An Evaluation Scheme for Safe Biometric Feature Extraction Algorithms

Ashly George
Studen
Meass College,
Areakode University of Calicut
Kerala, India

Shameem Kappan
Assistant Professor
Meass College,
Areakode University of Calicut
Kerala, India

Dr. R. Vijayakumar
Professor
School of Computer Sciences,
M. G University,
Kottayam, Kerala

Abstract - **Face recognition has been a fast growing, challenging and interesting area in surveillance, image analysis, access control, commercial security and pervasive computing. Face is the primary focused parts of human body that express most of feature which plays vital role to convey identity and emotions of an individual. It is a challenging task to build an automate face recognition system that has capabilities to recognize face as human do. This paper proposes a methodology for evaluation of algorithms for feature extraction in face recognition process. The paper also covers a survey on existing methodologies for face recognition algorithms available in literature namely, Principle Component Analysis, Linear Discriminant Analysis, Kernel Principle Component Analysis and Kernel Linear Discriminant Analysis. For classification the distance classifiers KNN-classifier and Euclidean distance classifiers are employed with ORL database and Faces94 database were to evaluate various degradation in image such as variation in pose, illumination, light effect etc. From the experimental result it is found that Linear Discriminant Analysis provide a better result of 99.68% with Faces94, when number of training set per person is minimum.**

Keywords: Face Recognition, Feature Extraction, principle component analysis(PCA), Linear Discriminant Analysis(LDA), KERNEL principle component analysis (KPCA), Kernel Linear Discriminant Analysis(KDA) ,Eigen value, fisher value, KNN-k Nearest Neighbour, Euclidean Distance, Discrete Cosine Transform(DCT), Fast Fourier Transform(FFT).

INTRODUCTION

The subject, Face Recognition is an emerging area in image processing because of its importance in security, identification and verification application. Although other biometric identification (such as fingerprint or iris scan) is available, Face recognition has always remained a priority topic for researches due to the intrusive nature and robustness.

To certifying people and for providing access to physical or virtual domain password, smart card, plastic card or key are used as identity mark. Biometric based techniques are the most promising option recognizing individuals. "Biometric technologies "are automated methods of verifying or recognizing the identity of a living person based on a physiological or behavioural characteristic[24,16].different biometrics currently used for automatic identification include fingerprints, voice, iris, retina, hand, face, handwriting, keystroke, finger shape, DNA, gait, signature and palm print etc. The ideal biometric system has the characteristic: robustness, distinctiveness, availability, accessibility and acceptability [15, 19] that appraise the performance of biometric recognition system. The most efficient Biometric method,' face recognition' is referred to identifying an unknown face image by computational algorithms. Face recognition can be done by comparing unknown face with face stored in database. The three stages of face recognition are Face detection, feature extraction and facial recognition by classification. Face detection as the process of extracting faces from still or moving image. Feature Extraction involves obtaining relevant facial features from the data. Feature extraction involves transforming the input data into a set of features which can uniquely represent an image [25]. These set of features are also called feature vector. Recognition is the identification task; the system would report an identity from the database. This phase involves a comparison method using classification Algorithm and accuracy measured [10].A generic face recognition system is shown below fig. 1.
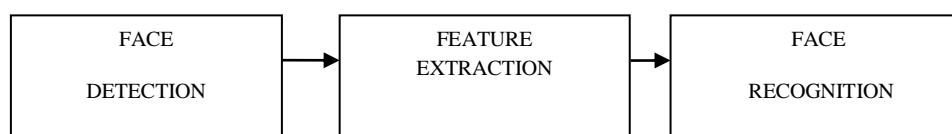

Fig1. Generic face recognition system

One of the pioneers of facial recognition, Woodrow Bledsoe in Palo Alto California, devised a technique called "man machine facial recognition" in the 1960s. Bledsoe's technique based on using computers to identify face and limits for variation in pose, illumination, and face expression also explained [10]. The most famous early face recognition system is due to Kohonen, who demonstrated that for aligned and normalized face images, a simple neural net could perform face recognition. But the system was not a practical success.

Recognition research started in the late 1970s and has become one of the most attractive and exciting research areas in computer science especially in image processing since 1990. Many face recognition algorithms have been developed during the last decade. Among that the appearance-based method, gives a promising result. Principal components analysis (PCA) and Linear Discriminant Analysis (LDA) are the two most popular linear methods in appearance-based approaches for face recognition. In 1988, Kirby and Sirovich [25] applied a standard statistic technique, principle component analysis, to the face recognition problem. This was considered somewhat of a milestone in face recognition. In 1991, based on Eigen-value Turk and Pentland [23] discovered that face in an image could be detect by the residual error, is a practical solution to face recognition system based on neural networks. Belhumeur P, Hespanha J., Kriegman D (1997) [4] make a study on Eigen faces and fisher faces to make an algorithm that is insensitive to variation in lighting and facial expression. Using Yale and Harvard database experimental result stands with fisher faces. A paper published by A.M. Martinez and A.C. Kak [3] make a comparative study of PCA and LDA .This paper describe the superiority of PCA over LDA under small number of samples per class or training data non uniformly sample the underlying distribution. For justifying their work the authors uses AR database of 3200 colour image of frontal images of faces of 126 subjects under tightly controlled condition of illumination and viewpoint .If the number of learning samples is large and representative for each class then LDA is better otherwise PCA gives better performance.

In late 1990's many kernel-based PCA and LDA methods have been developed and applied in pattern recognition tasks. In 1996, a nonlinear form of PCA, namely, kernel PCA (KPCA) that solves the Eigen-value problem is proposed by Schölkopf et al [22]. In 1999, V.Roth and V.Steinhage [17], Nonlinear Disciminant Analysis using kernel function that describes the kernel trick of representing dot products by kernel functions. To solve the problems in LDA, Baud at and Anouar [18] proposed the generalized discriminant analysis (GDA) method by extending the KFD method to multiple classes. They test the algorithm on iris, and seed and motivated from the experimental result the kernel-LDA based algorithm can solve the face recognition problem.

RELATED WORKS

The research done so far on this technique shows that it has a wide scope of research in the coming years by picking up the flaws and correcting them so that we get a better system to use this technique.

In order to represent higher order dependencies in image such as nonlinear relation, higher order statistic is needed. For that the researchers uses the method for extracting number of samples as features ,but also maximise the class separations when projected to lower dimensional spaces to recognition. Ming Hsuan Yang [14] describes kernel PCA and kernel LDA for learning low dimensional representations for face recognition. When project the features in lower dimensional space they maximizes the class separation for efficient algorithm. For that authors use AT&T and YALE DATABASE. Using AT&T the kernel PCA and LDA analysed with ICA over the variation in pose and scale. Kernel LDA provide lowest error rate, with YALE database ICA have less effect in varying light than the kernel methods. The authors concluded with suggestion that the performance can improve by using K-nearest neighbour and Preceptor and features extracted as nonlinear features.

Performance Analysis of PCA-based and LDA- based Algorithms for Face Recognition by Steven Fernandes and Josemin Bala [8] described performance of PCA and LDA algorithm with public database. The paper report performance analysis of various Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) for face recognition. Among various PCA algorithms analyzed, the best face recognition rate of 100% for Manual face localization used on ORL and SHEFFIELD database with 100 components. The next best was 99.70% face recognition rate using PCA based Immune Networks (PCA-IN) on ORL database. Illumination Adaptive Linear Discriminant Analysis (IALDA) on CMU PIE database gives the best face recognition rate of 98.9%, using Fuzzy Fisher face, through genetic algorithm on ORL database was 98.125% a next best for LDA algorithms analyzed.

The empirical evaluation of the popular appearance based feature extraction algorithms within face recognition system based on image face de-blurring by Hicham Mokhtari, Idir Belaidi and Said Alem [7],where evaluation done using the ORL databases. The image restoration performed using centralized sparse representation (CSR) and adaptive sparse domain selection with adaptive regularization (ASDS-AR) is compared by means of error equal rate and recognition rate. Then face recognition is done using PCA, LDA, KPCA and KFA and result plotted with ROC and CMC curve. The evaluation conducted using Gabor Wavelets and phase congruency where Gabor Wavelets provide better EER rate than phase congruency. Gabor linear discriminant analysis (GLDA) ensures the most consistent verification rates across the tested ORL databases for both methods CSR and ASDS-AR.

Experimental Assessment of LDA and KLDA for Face Recognition by Umesh Ashok Kamerikar [5] in 2014 analysed the performance of LDA and KLDA based feature extracted algorithm using Euclidian Distance classification .The evaluation of performance of the system's capability to distinguish between persons with different facial variations done on analyzing the facial expressions. System contains

main components – Feature Extraction (LDA and Kernel LDA algorithm) and Euclidean Distance as Classifier. To recognize the face for verification public database ORL (gray scale), Indian and Grimace (colour) database are used. Based on the experiment, as number of Train images per person increases recognition rate increases for ORL (gray scale) database, But LDA works well when less number of Train images per person. Projection of Train images for KLDA scatter within-classes and scatter between-classes are closer or farther than LDA. There for, Average recognition rate of KLDA performance is better than LDA.

FEATURE EXTRACTION ALGORITHMS

There are many approaches by which the face can be recognize. The approaches can be classified into two, the former one is geometric based and the other one is appearance based. The appearance based face recognition considered face as a raw intensity image, which are divided into two linear and nonlinear. Here, the preferred linear algorithms are PCA and LDA and nonlinear algorithms are KERNEL based PCA and LDA. Linear methods are simpler dimensionality reduction method while kernel methods are complex. Kernel learning is an important research topic in the machine learning area, and some theory and application fruits are achieved and widely applied in pattern recognition, data mining, computer vision, and image and signal processing areas. The nonlinear problems are solved at large with kernel function and system performances such as recognition accuracy, prediction accuracy largely increased. [24].

PRINCIPLE COMPONENT ANALYSIS (PCA)

PCA is the widely used mathematical approach based on the information theory concept that decomposes face image into small set of characteristic feature image called Eigen faces [10].PCA is a statistical procedure that uses an orthogonal transformation to convert a set of observation of possibly correlated variables into a set of values of uncorrelated variables called principle components [6].PCA was invented by Karl Pearson in 1901, also known as Kosambi-Karhunen-Loeve transform. Dimensionality reduction can be a very useful step for Visualising and processing high dimensional datasets.PCA provide dimensionality reduction by extracting principle components of multi-dimensional data. It is a very simple and efficient algorithm where no knowledge of geometry and reflectance of faces are required, and also data compressions are achieved by low dimensional subspace representation. The algorithm is based on Eigen faces, and its recognition rate decreases under illumination and pose variation.

This algorithm is very sensitive to scaling of variables. Lack of optimization in class separability (poor discriminating power within the class) and large computation are the well known common problems in PCA method [10]. This limitation is overcome by Linear Discriminant Analysis (LDA). LDA is the most dominant algorithms for feature selection in appearance based methods.

1. Consider the image matrix B of size (N x N) pixels is converted to the image vector T of size (P x1) where P = (N x N). Training Set:
$T' = [T'_1 \ T'_2 \ \ldots \ T'_M \ ]$. Then find the mean of training set, [7]

$$M = \frac{1}{P} \sum_{i=1}^{P} T_i$$

2. Average face image is calculated by each face differs from the average by

$A_i = T_i$ – M, where Difference Matrix: A = $[A_1 A_2 \ldots .. A_M]$

3. Using difference matrix A covariance matrix is constructed as: $C = \frac{1}{P} \sum_{i=1}^{P} A^T A$, of size (P×P). Due to its huge dimension covariance matrix is very hard to work with computational complexity. So, dimensionality of covariance matrix reduced as

$L = A^T A$, where size of L is (M×M).

4. In order to obtain the eigenvectors AX_i of the original covariance matrix with its corresponding Eigen values λ_i, it uses the following equations: $A^T AX_i = \lambda_i \ X_i$.

5. Then face image transformed into Eigen face component can be projected onto face space by

$$\Omega_k = U_k^T A_i$$

Where k=1, 2,..., M', M' Is the number of Eigen-faces used for the recognition.

To identify the best description of unknown image, the simplest method nearest neighbour method like Euclidean distance is used. To find the image k that minimizes the Euclidean distance $€_k$ [1].

To find the image k that minimizes the Euclidean distance $€_k$

$$€_k = \| (\Omega - \Omega_k) \|^2$$

LINEAR DISCRIMINANT ANALYSIS (LDA)

LDA, also known as Fishers Discriminant Analysis Or Multiple Discriminant Analysis is basically a statistical technique used in image recognition and classification.LDA is well known for feature extraction and dimensionality reduction based on fisher faces LDA project the data onto a lower-dimension vector space such that ratio of the between class distance to within class distance is maximise and therefore achieve maximum discrimination [3, 9].In Fisher face method Project the new image set to a lower-dimensional space defined by feature faces.LDA aims to find the optimal transformation of input space so that the it preserve maximum between class variance and minimum within class variance. Following steps describe Linear Discriminant Analysis:

Given data matrix A$\in R^{NXn}$ each column a_i of A maps to vector b_i in

1- Dimensional space. Partition Matrix A into k classes,

A = $\{C_1, C_2, \ldots, C_n\}$, C_i contain n_i data points from i^{th} class [4].

Calculate the mean of each class (μ_i) and mean image of all classes (µ).

$\mu_i = \frac{1}{C_p}\sum_{j=1}^{C_p} x_i$ where i =1, 2,....,, C_n and j=1, 2,..., C_p

$\mu = \frac{1}{C_n}\sum_{j=1}^{C_n} \mu_j$ where j=1, 2,..., C_n

Calculate within scatter matrix (S_w) & Between-class scatter matrix (S_p).

$S_w = \sum_{i=1}^{C_n}\sum_{j=1}^{C_p}(x_{j-\mu_i})((x_{j-\mu_i})^T$

$S_p = \sum_{i=1}^{C_n}(\mu_i-\mu)(\mu_i-\mu)^T$

Calculate Eigen vector of J (J_eig) from within scatter matrix & Between-class scatter matrix, here Eigen vector is known as fisher vector corresponding Eigen value is called fisher value.

Sorting of eigenvector of J

We have to sort out Fisher vector depending upon their corresponding Fisher values and neglect fisher vector corresponding to small fisher values.

Sorted eigenvectors of J = V_Fisher

Project data in fisher space

Ω_p is converted into Ω_F by projecting onto a Fisher subspace, so that images of same class or person move closer together & images of different classes move further apart.

$\Omega_F = (V\_Fisher^T X \Omega_p)$

For testing Acquire the Test face images (the training set). Suppose T be the test image of size (mxn), convert it from two dimensions (mxn) to one dimension (mxn) x1) and find deviation of Test image from mean image.

X= T – M (mnxp)

Transfer Test data to face space

$\Omega_{NEW} = V\_Fisher^T X\ V\_PCA^T X$ X

Calculate minimum Euclidean distance

$= \quad \in_i = \min(\| \Omega_{NEW} - \Omega_i\|)$

Find the index number of minimum Euclidian distance which is the best match for the Test image.

KERNEL PRINCIPLE COMPONENT ANALYSIS (KPCA)

KPCA is the modification of linear PCA to fulfil the high-dimensional gap that is constructed using a kernel function. In KPCA through the kernel trick, input data are mapped onto higher dimensional feature space [14].The KPCA method has been widely used for non-linear feature extraction and data projection. Kernel PCA have the advantage of can give a good re- encoding of the data when it lies along a non-linear manifold. The kernel PCA will have difficulties if we have lots of data points because of n x n kernel matrix.

In nonlinear method there is a transformation from D-dimensional feature   space to F-dimensional feature space , where F≥D. then each data point x_i is projected to a point φ(x_i). By using a kernel function, every linear algorithm that can be implicitly executed in F and constructing a nonlinear version of a linear algorithm by mapping from Φ: $R^N \rightarrow$ F. Kernel function that used in the project is the Radial-basis function: k $(X_i, X_j) = (x^T y + 1)^p$

Construct the kernel matrix from the training data set $\{X_j\}$ using the kernel function k (X_i, X_j).[22].

Given a set of centred data X^k , $k = 1,2, \ldots$ M:

$$\sum_{i=1}^{M} \Phi(X_i) = 0 \qquad (1)$$

The covariance matrix in F [21]:

$$C = \frac{1}{M}\sum_{j=1}^{M} \Phi(x_i)\, \Phi(x_j)^T \qquad (2)$$

We have to find the Eigen values ($\lambda_1 \le \lambda_2 \le \cdots \le \lambda_M$) $\lambda \ge 0$ and Eigen vectors ($\alpha^1, \alpha^2, \ldots, \alpha^M$)satisfying the condition

$$\lambda v = Cv. \quad (3)$$

All solutions v with $\lambda \ne 0$ must lie in the span of $\Phi(X_1)$, $\Phi(X_2)$, …, $\Phi(X_M)$. Hence equation become $\lambda (\Phi(X_k).v) = (\Phi(X_k)) . (v)$ For k=1, 2,..., M. (4)

Since v lie in the span of $\Phi(X_1)$, $\Phi(X_2)$,…, $\Phi(X_M)$. There exist coefficients α_i (i = 1, 2… M) Such that,

$$v = \sum_{j=1}^{M} \alpha_i \Phi(X_i) \qquad (5)$$

Combining Eq.(4)and Eq.(5), we get

$\lambda \sum_{i=1}^{M} \alpha_i \Phi(X_k)\, \Phi(X_i)$

$= \frac{1}{M}\sum_{i=1}^{M}(\alpha_i \Phi(X_k) \sum_{J=1}^{M} \Phi(X_k)) (\Phi(X_i)\, \Phi(X_j))$ (6)

By defining the kernel function as k (X_i, X_j), M x M kernel matrix K is given by,

$$k_{ij} = (\Phi(X_i) . \Phi(X_j)) \qquad (7)$$

We get   M$\lambda \alpha = k^2 \alpha$, where $\alpha = \{\alpha_1, \alpha_2, \ldots, \alpha_M\}$ is the column vector.

We solve the Eigen value problem as M$\lambda \alpha$ = kα.

$$(8)$$

Let λ_i be the Eigen value and α_i be the eigen vector in F have to be normalised.

The eq. (5) and (8) translate into

$1 = \frac{1}{M}\sum_{i,j=1}^{M}(\alpha_i^k \alpha_j^k \Phi(X_i)\, \Phi(X_j)$ $= \alpha_i^k \alpha_j^k k_{ij}$

$= \alpha^k . k\alpha^k = \lambda_k (\alpha^k . \alpha^k) \qquad (9)$

For the principal component extraction, the projections onto the eigenvectors v^k in F are needed. Given x as a test point and $\Phi(x)$ is its image in F, then its nonlinear principal components is,

$$(v^k . \Phi(X)) = \sum_{i=1}^{M} \alpha_i^k (\Phi(X_j)\Phi(X)). \qquad (10)$$

KERNEL LINEAR DISCRIMINANT ANALYSIS (KDA)

Nonlinear linear discriminant analysis in which there is a transformation from feature space to higher dimensional feature space F by Φ: $R^N \rightarrow$ F, X$\rightarrow$ Φ (x). In order to solve nonlinearity problem, the kernel LDA used. Here the between class scatter matrix and within scatter matrix are defined in kernel feature space F. Let λ be the Eigen value and W^Φ be the Eigen vector, then

transformation matrix W that maximize the objective function [17].

$$J(W) = \frac{W^T S_B^\Phi W}{W S_W^\Phi W}$$

The column of W is the generalised Eigen vector and corresponding Eigen values satisfy

$$S_B^\Phi W_i = \lambda_i S_w^\Phi W_i$$

Where, Between-class scatter matrix:

$$S_B^\Phi = \sum_{i=1}^c \sum_{j=1}^{M_i} \Phi(X_j^i - \mu_i^\Phi)(\Phi(X_j^i - \mu_i^\Phi)^T$$

And the within class scatter matrix

$$S_W^\Phi = \sum_{i=1}^c M(\mu_i^\Phi - \mu^\Phi)(\mu_i^\Phi - \mu^\Phi)^T$$

Mean image of class $X_i : \mu_i = \frac{1}{M}\sum_{J=1}^M X_j^i$

and $W^\Phi = \sum_{i=1}^c \sum_{j=1}^M \alpha_j^i \Phi(X_j^i)$

Optimal projection $W_{OPT} = argmax W^\Phi \frac{(W^\Phi)^T S_B^\Phi W^\Phi}{W)^T S_w^\Phi W^\Phi} = [w_1^\Phi, w_{2,}^\Phi ..... w_c^\Phi]$

That maximise the ratio of between class scatter to within class scatter matrix.

Finally Φ (x) projected to lower dimensional space that corresponding to Eigen vector W^Φ.

PROPOSED WORK

Firstly, the literature work on the various feature extraction techniques would be studied in detail, then reviewed the flow and refined in case any changes are required. Afterwards, the algorithm generated would be programmed in MATLAB to compare and analysis the result with existing works. Face recognition techniques that must have higher accuracy and low error rate (such as false error rate and false accept rate) in variation with pose, illumination and expression. To achieve this objective, our research will keep focused towards the implementation of best feature extraction algorithm for face recognition. Here, the developing system consisting of two phases which are the feature extraction phase where linear and non linear PCA and LDA used. The classification step chooses to be the simplest classifier; k-nearest neighbour with Euclidean distance. The performance of proposed algorithms are analysed with various public database such as AT&T and faces95.the major steps involve in the proposed method is given below fig.1
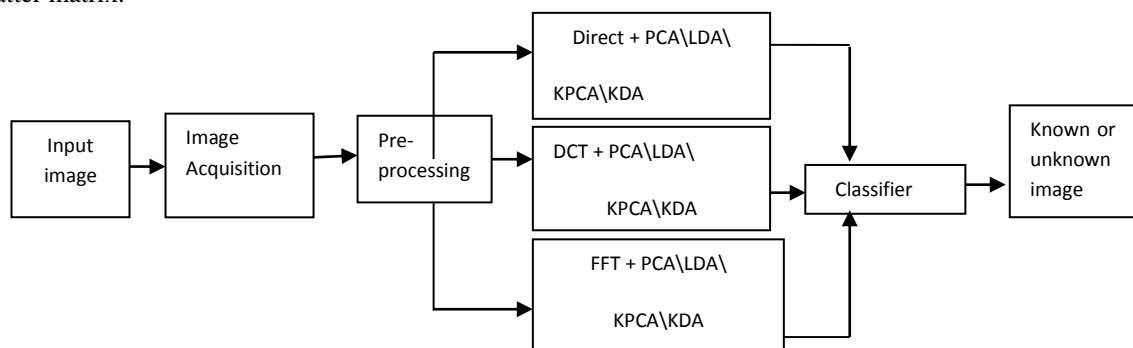

Fig.3.1 proposed face recognition system

1 .Image Acquisition

It's the entry point of the face recognition process. the module where the face image under consideration is presented to the system is the image acquisition module .here, the face image presented to the system by using standard dataset , namely, AT&T face dataset and faces94 dataset.

2 .Pre-processing

In this module, face images are normalized and if desired, they are enhanced to improve the recognition performance of the system. In Pre –processing module all images are converted to gray-scale by eliminating the hue and saturation information while retaining the luminance. For uniformity and faster execution we have resized image to 50x50 pixels resolution. Finally, for better processing, we converted 2d image matrix into 1d vector.

3. Feature extraction

After performing some pre-processing, the normalized face image is presented to the feature extraction module in order to find the key features that are going to be used for classification .In this proposed methodology the four feature extraction techniques, linear and nonlinear PCA and LDA are used .for extracting important features from face, the methods are tested by apply directly on pixel image intensities, apply on DCT Coefficients of image And apply on FFT Coefficients of image. One of the image processing tool, Discrete Cosine Transform (DCT) that concentrate most visually significant information about the image on few coefficient of DCT. DCT are independent of set of image are not applied on entire image, so it is most useful in lossy image compression like JPEG. Another processing tool, Fast Fourier Transform that computes the Discrete Fourier Transform and its inverse .Also, provides access to geometric characteristics of a spatial domain image.

Discrete Cosine Transform (DCT)

Well known compression standards DCT is a transform used to compress the representation of data by discarding redundant information. DCT convert image from spatial

domain to frequency domain. Most visually significant information about the image is concentrate on just few coefficients of DCT. SO, most useful lossy compression likes JPEG. Due to its compact representation power, data independent nature DCT is an important image processing tool [24]. It is an invertible linear transform that expresses a finite sequence of data points in terms of a sum of cosine functions oscillating at different frequencies. These properties make DCT to seek for face recognition [12] .General equation for 2D DCT is defined as

$$F(u,v) = \left(\frac{2}{N}\right)^{\wedge}\frac{1}{2}\left(\frac{2}{M}\right)^{\frac{1}{2}}\sum_{i=0}^{N-1}\sum_{j=0}^{M-1}\wedge(i).\wedge(j).\cos\left[\frac{\pi.u}{2.N}(2i+1)\right]\cos\left[\frac{\pi.u}{2.N}(2j+1)\right].f(i,j)$$

And the corresponding inverse 2D DCT transform is, $F^{-I}(u,v)$

Fast Fourier Transform (FFT)

The Fourier Transform is an important image processing tool. This Transformation is used if we want to access the geometric characteristics of a spatial domain image. The Fast Fourier Transform (FFT) is an efficient and fast algorithm to compute the discrete Fourier transforms (DFT) and its inverse [20].
The DFT is given in the following equation:

$$F(x, y) = \sum_{i=0}^{M-1}\sum_{j=0}^{N-1}Y(i,j)\,e^{-i2\pi(\frac{xi}{M}+\frac{yi}{N})}$$

Where Y (i, j) is the image in the spatial domain and the each point F(x, y) in the Fourier space. The following the steps are followed to calculate FFT of images:

- Apply FFT to image according to the equation above .In most implementations the Fourier image is shifted that the value
- (I.e. the image mean) F (0, 0) is displayed in the centre of the image.
- Use the abs and then log functions: abs (log (FFT)) to compute the magnitude of the combined components.
- We know that the second half of FFT carry no useful and duplicated information, so we can half the data to treat.

4. Classification
 In this module, with the help of a pattern classifier, extracted features of the face image is compared with the ones stored in a face library (or face database). After doing this comparison, face image is classified as either known or unknown. For classification, the well known classifiers, K-NN classifier and Euclidean Distance classifiers are used. KNN is a supervised learning method for classifying objects by finding the closest K neighbors in the feature space.

Euclidean Distance Classifiers
Euclidean Distance classifiers or nearest neighbor classifier is a non-parametric density estimation technique that find the distance of given feature vector x from all the training samples and find the closest sample in the training set. Here the value of k=1, it is very powerful when we have a large number of samples in our training set.

The Euclidean distance,

$$dist = \sqrt{\sum_{k=1}^{n}(p_k - q_k)^2}$$

Where n is the number of dimensions (attributes) and p_k and q_k are, respectively, the kth attributes (components) or data objects p and q.

K-NN Classifier
The k-nearest neighbour classifier is a very simple and intuitive method are classified based on their similarity with training data .For a given unlabeled samples $x_u \in R^D$, find the k "closest" labelled, (where k>1) samples in the training data set and assign x_u to the class that appears most frequently within the k- subset .The classifier only requires An integer k, A set of labelled samples and A measure of "closeness. The k instances are defined by calculating a certain distance such as: Euclidian distance, City block distance, etc.The benefits of classifier are they are analytically tractable with simple implementation, Uses local information, which can yield highly adaptive behaviour and also Lends itself very easily to parallel implementations .However, values of k are too large that become detrimental. It destroys the locality of the estimation in addition to increase in the computational burden [30].
The algorithm can be summarized as [11]:
- A positive integer k is specified, along with a new sample
- We select the k entries in our database which are closest to the new sample
- We find the most common classification of these entries
- This is the classification we give to the new sample.

EXPERIMENTAL RESULT
 Analysis of different feature extraction methods have implemented to find out performance of algorithms in terms of accuracy. Using this system choice of the algorithms was made one by one and then performance accuracy was generated for all the algorithms separately. For extracting important features from face, each method are tested by apply directly on pixel image intensities, apply on DCT Coefficients of image and apply on FFT Coefficients of image. The proposed system was evaluated using ORL database and faces94 database.

AT&T Dataset
 The AT&T Face database, sometimes also known as ORL Database of Faces, contains ten different images of each of 40 distinct subjects. For some subjects, the images were taken at different times, varying the lighting, facial expressions (open / closed eyes, smiling / not smiling) and facial details (glasses / no glasses). All the images were taken against a dark homogeneous background with the subjects in an upright, frontal

position (with tolerance for some side movement). The AT&T Face database is good for initial tests in linear method, but it's a fairly easy database [31][29].

Result of the performed experiment is shown in tables 1 in the case of selection features by applying of directly

on the images training data. Here the number of training set n<8 and the kernel function used here is the quadratic kernel φ=$K^2 + 1$.

Table 1: Accuracies of Direct Apply

DIRECT METHOD	EUCLIDEAN DISTANCE	KNN CLASSIFIER
PCA	93.75%	87.5%
LDA	97.5%	93.75%
KPCA	63.75%	50%
KLDA	88.75%	81.25%

Table 2 and table 3 describe the performance of PCA, LDA, KPCA and KDA by applying DCT and FFT on training data.

Table.2: Accuracies of DCT

DISCRETE COSINE TRANSFORM	EUCLIDEAN DISTANCE	KNN CLASSIFIER
PCA	91.25%	82.5%
LDA	96.25%	96.25%
KPCA	67.5%	53.7%
KLDA	83.75%	75%

The results of the study shown that the recognition rate of image is similar for the application of DCT and direct method. Euclidean distance as classifier gives a better result than KNN classifier.

Table 3: Accuracies of FFT

FAST FOURIER TRANSFORM	EUCLIDEAN DISTANCE	KNN CLASSIFIER
PCA	28.75%	27.5%
LDA	67.5%	60%
KPCA	20%	17.5%
KLDA	50%	46.25%

The ORL dataset is the simplest database for initial analysis by the application of FFT to the image that degrades the accuracy rate. Small changes are appeared in image due to the effect of noise or variation. Fourier Transform does not perform well so chance of ignoring a large amount of coefficient. From the above table it is given that the performance of LDA is better than the PCA and other nonlinear methods. Due to the fact that, number of training set and class separability make up LDA better than other methods. By applying DCT to the image there is no large variation in the recognition accuracy. But for the application of FFT to the image there is a large variation in the accuracy except for LDA.

Faces94 Dataset

Faces94 consists of face images 80 distinct subjects with 20 images per subject are taken in plain green background with mirror variation in head turn, tilt and slant. All images are in RGB format with the size of 180x200 pixels [31]. The performance of PCA\LDA\KPCA\KDA are described in below tables, where the number of training set per person is 8 and the kernel function used for analysis is the quadratic kernel function.

Table 4 accuracies of direct method

DIRECT	EUCLIDEAN DISTANCE	KNN CLASSIFIER
PCA	95%	95%
LDA	99.6875%	99.6875%
KPCA	83.5%	78%
KLDA	95.5208%	95.52085

FACES94 dataset is a large colour dataset in which linear method provide a better performance than non linear methods. The technique gives an accurate result for all methods with the accuracy greater than 85%.Like ORL database if the number of training set per dataset

and if we use polynomial kernel as kernel function the accuracy might be lower than linear method. Luminance is important for distinguishing picture features that is important in colour image. For both dataset the accuracy is better for LDA feature extraction algorithm.

Table 5: Accuracies of DCT

DISCRETE COSINE TRANSFORM	EUCLIDEAN DISTANCE	KNN CLASSIFIER
PCA	95%	95%
LDA	99.583%	99.479%
KPCA	81.3%	78.1%
KLDA	95.5208%	95.583%

Table 6: Accuracies of FFT

FAST FOURIER TRANSFORM	EUCLIDEAN DISTANCE	KNN CLASSIFIER
PCA	58.67%	57.3%
LDA	97.91%	97.812%
KPCA	53.4%	53.4%
KLDA	86.563%	85.9375%

From above table 6 it is clear that the application of Fast Fourier Transform to the image that greatly affect the analysis of image processing applications. From the above tables it is clear that performance of feature extraction algorithm greatly defences on the train set per person and the kernel function. If the number of training set per person is less, then it's better for linear methods especially for LDA. In my experiment I proposed the kernel function as quadratic kernel function. Kernel function that map image from lower dimensional space to higher dimensional space. The space become higher, it is better for nonlinear method. But here the space is in quadratic dimension, so better for linear method.

FUTURE WORK

Face recognition has a wide range of application in the field of image processing and pattern recognition. For a good recognition system it is necessary to have an excellent feature extraction algorithm and a classifier. In future, to improve the comparison work, I have decided to choose the kernel function as Gaussian function that improve the nonlinear extraction algorithm. Also perform comparison with the help of database like YALE and UMIST dataset. For classification the different classifier i.e., neural network classifiers also can yield to improve the experiment.

CONCLUTION

The paper presented a retrospective evaluation of the popular appearance based linear and nonlinear feature extraction algorithms within face recognition system based on image processing. This paper have analysed the works of researchers on many approaches of feature extraction based on various parameter. Year by year many different modified techniques and algorithms are being implemented are not hundred percent accurate. Still there are many chances of improvement in various algorithms to reach the ideal human face recognition system. From this experiment it is reveal that if the number of training set per person increases then the recognition accuracy is increases for nonlinear methods. If the number of training set per person is less the LDA provide a better result for face recognition system. So LDA can be suitable for those applications (like children transportation system or attendance monitoring system) where number of samples is least. The polynomial kernel is a parametric model where the size is fixed and giving more and more data won't help for representing features of an image. Hence it provides a better result to linear feature extraction techniques. Also the application of data compression method to image samples is irrelevant for face recognition system. I apologize to those researchers whose important contributions may have been inspected.

REFERENCE

[1] J. Prabin Jose P., P. Poornima Kukkapulli, Manoj Kumar et al (2013), A Novel method for colour Face recognition using KNN Classifier, IEEE.

[2] Jian Yang, Alejandro F. Frangi, Jing-yu Yang, David Zhang (February 2005) , Senior Member, IEEE, and Zhong Jin, KPCA Plus LDA: A Complete Kernel Fisher Discriminant Framework for Feature Extraction and Recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 27, No. 2.

[3] Aleix m. Martinez and Avinash c.kak, PCA versus LDA, 2001, IEEE transactions on pattern analysis and machine intelligence, 23(2), 228-233.

[4] Belhumeur, P., Hespanha J, Kriegman D. (1997), Eigen faces vs. Fisher faces: recognition using class specific linear projection. IEEE Transactions on Pattern Analysis and Intelligence, 19(7), 711–720.

[5] Umesh Ashok Kamerikar Dr. M.S. Chavan, February 2014 , Experimental Assessment of LDA and KLDA for Face Recognition , IJARCSMS,Volume 2, Issue 2, ISSN: 2321-778.

[6] Priyanka Dhoke, M.P. Parsai , (August 2014)A MATLAB based Face Recognition using PCA with Back Propagation Neural network , International Journal of Innovative Research in Computer and Communication Engineering (An ISO 3297: 2007 Certified Organization) Vol. 2, Issue 8,

[7] Hicham Mokhtari, Idir Belaidi and Said Alem, 2013, Performance Comparison of Face Recognition Algorithms based on face image Retrieval, Research Journal of Recent Sciences, ISSN 2277-2502 Vol. 2(12), 65-73.

[8] Steven Fernandes and Josemin Bala (2013), Performance Analysis of PCA-based and LDA- based Algorithms for Face Recognition , International Journal of Signal Processing Systems Vol. 1, No. 1 ,10.12720/ijsps.1.1.1-6.

[9] Sanjeev Kumar and Harpreet Kaur (2012), face recognition techniques: classification and comparisons, International Journal of Information Technology and Knowledge Management, Volume 5(2), pp. 361-363.

[10] Proyecto Fin de Carrera, June 16, 2010, Face Recognition Algorithms, Ion Marqu´es.

[11] L.Kozma [2008].'K Nearest Neighbours algorithm (KNN)', Helsinki University of Technology, T-61.6020 Special Course in Computer and Information Science.

[12] E. F. Glynn, [2007], 'Fourier Analysis and Image Processing', scientific programmer, Bioinformatics, S Towers Institute for medical Research.

[13] Hui Kong, Lei Wang, Eam Khwang Teoh, Xuchun Li, Jian Gang Wang, Ronda Venkateswarlu (2005) Generalised 2D principle component analysis for face image representation and recognition, Neural Networks,585-594.

[14] Ming hsuan yang (2002), kernel Eigen face vs. kernel Fisher face- face recognition using kernel methods, FG, 215-220.

[15] J. Wayman (2001), Fundamentals of biometric authentication technologies. Int. J. Imaging and Graphics, 1(1).

[16] J. Wayman (2000), A definition of biometrics National Biometric Test Centre Collected Works 1997–2000, San Jose State University.

[17] V.Roth and V.Steinhage (2000), nonlinear disciminant analysis using kernel functions in NIPS 12,568-574.

[18] Baudat, G., Anouar, F. (2000). Generalized Discriminant Analysis using a Kernel Approach. Neural Computation 12(10), 2385–2404.

[19] J. L. Wayman (1999), Technical testing and evaluation of biometric identification devices, in A. Jain, et al. (Eds) Biometrics: Personal Identification in Networked Society¸ Kluwer Academic Press.

[20] Haykin S., Neural Networks: A comprehensive foundation, Prentice Hall, 0-13-273350-1, New Jersey, [1999].

[21] V. Vapnik (1998), Statistical learning theory, Springer, Berlin Heidelberg, New York.

[22] Schölkopf, B., Smola, A., Müller, K.-R(1998), Nonlinear Component Analysis as A Kernel Eigen value Problem, Neural Computation, 10(5), 1299–1319 .

[23] M.Turk and A. Pentland (1991) "Eigen faces for recognition," J. Cognitive Neuroscience, vol. 3, 71-86.

[24] B. Miller (1988), everything you need to know about biometric identification. Personal Identification News Biometric Industry Directory, Warfel & Miller, Inc., Washington DC.

[25] L.Sirvoich and M.Kirby (1987) A low dimensional Procedure for Characterization of Human Faces, J.Optical SOC. Am. A, Vol. 4, No. 3, 519-524.

[26] Jun-Bao Li , Shu-Chuan Chu, Jeng-Shyang Pan ,Kernel Learning Algorithms for Face Recognition, , ISBN 978-1-4614-0160-5 ISBN 978-1-4614-0161-2 (eBook) DOI 10.1007/978-1-4614-0161-2 Springer New York Heidelberg Dordrecht London.

[27] Yuan Wang Yunde Jia, P.R.CHINA ,Changbo Hu ,Matthew Turk FACE RECOGNITION BASED ON KERNEL RADIAL BASIS FUNCTION NETWORKS Computer Science Department Beijing Institute of Technology Beijing 100081, Computer Science Department University of California Santa Barbara, CA 93106, USA

[28] Derzu Omaia, JanKees v. d. Poel, Leonardo V. Batista, 2D-DCT Distance Based Face Recognition Using a Reduced Number of Coefficients.

[29] Philipp Wagner, Face Recognition with GNU Octave/MATLAB

[30] Ricardo Gutierrez, Lecture 12: classification [ppt] , Intelligent Sensor Systems, Osuna Wright State University

[31] http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html