World Academy of Science, Engineering and Technology
International Journal of Electrical and Computer Engineering
Vol:3, No:6, 2009

A case study on appearance based feature extraction techniques and their susceptibility to image degradations for the task of face recognition

Vitomir Štruc and Nikola Pavešić, *Member, IEEE*

Abstract—Over the past decades, automatic face recognition has become a highly active research area, mainly due to the countless application possibilities in both the private as well as the public sector. Numerous algorithms have been proposed in the literature to cope with the problem of face recognition, nevertheless, a group of methods commonly referred to as appearance based have emerged as the dominant solution to the face recognition problem. Many comparative studies concerned with the performance of appearance based methods have already been presented in the literature, not rarely with inconclusive and often with contradictory results. No consent has been reached within the scientific community regarding the relative ranking of the efficiency of appearance based methods for the face recognition task, let alone regarding their susceptibility to appearance changes induced by various environmental factors. To tackle these open issues, this paper assess the performance of the three dominant appearance based methods: principal component analysis, linear discriminant analysis and independent component analysis, and compares them on equal footing (i.e., with the same preprocessing procedure, with optimized parameters for the best possible performance, etc.) in face verification experiments on the publicly available XM2VTS database. In addition to the comparative analysis on the XM2VTS database, ten degraded versions of the database are also employed in the experiments to evaluate the susceptibility of the appearance based methods on various image degradations which can occur in "real-life" operating conditions. Our experimental results suggest that linear discriminant analysis ensures the most consistent verification rates across the tested databases.

Keywords—Biometrics, face recognition, appearance based methods, image degradations, the XM2VTS database.

I. Introduction

FEATURE extraction techniques are an important part of each face recognition system with an crucial impact on the performance of such systems. Clearly, a detailed knowledge of the characteristics of the employed feature extraction techniques, especially their susceptibility to facial appearance changes induced by various environmental factors, is required to design effective recognition schemes and consequently to construct robust face recognition systems. Researchers have, therefore, conducted numerous studies comparing various feature extraction techniques and their robustness to facial appearance changes.

In a vast majority of these studies special attention is given to the so-called appearance based methods, which represent the dominant and most popular feature extraction techniques used in the field of face recognition. Appearance based methods, such as principal component analysis (PCA)[1], linear discriminant analysis (LDA)[2] or independent component analysis (ICA)[3], project facial images into a lower dimensional subspace and then classify the images based on some distance (or similarity) with images of known, i.e., enrolled, individuals. However, even though a number of comparative studies exits, researchers often report contradictory results regarding the performance of the appearance based methods.

Liu and Wechsler [4] and Bartlett et al. [3], for example, claim that ICA outperforms PCA, while Baek et al. [5] report to achieve better results with PCA [6]. Delac et al. [6] state that the performance of the appearance based methods is heavily dependent on the employed distance measure and that with the right combination of appearance based method and distance no claim regarding the superiority of any of the three techniques, i.e., PCA, LDA, ICA, can be made. Beveridge et al. [7], on the other hand, report that in their experiments PCA systematically outperformed LDA, whereas Belhumeur et al. [2] claim that LDA performs better than PCA in all of their tests. Li et al. [8] suggest that the performance of LDA depends on its implementation, i.e., on the numeric method used for solving LDAs eigenproblem, and that if implemented appropriately always outperforms PCA.

From the presented discussion we can see that no consent has been reached within the scientific community regarding the relative ranking of the appearance based methods for the face recognition task, let alone regarding their susceptibility to appearance changes induced by various environmental factors. In this paper we, therefore, perform a comparative analysis of the performance of the three dominant feature extraction techniques: PCA, LDA and ICA. We conduct face verification experiments using the popular XM2VTS database and assess the susceptibility of the appearance based methods to various image degradations encountered in real life operation conditions. The experiments are carried out within a fully automatic face recognition system, i.e., with an automatic face localization module, and, hence represent one of the few attempts to assess the appearance based feature extraction techniques in real working conditions.

The rest of the paper is structured as follows. Section II introduces the basic concepts of PCA, LDA and ICA. In Section III the problem of face verification is given. Section IV describes the employed XM2VTS database and its degradations. Section V presents the experiments and corresponding results, while Section VI concludes the paper.

V. Štruc and N. Pavešić are with the Faculty of Electrical Engineering, University of Ljubljana, Tržaška 25, SI-1000 Ljubljana, Slovenia e-mail: (see http://luks.fe.uni-lj.si/en/staff/index.html).

II. Appearance based feature extraction techniques

This section briefly reviews the theory underlying the three dominant appearance based feature extraction techniques, namely, principal component analysis, linear discriminant analysis and independent component analysis.

A. Principal component analysis

Principal component analysis (PCA), first introduced to the field of face recognition by Pentland and Turk in [1], is a powerful appearance based feature extraction technique, which applies the Karhunen-Loève transform to a set of training images and derives a number of projection axes that act as the basis vectors for the PCA subspace. Each face image is represented as a vector of projection coefficients in this subspace, in which information compression, dimensionality reduction and de-correlation of the pixel values of the face images is achieved [9].

Given a set of C $d-$dimensional training images[1] $\mathbf{x}_i$ (for $i = 1, 2, ..., C$) arranged into a $d \times C$ column data matrix $\mathbf{X} \in \mathbb{R}^{N \times d}$, principal component analysis derives a mapping $\mathbf{W}$ from the original $d-$dimensional space to a lower $d'-$dimensional space, i.e.:

$$\mathbf{y}_i = \mathbf{W}^T(\mathbf{x}_i - \boldsymbol{\mu}), \tag{1}$$

where $\mathbf{y}_i$ ($i = 1, 2, ..., C$) denotes the $d'-$dimensional PCA feature vector corresponding to the $i-$th facial image $\mathbf{x}_i$ and $\boldsymbol{\mu}$ stands for the global mean of al training images. The mapping is defined by a set of basis vectors which correlate to the maximum variance directions present in the training data. Hence, each training image can be projected into this subspace and again reconstructed with minimum error. From the mathematical point of view the mapping $\mathbf{W}$ corresponds to the leading eigenvectors of the covariance matrix of the training data, i.e.:

$$\mathbf{W} = [\mathbf{w}_1^T, \mathbf{w}_2^T, ..., \mathbf{w}_{d'}^T,], \tag{2}$$

where the eigenvectors $\mathbf{w}_j^T$, for $j = 1, 2, ..., d'$, represent the leading eigenvectors of the following eigenproblem:

$$\mathbf{K}\mathbf{w}_j = \lambda_j \mathbf{w}_j \tag{3}$$

and $\mathbf{K}$ stands for the covariance matrix of the training data.

PCA is often referred to as the Eigenface technique due to the fact that the eigenvectors comprising the transformation matrix $\mathbf{W}$ resemble facial images as shown in Fig. 1.

Fig. 1. The first five leading eigenvectors of the covariance matrix of the training data - first five Eigenfaces

[1]Note that for the purpose of PCA each image $I(x, y)_i$ of size $a \times b$ pixels is rearranged into a $d-$dimensional pattern vector, where $d = ab$.

B. Linear discriminant analysis

Unlike PCA, which considers only the variance of the training images to construct a subspace, linear discriminant analysis (LDA) aims at improving upon PCA by also taking the class-membership information of the training images into account when seeking for a subspace. From this point of view PCA is usually considered as being more appropriate for the task of data compression, while LDA is tailored more towards the classification task [10].

Given C d-dimensional training images arranged into a $d \times C$ column data matrix $\mathbf{X}$ and each belonging to one of N classes $\omega_1, \omega_2, ..., \omega_N$, one first computes the between-class and the within-class scatter matrices:

$$\boldsymbol{\Sigma}_B = \sum_{i=1}^{N} n_i(\boldsymbol{\mu}_i - \boldsymbol{\mu})(\boldsymbol{\mu}_i - \boldsymbol{\mu})^T \tag{4}$$

and

$$\boldsymbol{\Sigma}_W = \sum_{i=1}^{N} \sum_{\mathbf{x}_j \in \omega_i} (\mathbf{x}_j - \boldsymbol{\mu}_i)(\mathbf{x}_j - \boldsymbol{\mu}_i)^T \tag{5}$$

and then derives the LDA transformation matrix $\mathbf{W}$ which maximizes Fisher's discriminant criterion [2], [10]:

$$J(\mathbf{W}) = arg \max_{\mathbf{W}} \frac{|\mathbf{W}^T \boldsymbol{\Sigma}_B \mathbf{W}|}{|\mathbf{W}^T \boldsymbol{\Sigma}_W \mathbf{W}|}, \tag{6}$$

where n_i denotes the number of images in the i-th class, $\boldsymbol{\mu}_i$ stands for the class conditional mean and $\boldsymbol{\mu}$ represents the global mean of all training images. Fisher's discriminant criterion is maximized when $\mathbf{W}$ is constructed by a simple concatenation of the $d' \leq N - 1$ leading eigenvectors of the following eigenproblem:

$$\boldsymbol{\Sigma}_W^{-1} \boldsymbol{\Sigma}_B \mathbf{w}_i = \lambda_i \mathbf{w}_i, \quad i = 1, 2, ..., d', \tag{7}$$

that is $\mathbf{W} = [\mathbf{w}_1 \mathbf{w}_2 \cdots \mathbf{w}_{d'}]$. Once the transformation matrix $\mathbf{W}$ is calculated, it can be used to project a test image $\mathbf{x}$ into the LDA subspace, thus reducing the pattern's dimension from d to d':

$$\mathbf{y} = \mathbf{W}^T(\mathbf{x} - \boldsymbol{\mu}), \tag{8}$$

where $\mathbf{y}$ represents the d'-dimensional projection of the centered pattern $\mathbf{x}$.

Unfortunately, the number of training images available in the field of face recognition is commonly significantly smaller than the image's dimension rendering the within-class scatter matrix $\boldsymbol{\Sigma}_W$ singular and thus computation of the matrix $\mathbf{W}$ using the eigenproblem in (7) impossible. To overcome this problem, we used a modification of LDA, which first projects all images into the PCA subspace (to reduce their dimensionality and consequently ensures that the matrix $\boldsymbol{\Sigma}_W$ is invertible) and then performs LDA in the reduced space [10]. The presented technique is often referred to as the Fisherface method as the eigenvectors comprising the transformation matrix $\mathbf{W}$ again resemble faces as shown in Fig. 2.

Fig. 2. The first five Fisherfaces

C. Independent component analysis

Independent component analysis (ICA) [3] is an extension of PCA, which in addition to second-order statistics present in the training images tries to minimize higher-order dependencies in the training images as well. It does so by seeking basis vectors along which the projected images are statistically independent [11]. Unlike PCA and LDA, which represent orthogonal transformations (i.e., the basis vectors comprising the transformation matrices are orthogonal one to another), ICA represents a non-orthogonal transformation. Several algorithms have been developed to perform ICA, the most prominent definitely being Jade, InfoMax and FastICA. The latter was also employed in our experiments using the FastICA toolbox [12] for Matlab.

Two different architectures of ICA have been proposed for the purpose of face recognition. The first, commonly referred to as ICA Architecture I (ICA1) seeks statistically independent basis vectors, while the second, commonly referred to as ICA Architecture II (ICA2) seeks statistically independent features (or projection coefficients). Both ICA architectures can be implemented by subjecting either the transformation matrix of the PCA technique or the PCA feature vectors corresponding to the training images to the FastICA algorithm. For a more detailed description of the ICA technique the reader should refer to [3], [4] or [13].

Again we can visualize the basis vectors of both ICA architectures and notice that they again resemble facial images as shown in Fig. 3 and Fig. 4.

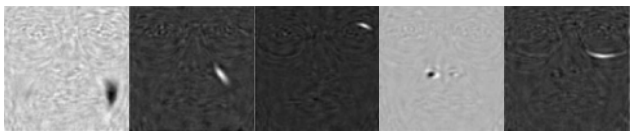

Fig. 3. Five examples of ICA1 basis vectors in image form

Fig. 4. Five examples of ICA2 basis vectors in image form

III. The verification procedure

In a face verification system the problem of verification can be defined as follows: given a $d'-$dimensional feature vector $\mathbf{y}$ and a claimed identity ω_i, for $i = 1, 2, ..., N$ (where N denotes the number of clients/identities enrolled in the system), classify the feature vector $\mathbf{y}$ into one of two classes w_1 or w_2 [14], i.e.,

$$(\omega_i, \mathbf{y}) \in \begin{cases} w_1, & \text{if } \{\delta(\mathbf{y}, \overline{\mathbf{y}}_i)\} \geq \Delta,\ i = 1, 2, ..., N \\ w_2, & \text{otherwise} \end{cases},$$

where w_1 denotes the client[2] class and w_2 stands for the impostor class.

The classification of the vector $\mathbf{y}$ is based on a matching procedure, in which the input (or "live") feature vector $\mathbf{y}$ is matched against the user-template $\overline{\mathbf{y}}_i$ corresponding to the claimed identity ω_i using a similarity measure $\delta(,)$. If the live feature vector and the user-template display a degree of similarity $\delta(\mathbf{y}, \overline{\mathbf{y}}_i)$ that is higher than a predefined decision threshold Δ, then the feature vector is classified as belonging to class w_1, otherwise it is classified as belonging to the impostor class w_2.

In this paper we use the normalized correlation coefficient (NC) as the similarity measure. The NC is a preferred choice when matching feature vectors extracted from facial images with appearance based feature extraction techniques and is defined as follows:

$$\delta_{NC}(\mathbf{y}, \overline{\mathbf{y}}_i) = \frac{|\mathbf{y} \cdot \overline{\mathbf{y}}_i|}{\| \mathbf{y} \| \| \overline{\mathbf{y}}_i \|}, \tag{9}$$

where the operator $\cdot$ denotes the dot product and $\| \cdot \|$ stands for the norm operator.

IV. The experimental databases

A. The XM2VTS database

For the comparative assessment of the appearance based feature extraction techniques we make use of the publicly available XM2VTS database, which was recorded at the University of Surrey, UK [17]. The database comprises 2360 color face images, which correspond to 295 distinct subjects. Thus, each subject in the database is accounted for with 8 facial images, which were captured during four separate recording sessions distributed over a period of approximately five months.

The recording conditions at each of the sessions were supervised and, hence, the variability in the appearance of the facial images of the same subjects is induced mainly by the temporal component (e.g., the facial images differ due to a different hairstyle, the presence/absence of glasses or make-up, different facial expressions, etc.), while other factors, such as the ambient illumination or the pose of the subject, are controlled. Some sample images from the database are presented in Fig. 5.

B. The degraded XM2VTS database

In the experiments presented in the next section the original XM2VTS database serves for assessing the baseline performance of the appearance based feature extraction

[2]In a face verification system the term client refers to a legitimate user for which a client model/template exists in the systems database, as opposed to the term impostor, which refers to a possible intruder claiming a false identity and, hence, trying to compromise the verification system.

Fig. 5. Sample images form the original (undegraded) XM2VTS database

techniques. To evaluate the robustness of the techniques to various image degradations, i.e., their susceptibility to image degradations, we employ the degraded form of the XM2VTS database presented by Pavešić et al. in [18]. The degradations used to alter the original database are presented in the remainder of this section.

Packet-loss simulation: The first degradation technique represents a simulation of loss of data packets, which can occur when a facial image is transmitted over a computer network via data packets. To simulate the loss of data packets Pavešić et al. employed a probabilistic model defined by the following two parameters:

- p - the probability that a correctly transmitted packet follows a lost packet, and
- q - the probability that a lost packet follows a correctly transmitted packet.

For the degradation (denoted as D1 in the experimental section) a value of $p = 0.1$ and $q = 0.3$ were selected.

Adding noise: The second degradation technique employed in [18] to render the facial images is an noise addition procedure. Here, the noise level is controlled by the parameter η, which defines the interval [-$\eta/2$,+$\eta/2$] from which integer values are randomly drawn and added to the pixel intensity values of each of the RGB color components individually. This process is repeated for all pixel locations and consequently results in noisy facial images. In Section V the degradations with the presented techniques for $\eta = 100$ and $\eta = 50$ are denoted as D2 and D3, respectively.

Background manipulation: The third degradation technique replaces the simple, uniform background of the original images from the XM2VTS database with a more complex background. To this end, 18 real-world images are used to replace the original background, many of which contain skin-color objects, which make the localization of the facial region more difficult. In the experimental section this degradation technique is denoted as D4.

JPEG compression: The fourth degradation technique used by Pavešić et al. represents the popular JPEG compression. JPEG compression is a lossy compression technique, where the amount of compression is adjusted as a compromise between the desired file size and the desired image quality. The JPEG quality parameter takes values from the integer-valued interval [0,100], however, for the degradation of the XM2VTS database, a value of 10 was selected. The error rates obtained in face verification experiments with JPEG compressed facial images can be found in Section V under the label D5.

Reducing the color depth: The fifth degradation technique used in our experiments reduces the color depth of the facial images. The color depth reduction is achieved by reducing the number of bits per image color component to less than eight. For each pixel and each color component, the predefined number of the most significant bits is left unchanged, while the remaining bits are set to zero. In the degraded form of the database solely two bits are left unchanged, al the remaining bits are set to zero. In the experimental section this technique is labeled with D6.

Blur: The sixth image degradation technique represents a blurring procedure. Here, a two dimensional zero-mean Gaussian is used for the degradation, i.e., the blurring, of the facial images:

$$\Psi(x, y) = \frac{1}{2\pi\sigma} \exp(-\frac{x^2 + y^2}{2\sigma^2}), \tag{10}$$

where σ denotes the standard deviation of the Gaussian.

Clearly, the degree of smoothing is controlled be the value of σ. In the work of Pavešić et al. [18] two different values of σ were used in the degradation process, namely, $\sigma = 2$ and $\sigma = 4$. These two degradation techniques are denoted as D7 and D8 in the experimental section, respectively.

Geometric transformations: The seventh degradation technique is comprised of random geometrical transformations (translation-, rotation- and scaling-procedures) of the facial images. The transformations are limited with the following parameters:

- translation: $x \pm 144$ and $y \pm 115$, where the translations are measures in pixels,
- rotation: 0 ± 30 degrees, and
- scaling: 1 ± 0.3.

Geometric transformations are denoted as D9 in Section V.

Partial occlusion: The last image degradation technique employed by Pavešić et al. [18] is an occlusion procedure, where part of the pixels comprising the facial region are set to zero (i.e., the color of the face is changed to black). The procedure denoted as D10 in the experimental section is limited to occlude at most 30% of the face.

Some examples of the visual effects of the presented degradation procedures on images from the XM2VTS database are shown in Fig. 6. Here, the first image in the upper left corner represents the original image from the XM2VTS database, while the remaining images are ordered in accordance with the introduction of the corresponding degradation techniques in this paper.

C. Image preprocessing

In the experimental section we perform experiments with automatically localized facial images. Even though, manually

Fig. 6. Examples of the visual effect of the image degradation techniques. Upper row (from left to right): original image, D1, D2, D3, D4, D5; lower row (from left to right): D6, D7, D8, D9, D10

marked eye-coordinates are often used in the literature to show the efficiency of the assessed feature extraction technique, we consider face recognition as a complex problem where the face localization procedure cannot simply be separated from the feature extraction and recognition stages. To this end, we use an automatic face localization procedure to find the location of the facial region in the images of the XM2VTS database. The employed localization procedure, a block diagram of which is shown in Fig. 7, is comprised of the following steps:

- the detection of the ”skin-color” pixels using a specially designed skin-color detector (Fig. 7b),
- a conversion of the retained skin-color region to its grey-scale form (Fig. 7c),
- an edge detection procedure (using Sobel operators - Fig. 7d) followed by an image thresholding procedure to produce a binary form of the facial region (Fig. 7e),
- an alignment procedure (using horizontal and vertical integral projections), which compensates for possible deviations of the face from the frontal position (Fig. 7f),
- an elliptical Hough transform to limit the image-region in which to search for the face (Fig. 7f), and finally
- a Hausdorff-distance-based localization procedure, which determines the location of the facial region by finding the shortest distance amongst the Hausdorff-distances computed between different sub-windows of the thresholded binary image and a preconstructed edge-based model of the face (Fig. 7e).

An example of the described procedure for a facial image degraded using geometrical transformations (D9) is shown in Fig. 7.

Once the face region is located, it is converted to a monochromatic, i.e., grey-scale, form and cropped to a standard size of 100×100 pixels. No additional geometric normalization is needed since the compensation for any head rotation potentially present in the facial image is already incorporated into the face localization procedure. After the cropping step, the facial region is photometrically normalized by removing the mean of the pixel intensities and scaling the result with their standard deviation.

It has to be noted that the verification errors produced by a face verification system do not depend only on the employed feature extraction technique, but are also subject to face localization errors. Hence, the error rates presented in Section V are influenced by both, the localization as well as the feature extraction procedures and, therefore, reflect the susceptibility of the whole face recognition system to the given image degradation.

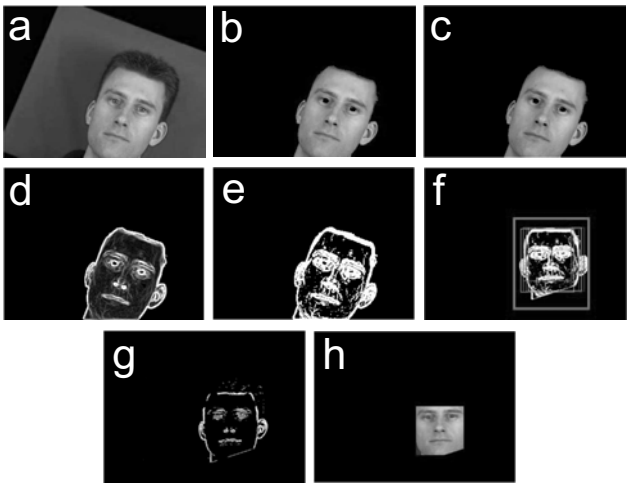

Fig. 7. An Example of the employed localization procedure: (a) a degraded image from the XM2VTS database using geometrical transformations, (b) detection of the ”skin-color” pixels, (c) grey-scale version of ”skin-color” pixels, (d) edge-image obtained with Sobel operators, (e) thresholded binary edge-image, (f) enframed facial region using a linear combination of frames that correspond to the top-five ellipses found with the elliptical Hough transform, (e) detected face using a face model and the Hausdorff distance, (f) final localized facial region

V. Experiments

This section presents the experiments and corresponding results obtained with the original as well as the degraded XM2VTS databases. First, the experimental setup is briefly introduced and then the actual experiments are described.

A. Experimental setup

Our experiments have been performed in accordance with the experimental protocol associated with XM2VTS database.

The protocol, known as the Lausanne protocol [17], partitions the database into image sets used for:

- training and enrollment - images, which are used to train the PCA, LDA and ICA feature extractors and build the client models/templates in form of mean feature vectors,
- evaluation - images, which are used to determine the operating point, i.e., the decision threshold, of the face verification system and to define any potential parameters of the feature extractor (e.g., number of features, selection of features, etc.), and
- testing - images, which are used to determine the verification error rates in real operating conditions and to assess the susceptibility of the face recognition system to the tested image degradations.

While the first image set features only images belonging to the client group, the latter two image sets comprise images belonging to both the client and the impostor groups. The client images are employed to assess the first kind of error a face verification system can make, namely, the false rejection error, whereas the impostor images are used to evaluate the second type of possible verification error, namely, the false acceptance error. The two errors are quantified by two corresponding error rates, namely, the false rejection and false acceptance error rates (FRR and FAR), which are defined as the relative frequency whit which a face verification system falsely rejects a legitimate- and falsely accepts an impostor-identity-claim, respectively.

As already presented in Section III, the decision regarding the validity of the identity claim is made based on the value of the matching score between the ”live” feature vector and the client model corresponding to the claimed identity, as well as the value of the decision threshold Δ. Unfortunately, both the FAR and the FRR depend on the value of Δ and, hence, selecting a threshold, which ensures a small value of FRR, inevitably results in a high value of the FAR and vice versa, selecting a threshold, which ensures a small value of the FAR results in a high value of the FRR. To effectively compare the performance of two face verification systems an operating point, i.e., a fixed value of Δ, has to be determined in advance. Commonly, the so-called equal error rate (EER) operating point (the EER operating point is defined with the decision threshold Δ which ensures equal values of the FAR and FRR on some evaluation image set) is chosen for this purpose. The mean value of the FAR and FRR, referred to as the half total error rate (HTER), is then used as a single comparative metric for different face verification systems.

The Lausanne protocol defines two distinct configurations on how to distribute the images into the training, evaluation and test sets. We, however, use the first configuration of the protocol, as it is considered the more difficult one of the two configurations. In accordance with the first experimental configuration, we perform 600 client and 40000 impostor verification attempts in the evaluation stage to determine the decision threshold at the EER operating point and 400 client and 112000 impostor verification attempts in the test stage to determine the final performance of the system and its robustness to the image degradations. Note that the training and evaluation stages are performed only with the original images from the XM2VTS database, since these two stages can always be supervised. The analysis of the performance and robustness of the system to the image degradations is conducted solely in the test stage, in which real operating conditions with a predefined decision threshold are simulated.

B. Assessing the localization procedure

Our first series of experiments focused on the assessment of the employed localization procedure. Since only images from the client and impostor test sets required a localization procedure (with the degraded databases), the total number of test images for this assessment equaled 960 (i.e., 400 facial images corresponding to the client test set and 560 facial images corresponding to the impostor test set). An facial image was assigned to the group of correctly localized images if the centers of both eyes differed from the location found with our face model by less than a quarter of the distance between the actual centers of the eyes. Or in other words, each of the the found eye-center positions (defined by the final location of the manually marked eye centers of our face model) has to lie within the outer circle of the reference eye. The same criterion was also used by Jesorsky et al. in [15]. Clearly, such an approach requires that all the reference eye positions in the tested images are marked by hand. This procedure, however, is susceptible to human error and, hence, the presented results are somewhat biased and only as accurate as the manually marked eye-positions. An example of a successfully localized face image and the employed localization criterion are shown in Fig.

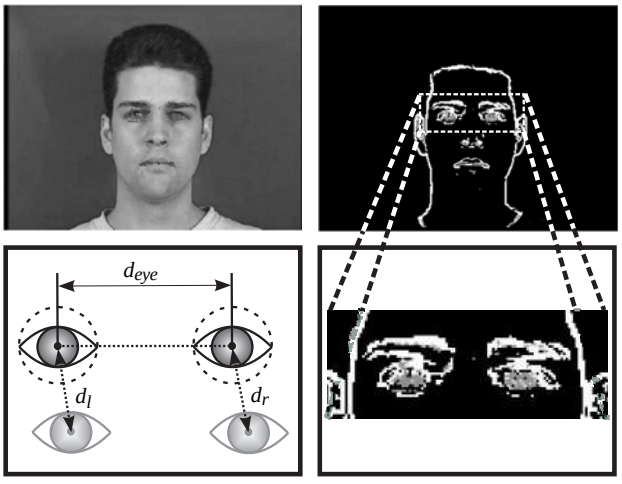

Fig. 8. An example of a successful face localization: the original image with the superimposed face model (upper left corner), the edge image with the superimposed face model (upper right corner), a visualization of the parameters used for determining the successfulness of the localization procedure (lower left corner), the eye-region of the face with the superimposed face model (lower right corner).

As already indicated above, the localization procedure was considered successful, if the following expression was less then 0.25:

$$\varepsilon_e = \frac{\max(d_l, d_r)}{d_{eye}}, \tag{11}$$

where d_{eye} denotes the distance between the reference eye-center locations and d_l and d_r represent the distances between the reference position and the located position of the left and right eye, respectively.

We again point out that the expression $\varepsilon_e < 0.25$ defines that the located eye-center locations are allowed to differ from the reference locations by less than half of an eye width.

With our localization procedure we correctly localized $97,8\%$ of all tested images from the original XM2VTS database. Jesorsky et al. [15] , for example, correctly localized $98,4\%$ of the facial images and Pozne reported in [16] to have achieved a localization rate of $99,9\%$. When our localization procedure was applied to the test images from the degraded XM2VTS database the localization rate dropped. The actual percentages of successfully localized images from the degraded databases are tabulated in Table V-B and graphically visualized in Fig. 9. Here, the label UD refers to the undegraded images. The results show that the employed

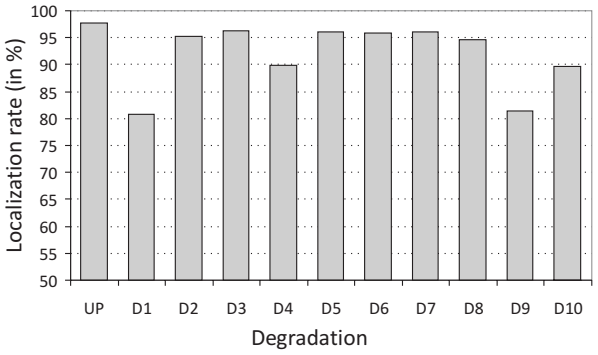

Fig. 9. The successfulness of the localization procedure on the original and degraded XM2VTS databases

localization procedure is rather insensible to the presence of noise, to image smoothing, JPEG compression and color-depth reduction. In all of these cases the localization rate dropped by at most 3.2% when compared to the original images. A larger drop in the localization rate was noticed for facial images with a complex background (D4) and partial occlusion (D10) of the face. Here, 89.9% and 89.6% of all tested images were localized correctly. The reason for the performance decrease with the degradation D4 can be found in the fact that many of the tested images contained background-objects whit "skin-color" and, therefore, interfered with the first step of our localization procedure. Partial occlusion of the face, on the other hand, influenced the final localization step, where a face model is fitted to the binary edge image of the face. The lowest localization rates were observed with images degraded using packet-loss simulation (D1) and geometric transformations (D2). Both degradations resulted in localization rates little above 80%.

C. Assessing the verification performance

In our second series of experiments we aimed at assessing the verification performance of the appearance based feature extraction techniques PCA, LDA, ICA1 and ICA2 and at testing their susceptibility to various image degradations. The techniques were applied to facial images localized with the procedure assessed in the previous section and optimized on the evaluation image set to yield the best possible performance. Specifically, the evaluation stage was used to determine the number of features to be employed with each of the techniques and to set the decision threshold Δ, as to ensure the EER operating point on the evaluation images.

We again point out that for the training and evaluation stages only images from the original (undegraded) XM2VTS database were used. The results of the experiments presented here[3] are, therefore, expected to deviate somehow from the EER operating point. Nevertheless, this represents the only way of assessing the appearance based feature extraction techniques in conditions close to real operational conditions.

As the localization procedure represents the front-end of our face verification system, the localization error is expected to propagate throughout the system and also influence the verification rates obtained on the test image sets. The results of the experiments are presented in Table V-C. Here, the HTER for the feature extraction technique which performed best for the given degradation techniques is printed in bold.

The presented results clearly show that amongst the tested feature extraction techniques the LDA approach performed the best on the majority of the tested databases. Only on the degraded databases where the loss of data packets was simulated and where parts of the faces have been occluded the ICA2 technique outperformed the LDA approach. If we look at the mean values and medians of the HTER for all tested techniques (shown in Fig. 10), we can see that the following ranking was obtained in our experiments: LDA performed the best, followed in order by the ICA2, ICA1 and PCA techniques.

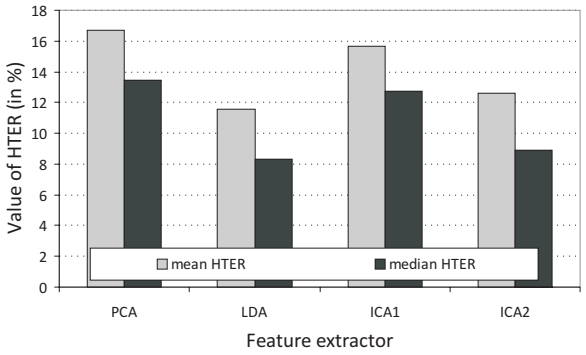

Fig. 10. The mean value and median of the HTER accross all tested databases

Generally, all tested techniques exhibit a similar susceptibility to image degradations. The largest deteriorations in the HTER when compared to the HTER obtained on the original database were found for the degradation techniques D1 (packet-loss simulation), D9 (geometrical transformations) and D10 (partial occlusion of the faces). While these degradations had the biggest impact on the HTER in has to be noted

[3]Note that the results presented in this section were obtained on the test image sets.

TABLE I
THE SUCCESSFULNESS OF THE LOCALIZATION PROCEDURE ON THE ORIGINAL AND DEGRADED XM2VTS DATABASES

Degradation	UD	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
Localization rate (in %)	97,8	80,7	95,3	96,2	89,9	96,1	95,8	96,0	94,6	81,4	89,6

TABLE II
THE VERIFICATION PERFORMANCE OF THE APPEARANCE BASED FEATURE EXTRACTION TECHNIQUES ON THE ORIGINAL AS WELL AS THE DEGRADED XM2VTS DATABASES

Degradation		UD	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
PCA	FRR (%)	10.50	40.75	11.50	10.75	17.00	11.75	19.00	13.25	15.50	33.00	79.75
	FAR (%)	11.28	7.12	10.52	10.62	11.90	10.87	7.91	11.52	11.83	7.12	4.64
	HTER (%)	10.89	23.94	11.01	10.69	14.45	11.31	13.46	12.39	13.67	20.06	42.20
LDA	FRR (%)	6.25	47.25	8.00	7.50	13.25	7.25	12.00	7.00	6.50	28.75	66.25
	FAR (%)	4.17	1.85	4.10	4.12	4.11	4.19	3.87	4.73	5.14	3.62	4.78
	HTER (%)	**5.21**	24.55	**6.05**	**5.81**	**8.68**	**5.72**	**7.94**	**5.87**	**5.82**	**16.19**	35.52
ICA1	FRR (%)	10.75	40.40	12.00	12.25	18.75	12.75	17.25	13.25	15.75	31.25	65.25
	FAR (%)	9.07	6.02	8.93	8.80	8.92	8.83	8.49	9.35	9.67	8.27	8.72
	HTER (%)	9.91	23.26	10.47	10.53	13.84	10.79	12.87	11.30	12.71	19.76	36.99
ICA2	FRR (%)	7.25	34.75	10.00	11.00	15.75	8.75	13.50	8.25	10.50	29.50	51.00
	FAR (%)	7.16	6.64	7.16	7.08	7.18	7.15	6.99	7.22	7.26	6.48	7.45
	HTER (%)	7.21	**20.70**	8.58	8.54	11.47	7.95	10.25	7.74	8.88	17.99	**29.23**

that exactly the same degradations also resulted in the lowest localization rates and, hence, represent the biggest challenges to face recognition systems.

The experimental results further suggest that all of techniques are rather insensitive to image degradations, such as the presence of noise (D2, D3), the presence of a complex background (D4), JPEG compression (D5), color-depth reduction (D6) and image smoothing (D7, D8), as there was only a small drop in the HTER observed. However, if we also consider the FAR and FRR and do not focus solely on the HTER we can see that there are noticeable deviations from the EER operating point. This finding can be directly linked to errors in the localization procedure, since miss-localization results in larger numbers of both client as well as impostor identity claims being rejected, thus, reducing the FAR and increasing the FRR.

VI. CONCLUSION

We have presented an empirical assessment of four popular appearance based feature extraction techniques within a face recognition system. The tested techniques (PCA, LDA, and ICA) were evaluated using the publicly available XM2VTS database and ten degraded version of the XM2VTS database created using various degradations techniques ranging from simple smoothing of the facial images, addition of noise or facial occlusion to the simulation of data packet loss as it occurs when images are transmitted over a computer network. Our experiments suggest that all of the assessed feature extraction techniques degrade similarly in their performance when applied to the degraded images. However, amongst the tested techniques, LDA was found to be the best - achieving the lowest error rates with the majority of tested degradations.

ACKNOWLEDGMENT

The research was partially supported by the national research program P2-0250(C) Metrology and Biometric Systems, the bilateral project with the People's Republic of China Bi-CN/07-09-019, the bilateral project with the Bulgarian Academy of Sciences - Face and Signature Biometrics, the national project AvID M2-0210, the COST Action 2101 Biometrics for Identity Documents and Smart Cards and the EU-FP7 project 217762 Homeland security, biometric Identification and personal Detection Ethics (HIDE).

REFERENCES

[1] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neurosicence*, vol. 3, no. 1, pp. 71-86, 1991.
[2] P.N. Belhumeur, J.P. Hespanha and D.J. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," in *Proc. of the 4th European Conference on Computer Vision, ECCV'96*, Cambridge, UK, 1996, pp. 45-58.
[3] M.S. Bartlett, J.R. Movellan ans T.J. Sejnowski, "Face recognition by independent component analysis," *IEEE Trans. on Neural Networks*, vol. 13, pp. 1450-1464, November 2002.
[4] C. Liu and H. Wechsler, "Comparative assessment of independent component analysis (ICA) for face recognition," in *Proc. of the 2nd International C onference on Audio- and Video-based Biometric Person Authentication, AVBPA'99*, Washington D.C., USA, 1999, pp. 211-216.
[5] K. Baek, B. Draper, J.R. Beveridge and K. She, "PCA vs. ICA: a comparison on the FERET data set," in *Proc. of the 4th International Conference on Computer Vision, ICCV'02*, Durham, 2002, pp. 824–827.
[6] K. Delac, M. Grgic and S. Grgic, "Independent comparative study of PCA, ICA, and LDA on the FERET data set," *International Journal of Imaging Systems and Technology*, vol. 15, no. 5, pp. 252-260, 2006.
[7] J.R. Beveridge, K. She, B. Draper and G.H. Givens, "A nonparametric statistical comparison of principal component and linear discriminant subspaces for face recognition," in *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR'01*, Kauai, 2001, pp. 535–542.
[8] Y. Li, J. Kittler and J. Matas, "Effective implementation of linear discriminant analysis for face recognition and verification," in *Proc. of the 8th International Conference on Computer Analysis of Images and Patterns, CAIP'99*, Ljubljana, 1999, pp. 234–242.
[9] V. Štruc and N. Pavešić, "Gabor-based kernel partial-least-squares discrimination features for face recognition," *International Journal Informatica*, to be published.
[10] V. Štruc and N. Pavešić, "The corrected normalized correlation coefficient : a novel way of matching score calculation for LDA-based face verification," in *Proc. of the 5th International Conference on Fuzzy Systems and Knowledge Discovery, FSKD'08*, Jinan, 2008, pp. 110–115.
[11] B. Batagelj and F. Solina, "Face recognition in different subspaces - a comparative study," in *Proc. of the 6th International Workshop on Pattern Recognition in Information Systems, PRIS'06*, Paphos, 2006, pp.71–80.

[12] The FastICA package for Matlab - accessed March 2009. Available: http://www.cis.hut.fi/projects/ica/fastica/
[13] A. Hyvarinen ans E. Oja, "Independent component analysis: algorithms and aplications," *Neural Networks* vol. 13, no. 4-5, pp. 411430, 2000.
[14] A.K. Jain, A. Ross and S. Prabhakar, "An introduction to biometric recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, pp. 4–20, January 2004.
[15] O. Jesorsky, K.J. Kirchberg and R.W. Frischholz, "Robust face detection using the Hausdorff distance," in: *Proc. of the 3rd Internatiojnal Conference on Audio- and Video-based Biometric Person Authentication, AVBPA'01*, Halmstad, 2001, pp.90–95.
[16] A. Pozne, "Extracting visual features for automated speech recognition," Ph.D. dissertation, Faculty of Electrical Engineering, University of Ljubljana, Ljubljana, 2005.
[17] K. Messer, J. Matas, J. Kittler, J. Luettin and G. Maitre, "XM2VTSDB: the extended M2VTS database," in *Proc. of AVBPA'99*, Washington D.C., 1999, pp. 72–77.
[18] N. Pavešić, I. Fratrić and S. Ribarić, "Degradation of the XM2VTS database face images," in *Proc. 2nd COST 275 Workshop*, Vigo, 2004, pp. 15–19.

Vitomir Štruc received his B.Sc. degree in electrical engineering from the University of Ljubljana in 2005. He is currently working as a researcher at the Laboratory of Artificial Perception, Systems and Cybernetics at the Faculty of Electrical Engineering of the University in Ljubljana. His research interests include pattern recognition, machine learning and biometrics.

Nikola Pavešić received his B.Sc. degree in electronics, M.Sc. degree in automatics, and Ph.D. degree in electrical engineering from the University of Ljubljana, Slovenia, in 1970, 1973 and 1976, respectively. Since 1970 he has been a staff member at the Faculty of Electrical Engineering in Ljubljana, where he is currently head of the Laboratory of Artificial Perception, Systems and Cybernetics. His research interests include pattern recognition, neural networks, image processing, speech processing, and information theory. He is the author and co-author of more than 200 papers and 3 books addressing several aspects of the above areas. Professor Nikola Pavešić is a member of IEEE, the Slovenian Association of Electrical Engineers and Technicians (Meritorious Member), the Slovenian Pattern Recognition Society, and the Slovenian Society for Medical and Biological Engineers.