Modular Image Principal Component Analysis for Face Recognition

José Francisco Pereira, George D. C. Cavalcanti and Tsang Ing Ren

***Abstract*— One of the most successful process to accomplish human face recognition are the methods based on the Principal Component Analysis (PCA), also known as Eigenfaces. Recently, novel PCA approaches have been proposed: modular (MPCA) and two-dimensional (IMPCA). These approaches have achieved outstanding result in feature extraction and recognition. IMPCA is used for feature extraction based on 2D matrix representation and MPCA is based on image division to improve face recognition with variations like facial expressions, light and head pose. In this work we use some aspects of these methods to build a new technique called Modular Image PCA (MIMPCA). The results achieved with the proposed method are superior in all experiments compared with the original techniques under different conditions of head pose angle, illumination and facial expression.**

I. INTRODUCTION

FACE recognition is a biometric method for identifying individuals using the features of their faces. It is considered as one of the fundamental problems in computer vision, and many scientists from different fields have being addressing this problem. Research in this area has been conducted for more than thirty years. A pioneer work is the well-known eigenfaces [6], i.e., the Karhunen-Loève transform (also known as Principal Component Analysis) applied to faces.

Currently, Principal Component Analysis, which is a statistical approach where faces are expressed as a subset of the eigenvectors, is one of the methods that yield the best results on frontal face recognition. PCA is also used in other research areas like handprint, object recognition, and industrial robotics [3]. Many other methods based on PCA have been developed to increase its accuracy and reduce its computational cost. These methods improve the feature extraction process and take advantage of partial variations of the face images.

Yang et al. [3] proposed a technique called Two-dimensional PCA (IMPCA or 2DPCA) that uses a 2-dimensional matrix representation for faces instead of the traditional 1-dimensional representation used by PCA. Using this technique the face representation is much smaller than the one necessary for the traditional PCA. Therefore, the method works with low dimensional data leading to a more statistical representative covariance matrix.

Other recent improvement was the Modular Principal Component Analysis (MPCA) developed by Gottmukkal and Asari [2]. In this technique, the face images are divided into smaller regions and the PCA approach is applied for each one of these sub-images. However, the MPCA technique does not simply apply PCA in each independent face region; since it keeps the relation between the regions and the global information of the face. In this way, the global information of the face is indirectly used for computing the vector of weights that will represent the images.

The principal objective of this paper is to improve face recognition subjected to varying facial expression, illumination and head pose. PCA based face recognition methods are not very accurate when the illumination and facial expression vary considerably. This work proposes a new technique that aims to combine the best aspects of Modular PCA and IMPCA. This new approach is called Modular Image Principal Component Analysis (MIMPCA, for short).

The development of the new proposed approach for facial feature extraction was motivated by the idea that the weakness in one of the discussed technique could be alleviated by the characteristics of the other. The image representation vector obtained by the IMPCA technique is statically more representative than the original PCA-based techniques [3], [11] but variations in illumination and facial expressions decrease considerably its recognition performance. On the other hand, the Modular PCA is ideal to be used in images that have local variation but its feature extraction is based on traditional PCA technique [2]. Therefore, we expected an improvement in performance by combining the modular features extraction with the image representation obtained with the IMPCA approach.

This work is organized as follows. In Section II, the Principal Component Analysis approaches for feature extraction of faces and the proposed technique is described. In Section III, the experimental study methodology is described. In Section IV, the obtained results are analyzed and discussed. Finally, in Section V some concluding remarks are presented.

II. PRINCIPAL COMPONENT ANALYSIS APPROACHES

PCA-based methods for face recognition have achieved a good performance to represent and to recognize frontal human faces. However, its accuracy is intensely affected by face image variations, like: illumination, facial expression and head pose. Furthermore, in PCA-based face recognition techniques, the $2D$ image matrices must be transformed into a $1D$ vector. As a result, the image representation leads to a high dimensional space. Consequently, it is difficult to evaluate the covariance matrix accurately due to its large size and the relative small numbers of training samples.

Researches in the area attempt to eliminate these limitations in the PCA-base techniques. The following subsections present recently developed approaches for PCA (II-A and II-B) and the strategy proposed here (II-C).

José Francisco Pereira, George D. C. Cavalcanti and Tsang Ing Ren are with Center of Informatics, Federal University of Pernambuco, Brazil. E-mail: {jfp,gdcc,tir}@cin.ufpe.br. Site: www.cin.ufpe.br/∼viisar

A. Two-Dimensional Principal Component Analysis (IMPCA)

IMPCA [3] is a technique based on PCA which treats the images as a two-dimensional matrix instead of the traditional $1D$ vector representation. In other words, the image matrix does not need to be previously converted into a vector, which is a simpler and more straightforward technique for feature extraction. Instead, an image covariance matrix can be constructed using directly the original image matrix, decreasing its computational cost and increasing the statistical representation of the image samples.

In this way, this approach considers the entire image matrix to compute the basis vector of the new feature spaces. So, considering an image matrix of size $m \times n$ the covariance matrix will be of size $m \times m$ instead of $m.n \times m.n$ obtained using the traditional PCA. Then, the covariance matrix can be obtained by

$$G_t = \sum_{j=1}^{M} (A_j - \bar{A})^T (A_j - \bar{A}) \tag{1}$$

where $\bar{A}$ is the average image of all training samples and A_j represents the j-th image of the training database.

Once the covariance matrix is defined, the optimal projection axis is given by the unitary vector that maximizes the generalized total scatter criterion, i.e., the eigenvector of G_t corresponding to the largest eigenvalues [2]. But, in general, it is not enough to have only one projection axis. Usually, it is necessary to select a set of projection axis, $V_1, V_2, \ldots, V_d$, corresponding to the d largest eigenvalues.

As a result, the final projection matrix can be defined as $P = [V_1, V_2, \ldots, V_d]$ where each column corresponds to an eigenvector of the covariance matrix. So, the test image can be projected on the new feature space using the expression:

$$Y_t = P \cdot I_t \qquad \forall t = 1 \ldots T \tag{2}$$

where Y_t represents the t-th test image and T represents the total number of images in the test database.

As a result, IMPCA has at least two important advantages over PCA. First, it is easier to calculate the covariance matrix accurately because of its low dimensionality. Second, less time is required to determine the corresponding eigenvectors. However, the technique is affected by image variations and it is not as efficient as PCA in terms of storage requirements. In IMPCA, each principal component is represented by a n-dimensional vector instead of a single scalar, in general, $n >> 1$.

In this technique the entire digit image, i.e., the global information is used in the feature extraction phase. So, important local features can be despised using this holistic approach. These local features can be used for improve the projected data representation and consequently enhance the global accuracy rate.

B. Modular Principal Component Analysis (MPCA)

The PCA-based techniques for face recognition are not efficient taking into account local variations of the face images. Variations like head pose, illumination and facial expression affect considerably the system performance, since the total information of the image is considered. Under these conditions the projections of the images vary significantly from the images in normal circumstances. Hence it is difficult to identify them correctly.

To address these problems of face invariant recognition under illumination, pose and facial expression, the MPCA technique was proposed by Gottumukkal and Asari [2]. This approach works as follows: the entire image is divided into smaller regions and the feature extraction is computed for each one of these regions. So, the local projection will be more representative to the area it covers, see Figure 1. Since some variations on face images do not affect the entire information of the faces, this technique takes advantage of the unaffected regions of the face to improve its accuracy rate.

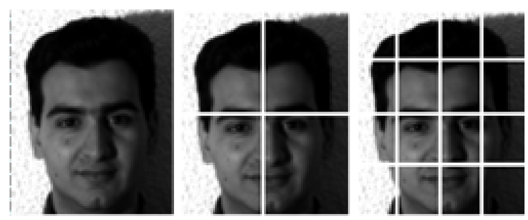

Fig. 1. Graphical representation of image subdivision based on MPCA technique

Consequently, in these unaffected regions, the projections of the face image will closely match the projections in that same region of an individual face image under normal conditions. Therefore, it is expected that improved recognition rates can be obtained by following the modular PCA approach. However, if the face images were divided into very small regions the global information of the faces may be lost and the accuracy of this method may deteriorate.

The original technique defines only one average matrix and only one covariance matrix for all training images, independently of the number of regions they were divided. This can be interpreted as a simple way of linking the smaller regions and keeping the global information of the original face image. As a result of this assumption, all the images are subtracted from the global mean and used to define the covariance matrix.

The basis vector of the new feature space is defined using the eigenvectors corresponding to the largest eigenvalues of the covariance matrix. This eigenvectors will be used to construct the final matrix projection for each region of the face.

Experiments performed by Yang et al. [3], with the UMIST and the Yale face databases, showed better performance of MPCA over PCA under conditions where illumination and facial expression varied. It is important to mention that even achieving a good recognition performance, this strategy uses the traditional PCA in each sub-image. However, this assumption decreases the representation power of the data, enhances its computational cost and harms the recognition performance of the system [2].

C. Modular Image Principal Component Analysis (MIMPCA)

The PCA-based face recognition method is not very effective under the conditions of varying pose and illumination, since it considers the global information of each face image. Under these conditions the weighted vectors that represent the image vary considerably from the normal face image. So, the accuracy of the technique is significantly affected by these changes.

Variation on facial expression and illumination generally affect only some regions of the faces. Meanwhile, other regions will remain the same as the face regions of the normal images. From this point of view, the technique will have good recognition performance comparing to the techniques based on traditional PCA due to its modular approach.

Another important attribute of feature extraction process is the quality or representativeness of the data extracted from the image. In traditional PCA-based methods the original matrix image must be previously transformed into a 1D image vector. The resulting image vectors of faces usually lead to a very high-dimensional image vector space, where it is difficult to evaluate the covariance matrix accurately due to its large size and relatively small number of training samples. The proposed technique uses feature extraction approach proposed in Image Principal Component (IMPCA) technique to improve the data representation.

Detailed explanation about the MIMPCA technique is described below:

Let $I_1, I_2, \ldots, I_M$ be the training set of face images. So, I_m denotes an image of size $K \times L$ in the training set, represented by a matrix of the same size. These images belong to Q different classes. The whole training set is divided in Q subsets: $C_1, C_2, \ldots, C_Q$. And each C_q represents a set of images I_m which belongs to the class q. $M = \sum_{q=1}^{Q} |C_q|$, where $|\cdot|$ is the cardinality of a set, i.e., the number of elements.

In this method, each image is divided into A pieces horizontally and B pieces vertically. Therefore, the original image is divided into N sub-image where $N = A \times B$ and the size of each sub-image is equal to $(KxL)/N$ pixels. These sub-images can be represented mathematically as

$$I_{mij}(x,y) = I_m\left(\frac{K}{A}(j-1)+x, \frac{L}{B}(i-1)+y\right) \forall i,j \quad (3)$$

where i varies from 1 to A and j varies from 1 to B, thus I_{mij} represents the sub-images of coordinates i,j of the m-th image in the training set.

In this technique just one average image is obtained for all sub-images. The average image is calculated as

$$\bar{A} = \frac{1}{Q}\sum_{q=1}^{Q} \bar{A}_q \quad (4)$$

where $\bar{A}_q$ corresponds to the average image of q-th class and is computed as:

$$\bar{A}_q = \frac{1}{(A \times B \times |C_q|)} \sum_{m=1}^{M}\sum_{i=1}^{A}\sum_{j=1}^{B} I_{mij}, \quad I_m \in C_q \quad (5)$$

The next step is to normalize all sub-images by subtracting them from the global mean

$$Z_{mij} = I_{mij} - \bar{A} \qquad \forall m,i,j \quad (6)$$

where Z_{mij} represents the normalized region vector with ij coordinates of the m-th image in the training set.

Based on the sub-images matrices the covariance matrix can be calculated as defined in Equation 7.

$$S = \frac{1}{Q}\sum_{q=1}^{Q} S_q \quad (7)$$

where S_q corresponds to the class covariance matrix of q-th class in dataset. This matrix can be computed as:

$$S_q = \frac{1}{(A \times B \times |C_q|)} \sum_{m=1}^{M}\sum_{i=1}^{A}\sum_{j=1}^{B} (Z_{mij} \cdot Z_{mij}^T), \;\; I_m \in C_q \quad (8)$$

It is important to highlight that only one covariance matrix is used for all the sub-images. The same happens with the mean. The intermediate mean ($\bar{A}_q$) and covariance (S_q) are only used for global mean and covariance matrix definition, respectively. Experimentally, it was observed that using only one average matrix and only one covariance matrix for all the training set, the final system precision is improved. This approach can capture global information.

The first V eigenvectors associated with the largest eigenvalues obtained from the covariance matrix S can be calculated. These eigenvectors are represented by $E_1, E_2, \ldots, E_V$. The image weights are computed by multiplying the normalized images (Z_{mij}) by the eigenvectors generating the principal components.

$$W_{mij} = (I_{mij} - \bar{A}) \cdot P \qquad \forall m,i,j,v \quad (9)$$

In general, it is not enough to have only one eigenvector to obtain the face image weights. Usually a set of eigenvector, $E_1, E_2, \ldots, E_V$, is used to represent these images. So, it is used V components to represent the image which results in a $(K/B) \times V$ matrix (W_{mij}) for each sub-image of the original image. In this way, the final projection matrix (P) is constructed using each eigenvector E_v as a column of the final projection matrix defined as $P = [E_1^T E_2^T E_3^T \ldots E_V^T]$. Now each projected sub-image sample can be computed as a simple matrix multiplication as define in Equation 9.

These weights are also computed for test sub-images $I_{test,i,j}$ using the same projection matrix P as shown in Equation 10.

$$W_{test,i,j} = (I_{test,i,j} - \bar{A}) \cdot P \qquad \forall i, j, m \tag{10}$$

Therefore, for representing an entire image, which is divided in N sub-images, it is necessary N matrix containing $L \times V$ coefficients. This is a disadvantage of MIMPCA compared to the original Modular PCA technique: its storage requirements. Since, it is necessary more coefficients to store the image representation.

Once extracted the weights (principal components) from training and test images a Nearest Neighbor based classifier is used for face classification. The distance between the two image representations is defined in Equation 11.

$$d(I_m, I_t) = \sum_{i=1}^{A} \sum_{j=1}^{B} (W_{mij} - W_{tij}) \tag{11}$$

where I_m is the reference image and I_t is the test image. This distance is computed for the test image against all patterns in the training set. The class of the test image is the same of the reference image closer to it.

III. Experiments

The experiments were conducted over three well-known face databases (Yale, ORL and UMIST). The three techniques (IMPCA, MPCA and MIMPCA) were tested under the same images and conditions. The ORL database was used to evaluate the performance of the techniques under conditions where the pose and the sample size vary. The Yale database was used to test the performance when facial expressions and illumination change. Finally UMIST face database was used to evaluate the performance over large change on face positions. All images in the Yale face database were cropped and normalized to 92×112 pixels.

In all the experiments, the first five images were used to train and the remaining images to test the techniques. Different subdivision configurations were applied to MPCA and MIMPCA ($N = 4, 9, 16$). Preliminary results showed that the recognition rate decreases for values of N greater than 16, this is probably due to the small size of each sub-image (Figure 2). At this condition (N greater or equals to 16), each sub-image will be smaller than 23×28 pixels and will not represent local information of the face efficiently.

Finally, a Nearest Neighbor classifier was used for classification. Threshold was not used for these experiments; hence there are no rejections, only correct recognition or false recognition. Note that the calculation of the distance between samples in each technique is slightly different. For MPCA the distance is based on N feature vector, where N is the number of subdivision and in the IMPCA technique the distance is calculated over matrices, while in the MIMPCA technique the distance is based on N feature matrices.

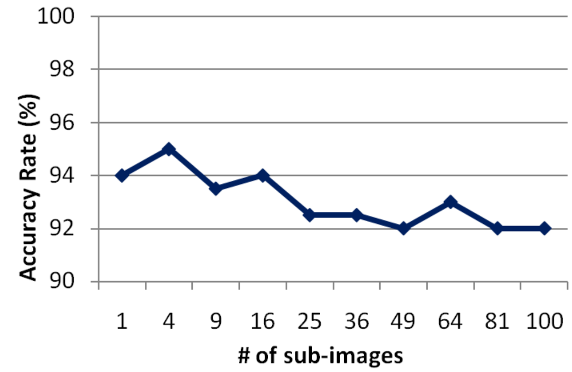

Fig. 2. IMPCA performance over ORL face database varying the number of sub-images per sample

IV. Results

The tests aim to compare the performance of the techniques under variations of illumination, head position and facial expression. The results presented for the UMIST database was achieved based on the experiments proposed by Gottmukkal and Asari [2] where training and test image samples were chosen to evaluate the techniques with head pose angles that lie outside the head pose they were trained with. Based on these experiments, the low performance over this database can be understood.

Note that the MPCA and the MIMPCA techniques are based on the image sub-division, hence in the next figures this information is denoted by: MPCA($A \times B$) or IM2PCA($A \times B$), where A and B were defined in Section II-C.

Experiments performed by varying the number of face regions per image sample show that the recognition rate keeps stable or decreases for values of N greater than 16. Figure 2 shows the recognition rate of the proposed technique over the ORL face database.

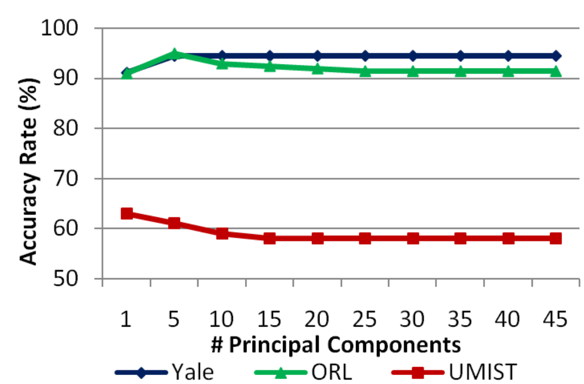

Fig. 3. MIMPCA recognition rate performance improvement varying the number of principal components

Figure 3 shows the MIMPCA global performance over all three databases varying the number of principal components used for face representation. There is not an improvement in recognition rate when 25 or more principal components are used for image representation. In other words, the proposed technique requires only the first components to find a good representation. Thus, using more than 25 components, the information added to face representation is irrelevant for accuracy rate enhancement.

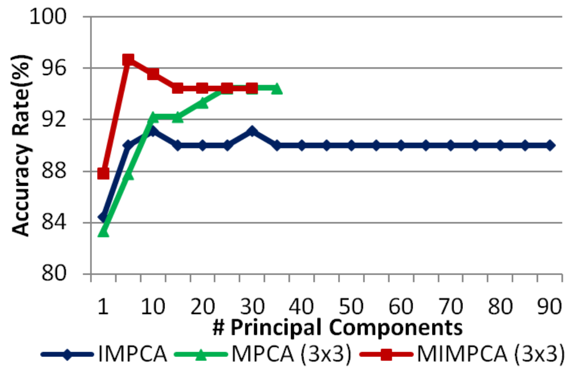

Fig. 4. Result of the three techniques performed over Yale face database

Figure 4 shows the performance of the techniques for the Yale face database. In this figure, it can be verified that MIMPCA shows a better result than IMPCA and MIMPCA for face variations like facial expression and illumination, which are characteristics explored in the Yale face database. In low dimensionality the proposed technique is more accurate than the original techniques but decreases its performance in high dimensional representations. In the MIMPCA technique, the first principal components store great part of the data information. Therefore, good precision is expected to be reached in low dimensionality. For these experiments, the best face recognition rate was obtained using nine sub-images per face image for the MPCA and MIMPCA techniques.

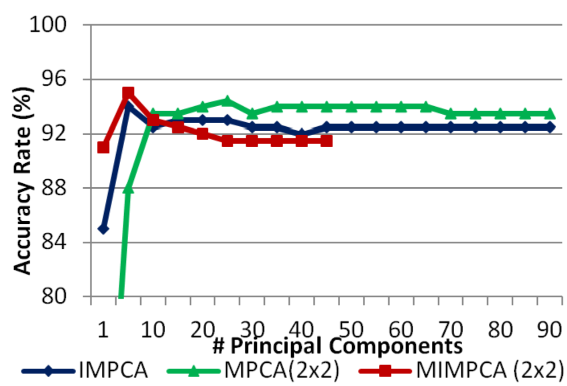

Fig. 5. Result of the three techniques performed over ORL face database

The results obtained over the ORL face database are shown in Figure 5. Under head pose and size variations, MIMPCA was slightly better than the other techniques in low dimension. When more than ten principal components were used, the performance decreased. However, it is important to highlight that the best recognition rate for the ORL face database was reached by the MIMPCA technique.

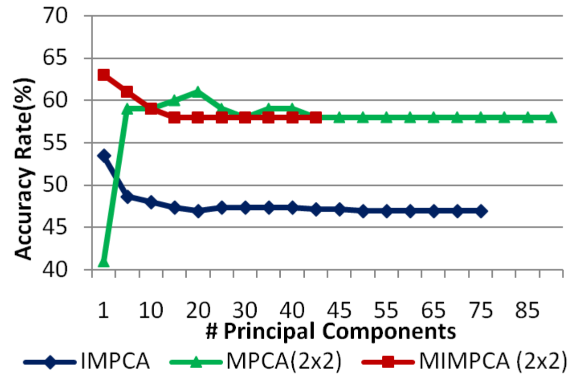

Fig. 6. Result of the three techniques performed over UMIST face database

Finally, the results using the UMIST face database are presented in Figure 6. The angles of the head pose in the test images are very different from the samples used to train the system. Due to the fact that variation on head pose angle is a hard problem for face recognition, it is expected that the accuracy rate should be slightly worse in these experiments. Once more, the MIMPCA technique reached the best performance for a low dimensionality. In Figure 6 the y-axis was zoomed to 40 to 70% scale for better result visualization.

Table I shows the best result of each technique analyzed in this paper applied over all databases. The best results are shown in italic. And the proposed technique achieved better results than MPCA and IMPCA.

TABLE I

BEST RECOGNITION RATES

	MPCA	**IMPCA**	**MIMPCA**
UMIST	63.00(10/09)	62.00(06/−)	*65.00*(02/09)
ORL	94.44(25/04)	93.00(05/−)	*95.00*(05/04)
Yale	94.44(25/09)	91.11(30/−)	*96.67*(05/09)

It is possible to see in Table II that the feature extraction implemented by MIMPCA takes much less time than the other ones. As the number of training samples per class is increased, the relative gain between the MIMPCA and the other methods becomes more evident.

TABLE II

AVERAGE TIME PROCESSING (SECONDS)

	MPCA	**IMPCA**	**MIMPCA**
Time (s)	443	34	19

Analyzing the experiments the proposed technique was better than MPCA and IMPCA for all the performed tests; in terms of accuracy rate and computational time. However, it was not more efficient in terms of storage requirements. One simple strategy to reduce this storage requirement is to use PCA for further data reduction after MIMPCA.

V. Conclusion

The Modular Image PCA method, which is a combination of the Modular PCA and Two-Dimensional PCA methods, was proposed as a feature extraction procedure to be applied in the problem of face recognition. For this application, the MIMPCA method performed better than the original methods. The database used contains several facial images under different conditions of varying facial expression, head pose and illumination.

The proposed technique is based on a modular approach which takes advantage of the regions in the face images that are not affected by local variation such as illumination, facial expression or head pose. Furthermore, for each region in the image, the two-dimensional PCA approach is applied aiming to extract more representative set of weights from the original data. Another advantage of the two-dimensional approach is that it reduces the size of the image representation and subsequent minimizing the computational cost.

However, the proposed technique is less efficient than the traditional PCA in terms of storage requirements because of the number of coefficients necessary to represent each principal component of the projected data. Each component is represented as a vector of coefficients and not just as a single scalar. Nevertheless, this storage limitation can be minimized by using a dimensionality reduction technique that can even be the traditional PCA.

Hence, MIMPCA is a good alternative to improve the recognition rate and to reduce the response time of face recognition systems.

Acknowledgments

This work was supported in part by the Brazilian National Research Council CNPq (Proc. 475911/2008-3) and by FACEPE - Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (Proc. APQ-0890-1.03/08).

References

[1] L. Wang, X. Wang, X. Zhang, and J. Feng, "The equivalence of two-dimensional PCA to line-based PCA," *Pattern Recognition Letters*, Vol.26, No.1, pp. 57-60, 2005.

[2] R. Gottmukkal and V. K. Asari, "An improved face recognition technique based on modular PCA approach," *Pattern Recognition Letters*, Vol.24, No.4, pp. 429-436, 2004.

[3] J. Yang, D. Zhang, A. F. Frangi, and J. 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.

[4] 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.

[5] M. E. Wall, A. Rechtsteiner, and L. M. Rocha, "Singular value decomposition and principal component analysis," *A Practical Approach to Microarray Data Analysis*, D.P. Berrar, W. Dubitzky, M. Granzow, eds. pp. 91-109, Kluwer: Norwell, MA (2003).

[6] M. Turk and A. Pentland, "Eigenfaces for Recognition," *Journal of Cognitive Neuroscience*. Vol.3, No.1, pp. 71-86, 1991.

[7] M. S. Barlett, 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.

[8] M. H. Yang, "Kernel Eigenfaces vs. Kernel Fisherfaces: Face Recognition Using Kernel Methods," *IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 215-220, 2002.

[9] D. Masip and J. Vitria, "Boosted discriminant projections for nearest neighbor classification," *Pattern Recognition*, Vol.39, No.2, pp. 164-170, 2006.

[10] D. Masip and J. Vitria, "Shared Feature Extraction for Nearest Neighbor Face Recognition," *IEEE Transactions on Neural Network*, Vol.19, No.4, pp. 584-596, 2008.

[11] J. Yang and J. Y. Yang, "From image vector to matrix: A straightforward image projection technique - IMPCA vs. PCA," *Pattern Recognition Letters*, Vol.35, No.9, pp. 1997-1999, 2002.

[12] S. Chen and Y. Zhu, "Subpattern-based principle component analysis," *Pattern Recognition Letters*, Vol.37, No.5, pp. 1081-1083, 2004.

[13] K. Tan and J. Yang, "Adaptively weighted sub-pattern PCA for face recognition," *Neurocomputing*, Vol.64, pp. 505-511, 2005.