



HAL
open science

A Novel Technique For Human Face Recognition Using Fractal Code and Bi-dimensional Subspace

Benouis Mohamed, Benkkadour Mohamed Kamel, Tlmesani Redwan, Senouci Mohamed

► **To cite this version:**

Benouis Mohamed, Benkkadour Mohamed Kamel, Tlmesani Redwan, Senouci Mohamed. A Novel Technique For Human Face Recognition Using Fractal Code and Bi-dimensional Subspace. 5th International Conference on Computer Science and Its Applications (CIIA), May 2015, Saida, Algeria. pp.31-42, 10.1007/978-3-319-19578-0_3. hal-01789950

HAL Id: hal-01789950

<https://inria.hal.science/hal-01789950>

Submitted on 11 May 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

A Novel Technique For Human Face Recognition Using Fractal Code And Bi-dimensional Subspace

Benouis Mohamed¹, Benkkadour Mohamed Kamel²,
Tlmesani Redwan², Senouci Mohamed¹

¹Computer Science Department, University Of Oran, Algeria

¹mhbenouis@yahoo.com, msenouci@yahoo.fr

²Computer Science Department, University Sidi Bel Abbes, Algeria

²INTTIC, ORAN, Algeria

²kamel.live@com, ²rtlemsani@ito.dz

Abstract. Face recognition is considered as one of the best biometric methods used for human identification and verification; this is because of its unique features that differ from one person to another, and its importance in the security field. This paper proposes an algorithm for face recognition and classification using a system based on WPD, fractal codes and two-dimensional subspace for feature extraction, and Combined Learning Vector Quantization and PNN Classifier as Neural Network approach for classification. This paper presents a new approach for extracted features and face recognition. Fractal codes which are determined by a fractal encoding method are used as feature in this system. Fractal image compression is a relatively recent technique based on the representation of an image by a contractive transform for which the fixed point is close to the original image. Each fractal code consists of five parameters such as corresponding domain coordinates for each range block. Brightness offset and an affine transformation. The proposed approach is tested on ORL and FEI face databases. Experimental results on this database demonstrated the effectiveness of the proposed approach for face recognition with high accuracy compared with previous methods.

Keywords: biometric, face recognition, 2DPCA, 2DLDA, DWT, PNN, WPD, IFS, fractal codes, LVQ

1 Introduction

The security of persons, goods or information is one of the major concerns of the modern societies. Face recognition is one of the most commonly used solutions to perform automatic identification of persons. However, automatic face recognition should consider several factors that contribute to the complexity of this task such as the occultation, changes in lighting, pose, expression and structural components (hair, beard, glasses, etc.) [1]. Several techniques have been proposed in the past in order to solve face recognition problems. Each of them evidently has their strengths and weaknesses, which, in most of the cases, depend on the conditions of acquiring in-

formation. Recently, several efforts and research in this domain have been done in order to increase the performance of the recognition, such as support vector machine (SVM), Markov hidden model (HMM), probabilistic methods (Bayesian networks) and artificial neural networks. This latter has attracted researchers because of its effectiveness in detection and classification of shapes, which has been adopted in new face recognition systems [2].

2 Face recognition system

A face recognition system is a system used for the identification and verification of individuals, which checks if a person belongs to the system's database, and identifies him/her if this is the case.

The methods used in face recognition based on 2D images are divided into three categories: global, local and hybrid methods.

- Local or analytical facial features approaches. This type consists on applying transformations in specific locations of the image, most frequently around the features points (corners of the eyes, mouth, nose,). They therefore require a prior knowledge of the images...
- Global approaches use the entire surface of the face as a source of information without considering the local characteristics such as eyes, mouth, etc.
- Hybrid methods associate the advantages of global and local methods by combining the detection of geometrical characteristics (or structural) with the extraction of local appearance characteristics.

This article is organized as follows: Basic notions concerning Two-dimensional subspace, wavelet transform theory are provided in Section 2. Fractal codes features are presented in section 3. Feature vectors results from two-dimensional subspaces is applied to a Combined LVQ and PNN classifier are described in Section 4. Section 5 provides face recognition system based on PNN, LVQ, the experimental results and Comparison between the several's types of features obtained using WPD, DWT, IFS, 2DPCA and 2DLDA. A comparison with other approaches is also done in section 6. Conclusion and future works are presented in Section 7.

2.1 Two-dimensional principal component approach analysis (2DPCA)

Proposed by Yang in 2004 [4], 2DPCA is a method of feature extraction and dimensionality reduction based on Principal Component Analysis (PCA) that deals directly with face images as matrices without having to turn them into vectors like as the traditional global approach.

2.2 The steps of face recognition by 2DPCA

Considering training set S of N face images, the idea of this technique is to project a matrix X of size $(n \times m)$ via a linear transformation like that:

$$Y_i = X \cdot R_i \quad (1)$$

Where Y_i is the principal component vector of size $(n - 1)$, and R_i is the base projection vector of size $(m - 1)$. The optimal vector R_i of the projection is obtained by maximizing the total generalized variance criterion

$$J(R) = R^T \cdot G_t \cdot R \quad (2)$$

Where G_t is the covariance matrix of size $(m - m)$ given by:

$$G_t = \frac{1}{M} \sum_{j=1}^M (X_j - \bar{X})^T (X_j - \bar{X}) \quad (3)$$

With X_j : The j^{th} image of the training set

\bar{X} : The average image of all the images in the training set.

$$\bar{X} = \frac{1}{M} \sum_{j=1}^M X_j \quad (4)$$

In general, one optimal projection axis is not enough. We must select a set of projection axes like:

$$\{R_1, R_2, \dots, R_d\} = \arg \max J(R) \quad (5)$$

$$R_i^T \cdot R_j = 0, i \neq j, i, j = 1, \dots, d$$

These axes are the eigenvectors of the covariance matrix corresponding to the largest "d" Eigenvalues. The extraction of characteristics of an image using 2DPCA is as follows

$$Y_k = X \cdot R_k \quad ; k=1 \dots d \quad (6)$$

Where $[R_1, R_2, \dots, R_d]$ is the projection matrix and $[Y_1, Y_2, \dots, Y_d]$ is the features matrix of the image X.

2.3 The 2DLDA approach

In 2004, Li and Yuan [5] have proposed a new two-dimensional LDA approach. The main difference between 2DLDA and the classic LDA is in the data representation model. Classic LDA is based on the analysis of vectors, while the 2DLDA algorithm is based on the analysis of matrices.

2.4 Face recognition using 2D LDA

Let X is a vector of the n-dimensional unitary columns. The main idea of this approach is to project the random image matrix of size $(m \times n)$ on X by the following linear transformation:

$$Y_i = A_j X \quad (7)$$

Y: the m-dimensional feature vector of the projected image A.

Let us suppose L: class numbers.

M: The total number of training images

The training image is represented by a matrix $m \times nA_j (j = 1, \dots, M)$

$\bar{A}_i (i=1 \dots L)$: The mean of all classes

N_i : Number of samples in each class

The optimal vector projection is selected as a matrix with orthonormal columns that maximizes the ratio of the determinant of the dispersion matrix of the projected inter-class images to the determinant of the dispersion matrix of the projected intra-class images;

$$J_{\text{FLD}}(X_{\text{opt}}) = \arg \max_W \frac{|X^T S_b X|}{|X^T S_w X|} \quad (8)$$

$$P_b = \text{trace}(S_b)$$

$$P_w = \text{trace}(S_w)$$

The unitary vector X maximizing J(X) is called the optimal projection axis. The optimal projection is chosen when X_{OPT} maximizes the criterion, as the following equation:

$$X_{\text{OPT}} = \arg \max_X J(X) \quad (9)$$

If S_w is invertible, the solution of optimization is to solve the generalized eigenvalue problem.

$$S_b X_{\text{opt}} = \lambda S_w X_{\text{opt}} \quad (10)$$

Like that λ is the maximum Eigenvalues of $S_w^{-1} S_b$

In general, it is not enough to have only one optimal projection axis. We need to select a set of projection axes, x_1, x_2, \dots, x_d under the following constraints:

$$\{x_1, x_2, \dots, x_d\} = \arg \max_X J(X) \quad (11)$$

Indeed, the optimal projection axes x_1, x_2, \dots, x_d are orthonormal eigenvectors of $S_w^{-1} S_b$ corresponding to the best first "d" eigenvalues permitting to create a new projection matrix X, which is a matrix of size $n \times d$: $X = [x_1, x_2, \dots, x_d]$

We will use the 2DLDA optimal projection vectors x_1, x_2, \dots, x_d to extract the image features; we use the equation (08).

3 Discrete wavelet transform

Discrete wavelet transform (DWT) is a well-known signal processing field tool; it is widely used in feature extraction and compression and de-noising applications. The discrete wavelet transform has been used in various face recognition studies. The main advantage of the wavelet transform over the Fourier transform is the time-scale

location. Mallat [8] shows that the DWT may be implemented using a filters bank including a low-pass filter (PB) and a high-pass filter (PH).

Discrete Wavelet Package Decomposition (D-WPD) is a wavelet transform where signal is passed through more filters than the Discrete Wavelet Transform (DWT). In the DWT, each level is calculated by passing only the previous approximation coefficients through low and high pass filters. However in the D-WPD, both the detail and approximation coefficients are decomposed [7] [8].

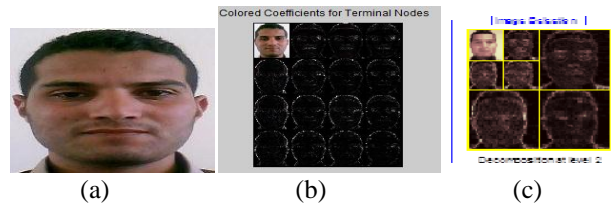


Fig. 1. Wavelet decomposition at different levels
 (a) Original image
 (b) 2-level wavelet decomposition using WPD
 (c) 2-levels wavelet decomposition using DWT

4 Fractal theory codes

Fractal theory of iterated contractive transformation has been used in several areas of image processing and computer vision. In this method, similarity between different parts of an image is used for representing of an image by a set of contractive transforms on the space of images, for which the fixed point is close to the original image. This concept was first proposed by Barnsley [9], [10]. Jacquin was the first to publish an implementation of fractal image coding in [11]. Despite the number of researchers and the proposed methods, several factors can significantly affect face recognition performances, such as the pose, the presence/absence of structural components, facial expressions, occlusion, and illumination variations. Different image compression methods have been focused for a long time to reduce this massive information, but fractal image compression is a relatively recent technique based on representation of an image by contractive transforms, for which the fixed point is close to original image.

Suppose we are dealing with a 64×64 binary image in which each pixel can have one of 256 levels (ranging from black to white). Let R_1, R_2, \dots, R_{256} be 4×4 non-overlapping sub-squares of the image (range blocks); and let D be the collection of all 8×8 pixel overlapping sub-squares of the image (Domain blocks) as depicted in Fig. 2. The collection D contains $57 \times 57 = 3249$ squares. For each R block, search through all of D blocks a $D_i \in D$ which minimizes equation (12). There are 8 ways to map one square onto another. Each square can be rotated to 4 orientations or flipped and rotated into 4 other orientations as shown in Fig. 2 having 8 different affine transformations means comparing $8 \times 3249 = 25992$ domain squares with each of the 256 range squares.

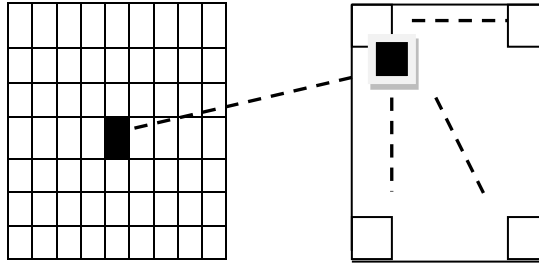


Fig2. .One of the block mapping in partitioned function systems representation (IFS)

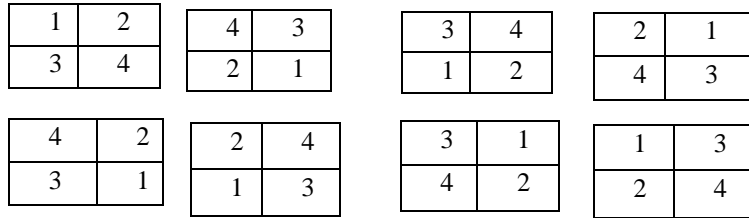


Fig3. Eight different affine transformations

$$collage\ Error = \min \|R_i - w(D_i)\|^2 \quad (12)$$

As mentioned before ,a D_i block has 4 times as many pixels as an R_i ,so we must either sub-sample (choose 1 from each 2*2 sub-square of D_i) or average the 2*2 sub-squares corresponding to each pixel of R when we minimize equation (12) .minimizing equation means two things .First it means finding a good choice for D_i second, it means finding a good contrast and brightness setting S_i and O_i for W_i . In equation (13)

$$w_i \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} a_i & b_i & 0 \\ c_i & d_i & 0 \\ 0 & 0 & s_i \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} c_i \\ f_i \\ o_i \end{bmatrix} \quad (13)$$

A choice of D_i , along with a corresponding S_i and O_i determines a map W_i . The type of image partitioning used for the range blocks can be so different. A wide variety of partitions have been investigated, the majority being composed of rectangular blocks. Different types of range block partitioning were described in[12].In this research we used the simplest possible range partition consists of the size square blocks, that is called fixed size square blocks (FSSB) partitioning. The procedure for finding a fractal model for a given image is called encoding; compression; or searching for a fractal image representation. After finding the best match ,fractal elements which of 6 real numbers (a, b, c, d, e, f) are selected as follows . (a, b, c, d) are (x, y) coordinates of the D block and its corresponding R block respectively . (e) is the index of affine transformation that makes the best match.(it is a number between 1 and 8) , (f) is the intensity is a number between 0 and 256.

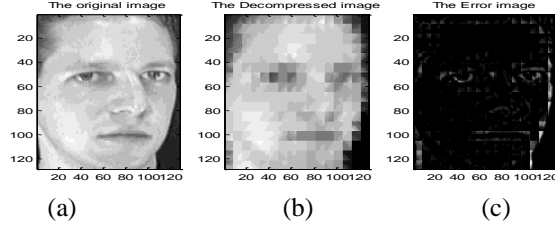


Fig. 4. Decoding algorithm results (IFS)

- (a) Original image
(b) Decoded image after 8 iteration for N=8
(c) The error image

In this paper, fractal code is introduced in order to extract the face features from the normalized face image based WPD. After fractal coding, where each domain is compared with all regions of the image, we obtain a set of transformations which can approximate the face image. Each transformation is represented by parameters of contrast S_i , brightness O_i , spatial coordinates of Range/Domain, and rotation W_i . The output of fractal code is the feature matrix with 2D-dimension used as a database of face which is applied two-dimensional subspace for reduction, discrimination and speed time.

5 Face Classification using neural networks

Several studies have shown improved face recognition systems using a neural classification compared to classification based on Euclidean distance measure [14].

5.1 Probabilistic Neural Networks

The probability neural network is proposed by D. F. Specht for solving the problem of classification in 1988 [15]. The theoretical foundation is developed based on Bayes decision theory, and implemented in feed-forward network architecture.

PNN represent mathematically by the following expression

$$a = \text{radbas}(\|IW - x\|/b) \quad (14)$$

$$y = \text{compet}(LW\alpha) \quad (15)$$

The structure PNN: The PNN architecture consists of two layers [15] [16]:

The first layer computes distances from input vector to the input weights (IW) and produces a vector whose elements indicate how close the input is to the IW.

The second layer sums these contributions for each class of inputs to produce as its net output a vector of probabilities. Finally a compete transfer function on the output of the second layer picks up the maximum of these probabilities and produces a 1 for that class and a 0 for the other classes. The architecture for this system is shown above.

The probability of neural network with back propagation networks in each hidden unit can approximate any continuous non linear function. In this paper, we use the Gaussian function as the activation function:

$$radbas = \exp[-n^2] \quad (16)$$

Finally, one or many larger values are chosen as the output unit that indicates these data points are in the same class via a competition transfer function from the output of summation unit [11], i.e.

$$compet(n) = e_i = [0 \dots 0 \ 1 \ 0 \dots 0], n(I) = MAX(n)$$

5.2 Linear vector quantization (LVQ)

The vector quantization technique was originally evoked by Tuevo Kohonen in the mid 80's [15] [17]. Both Vector quantization network and self organizing maps are based on the Kohonen layer, which is capable of sorting items into appropriate categories of similar objects. Such kinds of networks find their application in classification and segmentation problems.

LVQ network comprises of three layers: Input layer, Competitive layer and Output layer [17]. The number of neurons in each layer depends on the input data and the class of the system. The input neurons are as many as the input matrix features of the training pattern, and the number of the output neurons is equal to the number of person's to which face patterns are classified. The number of hidden neurons is heuristic. In order to implement a face recognition system by our approach, we follow this methodology:

- stage pre-processing using technique WPD
- coding image using fractal code
- feature extraction using 2DPCA/2DLDA
- classification using LVQ and PNN network

6 Results and Discussion

In order to evaluate and test our approach described for face recognition system, we chose three databases: ORL, FEI [18] [19] and our database of our laboratory. All experiences were performed in Matlab installed on a laptop with a dual core processor T5870 with 2.03 GHz and 2 GB of RAM.

To evaluate the performance of our proposed approach, we chose two test databases: ORL and FEI. The global performance of algorithms tested on the FEI database is not as better as that of the ORL database. There are two main reasons:

- The image quality of the ORL database is better than that of the FEI database.
- The FEI database is more complex due to variations in the face details and head orientations.

- After a series of experiences, we chose the best values of parameters in order to fix the choice of Eigenvalues, which give a better recognition rate.

Adding some effects: it is wanted to test our system with and without added noisy in the two data base in order to evaluate robustness of these approaches namely 2DPCA, 2DLDA, DWT,WPD ,fractal codes combined by using two classifier LVQ and PNN.

Noise: Two types of noise are used in this simulation: the Saltand Pepper type noise with a noise density $a=0.06$ (Figure 5 (a)) and Gaussian noise with mean $m=0$, variance $v=0.04$.Figure 5 illustrates these effects which are obtained as follows.



(a)Salt&pepper Noise (b) Gaussian Noise (c) Gaussian Noise $m=0$, $v=0.01$ $m=0$, $v=0.04$



(a)Salt&pepper Noise (b) Gaussian Noise (c) Gaussian Noise $m=0$, $v=0.01$ $m=0$, $v=0.04$

Fig. 5. Adding Noise (database face ORL &FEI)

The pre-processing stage: we proposed to add a preprocessing stage in order to improve our system's performance in speed by reducing the size and eliminating redundant information from the face images by the means of the DWT and WPD technique, and in other hand reduce the memory and compute of our neural network-training algorithm (PNN) and LVQ.

We performed face recognition analysis through WPD and DWT with various wavelet series: Daubechies, Gabor, Coiflets, Symlets and Gauss. In order to select a best wavelet to enhance a rate recognition.

The fractal code is used on WPD and DWT coefficients, derived from WPD to generate detailed high frequency features of animation which forms Feature set one. In order to have fractal feature vectors with the same length, the size of the face must be normalized (32×32). The normalized image is coded by 64 transformations using fractal code. Consequently, we obtained 320 fractal features as each transformation is coded on 5 parameters, as already explained in Section 3. Table 2 shows the performance of our system using fractal features for the two databases.

Feature extraction using 2DPCA/2DLDA: After reducing the dimensional of the face images using IFS. We used the 2DPCA and 2DLDA feature extraction approaches in order to extract the weight images (Features images in the new space) which

must be converted into vectors before implementing the classifier network (LVQ and PNN).

Choice of the number of eigenvalues: Two dimensional methods do not escape this problem, and the choice of the appropriate number depends on the used method and faces database. In our experiences, we have selected the best eigenvalues corresponding to the best variance values (eigenvectors)

Selection parameters and architecture system classifier:

PNN :our neural network training algorithm used in system face recognition is not require many parameters compared other neural networks (MLP,BP,LVQ.etc),that only parameter that is needed for performance of the network is the smoothing parameter σ .Usually, the researchers need to try different σ in a certain range to obtain one that can reach the optimum accuracy[16][17]. To get a higher recognition rate, we have made a series of experiments to choose the best smoothing parameter σ used in PNN. The probabilistic Neural Network used in our system is composed of two layers

Input layer: The first layer is the input layer and the number of hidden unit is the number of independent variables and receives the input data (number of feature extraction for each approach used in this paper)

Output layer: gives the number of faces used in the Database training (ex: ORL 200 person's).

LVQ: The changes of LVQ classifier parameters have a high effect on the classification results. In this paper, we found that the best learning rate increases the recognition rate of the system whereas the learning rate is a critical parameter that affected in the recognition process. We use a different number of learning rate (0.1, 0.2, 0.3, and 0.6) with 500 epochs and 80 hidden Neurons experiments.

Table 1. The recognition rate obtained by different methods on the database ORL with added noisy

Type of classifier	DWT-2DPCA	DWT-2DLDA	WPD-2DPCA	WPD-2DLDA
LVQ	93%	94%	94%	96%
PNN	94.8%	95%	96%	98%

Table 2. The recognition rate obtained by different methods on the database FEI with added noisy

Type of classifier	WPD-2DPCA	WPD-2DLDA	WPD/IFS-2DPCA	WPD/IFS-2DLDA
LVQ	90%	92.8%	95%	96%
PNN	95%	96%	99%	99%

Table 3. The running time (s) obtained by different methods on the database FEI with added noisy

	DWT-2DPCA	WPD-2DLDA	WPD&IFS-2DPCA	WPD&IFS-2DLDA
PNN	1.20	1.25	2.10	2.05
LVQ	1.25	1.45	2.08	1.98

Discussion:

After these series of experiments, we clearly see the superiority of the two-dimensional methods combined with a probabilistic neural classifier combining those of a LVQ classifier (table.1).

In table 2 ,we present the recognition rate obtained when using all fractal features ,and those reduced by the bi-dimensional subspace analysis .There is trade -off between encoding time and average of recognition rate because when N(domain range decreases ,size of features vector will increase so LVQ and PNN learns more details and its generalization ability become weak. As feature extraction is faster for N=8 and average of recognition rate is also fair so we encoded input faces with this R blocks size. The classification results for face is shown in table .2 for N=8.

We also note that the choice of optimal component and the choice smoothing parameter, which represents a better recognition, rate for methods, 2DPCA and 2DLDA and accuracy of classification PNN and LVQ.

In table 3, we present the running time obtained when using fractal codes .computational complexity of fractal encoding is the disadvantage of fractal features in our system which can be improving by adaptive search to speed-up fractal image compression.

7 Conclusion

In this paper, we propose an approach for face recognition based on the combination of two approaches, one used for the reduction of space and feature extractions in two dimensions and the other for classification and decision.

A hybrid approach is introduced in which, through the bi-dimensional subspace analysis, the most discriminating wavelet fractal features are extracted and used as the input of a neural network (LVQ, PNN). The performance of our method is both due to the fidelity of fractal coding for representing images, the WPD algorithm to speed up the features extraction step, and the 2DPCA and 2DLDA which highlights all discriminating features.

As a perspective, we propose to use this approach in an uncontrolled environment (video surveillance) based on video sequences (dynamic images) in order to make the task of face recognition more robust.

8 References

1. Handbook of Biometrics edited by Anil K. Jain Michigan State University, USA Patrick Flynn University of Notre Dame, USA Arun A. Ross West Virginia University, USA © 2008 Springer Science+Business Media, LLC
2. Biometric recognition challenges and opportunities Joseph N. Pato and Lynette I. Millett, Editors Whither Biometrics Committee Computer Science and Telecommunications Board Division on Engineering and Physical Sciences Copyright 2010 by the National Academy of Sciences
3. D. Zhang and Z. -H. Zhou. (2D) 2PCA: Two-directional two-dimensional PCA for efficient face representation and recognition. *Neuro computing*, Vol. 69, pp. 224-231, 2005.
4. N. Nguyen, Wanquan Liu and S. Venkatesh. Random Subspace Two-Dimensional PCA for Face Recognition. Department of Computing, Curtin University of Technology, WA 6845, Australia
5. Two-Dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition Jian Yang, David Zhang, 26, NO. 1, JANUARY 2004
6. S. Noushath, G.H. Kumar, and P. Shivakumara. (2D)LDA: An efficient approach for face recognition. *Pattern Recognition*, 39(7) :1396–1400, 2006
7. S. Mallat. A theory of multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7) :674–693, 1989
8. G.C. Feng, P.C. Yuen, D.Q. Dai. Human face recognition using PCA on wavelet sub-band. *SPIE Journal of Electronic Imaging*, Vol. 9, No.2, pp. 226–233, 2000
9. M. Barnsley, "Fractals Everywhere," Academic Press, San Diego, 1988
10. A. E. Jacquin, "Fractal image coding: A review," *Proc. of the IEEE*, vol. 81, pp. 1451-1465, 1993
11. A. E. Jacquin, "A Fractal Theory of Iterated Markov Operators with Applications to Digital Image Coding," PhD thesis, Georgia Tech, 1989. Y. Fisher, "Fractal Image Compression: Theory and Application," Springer-Verlag Inc., 1995
13. Ebrahimpour-Komleh, Hossein, "Face recognition using fractal codes". In *Proceedings of International Conference on Image Processing 2001*, IEEE, Thessaloniki, Greece.
14. Nazish, 2001. Face recognition using neural networks. *Proc. IEEE INMIC 2001*, pp: 277-281 (2007)
15. D. F. Specht, 1990. "Probabilistic neural network and the polynomial adaline as complementary techniques for classification" *IEEE Trans. Neural Networks*, 1(1): 111-121.
16. *Neural network toolbox matlab User's Guide* COPYRIGHT 1992 - 2002 by The MathWorks, Inc
17. *Computational intelligence paradigms: theory & applications using MATLAB* / S. Sumathi and Surekha Paneerselvam. © 2010 by Taylor and Francis Group
18. ORL. The ORL face database at the AT&T (Olivetti) Research Laboratory (1992)
19. FEI. The FEI face database at the Artificial Intelligence Laboratory of FEI in São Bernardo Campo, São Paulo, Brazil (June 2005 and March 2006)