Enhancing 3D Face Recognition based PCA by using Rough Set Theory

A.Sh. Ahmed
Dep. of Info. Technology,
Inst. of Gradu. Studies & Research,
Alexandria University, Egypt

Sh. K. Guirguis
Dep. of Info. Technology,
Inst. of Gradu. Studies & Research,
Alexandria University, Egypt

M. Z. Rashad
Dep. of Comp. sciences,
Faculty of Comp. and Info.,
Mansoura University, Egypt

ABSTRACT
Face recognition is a biometric authentication method that has become more and more relevant in the recent years. From being too inaccurate, it is becoming a more mature technology deployed in large scale systems like the new Visa Information System, etc. Sophisticated commercial systems have been developed that achieve high recognition rates. The proposed method of 3D facial recognition based on Rough set technique. In this paper PCA (Principal Component Analysis) approach has been used to reduce Feature vector, for selection of feature have been used the concept of Rough set approach that can be based on the minimal description length principle and tuning methods of parameters of the approximation spaces to obtain high quality classifiers. Finally, Classification of face applied by using Euclidean Distance (ED) and displaying the result to show efficiency and accuracy of proposed method.

General Terms
Pattern recognition, Principle component analysis PCA, Eigenfaces, Rough set theory, Euclidean distance.

Keywords
Pattern recognition, 3D Face recognition, Principle component analysis PCA, Eigenfaces, Rough set theory, Euclidean distance.

1. INTRODUCTION
Face recognition is one of the biometric techniques used in access control systems, surveillance systems, credit card payment systems, etc. Face recognition based on 2D face image has already been maturely developed [1]. In order to achieve higher accuracy, researchers introduce face recognition technique based on 3D data which appeared in the late 20th century and has been utilized widely recently. Although 3D image are more complicated than 2D images, they are invariant in illumination and accurate in geometric information which provides extra precision for the object recognition.

There are many researches using 2D color pictures adopted face recognition [2][3]. Many reasons led to this: first, it is possible to program a computer to identify humans from 2D pictures because human can identify other humans from 2D pictures also the computer can do. Second, till now 3D surface actuation hardware has been either low resolution or expensive and huge devices. However, the 2D data is no longer valid, new motivations which are interested in the use of 3D data for face recognition arisen. The stronger classifiers 3D data presented a high challenge for pure 2D techniques because of its construction which is more flexible to circumstances such as lighting conditions and cosmetics.

Even though 3D data in face recognition has its own challenges, its being used. Two of the largest challenges are, (1) The inaccurate global face matching due to changes in facial expressions that distort the global shape of the face and (2) the confused the registration process by occluded or non-frontal face scans which limit the utilized facial surface .

Face recognition can usually be used for either verification or identification. Individual in verification is already enrolled in the reference database or gallery i.e. it is a one-to-one matching mission but in identification, investigated image is matched with a biometric reference in the gallery i.e. it is a one-to-many problem.

Many technique of dimensionality reduction are available, Principle Component Analysis (PCA) [4] is one of a powerful technique for reducing a large set of correlated variables to a smaller number of uncorrelated components. Rough sets technique introduced by Pawlak [5]. It is another mathematical approach to solve imperfect knowledge problems. Its methodology is concerned with the classification and analysis of imprecise, uncertain or incomplete information and knowledge, and of is considered one of the first non-statistical approaches in data analysis. The main goal of the rough set analysis is to synthesize approximation of concepts from the acquired data and makes reduction of data to a minimal representation.

In this paper the authors attempt to enhance the performance of the PCA, we proposed system based on rough set theory, which take facial features as input data. Our experiment results show that rough set theory and method are effective in facial expression recognition, and high recognition rate is resulted.

2. PROPOSED METHOD
One of the most common challenges in classification problems is Curse of dimensionality; in this regard the number of extracted features is extremely high in comparison with the number of classes. There are several ways to overcome this problem such as: PCA, LDA, ICA, … etc. .

(Figure 1) shows the general block diagram of proposed method based 3D face recognition system achieving a better classification performance. This Figure is adapted from the 2D face recognition with modified blocks for 3D to 2D conversion. The figure shows that the system depends on PCA as data representation to project face patterns starting from higher-dimension image space to a lower dimensional space while it retains as much variation as possible in the data set.

To convert from 3D to 2D image, the mean is used (or max or min) of the image or other image distributed in 3 spatial dimensions.

xyImage = mean(volumeImage, 3);

xzImage = mean(volumeImage, 2);

yzImage = mean(volumeImage, 1);

The aim of this section is to introduce the proposed method for face recognition system which depends on combining three methods, principal component analysis, rough set theory and Euclidean distance.

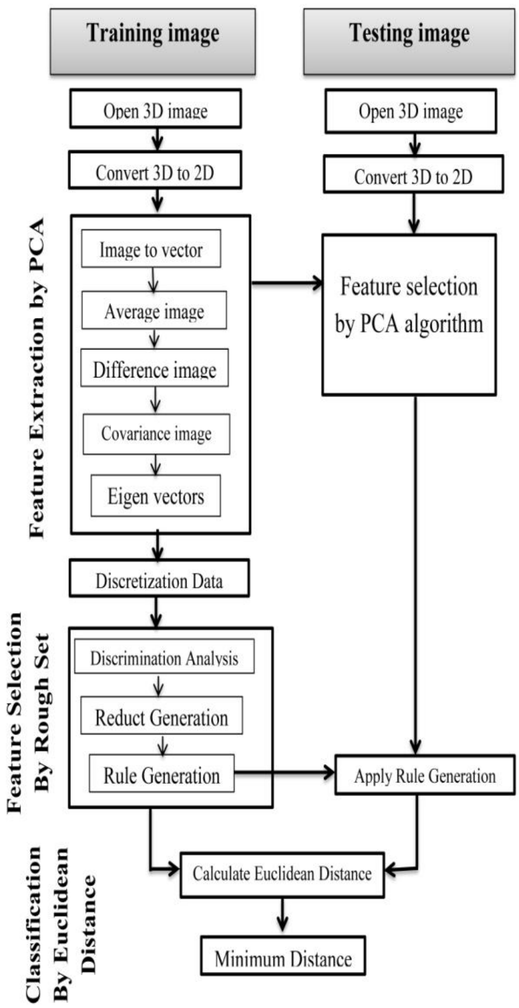

Fig 1: proposed system

2.1 Principal Component Analysis (PCA)for Dimensionality Reduction

Facial recognition can be decomposed into four phases: preprocessing phase, segmentation or localization phase, feature extraction phase and recognition phase. One of the most popular algorithms is principal component analysis (PCA) [6]. In PCA, the probe and gallery images must be the same size. Therefore, a normalization is needed to lineup the eyes and the mouths across all images. Each image is treated as one vector. All images of the training set are stored in a single matrix T and each row in the matrix represents an image. The average image has to be calculated and then subtracted from each original image in T. Then calculate the eigenvectors and eigenvalues of the covariance matrix S. These eigenvectors are called eigenfaces. The eigenfaces are the result of the reduction in dimensions which removes the unuseful information and decomposes the face structure into the uncorrelated components (eigenfaces). Each image may be represented as a weighted sum of the eigenfaces.

2.1.1 Eigenspace Projection

The Eigen Object Recognizer class applies PCA on each image [7][8], the results of which will be an array of Eigen values that a Euclidean Distance can be trained to recognize. PCA is a commonly used method of object recognition as its results, when used properly can be fairly accurate and resilient to noise. The method of which PCA is applied can vary at different stages so what will be demonstrated is a clear method for PCA application that can be followed. It is up for individuals to experiment in finding the best method for producing accurate results from PCA.

To perform PCA several steps are to be performed:

<u>Training Set</u>

- **Normalize the input images**

Contain all PCA method math work in program take the image wanted to recognition and compare it with database images. The training database consists of M images of the same size. The images are normalized by converting each image matrix to equivalent vector Γi . The training set matrix Γ is the set of image vector witch.

Training set $\Gamma = [\Gamma 1, \Gamma 2, \dots , \Gamma m]$ ……(1)

- **Averaging the face images**

$$\Psi=\frac{1}{M}\sum_{i=1}^{M}\Gamma i \quad \dots\dots\dots..(2)$$

(ψ) is the mean face

- **The difference images**

Calculate the difference images to remove average.

$$\Phi_{i} = \Gamma_{i} - \psi \quad \dots\dots\dots\dots.(3)$$

Subtract the Mean of the data from each variable (our adjusted data).Consider a difference matrix A= [$\Phi 1$, $\Phi 2$, $\Phi 3$, .ΦM] witch keeps only the distinguishing features for face images and removes the common features.

- **Covariance matrix**

In this stage eigenfaces are calculated by finding the Covariance matrix C of the training image vectors by:

$$C= A.A^{T} \quad \dots\dots\dots\dots..(4)$$

- **Calculate Eigenvectors**

As Dim (A) = (m*n) x N and Dim (AT) = N x (m*n) SO Dim(C) = (m*n) x (m*n) which is very large and computation will be difficult. So another matrix has been created by rearranging the eigenvectors of C (matrix U) can be obtained by using the eigenvectors of L (Matrix V) as given by:

$$U_i=AV_i \quad \dots\dots\dots\dots\dots.(5)$$

$$\text{Eigenface} = [U_1, U_2, U_3, \dots.., U_M] \dots\dots(6)$$

- **Discover the Eigen faces**

Instead of using M eigenfaces, the highest m, <=M is chosen as the eigenspace. Then the weight of each eigenvector ωi to represent the image in the eigenface space, as given by:

$$\omega_i = U_i^{T}(\Gamma-\psi) , \quad i =[1,2,\dots\dots m] \quad \dots\dots\dots..(7)$$

$$\text{Weight matrix } \Omega= [\omega_1, \omega_2,\dots,\omega_m]^{T} \dots\dots..(8)$$

Average class projection $\Omega_\psi = \frac{1}{xi}\sum_{i=1}^{xi}\Omega i \dots.(9)$

2.2 Discretization Algorithms

Rough set theory cannot deal with continuous attributes for the final feature selection/reduction of the reduced PCA continuous-valued patterns. Discretization of the continuous reduced PCA features we have applied the method of dividing each attribute value range into bins. The discretized training set was used to find relevant reducts, e.g., the minimal reduct. This reduct was used to form the final pattern. Real-value pattern attributes were reduced according to the selected reduct [9]. Used in this paper Algorithm (Boolean Reasoning Algorithm).

ROSETTA software [10] has been used to implement discretization algorithms.

2.3 Feature Selection using Rough Set

Till this point, we use all eigenvectors associated with non-zero eigenvalues, while making a subspace using Eigenspace projection. The computation time of Eigenspace projection is directly related to the number of eigenvectors used to create the eigenspace. Consequently, to decrease computation time some of the eigenvalues are removed if it is possible.Furthermore, by removing additional eigenvalues that do not contribute to the classification of the image, performance can be improved. The importance of feature selection is due to the potential for speeding up the processes of both concepts reducing the cost and improving the quality of classification. Eigenspace projection does not guarantee that the reduced space with transformed feature vector is minimal. To keep discriminative features from the principal components rough set theory can be applied [11]. Rough set approach to feature selection can be based on the minimal description length principle and tuning methods of parameters of the approximation spaces to obtain high quality classifiers based on selected features [12].

2.3.1 Rough set

Rough set theory has many applications in the areas like soft computing, machine learning, knowledge representation, decision making, data mining, expert systems, pattern classification and scene analysis.

To define the rough set [5][13] let us consider a knowledge base Information system is a tuple (U, A), where U consists of objects and A consists of features .Every a ∈ A corresponds to the function a :U →Va where Va is the value set of a. In the applications, we often distinguish between conditional features C and decision feature D, where C ∩ D =φ. In such cases, we define decision systems (U, C, D).

Any set X, which is a subset of U, can be characterized with respect to R as follows:

- **The B-*lower approximation*** of *X* is the set of all objects, which can be certainly classified as *X* with respect to R and that can be given a set B ⊆ A, the lower and upper approximations of a set Y ⊆U are defined by, respectively,

$$\underline{B}Y = \{x \mid [x]_B \subseteq X\} \ldots\ldots\ldots\ldots\ldots(10)$$

- **The B-*upper approximation*** of *X* is the set of all objects, which can be *possibly* classified as *X* with Respect to *R* and that can be given as:

$BY = \bar{B}(X) = \{x \mid [x]_B \cap X \neq \emptyset\}$…….(11) The positive region of X is defined as:

$$POS_C(D) = \bigcup_{X\in U/D} CX \ldots\ldots\ldots\ldots(12)$$

The expression POSC(D), called a positive region of the partition U/D with respect to C, is the set of all elements of U that can be uniquely classified to blocks of the partition U/D, by means of C. Summing up: D is totally (partially) dependent on C, if all (some) elements of the universe U can be uniquely classified to blocks of the partition U/D, employing C.

POS (D) is the set of all objects in U that can be uniquely classified by elementary sets in the partition U/IndD by means of C (10), the negative region $NEG_C(D)$ is defined by:

$$NEG_C(D) = U - \bigcup_{X\in U/D} \underline{C}X \ldots\ldots\ldots\ldots(13)$$

- The **B-*boundary region*** of a set *X* with respect to *R* is the set of all objects, which can be classified Neither as *X* nor as not-*X* with respect to *R* and that is given as:

$$BN_B(X) = \bar{B}X - \underline{B}X \ldots\ldots\ldots\ldots(14)$$

From the above definitions it is evident that it if the boundary region is empty then the set *X* is crisp i.e. exact with respect to *B* but if the boundary region is nonempty then the set *X* is rough i.e. inexact with respect to *B*.

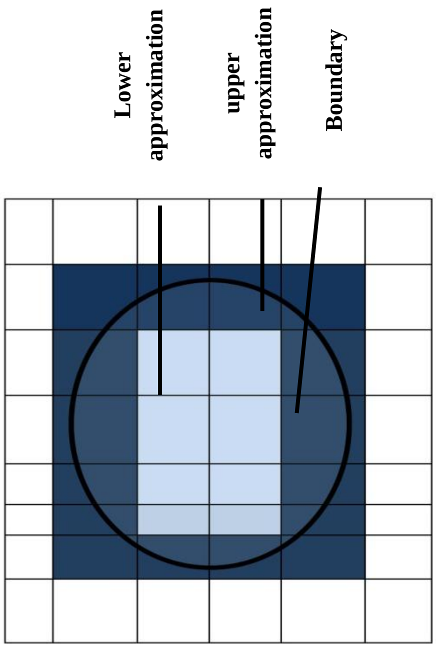

Fig 2: Rough set approximation

- **(Degree of Dependency)** Given a decision system, the degree of dependency of D on C can be defined as:

$$\gamma(C, D) = \text{card}\,(POS_C(D))/\text{card}(U) \ldots(15)$$

- **(Dispensable & Indispensable Attributes)** Let decision table T = (U, C, D) and c ⊆ C. Attribute c is dispensable in T if $POS_C(D) = POS_{(C-\{c\}}(D)$, otherwise attribute c is indispensable in T.

- **(Reduct)** The set of attributes R⊆C is called a reduct of C, if T' = (U, R, D) is independent and $POS_R(D) = POS_C(D)$. The set of all the condition attributes indispensable in T is denoted by CORE(C). A reduct is a subset R ⊆ C such that:

$$\gamma(C, D) = \gamma(R, D) \ldots\ldots\ldots\ldots(16)$$

The reduct set is a minimal subset of attributes that preserves the degree of dependency of decision attributes on full condition attributes. The intersection of all the relative reduct sets is called core.

$$CORE(T) = \cap RED(T) \ldots\ldots\ldots(17)$$

Where RED (T) is the set of all reducts of T.

2.4 Algorithm for Feature Selection using Rough Sets

In principle component space (eigenspace), after dimensional reduction of the input feature vectors the following algorithm must be applied for further reduction:-

Step1: obtain Yd, which is a discrete form of Y.

Step2: make the decision table using Yd as condition and attribute value representing corresponding class as decision.

Step3: From each one of the decision tables find the particular core.

Step4: Find the value that reduces from decision table containing these columns from the projected discrete matrix Yd which are corresponding to the selected feature set. Label the patterns by corresponding classes from the original data set.

Feature selection has been implemented using ROSETTA SOFTWARE [10]. Several methods have been tested to reduce the number of eignfaces attributes through applying them on each of the six reduction algorithms, namely, Genetic algorithm, Johnson's algorithm, Dynamic reducts (RSES), Exhaustive calculation (RSES), Johnson's algorithm (RSES) and Genetic algorithm (RSES). By generating each algorithm reduces and rules. The data is available as Microsoft Access database, loaded in ROSETTA as an ODBC and applied in each reduction algorithm.

Testing Images:-

On the other side a test image will be taken for comparison. This image will go through the same processing steps as the training set. The proposed system implements the Principle Component Analysis algorithm on the test images. Finally Rough set reduct generation will be applying on eigen vector of test image to be ready for matching and classification.

2.5 Classification

In this stage a probe image is compared against the gallery by measuring the distance between their represent vectors. Euclidean distance is used for classification and obtaining the level of similarity. The Euclidean distance δi is used to find out the distance between two face keys vectors and is given equation:-

$$\delta_i = \left\| \Omega - \Omega\psi i \right\| = \sum_{k=1}^{M} (\Omega - \Omega\psi ik) \quad (18)$$

The smallest distance is considered to be the face match score result. The Euclidean distance between two weight vectors d(i,j) provides a measure of similarity between the corresponding images i and j. If the Euclidean distance between Γ_{new} and other faces exceeds - on average - some threshold value θ, one can assume that Γ_{new} is no face at all. d(i,j) also allows one to construct "clusters" of faces such that similar faces are assigned to one cluster[14].

There are 3 cases of the distance (δ_i):

It is defined a threshold or level of acceptance below which it can be safely said that both faces are the same identity:-

- **Case 1**: (δ_i) is greater than the face threshold. -- This means the input image is not a face.
- **Case 2.**: (δ_i) is smaller than the face threshold, but still greater than some reorganization threshold. -- This means the input image is a face image, but not in the database so that we can't recognize the person.
- **Case 3**: (δ_i) is smaller than the reorganization threshold. -- This means we recognize the image. Then the one with smallest distance is the recognized person.

3. RESULTS AND DISCUSSION

The proposed method is implemented using Pentium(R) Dual-core CPU running at 2.30 GHz machine and MATLAB 8.2.0.701(R2013b). As in typical biometric systems, the proposed method includes two phases: the training phase and the testing phase. GAVAB 3D face database [15] that is divided into two subsets, which are the training set and probe set. The training set has 120 subjects (classes) with three subsets of each of 40 subjects and each subset contains 3 face images. The other 40 images are randomly chosen as probe set (testing set) from the databases. The sample 3D GAVAB database images which are used for experimentation are shown in the (Figure 3).

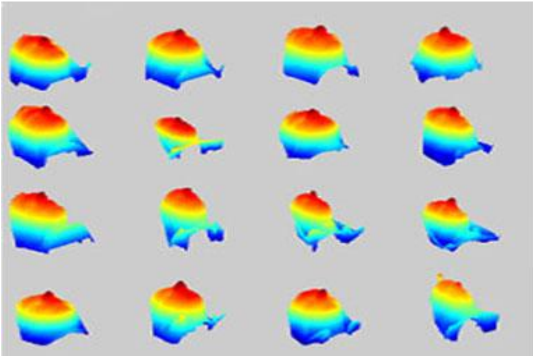

Fig 3: Sample of 3D face images

The performance comparison of the proposed method with the PCA+ED method, in terms of recognition accuracy is presented in the (Figure 4).

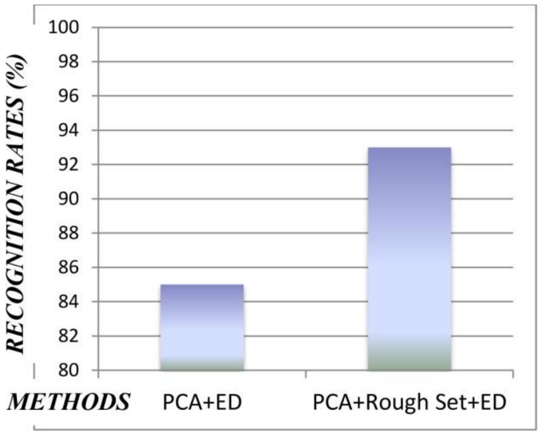

Fig 4: The recognition accuracy

4. CONCLUSION

Propose method presented a rough set method and its role in feature selection for pattern recognition. The authers proposed the sequence of 3D face recognition steps, including application of PCA, and rough sets for feature selection. This processing sequence has shown a potential for feasible feature extraction and feature selection in Euclidean distance for 3D face images. The discussed method provides substantial reduction of pattern dimensionality. Rough set methods have shown ability to reduce significantly the pattern dimensionality and have proven to be viable recognition techniques as a front end of Euclidean distance classifiers.

5. FUTURE WORKS

Face recognition system is designed, implemented and tested. Test results show that system has acceptable performance. On the other hand, system has some future works for improvements and implementation. It will be considered in future to:-

- Using large and other dataset in experiments.
- Using neural network classifier instead Euclidean distance.

6. REFRENCES

[1] W. Zhao, R. Chellappa, P. Phillips , A. Rosenfeld, Face recognition: A literature survey, *ACM Computing Survey*, Volume 35, No. 4, 2003, page 399-458

[2] Rajkiran Gottumukkal, Vijayan K.Asari, An improved face recognition technique based on modular PCA approach, Pattern Recognition Letters 25, 2004, page 429-436

[3] Arshi Shamsi, Waseem Ahmad, Saoud Sarwar, Face Recognition Using Principle Component Analysis Techniques, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 8, August 2013, page 755-757

[4] Jolliffe, I.T., Principal Component Analysis, Springer, Heidelberg, 1986.

[5] Zbigniew Suraj, Rough Sets, International Journal of Information and computer Science 11, 1982, page 341–356

[6] Matthew Turk, Alex Pentland, Eigenfaces for Recognition, Journal of Cognitive Neurosicence, Volume 3, No. 1, 1991, page 71-86

[7] M. Kirby and L. Sirovich, Application of the Karhunen-Loeve procedure for the characterization of human faces, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, Volume 12, No. I , JANUARY 1990, page 103-108

[8] Manal Abdullah, Majda Wazzan, Sahar Bo-saeed, OPTIMIZING FACE RECOGNITION USING PCA, International Journal of Artificial Intelligence & Applications (IJAIA), Volume 3, No.2, March 2012, page 23-31

[9] Cheng-Jung Tsai, Chien-I. Lee, Wei-Pang Yang, A discretization algorithm based on Class-Attribute Contingency Coefficient, Volume 178,Issue 3, February 2008, Pages 714–731

[10] http://www.lcb.uu.se/tools/rosetta/ROSETTA is a toolkit for analyzing tabular data within the framework of rough set theory.

[11] K. Thangavel , A. Pethalakshmi, Dimensionality reduction based on rough set theory: A review, Applied Soft Computing, Vol. 9, Issue 1, 2009, page 1-12

[12] M. Zhang J. T. Yao, Rough set methods in feature selection and recognition, Pattern Recognition Letters, Volume 24, Issue 6, 2003, page 833–849.

[13] Yasser Fouad Hassan, Nora Habeb, Hybrid System of PCA, Rough Sets and Neural Networks for Dimensionality Reduction and Classification in Human Face Recognition, International Journal of Intelligent Information Processing (IJIIP), Volume 3, No. 1, 2012, page 16-24

[14] Daniel Georgescu, A Real-Time Face Recognition System Using Eigenfaces, Journal of Mobile, Embedded and Distributed Systems, Volume 3, No. 4, 2011, page 193-204

[15] www.gavab.etsii.urjc.es/recursos.html Gavab DB 3D face database is created by Gavab Research Group Department of computing in Universidad Rey Juan Carlos, Madrid, Spain. Gavab DB 3D face database is created to be used for automatic facial recognition experiments and other possible facial applications like pose correction or register of 3D facial models.