the Euclidean distance and Point pattern matching techniques. In [75] a Graph Matching Technique is employed on the SIFT descriptors to deal with false pair assignment and reduce the number of SIFT features. In [76] SIFT features are ranked according to a discriminative criterion based on Fisher's Discriminant Analysis, so that the chosen features have the minimum within-class variation and maximum variation between classes. In [77] both global and local matching strategies are used. In order to reduce the identification errors, the Dempster-Shafer decision theory is applied to fuse the two matching techniques.

4.4 Comparison of SIFT features with other local features

There has been an extensive study done on the performance evaluation of different local descriptors, including SIFT, using a range of detectors. The main results are summarized below:

- SIFT and SIFT-like GLOH features exhibit the highest matching accuracies (recall rates) for an affine transformation of 50 degrees. After this transformation limit, results start to become unreliable.

- Distinctiveness of descriptors is measured by summing the eigenvalues of the descriptors, obtained by the Principal components analysis of the descriptors normalized by their variance. This corresponds to the amount of variance captured by different descriptors, therefore, to their distinctiveness. PCA-SIFT (Principal Components Analysis applied to SIFT descriptors), GLOH and SIFT features give the highest values.

- SIFT-based descriptors outperform other contemporary local descriptors on both textured and structured scenes, with the difference in performance larger on the textured scene.

- For scale changes in the range 2-2.5 and image rotations in the range 30 to 45 degrees, SIFT and SIFT-based descriptors again outperform other contemporary local descriptors with both textured and structured scene content.

- Introduction of blur affects all local descriptors, especially those based on edges, like shape context, because edges disappear in the case of a strong blur. But GLOH, PCA-SIFT and SIFT still performed better than the others. This is also true for evaluation in the case of illumination changes.

The evaluations carried out suggests strongly that SIFT-based descriptors, which are region-based, are the most robust and distinctive, and are therefore best suited for feature matching.

4.5 K nearest neighbour Classifier

Feature extraction method might introduce some distortion or addition of noise, which is unavoidable. In such cases, some disparity between extracted and enrolled feature vectors of the same person might occur. The above-mentioned system expects to match feature vectors exactly. If an exact match is not found it fails to recognize. So, we assume the feature extraction process to be perfect or propose some improvement thus for the purpose of indexing and matching of features K nearest neighbour classifier has been used in proposed work. If there is no exact match, it is possible to know what the nearest matches are. We used K-nearest neighbour classifier [9] to achieve that objective, as it is simple and effective [80].

Given an unknown sample, a k-nearest neighbour classifier searches the pattern space for the k training samples that are closest to the unknown sample. These k training samples are the k “nearest neighbours” of the unknown sample [81].

“Closeness” is defined in terms of Euclidean distance, where Euclidean distance between two feature vectors, X = (x1, x2, …, xn) and Y = (y1, y2,…, yn) is as in equation (11)

$$d_2(X,Y) = \sqrt{\sum_{i=1}^{n-1}(x_i - y_i)^2} \qquad (12)$$

It can be generalized to Minkowski similarity function, equation (12)

$$d_q(X,Y) = \sqrt[q]{\sum_{i=1}^{n-1} w_i |\, x_i - y_i|^q} \qquad (13)$$

If q=1, it gives the Manhattan distance as given in equation (13)

$$d_1(X,Y) = \sum_{i=1}^{n-1}|x_i - y_i| \qquad (14)$$

Any selected distance metric can be used to find the k nearest neighbours. It is fast and efficient technique for used matching.

Chapter 5

EXPERIMENTAL RESULTS

The ORL Database: has been used for testing the algorithm. The database consists of 10 different images of 40 different subjects. The images were taken at different times, varying the lighting conditions, facial expressions (eyes opened/closed, smiling/not smiling) and different facial details (wearing glass/ not wearing glass). All the images were taken against a dark homogeneous background with the subjects in an upright, frontal position (with tolerance for some side movement). The size of each image is 92*112 pixels, with 256 gray levels per pixels. Some of the faces of ORL Database are shown below:-

Fig. 5.1: ORL Database.

We can observe that each subject has 10 different orientation. All the images were taken against a dark homogenous background with the subjects in an upright, frontal position, with tolerance for some tilting and rotation of up to about 20 degrees. All the images were taken against a dark homogenous background with the subjects in an upright, frontal position, with tolerance for some tilting and rotation of up to about 20 degrees.

Input Image: The calculation of the LQ index for a given input image relies on the selection of window size, here we had selected 16×16 pixel size.

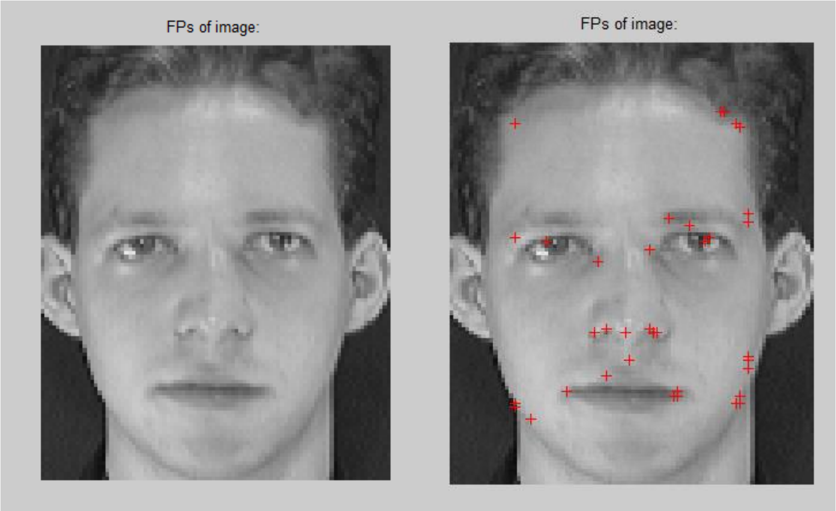

Fig. 5.2: (a) Training Image of Subject1, (b) Feature Point Extraction.

The above figure shows the extraction of various feature points in an image. The feature points are found by Harris Corner Detection.

Normalisation: The LQ index is evaluated using images from the ORL database, which consists of well-lit face images. However, variations in pose and face size are a characteristic of the images in this database. To demonstrate the appropriateness of LQ index for our purposes, we calculate it over the window of 16×16 pixel and normalize the window by using regional histogram equalization technique. A sensible threshold that could distinguish between images with good illumination and images with poor illumination is an LQ score of approximately 0.8.

Previously GLQ index as an objective measure of illumination quality of face images. However, in real-life scenarios, variations in illumination between database and test input images could be confined to a region of the face image due to the changes in the direction of the light source or pose. Therefore, it is sensible to measure the illumination quality on a region-by-region basis as shown in Fig. 5.3.

.

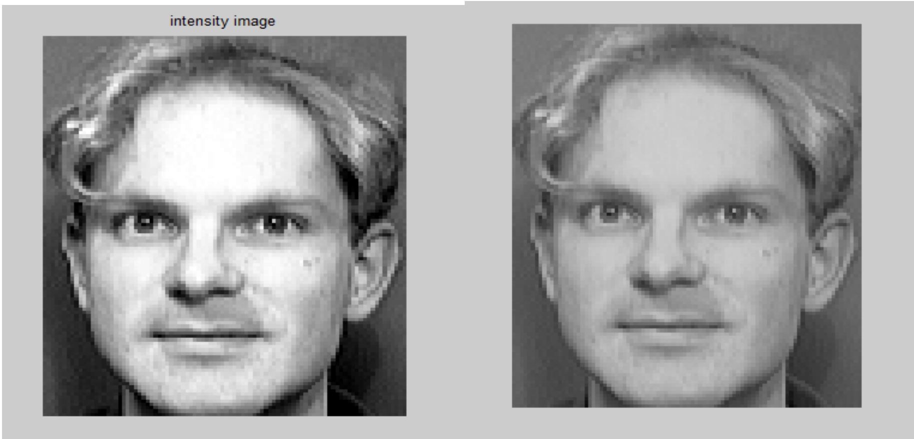

Fig.5.3: (a) Illuminated Test Image, (b) Normalized Test Image

After adaptive normalization of the input image, SIFT has been used for feature extraction of the pre-processed image. Scale invariant features have been extracted and these features are used for matching with low computational complexity. Further matching of features with the database images has been done by K nearest neighbour classifier and the results of matching has are shown in Fig. 5.4.

Fig.5.4: (a) Feature Points of Normalized Image, (b) User Matched Image

The above figures shows the different steps of Face Recognition. It shows that the test image matches with Subject No.18.

Using image quality measure for adaptive normalization and further using the image for SIFT reduces the error rate to 1.52 and increases the recognition rate to 96%.

Chapter 6

CONCLUSION

The system performs adaptive illumination invariant face recognition to achieve a good recognition rate and is applicable to any face image. It is reported to have a good efficiency as in real-life scenarios, variations in illumination in images could be confined to a region of the face image due to the changes in the direction of the light source or pose. Therefore, it is sensible to measure the illumination quality on a region-by-region basis instead on global basis. The analysis demonstrated that the RLQ measure identifies individual regions that have either good or poor illumination quality, thus it is a better representation of the illumination quality of a face image than the GLQ. Compared to the traditional use of HE, the proposed GQbHE further decreased the overall identification error by a further 1–2% across different feature representations. The proposed RQbHE further reduced the identification error. Use of SIFT leads to scale invariant feature extraction thus makes the system more accurate. Limitations to the method basically comprises of pose variation and occlusion. These create problem in feature extraction and matching thus leading to low recognition rate and reduced efficiency of the approach.

In this work, we present a image quality based adaptive approaches to face recognition. The effects of illumination variation on face recognition were found and analysed. An objective image quality measure which is luminance component of universal quality index has been selected for comparison of test and reference image. This measure is called the LQ index and thresholding of this measure to the value 0.8 leads to region quality based adaptive illumination normalization. This approach was applied on ORL database and the effectiveness over the previous approaches of normalization was demonstrated. Finally SIFT base recognition scheme was implemented on the image quality based approach of pre-processing. We have also presented a short review of certain latest and primitive face recognition techniques and some of the evaluation parameters are considered for comparing techniques with each other.

Table below shows the comparative analysis of these techniques:

TECHNIQUE	**APPROACH**	**DATABASE**	**RECOGNITION RATE(%)**
PCA	Holistic approach	ORL	90.56%
Artificial Neural Network	Feature based approach	CMU	90.10%
SVM	Holistic approach	ORL	95% (1® polynomial kernel)
HMM	Feature based approach	MIT	90%
MPCA+LPP	Holistic approach	FERET	96.5%
		ORL	96.5%
By using texture and depth information	Hybrid approach	FRAV3D	88.96% (2D+3D)

Table 1: Experimental results of some face recognition techniques.

Performance Measure:

METHODS	EQUAL ERROR RATE
Fisher face	4.58
Hankan’s	2.61
SIFT+SVM	2.58
Our Method	1.52

METHODS	RECOGNITION RATE
PCA	92.1
ICA	91.6
FISHER	92.8
2D_PCA	92.5
Our Method	96

Table 6.2: Experimental results with respect to other methods.

6.1 Future Work

Our future work will investigate other challenges to automated recognition system which effect the face image quality such as facial expression, pose, and occlusion etc. Certain other objective quality measures are to be used for a fully adaptive face recognition system, which will be able to perform better image pre-processing, face feature representation and classification. More suitable objective quality measure can be opted so that it can lead to a more efficient system.

Publications

Kritika Choudhary, Nidhi Goel; "**A Review on Face Recognition Techniques**", *Proc. SPIE* **8760**, International Conference on Communication and Electronics System Design, 87601E (January 28, 2013); doi:10.1117/12.2012238.

References

[1] Francis Galton, "Personal identification and description," *In Nature*,pp. 173-177, June 21, 1888.

[2] W. Zaho, "Robust image based 3D face recognition," Ph.D. Thesis,Maryland University, 1999.

[3] R.C. Flower. FINGERPRINT: an old touchstone decriminalized. *IEEE spectrum*, page 26, 1994.

[4] J.E. Siedlarz. IRIS: more detailed than a fingerprint. *IEEE spectrum*, page 27, 1994.

[5] J.G. Daugman. Two dimensional spectral analysis of cortical receotive field profile. Vision research, 20:847-856, 1980.

[6] R. Mandelbaum. SPEECH: just say the word. *IEEE Spectrum*, page 30, 1994.

[7] M. Yang, D. Kriegman, and N. Ahuja, "*Detecting Faces in Images: A Survey*," IEEE Transaction Pattern Analysis and Machine Intelligence, vol. 24, no. 1, pp. 34-58, 2002.

[8] E. Hjelmas, and B.K. Low, "*Face Detection: A Survey*," Computer Vision and Image Understanding, vol. 83, pp. 236-274, 2001.

[9] Mohammad Kabir Hossain, Abu Ahmed Sayeem Reaz, Rajibul Alam, and Dr.William Perrizo, "Automatic Face Recognition system using P tree and K-nearest neighbor classifier".

[10] M. Kirby, L. Sirovich, "Application of Karhunen–Loeve procedure for the characterization of human faces", IEEE Transactions on Pattern Analysis and Machine Intelligence vol. 12, pp.103–108, 1990.

[11] K.K. Sung and T. Poggio, "Learning human face detection in cluttered scenes," Computer Analysis of Image and patterns, pp. 432-439, 1995.

[12] Z. Pan, R. Adams, and H. Bolouri, "Image redundancy reduction for neural network classification using discrete cosine transforms," in Proc. IEEE-INNS-ENNS Int. Joint Conf. Neural Networks, vol. 3, Como, Italy, 2000, pp. 149–154.

[13] M. Turk and A. Pentland. Eigenfaces for recognition. Journal of Cog-nitive Neurosicence, 3(1):71–86, 1991.

[14] P. Belhumeur, P. Hespanha, and D. Kriegman, "Eigenfaces vs fisherfaces: Recognition using class specific linear projection," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 711–720, 1997.

[15] Beveridge, J. R., She, K, Draper, B.A., and Givens, G.H. 2001, A nonparametric statisical comparison of principal component and linear discriminant subspaces for face recognition. In Proceedings, IEEE Conference on Computer Vision and Pattern Recognition.

[16] Bartlett, M. S., Lades, H. M., and Sejnowski, T. 1998. Independent component representation for face recognition. In Proceedings, SPIE Symposium on Electronic Imaging: Science and Technology. 528–539.

[17] T. Kohonen. Self-organization and associative memory. Springer-Verlag, Berlin, 1989.

[18] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, “Eigenfaces versus fisherfaces: Recognition using class specific linear projection,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 19, no. 7, pp. 711–720, Jul. 1997.

[19] R. Lotlikar and R. Kothari, “Fractional-step dimensionality reduction,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 6, pp. 623–627, Jun. 2000.

[20] H. Yu and J. Yang, “A direct LDA algorithm for high-dimensional data—With application to face recognition,” Pattern Recognit., vol. 34, pp. 2067–2070, 2001. 690 IEEE TRANSACTIONS ON NEURAL NETWORKS, VOL. 16, NO. 3, MAY 2005.

[21] J. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, “Face recognition using LDA-based algorithms,” IEEE Trans. Neural Netw., vol. 14, no. 1, pp. 195–200, Jan. 2003.

[22] “Face recognition using Kernel direct discriminant analysis algorithms,” IEEE Trans. Neural Network., vol. 14, no. 1, pp. 195–200, Jan. 2003.

[23] Z. M. Hafed and M. D. Levine, “Face recognition using the discrete cosine transform,” Int. J. Comput. Vis., vol. 43, no. 3, pp. 167–188, 2001.

[24] C. M. Bishop, Neural Networks for Pattern Recognition. New York: Oxford Univ. Press, 1995.

[25] R. Chellappa, C. L.Wilson, and S. Sirohey, “Human and machine recognition of faces: A survey,” Proc. IEEE, vol.

[26] D. Valentin, H. Abdi, A. J. O’Toole, and G. W. Cottrell, “Connectionist models of face processing: A survey,” Pattern Recognit., vol. 27, pp.1209–1230, 1994.

[27] F.S.Samaria, "Face Recognition using Hidden Markov Models", PhD thesis, University of Cambridge ,1994.

[28] A. S. Tolba, A.H. El-Baz, and A.A. El-Harby. "Face Recognition: A Literature Review", International Journal of Information and Communication Engineering 2:2 2006.

[29] A.V.nefian and M.H.HIII, "Face Detection and Recognition Using Hidden Markove Models" proc.IEEE Int'l conf.Image Processing, vol, pp.141-145, 1998.

[30] T.J. Stonham, “Practical face recognition and verification with WISARD,” Aspects of Face Processing, pp. 426-441, 1984.

[31] Shatha K. Jawad, "Design a Facial Recognition System Using Multilayer Perceptron and Probabilistic Neural Networks Based Geometrics 3D Facial", European Journal of Scientific Research ISSN 1450-216X Vol.60 No.1 (2011), pp. 95-104.

[32] S. Lawrence, C.L. Giles, A.C. Tsoi, and A.D. Back, “Face recognition: A convolutional neural-network approach,” IEEE Trans. Neural Networks, vol. 8, pp. 98-113, 1997.

[33] E. Osuna, R. Freund, and F. Girosi. Training support vector machines: An application to face detection. Proceedings of the IEEE Conf. Com-puter Vision and Pattern Recognition, pages 130–136, June 1997.

[34] V.N. Vapnik, “The nature of statistical learning theory,” New York: Springverlag, 1995.

[35] L.R.Rabiner and B.H.Jung, "Fundamental of speech Recognition", Printice Hall, 1993.

[36] F. Samaria and A.C. Harter, “Parameterisation of a stochastic model for human face identification,” Proc. Second IEEE Workshop Applications of Computer Vision, 1994.

[37] J.Shermina and V.Vasudevan, "An Efficient Face Recognition System based on Fusion of MPCA and LPP", American Journal of Scientific Research ISSN 1450-223X Issue 11(2010), pp.6-19.

[38] X. He and P. Niyogi, “Locality preserving projections”, in Advances in Neural Information Processing Systems. Cambridge, MA: MIT Press, 2003.

[39] Fan R.K. Chung, “Spectral Graph Theory”, Regional Conference Series in Mathematics, number 92, 1997.

[40] S. Yan, D. Xu, B. Zhang, H. Zhang, Graph embedding: a general framework for dimensionality reduction, CVPR 2 (2005) 830–837.

[41] Minghua Zhao, Zhisheng You, Yonggang Zhao and Zhifang Liu, “Face Recognition Method Combining 3D Face Model with 2D Recognition,” in Proc. Fourth International Conference on Image and Graphics (ICIG2007), pp.655-660, 2007.

[42] A. Bossaghzade, M. Doostari and A. Behrad, ” Face recognition using non frontal face image”, in Proc. 17th Iranian Conference Electrical Engineering, pp. 196-201,2009. Published in Farsi.

[43] David G. Lowe, “Distinctive Image Feature from Scale Invariant Features," International Journal of computer Vition, vol. 60, no. 2, pp.91-110, 2004.

[44] F. Samaria and F. Fallside, “Face identification and feature extraction using hidden markov models,” Image Processing: Theory and Application, G. Vernazza, ed., Elsevier, 1993.

[45] Atefe Assadi and Alireza Behrad, " New Method for Human Face Recognition using Texture and Depth Information", in 10tb symposium on neural network applications in electrical engineering, NEUREL-2010, 2010 IEEE.

[46] G. Finlayson, S. Hordley, G. Schaefer, and G. Y. Tian. Illuminant and device invariant colour using histogram equalisation. Pattern Recognition, 38, 2005.

[47] K. Pearson. On lines and planes of closest fit to systems of points in space. Philosophical Magazine, 2(6):559–572, 1901.

[48] S.Watanabe. Karhunen-loeve expansion and factor analysis theoretical remarks and applications. In Proc. 4th Prague Conference on Infor-mation Theory, 1965.

[49] L. Sirovich and M. Kirby. Low-dimensional procedure for the characterization of human faces. Journal of the Optical Society of America A - Optics, Image Science and Vision, 4(3):519–524, March 1987.

[50] M. Turk and A. Pentland. Eigenfaces for recognition. Journal of Cog- nitive Neurosicence, 3(1):71–86, 1991.

[51] R. Brunelli and T. Poggio, “Face recognition: Feature vs. templates,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 10, pp. 1042–1053, Oct. 1993.

[52] L. Wiskott, J.-M. Fellous, N. Krüger, and C. von der Malsburg, “Face recognition by elastic bunch graph matching,” in *Proc. 7th Int. Conf. CAIP*, G. Sommer, K. Daniilidis, and J. Pauli, Eds., Kiel, Germany, 1997, vol. 1296, pp. 456–463.

[53] A. S. Georghiades, D. J. Kriegman, and P. N. Belhumeur, “Illumination cones for recognition under variable lighting: faces,” in *Proc. IEEE Comput. Vis. Pattern Recog.*, 1998, pp. 52–58.

[54] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman, “From few to many: Illumination cone models for face recognition under variable lighting and pose,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 6, pp. 643–660, Jun. 2001.

[55] T. Okabe and Y. Sato, “Object recognition based on photometric alignment using RANSAC,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Jun. 2003, vol. 1, pp. 221–228.

[56] K.-C. Lee, J. Ho, and D. J. Kriegman, “Acquiring linear subspaces for face recognition under variable lighting,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 5, pp. 684–698, May 2005.

[57] P. W. Hallinan, “A low-dimensional representation of human faces for arbitrary lighting conditions,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Jun. 1994, pp. 995–999.

[58] A. Shashua, “On photometric issues in 3D visual recognition from a single 2D image,” *Int. J. Comput. Vis.*, vol. 21, no. 1/2, pp. 99–122, Jan. 1997.

[59] P. N. Belhumeur and D. J. Kriegman, “What is the set of images of an object under all possible lighting conditions?” in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 1996, pp. 270–277.

[60] M. Turk and A. Pentland, “Eigenfaces for recognition,” *J. Cogn. Neurosci.*, vol. 3, no. 1, pp. 71–86, 1991.

[61] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, “Eigenfaces vs. fisherfaces: Recognition using class specific linear projection,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, Jul. 1997.

[62] K. Etemad and R. Chellappa, “Discriminant analysis for recognition of human face images,” *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 14, no. 8, pp. 1724–1733, Aug. 1997.

[63] S. Shan, W. Gao, B. Cao, and D. Zhao, “Illumination normalization for robust face recognition against varying lighting conditions,” in *Proc. IEEE Int. Workshop Anal. Model. Faces Gestures*, 2003, pp. 157–164.

[64] R. Gross and V. Brajovi, “An image preprocessing algorithm for illumination invariant face recognition,” in *Proc. 4th Int. Conf. AVBPA Audio-, Video-Based Biometric Person Authentication*, Jun. 2003, pp. 11–18.

[65] H. Sellahewa and S. Jassim, “Face recognition in the presence of expression and/or illumination variation,” in *Proc. 4th IEEE Workshop Autom. Identification Adv. Technol. AutoID,* Oct. 2005, pp. 144–148.

[66] H. Sellahewa, “Wavelet-based automatic face recognition for constrained devices,” Ph.D. dissertation, Univ. Buckingham, Buckingham, U.K., 2006.

[67] H. Sellahewa and S. A. Jassim, “Illumination and expression invariant face recognition: Toward sample quality-based adaptive fusion,” in *Proc. 2nd IEEE Int. Conf. Biometrics, Theory, Appl. Syst.*, Sep. 2008, pp. 1–6.

[68] Z.Wang and A. C. Bovik, “A universal image quality index,” *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.

[69] W. Chen, M. J. Er and S. Wu, “Illumination Compensation and Normalization Using Logarithm and Discret Cosine Transform”, *Proc. of IEEE Internacional Conference on Control, Automation, Robotics and Vision*, vol. 1, 2004, pp. 380-385.

[70] Lowe D.G.: Distinctive Image Features From Scale-Invariant Key-points, International Journal of Computer Vision, Vol. 60, pp. 91–110, 2004.

[71] Aly M.: Face Recognition using SIFT Features, CNS/Bi/EE report 186, 2006.

[72] Bicego M., Lagorio A., Grosso E., Tistarelli M.: On the Use of SIFT Features for Face Authentication, CVPR Workshop, pp. 35–35, 2006.

[73] Jun Luo, Ma Y., Takikawa E., Lao S., Kawade M., Bao-Liang Lu: Person-Specific SIFT Features for Face Recognition, ICASSP, pp. 593–596, 2007.

[74] Kisku D.R., Rattani A., Grosso E., Tistarelli M.: Face Identification by SIFT-based Complete Graph Topology, IEEE Workshop AIAT, pp. 63–68, 2007.

[75] Majumdar A., Ward R.K.: Discriminative SIFT Features for Face Recognition, Canadian Conference on Electrical and Computer Engineering, pp. 27–30, 2009.

[76] Kisku D.R., Tistarelli M., Sing J.K., Gupta P.: Face Recognition by Fusion of Local and Global Matching Scores Using DS Theory: An Evaluation With Uni-Classifier and Multi-Classifier Paradigm, CVPR Workshop, pp. 60–65, 2009.

[77] Geng C., Jiang X.: SIFT features for Face Recognition, IEEE Conference CSIT, pp. 598–602, 2009.

[78] Rattani A., Kisku D.R., Lagorio A., Tistarelli M.: Facial Template Synthesis based on SIFT Features, IEEE Workshop AIAT, pp. 69–73, 2007

[79] W. Perrizo, Q. Ding, and A. Denton, “Lazy classifiers using p-trees”. Available at: "http://www.citeseer.nj.nec.com/perrizo02lazy.html.

[80] J. Han and M. Kamber, “Data mining: concepts and techniques”,Ed., San Francisco, USA: Morgan Kaufmann Publishers, 2001, pp. 314-315.