Journal of King Saud University – Computer and Information Sciences (2018) **30**, 141–151

King Saud University

Journal of King Saud University – Computer and Information Sciences

www.ksu.edu.sa
www.sciencedirect.com

Face recognition using Angular Radial Transform

Bensenane Hamdan *, **Keche Mokhtar**

Laboratoire Signals and Images, Dept. of Electronique, Université des Sciences et de la Technologie d'Oran Mohamed Boudiaf, USTO-MB, BP 1505, 3100 Oran, Algeria

Received 8 May 2016; revised 24 October 2016; accepted 27 October 2016
Available online 5 December 2016

KEYWORDS

Angular Radial Transform (ART);
Legendre moment invariants (LMI);
Euclidean distance (ED);
Pseudo-Zernike moments (PZM);
Nearest Neighbor Classifier (NNC);
Support Vector Machines (SVM)

Abstract Moment-based Angular Radial Transform, Legendre moment invariants and Zernike moments are a family of orthogonal functions which allow the generation of non-redundant descriptors by the projection of an image onto an orthogonal basis. These descriptors can be used for classification, such as in face recognition. Zernike moments and Legendre moments have already been used for this purpose.

This paper proposes to use moment-based Angular Radial Transform for extracting the face characteristics that feed a Support Vector Machine or a Nearest Neighbor Classifier for face recognition. Facial images from the ORL database, Essex Faces94 database, Essex Faces96 database, and Yale database were used for testing the proposed approach. The experimental results obtained show that the proposed method is more efficient, in terms of recognition rate, than the methods based on Zernike and Legendre moments. It is also found that its performance is comparable to that of the best state-of-the-arts methods.

Contents

* Corresponding author.
E-mail addresses: hamdan.bensenane@univ-usto.dz (B. Hamdan), m_keche@yahoo.com (K. Mokhtar).

Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.jksuci.2016.10.006
1319-1578

1. Introduction

In recent years, security has become an international concern. Today, it is required to use appropriate data processing techniques to ensure global security. Security is to be ensured in several areas like access control to work or public places, access control to computers, e-commerce, banking system based on identification, means of transportation, etc. Nowadays, biometrics occupies a particular place as a means for ensuring security. It consists of identifying a person from one or more of his physiological characteristics (fingerprints, face, iris, hand's contour, DNA, etc.), or from his behavior (signature, gait, etc.).

Over the last decade, a great deal of research work has been done to improve the reliability of biometric systems. The choice of using facial recognition as a biometric modality is motivated by the fact that this kind of modality is contactless, natural, well accepted and requires only a very cheap sensor (webcam) which is available on a great number of electronic devices. However, it requires a little cooperation from the user during the facial image acquisition phase. Automatic face recognition is performed in two essential steps, namely extraction of facial features and classification.

Several research works have been carried out for the extraction of facial features. This research has led to the development of a multitude of methods, which can be classified into three categories, i.e. global, local, and hybrid.

The global methods for facial feature extraction use the whole image as input to the recognition system. The advantage of this representation is that it implicitly preserves the texture information and shape, which are required for face recognition. In addition, compared to local representations, this one allows a better appearance capture of the face (O'Toole et al., 1993). However, its major drawback is the very large storage space that it requires (Jain and Chandrasekaran, 1982). In practice, it is not necessary to have a large amount of data to develop an accurate model for the facial features of a person. Dimension reduction techniques, such as Eigenfaces (PCA) Turk and Pentland, 1991, Fisherfaces (LDA) Duda et al., 2001, are commonly used. In order not to lose information during the conversion process from 2D images to 1D image vectors, a 2D image-based PCA (2DPCA) method was proposed by Yang et al. (2004). Using similar 2D projections onto a subspace, Yang et al. proposed the 2DLDA method (Yang et al., 2005), whereas Niu et al. suggested the 2DLPP method (Niu et al., 2008). To improve the performance of the 2D projection method, Li et al. (2016) suggested a sequential three-way decision approach for cost-sensitive face recognition. The proposed method is based on a formal description of granular computing.

Local or geometric methods are based on the extraction of the relative position of the elements that make up the face (such as the nose, mouth and eyes). In the early 1990s, Brunelli and Poggio (1993)) described a facial recognition system that automatically extracts 35 geometric characteristics of the face. The similarity was calculated using Bayes classifiers. Manjunath et al. (1992) proposed the Elastic Bunch Graph Matching (EBGM), a local characteristics method for face recognition, based on Gabor Wavelet Transform (Lee, 1996). Zhang et al. (2015) presented a simple but efficient feature extraction method based on facial landmarks and multi-scale fusion features (FLWLD). They first extracted the local features using Weber Local Descriptors (WLD) Chang and Lin, 2001 and multi-scale patches centered at predefined facial landmarks, and then constructed fusion features by randomly selecting parts of the local features. However, the geometric characteristics are usually difficult to extract, especially in complex situations, such as variable illumination, and occlusions. The geometric characteristics alone are not sufficient to represent a face. Hybrid methods may be used in a modular manner for different facial areas. A global model may then be obtained from the combination of different local models. Consequently, the different facial regions are not affected in the same way by the various sources of variability. For example, wearing sunglasses considerably changes the appearance of the eyes, and a smile affects more the area around the mouth. The modular Eigenspace approach, introduced by Pentland et al. (1994), belongs to the above mentioned category. Another efficient feature extraction algorithm, called Discriminant Sparse Local Spline Embedding (D-SLSE), which can be considered as a hybrid approach for face recognition was proposed by Lei et al. (2015).

During the second phase of recognition, namely the classification phase, the system must decide whether the person belongs to the database and if so, to what class he belongs; in other words: who is the person? Of course, the answer of the system may be wrong. The approaches proposed in the literature, to solve such a problem, belong to the field of automatic data classification, a research field that has been widely explored in the recent decades, in many domains. The methods that can be applied in this step depend mainly on the technique used in the signature extraction step. These methods, include the methods for calculating distances (Cox et al., 1996) (Euclidean distance, Mahanalobis distance) and Support Vector Machines (SVM) Guo et al., 2000, which is the most effective technique for implicit identification of distribution models, underlying the data distribution in the feature space. The application of SVM for face image processing-related issues was presented for the first time in Osuna et al. (1997).

A new approach has emerged for the decomposition of a grayscale image, using the principle of radial polynomials. This approach uses the moments from Angular Radial Transformation (ART) to represent the image. Angular Radial Transformation (ART) is used in many pattern intelligent applications, such as video surveillance systems (Lee et al., 2011), logo recognition (Wahdan et al., 2011), and face detection with an 88.7% correct detection rate (Fang and Qui, 2003); ART is a region-based descriptor in MPEG-7 (Bober, 2001).

This approach is used to extract the feature vector in the proposed face recognition system.

2. Face recognition using the ART moments

As already mentioned above, a recognition system consists of two phases: the extraction of features phase and the classification phase (Fig. 1).

Many features can be used for face recognition.

These features include the polynomial moments, such as the Legendre moments invariants (LMI), polynomial circular moments, pseudo-Zernike moments (PZM), and the moments obtained by the ART. The first two have already been proposed for face recognition. This paper proposes to use ART moments, which are presented in the following.

2.1. Extraction of facial features using polynomial and circular moments

In the field of information processing, polynomial and circular moments are widely used for their orthogonal property which allows the generation of non-redundant descriptors, as well as for their translation, scale, and rotation invariance properties.

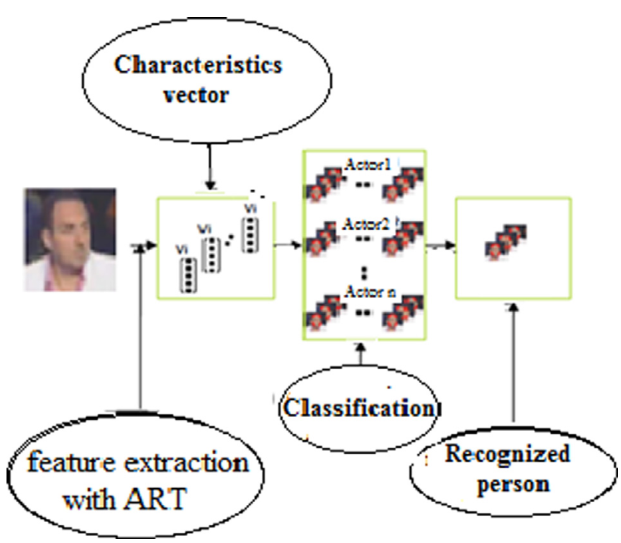

Figure 1 Schematic of a general face recognition system based on ART.

For example, their moments have been applied for recognition of the images of people, indexing images in databases, as well as for the analysis and description of 2D or 3D shapes of objects. Hu (1961) was the first to introduce the use of image moment invariants in 2D pattern recognition applications. Among the most popular moments used as features for face recognition, one may mention the Legendre moments and the circular moments. Both of these types of moments are briefly described in the following.

2.1.1. The Legendre moment invariants

Annadurai and Saradha (2004) used the Legendre moment invariants (LMI) for the polynomial decomposition of a grayscale image. The LMI, $L_{m,n}$ of a squared $N \times N$ image I (i, j), is given by the following equation:

$$L_{m,n} = \lambda_{m,n} \sum_{i=1}^{N} \sum_{j=1}^{N} I(i,j) \cdot P_m(x_i) \cdot P_n(y_j) \tag{1}$$

The normalization coefficient $\lambda_{m,n}$ is given by:

$$\lambda_{m,n} = \frac{(2m+1)(2n+1)}{(N-1)^2} \tag{2}$$

where the polynomial moment, $P_m(x)$, denotes the Legendre polynomial of order m; it is given by:

$$P_m(x) = \sum_{k=0}^{m} C_{mk}[(1-x)^k + (-m)^k(1+x)^k] \tag{3}$$

with:

$$C_{mk} = \frac{(-1)^k(m+k)!}{2^{k+1}(m+k)!(k!)^2} \tag{4}$$

Since Legendre polynomials are orthogonal over the interval $[-1, 1]$, a square image of $(N \times N)$ pixels with the intensity function $I(i, j)$ must be scaled to be within the region $-1 \leqslant x, y \leqslant 1$, i.e.:

$$x_i = \frac{2i-N-1}{N-1}, \quad y_j = \frac{2j-N-1}{N-1} \tag{5}$$

Fig. 2 represents the 2D Legendre basis, with $n = 0{:}2$ and $m = 0{:}4$.

2.2.2. The circular moments (ART, PMZ)

The feature vectors of face images, extracted with the help of ART or the PMZ (Nabatchian and Makaremi, 2008), are orthogonal projections of the face image on the radial basis function. This projection is defined by:

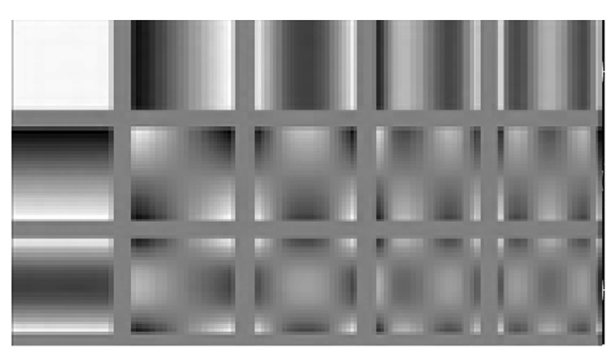

Figure 2 The 2D LMI basis, with $n = 0{:}2$ and $m = 0{:}4$.

$$W_{m,n} = \int_0^{2\pi} \int_0^1 f(r,\theta) \cdot V_{nm}(r,\theta) dr d\theta \tag{6}$$

The application of (6) to a discrete function $I(x, y)$ requires rewriting it as follows:

$$W_{m,n} = \sum_{r \leqslant 1} \sum_{\theta \leqslant 2\pi} I(r,\theta) \cdot [V_{nm}(r,\theta)]^* \tag{7}$$

where $I(r,\theta)$ represents the image intensity in polar coordinates, and $V_{n,m}(r,\theta)$ represents the orthogonal radial basis function forming the projection basis. The difference between ART and PMZ lies in this basis.

2.2.2.1. The PMZ projection basis. The radial basis function in PMZ consists of two functions that are written, in general, as follows:

$$V_{n,m}(r,\theta) = A_m(\theta).R_{n,m}(r) \tag{8}$$

where $\mathrm{A_m}(\theta)$ is an exponential function that ensures rotation invariance:

$$A_m(\theta) = \exp(-jm\theta) \tag{9}$$

and the radial basis function $R_{n,m}(r)$ is defined by:

$$R_{n,m}(r) = \lambda_n \sum_{k=0}^{n-|m|} \frac{(-1)^k \cdot (2 \cdot n + 1 - k)! \cdot r^{n-k}}{(k)!(n+|m|-k)!(n-|m|-k)!} \tag{10}$$

The normalization coefficient λ_n is given as:

$$\lambda_n = \frac{n+1}{\pi} \tag{11}$$

with $|\mathrm{m}| \leqslant \mathrm{n}$.

Fig. 3 represents the 2DPMZ basis, with n = 0:4 and m = 0:n.

2.2.2.2. The ART projection basis. The radial basis function in ART consists of two functions that are written, in general, as follows:

$$V_{n,m}(r,\theta) = A_m(\theta) \cdot R_n(r) \tag{12}$$

where $A_m(\theta)$ is an exponential function that ensures rotation invariance:

$$A_m(\theta) = \frac{1}{2} \cdot \exp(jm\theta), \tag{13}$$

and the radial basis function $R_n(r)$ is defined by a cosine function:

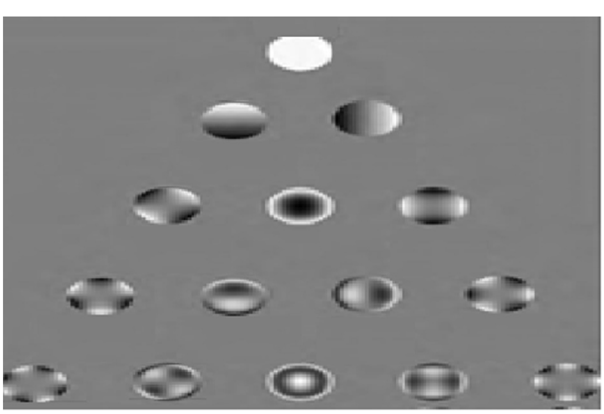

Figure 3 The 2D PMZ basis, with n = 0:4 and m = 0:n.

$$R_n(r) = \begin{cases} 1 & \text{for} \quad n = 0 \\ 2 \cdot cos(n \cdot r \cdot \theta) & \text{for} \quad n \neq 0 \end{cases} \tag{14}$$

Fig. 4 represents the 2D ART projection basis, with n = 0:2 and m = 0:4.

To preserve the orthogonality of the basis $W_{m,n}$, the function $I(x, y)$ must be recalculated inside the unit circle by rewriting it in polar coordinates (r, θ), in such a way that the image center is the center of the unit circle (Fig. 5). Transformation from Cartesian to polar coordinates is performed as follows:

$$\begin{cases} x = r \cdot cos\theta \\ y = r \cdot sin\theta \end{cases} \quad \text{and} \quad \begin{cases} r = \sqrt{x^2 + y^2} \\ \theta = tan^{-1}\left(\frac{y}{x}\right) \end{cases} \tag{15}$$

The shape of the image $I(x, y)$ is rectangular or square, which is incompatible with the shape of the unit circle. This requires making a choice between the elimination of certain points of the image (especially the corners) or the introduction of points that do not belong to the original image. The above relationship may be written again as follows:

$$\begin{cases} x_i = c + \frac{i(d-c)}{N-1} \\ y_j = d - \frac{j(d-c)}{M-1} \end{cases} \quad \text{and} \quad \begin{cases} R_{ij} = \sqrt{x_i^2 + y_j^2} \\ \theta_{ij} = tan^{-1}\left(\frac{y_j}{x_i}\right) \end{cases} \tag{16}$$

where i and j are the coordinates of a point in the original image, x_i and y_j are the new coordinates of this same point in the new coordinate system (the unit circle), M and N are the horizontal and vertical extents of this image, respectively, and c and d are the parameters that allow to choose between recalculating the function $I(x, y)$ entirely ($c = -\sqrt{1/2}$ and

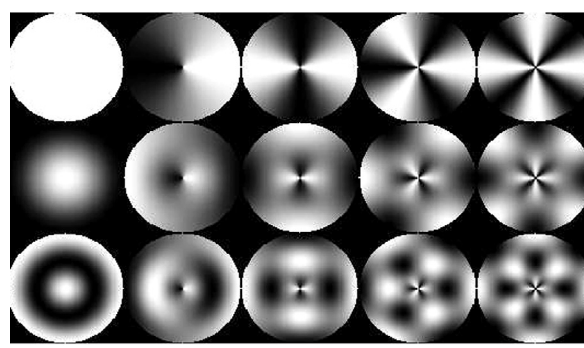

Figure 4 The 2D ART projection basis, with n = 0:2 and m = 0:4.

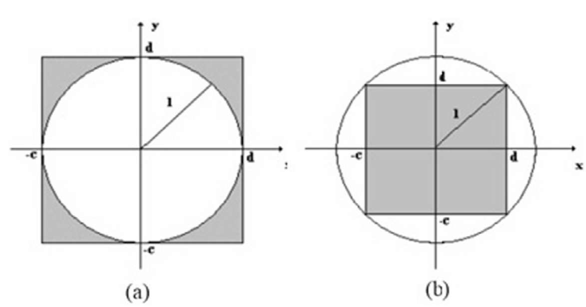

Figure 5 Recalculation of the function $I(x, y)$ (Gray rectangle) in the unit circle: a) $c = -1$ and $d = 1$, b) $c = -1/\sqrt{2}$and $d = 1/\sqrt{2}$.

$d = \sqrt{1/2}$) or partially ($c = -1$ and $d = 1$) inside the unit circle, as shown in Fig. 5.

The circular moments, $W_{m,n}$ obtained by the orthogonal projection of the face image onto the radial basis function, can be expressed as a feature vector of a face image in several ways:

(1) $\{\mathcal{R}(P_{nm}), \mathcal{J}(P_{nm})\}$: a one-dimensional complex number is converted into a two dimensional real number.
(2) $||P^2_{n,m}||$: the amplitude, which is the absolute value of the complex number.
(3) $\arg(P_{nm})$: the phase or the argument of the complex number.
(4) $\{||\{P^2_{n,m}||, \arg(P_{nm})\}$: the amplitude and the argument of the complex number.

The first representation was used in our work, as it combines both the real and imaginary parts in one feature vector. This representation preserves the phase and avoids complex calculations, compared to other conversions.

2.2.3. Comparison between the different projection bases

- The LMI projection basis is composed of two real polynomials, $P_m(x_i)$ and $P_n(y_j)$, which are independent. The resulting moments are thus real valued.
- The PMZ projection basis is the product of a complex function and a radial polynomial. The resulting moments are thus complex valued.
- The ART projection basis is the product of a complex function and a cosine function. The resulting moments are also complex valued.
- The LMI projection basis is orthogonal over the interval $[-1, 1]$.
- The ART and PMZ projection bases are orthogonal over the unit circle.
- The repeat order, m, is independent of the decomposition order, n, in the LMI and ART projection bases.
- All of these moments are originally invariant only to rotation. An initial normalization step is required for invariance to scaling and translation.

2.2. Classification

Two methods were tested, for the classification of face images. The first one uses a simple Euclidean distance to find the Nearest Neighbor. The second one is more efficient, but more complex; it uses Support Vector Machines (SVM) to improve the recognition rate.

The Euclidean distance may be defined from the Minkowski distance of order p in an Euclidean space R^N of dimension N.

Consider two vectors $X = (x_1, x_2, \ldots, x_N)$ and $Y = (y_1, y_2, \ldots, y_N)$; the pth order Minkowski distance, Lp, is thus defined by:

$$Lp = \left(\sum_{i=1}^{N} |x_i - y_i|^p\right)^{1/p} \tag{17}$$

For $p = 2$, the Euclidean distance is:

$$L2 = \sqrt{\sum_{i=1}^{N} |x_i - y_i|^2} \tag{18}$$

Support Vector Machines (SVM) may be used to solve discrimination problems, that is to say to decide to which class belongs a sample. Solving this problem requires the construction of a function which associates each input vector x with an output y, which characterizes its class.

$$y = h(x) \tag{19}$$

Support Vector Machines (SVM) use a supervised learning algorithm that aims to learn the function
$h(x)$ through a training set:
$(\{(x_0, y_0), \ldots, (x_k, y_k)\}$, where $x_k \subset \mathcal{R}^N$ and $y_k \in \{-1, 1\})$.
The class is given by Y, and is defined as:

$$Y = Sign(y_i \cdot h(x_i)) \tag{20}$$

Also:

$$\begin{cases} Y = 1 & \text{if } y_i \cdot h(x_i) \geqslant 0 \\ Y = -1 & \text{if } y_i \cdot h(x_i) < 0 \end{cases} \tag{21}$$

Theoretically, there is an infinite number of hyperplanes that separate the two classes, including the unique optimal hyperplane, defined as the hyperplane that maximizes its margin (distance) to the samples of the two classes (Fig. 6).

It can be shown that finding the optimal hyperplane is a quadratic optimization problem of dimension p (number of examples), under constraints, which can be solved by the conventional method of Lagrange multipliers. Its resolution gives the optimal Lagrange multipliers α_k^*, which can be used to obtain the equation of the optimal hyperplane:

$$h(x) = \sum_{k=1}^{p} \alpha_k^* \cdot y_k \cdot x_k \cdot x + w_0 \tag{22}$$

This formulation of SVM assumes that the data to classify are linearly separable. Extending the Support Vector Machines (SVM) to the case where the data are not linearly separable requires reformulating the problem in a higher dimension space, where the data become linearly separable. The problem can be solved using kernel functions that reduce the computational complexity. The equation of the separating hyperplane

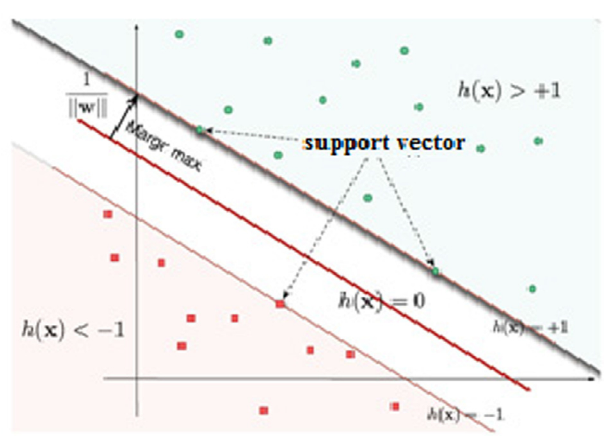

Figure 6 The optimal hyperplane (red) with the maximum margin. Samples indicated with arrows are support vectors.

may be expressed in terms of the kernel function $k(x \cdot x')$ as follows:

$$h(x) = \sum_{k=1}^{p} \alpha_k^* \cdot y_k \cdot K(x_k \cdot x) + w_0 \tag{23}$$

The most usual kernel functions are:

- Linear: $k(x, x') = x{\cdot}x'$
- Polynomial: $k(x, x') = (x \times x')^d$ or
- $k(x, x') = (c + x{\cdot}x')^d$
- Gaussian: $k(x,x') = e^{-|x-x'|^2/2\cdot\sigma^2}$
- Laplacian: $k(x,x') = e^{-|x-x'|/2\cdot\sigma^2}$

SVM is by nature a binary classification technique (two classes). Its extension to the multi-class case (M > 2 classes) may be performed using several methods. The most well-known are the one-versus-all (OVA) method that uses M binary classifiers, where each one compares one class to the rest, and the one-versus-one method, which uses M(M − 1)/2

binary classifiers, where each one compares two classes among M.

3. Performance evaluation of face recognition methods based on polynomial and circular moments

To evaluate the face recognition methods based on polynomial and circular moments, two databases were used, namely the Essex Faces94 database and the ORL database. Two classifiers, the Nearest Neighbor Classifier (NNC) and the SVM classifier were tested, using the one-against-all (OAA) strategy. Several kernel functions were tried for the SVM. The best results were obtained with the 5th order polynomial kernel function. Therefore, all results for the SVM presented below, correspond to that kernel function.

3.1. The Essex Faces94 database

The Essex Faces94 database was developed, at the University of Cambridge, as part of a work for the realization of a face recognition system with the PMZ. The authors obtained a good face recognition rate with this database. This database was used to compare our approach, which uses the ART, with the approach that uses the PMZ. The Faces94 database (Fig. 7) contains 72 classes, recorded without variations in facial orientation, and with 20 changes in facial expressions (Fig. 8). The size of the images in this database is 200 × 180 pixels.

3.2. The ORL database

The ORL database was collected between April 1992 and April 1994 by an AT&T laboratory, based in Cambridge. The ORL database (Fig. 9) contains 40 people, each one registered under 10 different views (Fig. 10). This base, which is considered as a reference for evaluating face recognition algorithms, was used to evaluate our proposed approach.

For some individuals, the images were collected at different times, with variations in lighting conditions and facial expressions (neutral expression, smile and closed eyes), and with partial occlusions by glasses. All the images of the database are labeled; this allows evaluating the performances of the face recognition methods. These images are in grayscale and consist of pixel values ranging from 0 to 255. The size of each image is 92 × 112 pixels.

3.3. The Yale face database

The Yale face database (Fig. 11) includes 165 images of 15 people, each registered under 11 different variations in lighting condition, facial expression, and with or without glasses (Fig. 12).

3.4. The Essex Faces96 database

The Essex Faces96 database (Fig. 13), developed at the University of Cambridge, is larger compared to the preceding databases. It contains 152 classes recorded with no great variation in facial orientation and with 20 changes in facial expressions (Fig. 14). The size of the images in this database is 196 × 196 pixels.

3.5. Preprocessing

Prior to extraction of the feature vectors, a preprocessing operation of the images was performed in the databases.

The first step consists in converting each image RGB color in the databases (Faces94, Faces96) to an image in grayscale. Then, the images in the Yale database and Faces96 database were cropped in order to keep the face region only. Finally, all images in the databases were resized to 64 × 48 pixels.

3.6. Influence of decomposition order on the performance of each method

The decomposition order n affects the response time and the recognition rate of the face recognition systems, which are based on the polynomial moments. Increasing the order of decomposition increases the recognition rate, as well as the response time. However, as shown by the results presented below, increasing the order of decomposition beyond a certain value, which depends on the used method, does not improve the recognition rate.

Figs. 15–22 illustrate the variation of the recognition rate as a function of the decomposition order n, with classification by

Figure 7 Image $n° = 1$ of the first ten individuals in the Faces94 database.

Figure 8 Extracts from the Faces94 database showing different facial expressions.

Figure 9 Image $n° = 1$ of the first ten individuals from the ORL database.

Figure 10 Extracts from the ORL database with different orientations.

Figure 11 Image $n° = 1$ of the first eleven individuals in the Yale database.

Figure 12 Extracts from the Yale database showing variations in lighting condition, facial expression, and with or without glasses.

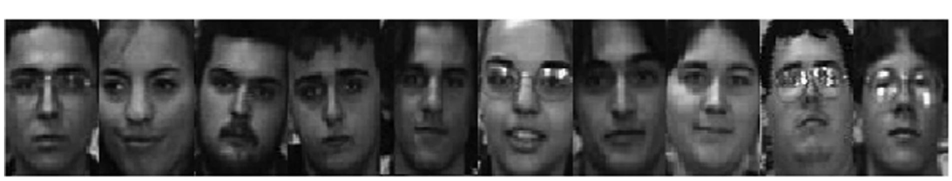

Figure 13 First image of ten individuals from the Faces96 database.

Figure 14 Extracts from the Faces96 database showing different facial expressions.

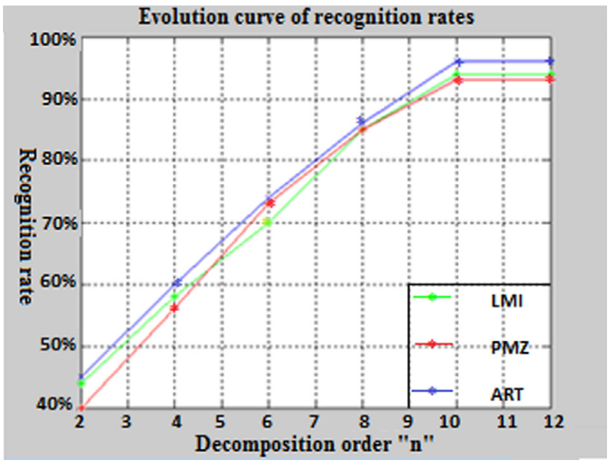

Figure 15 Recognition rate, with classification by SVM, for the Faces94 database.

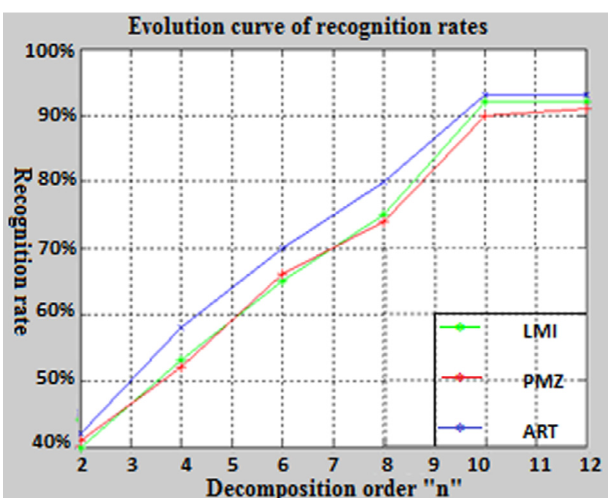

Figure 16 Recognition rate, with classification by NNC, for the Faces94 database.

SVM and NNC. Figs. 15 and 16 were obtained, using the Faces94 database, with 20 images per class (15 images for learning and 5 for the test), while Figs. 17 and 18 were obtained, using the ORL database, with 10 images per class (7 images for learning and 3 for the test).

Figs. 19 and 20 were obtained, using the Yale database, with 11 images per class (7 images for learning and 4 for the test).

Figs. 21 and 22 were obtained, using a larger database, i.e. the Faces96 database, with 20 images per class (15 images for learning and 5 for the test).

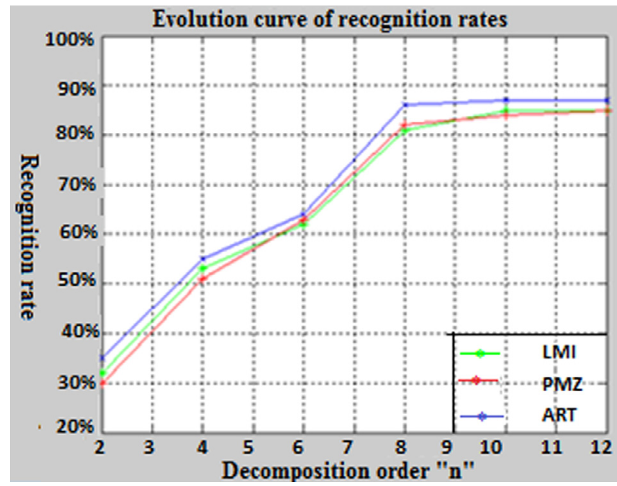

Figure 17 Recognition rate, with classification by SVM, for the ORL database.

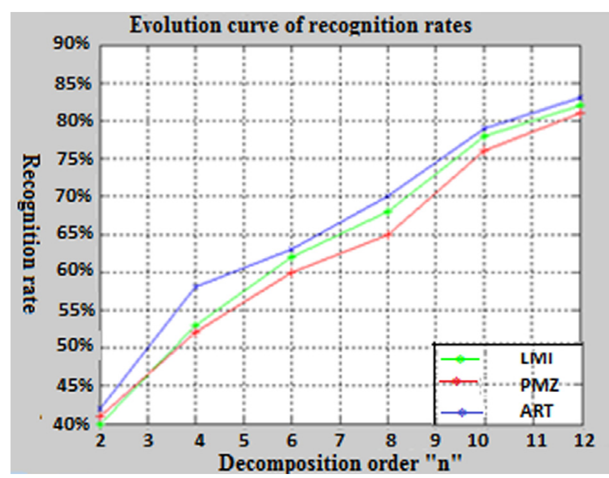

Figure 18 Recognition rate, with classification by NNC, for the ORL database.

3.7. The choice of the kernel function for the classification by SVM

As mentioned before, a polynomial kernel function with order d equal to 5, was chosen for the classification by SVM. This order was chosen by trial and error because there is no rule for such a choice. In Fig. 23, the recognition rate is plotted against the polynomial order of the SVM polynomial kernel. As can be seen from this figure, beyond the order 5, the recognition rate remains unchanged, which means that this is the

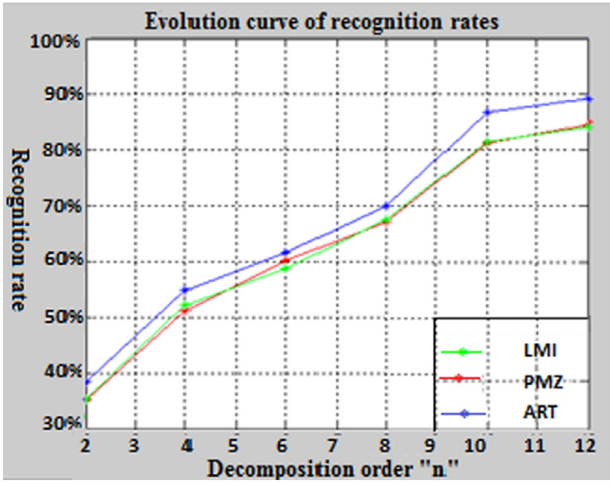

Figure 19 Recognition rate, with classification by SVM, for the Yale database.

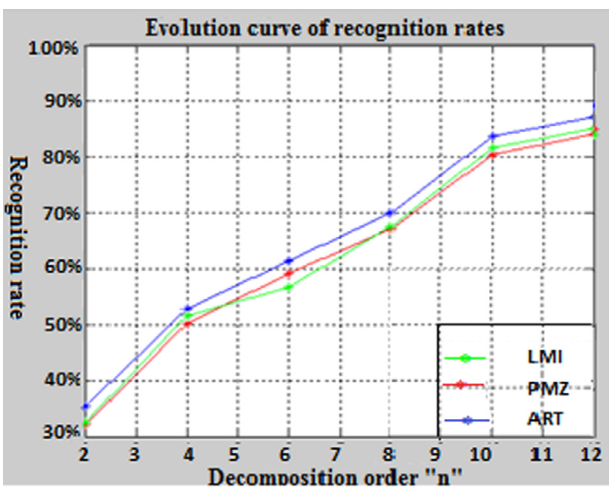

Figure 20 Recognition rate, with classification by NNC, for the Yale database.

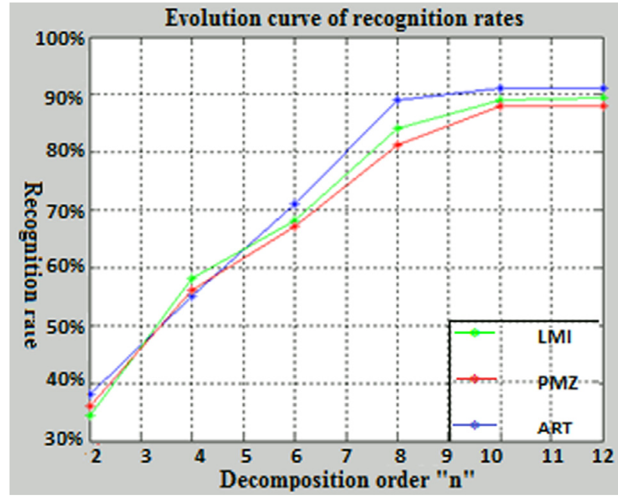

Figure 21 Recognition rate, with classification by SVM, for the Faces96 database.

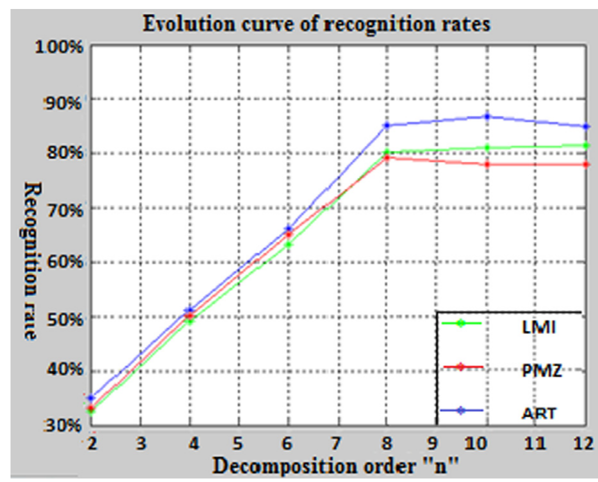

Figure 22 Recognition rate, with classification by NNC, for the Faces96 database.

minimum order that ensures separation between classes in our case.

To select the appropriate Kernel function, several tests were carried out, using the LIBSVM software package (Chang and Lin, 2001) and the Faces94 database. The polynomial Kernel gave us the highest recognition rate, compared to the linear and Gaussian kernels, as shown in Table 1.

3.8. Comparing the different moment-based face recognition methods

From the results presented in Figs. 13–20, the best rates achieved by the different moment-based face recognition methods were extracted and are given in Table 2, together with the order of the polynomial decomposition with which they were obtained.

It can be seen from the above table that the features extracted, using the ART and the PMZ, are, respectively, the more and the less discriminating features for the classification, which is always better achieved when the SVM is used.

The response time, as well as the required memory size and identification rate depend mainly on the order of decomposition and the classification method used. It was found that the optimal decomposition order is usually lower when the SVM classifier is employed, especially in the case of the ORL database. However, this classifier is more costly in terms of computation time than the NNC.

3.9. Comparing the proposed method with other methods

In Tables 3 and 4, the recognition rates obtained with our method are compared with those given by the other methods that used the ORL database and the Yale database. These methods are LDA (Duda et al., 2001), D-SLSE (Lei et al., 2015), LBP (Zhao et al., 2010), FLWLD (Zhang et al., 2015), and WLD (Chen et al., 2010). For a fair comparison, 5 images per person were used for training in all methods.

For a more reliable comparison, the Faces96 database was used as it is larger than the other two databases used (ORL and Yale). Table 5 displays the recognition rates obtained with

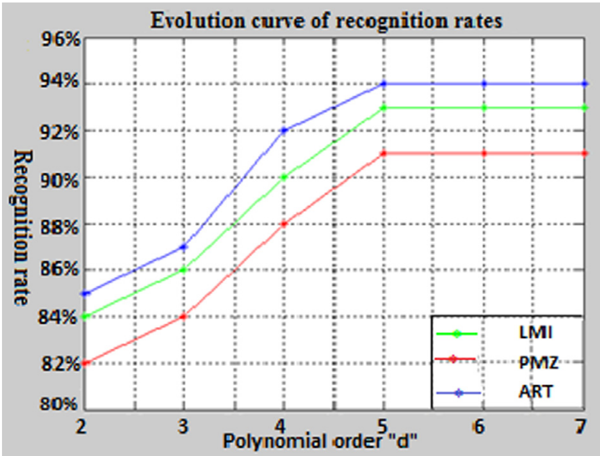

Figure 23 Recognition rate versus polynomial order of the SVM polynomial kernel.

Table 1 Comparison of recognition rates obtained with SVM, using different kernel functions.

	Gaussian (σ = 0.5) (%)	Linear (%)	Polynomial (d = 5) (%)
ART	83.0	89.1	96.0
LMI	83.5	87.5	93.1
PMZ	82.5	85.1	92.2

Table 2 Best recognition rates obtained with different moment-based methods.

	ORL (*n*)	Faces94 (*n*)	Yale (*n*)	Faces96 (*n*)
ART + NNC	83.1% (12)	94.1% (10)	88.0% (12)	87.4% (10)
ART + SVM	88.0% (10)	96.0% (10)	89.4% (12)	90.8% (12)
LMI + NNC	82.1% (12)	92.2% (10)	85.1% (12)	82.2% (12)
LMI + SVM	84.4% (10)	93.1% (10)	86.3% (12)	89.1% (12)
PMZ + NNC	81.3% (12)	91.3% (12)	84.6% (12)	79.3% (8)
PMZ + SVM	83.1% (10)	92.2% (10)	85.9% (12)	88.2% (12)

our method, and with the LDA, LBP, and WLD methods we programmed ourselves. These rates were obtained using 5 images per person for training in all methods.

From the results displayed in the above three tables, one can say that compared with other methods, the proposed method gives the best results, with the Yale and Faces96 databases, but with the ORL database, it performs worse than almost all other methods, especially with the FLWD and D-SLSE methods. However, compared to these two methods, our method is less complex, since it is a global method that extracts classification features using neither local processing

Table 3 Comparison of our approach with other approaches using the the ORL database.

Method	ORL (%)
LDA	85.6
D-SLSE	95.6
ART (proposed)	87.7
LBP	88.3
FLWLD	**97.5**
WLD	90.0

The values in bold are indicate the method that had the highest rate.

Table 4 Comparison of our approach with other approaches using the Yale database.

Method	Yale (%)
LDA	77.2
D-SLSE	81.1
ART(proposed)	**85.2**
LBP	78.6

The values in bold are indicate the method that had the highest rate.

Table 5 Comparison of our approach with other approaches, using the Faces96 database.

Method	Faces96 (%)
LDA	83.3
ART (proposed)	**87.4**
LBP	80.3
WLD	81.1

The values in bold are indicate the method that had the highest rate.

and fusion, like in the FLWLD method, nor projection, like in the LDA and D-SLSE methods.

As a conclusion one can say that this method is not the most suitable for face recognition when the facial images are tilted as in the ORL database, but it is appropriate when the facial images are frontal, as in Faces94, Faces96 and Yale databases, despite the change in facial expressions. Such a case is encountered in applications where the user of the face recognition system is cooperative, such as in an access control system.

4. Conclusion

A new method for face recognition is proposed in this paper. This method uses the moments based on ART as a feature vector, and SVM with the polynomial kernel function as a classifier. The performance of this method, in terms of recognition rate, was evaluated and compared with that of other methods using the polynomial approach, namely the LMI and the PMZ. The obtained results show that it is the most efficient.

Moreover, the proposed method was compared to various state-of-the-art methods, and it was found that our method outperforms these methods when the face images are frontal, despite the change in facial expressions. However, it is less suitable for the recognition of inclined faces.

The advantage of the polynomial approach, as compared to other statistical approaches, such as the PCA analysis, is that the size of the feature vector is smaller and does not depend on the size of the face image. As a result, a smaller memory size is needed for the storage of the database, and a shorter computational time is necessary.

The disadvantage of the polynomial approach is that the response time and the identification rate depend on the order of decomposition, hence the need for choosing a decomposition order that achieves a good compromise between these two criteria.

References

Annadurai, S., Saradha, A., 2004. Face recognition using legendre moments. In: Proceedings of the Fourth Indian Conference on Computer Vision, Graphics & Image Processing.

Bober, M., 2001. MPEG-7 visual shape descriptors. IEEE Trans. Circuits Syst. Video Technol. 11 (6), 716–719.

Brunelli, R., Poggio, T., 1993. Face recognition: features versus templates. Proc. IEEE Trans. Pattern Anal. Mach. Intell. 15, 1042–1062.

Chang, C.C., Lin, C.J., LIBSVM: A Library of Support Vector Machines, Software Available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm2001>.

Chen, J., Shan, S., He, C., et al, 2010. WLD: a robust local image descriptor. IEEE Trans. Pattern Anal. Mach. Intell. 32 (9), 1705–1720.

Cox, I., Ghosn, J., Yianilos, P., 1996. Feature-based face recognition using mixture-distance. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 209–216.

Duda, R.O., Hart, P.E., Stork, D.G., 2001. Pattern Classification. John Wiley & Sons Inc, New York etc.

Fang, J., Qui, G., 2003. Human face detection using angular radial transform and support vector machines. Int. Conf. Image Process. 1, 69–72.

Guo, G., Li, S.Z., Chan, K., 2000. Face recognition by support vector machines. In: Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 196–201.

Hu, M.K., 1961. Visual pattern recognition by moment invariants. IRE Trans. Inf. Theory 49, 179–187.

Jain, A.K., Chandrasekaran, B., 1982. Dimensionality and sample size considerations in pattern recognition practice. In: Krishnaiah, P. R., Kanal, L.N. (Eds.), Handbook of Statistics, vol. 2, pp. 835–855.

Lee, T.S., 1996. Image representation using 2-d Gabor wavelets. Proc. IEEE Trans. Pattern Anal. Mach. Intell. 18, 959–971.

Lee, S.H., Sharma, S., Sang, L., Park, J., Park, Y.G., 2011. An Intelligent Video Security System using Object Tracking and Shape Recognition, vol. 6915. ACVIS LNCS Springer-Verlag, Berlin Heidelberg, pp. 471–482.

Lei, Y., Han, H., Hao, X., 2015. Discriminant sparse local spline embedding with application to face recognition. Knowl. Based Syst. 89, 47–55.

Li, H., Zhang, L., Huang, B., Zhou, X., 2016. Sequential three-way decision and granulation for cost-sensitive face recognition. Knowl. Based Syst. 91, 241–251.

Manjunath, B.S., Chellappa, R., Malsburg, C.V.D., 1992. A feature based approach to face recognition. In: Proceedings, IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 373–378.

Nabatchian, A., Makaremi, I., 2008. Pseudo-Zernike moment invariants for recognition of faces using different classifiers in FERET database. In: Third International Conference on Convergence and Hybrid Information Technology, vol. 1, pp. 933–936.

Niu, B., Yang, Q., Shiu, S.C.K., Pal, S.K., 2008. Two-dimensional Laplacian faces method for face recognition. Pattern Recognit. 41 (10), 3237–3243.

O'Toole, A.J., Abdi, H., Deffenbacher, K.A., Valentin, D., 1993. Low-dimensional representation of faces in higher dimensions of the face space. Opt. Soc. Am. 10 (3), 405–411.

Osuna, E., Freund, R., Girosit, F., 1997. Training support vector machines: an application to face detection. In: Proceedings of the Conference on Computer Vision and Pattern Recognition, San Juan, PR, pp. 130–136.

Pentland, A., Moghaddam, B., Starner, T., 1994. View-based and modular eigenspaces for face recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 84–91.

Turk, M., Pentland, A., 1991. Eigenfaces for recognition. J. Cognitive Neurosci. 3 (1), 71–86.

Wahdan, O., Omar, K., Nasrudin, F., 2011. Logo recognition system using angular radial transform descriptors. J. Comput. Sci. 7, 1416–1422.

Yang, J., Zhang, D., Frangi, A.F., Yang, J.Y., 2004. Two-dimensional PCA: a new approach to appearance-based face representation and recognition. IEEE Trans. Pattern Anal. Mach. Intell. 26 (1), 131–137.

Yang, J., Zhang, D., Xu, Y., Yang, J.Y., 2005. Two-dimensional discriminant transform for face recognition. Pattern Recognit. 38 (7), 1125–1129.

Zhang, Z., Wang, L., Zhu, Q., Chen, S., Chen, Y., 2015. Pose-invariant face recognition using facial landmarks and Weber local descriptor. Knowl. Based Syst. 84, 78–88.

Zhao, G., Ahonen, T., Matas, J., et al, 2010. Rotation-invariant image and video description with local binary pattern features". IEEE Trans. Image Process. 21 (4), 1465–1477.