Face Recognition: A Literature Review

Nawaf Hazim Barnouti
Al-Mansour University College
Baghdad, Iraq

Sinan Sameer Mahmood Al-Dabbagh
Al-Mansour University College
Baghdad, Iraq

Wael Esam Matti
Al-Mansour University College
Baghdad, Iraq

ABSTRACT
Face recognition have gained a great deal of popularity because of the wide range of applications such as in entertainment, smart cards, information security, law enforcement, and surveillance. It is a relevant subject in pattern recognition, computer vision, and image processing. Two major methods are used for features extraction, which can be classified into appearance-based and Model-based methods. Appearance-based methods use global representations to identify a face. Model-based face methods aim to construct a model of the human face that capture facial variations. Image similarity is the distance between the vectors of two images. This paper contains Four sections. The first section discusses face recognition applications with examples. The second section discuss the common feature face recognition methods. The third section discuss distance measurement classifiers. The fourth section discuss different face recognition databases.

General Terms
Face Recognition, Face Recognition Methods, Distance Measurement, Face Recognition Database

Keywords
PCA, LDA, ICA, KPCA, KLDA, EBGM

1. INTRODUCTION
Face recognition is one of the most important applications of biometrics based authentication system in the last few decades. Face recognition is kind of recognition task pattern, where a face is categorized as either known or unknown after comparing it with the images of a known person stored in the database. Face recognition is a challenge, given the certain variability in information because of random variation across different people, including systematic variations from various factors such as lightening conditions and pose [1].

Computational methods of face recognition need to address numerous challenges. These type of difficulties appear because faces need to be represented in such a way that best utilizes the available face information to define a specific face from all the other faces in the database. Face pose is a specifically difficult problem in this aspect simply because all faces seem similar; specifically, all faces consist of two eyes, mouth, nose, and other features that are in the same location [1] [2].

The human face is an extremely complex and dynamic structure with characteristics that can significantly and quickly change in time. Face recognition involves a range of activities from various aspects of human life. Humans can recognize faces, but too many faces sometimes being hard to memorized, machine learning is now being improved to do this task. Scientists attempt to understand the architecture of the human face when building or developing face recognition systems.

The interest and focus on the methodology of human face recognition system can help researchers to understand the basic system. Human face recognition system utilizes some data obtained from a few or all of the senses, such as visual, auditory, and tactile. Each of these data are used either individually or together to memorize and store faces. In many cases, conditions around the person are also important in a human face recognition system. Handling sizable data and combining them are difficult for a machine recognition system. However, memorizing many faces is also difficult. Key advantage of a machine system is the memory capacity. Human features that may be used for face recognition are continuously being studied, and arguably. Both local and global features are needed for face recognition [3] [4].

The research on machine face recognition has developed independently from studies on human face recognition. During the 1970s typical pattern classification methods, which use measurements between facial features or face profiles are used. During 1980s, works on face recognition is almost stable. Since the early 1990s the research focus on machine face recognition has grown significantly [5]. The significant growth is because of the availability of real-time hardware, the growing need for surveillance applications, an increasing emphasis on commercial civilian research projects, and the studies on natural network classifiers, which stress real-time computation and adaptation. Face recognition system falls under two classifications: verification and identification. Face verification (one-to-one matching) that compares the face image against a template face images whose identity is being claimed. Face identification (one-to-many matching) that compares a query face image against all image templates in a face database. Several methods are used for facial features extraction, which can be broadly classified into appearance-based (Holistic) and model-based methods. The hybrid method is a combination of these two methods.

Regardless of the method used, the most important concern in face recognition is dimensionality. Suitable methods are needed to reduce the dimension of the studied space. Working on higher dimension cases overfitting, where the system starts to memorize. Computational complexity is also an important problem when working on large databases.

2. FACE RECOGNITION APPLICATIONS
Face recognition has received considerable attention because of its numerous applications, such as in entertainment, smart cards, information security, law enforcement, and surveillance [20]. Table 1 shows the most important face recognition applications.

Table 1. Face Recognition Applications

Areas	Applications
Security	Building access control, flight boarding system, email authentication on multimedia workstation, office access.
Criminal Justice System	Forensics and post-event analysis.
Image Database Investigation	National ID, welfare registration, searching image database of licensed drivers, benefit recipient.
Surveillance	Monitoring and searching for drug offenders, CCTV control, power grid surveillance, portal control.
Smart Card Applications	Face prints can be stored in smart card, bar code, or magnetic strip and authenticated by matching the live image with stored template.
Video Indexing	Labeling faces in video
Civilian Applications	e-booking and e-commerce
Human Computer Interactions	Interactive gaming and proactive computing.
Multimedia Environment with Adaptive Human Computer Interface	Part of context aware or ubiquitous systems, recognizing a customer and assessing the customer needs.

3. A GENERIC FACE RECOGNITION SYSTEM

The input of face recognition system is normally a digital image or video stream and the output is an identification or verification of the person that appear in the image or video stream [41]. Several approaches define a face recognition system as a three step process as shown in Figure 1.
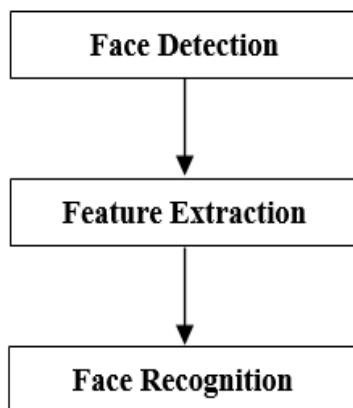

Fig 1: Face Recognition Process

4. FACE RECOGNITION METHODS

Two major methods are used for facial features extraction, which can be classified into appearance-based and Model-based methods. Figure 2 shows a summary of Holistic and Model-based methods.

4.1 Appearance-Based (Holistic) Methods

Apperance-Based methods attempt to identify faces using global representations that based on the entire image rather than local facial features. Many methods for object recognition and computer graphics are based directly on images without intermediate 3 dimensional models, most of these methods depends on image representation that induce a vector space structure and requires dense correspondance in principle [6] [7].

Global facial information is fundamentally represented by a small number of features that are directly derived from the pixel information of face images, these small number of features distinctly capture the variance among different individual faces and are used to identify unique individuals. In apperance-based method, the whole face region is considered as an input for face detection system to perform face recognition [8] [9]. Apperance-based methods, can be classified into linear and non-linear subspaces.

- **Linear Analysis**
  Three classic linear classifiers, PCA, LDA, and ICA with several other methods are explained in this paper, each classifier has its own representation of high dimensional face vector space based on different statistical viewpoints. The projection coefficients are used as the feature representation of each face image through the projection of the face vector onto the basis vectors. The matching score between the test face image and the training prototype is calculated between the coefficient vectors of the images, smaller matching score leads to a better matching process [7].

- **Non-Linear Analysis**
  The face manifold is more complicated than linear models. Linear subspace analysis is an approximation of this non-linear manifold. Direct non-linear manifold modeling schemes are explored to learn this non-linear manifold. One of the methods used is Kernel Principal Component Analysis (KPCA) [7].

4.2 Model-Based Methods

Model-based face recognition methods aims to construct a model of the human face that capture facial variations. Prior knowledge of the human face is highly utilized to design the model. For example, model-based matching derives the distance and relative position features from the placement of internal facial elements. Model-based methods can be made invariant to size, orientation, and lighting. The other benefits of these schemes are the compactness of the representation of face images and rapid matching [10] [6].

Three different extraction methods are distinguished (generic methods based on edges, lines, and curves; feature template-based methods, and structural matching methods that consider geometrical feature constraints) [9].

The major disadvantage of these methods is the difficulty of automatic feature detection. Implementation any of these methods needs arbitary decisions on which features are important. Probably if the features set lacks discrimination

ability, no amount of subsequent processing can compensate for the intrinsic deficiencey [6].
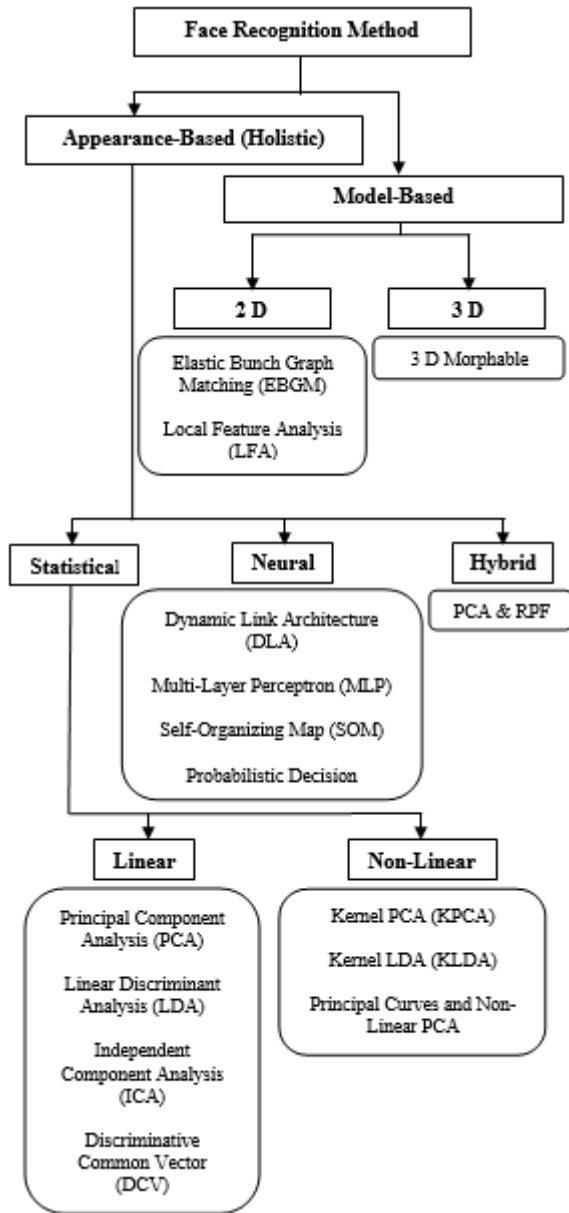

Fig 2: Holistic and Model-Based Methods

4.2.1 Principal Component Analysis (PCA)

PCA is feature extraction and dimension reduction method, which can be used to solve recognition and compression problems. PCA is a popular linear projection method, and is also known as Eigen space projection, Karhunen and Loeve (KL) transformation, or Hotelling [11]. PCA will reduces dimensionality by extracting the principal components (PCs) of multidimensional data [12] [39]. PCA can extract the important features, capture the nearly variable data components of samples, and then select numerous significant individuals from all the feature components [11]. Through PCA, an efficient and simple recognition process can be obtained compared with other methods. Recognitions that use PCA features also perform better in singular variation cases for each individual [13]. Raw intensity data are used for recognition and learning without mid-level or low-level processing. PCA is

unsupervised and does not require any knowledge about face image. Data compression can also be computed using the subspace of the low dimensional representation [20].

PCA was employed to extract the features of face images, and a sparse representation-based classification method is used for face recognition. PCA is suitable when measures on several observed variables exist, and when aiming to develop a smaller number of unknown variables that account for almost all of the variances in recognition variables. PCA steps are listed below [38] [39]:

a. Training set of total M images are used to compute the Average Mean as shown in the equation below:

$$AverageMean = \frac{1}{M} \sum_{n=1}^{M} TrainingImages\ (n) \qquad (1)$$

b. Subtract the original image from Average Mean as shown in the equation below:

$$S = TrainingImages - AverageMean \qquad (2)$$

c. Compute the Covariance Matrix as shown in the equation below:

$$Covariance = \sum_{n=1}^{M} S(n)\ S^T(n) \qquad (3)$$

d. Calculate Eigenvalues and Eigenvectors of the Covariance Matrix.

e. Sort and eliminate Eigenvalues.

f. Project the training samples onto the Eigenfaces.

(Nicholl and Amira, 2008) [24] study focus on automatically estimating the discriminative coefficients in a discrete wavelet transform (DWT)/PCA that deals with inter-class and intra class standard deviations. Consequently, Eigenfaces are selected based on the eigenvalues with discrepancies caused by the illumination factors among trained images.

(Kshirsagar et al., 2011) [25] proposed a scheme based on information theory approach, which decomposes face images into a small set of characteristic feature images called eigenfaces. These feature images are actually the PCs of the initial training set of face images. Recognition is performed by projecting a new image onto the subspace spanned by the eigenfaces (face space), and then classifying the faces by comparing the position of the faces in the face space with the positions of the known individuals. The eigenfaces approach provides an efficient way of determining this lower-dimensional space. Eigenfaces are eigenvectors that represent each of the dimensions of the face space, and they can be considered as various facial features. Any face can be expressed as linear combinations of the singular vectors of the set of faces, and these singular vectors are the eigenvectors of the covariance matrices. Eigenfaces have been proven to be capable of providing significant features and reducing the input size for neural networks.

(Ibrahim and Zin, 2011) [26] developed automated face recognition system to study the potential application for office door access control. The eigenfaces technique based on PCA and artificial neural networks is applied in this system. The training images can be obtained either offline using advance

captured and cropped face images, or online using face detection and recognition training modules on the real frontal face images of the system. The system can recognize faces at a reasonable rate and at the distance of 40 cm to 60 cm from the camera with the person's head rotational angle between −20° to +20°. The experimental results also confirm the influences of illumination and pose on the face recognition system.

(Abdullah et al., 2012) [23] performed three experements to improve PCA performance by decreasing the computational time while same performance. A first experiment is performed to find the best number of images for each person to be used in the training set that gives a highest recognition rate. In the second experiment, the analysis is tested using 28 images for each person with 6 images used for training process. the best results if obtained by varying the threshold value. In the third experiment the number of eigenvectors is decreased producing less computation time. The performance in terms of accuracy is same with second experiment with less computational time. This approach reduce the computation time by 35 % compared with the original PCA algorithm particularly with a large database.

(Kadam, 2014) [11] combined PCA and DCT and both aim at data dimension reduction. PCA is a very fast algorithm that uses relatively high validity and DCT to reduce time. This hybrid algorithm provides higher recognition rates than the simple PCA. However, (Chen et al, 2007) propose the adaptive PCA to improve PCA results by operating PCA on faces and face space rotation, and then applying warping between classes and within class covariance.

(Bakhshi et al., 2016) [42] presents a method which uses SIFT and SURF features to classify face images and PCA method is applied for better matching results in case of rotation as well as change in expressions. The Result show that the performance of the proposed method give better recognition rate than other methods.

(Poon et al., 2016) [43] different illumination invariant techniques were examined and identifies the one which works better with PCA for face recognition. Experimental results show that by applying the technique called Gradient faces at the pre- processing stage can greatly improve the recognition rates.

(Barnouti, N.H., 2016) [20] performed pre-processing techniques on the ORL database to increase the recognition rate for face recognition using PCA. The experiment result shows increasing in recognition rate when images are resized by 0.3 scale and increase image brightness.

(Barnouti, N.H., 2016) [39] present a method which uses PCA-BPNN with DCT. When BPNN is combined with PCA, the system will recognize faces easily. DCT is used to compress the face databases. Face94 and Grimace databases is used to test the system performance. This system has acceptance recognition rate is more than 90%.

4.2.2 Linear Discriminant Analysis (LDA)
LDA is used for feature extraction and dimension reduction, its widely used to search for linear groups of features while maintaining the separate of classes and also known as Fisherface. LDA is a supervised learning method that uses more than one training image for each class [38].

The original work of the Fisherface projection method is to solve the illumination problem by maximizing the ratio of between-class scatter to within-class scatter. Fisherface is less sensitive to pose, lighting, and expression variations. LDA optimizes the low dimensional representation of the objects [27] [38].

Singularity problem is the limitation of the classic LDA. Face recognition is challenged by the small sample size (SSS) problem through LDA when the total number of training images is less than the feature space dimension. Another LDA problem is the high computational cost in cases of huge data [14].

The images are divided and labeled into between-class and within-class. Between-class capture the image variations of the same individual. While, within-class capture the image variations among classes of individuals. Unlike PCA, LDA attempts to model the differences among classes. LDA obtains different projection vectors for each class, and multi-class LDA, which is widely used in biometric systems, can be performed on more than two classes [15]. The within-class scatter matrix SW and the between-class scatter matrix SB are defined in the equations bellow:

$$SB = \sum_{n=1}^{C} \sum_{m=1}^{L_n} (S_m^n - u_n)(S_m^n - u_n)^T \qquad (4)$$

$$SW = \sum_{n=1}^{C} (u_n - u)(u_n - u)^T \qquad (5)$$

Where S_m^n is the mth sample of class n, u_n is the mean of class n, C is the number of classes, L_n is the number of samples in class n, and u is the mean of all classes.

(Bhattacharyya and Rahul, 2013) [27] applied LDA algorithm on the ORL dataset. Images 1 to 40 are tested, and test images are compared with training images. After LDA, 37 out of 40 images are recognized correctly, while 3 images are recognized incorrectly.

(Murtaza et al., 2014) [14] proposed Adaptive Margin Fisher's Criterion (AMFC) LDA algorithm to overcome the inadequacy of conventional LDA and MMC, and to attack the singularity of the within-class scatter matrix under reasonable cost. The AMFC result is efficient, effective, and has fast convergence rate. False rejection and false acceptance rates are reduced because AMFC does not suffer from the SSS intraclass problem. Computational cost is likewise reduced.

(Barnouti, N.H., et al., 2016) [38] implemented an automatic face recognition system based on appearance-based methods. PCA-LDA implementation was successful. MUCT, Face94, and Grimace databases are used to test the system performance. The experiment results show that the recognition rates will be increased when increase the number of training images.

(Singh, N.A., et al., 2016) [44] proposed a novel method in face recognition using speed up robust feature to improve the quality parameters and linear discriminant analysis for optimize result. SURF is used for feature matching. The proposed method shows better result when comparing with previous method in all cases such as error rate, psnr, mse, accuracy and ssim.

4.2.3 Independnet Component Analysis (ICA)

ICA is used to transform data as linear combinations of statistically independent data points, and aims to supply an independent image representation rather than an uncorrelated one [8]. With Gaussian distribution, PCA treats the image elements as random variables, and the second order statistics are minimized. If PCA is non-Gaussian, dividing a large variation does not correspond to the PCA basics vector. ICA minimize both higher-order and second-order dependencies in the input data, and attempts to determine the basics together with statistically independent data. ICA is a statistical method of transforming observed multidimensional random vectors into component vector that are statistically independent [16]. However, PCA which relies on the pair wise relationship between pixels in the image database, can only represent second-order inter-pixel relationships or relationships that capture the amplitude spectrum of an image, not the phase spectrum of the image. ICA strives to exploit higher-order relationships among pixels. ICA is most widely used to solve the blind source separation problem [17]. It can be useful in various applications, such as image processing, reflection canceling, biomedical signal processing, time series forecasting, data mining, radio communications, text document analysis, and finding hidden factors in financial data [17].

Two different face recognition approaches are performed in ICA. In the first approach, images are considered as random variables and pixels as trials. In other words, ICA architecture initially attempts to search for a set of statistically independent basis images. In the second approach, pixels are considered as random variables and the images as trials. This approach may be more robust to variations, such as change in hair, facial expression, makeup, and lighting conditions, because its designed to maximize information transmission in the existence of noise [13].

(Naik and Kumar, 2011) [17] introduced an ICA/blind source separation (BSS) fundamentals. The mathematical framework of the source mixing problem that ICA and BSS addresses is examined in detail, including the general approach for solving the ICA/BSS. Some inherent ambiguities of the ICA/BSS framework are examined, including the two important preprocessing steps of centering and whitening. The specific details of the method are presented to solve the mixing problem. The application domains of this novel method are also shown. Some of the future works on the ICA method which need further investigation are discussed.

(Sharma and Dubey, 2014) [13] performed a Hybrid feature extraction technique for face recognition using PCA–ICA and training using neural networks to create a robust and reliable face recognition system. This technique aims to implement a PCA/ICA-based facial recognition system by extracting the invariant facial features.

(Bhat and Pujari, 2015) [45] proposed a hybrid technique with holistic-based PCA and ICA face recognition to recognize faces. PCA and ICA performance have a lesser accuracy though ICA is precise enough to consider a solution.

4.2.4 Kernel Principal Component Analysis (KPCA)

Non-linear PCA is also derived from different methods, different kernel methods are generalized to form one of the major types of non-linear PCA called Kernel PCA (KPCA). Primarily, KPCA calculates PCA by mapping the original input into a high dimensional feature space using different kernel methods.

(Liang et al., 2009) [28] used KPCA for face hallucination. The main concept of using KPCA is that this method can classify both linear and nonlinear data. Different resolution images have also been observed to contain similar features in the kernel subspace.

(Wang and Zhang, 2010) [29] outlined another approach for handling facial expressions and extracting suitable features. In this regard, the polynomial kernel is effectively employed to deal with nonlinear structures resulting from expression changes. In addition, the nearest neighbor and Euclidean distance are used for a compact representation. The experiment results are outstanding compared with those of traditional PCA-based approaches.

(Vinay et al., 2015) done comparison between GABOR-PCA and GABOR KPCA variants in order to ascertain the performance variation. ORL database is used to test the system performance. The results demonstrated that the GABOR-PCA method outperformed GABOR-KPCA by 6.67% (Euclidean), 0.83% (Cosine), 12.00% (City Block) and 4.17% (MAHCOS). This upsurge is contrary to the general assumption that KPCA is by default a better choice.

4.2.5 Kernel Linear Discriminant Analysis (KLDA)

Kernel is approach that consider higher-order statistics. First. The input space is mapped onto the feature space through non-linear mapping, and then the PCs in the feature space are determined. They can compute the dot products of two feature vectors even without knowing what the vectors are. The kernel method enables the construction of non-linear versions of any method that can be expressed solely in term of dot products. Given the increase in dimensionality, the mapping is made implicit and economical by using kernel functions that satisfy Mercer's theorem.

(Kamerikar and Chavan, 2014) [30] done a comparison between kernel LDA (KLDA) and LDA methods. This comparison show that LDA works better when a lesser number of training images are chosen per person. For any face database, whether colored or gray scale, LDA performs better than KLDA for a smaller number of training images per person. The projection of training images for the KLDA scatter within classes is closer than in LDA, but the scatter between classes are farther than in LDA. The average recognition rate of the KLDA performance is better than that of LDA.

4.2.6 Support Vector Machine (SVM)

In SVM based classification systems, the classification performed by constructing a new hyper plane that provides maximum separation among the data items. The margin of the hyper plane should be at the maximum. The vectors that found near the hyper plane are called support vectors. Every point in the input space is non-linearly mapped into a high-dimensional feature space. This mapping is done using kernel functions.

(Zagouras et al., 2007) [31] used Kernel functions for the experiments include linear, polynomial, and radial basis function kernels. The experiments are conducted using the ORL face database comprising 400 images of 40 individuals. A total of 200 randomly selected samples serve as the training set. This sample set is used to construct the Eigenfaces and to train the SVM. Changing the feature vector size alters the input vector size of the classifier because the output is given directly to the classifier after the PCA transformation. Based

on the experiments conducted, combining PCA and SVM provides better classification accuracy for face recognition.

(Pal and Foody, 2010) [32] discussed the need for dimensionality reduction as a preprocessing classification step. SVM is a widely used classification method. It constructs a hyper plane, where the margin among different classes is high, thus leading to better classification accuracy. The experiments are more conducted for the classification of hyper spectral data using SVM. The experimental results shown that the classification accuracy of SVM can be increased by reducing the dimensionality of the data. Thus, the results prove that dimensionality reduction is an essential preprocessing stage for SVM classification. By adding features, the SVM classification accuracy. However, even for a large training dataset, dimensionality reduction is important. The experimental results show that dimensionality reduction and SVM classification accuracy are significantly correlated. When the feature dimensions considered are 55 and 65, the classification accuracy are 92.24% and 91.76% respectively.

4.2.7 Hidden Markov Model (HMM)

The stochastic modeling of a non-stationary vector time series based on the midden Markov Model (HMM) is very successful for speech applications. Applying this method on human face recognition intuitvely divides the faces into different regions, such as the eyes, nose, and mouth, which can be associated with the states of an HMM. The images need to be converted into either a 1D temporal sequence or a 1D spatial sequence because HMMs require a 1D observation sequence and 2D images.

(Sun et al., 2010) [33] brefly analyzed different facial patterns via an adaptive 3D face model series. The role of the tracking-model-based method is usually to accommodate the lack of a feature vector. Similarly, a spatial temporal face model descriptor is used to evaluate the system. One of the major drawbacks of HMM is its sensitivity to geometrical shapes. To cope with such a problem, a confirmation-based HMM is appreciated, where the experimental results reveal that the proposed model is highly preferable than traditional HMMs and various standard-based HMM models.

4.2.8 Elastic Bunch Graph Matching (EBGM)

EBGM method identifies a person in a new face image by comparing the new face image to other faces stored in a database. The method extracts feature vectors called Gabor Jets from interest point on the face, and then those features are matched to corresponding features from the other faces in the database [8].

Change in one feature does not disable this method to recognize the person anymore. This method is relatively insensitive to variations in facial positions. EBGM is very sensitive to light and takes too much matching time; thus, it is inappropriate for actual environments. The graphs that are placed on the face should be compared to obtain better performance, but these graphs require a huge storage of convolution images [8].

EBGM method operates in three phases. First, important landmarks on the face located by comparing the Gabor Jets extracted from the new image to those taken from the training images. Second, each face image is processed into a smaller description of that face called a FaceGraph. In the last phase, the similarities among many FaceGraph are compared by calculating the similarities of the Gabor Jets features. The FaceGraph with the highest similarity has the highest possibility of belonging to the same person. Although, all of these phases affect the method performance.

4.2.9 3D Morphable Model

The 3D face recognition methods have become more popular; these methods use 3D sensors to capture information about face shape. They can be divided into two main categories (3D face reconstruction and 3D pose estimation) [24]. The human face is a surface that lies in a 3D space intrinsically. Therefore, 3D model can better represent faces, particularly in handling facial variations such as pose and lighting. The morphable face model is based on a vector space representation of faces constructed in such a way that any convex combination of the shape and texture vectors of a set of examples describes a realistic human face. Fitting a 3D morphable model to images can be used in two ways for recognition across different viewing conditions.

- **Paradigm 1:** After fitting the model, recognition can be based on the model coefficients, which represent intrinsic facial shapes and textures, and are independent of imaging conditions.

- **Paradigm 2:** 3D face reconstruction can be also employed to generate synthetic views from a gallery probe of images. The synthetic views are then transferred to a second viewpoint dependent recognition system.

(Hu et al., 2014) [47] present a novel Albedo Based 3D Morphable Model (AB3DMM). This model will remove the illumination component from the images using illumination normalization in a pre-processing step. A comparative study of 12 different illumination normalization methods are conducted. Multi- PIE database is used to evaluate SSR+LPQ with 86.76% recognition rate.

5. DISTANCE MEASUREMENT

Most face recognition methods from the last decade help in deciding according to the distance measurement. The distance between two images is a major concern in image recognition and computer vision. The final step of face recognition is measuring the distance between two images. Image similarity is the distance between the vectors of two images. The distances among feature space representations are used as the basis for recognition decisions [20]. Many existing image distance methods suffer from complicated measure computations, leading to a difficulty in combining the metric with some face recognition methods [18].

5.1 Euclidean Distance

Euclidean Distance is used to classify and obtain the similarity level, it is widely used because it is faster than other classifiers and is simple [19]. Euclidean Distance is defined as the straight line distance between two points, which examines the root of square differences between the coordinates of a pair of subjects [23]. A minimum Euclidean Distance classifier is the most favorable condition for normally distributed classes [21]. Suppose X is a test image and Y is a train image. Then, Euclidean Distance can be computed as shown in the equation bellow:

$$ED(X,Y)$$
$$= \sqrt{\sum_{i=1}^{No.Of\ Images} (X_i - Y_i)^2} \qquad (6)$$

(Jahanbin et al., 2008) [34] used iso-depth and iso-geodesic to identify facial landmarks (nose tip), and Euclidean distance is used to compare each facial feature. the Euclidean distance decision rule using features extracted from iso-geodesic curves can achieve a remarkable performance of EER= 2.58%. Comparing results appears that LDA followed by Euclidean distance verifier outperforms SVM.

(Malkauthekar et al., 2009) [35] applied Euclidean distance method to recognize the FERET face database. Facial images of two classes and three classes with different expressions and angles are used for classification. The experimental results have been demonstrated that performance of Fisher Discriminant Analysis for three classes is same as the performance for two classes.

(Pornpanomchai and Inkuna, 2010) [3] developed a face recognition system prototype. During the system development, several problems are found. One of these problems is the calculation of Euclidean distances between each facial organ. Out of 200 faces, 192 are recognizable while eight were not. The recognition average time is equal to 3.304 s/image.

(Sayeed et al., 2011) [36] developed a Face recognition system using three methods (DCT, PCA, and Fuzzy) and Euclidean distance and SVM are employed as the classification methods. The CG database is used, and the recognition rate of the segmental Euclidean distance method is 88% using PCA; when using the ORL, this method does not perform well.

5.2 City Block Distance
City Block Distance, which is also known as L1 Distance or Manhattan Distance classifier is the summation of the absolute difference between two vectors [23]. It is especially useful for discrete types of descriptions. City Block Distance is a true distance function because it responds to triangle inequality. It also assumes a triangular distribution [21]. City Block Distance can be computed as shown in the equation bellow:

$$CB(X,Y) = |X - Y| = \sum_{i=0}^{No.Of\ Images} |X_i - Y_i| \qquad (7)$$

5.3 Mahalanobis Distance
Mahalanobis Distance is a distance measure introduced by P. C. Mahalanobis in 1936. It is based on the correlations among variables, where different patterns can be identified and analyzed. This distance gauges the similarity of an unknown sample set to a known one. In contrast to Euclidean Distance, Mahalanobis Distance considers the correlations of the database and is scale-invariant [22]. It is very useful statistic in multivariate analysis and distance measurement for locating the similarity of a set of values from an unknown sample to a set of values measured from a collection of known samples. Mahalanobis Distance can be computed as shown in the equation bellow:

$$MD(X,Y) = \sqrt{(X - Y)^T\ S^{-1}\ (X - Y)} \qquad (8)$$

(Malkauthekar and Sapkal, 2014) [37] compare Mahalanobis distance with the city-block distance using 27 RGB images. Mahalanobis distance recognizes 12 out of 27 images. In addition, it generated better results than the city-block distance that recognizes 11 out of 27 images.

5.4 Square Euclidean Distance
Square Euclidean Distance (SED) measurement can be obtained without the square roots. The standard Euclidean Distance can be squared in order to place progressively greater weight on objects that are farther apart. In this case, the equation becomes as shown below:

$$SquareEuclidean(X,Y) = \sum_{n=1}^{No.of\ Images} (X_n - Y_n)^2 \qquad (9)$$

(Gawande and Agrawal, 2014) [22] examined PCA and generated better results with the square Euclidian distance classifier than the city-block distance classifier. Furthermore, the recognition rate using the square Euclidean distance has the same rate as the simple Euclidean distance, but is higher than when using the city-block distance. The square Euclidean distance is similar to Euclidean distance, but without the square root. The ORL database is used for the test results. It produces interesting results from the viewpoint of recognition success, robustness, and face recognition method rate.

5.5 Chebyshev Distance
Chebychev distance is a metric induced by the supermum norm or uniform norm. It is also called the maximum value distance and is an example of an injective metric. This distance is defined on a vector space, where the distance between two vectors is the greatest of the differences of the vectors along any coordinate dimension [22].

5.6 Hausdorff Distance
Hausdorff distance (HD) measure is the dissimilarity between two sets of points. It can also be applied on edge maps to compare shapes. This measure can be calculated without explicity pairing up the points from two sets because the proximity is considered instead of the exact superposition [40].

6. FACE RECOGNITION DATABASE
Several face databases are available to face recognition researchers. Some of the familiar databases are shown in Table 2. These databases different in size, scope, and purpose. The photographs in many of these databases are acquired by small research teams specifically to study face recognition [38] [39].

Table 2. Face Recognition Databases Features.

Name	Individuals	Image Resolution
AT&T (ORL)	40	92 x 112 Gray Scale
FERET	1199	256 x 384 Gray Scale/ RGB
AR Face Database	126 70 Male 56 Female	576 x 768 RGB
PIE Database, CMU	68	640 x 486 RGB

BioID Face Database	23	382 x 288 Gray Scale
The Yale Face Database	15 14 Male 1 Female	320 x 243 Gray Scale
The Yale Face Database B	10	640 x 480 Gray Scale
UMIST Face Database	20	92 x 112 Gray Scale
The MUCT Landmarked	276	480 x 640 RGB
Face 94	153	180 x 200 RGB
Indian Database	40	640 x 480 RGB
Grimace Database	20	180 x 200 RGB

6.1 AT&T (ORL) Database

The AT&T database, also called ORL database contain 40 different persons with 10 images for each person, with total of 400 images. Four females and 36 males are included in the database. Each image is 92x112 pixels, for a total of 10,304 pixels, and the files are in PGM format. The database is used in the context of a face recognition project performed in collaboration with the Speech, Vision, and Robotics Group of the Cambridge University Engineering Department in the United Kingdom [14] [48].

6.2 FERET Database

The FERET database is a standard dataset used for face recognition systems. The FERET program is managed by the Defense Advanced Research Projects Agency (DARPA) and the National Institute of Standards and Technology. A face imagery database was collected between December 1993 and August 1996. The images are separated into two sets: gallery and probe images. Gallery images are images with known labels, while probe images are matched with gallery images for identification [49] [50].

6.3 AR Face Database

The AR Face database was constructed in 1998 by Aleix Martinez and Robert Benavente in the Computer Vision Center (CVC) at the Universitat Autònoma de Barcelona in Barcelona, Spain [1]. This database is primarily used for face recognition, but may also be used for facial expression recognition. The AR database comprises over 4,000 frontal images of 126 subjects (70 males and 56 females). Each person has 26 different samples captured in two separate sessions in different days, with 13,576x768 pixels images per session [50] [51].

6.4 CMU-PIE Database

The Carnegie Mellon University (CMU) pose, illumination, and expression (PIE) database images were collected between October and December 2000. The PIE database has already made a significant effect on algorithm development for face recognition across poses and on the evaluation of face recognition algorithms. This database systematically considers numerous poses and illumination conditions along with

various facial expressions. This database has been distributed to more than 150 research groups to date [1] [14].

6.5 Yale Face Database B

The Yale face database B is collected to test the face recognition methods systematically under large pose and illumination variations. The images of the persons are captured inside a geodesic dome with 64 computer-controlled xenon strobes. The images of 10 persons are recorded in 9 poses (5 poses at 12°, 3 poses at 24°, and 1 frontal view from the camera axis) under 64 different lighting conditions [1] [14].

6.6 UMIST Face Database

UMIST database is created by the University of Manchester Institute of Science and Technology, which merged with the Victoria University of Manchester to form the University of Manchester. The UMIST database consists of images with varying poses. This database is primarily used for face recognition. A total of 20 persons are included in this database, with a total of 564 8-bit gray and 92x112 images. The image poses are captured from the profile and frontal views [52].

6.7 MUCT Face Database

The Milborrow/University of Cape Town (MUCT) database consists of 3,755 faces with 76 manual landmarks. The database is created to diversify lighting, age, and ethnicity. In this database, all the images captured in December 2008 are from the people around the University of Cape Town campus. The people in this database are university students, parents, high school teachers, and employees. Each person is photographed using five webcams, which makes the database useful for applications that require multiple occurring views of the person [38] [53].

The image format is RGB, with a resolution of 480x640 pixels. Images are captured using webcams with Sony ICX098BQ CCD sensors. Camera is directly in front of the person's face, which is directed up to some height variations of seated persons. No cameras are located on the left of the person because this view can be approximated for model-building purposes by mirroring the images of the cameras on the right side [53].

6.8 Face94 Database

The Face 94 database was collected by Dr. Libor Spacek. This database is a part of a collection of face images. The collection contains four folders of images, namely, face94, face95, face96, and grimace [23]. The images are stored in 24-bit RGB and JPEG formats, and the camera used is a S-VHS camcorder. Face 94 contains 153 images, each with a resolution of 180x200 pixels, and the directories comprise images of male and female persons in separate directories (20 females, 113 males, 20 male staffs). The images are mainly from first year undergraduate students; thus, the majority of the persons are between 18 and 20 years old, although some older persons are also present. The lighting is artificial and some of the images are captured with glasses, and a mixture of tungsten and fluorescent overheads [11] [38].

6.9 Indian Database

The Indian database contains images of 40 distinct persons, each with 11 different poses. The files are in 24-bit RGB and JPEG landscape formats. The size of each image is 640x480 pixels. The images are organized into two main directories

(males and females). All the images have a bright homogeneous background, and the persons are in an upright, frontal position. The poses vary with the face looking upfront, looking left, looking right, looking up, looking up toward the left, looking up toward the right, and looking down with various emotions [54].

6.10 Grimace Database

The Grimace database contains human face images with 20 persons each having 20 images with a small head scale variation. This database contains images of both males and females. The images are in 24-bit RGB JPEG format, and the size of each image is 180x200 pixel portrait formats with plain backgrounds. The lighting of the images minimally varies. The images of the people from various racial origins have major expression variations with beards and glasses. The camera used is a S-VHS camcorder. The lighting used is artificial, with a mixture of tungsten and fluorescent overheads [38] [54].

7. CONCLUSION

In this paper, the importance of face recognition and its numerous applications, algorithms, methods, face databases are discussed. Two major feature extraction methods for face recognition (Appearance-based and Model-Based). Appearance-Based includes PCA, LDA, ICA, and Model-Based includes EBGM and 3D Morphable Model. Different feature extraction methods with its previous work are discussed. Distance Measurement such as Euclidean Distance, City Block, and Mahalanobis Distance are important for recognition process, the Distance Measurement methods are discussed. Several face recognition databases are available and can be used to test the system performance.

8. REFERENCES

[1] Jain, A.K. and Li, S.Z., 2011. Handbook of face recognition. New York: springer.

[2] Dewi Agushinta, R. and Septadepi, I., Face Recognition System Using Eigenface Method based on Facial Component Region.

[3] Pornpanomchai, C. and Inkuna, C., 2010, February. Human face recognition by euclidean distance and neural network. In *Second International Conference on Digital Image Processing*. International Society for Optics and Photonics. pp. 754603-754603.

[4] Zhao, W., Chellappa, R., Phillips, P.J. and Rosenfeld, A., 2003. Face recognition: A literature survey. *ACM computing surveys (CSUR)*, 35(4), pp.399-458.

[5] Chellappa, R., Wilson, C.L. and Sirohey, S., 1995. Human and machine recognition of faces: A survey. *Proceedings of the IEEE*, 83(5), pp.705-741.

[6] Jafri, R. and Arabnia, H.R., 2009. A survey of face recognition techniques.*Jips*, 5(2), pp.41-68.

[7] Lu, X., 2003. Image analysis for face recognition, personal notes. *Dept. of Computer Science and Engineering*.

[8] Shah, D.H., Shah, D.J. and Shah, D.T.V., 2014. The Exploration of Face Recognition Techniques. *International Journal of Application or Innovation in Engg. And Management (IJAIEM) Web Site: www. ijaiem. org Email: editor@ ijaiem. org, editorijaiem@ gmail. com*, 3(2).

[9] Parmar, D.N. and Mehta, B.B., 2014. Face Recognition Methods & Applications. *arXiv preprint arXiv:1403.0485*.

[10] Kurmi, U.S., Agrawal, D. and Baghel, R.K., Study of different face recognition algorithms and challenges. *IJER, Volume*, (3), pp.112-115.

[11] Kadam, K.D., 2014. Face Recognition using Principal Component Analysis with DCT. *International Journal of Engineering Research and General Science, ISSN*, pp.2091-2730.

[12] Slavković, M. and Jevtić, D., 2012. Face recognition using eigenface approach. *Serbian Journal of Electrical Engineering*, 9(1), pp.121-130.

[13] Sharma, N. and Dubey, 2014. A.P.S.K., Face Recognition Analysis Using PCA, ICA And Neural Network. *International Journal of Digital Application & Contemporary research,* 2(9).

[14] Murtaza, M., Sharif, M., Raza, M. and Shah, J., 2014. Face recognition using adaptive margin fisher's criterion and linear discriminant analysis.*International Arab Journal of Information Technology*, 11(2), pp.1-11.

[15] Sahu, R.K., Singh, Y.P. and Kulshrestha, A., 2013. A Comparative Study of Face Recognition System Using PCA and LDA. *International Journal of IT, Engineering and Applied Sciences Research*, 2(10).

[16] Christry, C., Pavithra, J. and Sathya, F.M., 2014. Comparison on PCA ICA and LDA in Face Recognition. *International Journal of Computing Algorithm*,3, pp.917-922.

[17] Naik, G.R. and Kumar, D.K., 2011. An overview of independent component analysis and its applications. *Informatica*, 35(1).

[18] Wang, L., Zhang, Y. and Feng, J., 2005. On the Euclidean distance of images. *IEEE transactions on pattern analysis and machine intelligence*, 27(8), pp.1334-1339.

[19] Taghizadegan, Y., Ghassemian, H. and Naser-Moghaddasi, M., 2012. 3D Face Recognition Method Using 2DPCA-Euclidean Distance Classification.*ACEEE International Journal on Control System and Instrumentation*, 3(1).

[20] Barnouti, N.H., 2016. Improve Face Recognition Rate Using Different Image Pre-Processing Techniques. *American Journal of Engineering Research (AJER),* 5(4), pp. 46-53.

[21] Kapoor, S., Khanna, S. and Bhatia, R., 2010. Facial gesture recognition using correlation and mahalanobis distance. *arXiv preprint arXiv:1003.1819*.

[22] Gawande, M. P., and Agrawal, D. G., 2014. Face recognition using PCA and different distance classifiers. *IOSR Journal of Electronics and Communication Engineering (IOSR-JECE),* 9(1), pp. 1-5.

[23] Abdullah, M., Wazzan, M. and Bo-Saeed, S., 2012. Optimizing face recognition using PCA. *arXiv preprint arXiv:1206.1515*.

[24] Nicholl, P. and Amira, A., 2008. DWT/PCA face recognition using automatic coefficient selection. In *Electronic Design, Test and Applications, 2008. DELTA 2008. 4th IEEE International Symposium on*, pp. 390-393. IEEE.

[25] Kshirsagar, V.P., Baviskar, M.R. and Gaikwad, M.E., 2011. March. Face recognition using Eigenfaces. In *Computer Research and Development (ICCRD), 2011 3rd International Conference on*, 2, pp. 302-306. IEEE.

[26] Ibrahim, R. and Zin, Z.M., 2011. Study of automated face recognition system for office door access control application. In *Communication Software and Networks (ICCSN), 2011 IEEE 3rd International Conference on*, pp. 132-136. IEEE.

[27] Bhattacharyya, S.K. and Rahul, K., 2013. Face Recognition By Linear Discriminant Analysis. *International Journal of Communication Network Security, ISSN*, pp. 2231-1882.

[28] Liang, Y., Lai, J.H., Zou, Y.X., Zheng, W.S. and Yuen, P.C., 2009, October. Face hallucination through KPCA. In *Image and Signal Processing, 2009. CISP'09. 2nd International Congress on*, pp. 1-5. IEEE.

[29] Wang, Y. and Zhang, Y., 2010, November. Facial recognition based on kernel PCA. In *Intelligent Networks and Intelligent Systems (ICINIS), 2010 3rd International Conference on*, pp. 88-91. IEEE.

[30] Kamerikar, U.A. and Chavan, M.S., 2014. Experimental Assessment of LDA and KLDA for Face Recognition. *International Journal*, 2(2).

[31] Zagouras, A., Economou, G., Macedonas, A. and Fotopoulos, S., 2007, October. An application study of manifold learning-ranking techniques in face recognition. In *Multimedia Signal Processing, 2007. MMSP 2007. IEEE 9th Workshop on*, pp. 445-448. IEEE.

[32] Pal, M. and Foody, G.M., 2010. Feature selection for classification of hyperspectral data by SVM. *IEEE Transactions on Geoscience and Remote Sensing*, 48(5), pp. 2297-2307.

[33] Sun, Y., Chen, X., Rosato, M. and Yin, L., 2010. Tracking vertex flow and model adaptation for three-dimensional spatiotemporal face analysis. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 40(3), pp. 461-474.

[34] Jahanbin, S., Choi, H., Liu, Y. and Bovik, A.C., 2008, September. Three dimensional face recognition using iso-geodesic and iso-depth curves. In *Biometrics: Theory, Applications and Systems, 2008. BTAS 2008. 2nd IEEE International Conference on*, pp. 1-6. IEEE.

[35] Malkauthekar, M.D., Sapkal, S.D. and Kakarwal, S.N., 2009, March. Experimental Analysis of Classification of Facial Images. In *Advance Computing Conference, 2009. IACC 2009. IEEE International*, pp. 1093-1098. IEEE.

[36] Sayeed, F., Hanmandlu, M. and Ansari, A.Q., 2011. Face recognition using segmental euclidean distance. *Defence Science Journal*, 61(5), pp. 431-442.

[37] Malkauthekar, M. D., and Sapkal, S. D., 2014. Comparision of mahalanobis and manhattan distance measures in pca based face recognition. *International

Journal of Computer & Technology (IJCET)*, 5(5), pp 01-11.

[38] Barnouti, N.H., Al-Dabbagh, S.S.M., Matti, W.E., and Naser, M.A.S., 2016. Face Detection and Recognition Using Viola-Jones with PCA-LDA and Square Euclidean Distance. *International Journal of Advanced Computer Science and Applications(IJACSA)*, 7(5).

[39] Barnouti, N.H., 2016. Face Recognition using PCA-BPNN with DCT Implemented on Face94 and Grimace Databases. 2016. *International Journal of Computer Applications*, 142(6), pp. 8-13.

[40] Nigam, A. and Gupta, P., 2009. A new distance measure for face recognition system. In *Image and Graphics, 2009. ICIG'09. Fifth International Conference on*, pp. 696-701. IEEE.

[41] Kasar, M.M., Bhattacharyya, D. and Kim, T.H., 2016. Face Recognition Using Neural Network: A Review. *International Journal of Security and Its Applications*, 10(3), pp.81-100.

[42] Bakhshi, Y., Kaur, S. and Verma, P., 2016. An Improvement in Face Recognition for Invariant Faces. 6(2), pp. 423-426.

[43] Poon, B., Amin, M.A. and Yan, H., 2016. Improved Methods on PCA Based Human Face Recognition for Distorted Images. In *Proceedings of the International MultiConference of Engineers and Computer Scientists* (1).

[44] Singh, N.A., Kumar, M.B. and Bala, M.C., 2016. Face Recognition System based on SURF and LDA Technique. *International Journal of Intelligent Systems and Applications*, 8(2).

[45] Bhat, V.S. and Pujari, J.D., 2015. Human Face Recognition Using Combined Approaches PCA and ICA. *Digital Image Processing*, 7(10), pp. 301-307.

[46] Vinay, A., Shekhar, V.S., Murthy, K.B. and Natarajan, S., 2015. Face Recognition Using Gabor Wavelet Features with PCA and KPCA-A Comparative Study. *Procedia Computer Science*, 57, pp. 650-659.

[47] Hu, G., Chan, C.H., Yan, F., Christmas, W. and Kittler, J., 2014, September. Robust face recognition by an albedo based 3D morphable model. In *Biometrics (IJCB), 2014 IEEE International Joint Conference on*, pp. 1-8. IEEE.

[48] Hjelmås, E., 2000. Feature-based face recognition. In *NOBIM Proceedings (Norwegian Image Processing and Pattern Recognition Conference)*.

[49] Sodhi, K.S. and Lal, M., 2013. Comparative analysis of PCA-based face recognition system using different distance classifiers. *Int. J. of Appl. or Innovation in Eng. & Manag*, 2, pp. 341-348.

[50] Huang, G.B., Ramesh, M., Berg, T. and Learned-Miller, E., 2007. *Labeled faces in the wild: A database for studying face recognition in unconstrained environments*, 1(2). Technical Report 07-49, University of Massachusetts, Amherst.

[51] Zou, J., Ji, Q. and Nagy, G., 2007. A comparative study of local matching approach for face recognition. *IEEE

Transactions on image processing, *16*(10), pp. 2617-2628.

[52] Gottumukkal, R. and Asari, V.K., 2004. An improved face recognition technique based on modular PCA approach. *Pattern Recognition Letters*, 25(4), pp. 429-436.

[53] Milborrow, S., Morkel, J. and Nicolls, F., 2010. The MUCT landmarked face database. *Pattern Recognition Association of South Africa*, 201(0).

[54] Kamerikar, U.A. and Chavan, M.S., 2014. Experimental Assessment of LDA and KLDA for Face Recognition. *International Journal*, 2(2).