Expert Systems with Applications 40 (2013) 4971–4977

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Weighted Modular Image Principal Component Analysis for face recognition

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

Federal University of Pernambuco (UFPE), Center for Informatics (CIn), Av. Jornalista Anibal Fernandes s/n, Cidade Universitária 50740-560 Recife, PE, Brazil

ARTICLE INFO

Keywords:
Face recognition
Image feature extraction
Principal Component Analysis
Modular PCA

ABSTRACT

This paper proposes two feature extraction techniques that minimizes the effects of distortions generated by variations in illumination, rotation and, head pose in automatic face recognition systems. The proposed techniques are Modular Image Principal Component Analysis (MIMPCA) and weighted Modular Image Principal Component Analysis (wMIMPCA). Both techniques are based on PCA and they use the modular image decomposition to minimize local variation. Also, the covariance matrix is calculated directly from the original image matrix. This strategy generates a smaller matrix compared with traditional PCA and reduces the computational effort. wMIMPCA assumes that parts of the face are more discriminatory than others, so a Genetic Algorithm is used to obtain weights for each region in the face image. The proposed techniques are compared with Modular PCA and two-dimensional PCA using three well-known databases, showing better results.

1. Introduction

Automatic face classification systems aim to identify or verify the people in a given image. A general framework for face classification system is composed of the following phases: detection, feature extraction and recognition or verification. After the acquisition, the exact position of human faces in the scene is defined in the face detection phase. The output of the face detection is a set of matrices, in which each matrix contains a detected face image. The objective of the feature extraction phase is to find a new representation for each face image that: (i) has a smaller dimension than the original representation and (ii) improves or, at least, maintains the discriminatory information of the original representation. The last phase compares the input faces against the face models previously stored in a database of known faces. Recognition systems identify a person from a list of users. In contrast, verification systems confirm a person's claimed identity. Therefore, recognition is more difficult because it requires 1:N matching, while verification requires only 1:1 matching.

Face classification systems have shown interesting results when the image acquisition environments are controlled. This means that the environment has uniform illumination and the faces are captured in frontal view having small variation in rotation and some occlusion. In contrast, in uncontrolled conditions, the precision of the face classification system is drastically affected.

Among many possible alternatives to extract features from the face images, Principal Component Analysis (PCA) has been widely used (Barreto et al., 2012; Fan and Verma, 2009; Franco and Nanni, 2009; Gumus et al., 2010; Turk and Pentland, 1991; Yang, 2002; Zhao et al., 2003). However, PCA has a high sensitivity to illumination changes in the original spatial domain (Xudong and Kin-Man, 2006). Moreover, due to the holistic nature of PCA, the quality of the features is affected by different facial expression, occlusion and head poses. Some PCA-based methods have been proposed to improve face recognition rates either by minimizing variation in the acquired image or by increasing the representativeness of the data using local information. The Modular Principal Component Analysis (MPCA) (Gottmukkal and Asari, 2004) approach divides each face image into smaller regions of the same size and uses classical PCA in each region. This procedure obtains a relevant set of local features. Thus, if only part of the face is affected by changes in illumination, these regions have a restricted contribution in the final feature vector.

The two-dimensional PCA approach (2DPCA or IMPCA) (Yang et al., 2004) extracts features from the matrix of the image. Therefore, it is not required to transform each image into a one-dimensional vector as performed by the classical PCA approaches. The main idea of the 2DPCA is to construct an image covariance matrix directly using the original image matrices (Pereira et al., 2009; Pereira et al., 2011). This matrix has a much smaller size when compared with the classical PCA so, its calculation is significantly reduce, avoiding the singularity problem (Wang et al., 2006).

This work proposes two approaches to improve the face recognition rates in conditions where facial expression, local illumination and head pose vary. The proposed approaches are Modular Image Principal Component Analysis (MIMPCA) and weighted

* Corresponding author. Tel.: +55 81 2126 8430x4346; fax: +55 81 2126 8438.
E-mail addresses: gdcc@cin.ufpe.br (G.D.C. Cavalcanti), tir@cin.ufpe.br (T.I. Ren), jfp@cin.ufpe.br (J.F. Pereira).
URL: http://www.cin.ufpe.br/~viisar (G.D.C. Cavalcanti).

0957-4174/$ - see front matter
http://dx.doi.org/10.1016/j.eswa.2013.03.003

Modular Image Principal Component Analysis (wMIMPCA) that combines interesting characteristics of both MPCA and 2DPCA. Modular image decomposition is used to minimize local variation and the image covariance matrix is calculated directly from the matrix, improving the representativeness of the data. Distinct face regions can contribute differently to the final classification. In order to deal with this problem, wMIMPCA associates a weight to each face region. Therefore, higher importance is given to regions according to their discriminative power. This set of weights is found by an optimization process that is implemented using a Genetic Algorithm.

This work is organized as follows. Section 2 describes the proposed approaches: MIMPCA and wMIMPCA. The experimental methodology and the results are analyzed and discussed in Section 3. Section 4 presents the conclusions.

2. Proposed approaches

Principal Component Analysis is not very efficient when local features vary considerably, since it extract global information. In this context, a feature vector that represents one image in a scene poorly illuminated differs greatly from a feature vector of a face image captured from an environment with controlled illumination. Consequently, the accuracy rate of the classification algorithm is significantly affected by these changes. We proposed two techniques to deal with this problem: Modular Image Principal Component Analysis (MIMPCA) and weighted MIMPCA (wMIMPCA).

2.1. Modular Image Principal Component Analysis (MIMPCA)

A training set with n images is defined as $\boldsymbol{I}$ and $\boldsymbol{I}_i^t \in \boldsymbol{I}$ denotes an image of size $k \times l$ represented by a matrix of the same size, where $i = 1, \ldots, n$ and t represents its class. All images in the training set that have class equal to q is represented by $C_q = I_i^q$, i.e., $C_1 = \{I_1^1, I_2^1, \ldots, I_{n_1}^1\}, C_2 = \{I_1^2, I_2^2, \ldots, I_{n_2}^2\}$ and so on.

Figs. 1 and 2 show the training and test steps of the MIMPCA method, respectively. In the proposed method, each image is divided into a sections horizontally and b sections vertically. The original image is divided into m sub-images where $m = a \times b$ and the size of each sub-image is equal to $(k \times l)/m$ pixels. These sub-images is represented as:

$$\boldsymbol{I}_{i_{a'b'}}(x, y) = \boldsymbol{I}_i\left(\frac{k}{a}(a' - 1) + x, \frac{l}{b}(b' - 1) + y\right) \tag{1}$$

where a' varies from 1 to a and b' varies from 1 to b.$\boldsymbol{I}_{i_{a'b'}}$ represents the sub-images of coordinates a',b' of the ith image in the training set.

An average image is obtained for all sub-images. The average image is calculated as:

$$\bar{\boldsymbol{X}} = \frac{1}{c}\sum_{q=1}^{c}\bar{\boldsymbol{X}}_q \tag{2}$$

where c is the number of classes and $\bar{\boldsymbol{X}}_q$ corresponds to the average image of the qth class and is computed as:

$$\bar{\boldsymbol{X}}_q = \frac{1}{(|C_q| \cdot a \cdot b)}\sum_{i=1}^{|C_q|}\sum_{a'=1}^{a}\sum_{b'=1}^{b}\boldsymbol{I}^q_{i_{a'b'}} \tag{3}$$

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

$$\boldsymbol{Y}_{i_{ab}} = \boldsymbol{I}_{i_{a'b'}} - \bar{\boldsymbol{X}} \tag{4}$$

where $\boldsymbol{Y}_{i_{ab}}$ represents the normalized region matrix with a,b coordinates of the ith image in the training set.

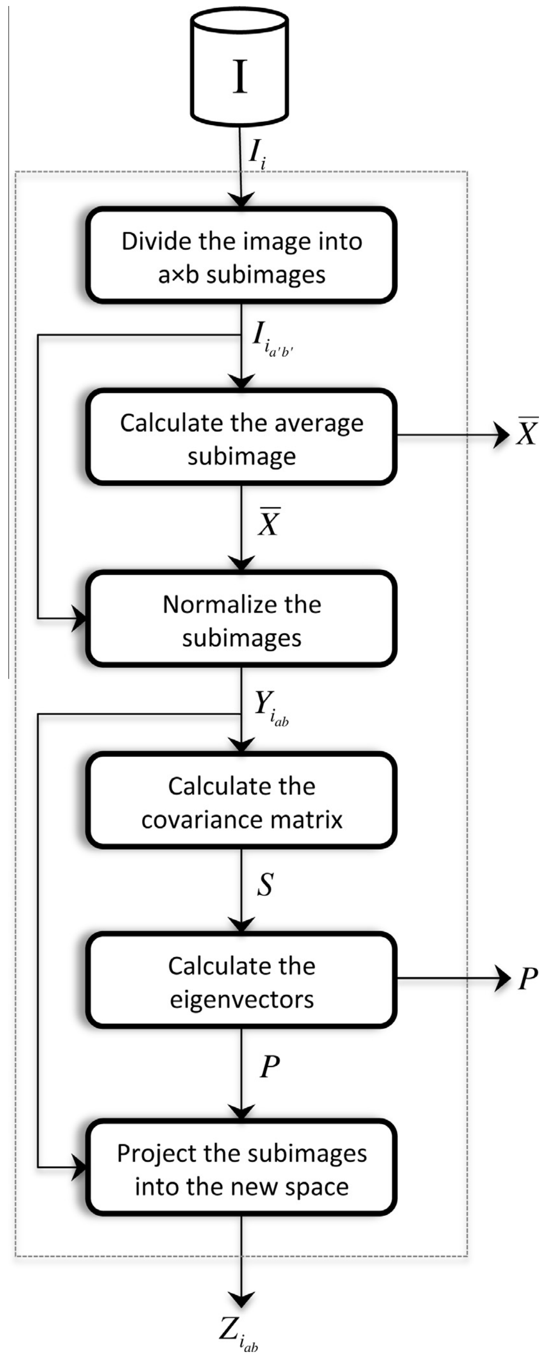


Fig. 1. The MIMPCA training procedure. Given a training dataset $\boldsymbol{I}$, the outputs are: the average sub-image ($\bar{\boldsymbol{X}}$), the transformation matrix ($\boldsymbol{P}$) and the sub-images after the projection to the new space ($\boldsymbol{Z}_{i_{ab}}$).

Based on the sub-images matrices, the covariance matrix is calculated as defined in Eq. (5).

$$\boldsymbol{S} = \frac{1}{c}\sum_{q=1}^{c}\boldsymbol{S}_q \tag{5}$$

where $\boldsymbol{S}_q$ corresponds to the covariance matrix of the qth class in the dataset. This matrix is computed as:

$$\boldsymbol{S}_q = \frac{1}{(|C_q| \cdot a \cdot b)}\sum_{i=1}^{|C_q|}\sum_{a'=1}^{a}\sum_{b'=1}^{b}\boldsymbol{Y}_{i_{ab}} \cdot \boldsymbol{Y}^T_{i_{ab}} \tag{6}$$

The first v eigenvectors, $e_1, e_2, \ldots, e_v$, associated with the largest eigenvalues, obtained from the covariance matrix $\boldsymbol{S}$, are used for

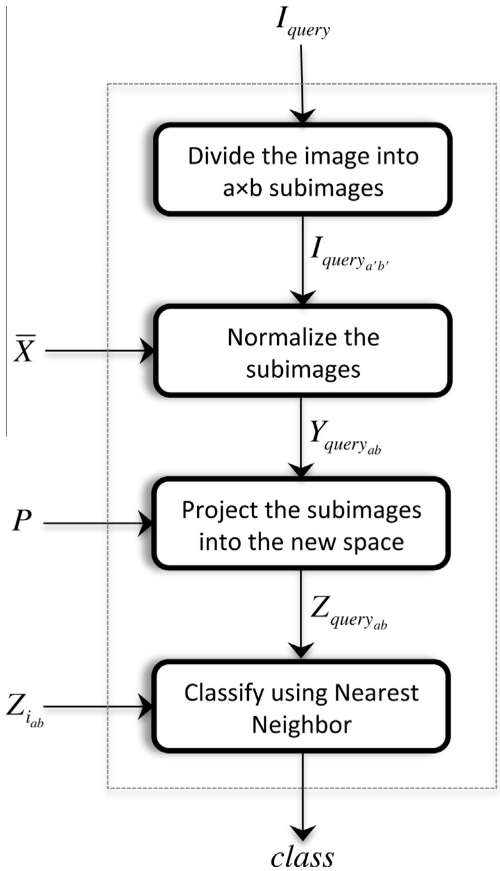


Fig. 2. The MIMPCA test procedure.

classification purposes. The images in the new space are computed multiplying these eigenvectors by the normalized images ($\boldsymbol{Y}_{i_{ab}}$). The result is an $l \times v$ matrix ($\boldsymbol{Z}_{i_{ab}}$) for each sub-image of the original image. In this way, the final projection matrix ($\boldsymbol{P} = [e_1^T e_2^T e_3^T \ldots e_v^T]$) is constructed using each eigenvector e_v as a column. Therefore, each projected sub-image can be computed as a simple matrix multiplication as define in Eq. (7).

$$\boldsymbol{Z}_{i_{ab}} = \boldsymbol{P} \cdot \boldsymbol{Y}_{i_{ab}} \tag{7}$$

To evaluate a query image $\boldsymbol{I}_{query}$, the following information of the training process (Fig. 1) are required: (i) the average sub-image ($\bar{\boldsymbol{X}}$); (ii) the transformation matrix ($\boldsymbol{P}$), and; the sub-images of the training dataset after the projection to the new space ($\boldsymbol{Z}_{i_{ab}}$).

Fig. 2 shows the MIMPCA procedure to evaluate a query image. The first three steps are similar to the ones of the training process (Fig. 1), they are: (i) the image is divided into sub-images, (ii) the sub-images are normalized and (iii) the sub-image are projected to the new space as shown in Eq. (8).

$$\boldsymbol{Z}_{query_{ab}} = \boldsymbol{P} \cdot (\boldsymbol{I}_{query_{ab}} - \bar{\boldsymbol{X}}) \tag{8}$$

After the projection of the query image, a Nearest Neighbor-based classifier is used. Given that each query image is divided into m sub-images, and each sub-image is a matrix with $l \times v$ coefficients, the distance between a reference image $\boldsymbol{I}_r$ and a query image $\boldsymbol{I}_{query}$ is defined in Eq. (9).

$$d(\boldsymbol{I}_{query}, \boldsymbol{I}_r) = \sqrt{\sum_{i=1}^{l}\sum_{j=1}^{v} \boldsymbol{M}_{ij}} \tag{9}$$

$$\boldsymbol{M} = \sum_{a'=1}^{a}\sum_{b'=1}^{b} (\boldsymbol{Z}_{r_{a'b'}} - \boldsymbol{Z}_{query_{a'b'}})^2 \tag{10}$$

This distance is computed for the query image against all patterns in the training set. The class of the query image is assigned as the same class of the reference image that is closer to it.

2.2. Weighted Modular Image Principal Component Analysis (wMIMPCA)

In the MIMPCA method, the feature extraction process uses only one mean and one covariance matrix to combine local and global face information. However, it does not consider that some areas of the face can provide more information than others. In this way, useful information present in different face regions may be minimized and this degrades the final performance of the system.

The Weighted Modular Image Principal Component Analysis (wMIMPCA) defines a set of weights in order to consider different contributions given by each face region in the final classification. The defined set of weights are used to increase or decrease the contribution of each face sub-image. Fig. 3 shows a possible set of weights for three different face database when each face is partitioned in nine regions.

The wMIMPCA training procedure can be divided into two parts: the first one is the training procedure performed by MIMPCA (Fig. 1), and the second one is an optimization procedure that searches for the best weights per image region. The optimization procedure is performed by a Genetic Algorithm (Fig. 4) and it is described in the next subsection.

2.2.1. Searching for the best weights using a genetic optimization

A genetic optimization is used to obtain the best weights per image region (Fig. 4). The first step is to create a vector of weights

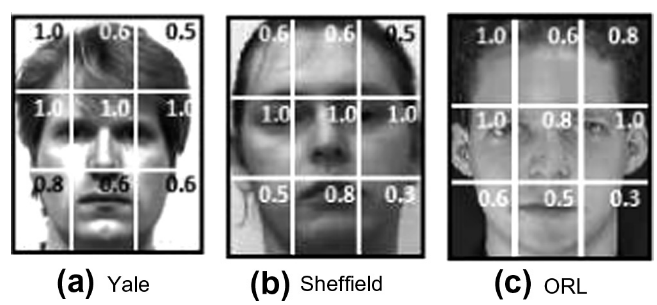


Fig. 3. Weights associated to each of the nine image face regions.

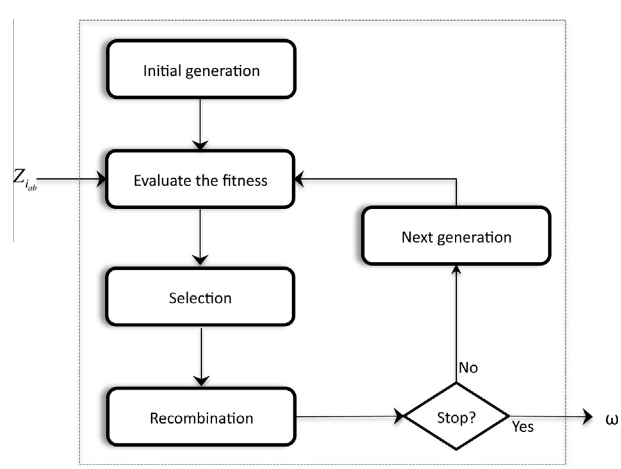
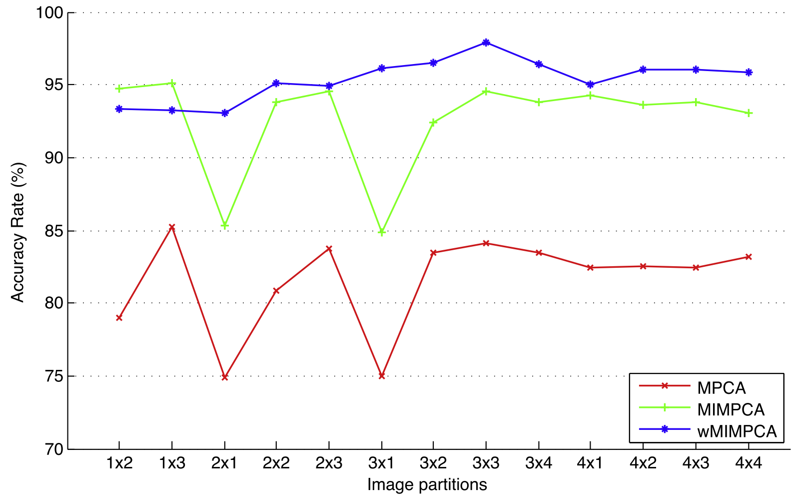
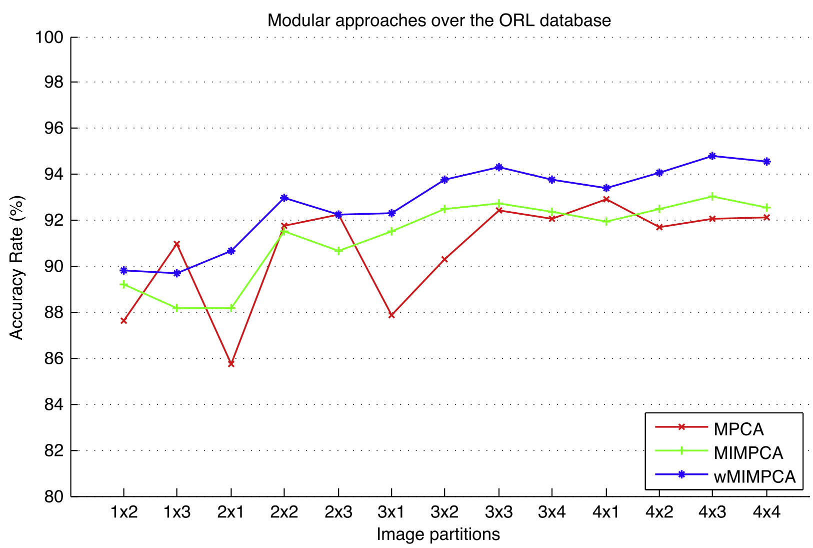


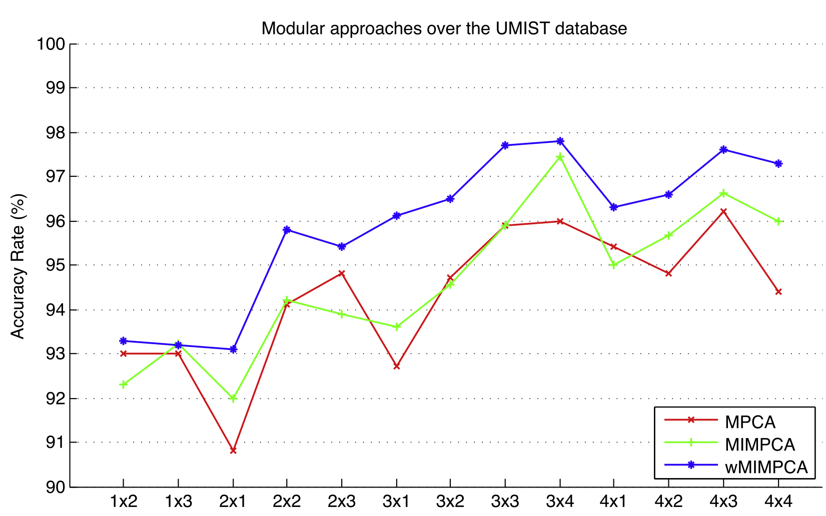
Fig. 4. wMIMPCA optimization procedure to search for the best weights per image region. $\boldsymbol{w}$ is a vector of weights



(a) Yale face database

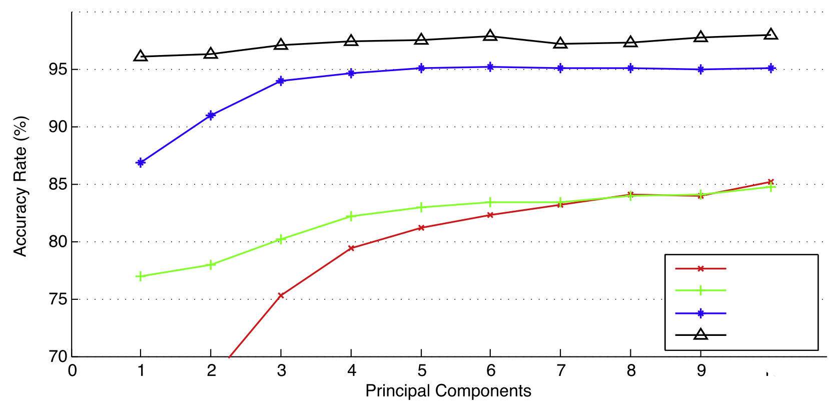


(b) ORL face database

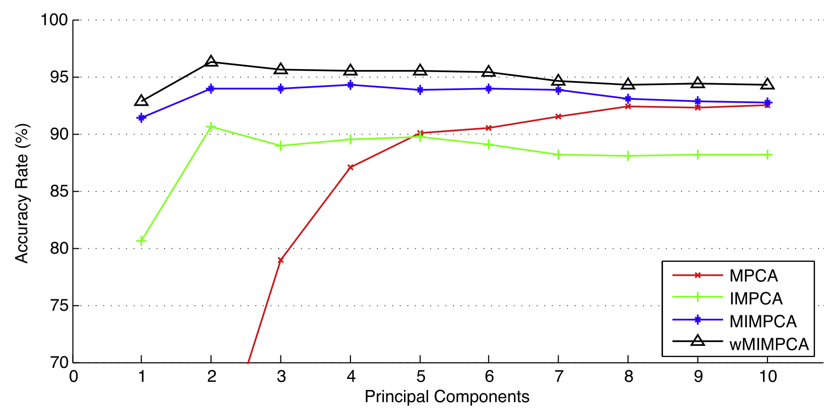


(c) Sheffield face database

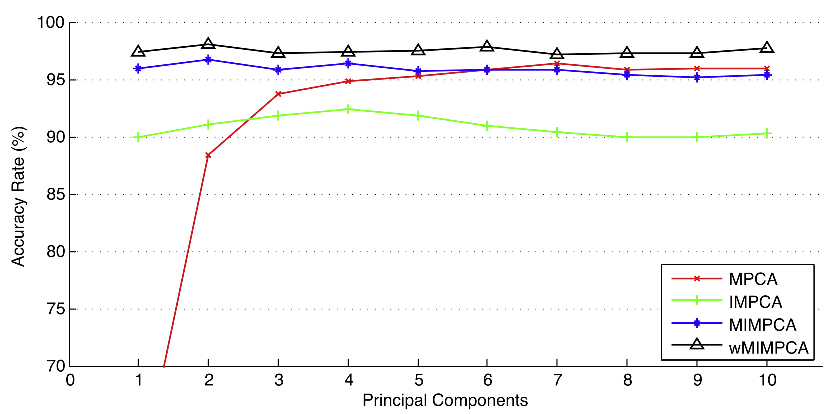
Fig. 5. Accuracy rates of the modular approaches using different image partition sizes (Vertical × Horizontal).



(a) Yale face database



(b) ORL face database



(c) Sheffield face database

Fig. 6. Accuracy rates of different techniques when the number of principal components vary.

Table 1
Best recognition rates: accuracy rates (partition).

	Face databases		
	Sheffield	**ORL**	**Yale**
MPCA (Gottmukkal & Asari, 2004)	96.40 (3×3)	92.45 (2×3)	85.22 (1×3)
IMPCA (Yang et al., 2004)	92.40 (–)	90.60 (–)	84.67 (–)
MIMPCA	96.70 (2×2)	94.33 (3×2)	95.22 (1×3)
wMIMPCA	98.10 (3×3)	96.31(3×3)	97.90 (3×3)

$\boldsymbol{w}$ that has the same size of the number of image regions ($m = a \times b$). Each instantiation of this vector is refereed to as chromosome or individual of the genetic optimization and each weight w_i is defined as a real number between 0 and 1.

A random uniform distribution is used to create the initial generation of individuals. After, the individuals are evaluated by the fitness function. The adopted fitness function is the accuracy rate of the Nearest Neighbor classifier (Eq. (13)) calculated using fivefold cross-validation over the training dataset. It is expected that individuals with higher values for the fitness function have a greater chance of survival than weaker ones. Thus, the selection process aims to choose the individuals from one generation to create the basis of the next generation. The selection function used is the Roulette. The next step is to cross over the selected individuals to produce the offspring and the technique used is the Two-points crossover. If the end condition is satisfied, the best solution in the current population ($\boldsymbol{w}$) is obtained.

2.2.2. Classification

The test procedure of the wMIMPCA method is similar to the one of MIMPCA (Fig. 2). The difference is that the distance is pondered by the weights calculated using the optimization procedure shown in Fig. 4. Thus, the distance between a reference image ($\boldsymbol{I}_r$) and a test image ($\boldsymbol{I}_{query}$) is given by:

$$d(\boldsymbol{I}_{query}, \boldsymbol{I}_r) = \sqrt{\sum_{i=1}^{l}\sum_{j=1}^{v}\boldsymbol{M}_{ij}} \tag{11}$$

$$\boldsymbol{M} = \sum_{a'=1}^{a}\sum_{b'=1}^{b} w_{a'b'} \cdot (\boldsymbol{Z}_{r_{a'b'}} - \boldsymbol{Z}_{query_{a'b'}})^2 \tag{12}$$

where $w_{a'b'}$ is the weight associated to each image region.

The final classification is defined as:

$$t = \arg\min_{\boldsymbol{I}_r \in I_i^t}(d(\boldsymbol{I}_{query}, \boldsymbol{I}_r)) \tag{13}$$

where t represents the class of the nearest neighbor of the test image I_{query}.

3. Experiments

The experiments were conducted using three well-known face databases (Yale, ORL and Sheffield) and the feature extraction techniques (IMPCA, MPCA, MIMPCA and wMIMPCA) were evaluated using the same conditions. The Yale database was used to test the performance of the proposed methods with different facial expressions and variation in illumination. The Sheffield (previously UMIST) face database was used to evaluate the performance of the methods on images with changes in face poses, from profile to frontal view. The face images in the ORL database have small changes in pose and in size.

In order to find the best set of weights required by the wMIMPCA technique, a Genetic Algorithm was used with the following initial parameters: (i) Initialization: Random uniform; (ii) Population size: 15; (iii) Number of generations: 30; (iv) Crossover: Two points using factor 0.7; and, (v) Selection function: Roulette. The performance of the methods was evaluated by stratified ten-fold cross-validation.

3.1. Results

The first experiment shows the behavior of the modular approaches (MPCA, MIMPCA and wMIMPCA) in the three databases when the image partition size changes (Fig. 5). The MPCA obtained the worse results while the proposed approaches obtained the best accuracy rates, especially the wMIMPCA. Independent of the database used, the partition size 3×3 presented the best performance. Thus for the next experiments this configuration was used as the partition parameter. In contrast, partition sizes 2×1 and 3×1 generated higher error rates for MPCA and MIMPCA, while the performance of the wMIMPCA was more stable.

Fig. 6 shows the accuracy rates of the feature extraction techniques varying the number of principal components. In all the experiments, MIMPCA and wMIMPCA presented a very stable behavior. This stability was not observed in MPCA neither in IMPCA. The difference in the accuracy rate between the proposed techniques and MPCA and IMPCA was more evident in the Yale face database (Fig. 6(a)). This database has facial images captured under different illumination conditions that generally affects only some regions of the faces. wMIMPCA obtained better results than MIMPCA because in the wMIMPCA technique, these unaffected regions are emphasized while the affected ones are de-emphasized in order to improve the final classification.

Experiments using the ORL face database (Fig. 6(b)) show that the improvement provided by the proposed techniques is more expressive at low dimensionality. Fig. 6(c) shows the results of the experiment using the Sheffield face database that explores different head pose angles. On the average, the accuracy rate of the wMIMPCA was one percentile point better than the MIMPCA technique.

Table 1 shows the best accuracy rates obtained by each of the evaluated techniques. The size of the partitions is shown in parenthesis. For the evaluated databases, the wMIMPCA obtained the best accuracy rates when compared with MPCA, IMPCA and MIMPCA.

An analysis of the computational time to process each feature extraction algorithm reveals that MPCA is three times slower than wMIMPCA. If we adopt that wMIMPCA requires s seconds to perform the feature extraction procedure of one face, on the average, the IMPCA takes 1.5s and the MIMPCA takes 0.84s to perform the same task.

4. Conclusion

We have proposed two feature extraction techniques: Modular IMage PCA (MIMPCA) and weighted Modular IMage PCA (wMIMPCA). The proposed techniques use the modular two-dimensional approach for feature extraction. This procedure takes advantage of the face regions that are not affected by local variations, such as illumination, facial expression and head pose. The wMIMPCA computes one weight per face region. These weights are calculated using a Genetic Algorithm and its cost function aims to minimize the relevance of regions that present local variations. Another advantage of the modular two-dimensional approach is the reduced size of the image representation and the subsequent decrease in the computational cost.

The experimental results show that wMIMPCA obtains better recognition rates when compared with MIMPCA, MPCA and IMPCA. Therefore, wMIMPCA is an alternative to improve the recognition rate and to reduce the response time of face recognition systems.

Acknowledgment

This work was supported in part by the Brazilian governmental agencies CNPq and Facepe

References

Barreto, R.M., Ren, T.I., & Cavalcanti, G.D.C., 2012. L2-norm metric learning applied to unconstrained face pair-matching. In International Conference on Image Processing (pp. 581–584).

Fan, X., & Verma, B. (2009]). Selection and fusion of facial features for face recognition. *Expert Systems with Applications, 36*, 7157–7169.

Franco, A., & Nanni, L. (2009]). Fusion of classifiers for illumination robust face recognition. *Expert Systems with Applications, 36*, 8946–8954.

Gottmukkal, R., & Asari, V. K. (2004]). An improved face recognition technique based on modular pca approach. *Pattern Recognition Letters, 24*, 429–436.

Gumus, E., Kilic, N., Sertbas, A., & Ucan, O. N. (2010]). Evaluation of face recognition techniques using PCA, wavelets and SVM. *Expert Systems with Applications, 37*, 6404–6408.

Pereira, J.F., Barreto, R.M., Cavalcanti, G.D.C., & Ren, T.I., 2011. A robust feature extraction algorithm based on class-modular image principal component analysis for face verification. In International Conference on Acoustics, Speech and Signal Processing. pp. 1469–1472.

Pereira, J.F., Cavalcanti, G.D.C., & Ren, T.I., 2009. Modular image principal component analysis for face recognition. In International Joint Conference on Neural Networks (pp. 2481–2486).
Turk, M., & Pentland, A. (1991]). Eigenfaces for recognition. *Journal of Cognitive Neuroscience, 3*, 71–86.
Wang, L., Wang, X., & Feng, J. (2006]). On image matrix based feature extraction algorithms. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 36*, 194–197.
Xudong, Xie., & Kin-Man, Lam. (2006]). An efficient illumination normalization method for face recognition. *Pattern Recognition Letters, 27*, 609–617.
Yang, J., Zhang, D., Frangi, A., & Yang, J. (2004]). Two-dimensional pca: A new approach to appearance-based face representation and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 26*, 131–137.
Yang, M.H., (2002). Kernel eigenfaces vs. kernel fisherfaces: Face recognition using kernel methods. In *International conference on automatic face and gesture recognition* (pp. 215–220).
Zhao, W., Chellappa, R., Phillips, P. J., & Rosenfeld, A. (2003]). Face recognition: A literature survey. *ACM Computing Surveys, 35*, 399–458.