# Evaluation of Face Recognition Techniques Using 2nd Order Derivative and New Feature Extraction Method Based on Linear Regression Slope

**Abdulbasit Alazzawi†,††, Osman N. Ucan†, Oguz Bayat†**

† Altinbas University, Department of Electrical and Electronics Engineering, Istanbul, Turkey
†† Diyala University, College of Science, Diyala, Iraq

**Summary**
Face recognition system has been widely utilized for various sensitive applications such as Airport gates, special monitoring, and tracking system. The performance of most face recognition systems would significantly decrease if there were several variations in the illumination of dataset images. In this paper the proposed a new algorithm based on a combination of edge detection operators, features extractors and artificial neural network ANN as a classifier. The Second based on Laplacian comprise Zero cross, Laplacian of gaussian LOG, and Canny edge detection filters. A segmentation process is used to segment each image to equaled size blocks treats face edge pixels precisely. A new features extractor technique based on Linear Regression Slope SLP with discrete wavelet transformation (DWT) and principle components analysis PCA used for features extraction. ANN used for the data set classification and all results obtained evaluated. We tried a combination of various techniques like (Zero cross, DWT, SLP-PCA, ANN),(LOG, DWT, SLP-PCA, ANN),(Canny, DWT, SLP-PCA, ANN ). The proposed method is examined and evaluated with different face datasets using ANN classifier. The experimental results were displaying the superiority of the proposed algorithm over the algorithms that used the state-of-art techniques where the combinations (Zero cross, SLP, ANN) gained the best results and could outperform all the other algorithms.

***Key words:***
*Face Recognition, SLP, PCA, Neural Network, ANN*

## 1. Introduction

Since from last decade, face recognition system has had vital effects in daily life, especially for security objectives. The system is the strong domain for human being authentication process in various authentication applications. The systems have been widely utilized for various sensitive applications such as Airport gates, special monitoring, and tracking system, criminal distinguishing, broad checkpoints and many other applications. Face recognition used for identity identification or verification process. Face identification process is nothing except one to one matching problem. Meanwhile, face verification is more complicated because it is one to many matching problems. It is utilized to identify the tested face against huge face images that saved in the database. Many approaches, and techniques used in face recognition process. Generally, we can divide it into two categories. First, one is to data reduction and features extraction such as holistic methods that include, eigenvectors based on PC [1], independent component analysis technique (ICA) [2], linear discriminant analysis (LDA) technique [3], and kernel LDA [4].The second one is utilized as a classifier to find the faces features which are most likely to be looking for such as Neural Network-based approaches [5] Support vector machine, and nearest distance [6]. Most of these approaches and techniques have trade-offs such as time for features extractions, the response time for training dataset, lack to update training dataset and hardware requirements.

In this study, we chose a gradient, Laplacian-based operators for edges detection, principal components analysis PCA, slope based method SLP for feature extraction, and ANN for a classification.

The edges are the crucial property that can be used for data reduction and feature detection. Data reduction comes from excluding all image data except edge. These data can be used to find face objects features. Many operators used for edge detection. Generally, we used second-order derivative operators that includes Zero cross, LOG and Canny filters based on Laplacian operator. All these filters have weaknesses during edge detection process such as weakness in noise depressing, find out some type of edge but not all and lack edge shape. Therefore, we connect those edge detection filters with discrete wavelet transformation DWT for edge getting optimization.

DWT is very well known in image decomposition by separation original data set image to approximation image and three detail images (horizontal, diagonal, and vertical image. Wavelet transformation used for dimensionality reduction. Also, it is used for time-space frequency analysis. Wavelet transformation provides time-frequency analysis for one or two-dimensional signal which is practically powerful in image analysis, computer vision,

Manuscript received March 5, 2018
Manuscript revised March 20, 2018

pattern recognition. Unlike Fourier transformation that used only for time analysis signals [7].

PCA is powerful and long-term studies technique applied for feature extraction from face images by creating a face feature space. PCA has another advantage that is low computation time because it depends on modified covariance matrix to find Eigenfaces. Whoever, because of being linear features extraction, PCA is less effective especially when nonlinearity issue take place in the underlying relationships [8].

Slope base method SLP is a new method used for features extraction and became dependable in IEEE based on [9]. The algorithm of SLP depends on linear regression slope.

The neural network is well known parallel processing units (neurons) container. The main challenges of the neural network are training dataset such that output same efficient rate in the testing phase (generalization).

The experimental analysis of proposed framework is done using MATLAB image processing toolbox with three different face datasets such as [10-13]. This dataset prepared under the different conditions and having various image qualities, illumination conditions, lightning effects, occlusions etc. Also, this study, we used 23 subjects classes in the face BIO ID database [14].

The sections of the paper are organized as follows: in sections 2, we present edge detection. Section 3 presents features extraction methods (DWT, SLP, PCA). These methods are discrete wavelet transformation DWT, eigenfaces based on principal components analysis PCA and slope based method SLP. Section3 introduces segmentation process based on divide and conquer principle. In section4, we introduce neural network as a classification method. Section 5 shows the experimental results and section 6 produces conclusions.

## 2. Methodology

The following diagram explains the steps of the proposed system .as we can see, the procedure of process starts with the select dataset for training ANN. Whoever, before reach to last step (training ANN) we performed some steps, these steps include applying DWT on the training data set images and select a LOW LOW image. performing edge detection operators on the approximating image is the second step. Many edge detection operators were used since each of which covers some types of edges, therefore we used multi filters to cover all the types of edges. segmentation process is used to segment the image that came from edge detection process. Each image is converted to many eaqeulad size blocks. Then, form these blocks, features are extracted using new suggested method with SLP name that depends on linear regression slope .to measure the efficiency of the SLP state of art are used, eigenface based on PCA. As a classifier, efficient Artifice neural network is used.

## 3. Edge Detection Methods

### 3.1 Zeroross Operator

The zero-crossing edge detector figures out zero-crossing regions in images where the values of the Laplacian cross through zero regions. According to this operator, the edge point pixels are mostly represented by points that have a value of the Laplacian passes through zero which means that in these points the Laplacian is changing its sign. For example, in face recognition algorithms, the points where the intensity of the face image edges changes suddenly is considered as edge points candidates. The scientists recommend dealing with this filter as feature extractor operator since the output of the zero-crossing algorithm is usually a binary image with one-pixel thickness [15-18].

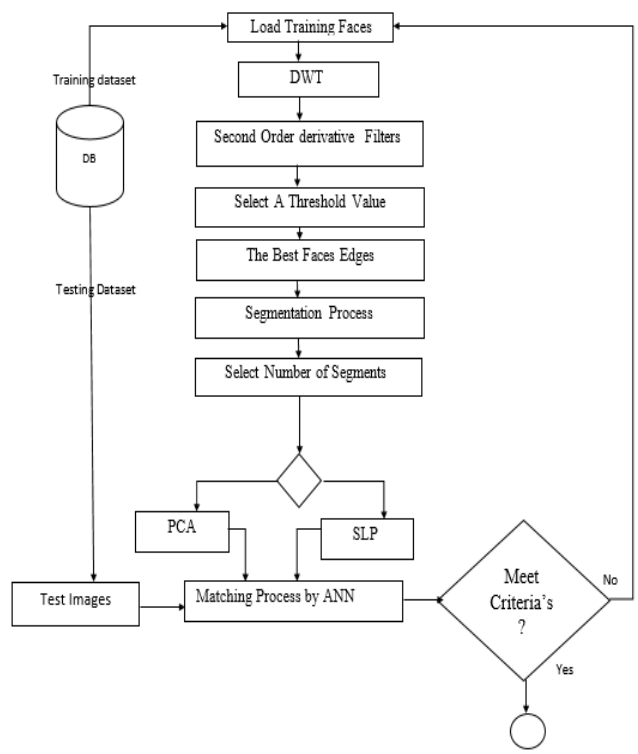


Fig.1 Main Diagram of the Proposed System

### 3.2 LOG Operator

The Laplacian of Gaussian (LOG) algorithm is proposed to solve the noise sensitivity of the regular Laplacian filter. The steps of this algorithm start with removing noise by applying image smoothing using a Gaussian blurring technique to obtain the best performance. After applying the Laplacian method, change in sign from positive to negative or vice-versa for any point (changes in sign) represents an edge. This algorithm supposes that important cross points almost be at edges in the images - i.e. where the intensity values of the images regions changes sharply. However, they also could occur at places that are not as easy to allocate with image edges because some of these regions are very like each other, especially in grayscale images. The lines that are detected by the LOG procedure have one-pixel thickness. The Laplacian L (x, y) of an image with pixel intensity values I (x, y) is given by:

$$L(x,y) = \left(\frac{\partial I^2}{\partial x^2}\right) + \left(\frac{\partial I^2}{\partial y^2}\right) \quad (1)$$

This an important pre-processing operation decreases the unwanted frequency noise prior to the differentiation operation. The2-D$ LOG function centered on zero and with Gaussian standard deviation has the form:

$$\text{LOG (x, y)} =(-\frac{1}{\pi\sigma^4} *((1- \frac{x^2}{y^2})* (\text{e} \frac{x^2}{2\sigma^2} )) ) \quad (2)$$

the Gaussian smoothing filter and the Laplacian filter f and then convolve this output result (hybrid filter) with the digital image to reach the desired output [18].

### 3.3 Canny Operator

Canny edge detection operator uses five different and sequential procedures to detect the edges; the first step is smoothing the image to remove possible noise using a Gaussian filter. The second step is computing gradient values in vertical (Gx) and horizontal (Gy) directions which is the same procedure that used in the Sobel algorithm. Canny filter eliminates pixels that are not part of image edges by using a non-maximum suppression process. In the fourth step, threshold values are used to determine the potential edges. In the last step of the Canny algorithm, all edge pixels that have weak value or isolated from other pixels are eliminated [19],[20]. For more explanation, the mathematical model of the canny filter includes the following equations.

$$G\ (m.n) = G\sigma\ (m, n) * f\ (m, n) \quad (3)$$

Where

$$G^{\sigma}=\frac{1}{\sqrt{2\pi\sigma^2}}\ \text{e-}\ (\frac{m^{2+}\ n^2}{2\pi\sigma^2}\ ) \quad (4)$$

Compute gradient of g (m, n) using any of the gradient operators to get the magnitude(M) and the direction ($\theta$) by

$$M(n,n) = \sqrt{g_m^2(m,n) + g_n^2(m,n)} \quad (5)$$

and

$$\theta(m,n) = tan^{-1}\ \frac{g_{n(m,n)}}{g(m,n)} \quad (6)$$

Where threshold is chosen so that all edge elements are kept while most of the noise is suppressed.

## 4. Segmentation

One of the basic steps in the proposed system is the segmentation of the image to equal size blocks. The idea behind dividing segment image to sub-images is inspired by the problem-solving technique. A popular technique that used in advance pattern recognition algorithm and we benefit from it is a divide and conquer technique [21] [22]. The dividing problem to many subproblems can help to solve complex problems such as pattern recognition. In this study, we divide the image into many sub-images for two reasons. One is to find the edge values positions precisely. Finding positions of the edge pixels is the basic step for using SLP features extraction method. Second is to separate noise pixels from edge pixels. Both of them we tried are to simplify feature extraction process. The number of segments parameter is a free number because there is no standard value of the number of segments. I.e., we tried many scenarios (it is related to other parameters) to select an optimal number of segments.

## 5. Features Extraction Methods

### 5.1 DWT Method

DWT is designed in a way such that we got good frequency resolution for low-frequency components.2D face signal translated into shifted and scaled levels according to these two operations we obtained 4 sub signals (sub-bands) viz; low -low LL, low high LH, high-low HL, and high-high HH. In the proposal system, we applied the multi-level discrete Wavelet transformation. Low pass filter applied on row and columns of the image to get LL subband. Low Low or approximation(A) presents smaller scaled from the input image. Low pass filter and high pass filter applied on the image to get the three sub-bands, LH, HL, and HH. The most important subband is low low(LL). We can construct all image from this subband since this sub-signal shows a general trend of pixel values. Also, if the details are small, we can ignore it [23] [24] [25]. representation of this process on the image as follows:

$$\text{I (i, j) = (LL) + (LH + HL + HH)} \quad (7)$$

Or:

$$I(i,j) = I_{A_1}^n + (I_{D_1}^n\ +\ I_{H_1}^n + I_{V_1}^n) \quad (8)$$

this formal used for (n) times DWT decompositions. In our study, we used:

n=21:23db (9)

-db is a Daubechies DWT.

we represented DWT decomposition levels that used in this algorithm like:

$$I_1(i,j) = I_{A_1}^1 + \left(I_{D_1}^1 + I_{H_1}^1 + I_{V_1}^1\right) \quad (10)$$

$$I_2(i,j) = I_{A_1}^2 + \left(I_{D_1}^2 + I_{H_1}^2 + I_{V_1}^2\right) \quad (11)$$

$$I_3(i,j) = I_{A_1}^4 + \left(I_{D_1}^4 + I_{H_1}^4 + I_{V_1}^4\right) \quad (12)$$

$$I_4(i,j) = I_{A_1}^8 + \left(I_{D_1}^8 + I_{H_1}^8 + I_{V_1}^8\right) \quad (13)$$

## 5.2 Eigenface Method

An Eigenface-finding based on Principal Component Analysis technique (PCA was also known as Karhunen Love) is specialized for face image data. In principal component analysis method, all dataset images are recognized with feature vectors that are created from projections of the training image to the basis in image space. In general, PCA classifies dataset images according to calculated distances among feature vectors. Typical classifiers comprise nearest distance measure, Euclidean distance, and nearest mean classification. Using PCA for the Eigenfaces method, feature vectors identifying each image is shown in figure (5), and these vectors can be obtained as follows:

(1) Assume that we have (n) faces with (m) rows and (m) column, as S = S1, S2,S...S T (14)
Represent the images in column vector space with ($m^2$ * 1) dimensions. According to the PCA algorithm, we must calculate a mean face image (Q) (common features for all dataset), from vectors by the following Eq.

$$Q = \frac{1}{T} \sum_{n=1}^{T} S_n \quad (15)$$

Where" Q" is the mean matrix (common properties of faces).

(2) The second step is to subtract the main data matrix from ($S_n$), which can be represented as:

$$L_n = S_n - Q \quad (16)$$

(3) Ln column vectors are gathered in the matrix:
R = [$L_1$, $L_2$, . . ., $L_N$] with dimension (m2 * N) and a covariance matrix C is formed as

$$C = \frac{1}{T} \sum_{n=1}^{T} L_n\, L_n^T = R.R^T \quad (17)$$

Because of high dimensionality and (considering the linear algebra principle) the great computational complexity of A AT multiplication, we use ATA instead.

$$C^* = \frac{1}{T} \sum_{n=1}^{T} L_n^T\, L_n \quad , C* = R^T.R \quad (18)$$

(4) At this step, we calculate N eigenvalues (λp) and N eigenvectors (VP) of C∗ to form the eigenface space. V = [$v_1$, $v_2$, . . ., $v_N$] represents a matrix including eigenvectors of (C) with a dimension of (N * N). We can obtain eigenface space U= $[u_1,u_2...,u_N]^T$ by

$$U = V.R^T \quad (19)$$

All row vectors of (U) are eigenfaces of face images in the training set. Face images with higher eigenvalues have a greater contribution to the eigenface space. For this reason, systems with low computational capability sort eigenvectors of face images according to their corresponding eigenvalues in decreasing order and choose eigenvectors to form a smaller eigenface space.

$$W = w_1, w_2, w_3..., w_N \quad (20)$$

Matrix with dimension (N*N) includes (N) column vectors corresponding to each face image in the training set. These vectors are called feature vectors and they represent each image specific characteristics. (W) can be obtained

$$W = U. R \quad (21)$$

After obtaining eigenface space and feature ectors, we can compare a test image with the faces in the training set by projecting a test image into faces space as follows:

(a) "$S_T$" is a column vector which represents our test image with ($m^2$ ∗1) dimension. At this stage, the distance of test image from mean face image should be calculated as LT column - vector,

$$L_T = S_T - Q \quad (22)$$

(b) After calculating $L_T$, we must project it on our eigenface space in order to obtain its feature vectors in the format of the column vector $w_T$ with dimension (N *1),

$$w_T = U.L_T \quad (23)$$

(c) find out which image in the training set resembles our test image, we need to find the similarity of $w_T$ to each ($w_i$) in the matrix (W). Various classifiers can be used at this stage. The technique we used is the artificial neural network. [26-29].

## 5.3 Simple Linear Regression Slope-Based Method

To efficiently extract the best features from edge images, a new feature extraction method that based on the slope of the estimated curve of each small segment is used. The proposed slope based feature extraction method (SLP) produces vectors of features that can identify each face successfully as shown in figure (6). The following steps explain in detail the process of extracting the features in the training phase:

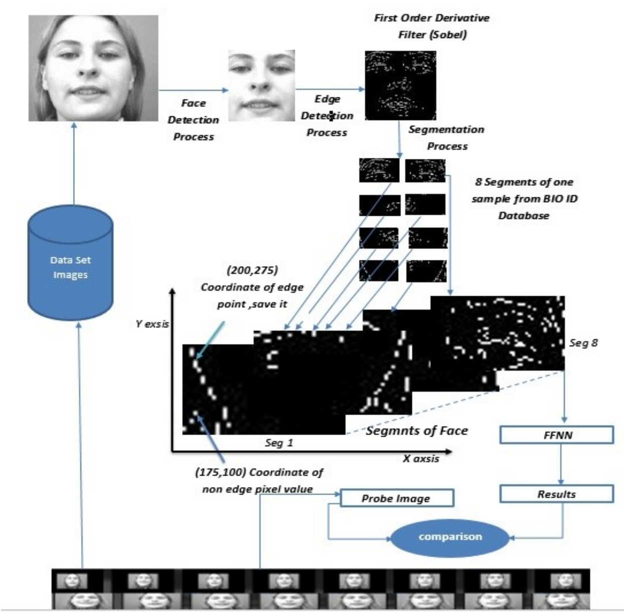


Fig. 2 SLP feature Extraction Method

Suppose that we have N face images (im1, im2, im3...imN) with m rows and n columns, where N is the number of images in the training set.

(1) Convert the face images ($im_1$, $im_2$, $im_3$...$im_N$) to binary images using one of the six edge detection filters that discussed in section 2. The binary images denoted as ($B_1$, $B_2$, $B_3$...$B_N$).
(2) For each binary image Bi, divide it into ($H * L$) segments, where H is the number of rows and L is the number of columns. The segments of image face denoted by ($Si_1$, $Si_2$...$Si_k$), where k is the number of segments with ($H * L$) dimension.
(3) For each segment Sij of each face image Bi, the value of pixels with ones are considered as points and a simple linear regression estimation method is used to find the best line that fit these points. Assume that the equation of the line is as
$$Y = m * x + b \quad (24)$$
Where b is the Y-intercept and m is the slope of the estimated line.
(4) Compute the slope of the estimated line of each segment Sij using the following equation which is derived from Pearson's correlation coefficient formula and simple linear regression:

$$m = \frac{(n\sum(xy) - \sum(x)\sum(y))}{(n\sum(x^2)) - (\sum x)^2} \quad (25)$$

Where m is the slope of Sij segment, x is the x-coordinates, y is the y-coordinates and n is the length of x, y vector in the binary image. In our work, we select the linear regression techniques because it is the most widely used statistical technique to model a relationship between two sets of variables and it is widely used to estimate the line equation from a set of points. Slopes of segments are collected in a matrix with dimension ($H * L$) corresponding to each face image in the training set to create features vectors which will be the input to the ANN. The usage of the simple linear regression techniques in our steps makes the proposed algorithm works very well since it generates very robust features that are not affecting by outlier points or noise points that may exist in the binary edge images.

## 6. Classification Process

The last step in the proposed system is training neural network on the face images dataset. The input matrix of the training dataset consists of the features vectors after we applied all the preprocessing operations, dimensionality reduction, and feature extraction. Steps of this algorithm are explained as follows.

(1) First, prepare training data set.
(2) calculate the weighted sum of the input vectors (dot product between the input vector and weights that connect the input and the hidden layers) [30].
$$NT_g = \sum_{k=1}^{n} X_k * W_{gk} \quad (26)$$
(3) The results that we obtain from the weighted sum calculation is passed to the activation function to normalize the output. In the proposed MLNN layers we used the sigmoid function.
$$H_g = \left(\frac{1}{1+\exp(NT_g)}\right) \quad (27)$$
(4) In the same way, we calculated the output of the final layer by calculation of the weighted sum between the hidden layer and an output layer.
$$out_v = \sum_{l=1}^{m} H_l * W_{lv} \quad (28)$$
$$TrueNOut = \left(\frac{1}{1+exp(-Out_{gv})}\right) \quad (29)$$
(5) After the output has been calculated, we calculate the differences between actual output and the desired output by using the minimum square error between them.
(6) $\text{error} = \sum_{i=1}^{V} TrueNOut_i * WD_i$ (30)
$$\Delta Wl_i = \alpha \delta_l H_l \quad (31)$$
$$\delta_l = (D_i - TrueNOut_i) * TrueNOut_i * (H_l - TrueNOut_i) \quad (32)$$

## 7. Results and discussions

We performed face recognition based on two groups of edge detection filters gradient and Laplacian-based operators, each of which includes three filters. The second order derivative operates includes LOG, Zero cross, and Canny Filters. A common feature of the second group is, all of them perform Laplacian process before they are starting edge detection process. For the features extraction, we used the eigenvectors-based method and suggested a method, SLP. The ANN is the classifier that used for the classification process. For DWT, PCA and SLP feature extraction methods a total of four training set were composed that include varying illumination, expression and pose factors for each person and remain factors are chosen as the test set. The rate of training and testing are:

Case 1: 15% of the dataset for testing and 85% for training.

Case 2:20% of the dataset for testing and 80% for training.

Case 3: 30% of the dataset for testing and 70% for training.

Case 4:40% of the dataset for testing and 60% for training.

According to experiments that included these edge detection filters, eigenfaces based on PCA technique, SLP method and ANN we registered the results. In the following paragraphs, we are showing it:

### 7.1 Laplacian Based Operators

The entire Laplacian Based Operators algorithms are designed to depress noise in the first step, and then extract images edges. This producer is best when we work on segments of images instead of the image to save fractions details. Those second order filters can save segments details more than first order operators. Zero crosses (ZF) using Laplacian function to discover the edges points. Where the Laplacian value crosses zero region, it signs this regain as an edge. Generally, zero crosses know as general edge detection class that includes log filter and another. Mathematically,

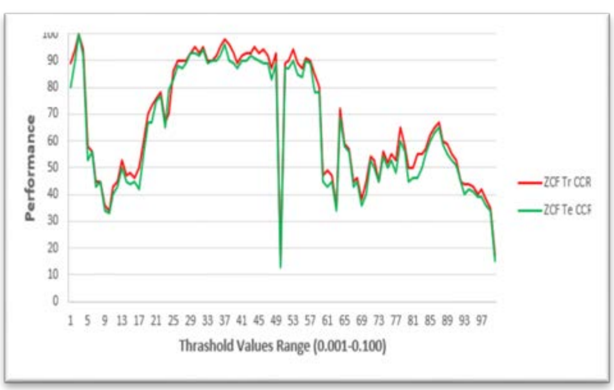


Fig. 3 Performance of Proposed System based on Zero cross

Laplacian of Gaussian defines the best kernel of edge detection by measuring the ratio of the signal to noise of the image. Edge detection by the Laplacian of an image means taking its double derivative process in both directions, horizontally and vertically. And because of it is a member of second-order derivative filters, it has a stronger response to smooth details. Consequently, well detection operators when we are using segmentation process. Table(2) includes classification results based on LOG, SLP, PCA, and ANN.

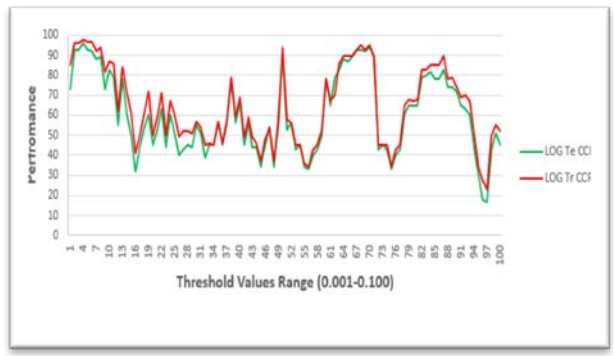


Fig. 4 Performance of Proposed System based on LOG, DWT, and SLP

Canny filter (CF)includes three processes. It is performing noise depressing process firstly by using the Gaussian function. Then, perform a process to detect the top value of first derivative operation also corresponds to the minimum of the second derivative by using threshold value. The power of canny filter algorithm comes from this point. This step enables canny algorithm to detect weak and strong edge points in all directions because it is not susceptible to noise as compared with other edge detection operators. Table (2) illustrate the best results according to experimental tests.

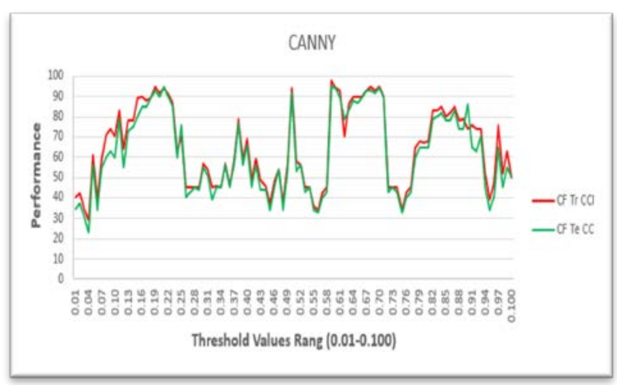


Fig.5 Performance of Proposed System based on Canny, DWT, and SLP

Daubechies wavelets were performed with level (21 -23) decomposition for getting the optimal choice to create feature vectors. Thus, we can see the differences between the algorithms (as it is explained in table 1 2 and 3) and make a fair comparison among them. Each of dataset images was converted to four sub-bands LL, LH, HL, and HH. The LL image represents the main image (approximate image) and it includes sufficient details that can be used to reconstruct the original image and ignores all three subbands, for data reduction in our proposed algorithm.

Table 1: ZF-Zeroross Filter, DWT-Discrete Wavelet Transformation SLP-Slope Based Method, ANN

| TECHNIQUE | TRAINING DB. (%) | T.C.R. (Training Classification Rate) | T.C.R. (Test Classification Rate) |
|---|---|---|---|
| ZF-DWT-SLP-ANN | 0.85 | 97.6534 | 96.6667 |
| ZF-DWT-SLP-ANN | 0.8 | 98.6564 | 93.5484 |
| ZF-DWT-SLP-ANN | 0.7 | 98.6784 | 92.5373 |
| ZF-DWT-SLP-ANN | 0.6 | 98.9744 | 92.2680 |
| ZF-DWT-SLP-ANN | 0.5 | 99.0854 | 90.6736 |

Table 2: LF-LOG Filter, DWT-Discrete Wavelet Transformation SLP-Slope Based Method, ANN

| TECHNIQUE | TRAINING DB. (%) | T.C.R. (Training Classification Rate) | T.C.R. (Test Classification Rate) |
|---|---|---|---|
| LF-DWT-SLP-ANN | 0.85 | 97.6534 | 96.6667 |
| LF-DWT-SLP-ANN | 0.8 | 98.6564 | 93.5484 |
| LF-DWT-SLP-ANN | 0.7 | 98.6784 | 92.5373 |
| LF-DWT-SLP-ANN | 0.6 | 98.9744 | 92.2680 |
| LF-DWT-SLP-ANN | 0.5 | 99.0854 | 90.6736 |

Table 3: CF-Canny Filter, DWT-Discrete Wavelet Transformation SLP-Slope Based Method, ANN

| TECHNIQUE | TRAINING DB. (%) | T.C.R. (Test Classification rate) | T.C.R. (Test Classification Rate) |
|---|---|---|---|
| CF-DWT-SLP-ANN | 0.85 | 97.6534 | 96.6667 |
| CF-DWT-SLP-ANN | 0.8 | 98.6564 | 93.5484 |
| CF-DWT-SLP-AN | 0.7 | 98.6784 | 92.5373 |
| CF-DWT-SLP-ANN | 0.6 | 98.9744 | 92.2680 |
| CF-DWT-SLP-ANN | 0.5 | 99.0854 | 90.6736 |

## PCA Method

The principal components analysis is a very powerful tool in feature extraction field, especially in the face recognition. It is wildly used in the face recognition technology for feature extraction and dimensionality reduction. We used it for same purposes, directionality redaction addition to feature extraction with different edge images. Applied 6 edge detection filters with PCA increasing the area of study to reach to the optimal pre-processing and features extractors combination. This cognition made the training of data set an efficient. BIO ID database is very complex database because of the similarity between foreground and background data. The PCA method that used in the proposed system has shown good result despite the change in light conditions, face expression, and complex background of the database Images.The idea of this method is very simple and efficient. According to the edge detection filters on the dataset images, the process of using edge point positions instead of edge values is the optimal solution to find face template. This is first and an important operation instead of using face detection methods that need more time, efforts, and storage. The output of edge detection filters is points or scatters with (0,1) value. Consequently, from these output points, we are visualizing faces by saving only 1s values. This process is data reduction process because of excluding background values from all dataset images. Instead of using these edge points values we used positions (x,y)of points values. Proposal system created positions matrix of pixels values to calculate the slope between every two points and save the results.

## 8. conclusion

we applied slope method as features vectors. This method is robust against environmental factors like light condition, complex background. Also, it doesn’t need complex normalization process such as dimensionality reduction. We applied this feature extractors method with same edge detection filters to measure the efficiency of the system from PCA, SLP techniques view. This paper found the importance of the relation between pre-processing operation, features extractors output vectors that used as input to the neural network. According to the results of proposal system in testing level, we registered correct classification rate practically. The inputs of ANN consist of features vectors that extracted by SLP and PCA techniques. A number of neurons that has critical effects in the hidden layer is an open parameter in the proposed ANN (it is changeable according to the differences between actual output and the desired output). The sigmoid function used in the hidden layer. A range of neural network cycle(1-10000 epochs )is to measure the rate of training, testing errors in each experiment. We tested 3000 experiments in this paper with different meta parameters like

- Many edge detection methods.
- The range of threshold values (0.01-0.9).

- No of hidden units (25,50,75,100,125,150) .
- several segments scenario (for example 4 ,6,8,10) .

The training rate of data set to test data set was:

- 85 % from the dataset for training 15% for testing.
- 80 % for training data set 20% for testing.
- 70 % for training data set 30% for testing.
- 60 % for training data set 40% for testing.

As a future work, we are studying to try various discrete wavelet transformation (DWT) with gradient, Laplacian filters to improve the data reduction process and for performance enhancement. Also, a different type of neural network like convolution neural network can take place in the new proposed algorithm. Our results indicate that better recognition rates are obtained with SLP method. This method can be used for features extraction in another field as well.

## References


[1] Turk, M. a., & Pentland, A. P. (1991). Face Recognition Using Eigenfaces. Journal of Cognitive Neuroscience https://doi.org/10.1109/CVPR.1991.139758

[2] Bartlett, M. S. (2001). Independent Component Representations for Face Recognition. Face Image Analysis by Unsupervised Learning, 3967. https://doi.org/ 10.1007/978-1-4615-1637-8-3

[3] Belhumeur, P. N., Hespanha, J. P., & Krieg man, D. J. (1997). Eigenfaces vs. fisher face: Recognition using class specific linear projection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7), 711720. https://doi.org/10.1109/34.598228

[4] Lu, J., Member, S., Platinoids, K. N., & Venetsanopoulos, A. N. (2003). Face Recognition Using Kernel Direct Discriminant Analysis Algorithms. Analysis, 14(1), 117126. http://doi.org/10.1109/ TNN. 2002.806629

[5] Lawrence, S., Giles, C. L., Ah Chung Tsoi, & Back, A. D. (1997). Face recognition: a convolutional neural-network approach. IEEE Transactions on Neural Networks, 8(1), 98113. https://doi.org/ 10.1109/ 72. 554195

[6] Karatzoglou, A., Meyer, D., & Hornik, K. (2006). Support Vector Algorithm in R. Journal of Statistical Software, 15(9), 128.

[7] DISCRETE WAVELET TRANSFORMS ̄ THEORY AND APPLICATIONS EDITED BY JUUSO OLKKONEN. (N.D.).

[8] UrÅun, O., & Favorov, O. V. (2004). SINBAD automation of scientific discovery: From factor analysis to theory synthesis. Natural Computing, 3(2), 207233. https://doi.org/ 10.1023 /B: NACO .0000027756 .50327.26

[9] Alazzawi, Osman, Bayat, face recognition based on multi features extractors" IEEE, (2017).

[10] http://vis-www.cs.umass.edu/lfw/

[11] http://www.pitt.edu/ emotion/ck-spread.htm

[12] http://www.jdl.ac.cn/peal/

[13] http://www.kasrl.org/jaffe.html

[14] BioID-Face-Database @ www.bioid.com. (n.d.). Retrieved from https:// www.bioid.com /About/ BioID-Face-Database

[15] Gonzalez, R., & Woods, R. (2002). Digital image processing. Prentice Hall. https://doi.org/10.1016 /0734-189X(90)90171-Q.

[16] Chaple, G. N., Daruwala, R. D., & Gonane, M. S. (2015). Comparisons of Robert, Prewitt, Sobel operator based edge detection methods for real time use on FPGA. Technologies for Sustainable Development (ICTSD), 2015 International Conference on, (1), 47. https://doi.org/10 .1109/ ICTSD.2015.7095920

[17] Maini, R., & Aggarwal, H. (2009). Study and comparison of various image edge detection techniques. International Journal of Image Processing, 3(1), 111. https://doi.org/ http://www. doaj.org /doajfunc-openurl&genre-article&issn-19852304&date-2009&volume-3&issue-1&spage-1

[18] Gllillon, s., & donias, m. (n.d.). directional second order derivatives: application to edge and corner detection i :, 2(3), 3–6.

[19] Wei-Bo, Y., & Zun-Sheng, D. (2011). An improved Kirsch human face image edge-detection method based on a canny algorithm. 2011 International Conference on Consumer Electronics, Communications and Networks,

[20] Spontón, H., & Cardelino, J. (2015). A Review of Classic Edge Detectors. Image Processing on Line, 5, 90–123. https://doi.org /10 .5201/ipol.2015.35

[21] Revisited, S., & Sorting, P. (2004). Introduction to Algorithms Part 1 : Divide and Conquer Sorting and Searching. Sort.

[22] Edelsbrunner, H., Assistant, T., & Gu, Z. (2008). Design and Analysis.

[23] Barford, L. A., Fazio, R. S., & Smith, D. R. (1992). An introduction to wavelets. Hewlett-Packard Labs, Bristol, UK, Tech. Rep. HPL-92-124, 2, 129. https://doi.org/10.1109/99.388960

[24] Yamamoto, A., & T. L. Lee, D. (1994). Wavelet Analysis: Theory and Applications. Hewlett-Packard Journal, (December), 44–52. https://doi.org/ 10. 1051/ jp1:1997114

[25] Discrete, T. H. E., & Transform, W. (n.d.). No Title, 1–15.

[26] Alorf, A. A. (2016). Performance evaluation of the PCA versus improved PCA (IPCA) in image compression, and in face detection and recognition. 2016 Future Technologies Conference (FTC), (December), 537–546. https://doi.org/10.1109/ FTC .2016 .7821659

[27] Kang, J., Lin, X., & Yang, G. (2015). Research of Multi-Scale PCA Algorithm for Face Recognition.

[28] Kumar, D. S. D., & Rao, P. V. (2015). Analysis and Design of Principal Component Analysis and Hidden Markov Model for Face Recognition. Procedia Materials Science, 10(Cnt 2014), 616–625. https://doi.org/10.1016/j.mspro.2015.06.014.

[29] Liu, K., & Moon, S. (2016). Robust dual-stage face recognition method using PCA and high-dimensional-LBP. 2016 IEEE International Conference on Information and Automation (ICIA), (August), 1828–1831. https://doi.org/10.1109/ ICInfA.2016 .7832115.

[30] Clarkson, T. G. (1996). Introduction to neural networks. Neural Network World (Vol. 6). https://doi.org/10.1016/S0140-6736(95)91746-2

**Abdulbasit K.S. Alazzwi** Completed M. Sc. from Iraqi Committee of Computes and Informatics in computer science. He is currently pursuing the Ph.D. degree at the electronic and computer engineering institute-Altinbas university. His research interests in the pattern recognition, image processing, and machine learning.

**Osman N. UCAN** received the B.SC., M.SC. and Ph. Degrees from Technique Istanbul University(IT), Turkey 1985,1988 and 1995, in respectively. he has published more than 250 papers in different fields. Now he is a Professor, Doctorate Supervisor and the Dean of the faculty of engineering, Altinbas university. His main research interests include biomedical, image processing and machine learning.

**Oguz Bayat** received the B.SC., degree from Istanbul Technique University(IT), Turkey in 2000.he has got M.Sc. from University of Hartford, Electrical Engineering, 2002. completed the Ph.D. in the North-eastern University, Electrical Engineering, 2006.extra, he has got CP from Massachusetts Institute of Technology, Management, and Leadership, 2009. he has published more many papers in different fields. Now he is assisting. Professor, Doctorate Supervisor and the Dean of the graduate institute, Altinbas university. His main research interests include signal processing, communication, and machine learning