Research Article

Face Recognition Method Based on Fuzzy 2DPCA

Xiaodong Li

School of Logistics, Linyi University, Linyi 276005, China

Correspondence should be addressed to Xiaodong Li; seuxdli@163.com

2DPCA, which is one of the most important face recognition methods, is relatively sensitive to substantial variations in light direction, face pose, and facial expression. In order to improve the recognition performance of the traditional 2DPCA, a new 2DPCA algorithm based on the fuzzy theory is proposed in this paper, namely, the fuzzy 2DPCA (F2DPCA). In this method, applying fuzzy K-nearest neighbor (FKNN), the membership degree matrix of the training samples is calculated, which is used to get the fuzzy means of each class. The average of fuzzy means is then incorporated into the definition of the general scatter matrix with anticipation that it can improve classification result. The comprehensive experiments on the ORL, the YALE, and the FERET face database show that the proposed method can improve the classification rates and reduce the sensitivity to variations between face images caused by changes in illumination, face expression, and face pose.

1. Introduction

Face recognition is one kind of biometrics, aiming at capturing and use of behavioral or physiological characteristics for individual verification or personal identification. Over the last several decades, it has been an important issue in image processing and pattern recognition, and, in many application fields, such as access control, card identification, mug shot searching, and security monitoring, it plays an important role.

In face recognition fields, many researchers have proposed a lot of feature extraction methods [1–8], of which the Eigenfaces method based on principal component analysis (PCA) introduced by Turk and Pentland [4] is one of the popular approaches and basic classification techniques. Recently, many improved methods were developed [9–11]. In fact, the PCA dwells on a linear projection of a high-dimensional face image space into a new low-dimensional feature space by finding a set of orthogonal basis images (called Eigenfaces) [2]. In this new basis, the image coordinates (the PCA coefficients) are uncorrelated.

There is an important problem in applying the above-mentioned methods which we should take into account, however. In the PCA-based face recognition technique, the 2D face image matrices must be previously transformed into 1-D image vectors [4], and the resulting image vectors of faces usually lead to a high-dimensional image vector space. Obviously, it is difficult to evaluate the covariance matrix accurately due to its large size and the relatively small number of training samples. On the other hand, the matrix-to-vector transform may also cause the loss of some useful structural information embedding in the original images. To overcome the problems, a straightforward image projection technique, called two-dimensional principal component analysis (2DPCA), is developed by Yang et al. [12] for image feature extraction. As opposed to the conventional PCA, the 2DPCA is based on 2-D matrices rather than 1-D vector; that is, the image matrix does not need to be previously transformed into a vector. Instead, an image covariance matrix can be constructed directly using the original image matrices. In contrast to the covariance matrix of the PCA, the size of the image covariance matrix using the 2DPCA is small. As a result, the 2DPCA has two important advantages over the PCA. First, the size of covariance matrix using the 2DPCA is much smaller, so it is easier to evaluate the covariance matrix accurately. Second, the 2DPCA computes the corresponding eigenvectors more quickly than that of the PCA, so less operation time is required [12].

However, there is an important problem in applying 2DPCA method which should be mentioned. In the 2DPCA model, the mean matrix, which is generally estimated by the class sample averages of all training samples, is used to

character the total scatter matrix, so the average of training samples plays a critical role in the construction of the total scatter matrix and finally affects the projection directions of the 2DPCA. Since face recognition is typically a small sample size problem, in which only a few of image samples are available for training per class, it is difficult to give an accurate estimate of the mean using the samples average, in particular when there are outliers in the sample set (e.g., the images with variations of the noise, occlusion, etc.) [13]. Inaccurate estimate of the mean must have a negative effect on the robustness of the 2DPCA models. Another major problem coming with the use of the Eigenface technique is that it can be affected by the variations in illumination, facial expressions, and pose. All these nonideal conditions make the distribution of samples uncertain. To improve the recognition performance of the 2DPCA and address these uncertainties, taking advantage of the fuzzy technology [14] is a good choice.

So far, a number of scholars have applied fuzzy theory to face recognition algorithm [14–19]; for example, Keller et al. [14] proposed a fuzzy fisherface method, and Yang et al. [17] extended it to a 2-D space. Zhai et al. [18] applied the fuzzy rough set to select the optimal discriminant feature, and Liu and Shi [19] incorporated the local features into the fuzzy weight scheme to improve the recognition performance of the traditional 2DPCA. Inspired by the successful application of them, we envision that the fusion of fuzzy theory and 2DPCA can improve the recognition rate to some extent.

Based on what have mentioned above, in order to make full use of the class information and the distribution information of the samples to make an accurate estimate of the training samples mean in the definition of the 2DPCA model, we incorporate the fuzzy theory and the class information into the computation of the mean matrix. We call this method fuzzy 2DPCA (F2DPCA) algorithm. In this method, the fuzzy membership degree matrix based on the Euclidean distance of all training samples is calculated firstly, and, then, the fuzzy mean of each class can be obtained. Finally, we perform average operation on these fuzzy means to get the center of all training samples. In the new definition of 2DPCA model, the average of all training samples is replaced with the center obtained above. The comprehensive experiments on the ORL, the YALE, and the FERET face databases show that the proposed method can improve classification rates and reduce sensitivity to variations between face images caused by changes in illumination and viewing directions.

The organization of this paper is as follows. We briefly reviewed the theory of 2DPCA in Section 2. The content of Section 3 is that the fuzzy K-nearest neighbor was introduced firstly, and then we described the idea of fuzzy 2DPCA in detail. In Section 4, experiments on face image databases were presented to demonstrate the effectiveness of the new method. Conclusions are summarized in Section 5.

2. 2DPCA

Suppose $\mathbf{A}$ is an n-dimensional unitary column vector. The idea of 2DPCA is to project a given image $\mathbf{X}$, an $m \times n$ matrix,

onto $\mathbf{A}$ to get an m-dimensional vector $\mathbf{Y}$ by the following linear transformation [12]:

$$\mathbf{Y} = \mathbf{XA}, \tag{1}$$

we call Y the projected feature vector of $\mathbf{X}$.

The procedure of 2DPCA to obtain the $\mathbf{A}$ vector can be characterized as follows.

Suppose there are c pattern classes in $R^{m \times n}$ space and there is a sample set including N face images $\{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_N\}$, where $\mathbf{X}_i \in R^{m \times n}$; each sample belongs to a class j, $j \in \{1, 2, \ldots, c\}$. Total scatter matrix can be defined as

$$\mathbf{S}_T = \frac{1}{N} \sum_{k=1}^{N} \left(\mathbf{X}_k - \overline{\mathbf{X}} \right) \left(\mathbf{X}_k - \overline{\mathbf{X}} \right)^T, \tag{2}$$

where N denotes the number of training samples and $\overline{\mathbf{X}} = (1/N) \sum_{k=1}^{N} \mathbf{X}_k$ is the mean of all training samples.

In 2DPCA algorithm, the optimal projection vector A_{opt} satisfies with

$$\mathbf{A}_{opt} = \arg \max \left| \mathbf{A}^T \mathbf{S}_T \mathbf{A} \right| = [\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_m], \tag{3}$$

where $\{\mathbf{A}_i \mid i = 1, 2, \ldots m\}$ are m orthogonal eigenvectors of $\mathbf{S}_T$ corresponding to the largest m eigenvalues, respectively.

3. Fuzzy 2DPCA

3.1. Fuzzy K-Nearest Neighbor (FKNN). In our method, the fuzzy membership degree and the class centers are obtained through the FKNN algorithm [14]. With FKNN, the computation of the membership degree can be realized through the following steps.

Step 1. Compute the Euclidean distance matrix between pairs of image matrices in the training set.

Step 2. Set the diagonal elements of the Euclidean distance matrix to infinity.

Step 3. Sort the distance matrix (treat each of its columns separately) in an ascending order. Collect the corresponding class labels of the patterns located in the closest neighborhood of the pattern under consideration (as we are concerned with "k" neighbor, this returns a list of "k" integers).

Step 4. Compute the membership degree to class "i" for the jth pattern using the method proposed in the literature [14] according to the following equation to construct the membership degree matrix $\mathbf{U}$:

$$u_{ij} = \begin{cases} 0.51 + 0.49 \times \left(\dfrac{n_{ij}}{k} \right), & \text{if } i = \text{the label of the } j\text{th pattern} \\ 0.49 \times \left(\dfrac{n_{ij}}{k} \right), & \text{if } i \neq \text{the label of the } j\text{th pattern}. \end{cases} \tag{4}$$

Consider $\mathbf{U} = [u_{ij}]$, $i = 1, 2, \ldots, c$; $j = 1, 2, \ldots, N$, which satisfies two obvious properties:

$$\sum_{i=1}^{c} u_{ij} = 1$$

$$0 < \sum_{j=1}^{N} u_{ij} < N, \tag{5}$$

where j is the number of the neighbors of the jth data (pattern) that belong to the ith class. After the examination of the membership allocation formula we conclude that the method attempts to "fuzzify" or refine the membership grades of the labeled patterns [14].

Intuitively, if there are very few neighbors of the pattern that belong to the same category, the membership grade is kept close to 0.51. Alternatively, if $n_{ij} = k$ means that all neighbors are in the same class as the pattern under consideration, then $u_{ij} = 1$ [14].

3.2. The Algorithm of Fuzzy 2DPCA. The key step of fuzzy 2DPCA is how to incorporate the contribution of each training sample into the total scatter matrix. Based on the fuzzy set theory, each sample can be classified into multiclasses with fuzzy membership degrees, instead of binary classification, so, in the redefinition of the fuzzy total scatter matrix, the membership degree of each sample (contribution to each class) and the class information should be considered. The idea of the fuzzy 2DPCA is that the means of each class are calculated with fuzzy membership degrees matrix of all training samples by (6) firstly, and then we average all class means to get the center of all training samples. In the definition of the new proposed method, the original mean $\overline{\mathbf{X}}$ of all samples is replaced with the new obtained fuzzy mean $\widetilde{\mathbf{X}}$. Based on what we have described above, the algorithm of the proposed supervised fuzzy 2DPCA can be summarized as follows.

Step 1 (FKNN). The fuzzy membership degree matrix $\mathbf{U}$ can be computed with the FKNN algorithm in the original training image space.

Step 2. Based on $\mathbf{U}$, compute the fuzzy class mean $\widetilde{\mathbf{X}}_i$ of each class; after that, the average procedure is performed on all fuzzy class means to get the fuzzy average $\widetilde{\mathbf{X}}$ of all training samples according to (6) and (7). Consider

$$\widetilde{\mathbf{X}}_i = \frac{\sum_{j=1}^{N} u_{ij} \mathbf{X}_j}{\sum_{j=1}^{N} u_{ij}}, \tag{6}$$

$$\widetilde{\mathbf{X}} = \frac{1}{c} \sum_{l=1}^{c} \widetilde{\mathbf{X}}_l. \tag{7}$$

Step 3. Redefine the total scatter matrix according to the new obtained fuzzy average vector $\widetilde{\mathbf{X}}$. Then the optimal projection

Figure 1: Sample images of some persons in the ORL database.

Table 1: Maximum recognition rate on the ORL database with 7 training samples.

Method	PCA	2DPCA	F2DPCA
Recognition rate	95.77	96.43	98.83
Dimension	90	112×6	112×6

matrix can be obtained by computing the optimal problem. Consider

$$\widetilde{\mathbf{S}}_T = \frac{1}{N} \sum_{k=1}^{N} \left(\mathbf{X}_k - \widetilde{\mathbf{X}} \right) \left(\mathbf{X}_k - \widetilde{\mathbf{X}} \right)^T \tag{8}$$

$$\widetilde{\mathbf{A}}_{\text{opt}} = \arg \max \left| \widetilde{\mathbf{A}}^T \widetilde{\mathbf{S}}_T \widetilde{\mathbf{A}} \right| = \left[\widetilde{\mathbf{A}}_1, \widetilde{\mathbf{A}}_2, \ldots, \widetilde{\mathbf{A}}_m \right].$$

Step 4 (classify). Project all samples into the obtained optimal discriminant matrix $\widetilde{\mathbf{A}}_{\text{opt}}$ and classify testing samples with nearest distance classifier.

4. Experiments

We compare the proposed algorithm (F2DPCA) with the traditional PCA [4] and the 2DPCA [12] on three face image databases, namely, the ORL database, the YALE database, and the FERET database, to illustrate the effectiveness of our method. To further show the performance of the F2DPCA, we also compare the F2DPCA algorithm with other subspace methods that were combined with the fuzzy theory, for example, the FKF [15], the CFLDA [16], the F2DLDA [17], the 2DPCAF [18], and the FW2DPCA [19]. The face image database introduction and experiment results are showed in the next subsections.

4.1. Experiment on the ORL Database. The ORL database [20] is a basic face database for testing face recognition method, which includes 400 face images from 40 subjects, each providing 10 different images with the size of 112×92 pixels and 256 grey levels per pixel. Some images were taken at different times, containing variation of the lighting, facial expressions, and facial details. Figure 1 shows some sample images of different individuals.

In our experiments, 7 random images of each individual are used for training, and the remaining 3 images are used for testing. The PCA, 2DPCA, and F2DPCA are, respectively, used for feature extraction. In order to make full use of

TABLE 2: Maximal recognition rate and time complexity comparison on the ORL database with 7 training samples.

Method	FKF	CFLDA	F2DLDA	2DPCAF	FW2DPCA	F2DPCA
Recognition rate (%)	98.80	99.21	99.30	96.87	97.11	98.83
Recognition time (s)	0.076	0.068	0.061	0.046	0.041	0.036

···▽··· PCA
——△—— 2DPCA
——◇—— F2DPCA
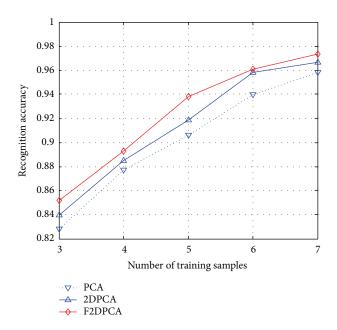
FIGURE 2: Recognition rate versus number of training samples on the ORL database.
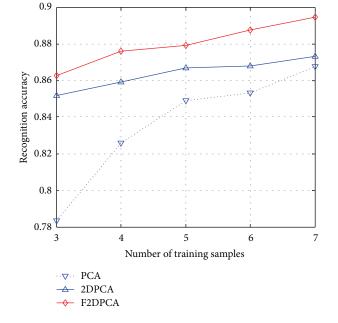

···▽··· PCA
——△—— 2DPCA
——◇—— F2DPCA

FIGURE 4: Recognition rate versus number of training samples on the YALE database.

FIGURE 3: Sample images of some persons in the YALE database.

FIGURE 5: Sample images of some persons in the FERET database.

TABLE 3: Maximum recognition rate on the YALE database with 6 training samples.

Method	PCA	2DPCA	F2DPCA
Recognition rate	85.33	86.48	88.5
Dimension	30	100×6	100×6

the available image samples and to evaluate the above algorithms more accurately, we adopt a cross-validation strategy and run the recognition procedure 10 times with different training sample set and testing sample set [17].

We keep nearly 95-percent image energy and the number of principal components. The FKNN parameter K is set as $k = n - 1$, where n denotes the number of training samples per class. Finally, a nearest neighbor classifier with the Euclidean distance is employed for classification. The recognition result versus the dimension is shown in Table 1, from which we can see that our method works well and the recognition performance outperforms that of the PCA and 2DPCA.

Figure 2 demonstrates that, with the number of training samples increasing, the maximal recognition rates of the PCA, 2DPCA, and F2DPCA all increase greatly, but, corresponding to the same number of training samples, the recognition rate of the proposed method is the best in the three algorithms, so the method we put forward is efficient.

TABLE 4: Maximal recognition rate and time complexity comparison on the YALE database with 6 training samples.

Method	FKF	CFLDA	F2DLDA	2DPCAF	FW2DPCA	F2DPCA
Recognition rate	88.58	88.67	89.11	86.41	87.2	88.5
Recognition time	0.082	0.068	0.071	0.059	0.051	0.046

TABLE 5: Maximum recognition rate on the FERET database with 4 training samples.

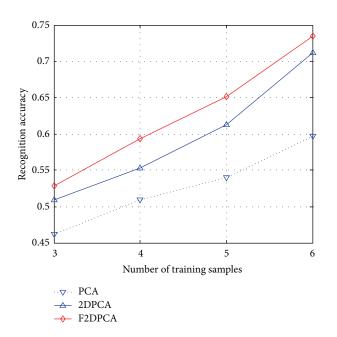| Method	PCA	2DPCA	F2DPCA
Recognition rate	50.95	55.12	59.67
Dimension	100	40×6	40×6

FIGURE 6: Recognition rate versus number of training samples on the FERET database.

The contents of Table 2 give the maximum recognition rate and time complexity comparison of different face recognition algorithm. The value of the time complexity is the time to recognize the remaining testing sample images. From Table 2, we can see that the recognition rate of the F2DPCA is slightly lower than that of the CFLDA and the F2DLDA, but the time complexity is much lower than that of the CFLDA and the F2DLDA. The ideas of the CFLDA and the F2DLDA are similar to the idea of the F2DPCA, but the CFLDA and the F2DLDA include complex inversion computation, so their computation complexities are higher. The core of the 2DPCAF and the FW2DPCA is still the idea of the 2DPCA algorithm, so the overall performance is poorer than that of the F2DPCA.

4.2. Experiment on YALE Database. There are 165 images from 15 individuals (each person has 11 different images) under various lighting conditions and facial expressions in the YALE face database [21]. In our experiment, each image was manually cropped and resized to 100×100 pixels. The processed sample images of some person are shown in Figure 3.

The same experiment procedure as that on ORL database is performed on the YALE database. The experiment was performed using 6 random images per class for training and the remaining 5 images for testing. The PCA, 2DPCA, and the proposed F2DPCA are, respectively, used for feature extraction. The recognition rate versus the dimension is illustrated in Table 3.

Table 3 shows that the F2DPCA significantly outperforms the PCA and the 2DPCA. The maximal recognition rate of F2DPCA is 89%, while that of 2DPCA is only 87%.

Besides, as illustrated in Figure 4, the experiment results also tell us that the recognition rates of the PCA, the 2DPCA, and the F2DPCA all increase significantly with the number of training samples increasing, but the F2DPCA has better performance than the others on the whole and the F2DPCA is more robust to outliers than PCA and 2DPCA.

The contents of Table 4 tell us that the overall performance of the F2DPCA is better than that of other methods.

4.3. Experiment on FERET Database. The proposed method was also tested on a subset of the FERET database. The FERET face image database is a standard database for testing and evaluating state-of-the-art face recognition algorithm [22, 23]. This subset includes 1400 images with variations in facial illumination, expression, and pose. All these images are from 200 people (each person has 7 images). In our experiment, the facial portion of each original image is cropped manually based on the location of eyes and resized to 40×40 pixels and without histogram equalization. Some processed images of one person are shown in Figure 5.

In this experiment, we applied the same experiment procedure as that on ORL database, and 4 random images of each individual are used for training, and the remaining 3 images are used for testing. The PCA, the 2DPCA, and the F2DPCA are, respectively, used for feature extraction. The recognition results are obtained by nearest neighbor classifier with the Euclidean distance. After repeating the recognition procedure 10 times with different training sets and testing sets, the average result is used as the final recognition rate. The recognition results versus the dimension of three methods are plotted in Table 5. The data in the table indicates that both the PCA and the 2DPCA are inferior to the F2DPCA in terms of recognition performance.

Figure 6 gives the situation of the recognition rate versus the number of training samples on the FERET database using three feature extraction methods, respectively. Obviously, the recognition performance of the proposed method is better.

TABLE 6: Maximal recognition rate and time complexity comparison on the FERET database with 4 training samples.

Method	FKF	CFLDA	F2DLDA	2DPCAF	FW2DPCA	F2DPCA
Recognition rate	59.86	60.27	60.81	57.76	58.23	59.67
Recognition time	0.068	0.061	0.058	0.04	0.035	0.028

The experiment results from Figures 2, 4, and 6 also show us that the proposed method works better on both the YALE and the FERET, because the samples in the YALE and the FERET contain more variations of expression, illumination, and pose, so the F2DPCA is robust to these nonideal conditions.

Table 6 shows the comprehensive analysis data of the 6 methods. From the table we can see that the overall performance of the F2DPCA is better than that of other methods.

5. Conclusion

In order to improve the recognition performance of the 2DPCA under nonideal conditions, a new method for feature extraction, namely, the fuzzy 2DPCA (F2DPCA), is proposed in this paper. Considering the important role that the fuzzy set theory plays in processing the uncertainty, in this method we fuse the 2DPCA and the fuzzy K-nearest neighbor. According to the fuzzy K-nearest neighbor algorithm, the fuzzy membership degree of all samples to each class is calculated to construct the membership degree matrix, firstly. The average of the fuzzy means of each class is computed which is used in the redefinition of the total scatter matrix. From the above discussion contents, we can see that the proposed method makes full use of both the class information and the sample's distribution information, so it can improve recognition results. Experiments on the ORL, the YALE, and the FERET face databases showed that the new method works effectively.

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

References

[1] X. Tan, S. Chen, Z.-H. Zhou, and F. Zhang, "Face recognition from a single image per person: a survey," *Pattern Recognition*, vol. 39, no. 9, pp. 1725–1745, 2006.

[2] K.-C. Kwak and W. Pedrycz, "Face recognition using a fuzzy fisherface classifier," *Pattern Recognition*, vol. 38, no. 10, pp. 1717–1732, 2005.

[3] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: a literature survey," *ACM Computing Surveys*, vol. 35, no. 4, pp. 399–458, 2003.

[4] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.

[5] X. Li, S. Fei, and T. Zhang, "Median MSD-based method for face recognition," *Neurocomputing*, vol. 72, no. 16–18, pp. 3930–3934, 2009.

[6] X. Li, S. Fei, and T. Zhang, "Weighted maximum scatter difference based feature extraction and its application to face recognition," *Machine Vision and Applications*, vol. 22, no. 3, pp. 591–595, 2011.

[7] W. Yang, J. Wang, M. Ren, J. Yang, L. Zhang, and G. Liu, "Feature extraction based on Laplacian bidirectional maximum margin criterion," *Pattern Recognition*, vol. 42, no. 11, pp. 2327–2334, 2009.

[8] W. Yang, C. Sun, and L. Zhang, "A multi-manifold discriminant analysis method for image feature extraction," *Pattern Recognition*, vol. 44, no. 8, pp. 1649–1657, 2011.

[9] M. S. Bartlett, J. R. Movellan, and T. J. Sejnowski, "Face recognition by independent component analysis," *IEEE Transactions on Neural Networks*, vol. 13, no. 6, pp. 1450–1464, 2002.

[10] J. Yang, Z. Jin, J.-Y. Yang, D. Zhang, and A. F. Frangi, "Essence of kernel Fisher discriminant: KPCA plus LDA," *Pattern Recognition*, vol. 37, no. 10, pp. 2097–2100, 2004.

[11] M. H. Yang, "Kernel Eigenfaces vs. Kernel fisherfaces: face recognition using Kernel methods," in *Proceedings of the 5th IEEE International Conference on Automatic Face and Gesture Recognition (RGR '02)*, pp. 215–220, 2002.

[12] J. Yang, D. Zhang, A. F. Frangi, and J.-Y. Yang, "Two-dimensional PCA: a new approach to appearance-based face representation and recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 1, pp. 131–137, 2004.

[13] J. Yang, D. Zhang, and J.-Y. Yang, "Median LDA: a robust feature extraction method for face recognition," in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, pp. 4208–4213, October 2006.

[14] J. M. Keller, M. R. Gray, and J. A. Givens, "A fuzzy k-nearest neighbor algorithm," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 15, no. 4, pp. 580–585, 1985.

[15] Y. Zheng, J. Yang, W. Wang, Q. Wang, J. Yang, and X. Wu, "Fuzzy Kernel fisher discriminant algorithm with application to face recognition," in *Proceedings of the 6th World Congress on Intelligent Control and Automation (WCICA '06)*, vol. 2, pp. 9669–9672, June 2006.

[16] W. Yang, H. Yan, J. Wang, and J. Yang, "Face recognition using complete fuzzy LDA," in *Proceedings of the 19th International Conference on Pattern Recognition (ICPR '08)*, December 2008.

[17] W. Yang, X. Yan, L. Zhang, and C. Sun, "Feature extraction based on fuzzy 2DLDA," *Neurocomputing*, vol. 73, no. 10–12, pp. 1556–1561, 2010.

[18] J.-H. Zhai, C.-Y. Bai, and S.-F. Zhang, "Face recognition based on 2DPCA and fuzzy-rough technique," in *Proceedings of the*

International Conference on Machine Learning and Cybernetics (ICMLC '10), pp. 725–729, July 2010.

[19] L. Liu and C. Shi, "Imerging local feature and fuzzy weighting for 2DPCA face recognition," in *Proceedings of the 2nd Annual Conference on Electrical and Control Engineering (ICECE '11)*, pp. 5295–5298, September 2011.

[20] The ORL Face Database, http://www.uk.research.att.com/face database.html.

[21] Yale Face Database, http://cvc.yale.edu/proiects/yalefaces/yale faces.html.

[22] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The FERET evaluation methodology for face-recognition algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, pp. 1090–1104, 2000.

[23] P. J. Phillips, "The Facial Recognition Technology (FERET) Database," 2004, http://www.itl.nist.Gov/iad/humanid/feret/feret_master.htmlS.