

## An Efficient Implementation of Weighted Fuzzy Fisherface Algorithm for Face Recognition using Wavelet Transform

<sup>1</sup>Esther Annlin Kala James and <sup>2</sup>S. Annadurai

<sup>1</sup>Faculty of ECE, Government College of Technology,  
Coimbatore, Tamilnadu, India

<sup>2</sup>Sri Krishna College of Engineering and Technology,  
Coimbatore, Tamilnadu, India

---

**Abstract: Problem statement:** The paper addresses the face recognition problem by proposing Weighted Fuzzy Fisherface (WFF) technique using biorthogonal transformation. The weighted fuzzy fisherface technique was an extension of Fisher Face technique by introducing fuzzy class membership to each training sample in calculating the scatter matrices. **Approach:** In weighted fuzzy fisherface method, weight emphasizes classes that were close together and deemphasizes the classes that are far away from each other. **Results:** The proposed method was more advantageous for the classification task and its accuracy was improved. Also with the performance measures False Acceptance Rate (FAR), False Rejection Rate (FRR) and Equal Error Rate (EER) were calculated. **Conclusion:** Weighted fuzzy fisherface algorithm using wavelet transform can effectively and efficiently used for face recognition and its accuracy is improved.

**Key words:** Linear Discriminant Analysis (LDA), Principal Component Analysis (PCA), Fisher Linear Discriminant (FLD), False Acceptance Rate (FAR), Equal Error Rate (ERR), Crossover Error Rate (CER), Equal Error Rate (ERR)

---

### INTRODUCTION

Face recognition involves two major steps. In the first step, some features of the image are extracted. In the second step, on the basis of the extracted features the classification is performed. There can be various features that can be extracted from the facial images.

In the field of pattern recognition, feature extraction for dimensionality reduction is an important topic of research, because in many practical technologies high dimensionality is a major cause of limitation. Also the large quantities of features, degrade the performances of the classifiers, when the size of the training set is small compared to the number of features. In the past several decades, many face recognition methods have been proposed, in which the most well-known methods are Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). Fisherface method outperforms the eigenface (Belhumeur *et al.*, 1997; Zhuang and Dai, 2007; Yuen *et al.*, 2009; Rizon *et al.*, 2006; Abusham *et al.*, 2008; Ma *et al.*, 2006; Chan *et al.*, 2010) method in case of large variation of lighting conditions, different face poses and different facial expressions. The basic idea of Fisher Linear Discriminant (FLD) is to calculate the fisher optimal discriminant vectors so that the ratio of the between-class to within-class scatter matrix is maximized. Fisherface combines

the techniques of Principal Component Analysis with the linear discriminant analysis.

Recently, Fuzzy Fisherface was proposed for face recognition (Shieh *et al.*, 2010). Fuzzy Fisherface computed fuzzy within-class scatter matrix and between-class scatter matrix by incorporating class membership. Although it was proved to be effective, Fuzzy Fisherface did not completely incorporate the class membership into the definition of between-class and within-class scatter matrices. One of the major drawbacks of the Fuzzy fisherface method is that it fails to consider the different contribution of class to the discrimination (Kwak and Pedrycz, 2005). In order to overcome the above drawbacks, the weighted form of the between-class scatter matrix is introduced, in which a weight is added to the well-known K-nearest neighbour classifier.

**Feature extraction:** Advantages of wavelets are very flexible: several basis exists and one can choose the basis which is more suitable for a given application and provide a spatial and frequency decomposition of the image at the same time. Wavelet transforms have advantages over traditional fourier transform for representing functions that have discontinuities and sharp peaks and for accurately decomposing and reconstructing finite, non-periodic and/or non-stationary signals (Rizon, 2010).

---

**Corresponding Author:** Esther Annlin Kala James, Faculty of ECE, Government College of Technology, Coimbatore, Tamilnadu, India

LL <sub>2</sub>	HL <sub>2</sub>	HL <sub>1</sub>
LH <sub>2</sub>	HH <sub>2</sub>	
LH <sub>1</sub>		HH <sub>1</sub>

Fig. 1: Two level wavelet decomposition

The original image is resized to 128×128 and applied to Discrete Wavelet Transformation (DWT). It is decomposed into 4 frequency bands which is one low-frequency band (LL) and three high-frequency bands (LH, HL, HH). If the information of low-frequency band is again transformed, the sub-level frequency band information will be obtained as shown in Fig. 1.

For the proposed Weighted Fuzzy Fisherface technique Biorthogonal wavelet transformation is used which supports both continuous wavelet transform and discrete wavelet transform. The special feature of this transform is less vanishing points, which removes fewer details and produce little distortion. Here for simplicity Biorthogonal wavelet 1.1 is used for feature extraction and dimensionality reduction.

**Fisherface method:** The eigenface algorithm takes advantage of the fact that, the variation within class lies in a linear subspace of the image space. Hence, the classes are convex and, therefore, linearly separable. In the face recognition problem, when one seeks insensitivity to lighting conditions, linear methods for dimensionality reduction is chosen.

Fisherface method performs dimensionality reduction using linear projection and still preserves linear separability. Fisher’s linear discriminant is a class specific method, as this method tries to shape the scatter in order to make it more reliable for classification. Also it maximizes the ratio of the between-class scatter and within-class scatter matrix. Fisher’s LDA (Belhumeur *et al.*, 1997) looks for a linear subspace, within which the projections of the different classes are best separated as defined by maximizing the discriminant criteria.

Let  $\{x_i \text{ where } i=1,2,\dots,n\}$  be a set of  $n$  samples in  $N$  dimensional space and  $c$  is the number of classes. Denote the  $i$ th class samples by  $n_i$  and  $n$  be the total number of class samples. Then the between-class, the within-class and the total-class scatter matrices are defined in Eq. 1-3 respectively:

$$S_B = \sum_{j=1}^c n_j (\mu_j - \mu)(\mu_j - \mu)^T \quad (1)$$

$$S_W = \sum_{j=1}^n (x_j - \mu_i)(x_j - \mu_i)^T \quad (2)$$

$$S_T = \sum_{j=1}^n (x_j - \mu)(x_j - \mu)^T \quad (3)$$

where,  $\mu_i$  is the mean of the  $i$ th class and  $\mu = \frac{1}{n} \sum_{j=1}^n x_j$  is the global mean of all samples.

The null space of the between-class scatter matrix  $S_B$  contains no useful information for recognition, hence it is discarded by diagonalization. The within-class scatter matrix  $S_W$  is then projected into the linear subspace of  $S_B$  and factorized using eigen analysis to obtain the solution. The solution of linear discriminant analysis method contains  $c-1$  eigenvectors with non-zero eigenvalues.

If any singular matrix in  $S_B$  or  $S_W$  is involved in finding the eigenvector and eigenvalue, the diagonalization must start from the non-singular matrix. Since the scatter matrix  $S_B$  has a maximal rank of  $c-1$ , it is often singular. For a singular scatter matrix  $S_W$ , Fisher’s LDA is under constrained.

**Nonsingular within-class scatter matrix:** If  $S_W$  is nonsingular, the optimal projection  $W_{opt}$  as in Eq. 4 is chosen as the matrix with orthonormal columns which maximizes the ratio of the determinant of the between-class scatter matrix of the projected samples to the determinant of the within-class scatter matrix of the projected samples i.e.:

$$W_{opt} = \underset{W}{\arg \max} \frac{|W^T S_B W|}{|W^T S_W W|} \quad (4)$$

$$= [w_1 w_2 \dots w_m]$$

where,  $\{w_i \mid i = 1, 2, \dots, m\}$  is the set of eigenvectors of  $S_B$  and  $S_W$  corresponding to the  $m$  largest eigenvalues  $\{\lambda_i \mid i = 1, 2, \dots, m\}$  arranged in the descending order, i.e., Eq. 5:

$$S_B W_i = \lambda_i S_W W_i \quad i = 1, 2, \dots, m \quad (5)$$

Since there are at most  $c-1$  nonzero eigenvalues, an upper bound on  $m$  is  $c-1$ , where  $c$  is the number of classes.

**Singular within-class scatter matrix:** If  $S_W$  is singular, first the PCA approach for the dimensionality reduction is used such that it becomes non-singular in the lower dimensional space. This is implemented by solving the principal eigenvectors of total scatter matrix  $S_T$ . By using PCA the dimension of the feature space is

reduced to n-c and then by applying FLD the dimension is reduced to c-1. Let  $W_{pca}$  be the PCA transform matrix and  $W_{fld}$  be the fisher LDA transform matrix.

The optimal transform matrix  $W_{opt}$ , in the case of singular  $S_W$  is given by Eq. 6-8:

$$W_{opt} = W_{pca} W_{fld} \quad (6)$$

Where:

$$W_{pca} = \arg \max_W |W^T S_T W| \quad (7)$$

$$W_{fld} = \arg \max_W \frac{|W^T W^T_{pcas} S_B W_{pca} W|}{|W^T W^T_{pca} S_W W_{pca} W|} \quad (8)$$

The optimization for  $W_{pca}$  is performed over  $(N \times (n-c))$  matrices with orthonormal columns, while the optimization for  $W_{fld}$  is performed over  $(n-c) \times m$  matrices with orthonormal columns. The smallest c-1 principal components are eliminated while computing  $W_{pca}$ . There are certainly other ways of reducing the within-class scatter while preserving between-class scatter. The method which is currently used chooses  $W$  to maximize the between-class scatter of the projected samples after first reducing the within-class scatter.

**Fuzzy fisherface method:** In the above fisherface method, it is noted that the scatter matrices  $S_B$  and  $S_W$  are computed under the assumption that each class is fully assigned to a given class. In face recognition, however, as the faces may be affected by large environmental (including the illumination, poses, expression,) variation conditions, it is advantageous to assign a class membership to each sample rather than merely use the binary class assignment.

In this study, a Fuzzy K-Nearest Neighbor (FKNN) algorithm, which makes use of the distribution of samples and considers the discriminative information in the null space of fuzzy within-class scatter matrix. Samples distribution information of every class is represented by fuzzy membership degree. Kwak and Pedrycz proposed to use the following fuzzy scatter matrices  $\tilde{S}_B$  and  $\tilde{S}_W$  to replace  $S_B$  and  $S_W$ . Eq. 9 and 10:

$$\tilde{S}_B = \sum_{j=1}^c n_i (\tilde{\mu}_i - \mu) (\tilde{\mu}_i - \mu)^T \quad (9)$$

$$\tilde{S}_W = \sum_{j=1}^n (x_i - \tilde{\mu}_{ij})(x_i - \tilde{\mu}_{ij})^T \quad (10)$$

where,  $\tilde{\mu}_i = \sum_{j=1}^n \mu_{ij} x_j / \sum_{j=1}^n \mu_{ij}$  is the mean of the  $i$ th class and  $\mu_{ij}$  is the class membership grade of the  $j$ th sample  $x_j$  to  $i$ th class.

**FKNN algorithm:** The class membership gradient can be computed by using the following steps (Kwak and Pedrycz, 2005):

- Step 1: Compute the Euclidean distance matrix
- Step 2: Set diagonal elements of this matrix to infinity
- Step 3: Sort the distance matrix in an ascending order. Collect the class labels of the patterns located in the closest neighborhood of the pattern under consideration.
- Step 4: Compute the membership grad  $u_{ij}$  as in Eq.11-14:

If  $i$  equals to the label of the  $j$ th sample:

Then:

$$\mu_{ij} = 0.51 + 0.49 \times n_{ij} / k \quad (11)$$

else then:

$$\mu_{ij} = 0.49 \times n_{ij} / k \quad (12)$$

where,  $n_{ij}$  stands for the number of the neighbors of the  $j$ th sample that belong to the  $i$ th class.

As usual,  $u_{ij}$  satisfies two obvious properties:

$$\sum_{j=1}^c \mu_{ij} = 1 \quad (13)$$

$$0 < \sum_{j=1}^n \mu_{ij} < n \quad (14)$$

where,  $u_{ij} \in [0,1]$ . Therefore the class center matrix  $\mu$  and the fuzzy membership matrix  $U$  can be achieved with the result of FKNN:

$$u = [u_{ij}] \quad i = 1, 2, \dots, c \quad j = 1, 2, \dots, n$$

$$\mu = [\mu_i] \quad i = 1, 2, \dots, c$$

Therefore, the class center matrix  $\mu$  and the fuzzy membership matrix  $U$  can be achieved with the result of FKNN. In this case, the optimal transform matrix of fuzzy fisherface denoted by  $W_{FF}$ , can be found by solving the eigenvector of  $\tilde{S}_B$  and  $\tilde{S}_W$  corresponding to the  $m$  largest eigenvalues, as in Eq.15 i.e.,:

$$\tilde{S}_B W_i = \lambda \tilde{S}_W W_i, i=1, \dots, m \quad (15)$$

If  $\tilde{S}_W$  is singular, the PCA approach for the dimensionality reduction is used. Then  $W_{FF}$  is given by Eq.16 and 17:

$$W_{FF} = W_{PCA} \tilde{W}_{FLD} \quad (16)$$

Where:

$$\tilde{W}_{FLD} = \arg \max_w \frac{|W^T W^T_{PCA} \tilde{S}_B W_{PCA} W|}{|W^T W^T_{PCA} \tilde{S}_W W_{PCA} W|} \quad (17)$$

### MATERIALS AND METHODS

Some drawbacks are there on the Fuzzy Fisherface method:

- The sample distribution information is not completely used in the definitions of fuzzy between-class and within-class scatter matrices
- In PCA transformed space, the fuzzy within-class scatter matrix still might be singular
- The null space of the fuzzy within-class scatter matrix contains discriminative information for classification

**Weighted fuzzy fisherface:** Hence, Fuzzy fisher criterion is modified to achieve better recognition rate in this proposed method. More specifically, if two of the class means are far away from each other, which means that they are well separated, then their contributions to the discrimination task is minor. However, if two of the class means are close together, which means that they are not well separated, then finding the discriminant vectors that can better separate them is used to improve the discriminant performance. The fuzzy fisherface method introduces the class membership to each training sample in order to enhance the discriminant ability. Since the global mean  $\mu$  is common to all the classes and is irrelevant to the class membership of the samples, this expression may not fully take advantage of the class membership.

The between-class scatter matrix,  $S_B$  is reformulated (Zhou *et al.*, 2009) as given in Eq.18:

$$S_B = \sum_{j=1}^{c-1} \sum_{i=j+1}^c (\mu_i - \mu_j)(\mu_i - \mu_j)^T \quad (18)$$

Then each class mean  $\mu_i$  is replaced with the fuzzy class mean  $\tilde{\mu}_i$ , to obtain the fuzzy between-class scatter matrix as in Eq.19:

$$\tilde{S}_B = \sum_{i=1}^{c-1} \sum_{j=i+1}^c (\tilde{\mu}_i - \tilde{\mu}_j)(\tilde{\mu}_i - \tilde{\mu}_j)^T \quad (19)$$

To control the contribution of the class mean difference between  $\tilde{\mu}_i$  and  $\tilde{\mu}_j$  to the between-class scatter matrix  $\tilde{S}_B$ , Weighted Fuzzy Fisher face algorithm is proposed where a weight denoted by  $\Delta_{ij}$  is introduced. Moreover, if  $\tilde{\mu}_i$  and  $\tilde{\mu}_j$  are far from each other, then  $\Delta_{ij}$  is given a small value, otherwise  $\Delta_{ij}$  will be given a larger value.

Thus, weight  $\Delta_{ij}$  is defined as given in Eq. 20:

$$\Delta_{ij} = \exp(-\|\tilde{\mu}_i - \tilde{\mu}_j\|^2 / a) \quad (20)$$

where,  $a$  is a parameter to be chosen.

Then, the weighted fuzzy between-class matrix is defined by Eq. 21:

$$\tilde{S}_B = \sum_{i=1}^{c-1} \sum_{j=i+1}^c \Delta_{ij} (\tilde{\mu}_i - \tilde{\mu}_j)(\tilde{\mu}_i - \tilde{\mu}_j)^T \quad (21)$$

Accordingly, the optimal transform matrix of weighted fuzzy fisherface, denoted by  $W_{WFF}$ , can be found by solving the eigenvector  $\tilde{S}_B$  of and  $\tilde{S}_W$  corresponding to the  $m$  largest eigenvalues as in Eq.22, i.e.,:

$$\tilde{S}_B W_i = \lambda_i \tilde{S}_W W_i, i=1, \dots, m \quad (22)$$

If  $\tilde{S}_W$  is singular, the PCA approach is used for the dimensionality reduction. Then  $W_{WFF}$  is given by Eq.23 and 24:

$$W_{WFF} = W_{PCA} \tilde{W}_{FLD} \quad (23)$$

$$\tilde{W}_{FLD} = \arg \max_w \frac{|W^T W^T_{PCA} \tilde{S}_B W_{PCA} W|}{|W^T W^T_{PCA} \tilde{S}_W W_{PCA} W|} \quad (24)$$

Figure 2 shows the block diagram of the general workflow of the weighted Fuzzy fisherface technique.

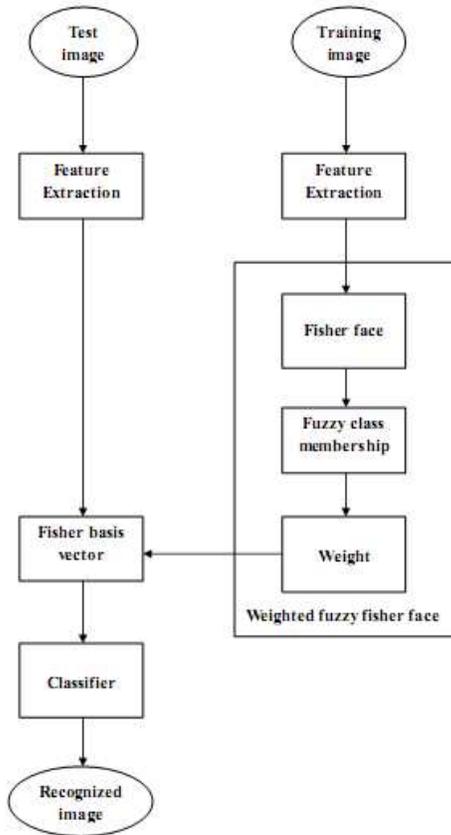


Fig. 2: A general flow for weighted fuzzy fisherface

### RESULTS

To evaluate the performance of the proposed method, the ORL face database (<http://www.cl.cam.ac.uk>) and the Yale face database (<http://cvc.yale.edu/projects/yalefaces/yalefaces.html>) face database is used. In face recognition both databases are widely used. The ORL database contains 40 distinct subjects, where each subject contains 10 different poses with varying lighting conditions. The original face images are all sized 112×92 pixels. Figure 3 shows ten face images of one subject in the ORL database.

The Yale face database contains 165 face images of 11 subjects that include variations in both facial expression and lighting condition. The original face images are sized 243×320 pixels. Figure 4 shows ten face images of one subject in the Yale database.

First the image is resized to 128×128. Then two level biorthogonal discrete wavelet transform is applied to get the feature extraction and also the dimensionality is reduced.



Fig. 3: Ten face images of one subject in the ORL face database



Fig. 4: Eight face images of one subject in the Yale face database

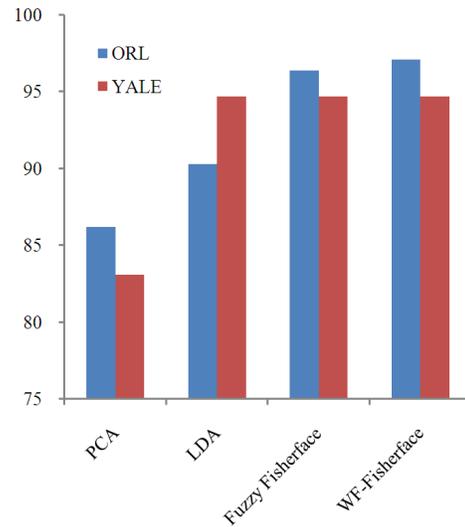


Fig. 5: Recognition rate of different face recognition methods on ORL and Yale database

Then experiment is performed by considering various values for  $k$  (K-nearest neighbour) and  $a$ . Better recognition rate with less computation is achieved by taking  $k = 5$  and  $a = 425$ . Table 1 shows that the weighted fuzzy fisherface method achieves better recognition result than fuzzy fisherface method using ORL and Yale databases.

Table 1: Comparison of recognition rate on ORL and YALE database

Database	Percentage of recognition rate				Weighted Fuzzy FF with DWT
	PCA	LDA	Fuzzy Fisherface	Weighted FuzzyFF	
ORL	86.2	90.3	92.4	94.8	98.1
YALE	83.1	87.7	89.5	92.6	96.7

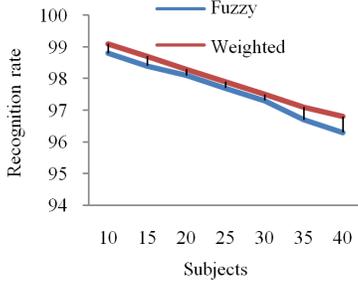


Fig. 6: Comparison of recognition rate for different subjects on ORL database for fuzzy and weighted fuzzy fisherface

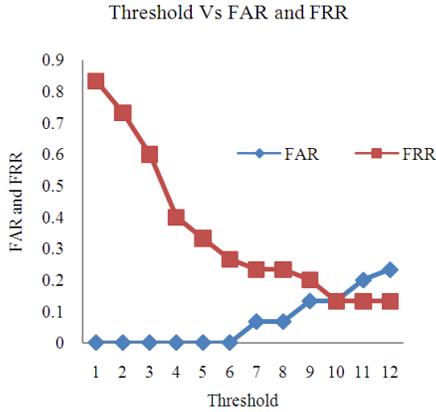


Fig. 7: Plot of various performance measures

Figure 5 shows the comparison rate of different face (PCA, LDA, Fuzzy Fisherface and Weighted Fuzzy Fisherface Weighted Fuzzy Fisherface with DWT) recognition methods on ORL and YALE database. Figure 6 shows the comparison of recognition rate for different subjects on ORL database for fuzzy and weighted fuzzy fisherface with DWT. Figure 7 shows the plot of various performance measures.

**DISCUSSION**

The accuracy of face recognition system is defined by two parameters False Acceptance Rate (FAR) and False Rejection Rate (FRR).

Table 2: Performance measures

Threshold	FAR	FRR	Accuracy
0.05	0.0000	0.833	99.58
0.06	0.0000	0.733	99.63
0.07	0.0000	0.600	99.70
0.08	0.0000	0.400	99.80
0.09	0.0000	0.333	99.83
0.10	0.0000	0.267	99.87
0.11	0.0670	0.233	99.85
0.12	0.0670	0.233	99.85
0.13	0.1330	0.200	99.83
0.14	0.1330	0.133	99.87
0.15	0.2000	0.133	99.83
0.16	0.2330	0.133	99.82

FRR measures how often an authorized user, who should be granted access, is not recognized, while FAR measures how often non-authorized user, who should not be granted access, is falsely recognized. For this a threshold value is fixed. The score above or below the threshold gives the acceptance or rejection. Therefore a threshold is set to compute the False Acceptance Rate (FAR) and False Rejection Rate (FRR). Table 2 shows the performance measures FAR and FRR at various thresholds. The control of FRR and FAR by adjusting the recognition threshold defines the accuracy of face recognition system. The FAR and FRR curves intersect at a point called Equal Error Rate (EER) or Crossover Error Rate (CER). At this point, the errors achieved are at its lowest resulting from combination of FAR and FRR. The effectiveness of the system is by setting the recognition threshold at this balance. The lower the ERR, the more reliable and accurate the system. Thus ERR is used as the standard setting for recognition threshold in face recognition system. For lowest risk of intrusion by non-authorized users, the FAR must be set at the lowest possible value.

$$FAR = \frac{NFA}{NIVA}$$

NFA is the number of false acceptance  
 NIVA is the number of imposter verification attempts:

$$FAR = \frac{NFA}{NIVA}$$

NFR is the number of false rejection  
 NEVA is the number of enrollee verification attempts.

From this the accuracy is calculated as in Eq.25:

$$Accuracy = [100 - \frac{FAR + FRR}{2}] \tag{25}$$

## CONCLUSION

The proposed weighted fuzzy fisherface method using Biorthogonal wavelet transform for face recognition is superior to fuzzy fisherface method by having full advantage of dimensionality reduction, fuzzy membership and the different contributions of the class means to the discrimination. Also it takes into account classification errors occurring between pairs of classes, unlike the fisher face. Several experiments are carried out successfully to confirm the effectiveness of the proposed method "Weighted Fuzzy Fisherface for face recognition using Wavelet Transform" with less computation time and better recognition rate. The performance measure is computed and its accuracy is calculated. The Equal Error Rate (EER) is found to be 0.133 and the threshold is fixed at 0.14 where the False Acceptance Rate (FAR) and False Rejection Rate (FRR) intersect. The lower value of FAR determines the effectiveness of the system.

## REFERENCES

- Abusham, E.E.A., A.T.B. Jin, W.E. Kion and G. Debashis, 2008. Face recognition based on nonlinear feature approach. *Am. J. Applied Sci.*, 5: 574-580. DOI: 10.3844/ajassp.2008.574.580
- Belhumeur, P.N., J.P. Hespanha and D.J. Kriegman, 1997. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Trans. Pattern Anal. Machine Intell.*, 19: 711-720. DOI: 10.1109/34.598228
- Chan, L.H. S.H. Salleh and C.M. Ting, 2010. Face biometrics based on principal component analysis and linear discriminant analysis. *J. Comput. Sci.*, 6: 693-699. DOI: 10.3884/jcssp.2010.693.699
- Kwak, K.C. and W. Pedrycz, 2005. Face recognition using a fuzzy fisherface classifier. *Pattern Recognition*, 38: 1717-1732. DOI: 10.1016/j.patcog.2005.01.018
- Ma, Y., C. Liu and H. Sun, 2006. A simple transform method in the field of image processing. *Proceedings of 6th International Conference on Intelligent Systems Design and Applications*, Oct. 16-18. IEEE Xplore Press, Jinan, pp: 429-432. DOI: 10.1109/ISDA.2006.253875
- Rizon, M., M.F. Hashim, P. Saad, S. Yaacob and M.R. Mamat *et al.*, 2006. Face recognition using eigenfaces and neural networks. *Am. J. Applied Sci.*, 3: 1872-1875. DOI: 10.3844/ajassp.2006.1872.1875
- Rizon, M., 2010. Discrete wavelet transform based classification of human emotions using electroencephalogram signals. *Am. J. Applied Sci.*, 7: 878-885. DOI: 10.3844/ajassp.2010.878.885
- Shieh, M.Y., C.M. Hsieh, J.Y. Chen, J.S. Chiou and J.H. Li, 2010. PCA and LDA based fuzzy face recognition system. *Proceeding of the SICE Annual Conference*, Aug. 18-21, IEEE Xplore Press, Taipei, pp: 1610-1615.
- Yuen, C.T., M. Rizon, W.S. San and T.C. Seong, 2009. Facial features for template matching based face recognition. *Am. J. Applied Sci.*, 6: 1897-1901. DOI: 10.3844/ajassp.2009.1897.1901
- Zhou, X., L. Zhao, W. Zheng and C. Zou, 2009. Face recognition using weighted pairwise fuzzy fisherface method. *Proceedings of the World Congress on Software Engineering*, May 19-21, IEEE Xplore Press, Xiamen, pp: 208-211. DOI: 10.1109/WCSE.2009.221
- Zhuang, X.S. and D.Q. Dai, 2007. Improved discriminate analysis for high-dimensional data and its application to face recognition. *Patt. Recog.*, 40: 1570-1578.