TWO DIMENSIONAL (2D) SUBSPACE CLASSIFIERS FOR IMAGE RECOGNITION

Hakan Cevikalp, Hasan Serhan Yavuz, and Atalay Barkana

Electrical-Electronics Engineering Department, Eskisehir Osmangazi University & Anadolu University
Eskisehir Osmangazi Universitesi, 26480, Eskisehir, Turkey
phone: + (90) 222 2393750, fax: + (90) 222 2290535, email: hakan.cevikalp@gmail.com
web: www.ogu.edu.tr

ABSTRACT

The Class-Featuring Information Compression (CLAFIC) is a pattern classification method which uses a linear subspace for each class. In order to apply the CLAFIC method to image recognition problems, 2D image matrices must be transformed into 1D vectors. In this paper, we propose new subspace classifiers to apply the conventional CLAFIC method directly to the image matrices. The proposed methods yield easier evaluation of correlation and covariance matrices, which in turn speeds up the training and testing phases. Moreover, experimental results on the AR and the ORL face databases also show that recognition performances of the proposed methods are typically better than recognition performances of other subspace classifiers given in the paper.

1. INTRODUCTION

The subspace classifier is a pattern recognition method, where the primary model for each class is a linear subspace of the Euclidean sample space [1]. In these methods, it is assumed that the vector distribution of a class corresponds to a lower-dimensional subspace of the original sample space. Even though this assumption is seldom valid, good recognition rates can be achieved when the dimensionality of the sample space is large enough [2]. The subspaces representing classes are defined in terms of basis vectors that are linear combinations of the sample vectors of each class. Once the basis vectors spanning those subspaces are computed, a test sample vector from an unknown class is classified based on the lengths of the projections of that sample onto each of the subspaces or, alternatively, on the distances of the test vector from these subspaces.

Watanabe *et al*. proposed the first subspace method, the Class-Featuring Information Compression (CLAFIC), for pattern classification [3]. This method employs the Principal Component Analysis (PCA) to compute the basis vectors spanning subspace of each class. Fukunaga and Koontz proposed a new method, which enabled to select the basis vectors in such a way that the projections onto the so-called rival subspaces are minimized [4]. Gulmezoglu *et al*. proposed the Common Vector (CV) method for classification tasks, where the number of samples in each class is smaller than or equal the dimensionality of the sample space [5]. Then, learning subspace methods, in which the subspaces are iteratively modified in order to diminish the number of misclassifications, have been proposed in [6]. Recently, the kernel based subspace methods, the Kernel CLAFIC [7] and the Kernel CV [8], have been proposed to extract nonlinear features of classes.

In order to apply the subspace methods to image recognition problems, 2D image matrices must be transformed into 1D vectors by concatenating rows or columns. The resulting image vectors typically lead to a high-dimensional sample space, which in turn forms a suitable environment for application of subspace classifiers. It is because most of the assumptions, upon which the subspace classifiers are founded, hold in high-dimensional sample spaces. However, some of the subspace classifiers cannot be applied in these high-dimensional sample spaces since they require the use of large correlation matrices or orthogonal projection operators explicitly. Fortunately, Yang *et al*. introduced a new method, coined the 2D-PCA method, which applies the classical PCA method directly to image matrices [9]. This procedure leads to easier evaluation of covariance matrices since the size of the image covariance matrices using 2D-PCA is much smaller. Additionally, it has been reported that the recognition performance of 2D-PCA is superior to the classical PCA (Eigenface) method [9]. In this paper, motivated by this technique, we propose new subspace classifier methods, which will be referred to as the 2D-CLAFIC and the 2D-CLAFIC-μ, in order to apply the classical CLAFIC method directly to the image matrices.

The remainder of the paper is organized as follows: In Section 2, we first review the CLAFIC method and its variant the CLAFIC-μ, and then describe our proposed methods. In Section 3, experimental results are given. Finally, our conclusions are presented in Section 4.

2. 2D SUBSPACE CLASSIFIERS

Before we introduce our proposed methods, we will first review the classical CLAFIC method and its variant, the CLAFIC-μ, briefly.

2.1 The CLAFIC Method

Suppose there are C classes denoted by $\omega^{(1)}, \omega^{(2)}, ..., \omega^{(C)}$ where the *i*-th class contains N_i samples. Let $x_j^i \in \Re^d$ be a *d*-dimensional column vector, which denotes the *j*-th sample of the *i*-th class. Let $L^{(1)}, L^{(2)}, ..., L^{(C)}$ are the subspaces representing classes. Each subspace is spanned by l_i orthonormal basis vectors $\{w_1^i, ..., w_{l_i}^i\}$ in $\Re^d$.

The CLAFIC method employs the PCA or the Karhunen-Loeve transform to compute the basis vectors $\{w_1^i,...,w_{l_i}^i\}$ spanning each subspace $L^{(i)}$. The basis vectors are computed through eigen-decomposition of class correlation matrices $R^{(i)} \in \Re^{d\times d}$ defined as

$$R^{(i)} = \frac{1}{N_i}\sum_{j=1}^{N_i} x_j^i (x_j^i)^T = \frac{1}{N_i}\Phi^{(i)}\Phi^{(i)^T}, \qquad i=1,...,C, \qquad (1)$$

where $\Phi^{(i)}$ is the matrix whose columns are the sample vectors of the *i*-th class. Note that the mean vectors μ_i of classes are not subtracted. The correlation matrix $R^{(i)}$ is a positive semi-definite matrix, hence all eigenvalues are larger than or equal to 0. The l_i eigenvectors corresponding to the largest eigenvalues of $R^{(i)}$ are chosen as basis vectors for the subspace $L^{(i)}$. The number of basis vectors determines the dimensionality of each subspace. In the CLAFIC method, the number of basis vectors cannot exceed $\min(d, N_i)$ for each class. There are different strategies to choose the subspace dimensions l_i. One way is to set all l_is to be equal to a fixed value *l*. Then, the optimal value of *l* can be chosen from the error curves as described in [2]. The other way employs eigenvalues for choosing the dimensions of subspaces. Let the eigenvalues of $R^{(i)}$ are ordered as

$$\lambda_1^i \geq \lambda_2^i \geq ... \geq \lambda_{r_i}^i > 0, \qquad (2)$$

where r_i is the rank of the matrix $R^{(i)}$. The dimension of $L^{(i)}$ is selected as the value by which the ratio of cumulative sums $\kappa_i = \sum_{j=1}^{l_i}\lambda_j^i / \sum_{j=1}^{r_i}\lambda_j^i, i=1,...,C,$ exceeds a threshold. Typical values of the threshold lie between $0.9 \leq \kappa_i \leq 1$. To classify a test sample vector, x_{test}, the squared norms of vectors for each class are found by

$$\| y_{test}^i \|^2 = \| W^{i^T} x_{test} \|^2, \quad i=1,...,C, \qquad (3)$$

and then the test sample is assigned to the class, in which it has the maximum norm value. Here W^i represents the transformation matrix whose columns are the basis vectors of the *i*-th class.

A variation of the CLAFIC, which is called the CLAFIC-μ method, uses the class-specific means μ_i in classification [2]. In this approach each class is modelled as a linear manifold centred at the mean of the corresponding class. Therefore, instead of using class correlation matrices, class covariance matrices are employed to compute the basis vectors. In this case, above classification rule, which is based on maximizing the projection length, cannot be used. Instead, the minimum distance of centred test vector from these subspaces determines class labels. In particular, we compute the distance

$$D_i = \| x_{test} - \mu_i \|^2 - \| W^{i^T}(x - \mu_i) \|^2, \quad i=1,...,C, \quad (4)$$

for each class, and the test sample is assigned to the class which gives the minimum distance.

2.2 2D-CLAFIC Methods

The application of the CLAFIC method to image recognition problems involves the transformation of original 2D image matrix data into 1D vectors. This transformation is usually performed by lexicographic ordering of image matrices into column vectors. As opposed to the conventional CLAFIC methods, our proposed methods utilize image matrices instead of image vectors. We employ the 2D-PCA method to compute the basis vectors spanning the subspaces of classes. Since 2D matrix form is preserved in calculations, the proposed approach is called as the 2D-CLAFIC method.

Assume that the image recognition problem consists of *C* classes denoted by $\omega^{(1)}, \omega^{(2)},...,\omega^{(C)}$ where the *i*-th class contains N_i samples. Let $X_j^i \in \Re^{m\times n}$ denote the *j*-th image matrix of the *i*-th class. In this case, each correlation matrix $R^{(i)} \in \Re^{n\times n}$ of the image matrices is defined as

$$R^{(i)} = \frac{1}{N_i}\sum_{j=1}^{N_i}(X_j^i)^T X_j^i, \quad i=1,...,C. \qquad (5)$$

Note that the size of the image correlation matrix, $n \times n$, is much smaller than the size of the correlation matrix, $d \times d$, obtained using the classical CLAFIC method since $d = mn$. The image correlation matrix $R^{(i)}$ is a positive semi-definite matrix, hence all eigenvalues are larger than or equal to 0. The most significant eigenvectors of $R^{(i)}$ are chosen as the basis vectors. The number of basis vectors will be limited by the number of columns of image matrices, which is equal to *n* for each class. After basis vectors of subspaces are obtained, to classify a test image matrix, X_{test}, we project the image matrix onto the basis vectors of classes to get the image feature matrix

$$Y_{test}^i = X_{test} W^i, \quad i=1,...,C. \qquad (6)$$

For the classification of X_{test}, we compute the *Frobenius* norms of feature matrices by

$$\| Y_{test}^i \|_F = \sqrt{\sum_{j=1}^{m}\sum_{k=1}^{l_i} | y_{jk} |^2}, \quad i=1,...,C, \qquad (7)$$

for each class. Then the test sample is assigned to the class, in which it has the maximum norm value.

Similar to the previous case mentioned in section 2.1, we can utilize the class-specific means in this approach. We obtain image covariance matrix of each class by using the following equation:

$$\Sigma^{(i)} = \frac{1}{N_i}\sum_{j=1}^{N_i}(X_j^i - \overline{X}_i)^T (X_j^i - \overline{X}_i), \quad i=1,...,C, \qquad (8)$$

where $\overline{X}_i$ represents the mean image matrix of the *i*-th class. The most significant eigenvectors corresponding to the largest eigenvalues of each image covariance matrix are used as basis vectors representing classes. Once the basis vectors are obtained, we compute the distance of the centred test image from each subspace by

$$D_i = \| X_{test} - \overline{X}_i \|_F - \| (X_{test} - \overline{X}_i) W^i \|_F, \quad i=1,...,C, \quad (9)$$

and we assign the test sample image to the class which minimizes this distance. We call this procedure as the 2D-CLAFIC-μ method.

3. EXPERIMENTAL RESULTS

We performed our experiments on two well-known face image databases, namely the AR [10] and the ORL (Olivetti-Oracle Research Lab) face [11] databases. The AR face database was employed to evaluate the recognition performances of the proposed methods under conditions where there is a variation over time, in facial expressions, and in lighting conditions, whereas the ORL face database was used to examine the recognition performances of the proposed methods under conditions where the pose is varied. Beside the proposed 2D-CLAFIC and 2D-CLAFIC-μ methods here, we tested CLAFIC and CLAFIC-μ subspace classifier methods for comparison. In addition, we also tested PCA and 2D-PCA methods for a better assessment of the recognition performances of our proposed methods. To determine subspace dimensions of the CLAFIC and the 2D-CLAFIC based methods, we set all subspace dimensions to be equal to a fixed value l. Then, the optimal value of l was chosen from the error curves. For the PCA and 2D-PCA methods, the most significant eigenvectors, which were used for feature extraction, were chosen such that corresponding eigenvalues contained 98% of the total energy. Then, the nearest-neighbor algorithm was employed during classification for these methods.

3.1 Experiments on the AR Face Database

The AR face database includes 26 images with different facial expressions, illumination conditions, and occlusions for 126 subjects. All individuals are in an upright, frontal position. Images were recorded in two different sessions 14 days apart. Thirteen images were recorded under controlled circumstances in each session. The size of the images in the database is 768x576 pixels, and each pixel is represented by 24 bits of RGB color values.

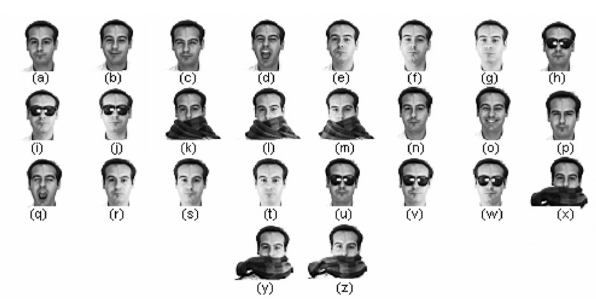

Figure 1: Images of an individual from the AR face database.

We randomly selected 50 individuals (30 males and 20 females) for the experiment. Only nonoccluded images ((a)-(g) and (n)-(t) as in Fig. 1) were chosen for every subject. Thus, our face database size was 700 with 14 images per subject. First, these images were converted to grayscale images. Second, we preprocessed these images by aligning and scaling them so that the distances between the eyes were the same for all images, and also ensuring that the eyes located in the same coordinates of the image. The resulting image was then cropped. The final size of the images was 222x299. Finally, based on empirical observations, we decreased the dimensionality of the sample space to 99x134 by down-sampling. The training set consisted of 7 images randomly selected from each subject, and the rest of the images were used for the test set. This process was repeated 15 times, and the final recognition rates for the experiment were found by averaging these 15 rates obtained in each trial. The computed recognition rates and corresponding standard deviations are shown in Table I.

TABLE I
RECOGNITION RATES ON THE AR FACE DATABASE

METHODS	RECOGNITION RATES (%)
PCA	74.71, $\sigma = 1.80$
2D-PCA	86.89, $\sigma = 1.59$
CLAFIC	82.77, $\sigma = 2.19$
CLAFIC-μ	92.26, $\sigma = 1.66$
2D-CLAFIC	90.31, $\sigma = 1.50$
2D-CLAFIC-μ	**95.97**, $\sigma = 0.53$

As can be seen in the table, our proposed method, 2D-CLAFIC-μ, achieved the best recognition rate among all methods tested here. It should be noted that all 2D-based methods outperformed their 1D-based counterparts. Using class-specific means significantly improved the recognition rates since the 2D-CLAFIC-μ method yielded better recognition rates than the 2D-CLAFIC method. In general, experimental results show that our proposed methods are the best suited subspace classifiers for image recognition tasks, where there is a variation over time, in facial expression, and in lighting conditions.

Testing time is the consumed time that is required to classify a new test image. To classify a test image, we have to compute the distances given in equations (3), (7), and (9) for this application. This process involves the projection of test samples onto the basis vectors spanning each subspace. The best recognition rates were obtained by fixing subspace dimensions to values between 2 and 4 for the CLAFIC method, whereas the best recognition rates were obtained by fixing subspace dimensions to values between 2 and 6 for the 2D-based CLAFIC methods. Therefore our proposed methods are also more practical than the CLAFIC method for real-time image recognition applications since the testing complexity of our proposed methods is given by $O(nd)$ whereas the testing complexity of the CLAFIC method is $O(d^2)$, where image size is $m \times n$ and $d = mn$.

3.2 Experiments on the ORL Face Database

The ORL face database contains 40 individuals, with 10 images per person. The images are taken at different time instances with different lighting conditions (slightly), facial expressions, and facial details. Some individuals from the ORL face database are shown in Fig. 2. The size of each image is 92x112.

We selected randomly five samples from each class for training and the remaining samples were used for testing. We did not apply any preprocessing to the images. Then, recognition rates were computed and this process was repeated 15 times. The recognition rates were found by averaging the recognition rates in each run. The computed recognition rates and standard deviations are shown in Table II.

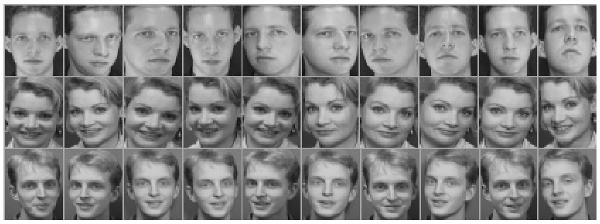

Figure 2: Images of some individuals from the ORL face database.

TABLE II
RECOGNITION RATES ON THE ORL FACE DATABASE

METHODS	RECOGNITION RATES (%)
PCA	94.03, $\sigma = 1.26$
2D-PCA	94.97, $\sigma = 1.48$
CLAFIC	93.63, $\sigma = 2.08$
CLAFIC-μ	**95.30**, $\sigma = 1.53$
2D-CLAFIC	92.50, $\sigma = 1.60$
2D-CLAFIC-μ	94.67, $\sigma = 1.21$

In terms of classification accuracy, the CLAFIC-μ method achieved the best recognition rate among all methods tested here for the ORL face database. Although the 2D-CLAFIC-μ outperformed the CLAFIC method and the 2D-PCA method outperformed the PCA method, the improvement in recognition rates is not significant. Therefore, we conclude that the 2D-based approaches do not offer a significant improvement over their 1D-based counterparts for image recognition problems where there is a variation in pose. However, it should be noted that as in the previous case, our proposed methods are better suited for real-time image recognition applications than the conventional 1D subspace classifiers because of their low computation cost.

4. CONCLUSION

In this paper we proposed new 2D subspace classifiers for image recognition problems. In contrast to the conventional subspace classifiers, our proposed methods can be applied directly to the image matrices. This process enables easier evaluation of correlation and covariance matrices, which in turn speeds up the training and testing phases. Therefore, our proposed methods are more practical than 1D based subspace classifiers for real–time image recognition tasks. In addition, experimental results demonstrated that our proposed methods are superior to other tested methods, especially when there is variation in lighting conditions. These results show that the proposed methods are robust to varying illumination conditions, which is a serious problem that PCA based feature extraction techniques cannot handle well.

REFERENCES

[1] E. Oja, *Subspace Methods of Pattern Recognition.* Research Studies Press, 1983.

[2] J. Laaksonen, *Subspace Classifiers in Recognition of Handwritten Digits*, Ph. D. thesis, Helsinki University of Technology, 1997.

[3] S. Watanabe, P.F. Lambert, C.A. Kulikowski, J.L. Buxton, and R. Walker, "Evaluation and selection of variables in pattern recognition," in Computer and Information Sciences II, pp. 91, 1967.

[4] K. Fukunaga and W. L. Koontz, "Application of the Karhunen-Loeve expansion to feature selection and ordering," *IEEE Transactions on Computers C-19(4)*, pp. 311-318, 1970.

[5] M. B. Gulmezoglu, V. Dzhafarov, and A. Barkana, "The common vector approach and its relation to principal component analysis," *IEEE Transactions on Speech and Audio Processing*, vol. 9, no. 6, September 2001.

[6] M. Kuusela and E. Oja, "The averaged learning subspace method for spectral pattern recognition," in *Proceedings of the 6th International Conference on Pattern Recognition,* Munchen, pp. 134-137, 1982.

[7] T. Balachander, and R. Kothari, "Kernel based subspace pattern classification," In *International Joint Conference on Neural Networks*, vol. 5, 3119-3122, 1999.

[8] H. Cevikalp, M. Neamtu, and A. Barkana, "A novel subspace classifier for pattern recognition," *IEEE Transaction on Systems, Man, and Cybernetics--Part B: Cybernetics*, in review.

[9] 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, January 2004.

[10]] A.M. Martinez and R. Benavente, "The AR face database," CVC Tech. Report #24, 1998.

[11] The ORL Database of Faces, AT&T Laboratories Cambridge. Available: http://www.uk.research.att.com/facedatabase.html.

[12] K. Tsuda, "Subspace classifier in the Hilbert space," *Pattern Recognition Letters*, vol. 20, pp. 513-519, February 1999.

[13] S.-W. Kim and B. J. Oommen, "On using prototype reduction schemes to optimize kernel-based nonlinear subspace methods," *Pattern Recognition*, vol. 37, pp. 227-239, 2003.