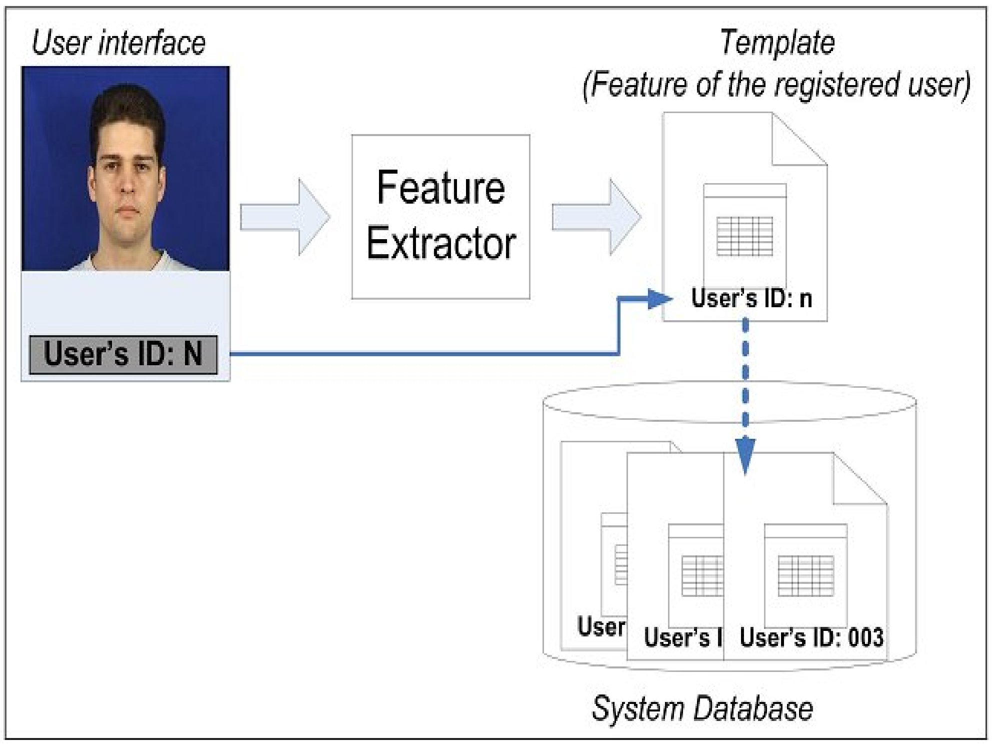
Chapter 1
Introduction

To date, the access to restricted systems has mostly been controlled by information-based or token-based security, such as passwords and identity cards. However, such security control can easily fail when a password is manipulated or a card is stolen. Furthermore, simple and short length passwords are easy to guess by a fraudulent user, while long and complex passwords may be hard to remember by a legitimate user. Therefore, the technologies of Biometric recognition are highly desired to address these problems. One of the biometric recognition modalities is face recognition which is non-intrusive, natural and easy to use. Therefore, it has a higher commercial value in the market. Now a days many commercial systems for face recognition are available. They have been summarised in [1].

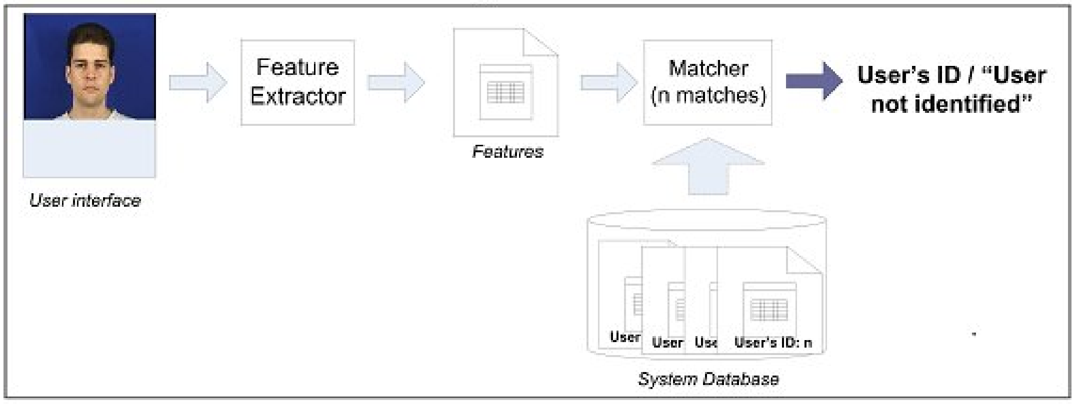
1.1 Basic System

A face recognition system can be either a verification system or an identification system depending on the context of an application. The verification system authenticate a user's identity by comparing the captured image with his/her own templates stored in the system. The system perform a one to one comparison to determine whether the person presenting herself/himself to the system is the person she/he claims to be. An identification system recognises a person by checking the entire template database for a match. It involves a one to many search technique. The system will either make a match and subsequently identify the person or it will fail to make a match.

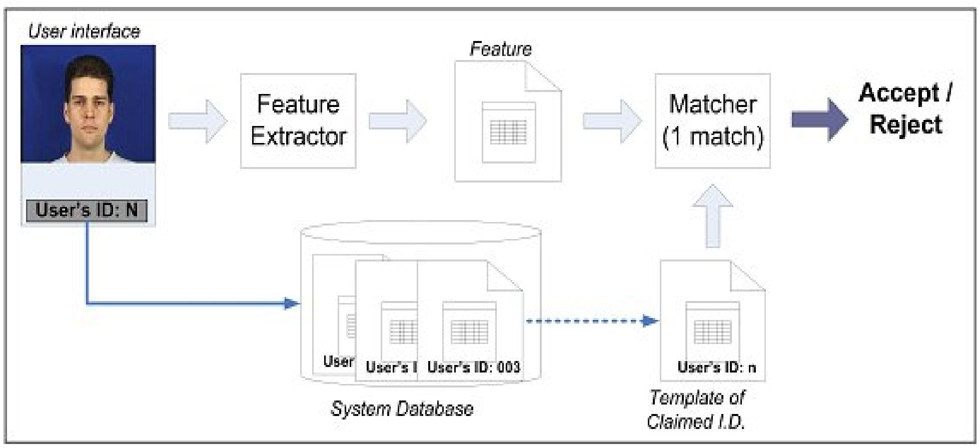
Block diagrams of the verification and identification systems respectively are presented in Figure 1.1. These systems consist of enrolment and matching. The Enrolment is the first stage of face recognition. The main objective of the enrolment is to register the person into the system database. The image of a person is captured by a sensor to produce a raw digital representation in the enrolment phase. The raw digital representation is then further processed by a feature extractor to generate a set of distinguishable features, called a template. The template can be stored in the central database of the system or be recorded on a magnetic card or smartcard depending on the application. In the task of verification, the user's name or PIN (Personal Identification Number) is read from the keyboard or the card. Then the image sensor captures the image of the person and the system converts it into a raw digital format. The feature extractor extracts then extracts the features from the raw format. The resultant features are fed into a one to one matcher;in order to determine whether the person should be accepted or rejected by comparing the extracted features against the template stored in the system database. In the identification task, PIN is not necessary and the matcher is a one to many, comparing the captured image with the templates of all users in the system database. The result is either an enrolled user's identity or a warning message such as''person not identified"



(a)Enrolment



(b) identification



(c) verification

Figure 1.1: Block diagrams of basic face recognition system

1.2 Applications of face recognition

There are innumerable applications of face recognition of which some have been mentioned below. There is potential for many more as soon as the problems due to pose variations, occlusion and lighting variations etc. are overcome.

1.2. Access Control

These days password protection is not enough and has increasingly become unreliable and unpredictable. The access control system based on **face recognition**

technology eliminates the need of the tradional and outdated security methods of implying security guards or maintaining a heap of papers containing the access details. The electronic security system captures the facial details of the individual and saves the relevant data automatically into the database of the computer. When the individual revisits the same for the next time, the system repeatedly captures the facial details and performs a match with the already stored patterns. Whenever it finds a match, it gives the permission for access right. This is an easy method of checking the access of the individuals to certain restricted areas or resources. This access system has many unique features like it ensures high security and is much more effective than the primitive methods. Also, there is no need of human touch.

1.3.2 Identification Systems

There may be many identification tasks, like sometimes a new applicant being enrolled must be compared against the entire database of previously enrolled claimants, to ensure that the new person is not claiming under more than one identity. In this case face recognition is most optimum than any other method because it is more acceptable and less intrusive. For example, face recognition can be employed to ensure that that people do not obtain multiple driving licenses.

1.2.3Surveillance

Surveillance is the application domain where most of the interest is being shown in face recognition. It can be applied without the person’s active participation, and indeed without his knowledge. Automated face recognition system can be applied ‘live’ to search for persons in the watch list or after using surveillance footage of a crime to search through the database of suspects.

1.2.4 Pervasive Computing

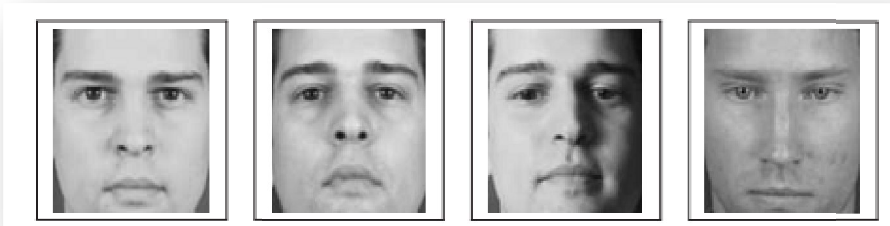
Pervasive or ubiquitous computing is another domain where face recognition is expected to become very useful, although it is still not commercially feasible. Computing devices, these days are already equipped with sensors of various kinds. These are found in our cars, phones and in many appliances in our homes. Networking of these devices is also becoming possible these days. We can imagine a future where many everyday gadgets will have some computational power, allowing them to change their behavior with time, user, user control and many other factors. Standard protocols are being developed which permit such devices to communicate with one another *e.g.* Bluetooth, ZigBee etc.

Most devices today can be controlled only by commands by the user. Some devices can also sense the environment, but it will be increasingly interesting and useful for such a pervasive, networked computing device to know about the physical world and the people within its region of interest. One of the most important parts of human awareness is knowing the identity of the user close to the device, and for this purpose, face recognition is most appropriate because of its passive nature.

Many examples can be found of pervasive face recognition tasks. Some devices such as Personal Digital Assistants (PDAs) already contain cameras for other such purposes and in good illumination conditions can be used to identify their users. Similarly, a domestic message centre may have user personalization that depends on identification by a built in camera.These kinds of devices when developed fully will not only make our lives easier but also prevent theft and intrusion of any kind by unwanted persons.

1.3Challenges of Face Recognition

Human visual system finds it easy to identify known human faces even under severely degraded viewing conditions, such as viewpoint, illumination, occlusion, expression degradation due to varied conditions and so on. Morever, automated face recognition is not yet able to achieve comparable results because measuring the similarity between two faces is based on the tradional measures of image similarity, such as, Normalised correlation or Eucledian metric. As Euclidean metric calcultes the distance between the images, the smaller the value of eucledian distance the greater the similarity. On the other side, Normalised correlation directly measures similarity between two images. It follows that these two measures are inverse to each other. Figure 1.2 illustrates the inadequacy of these measures for assessing similarity in face recognition. Image1 and Image 2 show the same person under even and uneven illumination, while Image 3 shows a different person. Template is a reference image belonging to the person in Image 1. Table1.1 clearly shows that similarity and distance measures would rate Image 3 to facial pose, illumination, and facial expression. The such variations are further enhanced by changes in the camera parameters, such as exposure time, aperture, sensor spectral response and lens aberrations.. As mentioned in [2, 3], the intrapersonal variations are usually larger than the image variation due to change in the face identity, called inter personal. This variability makes it difficult to build a simple model to describe.



(a) Template (b)image1 (c) image2 (d) image3

Figure 1.2: face images

	Image1	Image2	Image3
Normalised correlation	0.4334	-0.866	-0.2187
Eucledian distance	4,069	10,033	5424

Table 1.1: The similarity measure of the face images

This simple test demonstrates that the similarity measurements fail to generalise in the presence of image degradation. Zhao et al. [4] and others[31] have discussed extensively the challenges of face recognition which raise issues in computing,mathematics, engineering, neuroscience and psychophysics. The challenges can be summarised basically in two points: (1)A large variability in facial appearance of the same person and (2) High dimensionality of data and small sample size. A large variability in facial appearance of the same person is caused by variations of an individual from a small number of sample images or perform linear discriminant analysis(LDA) to separate different persons. Mathematically describing, the face manifold is highly complicated and non linear.

High dimensionality and small sample size: In general, the number of samples per person (typically less than 5) available is much smaller than the dimensionality of the image space. Therefore, the system cannot build reliable models of each individual to recognise the face identity from a probe image. This is called the generalisation problem. In addition, a small sample size may lead to numerical problems in matrix operations because of the singularity of within class covariance matrices [5]. In general, two directions, face image representation and pattern classification based on the extracted features, must be taken into account to deal with these challenges.

1.4 Contributions

The contributions of this dissertation to the methodological analysis of face recognition are summarised as follows:

The thesis presents a local binary pattern histogram (LBP) for face recognition under variant illuminated conditions.. The system offers considerable improvement in the face recognition performance in the presence of localisation errors because it takes advantages from the multiresolution information captured by the histogram. Previously the problem associated with a multiresolution analysis was the high dimensionality of the redundant representation combined with the small training sample size. This limits the total number of Local binary pattern (LBP) operators to at most of 3.

Illumination is cosidered to be the one of most significant problems in face recognition. The strategies for handling this problem can be summarised in two directions. The first is to convert the face image to a more canonical form in which illumination variations are suppressed. Other face modalities (such as 3D face shape or near-infrared face images), photometric normalisation and robust texture descriptor (such as Gabor filters and LBP) can help for this regard. The second way is to establish the robust classifier under illumination variations. In the thesis, the merit of different photometric normalisation techniques is investigated in the context of LBP face recognition.

1.5 Overview of Thesis

The overview of the dissetation is described below.

Chapter 2 the literature survey

In this section, the critical points of current knowledge on face recognition have been reviewed. Substantive findings as well as theoretical and methodological contributions in the field of face recognition have been included. This section summarizes the work done in the field of face recognition using various techniques in recent times.

Chapter3 Overview of Face Recognition: The structure of a generic face recognition described in this chapter. Some of the basic processing stages, including the face representation, the geometric and photometric normalisation, the feature selection and extraction, and the classifier are introduced.

Chapter 4. related work : the face recognition methods and correspondingly different approaches dealing with the constraints of illumination, pose and occlusion conditions has been explained.

Chapter 5. pre-processing has been defined in this chapter first. The operation of the LBP operator described as a face descriptor. The various classifiers for the purpose of classification of data has been explained.

Chapter 6. In this section, initially the dataset used for the work has been mentioned. Thereafter, all the experiments conducted along with the analysis of the results have been explained in

detailed. This includes two parts, feature extraction and classification. Figures and graphs have been provided wherever necessary.

Chapter7. Conclusions and Future Work: The dissertation is drawn to conclusion chapter where the directions of future work are also suggested.

References: This part includes the details of references used in the work.

Appendix A: Abbreviations

Appendix B: Overview of the MATLAB environment

Appendix C: MATLAB functions

Appendix D: Equal Error Rate

CHAPTER 2

THE LITERATURE SURVEY

Referring to a survey [1], face recognition systems can be clubbed into two categories : structure-based and appearance-based. In structure-based methods [6], a set of geometric face features, such as eyes, nose, etc., is extracted. The position of the different facial features of a face form a feature vector to put as input to a structural classifier in order to identify the particular subject. Morever, relevant facial feature detection and localisation methods employed are essential for this approach to be successful. Recently, the most systems, use the appearance of face as the input to decision making process and they can be further categorised as holistic and component based methods. The holistic appearance methods operate on the global features of the face image. As compared to structural methods, the face representation generally does not highly rely on accurate detection and localisation of specific facial points, and therefore these methods are usually more practical and easier to implement. Nowadays, appearance methods not only operate on the raw image space, but also other spaces, such as wavelet, local binary pattern (LBP) and ordinal pattern spaces. One of the reasons for using alternative face representations is that they simplify the face manifolds. Nevertheless, these kinds of representations exhibit high information redundancy and noise content, and information compression is needed to reduce the dimensionality of the representation to provide a concise and manageable feature space for classification. Several dimensionality reduction schemes have been developed to discover lower dimensional representation of human face by relying on statistical regularities. By reduction the dimensionality, it makes the face recognition system also computationally more tractable.

In general, good performance of holistic approaches can be achieved with well illuminated frontal face images. This is the direct response of the majority of algorithms relying on fundamentally linear analysis techniques. This performance of holistic approaches often degrades rapidly with pose changes, and background clutter,uneven illumination. Thus, an alternative to the holistic approach is to base face authentication on local facial components. The basic idea of component-based approach is to enhance the robustness to variations in illumination,pose and to face misalignment by allowing a flexible geometrical relation between the components in the classification stage. Heisele and his colleagues [9] have evaluated and compared the performance of holistic and component appearance systems with respect to their sensitivity to pose changes. Their experiments showed that the component-based system outperforms holistic systems even though more powerful classifiers are applied in the latter case. On the other hand, our work[8] and Ahonen et.al[7] also found that component-based approaches are more robust in the presence of face localisation errors.

In general, there are three traditional schemes to extract facial components. The simplest and most practical schemes[7, 8] divide the whole face image into non-overlapping or overlapping windows and regard them as the components. Another scheme[11, 9] is to extract the components centered on the facial features. The last scheme[10] is to apply the feature selection methods to select the components from a pool of over-complete local regions obtained by shifting and scaling a window on the face image.

Chapter 3

Overview of Face Recognition

In general, two directions, feature representation and pattern classification based on the features extracted from the face, must be pursued to tackle with the challenges mentioned in Section 1.3.The first is concerned with the representation of a face image in a “good” feature space where the face manifolds become simpler. Both image normalisation and face representation can help in this respect. The second way relates it to the design of a classifier to solve the difficult non linear classification and regression problems in the new face space and obtain good generalisation. In other words, the face image is segmented and then normalised by geometric and photometric normalisations which eliminate the effect of face rotation in plane, and scaling, and improve the face image quality. Then, a face representation, such as Gabor wavelets which reduce the non-linear behaviour of face data due to person to person variation, is extracted from the normalised image. The good normalisation and face representation methods help in reducing the degree of non-linearity, commonly the dimensionality of the face representation is increased. Thus, an effective dimensionality reduction method and a classifier are needed to deal with the above problem. The development of a successful algorithm requires the exploration of both directions. Many methods of face recognition pursuing the above directions have recently been proposed.

A block diagram of a generic face recognition system is presented in Figure 3.1.

Image Sensor: Most recent face recognition systems are based on face images captured in the spectrum of visible light. The main problem of these images is that the

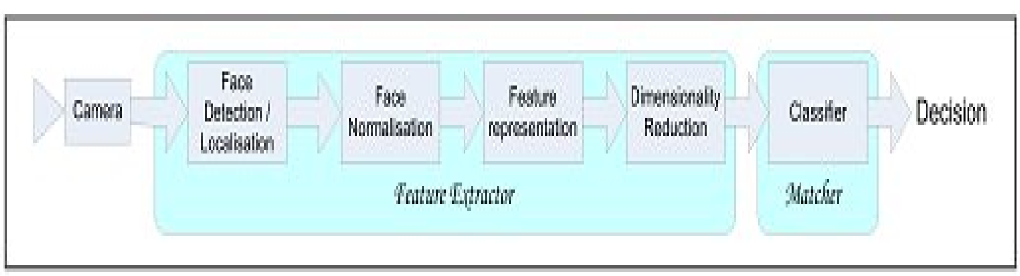


Figure 3.1: Configuration of a generic face recognition

variances in appearance of the particular person under different illumination conditions are larger than the changes due to different identities. This motivated the development of various special sensors to obtain different face modalities, such 3D face shape, near-infrared face images[12, 13], thermal face images, in order to eliminate dependence on illumination conditions. In this work, we focus on 2D image in the visible light spectrum only because this type of sensor is widely available.

3.1 Face Detection:The first step in the face recognition system is face detection. Its reliability has a major influence on the peformance and usability of a face recognition system. The purpose of this module is to provide the face location data for the face registration and normalisation module to segment the face region. Detecting a face in a complicated scene is very difficult because the system needs a set of reliable features which always appear when a face is present. Over the years, various methods have been reported. The reader can be referred to [14] for a comprehensive and critical survey of face detection methods. Up to now, perfect face localization is very difficult to achieve, and therefore a face recognition method capable of working well in the presence of localization errors is highly desired.

3.2 Face Normalisation: This module consists of geometric and photometric normalisation In general, the photometric normalisation is performed after the geometric normalisation. The objective of the geometric normalisation is to help the comparability of face images, while the aim of the photometric normalisation is to eliminate the illumination effects among different images.

3.2.1 Geometric Normalisation

Given an image, $\mathbf{I}$, and the eye coordination data, $\boldsymbol{i_{Leye}}$ and $\boldsymbol{i_{Reye}}$, with the predefined eye coordinates, $\boldsymbol{g_{Leye}}$ and $\boldsymbol{g_{Reye}}$ An affine warp can be applied for geometric normalisation. The affine warp equation relating the cropped face image to the image, called inverse mapping is presented below.

$$\boldsymbol{P = Aq + b} \text{......................................} \mathbf{3.1}$$

where p and q are locations of the input image and cropped face image respectively. Once the parameters of equations, A and b, are calculated first and then a cropped geometric face image, G can be obtained by the following equation.

$$G(q) = I((A^{-1}(p - b) \qquad 3.2$$

3.2.2 Photometric Normalisation

The objective of the photometric normalisation is to eliminate the illumination effect among different images. These techniques can be categorised into two groups. The first group employ training face samples to learn a global model of the possible illumination variations, for an instance, a linear subspace or an illumination cone[15], which eliminates the variations seen in the new images. The main drawback of this group is that it needs many training samples. The second group is to seek conventional image processing transformations which remove the influence of the face image illumination variations.The advantage of this group is that they do not require a training stage and training sample.

3.3 Face representation: As mentioned above, the main disadvantage of using the intensity image for face representation is its sensitivity to lighting variation, expression variation and a change of pose. Therefore many researchers have recently focused on developing more invariant face image representation. The features in these representations capture the local information that is difficult to learn using a small set of training data. In Section 3.1.2 and Chapter 3, Gabor features, ordinal features and Local Binary Patterns will be introduced which do not depend to the same extent on a large training set being available.

3.3.1 Gabor wavelets

Gabor wavelets were introduced to image analysis because of their familiarity to the receptive field profiles in cortical simple cells. The wavelets characterise the image as localised orientation selective and frequency selective features. Therefore, low level features, such as peaks, valleys and ridges are enhanced by 2-D Gabor filters. Thus, the eyes, nose and mouth, with other face details like wrinkles, dimples and scars are enhanced as key features to represent the face in higher dimensional space. Also, the Gabor wavelet representation of face image is robust to misalignment to some degree[80] because it captures the local texture characterised by spatial frequency, spatial position and orientation.

Recently, boosting algorithms have been widely accepted by the face research community. One of the reasons is that a boosting algorithm is a majority voting classifier.

Face recognition is a multi-class problem, but binary adaboost can solve only the two class problem. In order to avoid the need for a complex training process, the training samples can be remapped to intra-personal and inter-personal differential populations. An ideal intra-personal difference is an image with all pixel values set to zero, while an inter-personal difference should have much larger pixel values. Several ways of implementing this mapping have been suggested in the literature[16, 17]. In the Gabor feature space[18], the positive examples are derived from the pair of intrapersonal differences on the magnitude images and phase images in their corresponding scale and orientation space, whereas the negative samples are from the pair of interpersonal differences.

In the LBP histogram[19], an image pair is first split into sub-regions. The similarity score of each local LBP histogram pair is measured using the similarity function which will be discussed in below. Then similarity scores are concatenated to form an input feature vector for the process of feature selection . Over-complete features can be provided by shifting and scaling the local regions. In general, the total sample size of inter-person pairs is larger than that of intra-person pairs. This will give rise to a bias for feature selection process. The particular two approaches to solve this problem. One approach is to employ multiple feature selectors, each one of them using the whole set of intra-person samples with a portion[20] of the inter-person samples determined by randomly sampling. Another[21, 22] is to devise a cascaded Adaboost system with predefined false positive rate and detection rate (or called recognition rate) in each stage, and a predefined final false positive rate. This sample size ratio of intra-person pairs to inter-person pairs is fixed in each stage, and therefore inter-person samples are randomly sampled in the pool. During each stage after the training phase of test images, evaluation samples are involved to measure the false positive rate of the strong classifier so as to fulfil the predefined detection rate in decreasing the threshold in the last stage of the strong classifier. If the false positive rate does not meet the predefined rate, the Adaboost feature selection process will be iterated. Otherwise, the misclassified inter-person samples in the current stage, the full set of intra-person sample will be used to design the next stage of the cascaded classification process. In case the inter-person samples do not meet the sample size ratio, the remain will be added by randomly sampling in the pool. The process is iterated until the false positive rate meets the predefined final false positive rate. After that, the selected features will be stored.

3.4 Dimensionality Reduction: The main problem of face recognition methods is the high-dimensionality of feature space with commonly a small sample size dataset available for training. A simple straightforward implementation is computationally expensive. Therefore, techniques of feature selection or feature extraction are highly desired. One of the simple feature selection methods is based on the human perception.In the Elastic Bunch Graph Matching approach, the features are selected based on specific facial points chosen by a human expert. In another alternative, Adaboost will be described.it will introduce the techniques of feature extraction.

Feature Extraction

Classification methods operating in the image-space or feature-space representation suffer from a number of potential disadvantages, most of which root in the curse of dimensionality. Most of the face surface is smooth and has regular texture. A particular pixel value is typically highly correlated with the values of the surrounding pixels. Moreover, the face appearance is highly constrained; for example, the frontal view of a face is roughly symmetrical. Therefore the natural constraints dictate that the face images are confined to a subspace. In order to solve the curse of dimensionality problem, the feature selection and the feature extraction in the current section can assist to reduce the dimensionality. These feature extraction methods can be linear or nonlinear. They project the high-dimensional raw vector, $x \in R^n$ such as concatenated pixels in the feature space or selected feature space,image space into a low dimensional space in which a new feature vector, $\mathbf{y} \in \mathbf{R}^{\mathbf{v}}$ is given as.

$$\mathbf{y} = \boldsymbol{W}^{\boldsymbol{T}}\mathbf{X}$$

where $\mathbf{W} \in \mathbf{R}^{\mathbf{n}\times\mathbf{v}}$ is a transformation matrix. In this section, linear combination methods such as the well known Principle Component Analysis (PCA) and Linear Discriminant Analysis (LDA) will be described.

PCA

PCA is a standard decorrelation technique which projects the input signal into a space where features have no correlation with each other. It is a common technique for signal representation or signal compression because PCA can reduce the dimensionality by keeping the space which encapsulates the maximum amount of signal variation and throwing out dimensions with small variation which are regarded as noise. Pentland et al. [23] applied PCA to face recognition and called the face subspace as Eigenfaces. In

PCA-based training algorithm, the input is a training set, $X = [x_1, \ldots, x_m]$ of m facial images such that the mean of the training set is zero. The dimension of x is the total number of features used for describing the face.

$$\sum_x \Psi = \Psi \Lambda$$

where $\Psi = [\psi_1, \ldots, \psi_n]^T$ is the matrix of eigenvectors of the train set covariance matrix, Σ_x, and Λ is the diagonal matrix with eigenvalues $\lambda_1 \geq \ldots \lambda_n$ on its main diagonal, so ψ_j is the eigenvector corresponding to the jth largest eigenvalue. Thus it can be shown that the eigenvalue λ_i is the variance of the data projected on ψ_i variable. Thus, the lower order eigenvectors encode to larger variations of the described training set, although the higher order eigenvectors encode smaller variations of the training set. As reported by Zheng[24], ordering eigenvectors based on the descending order of eigenvalues is good to represent or compress the information, but it may not be good for signal classification. Therefore, the eigenvectors can be reordered based on the distance between image pairs of the same persons projected into Eigenspace, so-called Like-image different ordering[25].

In general, some portion of the higher-order eigenvectors is removed because it does not contribute to face recognition and the computation time can also be saved. There are six variants of eigenvector selection as shown below.

Standard eigenspace projection[25]: All eigenvectors corresponding to non-zero eigenvalues are kept to establish the subspace. Remove the last 40% of the eigenvectors[26]: The eigenvectors are sorted by the corresponding descending non-zero eigenvalues and this method only keeps 60% of the lower-order eigenvectors.

Energy dimension[25]: This method uses the minimum number of eigenvectors to guarantee that the retained energy is greater than a threshold. A typical threshold is 0.9. The PCA and LDA functions adopt this method to choose the eigenvectors.Therefore, The energy, e_i, of the ith eigenvector is the ratio of the sum of the first i eigenvalues over the sum of all the eigenvalues.

LDA

Although the eigenface method is useful to represent the face image, there is no reason to assume that this method enhances face recognition and the majority of face recognition papers have already argued this particular point. As Motivated by this observation Belhumeur et al. proposed the class specific linear method, called Fisher's Linear Discriminant analysis, FLD, to achieve better face recognition. The theoretical framework for the FLD is to maximise the ratio of between-class scatter to that of within-class scatter.

During face recognition, the number of images in the training set m is much smaller than the number of pixels in the image, n, making the within-class scatter matrix singular. It means that it is possible to choose matrix W such that the within-class scatter of the projected samples is zeros. In current years, many solutions have been proposed to deal this problem yet. Belhumeur et al. have proposed the method called Fisherface, that avoids the problem by projecting the image set to a lower dimensional space so that the resulting within-class scatter matrix is non-singular. This process is done by using PCA to reduce the dimension of the feature space so that the within-class scatter in the PCA space is non-zero, and then using the standard FLD to reduce the dimension to C-1.Therefore, In order to improve the generalisation capability of FLD, other researchers[28, 29, 30] have suggested the Enhanced Fisher Linear Discriminant Model (EFLDM). This method decomposes the FLD procedure into a simultaneous diagonalisation of the two within- and between-class scatter matrices. It first diagonalises the within-class scatter matrix and then the between-class scatter matrix. With this approach, more discriminating features in the PCA space can be kept for EFLDM.

Chen et al. [27] have developed a new LDA-based face recognition, called Null Space LDA (N-LDA), which can solve the small size problem. It chooses the projection vectors (transformation matrix) maximising between-class scatter with the constraint that the within-class scatter is zero; as null space of the within-class scatter matrix has been shown containing discriminative information. In a similar vein, Yu and Yang proposed the socalled Direct LDA (D-LDA). The key idea is to discard the null space of between-class scatter which contains no useful first order information. This process can be achieved by diagonalising the between-class scatter matrix and then diagonalising the within-class scatter matrix. Recently,Ye and Li[31] have suggested a two-stage FDA via the QR-decomposition. The present stage of the QR decomposition method, as a dimension reduction, maximises the separation between different classes. The second stage of QR is to perform FDA.

Other Subspace methods

Other subspace methods, such as Independent Component Analysis (ICA), Kernel PCA(KPCA), Kernel LDA (KDA), Discriminant Common Vector Approach (DCV) and Discrete Cosine transform (DCT) have been proposed. KPCA and KLDA are the kernel versions of PCA and LDA where a nonlinear mapping is applied to the original space before a PCA or LDA projection. In the Independent component analysis, a non-orthogonal transformation is selected such that the variables in the feature space are statistically independent.

3.5 Classifier: Once the images are projected to a subspace then the similarity of the image
and the template(s) will be measured to determine the person's identity.The goal of a classifier is to compare the features of a face probe image with those of the template and report the degree of match in terms of some match or similarity measure. Since face recognition is a multiclass problem often involving a small sample size, most systems apply a Nearest Neighbor(NN) classifier to make the decision. An important issue of the NN classifier design is how to measure similarity. In general, there are two methods to measure similarity. The first one is to measure the distance between these image features. The another possibility is to measure how similar they are. These two measures are the inverse of each other. There are many possible similarity and distance measures and some of them are presented below.Some researchers have applied classifiers, such as boosting classifier,SVM for recognition. These are naturally defined as two-class discriminant classifiers. There are two approaches to convert the multiclass problem into a binary problem. The first approach, called intra-interpersonal difference method, is to evaluate the difference between two images as a basis for determining whether the images are of the same person. The second approach, called client-specific method or one-vs-all method, is to establish classifiers each of which separates a single class from all remaining classes. There are two ways to perform the classification In component-based approach. The most practical and simplest one, called score-based classifier, is to build a classifier for each component and then combine the output scores by applying fusion techniques. The second method, called feature-based classifier, is to apply a single classifier on the component features. Researchers[32, 33, 34] applied a Hidden Markov Model(HMM) classifier or Gaussian mixture models(GMM) classifier in which the components features, such as features located on the eyes, chin and mouth regions, are represented by a multivariate probability distributions. Tan and his colleagues[35] concatenate the features of individual components into a single feature vector, and then apply the technique of dimensionality reduction to determine the final discriminative feature vector.

3.6 Summary

Generic face recognition systems can be classified as structure-based or appearance based. Recently, the appearance-based approach has been used in most of the face recognition systems. The chapter describes the appearance-base methods have further been categorised into holistic-and component-based methods. Face recognition is a multiclass problem and potentially requires a vast quantity of training data to design .This problem is mitigated by the intra-interperson difference approach and the client-specific approach, described in this chapter, which transforms the multiclass problem to a more manageable binary problem. In order to improve the performance of automatic face recognition, the techniques of face representation and pattern classification have been introduced to simplify the human face a number of times. The researchers have focused on developing a face representation capturing the local information which achieves invariance to illumination and facial expression. Motivated by a simple but powerful texture descriptor, called Local Binary Pattern, our proposed system extends this descriptor to multiresolution and multispectral analysis for face recognition.

CHAPTER 4

RELATED WORK

This work has presented the face recognition area, explaining different approaches, methods, tools and algorithms used since the 60's. Some algorithms are better, some are less accurate, some of these are more versatile and others are too computationally costly. Despite this variety, face recognition faces some issues inherent to the problem definition, environmental conditions and hardware constraints. Some specific face detection problems are explained in previous chapter. In fact, some of these issues are common to other face recognition related subjects. Nevertheless, those and some more will be detailed in this section.

4.1 Illumination

Many algorithms rely on color information to recognize faces. Features are extracted from color images, although some of them may be gray-scale. The color that we perceive from a given surface depends not only on the surface's nature, but also on the light upon it. In fact, color derives from the perception of our light receptors of the spectrum of light -distribution of light energy versus wavelength. There can be relevant illumination variations on images taken under uncontrolled environment. That said, the chromacity is an essential factor in face recognition. The intensity of the color in a pixel can vary greatly depending on the lighting conditions.

Is not only the sole value of the pixels what varies with light changes. The relation or variations between pixels may also vary. As many feature extraction methods relay on color/intensity variability measures between pixels to obtain relevant data, they show an important dependency on lighting changes. Always take into consideartion that, not only light sources can vary, but also light intensities may increase or decrease, new light sources added. Entire face regions be obscured or in shadow, and also feature extraction can become impossible because of solarization. The big problem is that two faces of the same subject but with illumination variations may show more differences between them than compared to another subject. Summing up, illumination is one of the big challenges of automated face recognition systems. Thus, there is much literature on the subject. However, it has been demonstrated that humans can generalize representations of a face under radically different illumination conditions, although human recognition of faces is sensitive to illumination direction.

Zhao et al. [36] illustrated this problem plotting the variation of eigenspace projection coeffcient vectors due to differences in class label along with variations due to illumination changes of the same class .

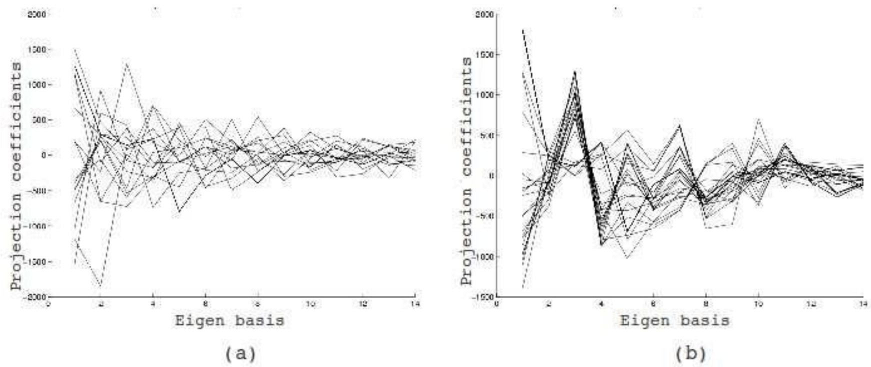


Figure 4.1: Variability due to class and illumination difference.

It has been demonstrated that the most popular image representations such as edge maps or Gabor-filtered images are not capable of overcoming illumination changes [37]. This illumination problem can be faced employing different approaches:

4.1.1 Heuristic approach

The observation of face recognition algorithms behavior can provide relevant clues in order to prevent illumination problems.

It has been suggested that, within eigenspace domain, three most significant principal component can be discarded [36]. However, we must maintain the system's performance with normally illuminated face images. Thus, we must assume that the first three principal components capture the variations only due to illumination.

An approach based on symmetry is proposed by Sirovich et. al in 2009 [38]. Their method is based on the natural symmetry of human faces. They conclude that odd eigenfaces are caused by illumination artifacts. Therefore, they discard them from their syntactic face construction procedure. This algorithm shows a nearly perfect accuracy recognizing frontal face images under different lighting conditions.

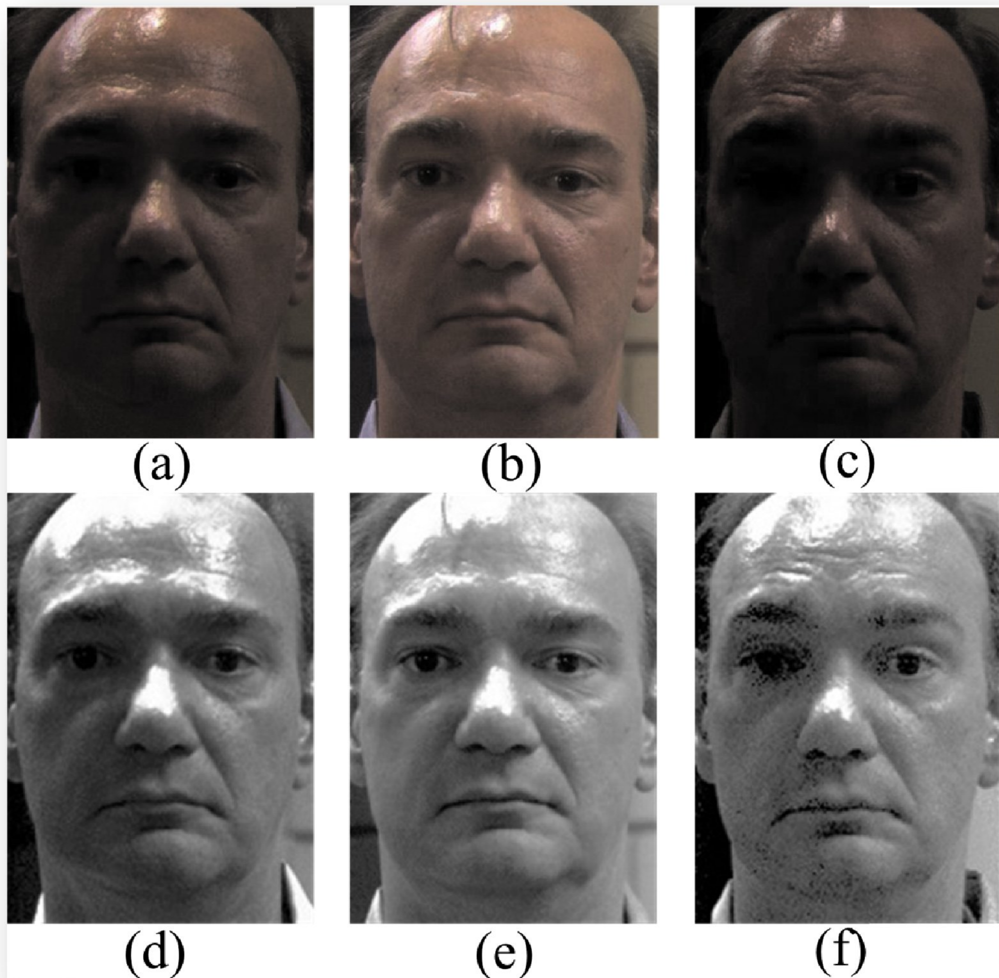


Figure 4.2 Faces with variant illumination

4.1.2 Statistical approach

Statistical methods for feature extraction can offer better or worse recognition rates. Moreover, there is extensive literature about those methods. Some papers test these algorithms in terms of illumination invariance, showing interesting results [39].

The class separability that provides LDA shows better performance that PCA, a method very sensitive to lighting changes. Bayesian methods allow to define intra-class variations such as illumination variations, so these methods show a better performance than LDA. However, all this linear analysis algorithms do not capture satisfactorily illumination variations.

Non-linear methods can deal with illumination changes better than linear methods. Kernel PCA and non-linear SVD show better performance than previous linear methods. Moreover, Kernel LDA deals better with lighting variations that KPCA and non-linear SVD. However, it seems that although choosing the right statistical model can help dealing with illumination, further techniques may be required.

4.1.3 Light-modeling approach

Some recognition methods try to model a lighting template in order to build illumination invariant algorithms. There are several methods, like building a 3D illumination subspace from 3 images taken under different lighting conditions, developing an illumination cone or using quotient shape-invariant images [36].

Some other methods try to model the illumination in order to detect and extract the lighting variations from the picture. Gross et. al develop a Bayesian sub-region method that regards images as a product of the reflectance and the luminance of each point. This characterization of the illumination allows the extraction of lighting variations, i.e. the enhancement of local contrast. This illumination variation removing process enhances the recognition accuracy of their algorithms in a 6.7%.

4.1.4 Model-based approach

The most recent model-based approaches try to build 3D models. The idea is to make intrinsic shape and texture fully independent from extrinsic parameters like light variations. Usually, a 3D head set is needed, which can be captured from a multiple-camera rig or a laser-scanned face [41]. The 3D heads can be used to build a morphable model, which is used to fit the input images [40]. Light directions and cast shadows can be estimated automatically. This approaches can show good performance with data sets that have many illumination variations like CMU-PIE and FERET [40].

4.2 Pose

Pose variation and illumination are the two main problems face by face recognition researchers. The vast majority of face recognition methods are based on frontal face images. These set of images can provide a solid research base. It can be mentioned that maybe the 90% of papers referenced on this work use these kind of databases. Image representation methods, dimension reduction algorithms, basic recognition techniques, illumination invariant methods and many other subjects are well tested using frontal faces.

On the other hand, recognition algorithms must implement the constraints of the recognition applications. Many of them, like video surveillance, video security systems or augmented reality home entertainment systems take input data from uncontrolled environment. The uncontrolled environment constraint involves several obstacles for face recognition: The aforementioned problem of lighting is one of them. Pose variation is another one.

Figure 4.3 ORL faces with pose variations.

There are several approaches used to face pose variations. Most of them have already been detailed, so this section won't go over them again. How-ever, it's worth mentioning the most relevant approaches to pose problem solutions:

4.2.1 Multi-image based approaches

These methods require multiple images for training. The idea behind this approach is to make templates of all the possible pose variations [36]. So, when an input image must be classified, it is aligned to the images corresponding to a single pose. This process is repeated for each stored pose until the correct one is detected. The restrictions of such methods are firstly, that many images are needed for each person. Secondly, this systems per-form a pure texture mapping, so expression and illumination variations are an added issue. Finally, the iterative nature of the algorithm increases the computational cost.

Multi-linear SVDs are another example of this approach . Other approaches include view-based eigenface methods, which construct an individual eigenface for each pose [36].

4.2.2 Single-model based approaches

This approach uses several data of a subject on training, but only one image at recognition. There are several examples of this approach, from which 3D morphable model methods are a self-explanatory example. Data may be collected from many different sources, like multiple-camera rigs or laser-scanners [42]. The goal is to include every pose variation in a single image. The image, on this case, would be a three dimensional face model. Theoretically, if a perfect 3D model could be built, pose variations would become a trivial issue. However, the recording of sufficient data for model building is one problem. Many face recognition applications don't provide the commodities needed to build such models from people. Other methods mentioned in [36] are low-level feature based methods and invariant feature based methods.

Some researchers have tried an hybrid approach, trying to use a few images and model the rotation operations, so that every single pose can be deducted from just one frontal photo and another profile image. This method is called Active Appearance Model (AAM) in [43].

4.2.3 Geometric approaches

There are approaches that try to build a sub-layer of pose-invariant information of faces. The input images are transformed depending on geometric measures on those models. The most common method is to build a graph which can link features to nodes and define the transformation needed to mimic face rotations. Elastic Bunch Graphic Matching (EBGM) algorithms have been used for this purpose [44, 45].

4.3 Occlusion

We understand that occlusion as the state of being obstructed. In the face recognition context, it involves that some parts of the face that can't be obtained. Considering an example, a face photograph taken from a surveillance camera could be partially hidden behind the column. The recognition process can rely heavily on the availability of a full input face. Therefore, the absence of some parts of the face may lead to a bad classification. This problem speaks in favor of a piecemeal approach to feature extraction, which do not depend on the whole face. There are also certain objects that can occlude facial features -, certain hair cuts, beard,hats and glasses etc.

Figure4.4 faces with variant occlusion faces

Optical technology: A face recognition system should be aware of the format in which the input images are provided. There are different cameras, with different features, different weaknesses and problems. Usually, most of the recognition processes involve a preprocessing step that deals with this problem.

4.4 Expression

Facial expression is another variability provider. However, it isn't as strong as illumination or pose. Several algorithms don't deal with this problem in a explicit way, but they show a good performance when different facial expressions are present

On the other hand, the addition of expression variability to pose and illumination problems can become a real impediment for accurate face recognition.

4.5 Algorithm evaluation

It's not easy to evaluate the effectiveness of a recognition algorithm. Several core factors are unavoidable:

- Hit ratio.
- Error rate.
- Computational speed.
- Memory usage.

Then, depending on the system requirements, there could be other important factors.

- Pose, Illumination,occlusion,expression, invariability.
- Scalability.
- Adaptability (to variable sizes, input image formats, etc.).
- Automatism (unsupervised processes).

CHAPTER 5

PROPOSED FACE RECOGNITION SYSTEM

5.1 Pre-processing

All of input images are the cropped faces which were automatically detected in the images. Pre-processing of the facial images is composed of four steps:

1) RGB to Gray conversion;

2) resizing the face;

3) illumination normalization;

4) sub-block division.

5.1.1 RGB to Gray Conversion

Color information may be useful for ethnicity classification. However, for the case that the person is wearing make-up, the performance of classification may be degraded. Hence, color information is discarded in our system. This is done by converting color image to YCbCr image and discarding Cb and Cr channels.

5.1.2 Resizing faces

Query faces are resized into the resolution among 128*150, 256*300, and 512*600 which is closest one to their original resolution. In order to cope with scale change of query images, key images are resized into the resolutions of 128*150, 256*300, and 512*600.

5.1.3 Illumination Normalization

Various lighting conditions make face recognition difficult. Even though they can't be perfectly controlled, the effect of illumination can be reduced by applying illumination normalization. After RGB to Gray conversion, illumination normalization is applied to the facial image. Several illumination normalizing methods were tested in our face recognition framework for comparison.

• **Histogram Equalization:** The goal of histogram equalization is to obtain uniform histogram. By spreading out the range of intensities with higher probability in histogram, the images with poor contrast can be visually enhanced.

• **Contrast stretching ends-in search:** Histogram of image is stretched according to (1). I_{new} and I_{old}denote the intensity of the processed and the original , respectively. H and L denote the higher and lower threshold of intensity respectively.

$$I_{new} = \begin{cases} 0 & if \;\; I_{old} \leq L \\ \dfrac{I_{old} - L}{H - L} & if \; L \leq I_{old} \leq H \\ 255 & if \;\; I_{old} \geq H \end{cases}$$

• **Retinex [46]:** It decomposes intensity of the pixel into illumination and reflection component and then reduces the effect of illumination.

5.1.4 Sub-block Division

Sub-block division is applied to the facial image after illumination is normalized. Facial images are divided into 1 (whole image), 4 (eyes, nose and lip cropping according to fixed proportion of face obtained by heuristic), 10 (2x5), 16 (4x4), 40(4x10), 49 (7x7), 80 (4x20) or 100 (10x10) blocks. When the facial images are divided into N blocks, the notation of i-th block is B_i

5.2 LOCAL BINARY PATTERN

The local binary pattern operator is an image operator which transforms an image into an array or image of integer labels describing small-scale appearance of the image. These particular labels or their statistics of their labels , the most commonly the histogram, are then used for further image analysis. Most widely used versions of the operator are designed for monochrome still images but it has been extended also for color (multi channel) images as well as videos and volumetric data. The present chapter covers the different versions of the actual LBP operator in spatial domain [47,48,49]].

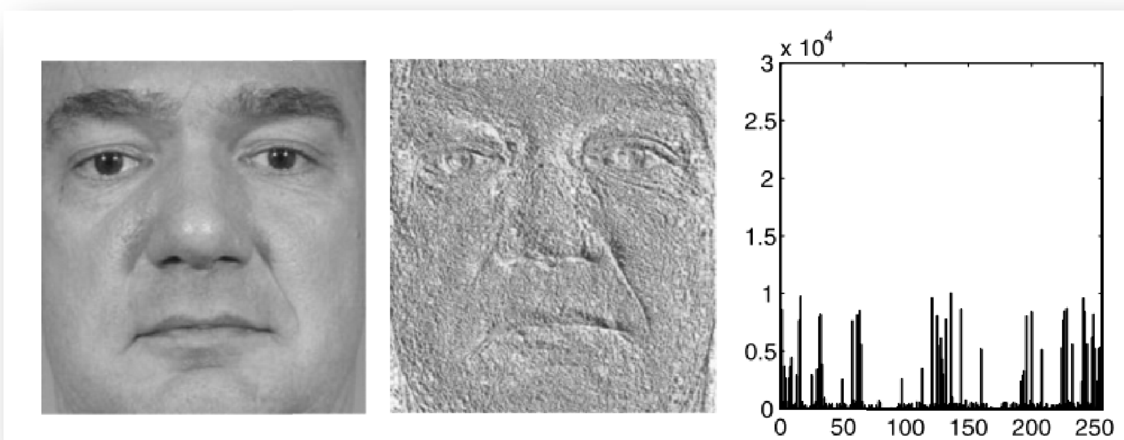
5.2.1 Basic LBP

The basic local binary pattern operator, introduced by Ojala et al.[50], was based on the assumption that texture has locally two complementary aspects, a pattern and its strength. In that work, the LBP was proposed as a two-level version of the texture unit [51] to describe the local textural patterns.

The original version of the local binary pattern operator works in a 3×3 pixel block of an image. The pixels in this block are thresholded by its center pixel value, multiplied by powers of two and then summed to obtain a label for the center pixel. As the neighborhood consists of 8 pixels, a total of = 256 different labels can be obtained depending on the relative gray values of the center and the pixels in the neighborhood. See Fig. 1.1 for an illustration of the basic LBP operator. An example of an LBP image and histogram are shown in Fig. 5.1.

5.2.2 Derivation of the Generic LBP Operator

Several years after its original publication, the local binary pattern operator was presented in a more generic revised form by Ojala et al. [50]. In contrast to the basic



Input image LBP image LBP histogram

Fig. 5.1 Example of an input image, the corresponding LBP image and its histogram

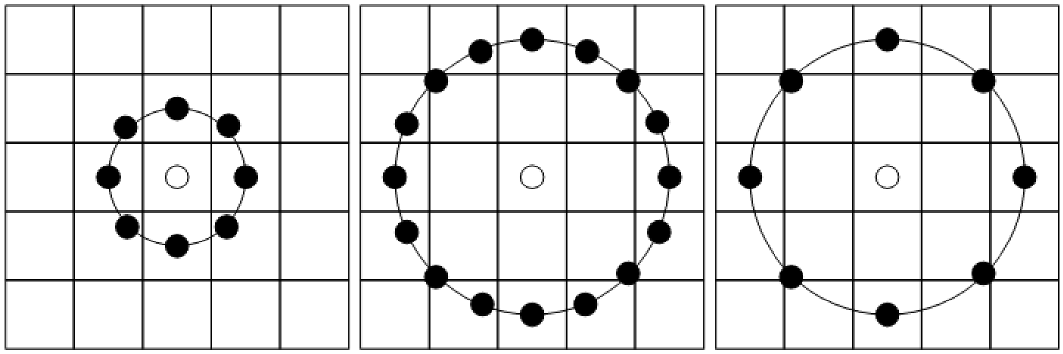


Fig. 5.2 The circular (8, 1), (16, 2) and (8, 2) neighborhoods. The pixel values are bilinearly interpolated whenever the sampling point is not in the center of a pixel.

LBP using 8 pixels in a 3 × 3 pixel block, this generic formulation of the operator puts no limitations to the size of the neighborhood or to the number of sampling points. The derivation of the generic LBP presented below follows that of [47, 48,49].

Consider a monochrome image I (x ,y) and let g_c denote the gray level of an arbitrary pixel (x, y),i.e. $g_{c=}$ I(x ,y).

Moreover, let $g_p$denote the gray value of a sampling point in an evenly spaced

circular neighborhood of P sampling points and radius R around point (x, y):

$$g_p = I(x, y), \quad p=0,\ldots\ldots\ldots,p-1 \text{ and}$$

$$x_p = x + R\cos(\tfrac{2\pi p}{P}),$$

$$y_p = y - R\sin(\tfrac{2\pi p}{P}).$$

See Fig. 5.2 for examples of local circular neighborhoods.

Assuming that the local texture of the image I(x,y) is characterized by the joint

distribution of gray values of P + 1 (P > 0) pixels:

$$T = t\,(g_{c,}g_0,g_1,\ldots\ldots\ldots\ldots\ldots\ldots g_{p-1}) \qquad 5.1$$

Without loss of information, the center pixel value can be subtracted from the neighborhood:

$$T = t(g_c, g_0 - g_{c,}, g_1 - g_{c,}, \ldots\ldots\ldots g_{p-1} - g_{c,}) \quad 5.2$$

In the next step the joint distribution is approximated by assuming the center pixel to be statistically independent of the differences, which allows for factorization of the distribution:

$$T \approx t(g_{c,})t(g_0 - g_{c,}, g_1 - g_{c,}, \ldots\ldots\ldots g_{p-1} - g_{c,}) \quad 5.3$$

Now the first factor $t(g_{c,})$ is the intensity distribution over I(x,y). From the point of view of analyzing local textural patterns, it contains no useful information. Instead the joint distribution of differences

$$t(g_0 - g_{c,}, g_1 - g_{c,}, \ldots\ldots\ldots g_{p-1} - g_{c,}) \quad 5.4$$

can be used to model the local texture. However, reliable estimation of this multi dimensional distribution from image data can be difficult. One solution to this problem, proposed by Ojala et al. in [52], is to apply vector quantization. They used learning vector quantization with a codebook of 384 code words to reduce the dimensionality of the high dimensional feature space. The indices of the 384 code words correspond to the 384 bins in the histogram. Thus, this powerful operator based on signed gray-level differences can be regarded as a text on operator, resembling some more recent methods based on image patch exemplars (e.g. [53]).

The learning vector quantization based approach still has certain unfortunate

properties that make its use difficult. First, the differences g_p-$g_c$are invariant to changes of the mean gray value of the image but not to other changes in gray levels. Second, in order to use it for texture classification the codebook must be trained similar to the other texton-based methods. In order to alleviate these challenges, only the signs of the differences are considered:

$$t(s(g_0 - g_{c,}, s(g_1 - g_{c,}), \ldots\ldots\ldots s(g_{p-1} - g_{c,})) \quad 5.5$$

where s(z) is the thresholding (step) function

$$s(z) = \begin{cases} 1, & z \geq 0 \\ 0, & z < 0 \end{cases}$$

The generic local binary pattern operator is derived from this joint distribution. As in the case of basic LBP, it is obtained by summing the thresholded differences

weighted by powers of two. The $LBP_{P,R}$ operator is defined as

$$LBP_{P,R}(x_c\ ,y_c) = \sum_{i=0}^{p-1} s(g_p\text{-}g_c)2^p \qquad 5.6$$

In practice, Eq. 2.10 means that the signs of the differences in a neighbourhood are interpreted as a P -bit binary number, resulting in 2P distinct values for the LBP code. The local gray-scale distribution, i.e. texture, can thus be approximately described with a 2P -bin discrete distribution of LBP codes:

$$T = t(LBP_{P,R}(x_c\ ,y_c)). \qquad 5.7$$

In calculating the $LBP_{P,R}$ distribution (feature vector) for a given N × M image

sample ($x_c \in \{0,...,N-1\}$, $y_c \in \{0,...,M-1\}$), the central part is only considered because a sufficiently large neighborhood cannot be used on the borders. The

LBP code is calculated for each pixel in the cropped portion of the image, and the distribution of the codes is used as a feature vector, denoted by S:

$$S= t(LBP_{P,R}(x_c\ ,y_c)). \qquad 5.8$$

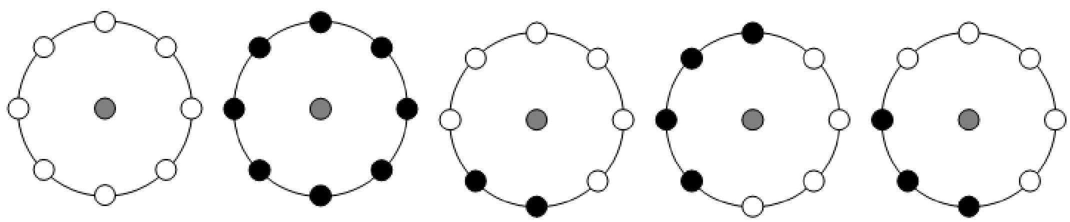
$$X \in \{ \lceil R \rceil,\ldots\ldots,N\text{-}1\text{-} \lceil R \rceil,\ \ Y \in \{\lceil R \rceil,\ldots\ldots\ldots\ldots\ldots M\text{-}\ 1\text{-}\lceil R \rceil\}.$$

The original LBP (Fig. 1.1) is very similar to $LBP_{8,1}$, with two differences. First, the neighborhood in the general definition is indexed circularly, making it easier to derive rotation invariant texture descriptors. Second, the diagonal pixels in the 3×3 neighborhood are interpolated in $LBP_{8,1}$.

5.2.3 Mappings of the LBP Labels: Uniform Patterns

In many texture analysis applications it is desirable to have features that are invariant or robust to rotations of the input image. As the $LBP_{P,R}$ patterns are obtained by circularly sampling around the center pixel, rotation of the input image has two effects: each local neighborhood is rotated into other pixel location, and within each neighborhood, the sampling points on the circle surrounding the center point are rotated into a different orientation.

Another extension to the original operator uses so called uniform patterns [53]. For this, a uniformity measure of a pattern is used: U ("pattern") is the number of bitwise transitions from 0 to 1 or vice versa when the bit pattern is considered circular. A local binary pattern is called uniform if its uniformity measure is at most 2.For example, the patterns 00000000 (0 transitions), 01110000 (2 transitions) and 11001111 (2 transitions) are uniform whereas the patterns 11001001 (4 transitions) and 01010011 (6 transitions) are not. In uniform LBP mapping there is a separate output label for each uniform pattern and all the non-uniform patterns are assigned to a single label. Thus, the number of different output labels for mapping for patterns



SPOT SPOT/FLAT LINE END EDGE CORNER

Fig. 5.3 Different texture primitives detected by the LBP

of P bits is $P(P - 1) + 3$. For instance, the uniform mapping produces 59 output labels for neighborhoods of 8 sampling points, and 243 labels for neighborhoods of 16 sampling points.

The reasons for omitting the non-uniform patterns are two fold. First, most of the local binary patterns in natural images are uniform. Ojala et al. noticed that in heir experiments with texture images, uniform patterns account for a bit less than 90% of all patterns when using the (8, 1) neighborhood and for around 70% in the (16, 2) neighborhood. In experiments with facial images [54] it was found that 90.6% of the patterns in the (8, 1) neighborhood and 85.2% of the patterns in the(8, 2) neighborhood are uniform.

The second reason for considering uniform patterns is the statistical robustness. Using uniform patterns instead of all the possible patterns has produced better recognition results in many applications. On one hand, there are indications that uniform patterns themselves are more stable, i.e. less prone to noise and on the other hand, considering only uniform patterns makes the number of possible LBP labels significantly lower and reliable estimation of their distribution requires fewer samples.

The uniform patterns allows to see the LBP method as a unifying approach to the traditionally divergent statistical and structural models of texture analysis [45]. Each pixel is labeled with the code of the texture primitive that best matches the local neighborhood. Thus each LBP code can be regarded as a micro-texton. Local primitives detected by the LBP include spots, flat areas, edges, edge ends, curves and so on. Some examples are shown in Fig. 5.3 with the $LBP_{8,R}$ operator. In the figure, ones are represented as black circles, and zeros are white.

The combination of the structural and statistical approaches stems from the fact that the distribution of micro-textons can be seen as statistical placement rules. The LBP distribution therefore has both of the properties of a structural analysis method: texture primitives and placement rules. On the other hand, the distribution is just a statistic of a non-linearly filtered image, clearly making the method a statistical one. For these reasons, the LBP distribution can be successfully used in recognizing a wide variety of different textures, to which statistical and structural methods have normally been applied separately.

5.3 Face descriptor extraction

In each block of facial images, histogram is independently obtained based on LBP operator and its variants. If the facial image is divided into N blocks, N histograms are calculated. Finally, N histograms are concatenated to yield the face descriptor. In the same way, 2D LBP histogram, which is composed of two of LBP and its 3 variants, 3D histogram (MGST, AP, Shape) proposed in the earlier work by the author [55], and 4D histogram, which combined proposed 3D histogram and LBP, are also extracted for comparison. Additionally, 3D histogram which combined three of MGST, AP [56], Shape descriptor based on FAST corner detector [55] and LBP is also extracted. MGST measures symmetrical similarity in local region and AP measures directionality of the pixel intensity change. Both of them are robust to rotation. The proposed shape descriptor in the earlier work by the author was refined to have 16 levels instead of 32 levels.

5.4 Matching

The similarity of face descriptor formed by combining N histograms was measured by several dissimilarity measures: Bhattacharyya distance, Histogram intersection, log-likelihood, Chi square, and correlation distance. F_j and $F_j(\mathrm{i})$ denotes the face descriptor of j-th facial image and its i-th value, respectively. E() and S() denotes averaging operator, andstandard deviation operator respectively.

Minimum distance classifiers

The minimum distance classifiers, as the name suggests, classify the data on the basis of the distance between the data points in the feature space. It classifies in such a manner, so as to minimize the distance between the data and the class in multi-feature space. The index of similarity here is the distance so that the minimum distance is identical to the maximum similarity. The most common distances often used in this procedure are:

- Bhattacharya distance

$$d(F_1,F_2) = \sqrt{1 - \sum_i \sqrt{\frac{F_1(\mathrm{i}) * F_2(\mathrm{i})}{E(F_1) * E(F_2)}}}$$

- Chi square statistics

$$d(F_1,F_2) = \sum_i \frac{((\mathrm{F}_1(\mathrm{i}) - F_2(\mathrm{i}))^2}{((\mathrm{F}_1(\mathrm{i}) + F_2(\mathrm{i}))^2}$$

- Histogram intersection

$$d(F_1,F_2) = \sum_i \min\left(F_1(\mathrm{i}), F_2(\mathrm{i}) \right)$$

- Correlation distance

$$d(F_1,F_2) = \frac{E((F_1 - E(F_1) * (F_2 - E(F_2)))}{S(F_1) * S(F_2)}$$

- Log – likelihood statistics

$$d(F_1,F_2) = - \sum F_1(\mathrm{i}) * \log(F_2(\mathrm{i}))$$

The weight for B_i , $w_i$can be applied to the dissimilarity results of histogram in B_i to form the final dissimilarity. It is because that each block of the facial image has not valuable information equivalently.

Query images are classified into the person with the lowest dissimilarity according to matching results.

CHAPTER 6

EXPERIMENTS AND RESULTS

The LOCAL BINARY PATTERN (LBP) is used for the purpose of texture feature extraction. The standard ORL database is used for conducting all the experiments. It consists of images of 20 subjects of size 92 x 112 as shown in Figure 6.1

Figure 6.1: Subjects of the ORL database

Each of the 20 subjects has 10 orientations. Some orientations of sample image (subject 2) are shown in figure 6.2.

Figure 6.2 : Different orientations of subject 2 in ORL database

For all the methods implemented implemented, 8 orientations (of each subject) were used and two orientations for testing. The entire programming is done using the software 'MATLAB'. An introduction to the programming in matlab has been provided in appendix A.

6.1 IMAGE PROCESSING

Image normalization is done in order to reduce computational overhead and improve the performance of the method. The images in the ORL database are grayscale images which contain intensity values ranging from 0 to 255. The pixel values of all the images are divided by maximum pixel value of the image. After normalisation, the values of all the pixels lie between 0 and 1

6.1.1 Feature extraction The facial image is divided into local regions and texture descriptors are extracted from each region independently. The descriptors are then concatenated to form a global description of the face. The LBP labels for the histogram contain information about the patterns on a pixel-level. The labels are summed over a small region to produce information on a regional level The regional histograms are concatenated to build a global description of the face.

6.3 Result:

MATCHING AND CLASSIFICATION

Euclidian distance

The Euclidean distance is also known as the Euclidean metric that is distance between two points that is measured with a ruler ,and is given by the Pythagoras formula. By using this formula as distance, Euclidean space or any other inner product space becomes a metric space. The norm of the Euclidian distance is called the Euclidean norm.

It is used in those problems where the variances of the population classes are different from each other. Theoretically the Euclidian distance is identical to the similarity index. The Euclidean distance between two points ‘p’ and ‘q’ is the length of the line segment connecting them.

If p = (p_1, p_2,...,p_n) and q = (q_1, q_2,..., q_n) are two points in Euclidean n dimensional space, then the Euclidian distance from p to q is given by:

$$\text{d (p, q)} = \text{d (q, p)} = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 \ldots\ldots\ldots\ldots + (p_n - q_n)^2}$$

$$+\sqrt{\sum_{i=1}^{n}(p_i - q_i)^2}$$

The position of a point in a Euclidean *n*-space is shown by a Euclidean vector. So, Euclidean vectors p and q, starting from the origin of the space. The Euclidean length, or Euclidean norm, or magnitude of a vector measures the length of this vector and is given by:

$$\text{x(t)},\frac{1}{\sqrt{A}}y\left(\frac{t-\tau}{A}\right),z(t)$$

Here the last equation involves the dot product.

K-Means Classifier

The k-means algorithm is a fast iterative algorithm that has been used in many clustering applications. It finds locally optimal solutions with respect to the clustering error. Let us consider a data set X={ $x_1, x_2, x_3, \ldots\ldots\ldots\ldots x_n$} X $\in R^d$where d is the dimension. The K-clustering problem aims at partitioning this data set into K independent clusters $\{C_{1,}C_2, C_3, \ldots\ldots\ldots\ldots C_k\}$, such that squared Euclidean distances between each data point x_i and the centroid m_c (cluster center) of the subset C_K is minimized. This criterion is called clustering error and depends on the cluster centers { $m_1 m_2 m_3 \ldots\ldots\ldots.. m_K$}

$$\text{E } (m_1 m_2 m_3 \ldots\ldots\ldots.. m_K) = \sum_{i=1}^{M}\sum_{j=1}^{N} I\left(x_i \in C_j\right)||x_i - m_j||^2$$

Where I(X) =1 if X is true and 0 otherwise.

6.2.1 Matching with pose image

a) Subject 2

b) Subject 9

Figure 6.3.1 : matching with pose image with subject 2 and 9.

6.2.2 matching result with illumination or intensity variation

a) Subject 8

b) Subject 12

Figure 6.3.2 : matching with pose image with subject 8 and 12.

6.3.3 Matching result with Partial Occlusion

a) Subject 4

b) Subject 1

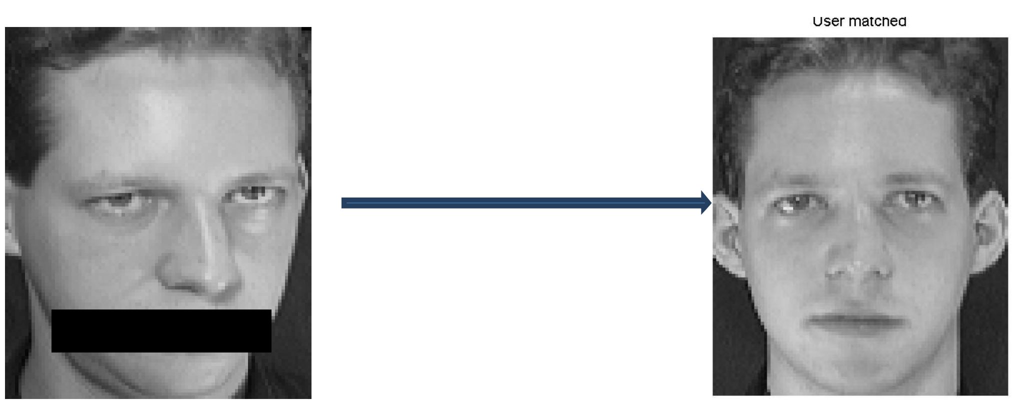


Figure 4.3.3 : matching with pose image with subject 4 and 1.

6.4 EQUAL ERROR RATE

Performance measures in Face Verification

The verification systems make two different types of error: 1)mistaking biometric measurements from two different persons to be from the same person, namely False Acceptance (FA). 2) mistaking two biometric measurements from the same person to be from two different persons , namely False Rejection (FR). The peformance is measured in terms of False Acceptance Rate (FAR) and False Rejection Rate (FRR), defined as:

$$FAR = \frac{number\ of\ FAs}{number\ of\ imposter\ accesses}$$

$$\mathrm{FRR} = \frac{Number\ of\ FRs}{Number\ of\ total\ true\ client\ accesses}$$

There is a tradeoff between FAR and FRR in every verification system, as both FAR and FRR are a function of the threshold (T). For a given value of the threshold (T), there is a pair of FAR(T) and FRR(T). They can be plotted against each other as a curve known as Receiver Operating Characteristic(ROC) to express the behavior of FAR and FRR. In XM2VTS experiments, the threshold is usually chosen on the evaluation set at FAR=FRR, called Equal Error Rate (EER). It is then applied to the test set to obtain FAR and FRR and consequently sum of both to get the Total Error Rate (TER). By comparing the TER with other systems, our systems can be benchmarked. On the other hand, the verification rate (i.e. 1-FRR) at 0.1% FAR is generally used to represent the system accuracy in ORL experiments.

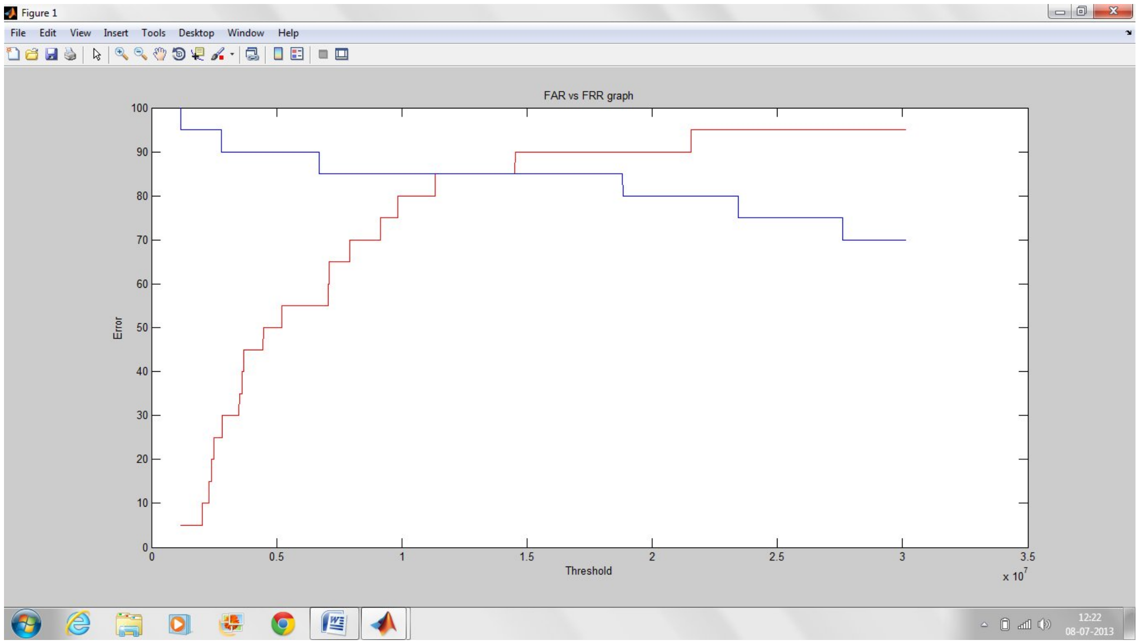


Fig 6.4 receiver operator characterstic (ROC) curve

- False Acceptance Ratio (FAR) is 0.75
- False Rejection Ratio is 0.95
- Equal error rate is 0.85.
- Total time in recognising is 0.926835 seconds.
- Reconition rate is 98.12 %.

CHAPTER 7

CONCLUSION AND FUTURE WORK

Conclusion

In the presented work, a robust face recognition method using Local Binary Pattern (LBP) has been proposed. It performed satisfactorily under illumination variation, pose variation and under occlusion conditions. Earlier, LBP is one of the most used texture descriptor in image analysis. one of the best performing texture descriptors and widely used in various applications. The LBP method has proven to be highly discriminative and because its invariance to monotonic gray level changes and better computational efficiency, make it suitable for high demanding image analysis tasks. Face can be seen as a composition of micro-patterns which can be well described by LBP operator.

The whole facial image is divided into local regions and then texture descriptors are extracted from each region independently. The facial descriptors are then concatenated to form a global description of the whole face. The LBP labels for the histogram contain information about the patterns on a pixel-level. The labels are concatenated over a small region to produce information on a regional level The regional histograms are concatenated to build a global description of the face.

Its tolerance to monotonic gray-scale changes. The computational efficiency of the LBP operator and that no gray-scale normalization is needed prior to applying the LBP operator to the face image.Real-world face recognition systems need to perform face detection prior to face recognition. The automatic face localization may not be completely accurate so it is desirable that face recognition works under small localization errors. It can be sobserved that when no error or only a small error is present, then LBP with small local regions works well but as the localization error enhances, using larger local regions produces better recognition rate.

Future work

Studying more advanced methods for dividing the facial image into local regions and finding the weights for them. Looking for image preprocessing methods and descriptors that are more robust against image transformations that change the appearance of the surface texture. LBP features for facial expression recognition has been studied. A novel and efficient facial representation is proposed precisely. LBP based face description has already attained an established position in face analysis research and many research group already study about it. The recognition rates of the LBP perform pretty good than other comparison algorithm presented in this paper.The recognition rates of the LBP maintain high level under the effect of localization errors.Further research is still needed to achieve even better performance in the recognition rates of LBP. Didnt give the method whether we can improve the recognition rates of the LBP under the effect of localization errors.

REFERNCES

[1] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. “Face recognition: A literature survey”. ACM Comput. Surv. , 35(4):399–458, 2003.

[2] Anil K. Jain and Stan Z. Li.” Handbook of Face Recognition”. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2004.

[3] Tae-Kyun Kim and Josef Kittler.” Locally linear discriminant analysis for multimodally distributed classes for face recognition with a single model image”. IEEE Transactions on Pattern Analysis and Machine Intelligence , 27(3):318–327, 2005.

[4] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. “Face recognition: A literature survey”. ACM Comput. Surv. , 35(4):399–458, 2003.

[5] P.N. Belhumeur, J.P. Hespanha, and D.J. Kriegman. “Eigenfaces vs. fisherfaces: recognition using class specific linear projection”. Transactions on Pattern Analysis and Machine Intelligence , 19(7):711–720, Jul 1997.

[6] Laurenz Wiskott, Jean-Marc Fellous, Norb ert Kruger, and Christoph von derMalsburg. “Face recognition by elastic bunch graph matching”. IEEE Trans. onPattern Analysis and Machine Intelligence , 19(7):775–779, 1997.

[7] Timo Ahonen, Ab denour Hadid, and Matti Pietik¨ainen.” Face description with local binary patterns: Application to face recognition”. IEEE Trans. Pattern Anal.Mach. Intell. , 28(12):2037–2041, 2006.

[8] Chi-Ho Chan, Josef Kittler, and Kieron Messer. “Multi-scale local binary pattern histograms for face recognition”. In Seong-Whan Lee and Stan Z. Li, editors, ICB ,volume 4642 of Lecture Notes in Computer Science , pages 809–818. Springer,2007.

[9] B. Heisele, P. Ho, and T. Poggio.” Face recognition with supp ort vector machines: global versus component-based approach”. In Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on , volume 2, pages 688–694 vol.2, 2001.

[10] Stan Z. Li, ChunShui Zhao, Meng Ao, , and Zhen Lei.” Learning to fuse 3d+2dbased face recognition at both feature and decision levels. In Analysis and Model ling of Faces and Gestures , pages 44–54”. Springer Berlin / Heidelb erg, 2005.

[11] A. Pentland, B. Moghaddam, and T. Starner. “View-based and modular eigenspaces for face recognition”. In Computer Vision and Pattern Recognition, 1994. Proceedings CVPR ’94., 1994 IEEE Computer Society Conference on , pages 84–91, 21-23 Jun 1994.

[12] Stan Z. Li, Rufeng Chu, ShengCai Liao, and Lun Zhang.” Illumination invariant face recognition using near-infrared images”. IEEE Trans. Pattern Anal. Mach.Intel l. , 29(4):627–639, 2007.

[13] Xuan Zou, Josef Kittler, and Kieron Messer.” Motion compensation for face recognition based on active differential imaging”. In ICB , pages 39–48, 2007.

[14] Ming-Hsuan Yang, David J. Kriegman, and Narendra Ahuja. “Detecting faces in images: A survey”. IEEE Trans. Pattern Anal. Mach. Intel l. , 24(1):34–58, 2002.

[15] P.N. Belhumeur and D.J. Kriegman.” What is the set of images of an object under all ossible lighting conditions?” In CVPR , page 270, Los Alamitos, CA, USA,1996. IEEE Computer Society.

[16] Xiaogang Wang and Xiaoou Tang.” Bayesian face recognition using gabor features.In WBMA ’03: Proceedings of the 2003 ACM SIGMM workshop on Biometrics methods and applications , pages 70–73”, New York, NY, USA, 2003. ACM.

[17] B. Moghaddam, C. Nastar, and A. Pentland. “ A bayesian similarity measure for direct image matching”. icpr , 02:350, 1996.

[18] Lei Zhang, S.Z. Li, Zhi Yi Qu, and Xiangsheng Huang. “Boosting local feature based classifiers for face recognition. In Computer Vision and Pattern Recognition Workshop, 2004 Conference on , pages 87–87”, 27-02 June 2004.

[19] Stan Z. Li, ChunShui Zhao, Meng Ao, and Zhen Lei. “ Learning to fuse 3d+2d based face recognition at both feature and decision levels”. In AMFG , pages 44–54,2005.

[20] X. Chen, S. Shan, B. Cao, X. Chen, and W. Gao. “Multiple fisher classifiers combination for face recognition based on grouping adaboosted gab or features.In BMVC05 , pages 569–578”, 2005.

[21] Shiguang Shan, Peng Yang, Xilin Chen, and Wen Gao. “Adaboost gab or fisher classifier for face recognition”. In AMFG , pages 279–292, 2005.

[22] Paul Viola and Michael J. Jones. “ Robust real-time face detection. Int. J. Comput. Vision , 57(2):137–154”, 2004.

[23] Tae-Kyun Kim, Hyunwo o Kim, Wonjun Hwang, and Josef Kittler. “Component-based lda face description for image retrieval and mp eg-7 standardisation. Image Vision Comput. , 23(7):631–642”, 2005.

[24] P.C. Wei-Shi Zheng; Jian-Huang Lai; Yuen. "Ga-fisher: a new lda-based face recognition algorithm with selection of principal components. Systems, Man, and Cybernetics, Part B, IEEE Transactions on , 35(5):1065–1078", Oct. 2005.

[25] Wendy S. Yambor. " Analysis of pca-based and fisher discriminant-based image recognition algorithms". Master's thesis, CSU Computer Science Department, July 2003.

[26] Jonathon Phillips Hyeonjoon Moon. "Analysis of pca-based face recognition algorithms. Empirical Evaluation Techniques in Computer Vision" , 1998.

[27] Li-Fen Chen, Hong-Yuan Mark Liao, Ming-Tat Ko, Ja-Chen Lin, and Gwo-Jong Yu. " A new lda-based face recognition system which can solve the small sample size problem. Pattern Recognition , 33(10):1713–1726", 2000.

[28] AnaghaV.Malkapurkar, Rupali Patil, Prof. Sachin Murarka. "A New Technique for LBP Method to Improve Face Recognition". International Journal of Emerging Technology and Advanced Engineering Website: www.ijetae.com (ISSN 2250-2459, Volume 1, Issue 1, November 2011) .

[29] H. Cheng jun Liu; Wechsler. " Gabor feature based classification using the enhancedfisher linear discriminant model for face recognition. Image Processing, IEEETransactions on , 11(4):467–476", Apr 2002.

[30] Bao chang Zhang, Peng Yang, Shiguang Shan, and Wen Gao. " Discriminant gab or- faces and support vector machines classifier for face recognition". In ACCV, pages 37–41, 2004.

[31] Jieping Ye and Qi Li . "A two-stage linear discriminant analysis via qr-decomposition." IEEE Trans. Pattern Anal. Mach. Intel l. , 27(6):929–941, 2005.

[32] Guillaume Heusch, Yann Ro driguez, and Sebastien Marcel. " Local binary patterns as an image preprocessing for face authentication. In FGR '06: Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition , pages 9–14," Washington, DC, USA, 2006. IEEE Computer Society.

[33] A.V. Nefian and I I I Hayes, M.H. "Hidden markov models for face recognition.Acoustics, Speech and Signal Processing", 1998. Proceedings of the 1998 IEEE International Conference on , 5:2721–2724 vol.5, 12-15 May 1998.

[34] Yann Ro driguez and S´ebastien Marcel. " Face authentication using adapted local binary pattern histograms". In ECCV (4) , pages 321–332, 2006.

[35] Xiaoyang Tan and Bill Triggs. "Fusing gab or and lbp feature sets for kernel-based face recognition. In AMFG , pages 235–249", 2007.

[36] Guoying Zhao and Matti Pietikainen. " Dynamic texture recognition using local binary patterns with an application to facial expressions". IEEE Transactions on Pattern Analysis and Machine Intel ligence , 29(6):915–928, 2007.

[37] Timo Ahonen, Ab denour Hadid, and Matti Pietik¨ainen. "Face recognition with local binary patterns". In ECCV (1) , pages 469–481, 2004.

[38] P.C. Wei-Shi Zheng; Jian-Huang Lai; Yuen. " Ga-fisher: a new lda-based face recognition algorithm with selection of principal comp onents. Systems, Man, and Cybernetics", Part B, IEEE Transactions on , 35(5):1065–1078, Oct. 2005.

[39] A.P. Turk, M.A.; Pentland. "Face recognition using eigenfaces. In Computer Vision and Pattern Recognition", 1991. Proceedings CVPR '91., IEEE Computer Society Conference on , pages 586–591, 3-6 Jun 1991.

[40] Li-Fen Chen, Hong-Yuan Mark Liao, Ming-Tat Ko, Ja-Chen Lin, and Gwo-Jong Yu. " A new lda-based face recognition system which can solve the small sample size problem. Pattern Recognition , 33(10):1713–1726", 2000.

[41] Anton Schwaighofer. Svm tool box for matlab, January 2002.http://ida.first.fraunhofer.de/ anton/software.html.

[42] Timo Ojala, Matti Pietikainen, and Topi Maenpaa. " Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence ", 24(7):971–987, 2002.

[43] Xiangsheng Huang, Stan Z. Li, and Yangsheng Wang. "Jensen-shannon boosting learning for object recognition". In CVPR (2) , pages 144–149, 2005.

[44] Hua Yu and Jie Yang. "A direct lda algorithm for high-dimensional data – with application to face recognition". Pattern Recognition , 34(10):2067–2070, 2001.

[45] Ramin Zabih and John Woodfill. " Non-parametric local transforms for computing visual correspondence". In ECCV (2) , pages 151–158, 1994.

[46] D.J. Jobson, Z. Rahman, and G.A. Woodell, "Properties and performance of a center/surround retinex", IEEE Trans. Image Processing, vo.6, no.3, pp.451-462, 1997

[47.] Mäenpää, T., Pietikäinen, M.: "Texture analysis with local binary patterns". In: Chen, C.H.,Wang, P.S.P. (eds.) Handbook of Pattern Recognition and Computer Vision, 3rd edn., pp.197–216. World Scientific, Singapore (2005)

[48]. Maria De Marsico, Member, IEEE , Michele Nappi, Daniel Riccio, and Harry Wechsler, Fellow, IEEE. " Robust Face Recognition f or Uncontrolled Pose and Illumination Changes"IEEE TRANSAC TIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS, VOL. 43, NO. 1, JANUARY 2013.

[49]. Ojala, T., Pietikäinen, M., Mäenpää, T.: "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns". IEEE Trans. Pattern Anal. Mach. Intell. 24(7),971–987 (2002)

[50]. Ojala, T., Pietikäinen, M., Harwood, D.: "A comparative study of texture measures with classification based on feature distributions". Pattern Recognit. 29(1), 51–59 (1996)

[51]. Wang, L., He, D.C.: "Texture classification using texture spectrum. Pattern Recognit". 23, 905–910 (1990)

[52] Ojala, T., Valkealahti, K., Oja, E., Pietikäinen, M.: "Texture discrimination with multidimensional distributions of signed gray-level differences". Pattern Recognit. 34(3), 727–739 (2001)

[53]. Varma, M., Zisserman, A.: " A statistical approach to materials classification using image patchexemplars." IEEE Trans. Pattern Anal. Mach. Intell.31, 2032–2047 (2009)

[54]. Ahonen, T., Hadid, A., Pietikäinen, M.: " Face description with local binary patterns: Application to face recognition". IEEE Trans. Pattern Anal. Mach. Intell. 28(12), 2037–2041 (2006)

[55] Keun-Dong Lee, Seon-Tae Kim, Sung-Kwang Je, and Weon-Geun Oh, "Face recognition for personal image management system," in Proceedings of the 16th Korea-Japan Joint Workshop on Frontiers of Computer Vision, 2009

[56] Won-Keun Yang, Ayoung Cho, Dong-Seok Jeong, and Weon-Geun Oh, "Image Description and Matching Scheme for Identical Image Searching", Conf. on Advanced Service Computing, IEEE, pp.669-674, 2009

APPENDIX A

ABBREVIATIONS

LBP	Local binary pattern
PCA	Principal component analysis
LDA	Linear discriminant analysis
EGBM	Elastic group bunching methods
AAM	Active appearance models
FAR	False acceptance rate
FRR	False rejection rate
ROC	Receiver operating characterstic

APPENDIX B

OVERVIEW OF MATLAB ENVIRONMENT

MATLAB is a high-level technical computing language and interactive environment for algorithm development, data visualization, data analysis, and numeric computation. Using the MATLAB product, you can solve technical computing problems faster than with traditional programming languages, such as C, C++, and Fortran.

You can use MATLAB in a wide range of applications, including signal and image processing, communications, control design, test and measurement, financial modeling and analysis, and computational biology. Add-on toolboxes (collections of special-purpose MATLAB functions, available separately) extend the MATLAB environment to solve particular classes of problems in these application areas.

MATLAB provides a number of features for documenting and sharing your work. You can integrate your MATLAB code with other languages and applications, and distribute your MATLAB algorithms and applications. Features include:

- High-level language for technical computing
- Development environment for managing code, files, and data
- Interactive tools for iterative exploration, design, and problem solving
- Mathematical functions for linear algebra, statistics, Fourier analysis, filtering, optimization, and numerical integration
- 2-D and 3-D graphics functions for visualizing data
- Tools for building custom graphical user interfaces
- Functions for integrating MATLAB based algorithms with external applications and languages, such as C, C++, Fortran, Java™, COM, and Microsoft® Excel®

The MATLAB System

The MATLAB system consists of these main parts:

Desktop Tools and Development Environment

This part of MATLAB is the set of tools and facilities that help you use and become more productive with MATLAB functions and files. Many of these tools are graphical user interfaces. It includes: the MATLAB desktop and Command Window, an editor and debugger, a code analyzer, and browsers for viewing help, the workspace, and folders.

Mathematical Function Library

This library is a vast collection of computational algorithms ranging from elementary functions, like sum, sine, cosine, and complex arithmetic, to more sophisticated functions like matrix inverse, matrix eigenvalues, Bessel functions, and fast Fourier transforms.

The Language

The MATLAB language is a high-level matrix/array language with control flow statements, functions, data structures, input/output, and object-oriented programming features. It allows both "programming in the small" to rapidly create quick programs you do not intend to reuse. You can also do "programming in the large" to create complex application programs intended for reuse.

Graphics

MATLAB has extensive facilities for displaying vectors and matrices as graphs, as well as annotating and printing these graphs. It includes high-level functions for two-dimensional and three-dimensional data visualization, image processing, animation, and presentation graphics. It also includes low-level functions that allow you to fully customize the appearance of graphics as well as to build complete graphical user interfaces on your MATLAB applications.

External Interfaces

The external interfaces library allows you to write C/C++ and Fortran programs that interact with MATLAB. It includes facilities for calling routines from MATLAB (dynamic linking), for calling MATLAB as a computational engine, and for reading and writing MAT-files.

APPENDIX C

The most important matlab functions used in the presentation of the present thesis work are defined below with their proper description.

Fspecial

h = fspecial(*type*) creates a two-dimensional filter h of the specified type. fspecial returns h as a correlation kernel, which is the appropriate form to use with imfilter. *type* is a string having one of these values.

Imfilter

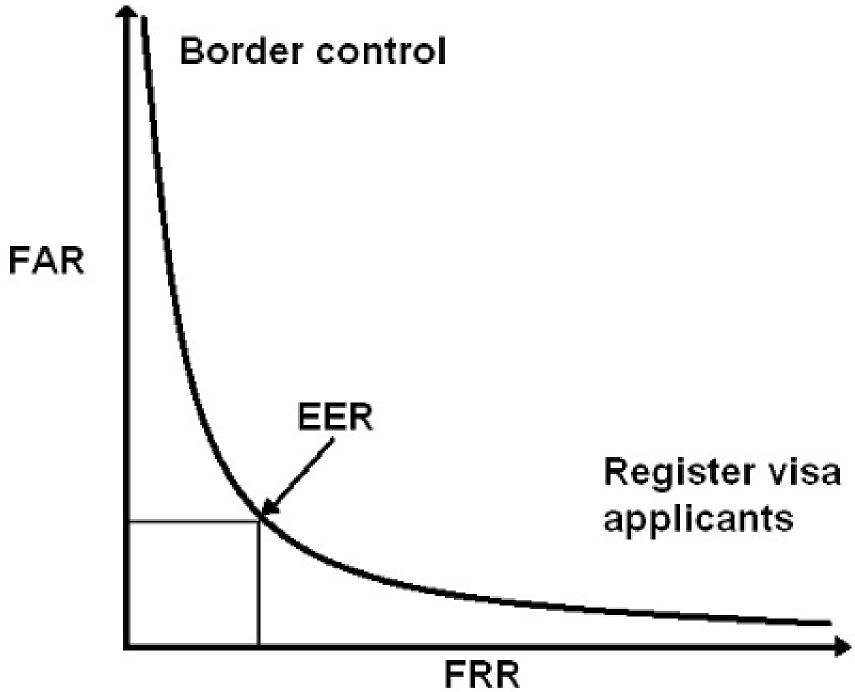
B = imfilter(A, H) filters the multidimensional array A with the multidimensional filter H. The array A can be logical or a nonsparse numeric array of any class and dimension. The result B has the same size and class as A.

Imhist( )

`imhist(I)` displays a histogram for the image `I` above a grayscale colorbar. The number of bins in the histogram is specified by the image type. If `I` is a grayscale image, `imhist` uses a default value of 256 bins. If `I` is a binary image, `imhist` uses two bins.

APPENDIX D

Critical importance of false acceptance rate.



The figure provides a graphical illustration of the relation between false acceptance and false rejection in a border control environment. In a border control we want a small False Acceptance Rate to prevent impostors using stolen visa's to pass themselves off as some one with a legal visa. At the registration of the visa, we want a small False Rejection Rate when we search to find whether or not the applicant previously have been issued a visa with another identity or whether or not the applicant is registered in a watchlist.