

Age Invariant Face Recognition Based on DCT Feature Extraction and Kernel Fisher Analysis

Leila Boussaad*, Mohamed Benmohammed**, and Redha Benzid***

Abstract

The aim of this paper is to examine the effectiveness of combining three popular tools used in pattern recognition, which are the Active Appearance Model (AAM), the two-dimensional discrete cosine transform (2D-DCT), and Kernel Fisher Analysis (KFA), for face recognition across age variations. For this purpose, we first used AAM to generate an AAM-based face representation; then, we applied 2D-DCT to get the descriptor of the image; and finally, we used a multiclass KFA for dimension reduction. Classification was made through a K-nearest neighbor classifier, based on Euclidean distance. Our experimental results on face images, which were obtained from the publicly available FG-NET face database, showed that the proposed descriptor worked satisfactorily for both face identification and verification across age progression.

Keywords

Active Appearance Model, Age-Invariant, Face Recognition, Kernel Fisher Analysis, 2D-Discrete Cosine Transform

1. Introduction

Face recognition across age progression is a recent topic of growing interest, especially for applications in which age compensation is required, such as for identifying missing children, conducting surveillance and detecting multiple enrolments, where subjects are either not available or are trying to hide their identity. In such applications, there may be a significant age difference between the query image and those stored in the database, and it may be impossible to obtain the subject's recent face images to update the database. Developing age-invariant face recognition systems would avoid the necessity of updating large facial databases with more recent images.

Aging-related variations have two main characteristics that influence the process of dealing with these types of variations and that make it a challenging task. First, the effects of aging cannot be controlled because it is not possible to eliminate aging variations during the capturing of a face image and collecting the appropriate training data for studying the effects of aging requires a long time. Second, both the rate and kind of age-related effects differ for different individuals and it is in combination with external factors, like health conditions and lifestyle, which have been shown to

* This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
Manuscript received May 15, 2015; accepted October 15, 2015.

Corresponding Author: Leila Boussaad (boussaad.mous@gmail.com)

* Dept. of Computer Science, University of Batna, Algeria (boussaad.mous@gmail.com)

** Dept. of Computer Science, University of Constantine, Algeria (ben_moh123@yahoo.com)

*** Dept. of Electronics, University of Batna, Algeria (redha.benzid@gmail.com)

contribute to facial aging effects [1]. These particular features make the aging problem a challenging research task.

Unlike studies on other effects on face recognition systems, like illumination, facial expressions, and pose changes, there are few studies that have been carried out on face recognition across age progression. One of the principal causes for this has been the lack of public databases containing images of individuals at different ages. To the best of our knowledge, the only two largest publicly available face datasets are the MORPH [2] and FG-NET (Face and Gesture Recognition Research Network) [3] databases. Unfortunately, these databases include other issues like poses, illumination, and expressions, along with the poor quality of old photos. Therefore, proposed approaches must take these problems into account.

The existing age-dependent face recognition methods include two major categories [4]. The approaches in the first category are called “generative” since they apply a computational model to simulate the aging process to offset the impact of aging on face texture and shape, and then later apply recognition algorithms to obtain the query identity. The other category includes “non-generative” approaches [4], which concentrate on defining discriminatory features and projection methods that are minimally affected by temporal variations to allow for the accurate identification of an individual.

In this paper, we focus on non-generative methods. Our proposed method integrates Active Appearance Models (AAMs) [5] to extract a free-shape representation of the image texture, a two-dimensional discrete cosine transform (2D-DCT), and the multiclass Kernel Fisher Analysis (KFA) [6] for dimension reduction.

The rest of the paper is organized as follows: Section 2 gives a brief review of some works dealing with aging problems, and Section 3 presents the steps constituting our proposed approach. Section 4 provides information on the experimentation we carried out, and presents results and comparisons. Section 5 concludes the paper.

2. Related Works

This section outlines some earlier works that have been carried out in regards to the aging impact on face recognition systems. As we mentioned earlier, age-invariant face recognition methods are divided into two parts: generative and non-generative methods.

As an example of the first class, Ramanathan and Chellappa [7] built a system for modeling the aging process in adulthood. They built a muscle-based geometrical change model that describes changes throughout adulthood, and they characterized facial wrinkles and other skin traits that can be observed during different ages by using an image gradient-based texture transformation function.

Park et al. [8] proposed a 3D generative facial aging model for age-invariant face recognition. In this modeling technique, the query image is transformed to the same age as the database images using the trained 3D aging model.

As example of methods belonging to the second category, Ling et al. [9,10] proposed a face descriptor based on image gradient orientations taken from multiple resolutions combined with a support vector machine (SVM) as a classifier for face verification throughout all stages of the aging process. The proposed approach was tested using two private passport databases to study how increasing age gaps affects the verification performance. They also studied the effects of age-related issues, such as image

quality, presence of spectacles, and facial hair.

In the same context, Mahalingam and Kambhamettu [11] developed a non-generative approach in which they defined a face operator by combining the local binary pattern (LBP) at each level of the Gaussian pyramid constructed for each face image, and the classification was performed using the AdaBoost classifier. Singh et al. [12] presented an age transformation algorithm for minimizing the variation in facial features caused by aging that registers the gallery and query face images in a polar coordinate system. Biswas et al. [13] created a verification algorithm that analyzes the coherency of the drifts in the various facial features to check if two images taken at different ages are for the same individual or not.

Additionally, Mahalingam and Kambhamettu [14] proposed an age-invariant face recognition approach using a graph-based face representation that contains geometry and the appearance of facial feature points. The aging model was constructed by using Gaussian mixture models for each individual to model their age variations in shape and texture. Their recognition approach consists of two steps: In the first step, the search space was reduced and the potential candidates were effectively selected by using a maximum a posteriori (MAP) for each individual aging model. They exploited the spatial similarity between graphs using a simple deterministic algorithm for matching in the second step.

Bereta et al. [15] evaluated the local descriptors that are commonly used in face recognition, such as LBP, multi-scale block LBP (MBLBP), and Weber local descriptor (WLD), in the context of age progression. In their study, classification was carried out by calculating a distance between feature vectors containing local textures involving various similarity measures, such as Euclidean, cosine, correlation, etc. Their results showed that the Gabor coded MBLBP feature combined with the Euclidean distance yielded the highest recognition accuracy.

Sungatullina et al. [16] presented a multiview discriminative learning (MDL) method that learns about a latent low dimensional subspace. It does so by projecting three local features (SIFT, LBP, and GOP) into a common feature space, such that the correlations of different feature representations of each sample are maximized, the within-class variation of each feature is minimized, and the between-class variation of each feature is maximized.

Inspired by Elastic Bunch Graph Matching (EBGM) [17], Yang et al. [18] proposed a method called texture embedded discriminative graph matching, which formulated age-invariant face recognition as a graph matching problem. In their approach, each face is represented as a texture embedded graph, the nodes of the graph present the texture of a face area around a set of fiducial landmarks of the face, and the edges correspond to the geometric topology of the face (they represent the relative distances between the nodes). For each area, they extracted discriminative age-invariant features using the local Gabor binary pattern histogram sequence (LGBPHS) that was projected in a Linear Discriminant Analysis (LDA) subspace. An objective function was then designed to match graphs for registration and identification.

For the same purpose, Juefei-Xu et al. [19] demonstrated that the periocular region is more stable than the full face for all ages. They presented a feature extraction approach for age-invariant face recognition by applying a robust Walsh-Hadamard transform encoded local binary pattern (WLBP) on a preprocessed periocular region, followed by an unsupervised discriminant projection (UDP) for subspace modeling.

3. Proposed Approach Description

In this section, we present the details of our proposed method so as to get age-invariant features. After the crucial step of pose correction using AAMs [5], we applied a 2D-DCT on entire face images and some of the first low frequency coefficients were discarded. The remaining coefficients were used as feature vectors, and then we performed a KFA [6] to reduce dimensionality and obtain an age-robust feature. The K-nearest neighbor classifier performed the classification. The selection of coefficients was made in zigzag manner.

3.1 Aging Database

For the experiment described in this paper, we used the FG-NET aging database [3], which is a well-known database that is used to evaluate facial aging models. The FG-NET aging database is a publicly available image database containing face images showing a number of subjects at different ages. The database has been developed as an aide for researchers who study the effects of aging on facial appearance and their effects on the performance of face recognition systems. The database contains 1,002 images from 82 different subjects ranging in age from newborns to 69 years old. Typical images from the database are shown in Fig. 1. Data files containing the locations of 68 facial landmarks that were identified manually and the age of the subject in each image are also available.

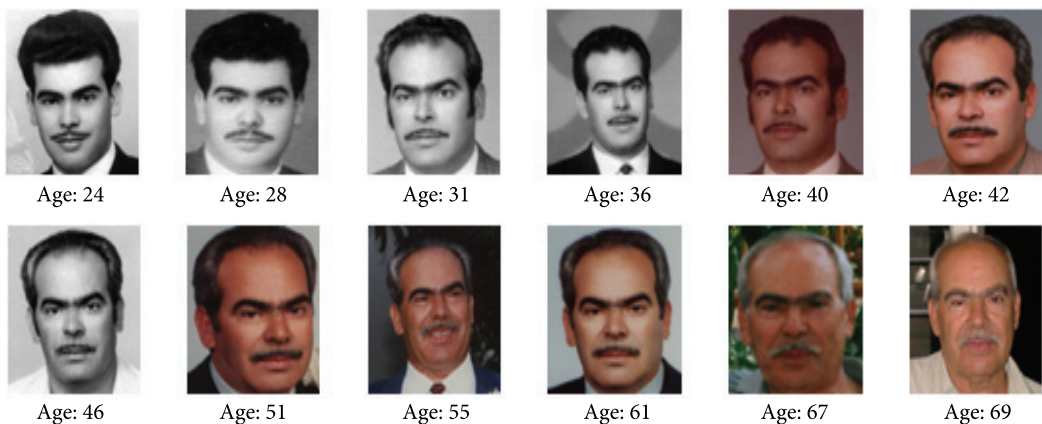


Fig. 1. Examples of images of a person taken at different ages. Images are from the FG-NET aging database [3].

3.2 Face Normalization in Images

The normalization step for the original images is an important first step for obtaining successful results. We used the same approach as for AAMs [5,20], which is described below.

All the face images in the FG-NET database are annotated with 68 facial landmarks. A set of these landmarks is called “shape,” and the set of pixels in a gray level inside this shape is called “texture,” as shown in Fig. 2.

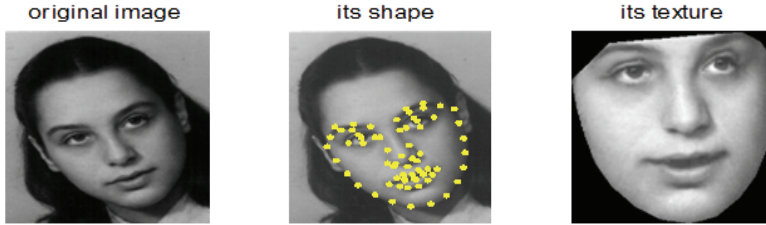


Fig. 2. Original image, its shape and texture inside the shape.

Shape alignment: This includes the translation, scaling, and rotation of all shapes to represent them in the same referential (Fig. 3). The classical solution to align shapes in a common referential is the Procrustes analysis method [20]. A simple iterative approach is described in Fig. 4.

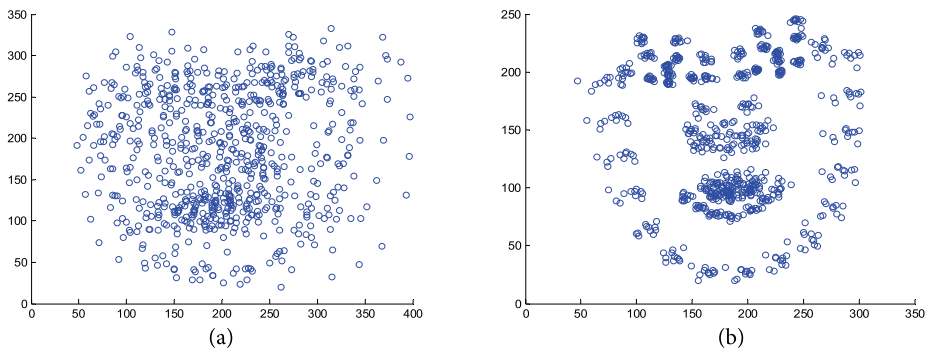


Fig. 3. Shapes alignment. (a) Original shapes and (b) aligned shapes.

Input: Set of unaligned shapes.

Output: Set of aligned shapes and mean shape.

1. Choose a reference shape (for example the first shape)
2. Translate each shape so that is centered on the origin
3. Scale the reference shape to unit size, call this shape \bar{x}_0 , the initial mean shape;
4. Align all shapes to the mean shape;
5. Recalculate the mean shape from the aligned shapes;
6. Constrain the current mean shape (align to \bar{x}_0 , scale to unit size)
7. Go to step 4 until convergence (i.e., mean shape does not change much).

Fig. 4. Different steps involved to aligning shapes [20].

Aligning two shapes (Step 4 in Fig. 4), s_1 to s_2 , consists of finding the parameters of the transform T , (i.e., scale coefficient s , rotation angle θ , and translation (t_x, t_y)) that when applied to s_1 best align it with s_2 , and thereby minimizing the Procrustes distance metric:

$$D_{procrustes}(s_1, s_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (1)$$

with respect to s , θ and (t_x, t_y) .

Cootes and Taylor [20] proposed a simple method to estimate s and θ , as follows:

$$a = (s1^T s2) / |s1|^2 \quad (2)$$

$$b = \sum_{i=1}^{68} (x_{i,1} y_{i,2} - y_{i,1} x_{i,2}) / |s1|^2 \quad (3)$$

then, the rotation angle and the scale coefficient are given by:

$$\theta = \arctan(b/a) \quad (4)$$

$$s = \sqrt{a^2 + b^2} \quad (5)$$

1. **Warping image texture:** Once all of the shapes are aligned into a common frame, this step makes it possible to make texture independent from the variations of the shape. In other words, we wanted to get a free shape representation of texture. For a given shape in an image, we extracted the texture and warped it to the mean shape, which was calculated in the previous step. Delaunay triangulation was used in the mean shape to establish triangles that will be used to warp texture (see Fig. 5). The method that we used was a piece-wise affine transformation, where each pixel belonging to a triangle is mapped into the corresponding triangle in the mean shape using barycentric coordinates [20].

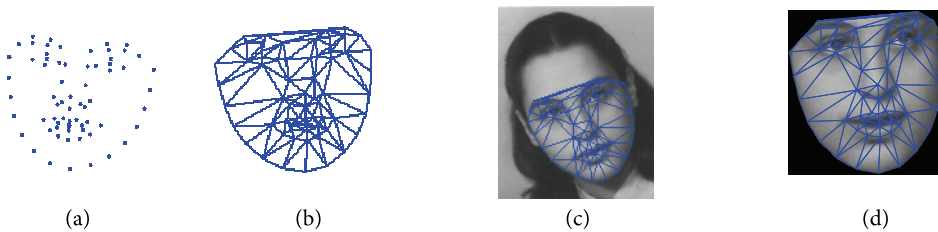


Fig. 5. Warping texture example. (a) The mean shape, (b) Delaunay triangulation of the mean shape, (c) current image, and (d) warped texture.

2. Images were rotated to get eyes at fixed points in an image so that the inter-ocular segment was horizontal. This was based on files of eye coordinates that were provided with the original FG-NET database.
3. Images were resized to 128×128 . Fig. 6 shows the face images in Fig. 1 after normalization.

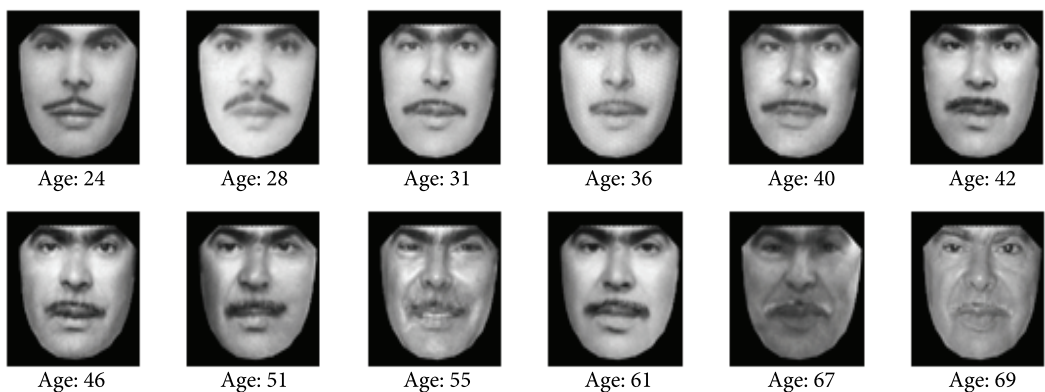


Fig. 6. Normalized images of a person at different ages.

3.3 Feature Extraction

In order to obtain the feature vector from a preprocessed image using AAM, we applied a 2D-DCT, and we only kept a subset of coefficients. The number of selected coefficients was chosen such that they could represent a face. Our purpose was not to reduce the dimensionality of the data, but to discard coefficients that represented too much of the texture details. The selected features that construct the feature vector were represented as points in high dimensional space, and we performed KFA to reduce dimensionality. The description of KFA is given in Subsection 3.5. The details of the DCT method are presented in the following subsection.

3.4 Two-Dimensional Discrete Cosine Transform

DCT is a powerful mathematical transform used for numerous applications in image processing applications, such as image coding, and is used for feature extraction in several studies on face recognition [21,22]. Diverse classes of DCT have been proposed [23]. In particular, the DCT was categorized [23] into four classes: DCT-I, DCT-II, DCT-III, and DCT-IV. DCT-II is the most common variant of DCT applied in signal coding, especially in compression, since it was the main idea in JPEG compression [24].

The DCT transforms an image from the spatial domain to the frequency domain. The 2D-DCT for input image A is defined as follows:

$$B(p, q) = \alpha_p \alpha_q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A_{mn} \cos \frac{\pi(2m+1)p}{2M} \cos \frac{\pi(2n+1)q}{2N}, \quad 0 \leq p \leq M-1, \quad 0 \leq q \leq N-1 \quad (6)$$

and the inverse transform is defined as:

$$A(m, n) = \sum_{p=0}^{M-1} \sum_{q=0}^{N-1} \alpha_p \alpha_q B_{pq} \cos \frac{\pi(2m+1)p}{2M} \cos \frac{\pi(2n+1)q}{2N}, \quad 0 \leq p \leq M-1, \quad 0 \leq q \leq N-1 \quad (7)$$

where:

$$\alpha_p = \begin{cases} \frac{1}{\sqrt{M}}, & p = 0 \\ \sqrt{\frac{2}{M}}, & 1 \leq p \leq M-1 \end{cases} \quad (8)$$

and:

$$\alpha_q = \begin{cases} \frac{1}{\sqrt{N}}, & q = 0 \\ \sqrt{\frac{2}{N}}, & 1 \leq q \leq N-1 \end{cases} \quad (9)$$

M and N are the row and column size of A .

The DCT has a very important feature—it helps to separate the image into parts of differing importance. After the original image has been transformed, its DCT coefficients reflect the importance of the frequencies that are present in it. The DC-coefficient (the very first coefficient) shows the total illumination of the image, low frequency DCT coefficients are closely related to illumination variation, and high frequency coefficients show detail and fine information that have possibly been caused by noise. The coefficients located between the first and the last coefficients present different information levels for the original image [25].

For the JPEG compression standard, the image is initially partitioned into no overlapping blocks (8×8 blocks), and the DCT is performed independently on the sub-image blocks [24]. However, in our approach, the DCT is computed on the entire representation of the face image and the DCT coefficients are arranged in a zigzag scanning manner, in order to map the $M \times N$ dimension image to the $1 \times (N \times M)$ dimension vector and group low frequency coefficients at the top of the vector (see Fig. 7).

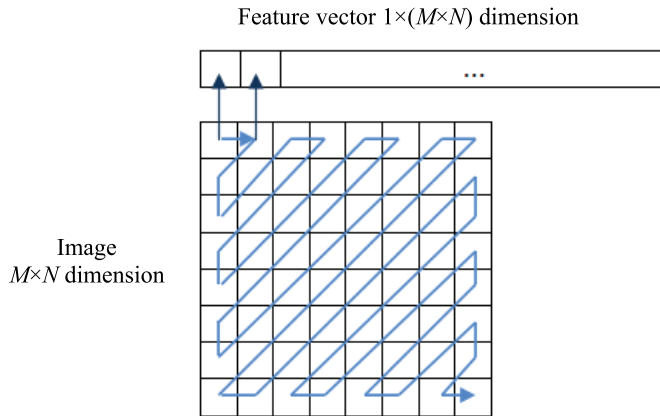


Fig. 7. Feature vector formation in DCT domain.

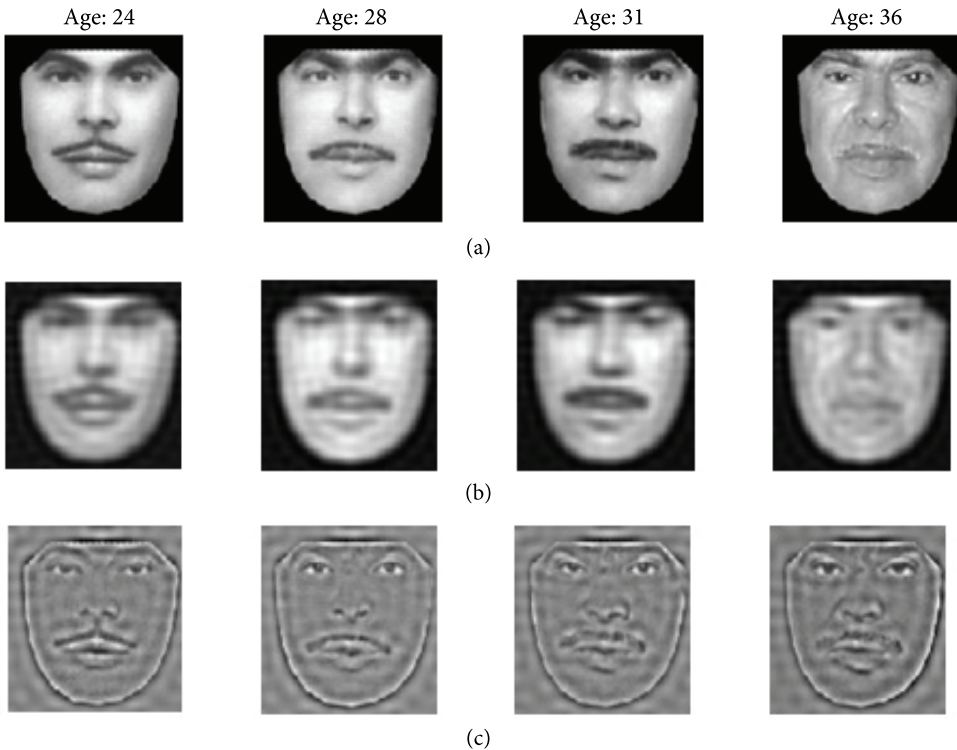


Fig. 8. Normalized images of a person at different ages and their reconstructed images. (a) Normalized images of a person at different ages, (b) reconstructed images by using 500 low frequency coefficients, and (c) reconstructed images by discarding 500 low frequency coefficients.

Since our purpose is to get age-invariant features, we know the following:

- Usually, facial aging consists of facial shape and skin texture variations.
- The appearance of a surface texture depends on illumination; thus, it changes under different illumination conditions.
- Illumination variations lie in the low frequency band.
- Generally, to recognize a person, we look for the principal characteristics of the face, such as the shapes and geometrical relationships of the main components of the face, including the eyes, nose and mouth, and we almost ignore some details related skin texture.
- As an example, the first row (a) in Fig. 8 displays four face images of the same person at different ages (24, 31, 42, and 61 years old) after normalization. The second row (b) in Fig. 8 shows reconstructed images of the same person after applying the DCT, taking a small set of low frequency coefficients, and applying the inverse DCT. The third row (c) in Fig. 8 shows reconstructed images of the same person after applying the DCT, setting a small set of low frequency coefficients to zero, and applying the inverse DCT. From the second and the third rows, we can see that reconstructed images discarding a set of low frequency coefficients present a diminution in some details (mainly, texture). However, the important features that characterize the face, such as the eyes, nose, etc., were preserved.

From these remarks, it can be concluded that the use of the DCT discarding some low frequency coefficients is crucial for successful feature extraction in face recognition across age progression.

3.5 Kernel Fisher Analysis

The Kernel Fisher Analysis, or the KFA method, is a kernelized version of linear discriminant analysis. It has been successfully applied to biometric recognition, such as palmprints [26] and face identification and verification [6,27,28].

In this approach, the input data is projected from the input space, \mathcal{R}^n , into an implicit high dimensional space, \mathcal{R}^f , known as a feature space, by a nonlinear kernel mapping function: $\Phi : \mathcal{R}^n \rightarrow \mathcal{R}^f$, $f > n$, and then, the Fisher Linear Discriminant Analysis is adopted to this feature space. Using the same terminology and algorithm to perform KFA as in [6].

Assuming that projected samples $\Phi(x_i)$ are centered in \mathcal{R}^f with the information of all samples and their classes, the between-class scatter matrix S_B^Φ and the within-class scatter matrix S_W^Φ are defined as:

$$S_B^\Phi = \sum_{i=1}^c l_i \mu_i^\Phi (\mu_i^\Phi)^T \quad (10)$$

$$S_W^\Phi = \sum_{i=1}^c \sum_{j=1}^{l_i} \Phi(X_{ij}) \Phi(X_{ij})^T \quad (11)$$

where l_i is the number of samples in class i , c is the number of classes, and μ_i^Φ is the mean of a class i in \mathcal{R}^f .

To apply LDA in a kernel space, we need to find the eigenvalues λ and eigenvectors W^Φ of:

$$\lambda S_W^\Phi W^\Phi = S_B^\Phi W^\Phi \quad (12)$$

which can be obtained by finding a set of vectors W_{OPT}^Φ that maximize the criterion:

$$W_{OPT}^\Phi = \arg \max_{W^\Phi} \left| \frac{(W^\Phi)^T \cdot S_B^\Phi \cdot W^\Phi}{(W^\Phi)^T \cdot S_W^\Phi \cdot W^\Phi} \right| = [W_1^\Phi \dots W_m^\Phi] \quad (13)$$

where, $\{W_i^\Phi, i = 1, 2, \dots, m\}$ are the eigenvectors corresponding to the m largest eigenvalues $\{\lambda_i, i = 1, 2, \dots, m\}$.

Consider c to be the number of classes and let the r^{th} sample of a class t and the s^{th} sample of a class u be X_{tr} and X_{us} , respectively, and the class t has l_t samples and the class u has l_u samples. The kernel function is then defined as:

$$(k_{rs})_{tu} = k(X_{tr}, X_{us}) = \Phi(X_{tr}) \cdot \Phi(X_{us}) \quad (14)$$

Let K be a $m \times m$ symmetric matrix defined by the elements $(k_{tu})_{u=1, \dots, c}^{t=1, \dots, c}$, where, k_{tu} is a $l_t \times l_u$ matrix composed of dot products in the feature space \mathcal{R}^f . For example:

$$K = (k_{tu})_{u=1, \dots, c}^{t=1, \dots, c}, \text{ where } k_{tu} = (k_{rs})_{s=1, \dots, l_u}^{r=1, \dots, l_t} \quad (15)$$

Let Z be a $m \times m$ block diagonal matrix and each block (Z_t) be a $l_t \times l_t$ matrix with terms all equal to $\frac{1}{l_t}$:

$$Z = (Z_t)_{t=1, \dots, c} \quad (16)$$

Any solution $W^\Phi \in \mathcal{R}^f$ must lie in the span of all training samples in \mathcal{R}^f . For example:

$$W^\Phi = \sum_{p=1}^c \sum_{q=1}^{l_p} \alpha_{pq} \Phi(X_{pq}) \quad (17)$$

It follows that the solution for (17) can be obtained by solving:

$$\lambda K K \alpha = K Z K \alpha \quad (18)$$

Therefore, Eq. (13) can be written as:

$$W_{OPT}^\Phi = \arg \max_{W^\Phi} \left| \frac{(W^\Phi)^T \cdot S_B^\Phi \cdot W^\Phi}{(W^\Phi)^T \cdot S_W^\Phi \cdot W^\Phi} \right| = \arg \max_{W^\Phi} \left| \frac{\alpha K Z K \alpha}{\alpha K K \alpha} \right| = [W_1^\Phi \dots W_m^\Phi] \quad (19)$$

To avoid singularity in computing W_{OPT}^Φ , we used the same technique in Fisherfaces [29].

4. Experimentation

4.1 Experimentation Methodology Description

All algorithms and normalization steps were implemented in MATLAB (R.2012a). To evaluate our

method, we used the entire FG-NET database. It included 1,002 images divided into 82 classes (ranging in age from 0 to 69), for which there were a different number of images per class, and these numbers varied from 4 to 12. The database was used for face verification or face identification experiments to verify and test the invariance of the proposed approach.

For easy comparison with other results obtained on the FG-NET database, we adopted the same leave-one-person-out (LOPO) evaluation scheme adopted by [8,18,19], since this is the evaluation scheme used most often for the FG-NET database to divide the database into training and testing sets. This strategy selected all of the images of one person for testing and used all remaining images as a training set. This was repeated for all of the persons in the database. Thus, in the case of the FG-NET database the experiment was done 82 times. Hence, all images of one person were either in the training or the testing set. This strategy was adopted in order to ensure that the training and the testing sets were separated.

As we have already mentioned, there is a large variation in expressions, poses, and illumination between images in the FG-NET database. To evaluate our method, all of the images were preprocessed using the steps already presented in Section 3.2. Then, we transformed all of the images in a frequency domain using the DCT, we ordered the coefficient results in zigzag order, and we set a subset of the first low frequency coefficients to zero. Then, without doing inverse DCT, we applied a KFA to the results for dimensionality reduction, we used the polynomial kernel in our experiments, and we empirically chose the corresponding parameter of the kernel (Polynomial degree). Lastly, recognition was carried out by using a K -nearest neighbor classification, where K was selected equal to one ($K=1$) to calculate the Rank-1 recognition rates.

4.2 Results

The results of our experiments can be shown in two ways:

- Table showing the identification rate at rank one (Rank-1 recognition rate) in the identification case and the equal error rate (EER) in the verification case. The results are the average of the 82 results we obtained from experiments using the LOPO strategy.
- Cumulative match characteristic (CMC) curve [30] to show the recognition rate for ranks one and higher.

The performance measures were calculated by using the MATLAB PhD (Pretty Helpful Development) toolbox for face recognition [31,32].

We performed three experimentations to evaluate the performance of our proposed method. In all of the experiments extraction features were followed by a subspace KFA for dimension reduction.

In the first experiment, we studied the effect of using the AAM pose correction method, and we compared the obtained results using the AAM pose correction with those obtained using the approach described below.

Images were first rotated to get the eyes at fixed points so that the inter-ocular segment was horizontal based on files of eye coordinates that were provided with the original FG-NET database. The images were then cropped (using the eyes coordinates) to remove the background and were resized to 128×128 . This approach is referred to in the following as manual alignment (MA). The results achieved clearly indicate the benefit of using the AAM as a pose correction method, where there is an

improvement of 19.15% for the identification rate and a decrease in EER from 22.68% to 7.2% for verification in the case where we used the DCT. There was also an improvement of 23.14% for the identification rate and a decrease in EER from 37.70% to 17.23% without using DCT.

In the second experiment, we studied the importance of using the DCT. First, we compared the obtained results by applying the DCT on manually aligned images with those without the DCT. Second, we compared the results by combining the AAM and DCT with those using only the AAM. From the results achieved in these two cases we observed that adding the DCT increased the recognition rate by 14.22% and decreased the EER by 15.02% for verification in the first case. For the second case, the recognition rate was increased by 10.23% and the EER was decreased by 10.03% for verification. This led to the conclusion that using the DCT provides useful information for overcoming the variances in a face recognition system that are caused by variations in age.

In the third experiment, we evaluated the performance of the DCT in the presence of age variations. We compared the results obtained by using the DCT on preprocessed images and set a few number of first coefficients to zero with those obtained without discarding coefficients. From Table 1 we can observe that applying DCT and discarding some coefficients achieved a rate of 61.41% in the case of identification and an EER of 7.2% in the case of verification. This had a 7.85% improvement for identification and 6.03% for verification, which demonstrates the impact of low frequency coefficients on the performance of face recognition in the presence of age variations.

To select the best percentage of low frequency coefficients to be discarded, we performed a succession of experiments. In each one we set a percentage of low frequency coefficients (0%, 5%, 10%, 15%, ..., 90%) to zero. The best recognition rate was between 10% and 15% (see Fig. 9). The recognition rate decreased when we chose high percentages, as seen in the last experiment where the obtained recognition rate was 0.6%. Obviously, there was a significant loss of information, which can be seen in Fig. 10. Fig. 10(a) represents the original image, Fig. 10(b)–(e) represent, respectively, reconstructed images after discarding 5%, 10%, 30%, and 85% of the low frequency coefficients. The last image Fig. 10(e) barely contains any information.

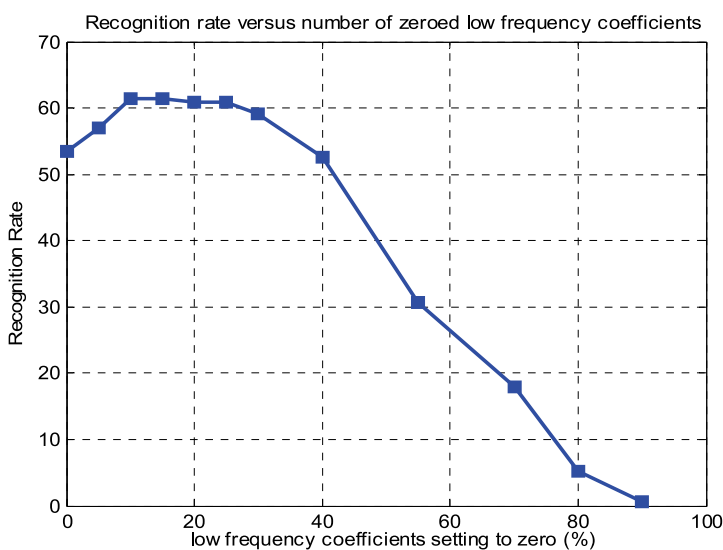


Fig. 9. Recognition rate versus number of zeroed low frequency coefficients.

Additionally, we demonstrated the importance of using the kernel method for dimensionality reduction. For this purpose we repeated the three experiments by using a Linear Fisher Analysis instead of the KFA. As seen in the results (see Table 1), the obtained rates with KFA are always higher than those obtained by using LDA.

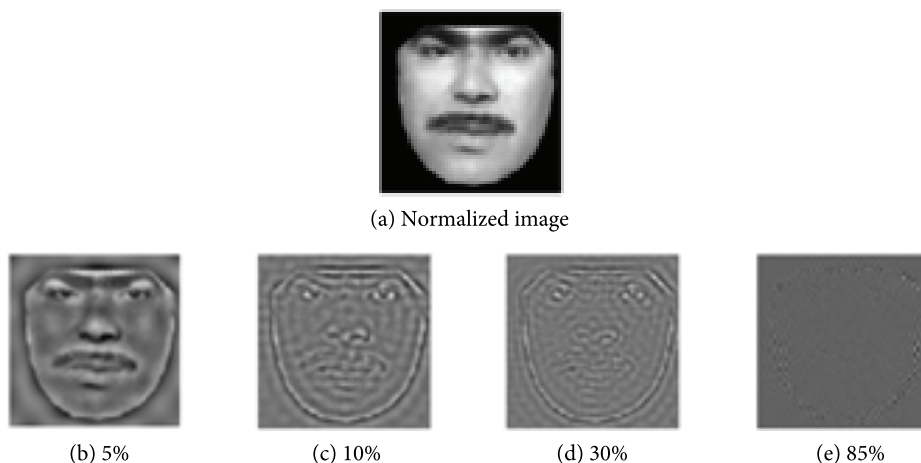


Fig. 10. Normalized image of a person and its reconstructed images by setting to zero a percentage of low frequency coefficients.

Table 1. Rank-1 identification rate and EER for verification for FG-NET evaluation

Method	Rank-1 identification rate (%)		EER (%)	
	LDA	KFA	LDA	KFA
MA	22.16	28.04	43.29	37.70
MA+DCT (-10% of coeff.)	36.85	42.26	27.28	22.68
AAM	46.51	51.18	19.45	17.23
AAM+DCT	47.70	53.56	16.25	13.23
AAM+DCT (-10% of coeff.)	51.58	61.41	13.85	7.20

EER=equal error rate, FG-NET=Face and Gesture Recognition Research Network, LDA=Linear Discriminant Analysis, KFA=Kernel Fisher Analysis, MA>manual alignment, DCT=discrete cosine transform, AAM=Active Appearance Model.

From the results presented in Table 1, we notice that for temporal change tasks, the DCT discards some low frequency coefficients when coupled with the AAM normalization, and the KFA projection method still produced the best performances out of all of the cases studied.

In Table 2, we summarize the results achieved by some approaches that have been proposed in the field of age-invariant face recognition and compared their results with ours.

From the results presented in Fig. 11, it is seen that our proposed approach provides the best result at rank-1 and always still the only best one at higher ranks. The importance of the preprocessed step, where always the combinations in which we use the AAM method (red, magenta and blue curve in Fig. 11) give the highest results, can also be seen.

It is worth noting that better results can probably be achieved if a component-based approach is used. The same process was applied for each component of the face (eyes, nose, and mouth) separately.

Table 2. Performance comparison for age-invariant face recognition

Author	Approach	Database	No. of images	No. of subjects	Rank-1 recognition rate (%)
Lanitis et al. [33]	Build an aging function in terms of PCA coefficients of shape and texture	Private database	85	12	68.5
Ramanathan and Chellappa [34]	Shape growth modeling up to age 18	Private database	109	109	15
Wang et al. [35]	Build an aging function in terms of PCA coefficients of shape and texture	Private database	NA	2,000	63
Geng et al. [36]	Learn aging pattern on concatenated PCA coefficients of shape and texture across a series of ages	FG-NET	10	10	38.1
Park et al. [8,37]	Learn aging pattern based on PCA coefficients in separate 3D shape and texture spaces from the given 2D database	FG-NET	1002	82	37.4
		MORPH I	612	612	66.4
		MORPH II	20,000	10,000	79.8
		Browns	100	100	28.1
Juefei-Xu et al. [19]	Walsh-Hadamard transform encoded local binary patterns (WLBP) on preprocessed periocular region Unsupervised discriminant projection (UDP) as subspace modeling technique	FG-NET	1,002	82	100
Yang et al. [18]	Texture embedded discriminative graph matching Local Gabor Binary Pattern Histogram sequence (LGBPHS) projected in LDA subspace	FG-NET	1,002	82	64.47
Li et al. [38]	Multi-feature discriminant analysis (MFDA) combined with scale invariant feature transform (SIFT) and multi-scale local binary patterns (MLBP) as densely sampled local descriptors	FG-NET	1,002	82	47.5
		MORPH II	20,000	10,000	83.9
Bereta et al. [15]	Divers local texture descriptors combined with several distances based classifiers	FG-NET	1,002	82	<45
Proposed approach	DCT on preprocessed image using AAM, projected in KFA subspace	FG-NET	1,002	82	61.41

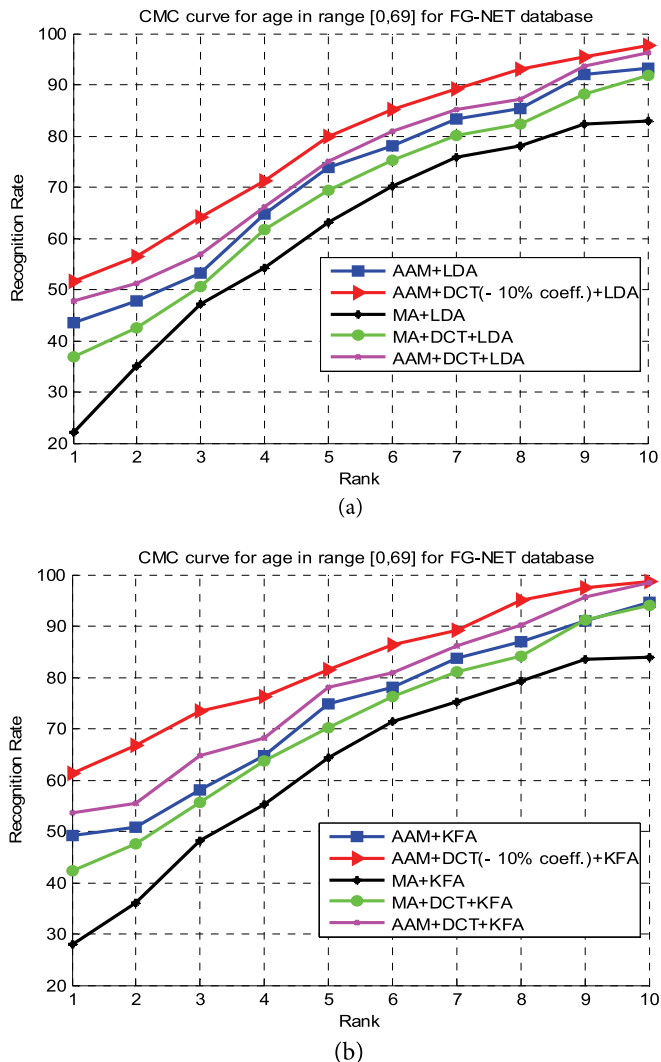


Fig. 11. Cumulative match characteristic (CMC) curve for experiments on FG-NET database. (a) CMC curve by using Linear Discriminant Analysis (LDA) and (b) CMC curve by using Kernel Fisher Analysis (KFA).

5. Conclusions

In this paper, we have presented an age-invariant feature extraction method for face identification and verification across age variations that consists of preprocessing steps that use the AAM, combined with the DCT, and followed by the KFA projection method. The proposed method provided a 61.41% Rank-1 identification and 7.2% EER for verification. The obtained results from experiments on the FG-NET database encourage the use of a combination of the AAM and DCT for face recognition across age progression. In the future, we will test our method on the MORPH database, study its robustness over large time lapses, and compare the obtained results with those obtained using other classifiers, such as SVM.

References

- [1] A. Lanitis, "Facial biometric templates and aging: problems and challenges for artificial intelligence," in *Proceedings of 5th IFIP Conference on Artificial Intelligence Applications & Innovations (AIAI)*, Thessaloniki, Greece, 2009, pp. 142-149.
- [2] K. Ricanek and T. Tesafaye, "Morph: a longitudinal image database of normal adult age-progression," in *Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition*, Southampton, UK, 2006, pp. 341-345.
- [3] FG-NET aging database [Online]. Available: <http://www.fgnet.rsunit.com>.
- [4] N. Ramanathan, R. Chellappa, and S. Biswas, "Age progression in human faces: a survey," *Journal of Visual Languages and Computing*, vol. 15, pp. 3349-3361, 2009.
- [5] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 681-685, 2001.
- [6] M. H. Yang, "Kernel eigenfaces vs. kernel fisherfaces: face recognition using kernel methods," in *Proceedings of the 5th International Conference on Automatic Face and Gesture Recognition*, Washington, DC, 2002, pp. 215-220.
- [7] N. Ramanathan and R. Chellappa, "Modeling shape and textural variations in aging faces," in *Proceedings of the 8th International Conference on Automatic Face and Gesture Recognition*, Amsterdam, Netherlands, 2008, pp. 1-8.
- [8] U. Park, Y. Tong, and A. K. Jain, "Age-invariant face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 5, pp. 947-954, 2010.
- [9] H. Ling, S. Soatto, N. Ramanathan, and D. W. Jacobs, "A study of face recognition as people age," in *Proceedings of the 11th International IEEE Conference on Computer Vision*, Rio de Janeiro, 2008, pp. 1-8.
- [10] H. Ling, S. Soatto, N. Ramanathan, and D. W. Jacobs, "Face verification across age progression using discriminative methods," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 1, pp. 82-91, 2010.
- [11] G. Mahalingam and C. Kambhamettu, "Face verification with aging using AdaBoost and local binary patterns," in *Proceedings of the 7th Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP)*, Chennai, India, 2010, pp. 101-108.
- [12] R. Singh, M. Vatsa, A. Noore, and S. K. Singh, "Age transformation for improving face recognition performance," in *Proceedings of 2nd International Conference on Pattern Recognition and Machine Intelligence (PReMI)*, Kolkata, India, 2007, pp. 576-583.
- [13] S. Biswas, G. Aggarwal, N. Ramanathan, and R. Chellappa, "A non-generative approach for face recognition across aging," in *Proceedings of the 2nd IEEE International Conference on Biometrics: Theory, Application and Systems (BTAS)*, Arlington, VA, 2008, pp. 1-6.
- [14] G. Mahalingam and C. Kambhamettu, "Age invariant face recognition using graph matching," in *Proceedings of the 4th IEEE International Conference on Biometrics: Theory, Application and Systems (BTAS)*, Washington, DC, 2010, pp. 1-7.
- [15] M. Bereta, P. Karczmarek, W. Pedrycz, and M. Reformat, "Local descriptors in application to the aging problem in face recognition," *Pattern Recognition*, vol. 46, no. 10, pp. 2634-2646, 2013.
- [16] D. Sungatullina, J. Lu, G. Wang, and P. Moulin, "Multiview discriminative learning for age-invariant face recognition," in *Proceedings of the 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition*, Shanghai, China, 2013, pp. 1-6.
- [17] L. Wiskott, J. M. Fellous, N. Kuiger, and C. Von Der Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 775-779, 1997.

- [18] H. Yang, D. Huang, and Y. Wang, "Age invariant face recognition based on texture embedded discriminative graph model," in *Proceedings of 2014 IEEE International Joint Conference on Biometrics (IJCB)*, Clearwater, FL, 2014, pp. 1-8.
- [19] F. Juefei-Xu, K. Luu, M. Savvides, T. D. Bui, and C. Y. Suen, "Investigating age invariant face recognition based on periocular biometrics," in *Proceedings of 2011 IEEE International Joint Conference on Biometrics (IJCB)*, Washington, DC, 2011, pp. 1-7.
- [20] T. F. Cootes and C. J. Taylor, "Statistical models of appearance for computer vision," Department of Imaging Science and Biomedical Engineering, University of Manchester, UK, 2004.
- [21] Z. M. Hafed and M. D. Levine, "Face recognition using the discrete cosine transform," *International Journal of Computer Vision*, vol. 43, no. 3, pp. 167-188, 2001.
- [22] J. Jiang and G. Feng, "Robustness analysis on facial image description in DCT domain," *Electronics Letters*, vol. 43, no. 24, pp. 1354-1355, 2007.
- [23] K. R. Rao and P. Yip, *Discrete Cosine Transform: Algorithms, Advantages, Applications*. Boston, MA: Academic Press, 1990.
- [24] W. B. Pennebaker and J. L. Mitchell, *JPEG: Still Image Data Compression Standard*. New York: Springer, 1993.
- [25] W. Chen, M. J. Er, and S. Wu, "Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 36, no. 2, pp. 458-466, 2006.
- [26] Y. Wang and Q. Ruan, "Kernel fisher discriminant analysis for palmprint recognition," in *Proceedings of the 18th International Conference on Pattern Recognition*, Hong Kong, 2006, pp. 457-460.
- [27] C. Liu, "Capitalize on dimensionality increasing techniques for improving face recognition grand challenge performance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 5, pp. 725-737, 2006.
- [28] Q. Liu, H. Lu, and S. Ma, "Improving kernel Fisher discriminant analysis for face recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 42-49, 2004.
- [29] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: recognition using class specific linear projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711-720, 1997.
- [30] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The FERET evaluation methodology for face-recognition algorithms," *IEEE Transactions on Pattern Recognition and Machine Intelligence*, vol. 22, no. 10, pp. 1090-1104, 2000.
- [31] V. Struc and N. Pavesic, "The complete gabor-fisher classifier for robust face recognition," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, article ID. 847680, 2010.
- [32] V. Struc, and N. Pavesic, "Gabor-based kernel partial-least-squares discrimination features for face recognition," *Informatica*, vol. 20, no. 1, pp. 115-138, 2009.
- [33] A. Lanitis, C. J. Taylor, and T. F. Cootes, "Toward automatic simulation of aging effects on face images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 4, pp. 442-455, 2002.
- [34] N. Ramanathan and R. Chellappa, "Modeling age progression in young faces," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, New York, NY, 2006, pp. 387-394, 2006.
- [35] J. Wang, Y. Shang, G. Su, and X. Lin, "Age simulation for face recognition," in *Proceedings of the 18th IEEE International Conference on Pattern Recognition*, Hong Kong, 2006, pp. 913-916.
- [36] X. Geng, Z. H. Zhou, and K. Smith-Miles, "Automatic age estimation based on facial aging patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 12, pp. 2234-2240, 2007.
- [37] U. Park, Y. Tong, and A. K. Jain, "Face recognition with temporal invariance: a 3D aging model," in *Proceedings of the 8th IEEE International Conference on Automatic Face & Gesture Recognition*, Amsterdam, 2008, pp. 1-7.
- [38] Z. Li, U. Park, and A. K. Jain, "A discriminative model for age invariant face recognition," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 1028-1037, 2011.



Leila Boussaad

She received the Dipl. Ing. degree in Computer Science in 1999 and the M.Sc. degree in computer science in 2009, both from Batna University, Algeria. She is now a PhD candidate in computer science at the same university. Currently, she works as an assistant professor in the department of Economics at Batna university, Algeria. Her research interests include Content-Based Image Retrieval, Biometry, and Pattern Recognition.



Mohamed Benmohammed

He is a professor in computer science at university of Constantine 2, Algeria. He received the Ph.D. degree from university of Sidi Bel-Abbes, Algeria, in 1997. His research interests include: Microprocessor, Computer Architecture, Embedded system and Computer Networks.



Redha Benzid

He received the Dipl. Ing. degree in electronics in 1994, the M.Sc. degree in electronics in 1999, and the Ph.D. degree in electronics in 2005, all obtained from Batna University, Algeria. Currently, he is a full Professor in the Department of Electronics at Batna University, Algeria. He co-authored several papers in refereed international journals and served as reviewer in Biomedical Signal Processing and control, IET Electronics letters, Digital Signal Processing and many other journals. His research interests include: Data Compression, Biomedical Engineering, Biometry, Digital Watermarking and Visual Cryptography.