OCCLUSION INVARIANT FACE RECOGNITION USING TWO-DIMENSIONAL PCA

Tae Young Kim
Samsung Electronics Co. Ltd.
Suwon, Korea

Kyoung Mu Lee and Sang Uk Lee
School of Electrical Eng., ASRI, Seoul National University
151-600, Seoul Korea

Keywords: Face recognition, occlusion invariance, 2DPCA.

Abstract: Subspace analysis such as Principal Component Analysis(PCA) and Linear Discriminant Analysis(LDA) are widely used feature extraction methods for face recognition. However, most of them employ holistic basis so that local parts can not be efficiently represented in the subspace. Therefore, they cannot cope with occlusion problem. In this paper, we propose a new method using two-dimensional principal component analysis (2D PCA) for occlusion invariant face recognition. In contrast to PCA, 2D PCA is performed by projecting 2D image directly onto the 2D PCA subspace, and each row of feature matrix represents the distribution of corresponding row of the image. Therefore by classifying each row of the feature matrix independently, we can easily identify the locally occluded parts in the face image. The proposed occlusion invariant face recognition system consists of two steps: occlusion detection and partial matching. To detect occluded regions, we apply a new combined k-NN and 1-NN classifier to each row or block of the feature matrix of the test face. For partial matching, similarity between feature matrices is evaluated after removing the rows identified as the occluded parts. The experimental results on AR face database demonstrate that the proposed algorithm outperforms other existing approaches.

1 INTRODUCTION

Face recognition has been one of the most challenging and active research topics in computer vision for several decades (Zhao, 2000). The goal of face recognition is to identify one or more persons, given still or video scenes using stored faces in a database. Face recognition system should recognize a face robustly and independently as possible to the image variations such as pose, illumination changes, expression, and occlusion.

Face recognition approaches can be divided into two categories: feature based methods (Gao, 2002)-(Park, 2005) and appearance based methods (Turk, 1991)-(Georghiades, 2001). In feature based methods, some features such as eyes, nose, and mouth are extracted, and the geometrical relationship between them are analyzed for the recognition. This approach has advantages such as low memory requirement and robustness to illumination changes. However, during the process of extracting low-level features, some distortions may arise. On the other hand, in appearance based methods, the holistic intensity information of a face image is represented in terms of principal modes on a compact low dimensional subspace (Turk, 1991)-(Belhumeur, 1997). Appearance based approaches are known to be sensitive to illumination changes and needs more memory than feature based approach. Up to now, many techniques have been introduced using these two approaches. Still, if there are occluded parts on the face image, recognition rate remains relatively low. Occluded faces wearing sunglasses or scarfs are examples of partial information loss. These damaged regions usually degrade the performance of face recognition system severely. Recently, some methods for reconstructing partially damaged face have been developed (Saito, 1999) (Hwang, 2003). They reconstructed damaged regions by interpolation or extrapolation using linear subspace analysis or a morphable face model. Also, occlusion problem can be handled by partial matching after detecting and removing the lost features without direct reconstruction of the lost information. Leonardis *et al.* (Leonardis, 1996) rejected outliers and dealt with occlusions through a hypothesize-and-test paradigm using subsets of image points. On the other hand, Black *et al.*(Black, 1998) calculated the coefficients by means of a conventional robust M-estimator to eliminate outliers. However,

Young Kim T., Mu Lee K. and Uk Lee S. (2006).
OCCLUSION INVARIANT FACE RECOGNITION USING TWO-DIMENSIONAL PCA.
In *Proceedings of the First International Conference on Computer Vision Theory and Applications*, pages 56-61
DOI: 10.5220/0001372500560061

these methods either need extensive training images or must satisfy some prior conditions, so they are not easily applicable for general situation.

In this paper, we propose a novel occlusion invariant face recognition method based on 2D PCA technique (Yang, 2004). In a subspace obtained after a transformation by 2D PCA, we can detect occluded parts by applying a combined k-NN (Earl, 1996) and 1-NN (Dick, 1998) classifier, that consider the relative distance from test data and its nearest neighbor, and conduct partial matching only on the non-occluded parts after eliminating occlusion effect. The proposed algorithm recognizes a face by excluding unreliable and inconsistent features without estimation or reconstruction of the lost information. In addition, unlike most other algorithms, our algorithm requires only a single training face image per person; it can be applied to more general situations where many training samples are not available.

2 TWO-DIMENSIONAL PCA

In order to perform the conventional PCA for face recognition, the 2D face image must be transformed to 1D vectors in advance. The resulting image vectors of faces usually lead to a high-dimensional vector space (Turk, 1991). Contrast to the conventional PCA, 2D PCA uses 2D matrices directly rather than 1D vectors. That is, the image matrix does not need to be transformed into a vector. Also, the image covariance matrix can be constructed directly using the original image matrices, and the size of it is much smaller than that of PCA. The details of 2D PCA can be found in (Yang, 2004).

Let $\mathbf{X}$ denote an n-dimensional unit column vector. The main idea of 2D PCA is to project image $\mathbf{A}$, an $m \times n$ random matrix, onto $\mathbf{X}$ by the following linear transformation.

$$\mathbf{Y} = \mathbf{A}\mathbf{X}. \tag{1}$$

Thus, we obtain an m-dimensional projected vector $\mathbf{Y}$, which is called the projected feature vector of image $\mathbf{A}$. To make a performance of 2D PCA better, we have to determine a good projection vector $\mathbf{X}$. In fact, the total scatter of the projected samples can be introduced to measure the discriminative power of the projection vector $\mathbf{X}$. Moreover, the total scatter of the projected samples can be characterized by the trace of the covariance matrix of the projected feature vectors. Therefore, by maximizing the total scatter of the projected samples, we can determine a good projection vector $\mathbf{X}$. The physical significance of a good projection vector is to find a projection direction $\mathbf{X}$, onto which all samples are projected, so that the total scatter of the resulting projected samples is maximized. The covariance matrix of the projected feature vectors of the training samples can be denoted by

$$\begin{aligned} \mathbf{S}_x &= E[(\mathbf{Y} - E\mathbf{Y})(\mathbf{Y} - E\mathbf{Y})^T] \\ &= E[(\mathbf{AX} - E(\mathbf{AX}))(\mathbf{AX} - E(\mathbf{AX}))^T] \\ &= E[((\mathbf{A} - E\mathbf{A})\mathbf{X})((\mathbf{A} - E\mathbf{A})\mathbf{X})^T]. \end{aligned} \tag{2}$$

Therefore,

$$tr(\mathbf{S}_x) = \mathbf{X}^T E[(\mathbf{A} - E\mathbf{A})^T(\mathbf{A} - E\mathbf{A})]\mathbf{X}. \tag{3}$$

Let us define the following *image covariance matrix*.

$$\mathbf{G}_t = E[(\mathbf{A} - E\mathbf{A})^T(\mathbf{A} - E\mathbf{A})]. \tag{4}$$

We can calculate $\mathbf{G}_t$ directly using the training image samples. Suppose that the training set contains M samples in total, the *j*-th training image is denoted by an $m \times n$ matrix $\mathbf{A}_j (j = 1, 2, \cdots, M)$, and the average image of all training samples is denoted by $\bar{\mathbf{A}}$. Then, $\mathbf{G}_t$ can be evaluated by

$$\mathbf{G}_t = \frac{1}{M}\sum_{j=1}^{M}(\mathbf{A}_j - \bar{\mathbf{A}})^T(\mathbf{A}_j - \bar{\mathbf{A}}). \tag{5}$$

Alternatively, the Eq. (3) can be expressed by

$$tr(\mathbf{S}_x) = \mathbf{X}^T\mathbf{G}_t\mathbf{X}. \tag{6}$$

The unit column vector $\mathbf{X}$ that maximizes Eq. (6 is called the optimal projection axis. This means that the total scatter of the projected samples are maximized after the projection of an image matrix onto $\mathbf{X}$, so that the discriminative power of the projection vector $\mathbf{X}$ is also maximized.

The optimal projection axis $\mathbf{X}_{opt}$ is the unit column vector that maximizes Eq. (6), *i.e.*, the eigenvector of $\mathbf{G}_t$ corresponding to the largest eigenvalue (Yang, 2002). In general, since it is not enough to have only one optimal projection axis, we usually need to select a set of projection axes, $\mathbf{X}_1, \cdots, \mathbf{X}_d$, satisfying the following criterion,

$$\{\mathbf{X}_1, \cdots, \mathbf{X}_d\} = \arg\max tr(\mathbf{S}_x). \tag{7}$$

In fact, the optimal projection axes, $\mathbf{X}_1, \cdots, \mathbf{X}_d$, are the eigenvectors of $\mathbf{G}_t$ corresponding to the first d largest eigenvalues.

After finding the optimal projection axes of 2D PCA, they are used for feature extraction. For a given image sample $\mathbf{A}$, we can obtain the following *principal component vectors*

$$\mathbf{Y}_k = \mathbf{A}\mathbf{X}_k, k = 1, 2, \cdots, d. \tag{8}$$

These principal component vectors are used to form an $m \times d$ matrix $\mathbf{B} = [\mathbf{Y}_1, \cdots, \mathbf{Y}_d]$ called the *feature matrix*, which characterize the image sample $\mathbf{A}$ in the 2D PCA space.

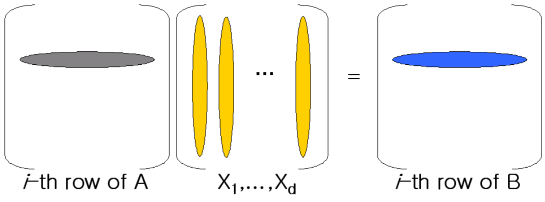

Figure 1: Matrix multiplication process of 2D PCA.

3 OCCLUSION DETECTION AND PARTIAL MATCHING USING 2D PCA

3.1 Occlusion Detection

One interesting property of 2D PCA is that each row of an image $\mathbf{A}$ is projected onto the optimal projection axes and produces a corresponding row of the feature matrix (Fig. 1), *i.e.*, the *i*-th row of a feature matrix represents the projection of the *i*-th row of the image in 2D PCA subspace. Therefore, by analyzing each row of the feature matrix statistically and independently, we can identify the occluded rows or local regions in the image.

The process of occlusion detection is a type of one-class classification problem which discriminates a face region from non-face ones. In this case, a face regions belong to a target class and the occluded face regions belong to an outlier class. One-class classification techniques can be categorized into two types (Dick, 1998). One is the *unsupervised* classifiers that use only the samples of the target class for training. The other is the *supervised* classifiers that employ the sample training objects of both target and outlier classes. Although it needs additional efforts for providing outlier samples during the training process, generally the supervised method gives better result than unsupervised one. Therefore, after obtaining the distributions of every row of both normal and occluded face images' feature matrices, we can apply a supervised one class classifier to each row for the detection of occluded face parts.

In this paper, we developed and used a new supervised one class classifier which combines the k-NN and a modified 1-NN classifier (Dick, 1998) sequentially. Note that usually the k-NN classifier shows good performance for one class classification. However, since no distance constraint is considered, misclassification may occur when a test data is located far away from the training samples and most of the nearest neighbors belong to a target class. In this case, the test data is assigned to the target class even though it is an outlier object. In order to resolve this problem, we employ the relative distance-based 1-NN classifier

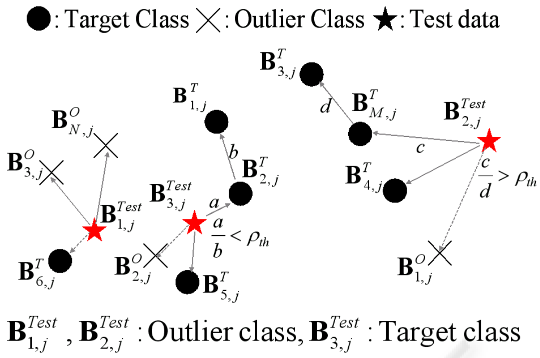

Figure 2: Occlusion detection using a combined k-NN and 1-NN classifier.

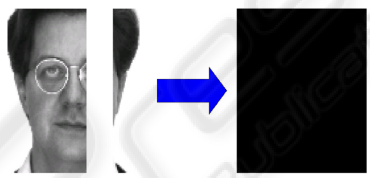

Figure 3: The problem of occlusion detection by row.

(Dick, 1998) for a post verification for the decision of target class.

The proposed classifier works as follows: For a test data x, k-NN classifier first seeks the k nearest samples among the training samples. Among these k closest samples, if the number of outlier class samples is more than that of target class samples, the test data is classified as an outlier object. Otherwise, we apply 1-NN classifier to the target class samples only by using the relative distance from x to its first nearest neighbor in the training set defined by

$$\rho_{NN}(x) = \frac{||x - NN^{tr}(x)||}{||NN^{tr}(x) - NN^{tr}(NN^{tr}(x))||}, \quad (9)$$

where $NN^{tr}(x)$ denotes the nearest neighbor of object x in the training set. If it is smaller than a prespecified threshold value, then the test data x is classified as a target object, otherwise, it is assigned to the outlier class. Let us denote $\mathbf{B}^T_i$ and $\mathbf{B}^O_i$ as the *i*-th feature matrices of training images in target class and outlier class, respectively. For a given test image $\mathbf{A}_i$, we obtain its feature matrix denoted by

$$\mathbf{B}^{Test}_i = \begin{bmatrix} \mathbf{B}^{Test}_{i,1} \\ \mathbf{B}^{Test}_{i,2} \\ \vdots \\ \mathbf{B}^{Test}_{i,m} \end{bmatrix}. \quad (10)$$

The occlusion detection is done for each row vector, $\mathbf{B}^{Test}_{i,j}, j = 1, \ldots, m$ that corresponds to image row,

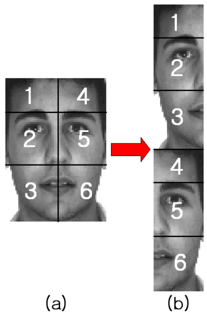

Figure 4: Image partitioning and transformation.

independently by the method described above. Fig. 2 shows the examples of the proposed classification results for some row features. Note that although the row-based occlusion detection scheme shows a good result, it cannot cope with vertical occlusions efficiently. For example, If we apply the row-based occlusion detection scheme to the occluded face image in Fig. 3, the entire image is considered to be occluded since all the rows contain occluded regions. To overcome this horizontal localization problem, we divide a face image into two columns and concatenate them into a single one as shown in Fig. 4. After transforming all images in this way and performing the training with these transformed images, we apply a combined k-NN and 1-NN classifier to each row of the input feature matrix. Alternative method is to partition the column into several blocks as shown in Fig. 4 (b), and investigate the block-based occlusion by analyzing corresponding features of each block.

3.2 Partial Matching

We perform partial matching of faces using the corresponding feature matrices after removing the rows detected as occluded regions. Fig. 5 shows the idea. If feature matrices $\mathbf{B}_1$ and $\mathbf{B}_2$ are matched directly, the matching result may not be reliable due to the effect of the occluded part. While, by using $\mathbf{B}_3$ and $\mathbf{B}_4$ obtained after removing the occluded rows, we can achieve occlusion invariant matching results.

The dissimilarity measure between a target feature matrix $\mathbf{B}_i^T = [\mathbf{B}_{i,1}^T, \cdots, \mathbf{B}_{i,m}^T]^T$ and a test feature matrix $\mathbf{B}_j^{Test} = [\mathbf{B}_{j,1}^{Test}, \cdots, \mathbf{B}_{j,m}^{Test}]^T$, is defined by

$$d(\mathbf{B}_i^T, \mathbf{B}_j^{Test}) = \sum_{k=1}^{m} \omega_k \left|\left|\mathbf{B}_{i,k}^T - \mathbf{B}_{j,k}^{Test}\right|\right|_2, \quad (11)$$

where

$$\omega_k = \begin{cases} 0 & \text{if } \mathbf{B}_{j,k}^{Test} \text{ is an occluded row} \\ 1 & \text{otherwise} \end{cases} \quad (12)$$

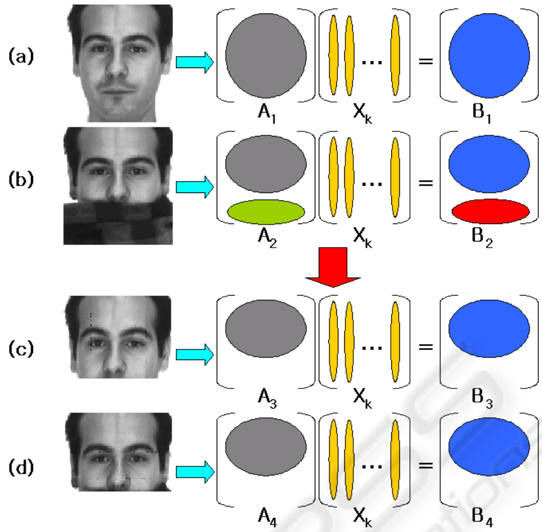

Figure 5: Partial matching after removing occlusion.

Figure 6: The result of occlusion detection by row.

and $||\mathbf{B}_{i,k} - \mathbf{B}_{j,k}||_2$ denotes the Euclidean distance between the two k-th rows of feature matrices $\mathbf{B}_i$ and $\mathbf{B}_j$.

4 EXPERIMENTAL RESULTS

4.1 Experimental Environment

To evaluate the performance of the proposed algorithm, we tested it on AR face database (AR, 1998). Specifically, we used neutral frontal images and occluded face images wearing sunglasses or scarfs. The performance of the proposed algorithm has been compared to those of the conventional approaches including 1-NN, Eigenface (Turk, 1991), NMF method (Lee, 1999), and modified local NMF (LNMF) (Li, 2001) method. In addition to these methods, it has been also compared to those of the LEM based

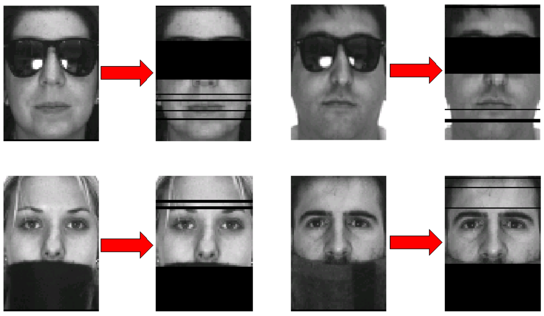

Figure 7: Some false alarms.

method (LEM) (Gao, 2002) by Gao *et. al.*, the technique proposed by Martinez(AMM) (Martinez, 2002), and Park's face-ARG technique (Park, 2005). Similar to our method, these three methods used a single frontal view image per person as a reference model, and the performances on the AR face database were reported in (Gao, 2002), (Martinez, 2002), and (Park, 2005), respectively. For comparison, we have referred their recognition rates. All 135 people (76 men and 59 women) in the AR face database were used. Among these, all 135 normal face images and 70 occluded face images (35 sunglass images and 35 scarf images of 20 men and 15 women) were used for the training the target class and the outlier class, respectively. The remaining 100 sunglass images and 100 scarf images were used for probes and all the normal frontal faces were used for the gallery.

4.2 Occlusion Detection and Partial Matching Results

Fig. 6 and 7 show some results of occlusion detection by the proposed combined k-NN and 1-NN classifier to every row of feature matrix, in which the occluded rows are displayed by black lines. For most figure we can detect occlusions accurately. However, there are some false alarms as shown in Fig. 7. Fig. 8 shows the results of the block-based occlusion detection with 6 regions, and Fig. 9 presents the results of occlusion detection to each individual row of the transformed images in Fig. 4 (b). We observed that the 'row after transformation' method gave the best detection results.

Face recognition test was conducted using the proposed algorithm explained in section 3, with the dissimilarity measure defined in Eq. (11). The recognition results of the proposed algorithm are compared to other algorithms and summarized in Table 1. Form these results, we can conclude that the proposed partial face recognition algorithm outperforms the conventional face recognition techniques.

Figure 8: The result of occlusion detection by region.

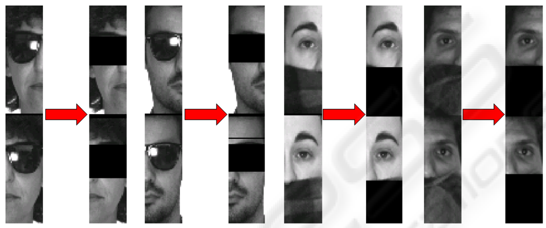

Figure 9: The result of occlusion detection by row after image transformation.

4.3 Classifier Test to Synthetic Occlusions

Note that the occlusion patterns in AR database are limited to sunglasses and scarfs. Therefore, we have tested the proposed occlusion detection algorithm to other types of occlusions. Fig 10 (a) shows the results for the synthetic white occlusion masks, and Fig 10 (a) presents those for the occlusion masks generated by random noise. These experimental results demonstrate that the new combined k-NN and 1-NN classifier can work satisfactory to other types of occlusion patterns.

5 CONCLUSION

In this paper, we proposed a novel occlusion invariant face recognition algorithm using 2D PCA. In 2D PCA subspace, a face is described by a feature matrix. Therefore, by finding occluded parts by every row in feature matrix accurately, we are able to remove severe distortions caused by occlusions. Since the proposed algorithm can detect and exclude unreliable and inconsistent parts by combining k-NN and 1-NN classifier sequentially, it recognizes a face very accurately. The performance of the proposed algorithm has been tested on the AR face database. The results show that for the faces with occlusions by sunglasses or scarfs, the proposed algorithm produces more robust and reliable results over other existing methods.

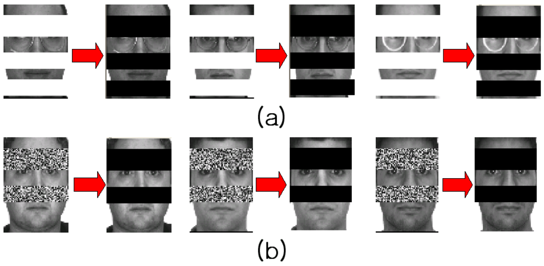

Figure 10: Classifier test to virtual occlusions.

Table 1: The recognition rate under occlusion on the AR face database : Proposed method (a) occlusion detection by row, (b) occlusion detection by 6 regions, and (c) occlusion detection by row after image transformation as shown in Fig. 4 (b).

Detection Method	Sunglasses	Scarfs
Proposed Method (a)	98.00%	99.00%
Proposed Method (b)	96.00%	98.00%
Proposed Method (c)	98.00%	98.00%
1-NN	43.18%	20.45%
PCA	43.18%	20.45%
NMF	25.00%	2.27%
LNMF	43.18%	13.64%
AMM	80.00%	82.00%
LEM	68.18%	63.64%
Face-ARG	73.48%	87.88%

ACKNOWLEDGEMENTS

This work has been supported in part by the 3DRC (3D Display Research Center) under the ITRC (Information Technology Research Center) program of Korean government.

REFERENCES

Zhao, W. Y., Chellappa, R., Rosenfeld, A. and Phillips, P. J. (2000). Face Recognition : A Literature Survey. In *UMD CfAR Technical Report CAR-TR-948*.

Gao, Y. and Leung, M. K. H. (2002). Face Recognition Using Line Edge Map. In *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.24, no.6, pp.764-779.

Park, B. G., Lee, K. M. and Lee, S. U. (2005). A Novel Face Recognition Technique Using Face-ARG Matching. In *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol 27, no. 12, pp.1982-1988.

Turk, M., Pentland, A. (1991). Eigenfaces for Recognition. In *Journal of Cognitive Neuroscience*, vol.3, pp.71-86.

Belhumeur, P. N., Hepanha, J. P. and Kriegman, D. J. (1997). Eigenfaces vs. Fisherfaces : Recognition Using Class Specific Linear Projection. In *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.19, no.7, pp.711-720.

Georghiades, A. S., Belhumeur, P. N. and Kriegman, D. J. (2001). From Few to Many : Illumination Cone Models for Face Recognition under Variable Lighting and Pose. In *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.23, no.6, pp.643-660.

Saito, Y., Kenmochi, Y. and Kotani, K. (1999). Estimation of eyeglassless facial images using principal component analysis. In *IEEE International Conference on Image Processing*, vol.4, pp.192-201.

Hwang, B. W. and Lee, S. W. (2003). Reconstruction of partially damaged face images based on a morphable model. In *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.25, no.3, pp.365-372.

Leonardis, A. and Bischof, H. (1996). Dealing with Occlusions in the Eigenspace Approach. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*.

Black, M. and Jepson, A. (1998). Eigentracking : Robust matching and tracking of articulated objects using a view-based representation. In *International Journal of Computer Vision*, vol.26, no.1, pp.63-84.

Yang, J., Zhang, D., Frangi, A. F. and Yang, J. Y. (2004). Two-Dimensional PCA : A New Approach to Appearance-Based Face Representation and Recognition. In *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.26, no.1.

Gose, E., Johnsonbaugh, R. and S. Jost (1996). *The Book. Pattern Recognition and Image Analysis*, Prentice Hall.

Yang, J., Yang, J. Y. (2002). From Image Vector to Matrix : A Straightforward Image Projection Technique-IMPCA vs. PCA. In *Pattern Recognition*, vol.35, no.9, pp.1997-1999.

Ridder, D., Tax, D. M. J. and Duin, R. P. W. (1998). An Experimental Comparison of One-Class Classification Methods. In *Proceedings of the Fourth Annual Conference of the Advanced School for Computing and Imaging*, Delft.

Martinez, A. M. and Benavente, R. (1998). The AR Face Database. In *CVC Technical Report*, no.24.

Lee, D. D. and Seung, H. S. (1999), Learning the parts of objects by non-negative matrix factorization. In *Nature*, vol.401, pp.788-791.

Li, S. Z., Hou, X. W., Zhang, H. J. and Cheng, Q. S. (2001). Learning spatially localized, part-based representation. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp.207-212.

Martinez, A. M. (2002). Recognizing Imprecisely Localized, Partially Occluded, and Expression Variant Faces from a Single Sample per Class. In *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.24, no.6, pp.748-763.