



A Robust High Performance Face Recognition using Decorrelation of Local Features by Discrete Cosine Transforms

Vasudha S

Research Scholar and Member, IEEE, vasudha.dsce@gmail.com

Neelamma K. Patil

Lecturer and Member, IEEE, neelammakletc@gmail.com

Dr. Chandrasekhar B,

Visvesvaraya Technological University, drchandrasekhar.b@gmail.com

ABSTRACT

This paper proposes a novel method of face recognition using de-correlation of local features using Discrete Cosine Transforms (DCT). The impulse for the proposed idea is with the fact that histograms DC constituent of local Gabor binary patterns constitute low frequency components which will sparsely help in actual recognition, because information in face resides in low frequency bands and is similar to all the images and also when these histograms are concatenated, it becomes difficult to differentiate and segregate actual frequency variations which add value for accurate recognition. A high correlation exists in between these histograms. This high correlation affects the recognition accuracy, hence de-correlation is achieved with the help of DCT application for individual histogram bins which aids in identification of actual frequency variations and highlights the changes in between two histograms thus improving the recognition accuracy. This method employs a non-statistical procedure which avoids training step for face samples thereby avoiding generalizability problem which is caused due to statistical learning procedure. The performance modeling is carried out by varying both internal and external factors of face recognition system and improvement is shown considerably high in terms of recognition accuracy and reduction in storage space by storing train images in compressed domain.

Index Terms— Discrete Cosine Transforms, External Factors, False Acceptance Ratio, False Rejection Ratio, Gabor filter, Internal factors, Local Binary Pattern, Performance modeling.



Council for Innovative Research

Peer Review Research Publishing System

Journal: INTERNATIONAL JOURNAL OF COMPUTERS & TECHNOLOGY

Vol 11 No 7

editor@cirworld.com

www.cirworld.com, member.cirworld.com

INTRODUCTION

In image analysis, face recognition is one of the most representative applications. In law enforcement and security applications, it has gained major importance. Face as a biometric is most popular because face image of a particular subject can be taken in a non-intrusive way without their co-operation. In general, problem statement of automated recognition of faces can be formulated as ‘Given still or video image of a scene, identify or verify one or more persons in the scene using a stored database of faces’ [17]. Solution to the above problem statement involves segmentation of faces - face detection from cluttered scenes as a first step followed with feature extraction from the face regions and recognition/verification depending upon the application it is used for. Face recognition flow chart is depicted in Fig. 1.

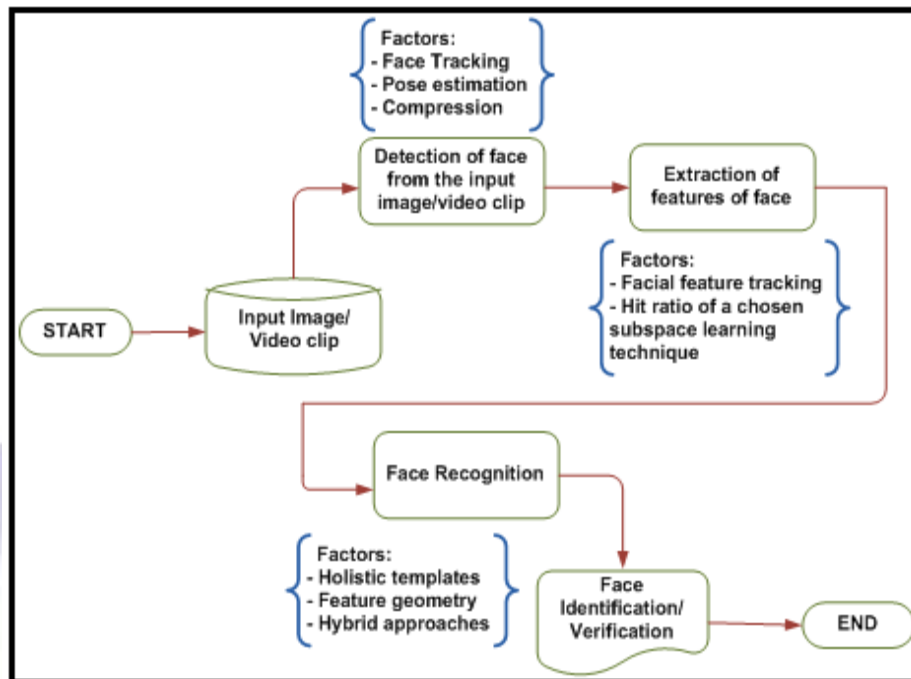


Fig. 1 A generic Face recognition system

In face recognition, crucial step is the assessment and benchmarking of algorithms. Although many face recognition techniques have been proposed and have shown appreciably promising results, design of robust face recognition system is still an open challenge. Illumination variations, pose/ alignment and recognition in outdoor imagery are three major challenges which affects the performance of face recognition. For example, pose discrimination is not that complex compared to accurate pose estimation. In addition to these two problems, there are bigger challenges such as, recognition of a person from images acquired years apart, involving variations in the overall appearance. The performance of face recognition system is limited due to various challenges such as modeling and normalization of illumination, pose variance / pose estimation, accurate landmark detection, feature extraction and model building, matching methodology / determination of distance measures, vast database to be handled (in most applications), subspace learning and so on. The performance of face recognition systems in different stages is depicted in Fig. 2. The factors which are affecting face detection phase are pose, alignment of face, number of frames in video input, background and so on. In feature extraction phase many factors like shape of the eyes, eyebrows, texture of skin, shape of the face needs to be considered strictly. In actual recognition, many challenges arise like rate of recognition, storage space for datasets and complexity of algorithm and so on.

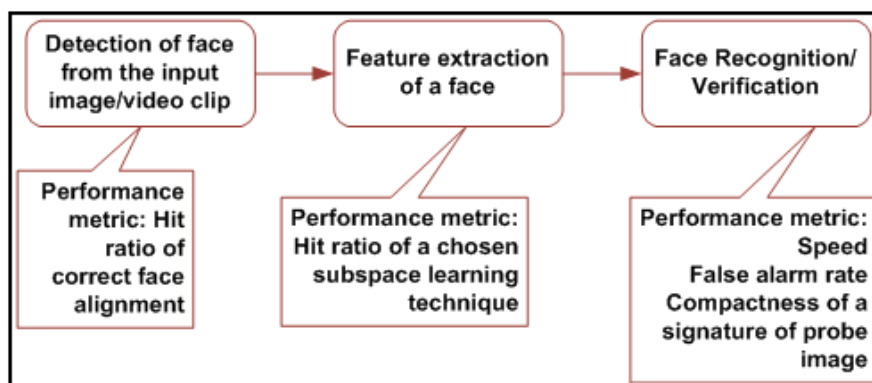


Fig. 2 Performance metric details at various stages of Face recognition

This paper proposes non-statistical method for face recognition which not only is robust to the variations of imaging condition but also embeds much discriminating power. It uses effective Gabor filtering of local features and forming histograms as the primary step. Successively DCT operator is applied for feature vector and thereby recognizing faces using similarity measures. It combines local intensity distribution and considers maximum values, which aids in segregating high frequency variations, making this approach a robust recognition technique. Also, storing train images in compressed domain solves security concerns of image splicing and aids in storage space reduction for the datasets. The performance of the proposed method is compared to the state-of-art method and results are documented in experimentation summary section.

The rest of the paper is organized as follows. Section II outlines literature survey, Section III describes the application of Gabor filters to face images and formation of local histogram feature vectors. Section IV details about Local Binary Patterns for GMPs, Section V describes about image compression using Discrete Cosine Transform (DCT) technique, Section VI provides the details of the proposed method of face recognition using DCT coefficient vectors. Experimental results and performance modeling based on the analysis of accuracy of recognition are discussed in Section VII, ending with conclusions in Section VIII and references.

LITERATURE SURVEY

In the course of studying papers which are referenced [2] through [7] in existing state-of-art literature, the concepts such as face image variations caused by different poses, inter-person differences used in distinguishing identities, related potentials in many applications dealing with uncooperative subjects, full power of face recognition as a passive biometric technique, problems due to varying illumination, generalizability problems, problems related to image splicing, performance limitations caused due to vast datasets have been understood.

Previous work carries many face representation approaches, most of them are based on statistical learning, such as linear discriminant analysis, Support Vector Machines or AdaBoost technique which inherently suffers from the generalizability problem due to the difference in between unpredictable distribution of the 'testing' face images from real world and the 'training' face images. In subspace approaches, a training set is developed in order to extract the most discriminating representation for final classification, but the discriminant subspace is greatly dependent on the training set [17]. If in case the testing images are captured under different environment from those of the training ones, already developed discriminant subspace would be inapplicable to those probe faces.

Though, generalizability problem has been theoretically discussed through Structural Risk Minimization (SRM) principle [1], in most real-world face recognition applications, generalizability problem remains as a huge bottleneck for most face recognition systems. Though there are many methods available for face recognition which avoids generalizability problem such as histogram method, multi resolution histogram method, Gabor feature based method, all these methods has its own pros and cons as listed in Table.1. A thought process map for generalizability problem is depicted in Fig.3.

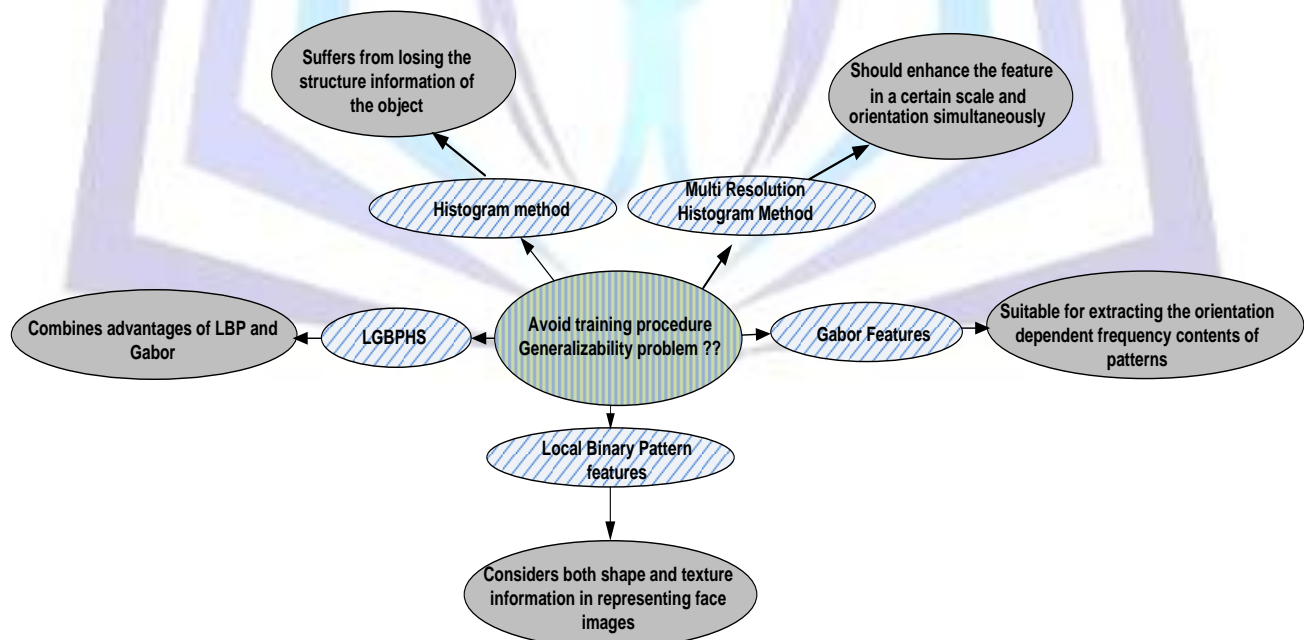


Fig. 3 Feature descriptors which avoids generalizability problem

Currently, many methods in the existing literature focus mainly on actual recognition algorithms. But, face images are extremely sensitive to illumination variation. This paper shows that, if the same importance is given to image preprocessing steps such as cropping and illumination normalization, a better performance can be achieved for any given face recognition system. Also, in certain applications such as airport surveillance systems, the systemic issues and computational complexity involved in storing huge train datasets poses a major challenge in performance of real-time face



recognition systems. Although there are commercial systems available in the market, it is still an active and challenging topic, due to the fact that the appearance of the same face looks dramatically different in uncontrolled environments with rich variations of pose, expression, illumination and occlusion.

TABLE I
PROS-CONS FOR VARIOUS FEATURE DESCRIPTORS

	Pros	Cons
Histogram Method	It can be calculated easily and efficiently and robust to the noise and local image transformations	The histogram is not adequate for many applications as it suffers from losing the structure information of the object
Multi-resolution histograms	Encodes the structure information by convolving the image with Gaussian filters	For a powerful descriptor, the feature extractor should be anisotropic - it should enhance the feature in a certain scale and orientation simultaneously.
Gabor transformation which is an effective element in the image processing and pattern recognition tasks.	It can characterize the spatial frequency structure in the image while preserving information of spatial relations and thus, it is suitable for extracting the orientation-dependent frequency contents of patterns.	Texture information not captured
Local Binary Patterns	Local Binary Pattern (LBP) histograms for face recognition considers both shape and texture information in representing face images.	Spatial relationship of images not captured
Gabor+LBP=LGBPHS	Combines the magnitude part of Gabor feature and the LBP operator with which face image can be seen as a composition of micro-patterns.	Better representation which will be robust to noise is required Extension of the neighborhood size should be addressed To reduce length of histogram sequence feature by removing the redundancy in original histogram features to make the representation more effective and compact.

GABOR MAGNITUDE PICTURES FOR FACE IMAGE

This section describes the filtering of face image and formation of GMPs using Gabor filters. Starting with the introduction of Gabor filters and its construction, the basic concepts of feature extraction for a face using the Gabor filter bank is explained and proceeds with the presentation of the Gabor Magnitude Picture (GMP of a face) representation. Gabor filters which are also called as Gabor wavelets or kernels have been proven to be efficient tool for facial feature extraction and robust face detection and recognition. They represent complex band-limited filters with an optimal localization in both the spatial and frequency domain. Hence, when employed for facial feature extraction, they extract multiresolutional, spatially local features of a confined frequency band [2]. Like all filters operating in the scale-space, Gabor filters also relate to the simple cells of the mammalian visual cortex and hence relevant from the biological point of view as well. In general, the family of 2D Gabor filters can be defined in the spatial domain as follows [8]:

$$g_{p,q}(a, b) = \frac{f_p^2}{\pi^{i+c+\eta}} \exp\left(-\left(\frac{f_p^2}{c^2} a^2 + \frac{f_p^2}{\eta^2} b^2\right)\right) * \exp(j * 2 * \pi * f_p * a') \tag{1}$$

$$a' = a \cos \theta_q + b \sin \theta_q \tag{2}$$

$$b' = -a \sin \theta_q + b \cos \theta_q \tag{3}$$

$$f_p = \frac{f_{max}}{2^{\frac{p}{2}}} \tag{4}$$

$$\theta_q = \frac{q\pi}{8} \tag{5}$$

Each Gabor filter represents a Gaussian kernel function modulated by a complex plane wave whose center frequency is f_p and orientation is θ_q . The parameters 'c' and 'η' determine the ratio between the center frequency and the size of the Gaussian envelope. While different choices of the parameters determining the shape and characteristics of the filters define different families of Gabor filters, the most common parameters used for face recognition are $c = \eta = \sqrt{2}$ and

$f_{max}=0.25$ [9, 10, 11]. When using the Gabor filters for facial feature extraction, researchers typically construct a filter bank featuring filters of five scales and eight orientations, that is, $p = 0, 1, \dots, r - 1$ and $q = 0, 1, \dots, s - 1$, where $r = 5$ and $s = 8$. The real parts of the entire filter bank will be used for facial feature extraction (comprised of 40 filters) as shown in Fig.4.

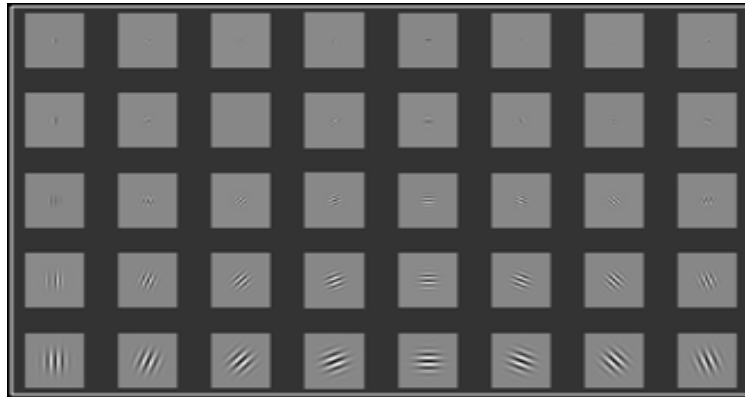


Fig. 4 Real parts of the Gabor filter bank.

Let $I(a, b)$ stand for a grey-scale face image of size 128×128 pixels and $g_{p,q}(a, b)$ denote a Gabor filter given by its center frequency is f_p and orientation is θ_q . The feature extraction is defined as a filtering operation of the given face image $I(a, b)$ the Gabor filter $g_{p,q}(a, b)$ as shown below.

$$G_{p,q}(a, b) = I(a, b) * g_{p,q}(a, b) \tag{6}$$

$G_{p,q}(a, b)$ is decomposed into real and imaginary parts as $Re[G_{p,q}(a, b)]$ and $Im[G_{p,q}(a, b)]$ respectively. Based on this, magnitude and phase responses are given by following equations.

$$M_{(p,q)}(a, b) = \sqrt{Re^2[G_{p,q}(a, b)] + Im^2[G_{p,q}(a, b)]} \tag{7}$$

$$\Phi_{(p,q)}(a, b) = \arctan \left(\frac{Im[G_{p,q}(a, b)]}{Re[G_{p,q}(a, b)]} \right) \tag{8}$$

The Gabor Magnitude Picture (GMP) for a face image formation includes the construction of the Gabor filter bank and filtering the face image with all constructed Gabor filters from the filter bank. An example of the Gabor magnitude output for a sample image from FERET database and the magnitude output of the filtering operation with the entire Gabor filter bank of 40 Gabor filters is depicted in the following Fig.5. All the Gabor magnitude then concatenated into the final Gabor magnitude face representation named by Liu and Wechsler [4], into the augmented Gabor feature vector

$$x = (g_{0,0}^T, g_{0,1}^T, g_{0,2}^T, \dots, g_{r-1,s-1}^T)^T$$

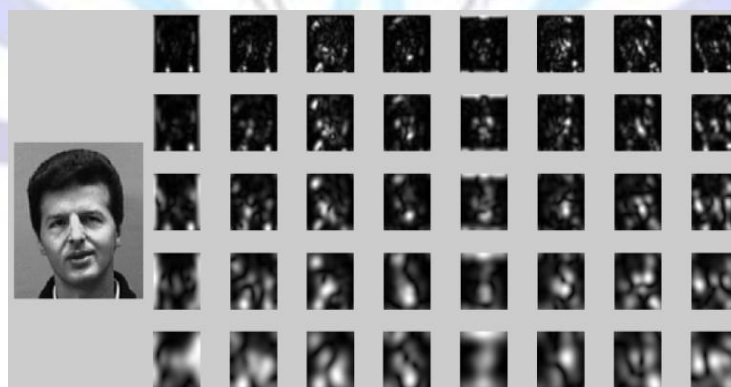


Fig. 5 Example sample face image from FERET database and the respective GMP.

LOCAL BINARY PATTERN FOR GABOR MAGNITUDE PICTURES OF FACE IMAGE

Local Binary Pattern (LBP) was firstly designed for texture classification. Ahonen et al. [12] successfully applied it to represent faces. LBP considers both shape and texture information to represent the face images. In this approach the textures of the facial regions are locally encoded by the LBP patterns while the whole shape of the face is recovered by the construction of the face feature histogram. The idea behind using the LBP features is that the face images can be seen as composition of micro-patterns such as flat areas, spots, lines and edges which are invariant with respect to monotonic grey scale transformations [13, 14]. Combining these micro-patterns, a global description of the face image is

obtained. The basic LBP operator labels the pixels of an image by thresholding 3x3 neighborhood of each pixel with the center value and considering the result as a binary number as depicted in Fig.6. Later histogram of the labels is used as a texture descriptor.

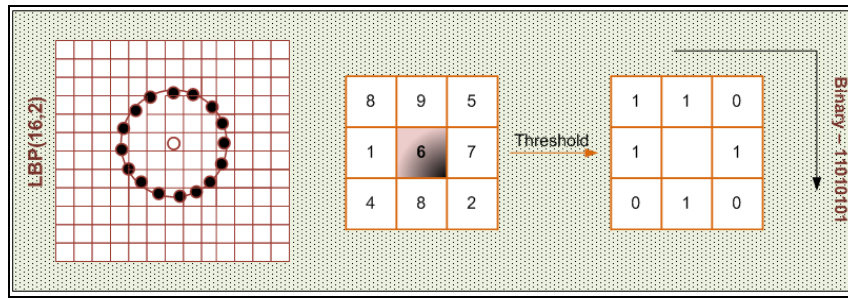


Fig.6 Basic LBP operator for 8 sampling points on a circle of radius 1.

A histogram of the labeled image $I(a, b)$ can be defined as below, where 'm' is the number of different labels produced by the LBP operator. This histogram contains information about the distribution of the local micro patterns, such as edges, spots and flat areas, over the whole image. In this paper, the proposed method applies LBP operator for all 40 GMPs and respective histogram are concatenated to build a global description of the face.

$$H_j = \sum_{(a,b)} P[I(a, b)] = j, j = 0, 1, \dots, m - 1 \tag{9}$$

$$P[B] = \begin{cases} 1, & b \geq 0 \\ 0, & b < 0 \end{cases} \tag{10}$$

IMAGE COMPRESSION USING DISCRETE COSINE TRANSFORMS

This section explains briefly about image compression using Discrete Cosine Transform (DCT) technique. The DCT helps separate the image into parts (or spectral sub-bands) of differing importance with respect to the image's visual quality. The general equation for a 2D DCT of an $N \times M$ image is defined by the following equation.

$$D(p, q) = \frac{1}{\sqrt{2N}} c(p) c(q) \sum_{a=0}^{N-1} \sum_{b=0}^{N-1} I(a, b) \cos \left[\frac{(2a+1)p\pi}{2N} \right] \cos \left[\frac{(2b+1)q\pi}{2N} \right] \tag{11}$$

$$c(r) = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } r = 0 \\ 1 & \text{if } r > 0 \end{cases} \tag{12}$$

Image $I(a, b)$ is the a th and b th element of image represented by matrix 'I'. 'N' is the size of the block on which DCT is operated. The 2D DCT equation computes one entry of transformed image using pixel values of original image matrix. For standard 8×8 block of an image, $N=8$, (a, b) range from 0-7, DCT is calculated using $D=SCS^T$, where 'C' is the 8×8 block of an image and 'S' is represented using the following equation.

$$S_{(p,q)} = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } p = 0 \\ \sqrt{\frac{2}{N}} \cos \left[\frac{(2a+1)p\pi}{2N} \right] & \text{if } p > 0 \end{cases} \tag{13}$$

First 8×8 block of an image is multiplied by 'S' to transform rows and then it is multiplied by 'S' to transform columns. This results in a block matrix with 64 DCT coefficients $C_{p,q}$ where p and q range from (0, 7), in which the top-left coefficient c_{00} correlates to low frequencies of original image block and c_{00} corresponds to higher frequencies. It must be noted that human eye is most sensitive to low frequency components.

Block coefficient matrix which is derived from previous step is compressed using quantization. Varying level of image compression can be chosen using specific quantization matrices depending on the intended image quality intended for future processing. Hence there will be a trade-off between quality and compression to suit different needs. There exists a standard image compression matrix 'Q₅₀' which considers human visual system and helps to achieve quality level of 50 and high compression with excellent decompression quality as well.

Quantization is achieved by dividing each element in the block coefficient matrix by corresponding element in the Q₅₀ matrix and rounding off the results to the nearest integer value. In this course of division, many of the coefficients become zeroes. Also coefficients situated near the upper left corner corresponds to lower frequencies to which human eye is most sensitive and can be retrieved and lower right coefficients which includes most of the zeroes corresponds higher frequencies which can be discarded. Also only non-zero coefficients are used for image reconstruction.

$$R = \text{round} \left[\frac{c_{(p,q)}}{Q_{(i,j)}} \right] \tag{14}$$

In the final step of compression, all coefficients of 'R' are encoded to a stream of binary data. It must be noted that after quantization most of the coefficients will become zero and hence JPEG takes advantage of this fact and follows

encoding in zig-zag manner as depicted in the Fig.7. Decompression is carried out by multiplying the corresponding element of quantized matrix 'R' with quantization matrix Q_{50} .

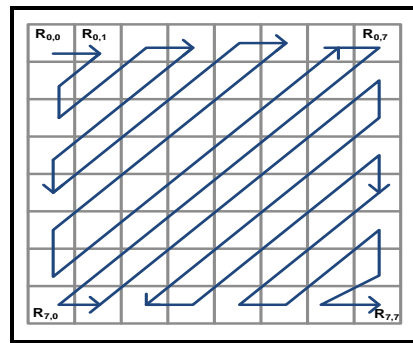


Fig.7 Encoding scheme in zig-zag manner.

IMAGE COMPRESSION

The main aim of this method is to achieve performance improvement of Local Gabor Binary Pattern Histogram Sequence method in terms of recognition accuracy. To achieve this, image preprocessing is also considered as the main factor. The decorrelation of local features using DCT helps in recognition of faces with expression and pose variations accurately. The secondary aim of this method is to alleviate the problems caused by vast datasets. To achieve this, train dataset is stored in compressed domain. This helps in reducing the storage space and avoids image splicing and preserve the privacy of datasets. Detailed flowchart of the proposed method is depicted in Fig.8. The following sub sections explain all the steps in detail.

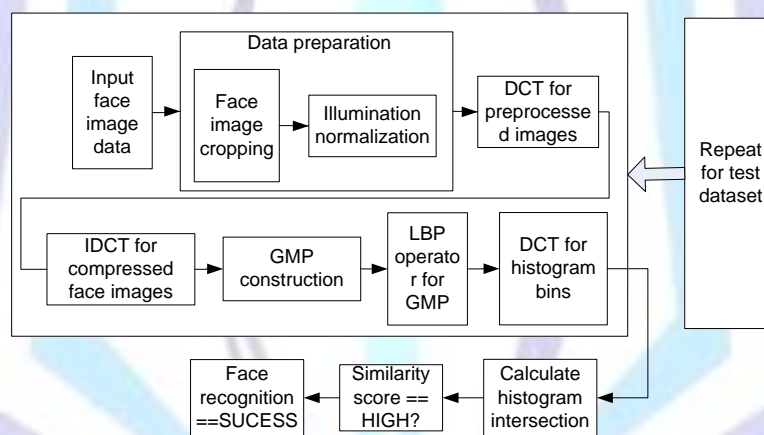


Fig.8 Block diagram for the proposed method.

A. Preparation of input data using preprocessing steps

The preprocessing is a very essential step in image processing. Many literatures and recognition methods focus more on actual recognition algorithms rather than cleaning up the input data such that the expected result matches the actual result thereby improving the accuracy. In this paper the experimentation is carried out after careful preprocessing of the images and a clear difference is observed with accuracy of face recognition results. Data preparation is carried out in two steps. They are cropping and illumination normalization.

The images are cropped according to the provided eye positions, resized to 50 x 80, succeeded by illumination normalization. The reason behind taking 50x80 [width x height] sized images is that face has more valuable information in vertical direction. Also with resized 50x80 images, memory required for storage and processing time will be reduced considerably which leads to performance improvement to face recognition system.

Main aim of image cropping step is to extract more information from face images which actually add importance and enhances the accuracy of face recognition. In dynamic handling of real time systems, some typical situations occur such that, the distance between camera and person will not be uniform always. This may lead to capturing of face image of a person along with hairs, head, neck, shoulders, hands, etc. It is essential to remove this information that fetches less or no value for actual recognition but may mislead the system. Hence, cropping is an essential step which helps to retain only necessary image information. The steps involved in cropping are as follows.

- The Canny edge detector [15] is applied to the original face image of size 256x384 and gray-scale image is converted into binary image. In binary images edges are clearly visible.
- Horizontal and vertical projections of binary image are taken.

- The maximum bins in horizontal and vertical projection are calculated. This indicates the location of eyes and aids in marking the same.
- Once eye position is marked, width of the face is calculated as the face is symmetrical with respect to eye positions.
- Finally, a border of face is marked with the help of width of face, eye location and edges in binary image.
- This cropped face image is resized to 50 x 80.



Fig.9 Sample face images from FERET database (Cropped and resized to 50 x 80).

This section explains details of illumination normalization based on histogram of an image and scaling function. In order to develop adaptive contrast enhancement method which suits for all kinds of images of differing illumination levels, we studied the histogram of couple of images (considerably more images with different illumination variations) and concluded that histogram is not concentrated uniformly throughout. As illumination varies, histogram distribution in each bin changes and it is concentrated at particular left, right and corner regions as shown in Fig.10. Analyzing this, we propose a novel enhancement algorithm and it is providing promising result for low, medium and high contrast images. It balances the images in contrast and illumination variations with improved and enhancing edges and preserving the same. This leads to explore more information in enhanced images and helps in improvement of face recognition accuracy.

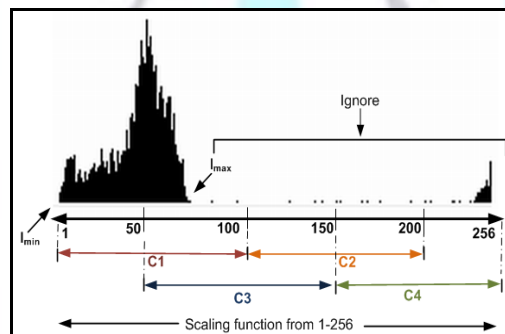


Fig.10 Study of histograms.

The steps involved for the proposed method are as follows.

- Histogram for cropped image is taken and its maximum value is calculated and it assigned as max_1 . Second maximum value is assigned as max_2 and threshold is set equal to $max_2 / 4$.
- Based on observation of histogram concentration in different ranges, histogram intensity range is divided into different ranges such as 1-100, 50-150, 100-200 and 150-256.
- The values which are greater than or equal to threshold in particular region of interest are counted, and is set as C1, C2, C3, C4 corresponding to different regions.
- Maximum value among count values of previous step is identified and set as $C_{max} = \max(C1, C2, C3, C4)$.
- It so happens that there can be two count values which are very near or equal. In such case, higher range is selected.
- Based on maximum count value and particular range, the limit for histogram is calculated by searching for histogram values less than two on left and right side of histogram. The left minimum value and right minimum values are denoted by I_{min} and I_{max} which represent the intensity values. This acts as a terminator on both the sides to set I_{min} and I_{max} .
- Values which are below intensity I_{min} in the original image are set to I_{min} and to intensity I_{max} which are greater than I_{max} in the original image.
- Finally scaling function is applied, which scales function from 1 to 256 to create missing intensity values.

Histogram of an image after applying proposed above algorithm may not have all the intensity values in it. In order to enhance and make all the intensity levels to be present which makes uniform variations in illumination, we use scaling function. The scaling function scales given range of input from 1 to 256. The algorithm is as follows and the sample preprocessed face image is shown in Fig. 11.

- The maximum and minimum value present in an image is calculated.
- If the difference between maximum and minimum value is zero, then pixel values of image is assigned to zero.

- If not, subtract minimum value from each pixel value in an image and divide by the difference of maximum and minimum value and finally multiply by 255 and the result of is rounded off.
- The above steps are repeated for all pixel values in an image.

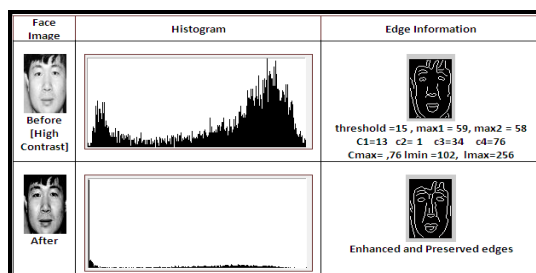


Fig. 11 Sample per-processed face image from FERET database.

B. Storing preprocessed datasets in compressed domain

In most of the face recognition/verification applications, problems caused by vast datasets are still an open challenge. This memory constraint is solved by storing train and test images in compressed domain. During runtime decompression is applied to the compressed images. The preprocessed images are compressed using 2D-DCT and encoded using Q50 quantizer from JPEG standard as shown in the Fig.12. The non-zero significant DCT coefficients which are less in number are stored in the database which leads to huge reduction in storage space. The fact that face images carry information in lower frequency band helps in making most of the DCT coefficients zero. Exploiting this factor aids to achieve high compression and to save memory. The computation of feature vector is performed by decompressing stored DCT coefficients followed by de-quantization and reconstruction of train images as shown in Fig.13. As and when new images added up with train dataset, the process of compression and decompression can be adopted as an automated process.

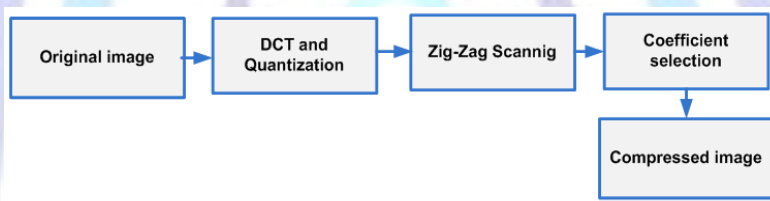


Fig. 12 Compression of train database images

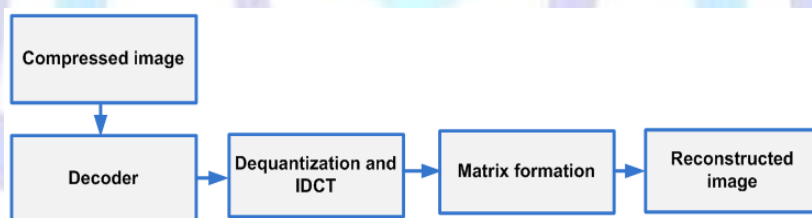


Fig. 13 Decompression of DCT coefficients

The pre-processed face image as shown in the Fig.14 is divided into 8x8 blocks and only non-zero coefficients are saved out of all coefficients in proper order as depicted in Fig.15 and 16. At the receiving end, with the help of de-quantization followed IDCT, original image is reconstructed and no noticeable losses are observed.

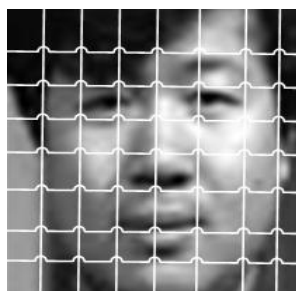


Fig. 14 DCT for sample pre-processed face image.

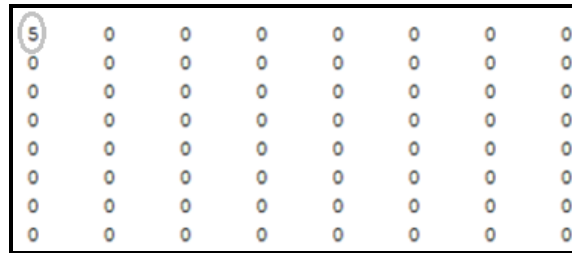


Fig. 15 Only one non-zero coefficient is retained.

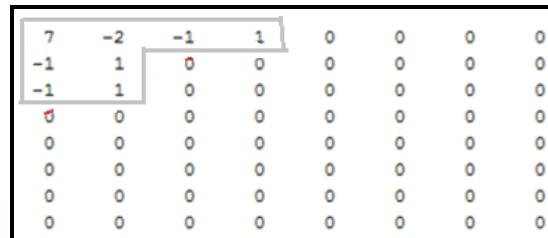


Fig. 16 Non-zero coefficients retained.

High compression is achieved by applying DCT followed by Q50 quantizer as most of the DCT coefficients zero. The same is explained in the in the form of blocks of DCT coefficients in Fig.15. Depending on the information present in an image block, it is essential to select non-zero DCT coefficients by ignoring zero coefficients. At the time of decompression a matrix is formed based on this non-zero coefficients and making remaining places filled with zeros. Then dequantization is applied followed by IDCT to reconstruct an image block.

C. Application of DCT to local features

During the runtime, the Gabor operator is applied for decompressed train images and Gabor Magnitude Pictures (GMP) are constructed. Successively Local Binary Pattern is applied on GMPs and local histograms are formed. The conceptual definition for the proposed idea lies in the fact that the DC component of histograms of local Gabor binary patterns constitutes low frequency components which will not help in recognition and also when these histograms are concatenated, it becomes difficult to differentiate and segregate actual frequency variations which add value for accurate recognition. There exists a high correlation in between these histograms. This high correlation affects the recognition accuracy, hence de-correlation is achieved with the help of DCT application for individual histogram bins which aids in identification of actual frequency variations and highlights the changes in between two histograms thus improving the recognition accuracy. This also exploits full utilization of Gabor filtering and face recognition among faces which has expression variations is accurate. And also one important point is to be noted that if recognition is carried out with application of DCT after concatenation of histograms hits accuracy of recognition, because there is high a chance of mixing of various frequencies from the Gabor transform.

This section explains DCT coefficients used for decorrelation of local features. DCT helps to separate the image into parts (or spectral sub-bands) of differing importance with respect to the image's visual quality [16]. The general equation for a 1D DCT of an NxM image is defined by the following equations.

$$DCT(p) = \sqrt{\frac{2}{N}} \sum_{\alpha=0}^{N-1} c(r) \cos \left[\frac{(2\alpha+1)p\pi}{2N} \right] H_j \tag{15}$$

$$c(r) = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } r = 0 \\ 1 & \text{if } r > 0 \end{cases} \tag{16}$$

After application of 1D-DCT, the left coefficient c_0 corresponds to low frequencies of original vector and right coefficient c_7 corresponds to higher frequencies. It must be noted that human eye is most sensitive to low frequency components. In the case of face images, low frequency components are same for all images, hence these does not add any value for matching. After applying DCT for individual histograms, they are rounded using $R=\text{round} [DCT(p)]$ and concatenated and then it is stored as a single feature vector. The variation in frequency leads to better result in face recognition system.

Finally histogram intersection is calculated and it is used as similarity measurement for the proposed method. This similarity measure is chosen because comparison of proposed method with the state-of-art method will be straightforward. The common part of histograms is calculated using the following formula, where H^G and H^P are local feature histograms of gallery and probe images.

$$\Delta[H^G, H^P] = \sum_{j=1}^N \min [h_j^G, h_j^P] \tag{17}$$

The sample preprocessed face image from FERET database and corresponding retrieved identity which is having a different expression from test database is shown in Fig.17. Fig. 18 shows the state of art method in which local feature histograms are concatenated and in turn histogram intersection is calculated as similarity measure in Fig.19. In this case

disadvantage is highly correlated histograms may not lead to good recognition accuracy since utilization of Gabor with different scales and orientation are complete.

Hence in the proposed method, to alleviate this problem, DCT is applied to individual histogram bins as shown in the Fig.20 and histogram intersection is calculated as similarity measure as depicted in Fig.21. An alternative method is also proposed in which DCT is applied after concatenation of local feature histogram as shown in Fig. 22 and 23. The advantage of this method is, clear differentiation among the train and test histograms which are decorrelated by DCT is achieved and this leads to an accurate recognition even in case of facial expression changes. In both the cases, recognition accuracy is improved because actual frequency variations are identified and changes in between two histograms are prominent.

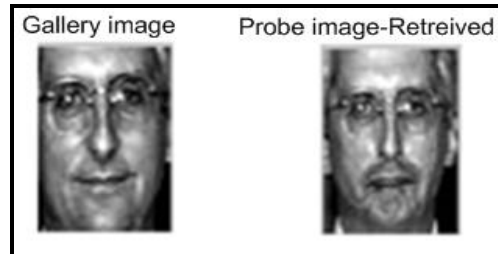


Fig. 17 The sample pre-processed face image from FERET database and corresponding retrieved identity which is having a different expression (from test database)

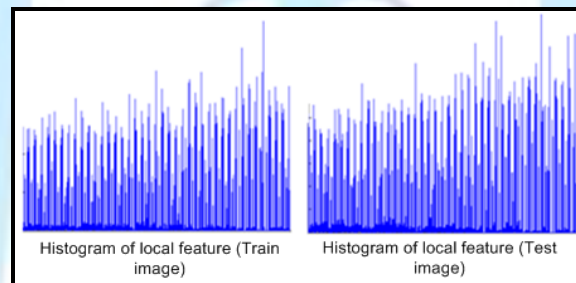


Fig. 18 Local feature histogram

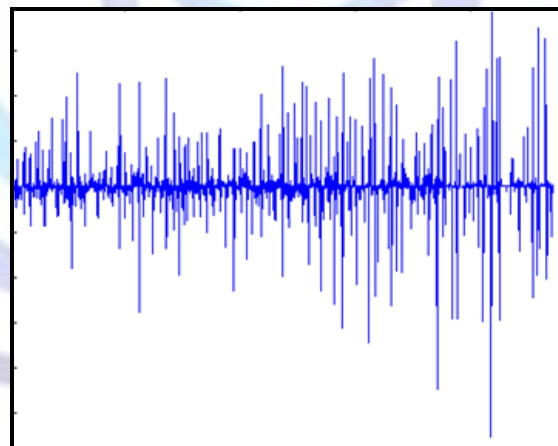


Fig. 19 Histogram Intersection

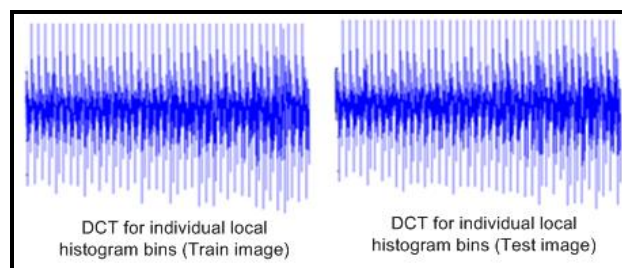


Fig. 20 DCT applied for local feature histogram bins

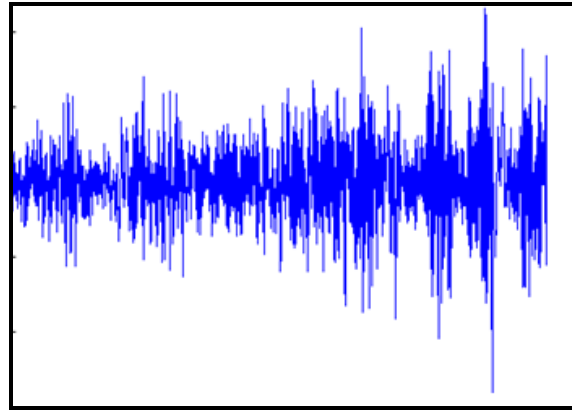


Fig. 21 Histogram Intersection

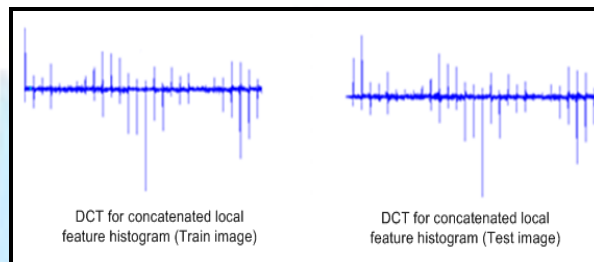


Fig. 22 DCT applied after concatenation of local feature histogram

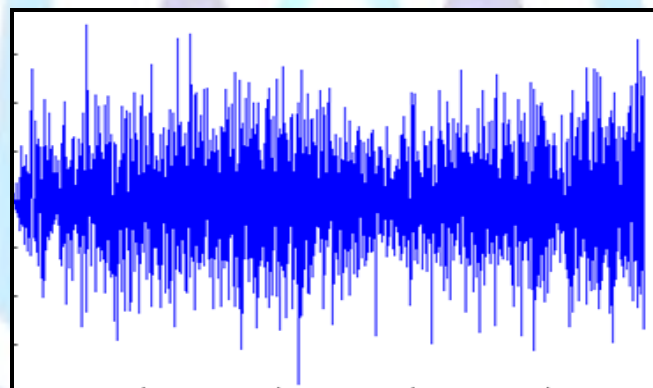


Fig. 23 Histogram Intersection

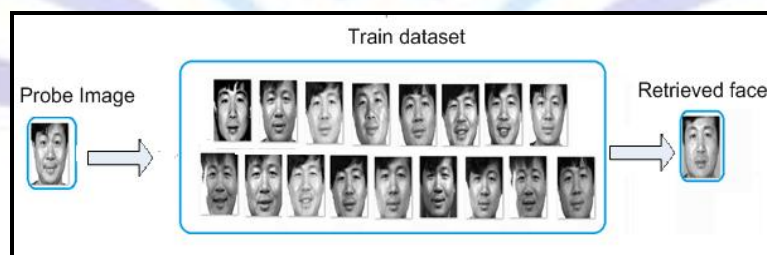


Fig. 24 Face recognition for sample FERET pre-processed image

EXPERIMENTAL ANALYSIS AND PERFORMANCE MODELING

A. Face Variations in terms of appearance

Face variations includes various aspects such as variations with respect to pose, alignment, expression, variations with respect to ethnicity, variations with respect to illumination variation, variations with respect to gender and aging and so on. A detailed study on face variations is carried out for the sake of our experimentation which helped extensively for the analysis of feature histograms. Faces, in general can be classified into two broad categories, that is, static faces with no pose variations and dynamic poses where face alignment and illumination variation exists. In the static faces, there are many variations with respect to ethnicity, gender and aging. The variations of faces according to different races can be categorized using mathematical proportions. One of the ratios defined by ancient Greeks is the concept of “phi,” the golden ratio of 1:1.618. Phi is the unique point on a line that divides the line into two lines in such a manner that the ratio

of the smaller portion to the larger portion is the same as the ratio of the larger portion to the whole line. Different geographic races with particular genetic similarities are Europeans, Eastern Indians, Asians, American Indians, Africans, Melanesians, Micronesians, Polynesians, Australians and Aborigines. But these races can be accommodated into three basic groups as Caucasians, Asians and Africans. Based on the anatomic structural differences and phenotypic characteristic, certain faces will have a broader nasal base, decreased nasal projection, protrusion of upper and lower jaws, bulging of eyes - orbital proptosis, increased soft tissue of the mid face, prominent lips, and increased facial convexity, faces with a high / low dorsum – bridge of the nose feature, distance between the eyes can be small/medium/large, shape of the chin, size of the lips and so on.

When face recognition is carried out with gallery images which are taken years apart from probe images, aging factors are required to be considered. Due to several physiologic changes that occur with time, the collagen in the face becomes less elastic and allows the soft tissue features of the face to sag inferiorly. In addition facial fat is generally lost with age making the facial soft tissues thinner. Finally the facial soft tissues gradually wither making the facial bones appear more prominent. Hence the sagging of cheek results in cheeks, tissues around eyes, eye lids, drift of tissues on forehead, wrinkled around eyebrows, elongated nose and so on.

Along with natural face variations, manmade appearance changes to faces for their disguise also need to be considered. For example face with mustache, beard, clean face without mustache / beard, faces with different hair styles like loose hairs, tied hairs, balded, with / without eye glasses, variations due to eyebrow shapes, variations due to hat / caps and so on.

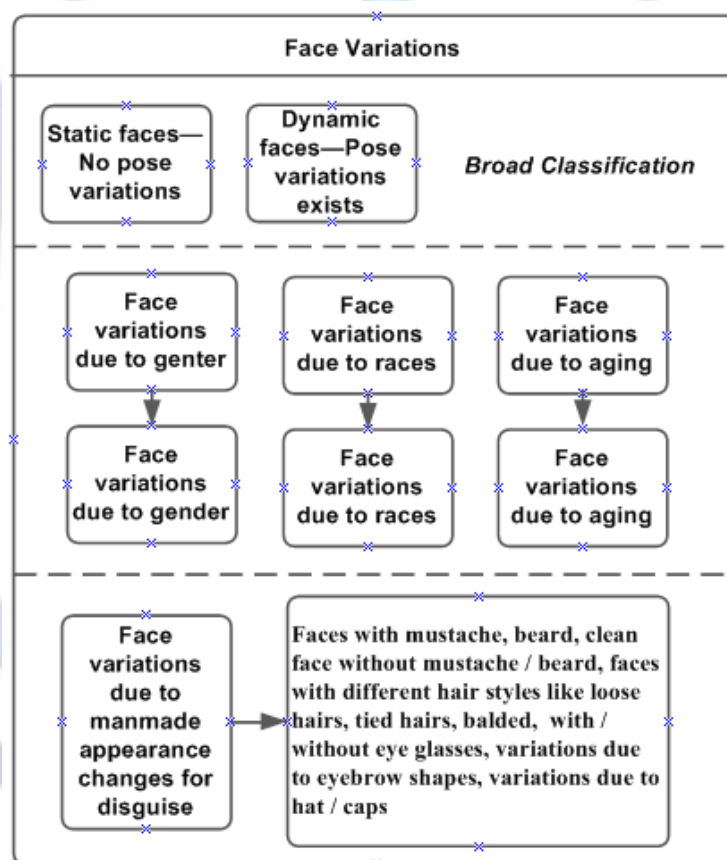


Fig. 25 Face Variations

B. Experimental dataset details

This section presents the experimental dataset details and performance measures used to evaluate the feasibility of the proposed method. FERET database is selected for the experiments to ensure that our results are comparable to other existing literatures. All the images are cropped to 50x80 according to the provided eye positions succeeded by histogram equalization preprocessing. Three databases are formed from FERET database, BioID database and MIT-CBCL-facerec database and labeled as DB1, DB2 and DB3 respectively. Details of FERET database is depicted in the Fig.26. Sample face images from BioID and MIT database are showed in Fig.27 and 28 respectively.

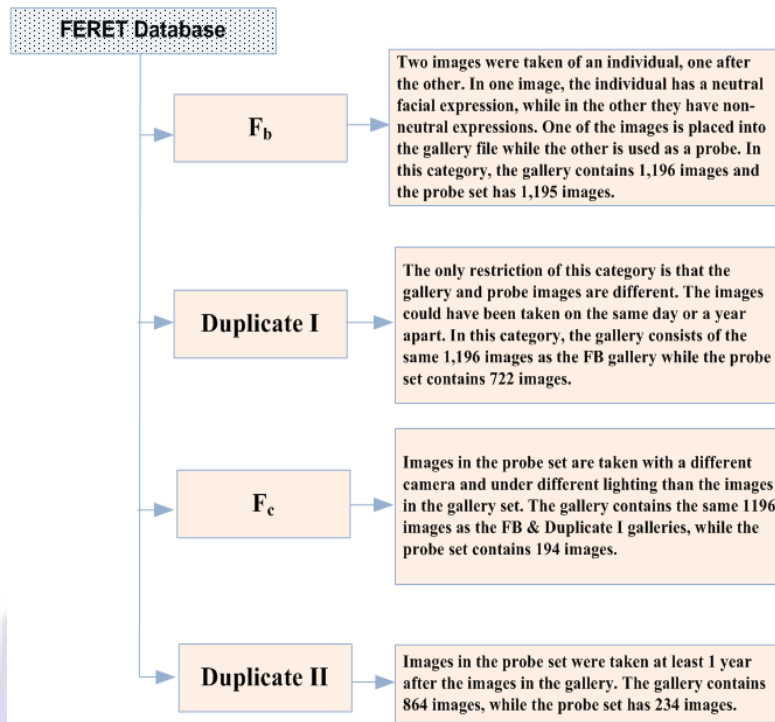


Fig. 26 FERET Database



Fig. 27 Sample face images from BioID database.



Fig.28 Sample face images from MIT-CBCL-facerec database.

C. Performance modeling of face recognition

Performance of face recognition is dependent on both internal and external factors. Internal factors include the actual algorithm used for recognition and train image dataset. External factors include variation of test image dataset in terms of illumination, pose, expression and so on. In a face recognition application, the function of system is to map a query image to a label that represents its identity as depicted in Fig.29. [18]. The gallery dataset is denoted as $I(G) = [i_{g_1}, i_{g_2}, i_{g_3}, \dots, i_{g_x}]$ consisting of 'x' gallery images whose identity is known to the face recognition algorithm. The query set is denoted as $I(Q) = [i_{q_1}, i_{q_2}, i_{q_3}, \dots, i_{q_y}]$ consisting of 'y' query images whose identity is unknown to the face recognition algorithm. A face recognition algorithm measures the similarity between query images and each gallery image. The measure which is used in calculating the similarity in between query image and the probe image plays an important role since it connects both internal and external factors of an recognition system [18]. The similarity score is denoted by $S[i_{g_x}, i_{q_y}]$ for the comparison between the query image i_{q_y} and the gallery image i_{g_x} . For a perfect match, similarity score should be potentially larger quantitatively.

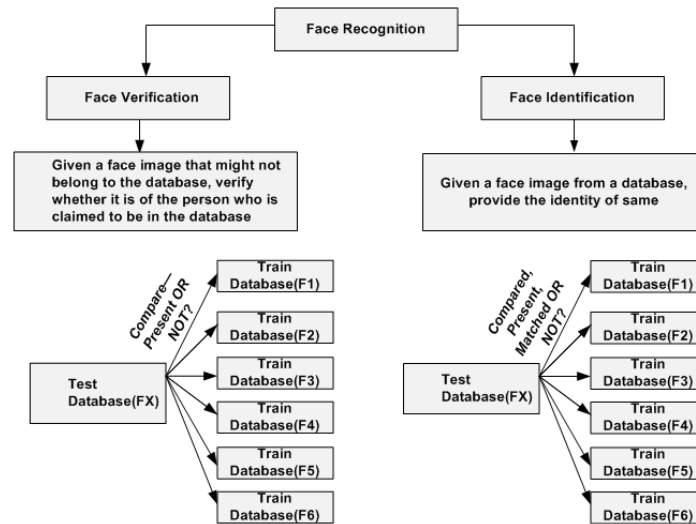


Fig.29 Difference between face recognition and face verification

For the sake of performance modeling, all the similarity scores of a query image i_{qy} are sorted in descending order and normalized to the range [0, 1]. With this, the similarity scores of using different internal factor (in this case it is different recognition algorithm) can be compared. After sorting and normalization, the set of similarity scores for data i_{qy} are represented by following equation.

$$S_{i_{qy}} = [S(i_q, i_{g_1}), S(i_q, i_{g_2}), S(i_q, i_{g_3}), \dots, S(i_q, i_{g_x})] \tag{18}$$

where i_{g_x} is the index of the gallery image corresponding to the x^{th} sorted similarity score. For each query image it is assumed that there is one and only one gallery image from the same person and hence the largest similarity score is termed as ‘matching’ score and the remaining similarity scores are termed as ‘non-matching’ scores [18].

A term called “perfect recognition” is introduced for performance modeling of proposed face recognition method [18]. The gallery dataset is duplicated to produce the query dataset. The duplicated set is used for recognition and similarity scores of the each query image are obtained. Such similarity scores are called “Perfect Recognition” (PR) [18]. Quantitative characterization of the difference between matching and non-matching similarity scores for data i_q , a performance metric ‘P’ is defined as

$$P = e^{\left[\frac{S(i_q, i_{g_1}) - \mu_{i_q}^{xy}}{\sigma_{i_q}^{xy}} \right]} \tag{19}$$

where, $S(i_q, i_{g_1})$ is the matching score, $\mu_{i_q}^{xy}$ and $\sigma_{i_q}^{xy}$ are mean and standard deviation of non-matching scores of $S(i_q, i_{g_2}), S(i_q, i_{g_3}), S(i_q, i_{g_4}), \dots, S(i_q, i_{g_x})$. The mean of ‘P’ is defined as

$$p = \frac{\sum i_q P i_q}{x} \tag{20}$$

For the quantitative demonstration, the following experimental steps are carried out. The internal factor being the face recognition algorithm is varied and ‘P’ is calculated using only gallery data. As a second step, actual recognition rate ‘β’ is calculated using different recognition algorithms. It shows that the system with better recognition accuracy has a larger difference between matching and non-matching similarity scores, thus corresponding ‘P’ and ‘β’ are larger. The value of ‘P’ is bigger with proposed method when compared with that of state of art methods thus concluding the accuracy improvement.

For the experimental purpose, 500 images from BioID database are selected randomly and gallery set is formed with neutral expression and the partially occluded images are used as the probes. From MIT-CBCL-facerec database, 200 images per 10 subjects are taken and randomly divided into gallery and probe images. From FERET database - f_b set, 800 faces with expression variation and from f_c set, 500 faces with lighting variation are selected. The outcome of the experiments shows that proposed method remarkably highlights the robustness and high discriminating capacity of face features and improved performance. Recognition results are tabulated in Table II in comparison with existing state-of-art methods available in literature.

TABLE II
COMPARISON OF RECOGNITION RESULTS

Methods	DB1	DB2	DB3
Principal Component Analysis method	0.92	0.85	0.88
LGBPMS method	0.94	0.89	0.89
Proposed method	0.96	0.91	0.93

The recognition performance for the proposed method is modeled by the standard error rate and recognition rate commonly used in the face recognition field. False Acceptance is defined as an incorrect prediction that recognizes an incorrect match as a correct match. The false negative, on the other hand, is incorrect prediction that recognizes a correct match as an incorrect match. The plots of FAR, FRR and HTER are shown in Fig.30 and 31 respectively.

$$FAR = \frac{\text{Total No. of false positives}}{\text{Total no. of images in test database}} \tag{21}$$

$$FRR = \frac{\text{Total No. of false negatives}}{\text{Total no. of images in test database}} \tag{22}$$

$$\text{Half Total Error Rate} = \text{HTER} = 0.5(FAR + FRR) \tag{23}$$

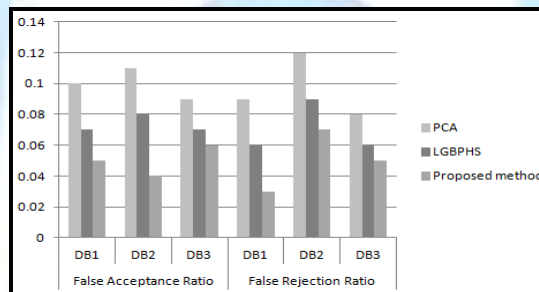


Fig.30 FAR and FRR plots for DB1, DB2 and DB3 databases.

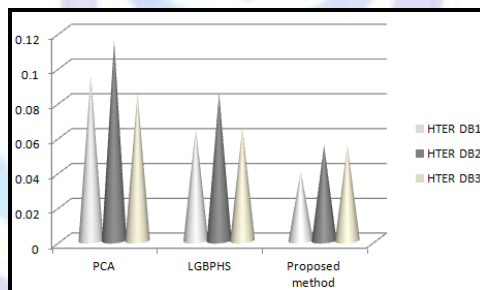
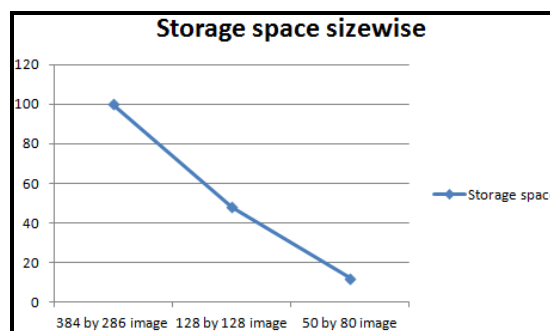


Fig.31 HTER plots for DB1, DB2 and DB3 databases.

Another performance measure taken into account is the storage space for train and test databases and processing time for the recognition. It is being taken care that all the images are preprocessed and cropped to 50x80 such that it occupies only 1/4th space that of original 128x128 sized images. If the database is stored in compressed domain storage space reduces 1/3rd that of uncompressed images storage space and processing time is reduced by a factor of 3. The performance plots of storage and processing time is depicted in Fig.32 and 33 respectively.



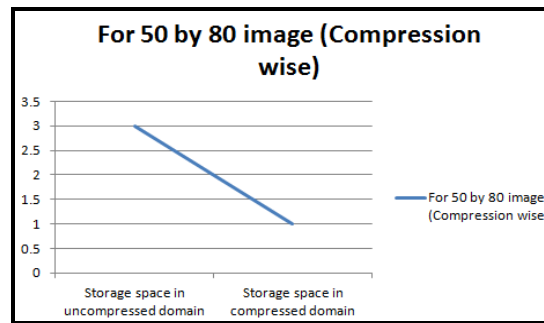


Fig.32 Storage space for different sized face images.

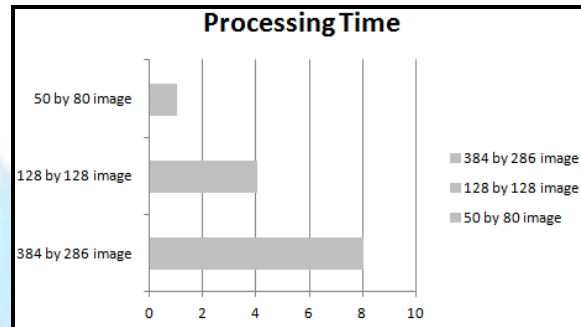


Fig.33 Processing time for different sized face images.

Conclusion

This paper proposes a novel method of face recognition using de-correlation of local features using Discrete Cosine Transforms (DCT). The DC component of histograms of local Gabor binary patterns which constitutes low frequency components which will not help in actual recognition are de-correlated using DCT technique. This differentiation and segregation of actual frequency variations adds value for accurate recognition. This method employs a non-statistical procedure which avoids training step for face samples thereby avoiding generalizability problem which is caused due to statistical learning procedure. Face image preprocessing steps are given more stress to achieve recognition accuracy. The performance modeling is carried out by varying both internal and external factors of face recognition system and improvement is shown considerably high in terms of recognition accuracy and reduction in storage space by storing train images in compressed domain. All the images are preprocessed and cropped to 50x80 such that it occupies only 1/4th space that of original 128x128 sized images and the database is stored in compressed domain such that storage space reduces 1/3rd that of uncompressed images storage space and processing time is reduced by a factor of 3.

REFERENCES

- [1] Y. Gao and M. K. H. Leung, "Face recognition using line edge map," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 6, pp. 764–779, 2002.
- [2] Haitao Wang, Stan Z. Li, Yangsheng Wang, Weiwei, Zhang, "Illumination Modeling and Normalization for Face Recognition," *Handbook of Face Recognition*, Springer, 2004.
- [3] John Wright, Gang Hua, "Implicit Elastic Matching with Random Projections for Pose-Variant Face Recognition," *IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 1502 – 1509, June. 2009.
- [4] Michal Uricar, "Detector of facial landmarks", Master's Thesis, Center for Machine Perception, Department of Cybernetics, Czech Technical University, 2011.
- [5] Jie Zou, Qiang Ji, George Nagy, "A Comparative Study of Local Matching Approach for Face Recognition," *IEEE Transactions on Image Processing*, vol. 16, no.10, pp. 2617 – 2628, 2007.
- [6] Xiaodong Liu, Guangda Su, "A Cluster-Based Parallel Face Recognition System," *Proceedings of the Intl. Conference on Parallel and Distributed Processing Techniques and Applications & Conference on Real-Time Computing Systems and Applications*, vol. 2, pp. 26-29, 2006.
- [7] Haiping Lu, "Multilinear Subspace Learning for Face and Gait Recognition," A Thesis, Department of The Edward S. Rogers Sr. Department of Electrical and Computer Engineering Repository, 2009.
- [8] V. Štruc, B. Vesnicer, and N. Pavešić, "The phase-based Gabor Fisher classifier and its application to face recognition under varying illumination conditions," *Proceedings of the 2nd International Conference on Signal Processing and Communication Systems (ICSPCS '08)*, pp. 1–6, Gold Coast, Australia, 2008.



- [9] L. Shen and L. Bai, "A review of Gabor wavelets for face recognition," *Pattern Analysis and Applications*, vol. 9, no. 2, pp. 273–292, 2006.
- [10] C. Liu and H. Wechsler, "Gabor feature based classification using the enhanced Fisher linear discriminant model for face recognition," *IEEE Transactions on Image Processing*, vol. 11, no. 4, pp. 467–476, 2002.
- [11] L. Nanni and D. Maio, "Weighted sub-Gabor for face recognition," *Pattern Recognition Letters*, vol. 28, no. 4, pp. 487–492, 2007.
- [12] Timo Ahonen, Abdenour Hadid and Matti Pietikainen, "Face Recognition with Local Binary Patterns," *ECCV*, pp. 469–481, 2004.
- [13] Wenchao Zhang, Shiguang Shan, Wen Gao, Xilin Chen, Hongming Zhang, "Local Gabor Binary Pattern Histogram Sequence (LGBPHS): A Novel Non-Statistical Model for Face Representation and Recognition," *IEEE International Conference on Computer Vision (ICCV'05)*, 2005.
- [14] Di Huang, Caifeng Shan, Mohsen Ardabilian, "Local Binary Patterns and Its Application to Facial Image Analysis: A Survey," *IEEE Transactions on systems, man, and cybernetics-Part C: Applications and Reviews*, vol. 41, no. 6, 2011.
- [15] Canny, John, "A Computational Approach to Edge Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence.*, vol. 8, no. 6, pp. 679 – 698, Nov. 1986.
- [16] Neelamma K.Patil, Suresh F. Murgod, Lokesh Boregowda, V.R.Udupi, "Adaptive Significant DCT Coefficients For Texture-Rich Grayscale Image Compression", *International Conference on Communication, VLSI and Signal Processing (ICCVSP- 2013)*,Feb. 2013.
- [17] Vasudha S, "ICA And PCA Based Face Recognition Systems-A Survey", *Proceedings of International Conference, ICERECT Series: Lecture Notes in Electrical Engineering*, Springer, vol. 248, June 2013.
- [18] Peng Wang, Member, Qiang Ji, and James L. Wayman, "Modeling and Predicting Face Recognition System Performance Based on Analysis of Similarity Scores," *IEEE Transactions on Pattern Analysis and Machine Intelligence.*,vol.29,no.4,APRIL.2007

