Available online at www.sciencedirect.com
ScienceDirect
Pattern Recognition 40 (2007) 2798–2810
PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY
www.elsevier.com/locate/pr

The discriminant elastic graph matching algorithm applied to frontal face verification

Stefanos Zafeiriou*, Anastasios Tefas, Ioannis Pitas

Department of Informatics, Aristotle University of Thessaloniki, Box 451, 54124 Thessaloniki, Greece

Received 22 August 2006; received in revised form 24 January 2007; accepted 31 January 2007

Abstract

In this paper a generalized framework for face verification is proposed employing discriminant techniques in all phases of elastic graph matching. The proposed algorithm is called *discriminant elastic graph matching* (DEGM) algorithm. In the first step of the proposed method, DEGM, discriminant techniques at the node feature vectors are used for feature selection. In the sequel, the two local similarity values, i.e., the similarity measure for the projected node feature vector and the node deformation, are combined in a discriminant manner in order to form the new local similarity measure. Moreover, the new local similarity values at the nodes of the elastic graph are weighted by coefficients that are derived as well from discriminant analysis in order to form a total similarity measure between faces. The proposed method exploits the individuality of the human face and the discriminant information of elastic graph matching in order to improve the verification performance of elastic graph matching. We have applied the proposed scheme to a modified morphological elastic graph matching algorithm. All experiments have been conducted in the XM2VTS database resulting in very low error rates for the test sets.

Keywords: Elastic graph matching; Discriminant analysis; Frontal face verification

1. Introduction

A well known technique for face recognition and verification is the *elastic graph matching* (EGM) algorithm [1]. In EGM, a reference object graph is created by overlaying a rectangular elastic sparse graph on the object image and calculating a Gabor wavelet bank response at each graph node. The graph matching process is implemented by a stochastic optimization of a cost function which takes into account both jet similarities and grid deformations. A two stage coarse-to-fine optimization procedure suffices for the minimization of such a cost function.

Since its invention, EGM for face verification and recognition has been a very active research field [2–11]. In Ref. [2], it has been shown that EGM outperforms eigenfaces and autoassociation classification neural networks for face recognition. In Ref. [3] the graph structure has been enhanced by introducing a stack like structure, the so-called *bunch graph*, and has been tested for face recognition. In the bunch graph structure for every node a set of jets has been measured for different instances of a face (e.g., with mouth opened or closed, eyes opened or closed). That way, the bunch graph representation could cover a variety of possible changes in the appearance of a face. In Ref. [4], the bunch graph structure has been used for determining facial characteristics such as beard, presence of glasses or a person's sex.

Practical methods for increasing the robustness of EGM against translations, deformations and changes in background have been presented in Ref. [5]. In Ref. [6], EGM has been proposed and tested for frontal face verification, where the different choices for the elasticity of the graph have been investigated. A variant of the standard EGM, the so-called *morphological elastic graph matching* (MEGM), has been proposed for frontal face verification and tested for various recording conditions [7–9]. In MEGM, the Gabor features have been replaced by multiscale morphological features obtained through dilation–erosion of the facial image by a structuring function [12]. In Refs. [7,9] the standard coarse-to-fine approach [6] for

* Corresponding author. Tel.: +30 231 099 63 04; fax: +30 231 099 63 04.
E-mail addresses: dralbert@aiia.csd.auth.gr (S. Zafeiriou), tefas@aiia.csd.auth.gr (A. Tefas), pitas@aiia.csd.auth.gr (I. Pitas).

0031-3203/$30.00 © 2007 Published by Elsevier Ltd on behalf of Pattern Recognition Society.
doi:10.1016/j.patcog.2007.01.026

elastic matching has been replaced by a simulated annealing method that optimizes a cost function of the jet similarity distances subject to node deformation constraints. The multiscale morphological analysis has been proven to be suitable for facial image analysis and MEGM has given comparable verification results with the standard EGM approach, without having to compute the computationally expensive Gabor filter bank output. Another variant of EGM has been presented in Ref. [10], where morphological signal decomposition has been used instead of the standard Gabor analysis [6]. In Ref. [13] the use of EGM has been extended in order to treat the problem of hand posture recognition.

Discriminant techniques have been employed in order to enhance the recognition and verification performance of EGM. The use of linear discriminating techniques at the feature vectors for selecting the most discriminating features has been proposed in Refs. [6,7,9]. Several schemes that aim at weighting the graph nodes according to their discriminatory power have been proposed [7,9,11,14]. In Ref. [11] the selection of the weighting coefficients has been based on a nonlinear function that depends on a small set of parameters. These parameters have been determined on the training set by maximizing a criterion using the simplex method. In Refs. [7,9,10] the set of node weighting coefficient was not calculated by some criterion optimization but by using the first and second order statistics of the node similarity values. A Bayesian approach for determining which nodes are more reliable has been used in Ref. [4]. A more sophisticated scheme for weighting the nodes of the elastic graph by constructing a modified class of support vector machines [15] has been proposed in Ref. [14]. In Ref. [14], it has been also shown that the verification performance of the EGM can be highly improved by proper node weighting strategies.

In this paper we introduce a methodology for applying discriminant analysis techniques at all phases of EGM for face verification. This methodology (abbreviated as DEGM) can be applied to all EGM algorithms like Refs. [6,7]. More precisely, in the DEGM, each node is considered as a local expert and discriminant feature selection techniques are employed for enhancing its verification performance. The deformation of each node is considered as a second local similarity metric that can quantify the relationships between its neighboring nodes. The new local similarity value at each node is produced by discriminant weighting of both the feature vector similarity measure and the node deformation. Finally, a discriminant node weighting step is used in order to form the similarity measure between face graphs. In the proposed method the individuality of human faces is harnessed for improving the verification performance of the EGM algorithm. The motivations of this work are also supported by the fact that in many methods it has been shown that the verification performance is increased by exploiting the individuality of human features [16,17].

The novelty of the proposed approach is the exploitation of the discriminant information in every phase of EGM. Other approaches like Ref. [7] consider discriminant analysis only for feature selection in the node jets or only for weighting the local similarity measures. The combined use of discriminant analysis for the feature selection and afterwards for the node deformation along with the node weighting has not been proposed in the literature. Indeed, in Ref. [7], the node weighting strategies have been considered as alternative schemes to linear discriminant analysis of feature vectors and they have never been combined in a more powerful discriminant scheme. The DEGM algorithm employs a node weighting step not as an alternative but in order to enhance the discriminant power of the new local similarity measures and form a single distance between two face graphs. Moreover, a novel morphological analysis is proposed in order to robustify the morphological analysis against illumination changes. All the discriminant steps in DEGM have a physical meaning in the verification procedure which will be analyzed in detail in Section 2.

The outline of this paper is as follows. The problem is stated in Section 2. The steps of the DEGM are described in Section 3. The modified multiscale morphological analysis that has been used in our experiments is discussed in Section 4. Experimental results are depicted in Section 5. Finally, conclusions are drawn in Section 6.

2. EGM revisited

In the first step of the EGM algorithm a sparse graph suitable for face representation is selected [3,6,7]. The facial image region is analyzed and a set of local descriptors is extracted at each graph node. Analysis is usually performed by building an information pyramid using scale-space techniques. In the standard EGM, a two-dimensional Gabor based filter bank has been used for image analysis [1]. The output of multiscale morphological dilation–erosion operations or the morphological signal decomposition at several scales are nonlinear alternatives of the Gabor filters for multiscale analysis and both have been successfully used for facial image analysis [7,9,10,18]. At each graph node l that is located at image coordinates $\mathbf{x}^l$, a jet (feature vector) $\mathbf{j}(\mathbf{x})$ is formed:

$$\mathbf{j}(\mathbf{x}^l) = [f_1(\mathbf{x}^l), \ldots, f_M(\mathbf{x}^l)]^{\mathrm{T}}, \tag{1}$$

where $f_i(\mathbf{x}^l)$ denotes the output of a local operator applied to the image f at the ith scale or at the ith pair (scale, orientation) and M is the dimensionality of the jet. The next step of the EGM is to translate and deform the reference graph on the test image in order to find the correspondences of the reference graph nodes on the test image. This is accomplished by minimizing a cost function that employs node jet similarities and, in the same time preserves, the node neighborhood relationships. Let the superscripts t and r denote a test and a reference facial image (or graph), respectively. The L_2 norm between the feature vectors at the lth graph node of the reference and the test graph is used as a similarity measure between jets, i.e.,

$$C_f(\mathbf{j}(\mathbf{x}_t^l), \mathbf{j}(\mathbf{x}_r^l)) = \|\mathbf{j}(\mathbf{x}_r^l) - \mathbf{j}(\mathbf{x}_t^l)\|. \tag{2}$$

Let $\mathscr{V}$ be the set of all graph vertices[1] of a certain facial image. The graphs considered in this work are rectangular graphs

[1] The vertices of the graphs will be referred to as nodes throughout the paper.

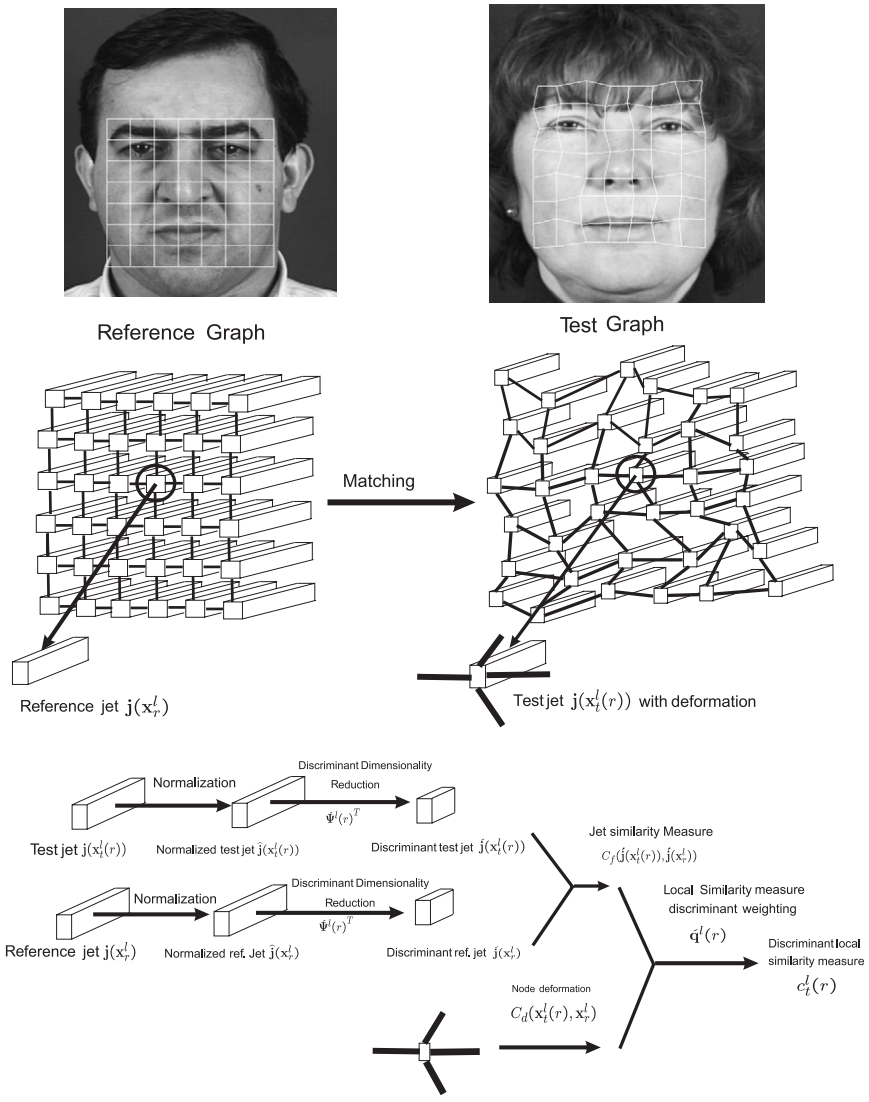

Fig. 1.

which are topologically equivalent to a rectangular subset of $\mathscr{Z}^2$ ($\mathscr{Z}$ is the set of integers). Thus, all nodes, except from the boundary nodes, have exactly four-connected nodes. Fig. 1 shows a typical reference rectangular graph used in this work. Let $\mathscr{H}(l)$ be the four-connected neighborhood of node l. In order to quantify the node neighborhood relationships using a metric, the local node deformation is used:

$$C_d(\mathbf{x}_t^l, \mathbf{x}_r^l) = \sum_{\xi \in \mathscr{H}(l)} \|(\mathbf{x}_t^l - \mathbf{x}_r^l) - (\mathbf{x}_t^\xi - \mathbf{x}_r^\xi)\|. \tag{3}$$

The objective is to find a set of vertices $\{\mathbf{x}_t^l(r), l \in \mathscr{V}\}$ in the test image that minimizes the cost function:

$$C(\{\mathbf{x}_t^l\}) = \sum_{l \in \mathscr{V}} \{C_f(\mathbf{j}(\mathbf{x}_t^l), \mathbf{j}(\mathbf{x}_r^l)) + \lambda C_d(\mathbf{x}_t^l, \mathbf{x}_r^l)\}. \tag{4}$$

The jet of the lth node that has been produced after the matching procedure of the graph of the reference person r in the image of the test person t is denoted as $\mathbf{j}(\mathbf{x}_t^l(r))$. This notation is used due to the fact that different reference graphs r result to different test jets $\mathbf{j}(\mathbf{x}_t^l(r))$. Thus, the jet of the lth node of the test graph t is a function of the reference graph r. The notation $\mathbf{j}(\mathbf{x}_r^l)$ is used only when the lth node is in a preselected position of a facial image.

The optimization of Eq. (4) has been interpreted as a simulated annealing with additional penalties imposed by the graph deformations, in Ref. [7]. Accordingly, Eq. (4) can be simplified to minimization of

$$D_t(r) = \sum_{l \in \mathscr{V}} \{C_f(\mathbf{j}(\mathbf{x}_t^l), \mathbf{j}(\mathbf{x}_r^l))\} \quad \text{subject to}$$
$$\mathbf{x}_t^l = \mathbf{x}_r^l + \mathbf{s} + \boldsymbol{\delta}_l, \quad \|\boldsymbol{\delta}_l\| \leqslant \delta_{\max}, \tag{5}$$

where $\mathbf{s}$ is a global translation of the graph and $\boldsymbol{\delta}_l$ denotes a local perturbation of the graph nodes. The choices of λ in Eq. (4) and of $\delta_{\max}$ in Eq. (5) control the rigidity/plasticity of the graph [6,7]. Obviously, both functions (4) and (5) define a similarity measure between two faces. After the matching

procedure the distance $D_t(r)$ is used as a quantitative measure for the similarity of two faces [6–8].

By examining carefully the EGM procedure from a pattern recognition perspective, the following questions arise: Do all the dimensions of the jet possess discriminant information? Does the node deformation posses any discriminant information? Are all the graph nodes equally significant for verifying the identity of a facial image?

In order to answer to all these questions, a general framework that enhances the verification performance of the EGM algorithm in a supervised manner is proposed. In more detail, discriminant techniques are used for selecting the most discriminant features of every facial image class. The jet similarity measure is combined in a discriminant manner with the node deformation in order to form a local discriminant similarity measure between nodes. The use of deformation in a discriminant manner can be explained intuitively as follows. The face graph has nodes that may correspond to landmarks (the landmarks correspond to facial points) whose deformation can be considered either as rigid or elastic for a particular face. For example, nodes corresponding to a person's face scars that are in some rigid region like forehead or nose cannot be easily moved, whereas some nodes corresponding to landmarks in lips can be moved much more freely.

If we had available the information about the elasticity/rigidity of each facial region a priori, we could have incorporated it in the grid matching procedure. However, this information is person specific and thus it should be retrieved using a training procedure and taken it into account when forming the local similarity measure between nodes. The last step of DEGM is to learn which nodes contain significant discriminant information and thus, to use proper weights when forming the similarity measure between faces. This step is motivated by the fact that certain facial features (e.g., beauty spots) are more discriminant than others.

3. Discriminant Elastic Graph Matching

Unlike the other proposed approaches [6,7,14] where discriminant analysis is used only in one step we propose a combined discriminant architecture. The steps of this architecture (DEGM) are the following:

- Use of discriminant techniques for feature selection at every node. This issue will be treated in Section 3.1.
- Combination of both the node jet similarity measure and the node deformation in a manner that exploits discriminant information to form the new local similarity measure at each node. This problem will be discussed in Section 3.2.
- Weighting the new similarity values at the nodes of the elastic graphs by coefficients that are also derived from discriminant analysis. This step is explained in detail in Section 3.3.

A brief conversation concerning the effect of the choice of discriminant functions in the various steps will be given in Section 5.2. In the following $\mathbf{m}(\mathscr{X})$ denotes the mean vector of a set of vectors $\mathscr{X}$ and $N(\mathscr{X})$ denotes the cardinality of a set $\mathscr{X}$.

3.1. Feature vector discriminant analysis

The first step of the DEGM is to learn a face and node specific discriminant function $\mathbf{g}_r^l$, for the lth node of the reference face r, that transforms the jets $\mathbf{j}(\mathbf{x}_t^l(r))$:

$$\acute{\mathbf{j}}(\mathbf{x}_t^l(r)) = \mathbf{g}_r^l(\mathbf{j}(\mathbf{x}_t^l(r))). \tag{6}$$

The transform $\mathbf{g}_r^l$ can be any linear or nonlinear discriminant feature transform, like the ones used for face recognition and verification [7,19–22]. We will use linear techniques in the remaining of the section. Alternatively, nonlinear techniques could also be used.

Before calculating the linear projections, we normalize all the jets that have been produced during the matching of the graphs of the reference person r to all other facial images in the training set in order to have zero mean and unit magnitude jets. Let $\widehat{\mathbf{j}}(\mathbf{x}_t^l(r))$ be the normalized jet at lth node. Let $\mathscr{F}_C^l(r)$ and $\mathscr{F}_I^l(r)$ be the sets of the normalized jets of the lth node that correspond to genuine and impostor claims related to person r, respectively. In the Fisher's Linear Discriminant Analysis (FLDA), the within-class and between-class scatter matrices are used to formulate criteria of class separability [23]. For a two class problem the within-class scatter for the vectors $\widehat{\mathbf{j}}(\mathbf{x}_t^l(r))$ is defined as

$$\begin{aligned}\mathbf{F}_W^l(r) = {} & \widehat{P}_C \frac{1}{N(\mathscr{F}_C^l(r))} \\ & \times \sum_{\widehat{\mathbf{j}}(\mathbf{x}_t^l(r)) \in \mathscr{F}_C^l(r)} (\widehat{\mathbf{j}}(\mathbf{x}_t^l(r)) - \mathbf{m}(\mathscr{F}_C^l(r)))(\widehat{\mathbf{j}}(\mathbf{x}_t^l(r)) \\ & - \mathbf{m}(\mathscr{F}_C^l(r)))^{\mathrm{T}} + \widehat{P}_I \frac{1}{N(\mathscr{F}_I^l(r))} \\ & \times \sum_{\widehat{\mathbf{j}}(\mathbf{x}_t^l(r)) \in \mathscr{F}_I^l(r)} (\widehat{\mathbf{j}}(\mathbf{x}_t^l(r)) - \mathbf{m}(\mathscr{F}_I^l(r)))(\widehat{\mathbf{j}}(\mathbf{x}_t^l(r)) \\ & - \mathbf{m}(\mathscr{F}_I^l(r)))^{\mathrm{T}}\end{aligned} \tag{7}$$

whereas the between-class scatter is

$$\begin{aligned}\mathbf{F}_B^l(r) = {} & \widehat{P}_C \widehat{P}_I (\mathbf{m}(\mathscr{F}_I^l(r)) - \mathbf{m}(\mathscr{F}_C^l(r)))(\mathbf{m}(\mathscr{F}_I^l(r)) \\ & - \mathbf{m}(\mathscr{F}_C^l(r)))^{\mathrm{T}},\end{aligned} \tag{8}$$

where $\widehat{P}_C$ and $\widehat{P}_I$ are the a priori probability estimates for the genuine and impostor class, respectively.

The most common criterion used for transforming linearly the feature vectors is the one that projects the feature vectors in the direction of $\boldsymbol{\psi}^l(r)$ so that the Fisher's discriminant ratio:

$$J(\boldsymbol{\psi}^l(r)) = \frac{\boldsymbol{\psi}^l(r)^{\mathrm{T}} \mathbf{F}_B^l(r) \boldsymbol{\psi}^l(r)}{\boldsymbol{\psi}^l(r)^{\mathrm{T}} \mathbf{F}_W^l(r) \boldsymbol{\psi}^l(r)} \tag{9}$$

is maximized [23]. The optimal projection $\boldsymbol{\psi}^l(r)$ is given by [23]

$$\boldsymbol{\psi}^l(r) \triangleq \frac{\mathbf{F}_W^l(r)^{-1}(\mathbf{m}(\mathscr{F}_I^l(r)) - \mathbf{m}(\mathscr{F}_C^l(r)))}{\|\mathbf{F}_W^l(r)^{-1}(\mathbf{m}(\mathscr{F}_I^l(r)) - \mathbf{m}(\mathscr{F}_C^l(r)))\|}. \tag{10}$$

It is assumed that $\mathbf{F}_W^l(r)$ is invertible, which is true in most implementations of EGM [6,7,9,10] where the feature vector has not more than 20 dimensions and most databases provide a relative large number of impostor claims. Eq. (10) indicates that, for the face verification problem, the original multidimensional feature space is projected to a one-dimensional feature space. The jet, $\widehat{\mathbf{j}}(\mathbf{x}_t^l(r))$, is projected to one dimension by

$$\acute{j}(\mathbf{x}_t^l(r)) = \boldsymbol{\psi}^l(r)^{\mathrm{T}}\widehat{\mathbf{j}}(\mathbf{x}_t^l(r)). \tag{11}$$

It is obvious that the one-dimensional feature space derived by Eq. (11) is only a very limited solution to the problem of discovering discriminant projections in a multidimensional feature space. Recently, it was shown [24] that alternative LDA schemes that give more than one discriminative dimensions, in a two class problem, have better classification performance. In the proposed method, we use the same criterion as Refs. [6,7] that can also give more than one discriminant directions. Let $\mathbf{W}^l(r)$ and $\mathbf{B}^l(r)$ be the matrices:

$$\mathbf{W}^l(r) = \sum_{\widehat{\mathbf{j}}(\mathbf{x}_t^l(r))\in\mathscr{F}_l^l(r)} (\widehat{\mathbf{j}}(\mathbf{x}_t^l(r)) - \mathbf{m}(\mathscr{F}_C^l(r)))(\widehat{\mathbf{j}}(\mathbf{x}_t^l(r)) - \mathbf{m}(\mathscr{F}_C^l(r)))^{\mathrm{T}} \tag{12}$$

and

$$\mathbf{B}^l(r) = \sum_{\widehat{\mathbf{j}}(\mathbf{x}_t^l(r))\in\mathscr{F}_C^l(r)} (\widehat{\mathbf{j}}(\mathbf{x}_t^l(r)) - \mathbf{m}(\mathscr{F}_C^l(r)))(\widehat{\mathbf{j}}(\mathbf{x}_t^l(r)) - \mathbf{m}(\mathscr{F}_C^l(r)))^{\mathrm{T}}, \tag{13}$$

the trace of the matrix $\mathbf{W}^l(r)$ denotes the dispersion of the impostor jets from the center of the genuine class while the trace of the matrix $\mathbf{B}^l(r)$ denotes the dispersion of the jets of the genuine class from the center of the genuine class. The optimal discriminant directions are the columns of the matrix $\acute{\boldsymbol{\Psi}}^l(r)$ which is given by the maximization of the criterion:

$$J(\boldsymbol{\Psi}^l(r)) = \frac{\mathrm{tr}[\boldsymbol{\Psi}^l(r)^{\mathrm{T}}\mathbf{W}^l(r)\boldsymbol{\Psi}^l(r)]}{\mathrm{tr}[\boldsymbol{\Psi}^l(r)^{\mathrm{T}}\mathbf{B}^l(r)\boldsymbol{\Psi}^l(r)]}, \tag{14}$$

where tr[$\mathbf{R}$] is the trace of the matrix $\mathbf{R}$. This criterion is well suited for the face verification problem due to the fact that it tries to find the feature projections that maximize the distance of the impostor jets from the genuine class center, while minimizing the distance of the genuine jets from the genuine class center. If $\mathbf{B}^l(r)$ is not singular, then (14) is maximized when the column vectors of the projection matrix, $\acute{\boldsymbol{\Psi}}^l(r)$, are the eigenvectors of $\mathbf{B}^l(r)^{-1}\mathbf{W}^l(r)$.

In order to proceed to feature dimensionality reduction in $P < M$ dimensions the matrix $\acute{\boldsymbol{\Psi}}^l(r)$ should be comprised by the eigenvectors of $\mathbf{B}^l(r)^{-1}\mathbf{W}^l(r)$ that correspond to the P greatest eigenvalues. It is obvious that $\mathbf{B}^l(r)$ is not always invertible since the training sets usually provide less samples than required. Numerous methods have been proposed in order to solve such optimization problems, like the maximization of Eq. (14), when the matrix in the denominator is singular [25–27]. The feature vector after discriminant dimensionality reduction is

$$\acute{\mathbf{j}}(\mathbf{x}_t^l(r)) = \mathbf{g}_r^l(\widehat{\mathbf{j}}(\mathbf{x}_t^l(r))) = \acute{\boldsymbol{\Psi}}^l(r)^{\mathrm{T}}\widehat{\mathbf{j}}(\mathbf{x}_t^l(r)). \tag{15}$$

The similarity measure of the new feature vectors can be given by a simple distance metric. We have used the L_2 norm for forming the new feature vector similarity measure in the final multidimensional space. Other choices for the distance metric are the L_1 norm, the normalized correlation or the Mahalanobis distance. Another alternative distance metric that has been recently introduced for LDA subspaces is the gradient direction metric along the most discriminant direction [28].

3.2. Local similarity measure discriminant weighting

The second step of the DEGM is to combine the feature vector similarity measure and the node deformation in a discriminant manner in order to form the new local similarity measure. In all approaches proposed so far [1,6,7], the node deformation was only employed implicitly in the EGM stage by imposing additional rigidity/plasticity penalties.

The node feature similarity measure between the reference person r and the test person t for the lth node is $C_f(\acute{\mathbf{j}}(\mathbf{x}_t^l(r)), \acute{\mathbf{j}}(\mathbf{x}_r^l))$ and the node deformation is $C_d(\mathbf{x}_t^l(r), \mathbf{x}_r^l)$. Let $\mathbf{d}_t^l(r) \in \Re^2$ be a column vector that is comprised by the two similarity measures for the node l between the test person t and the reference person r, i.e.,

$$\mathbf{d}_t^l(r) = [\, C_f(\acute{\mathbf{j}}(\mathbf{x}_t^l(r)), \acute{\mathbf{j}}(\mathbf{x}_r^l)) \quad C_d(\mathbf{x}_t^l(r), \mathbf{x}_r^l)\,]^{\mathrm{T}}. \tag{16}$$

According to the standard EGM [1,6] the node similarity value after the matching procedure can be given by

$$c_t^l(r) = C_f(\acute{\mathbf{j}}(\mathbf{x}_t^l(r)), \acute{\mathbf{j}}(\mathbf{x}_r^l)) + \lambda C_d(\mathbf{x}_t^l(r), \mathbf{x}_r^l) = [\,1 \quad \lambda\,]\mathbf{d}_t^l(r) = \mathbf{p}^{\mathrm{T}}\mathbf{d}_t^l(r), \tag{17}$$

where λ is the constant that controls the rigidity/plasticity of the graph [6]. In some cases λ is set equal to zero [6,7] when forming the local similarity measure. In this approach we propose to substitute the vector $\mathbf{p}$, which in general contains no discriminant information, with a discriminant transform. The two similarity measures $C_f(\acute{\mathbf{j}}(\mathbf{x}_t^l(r)), \acute{\mathbf{j}}(\mathbf{x}_r^l))$ and $C_d(\mathbf{x}_t^l(r), \mathbf{x}_r^l)$ could be considered as similarity scores that occurred from different sensors for the same modality. Thus, its values may range in different intervals and for robust fusion of these scores normalization techniques can be used as Ref. [29].

Let $\widehat{\mathbf{d}}_t^l(r)$ be the vector with the normalized scores. When performing discriminant local similarity measure weighting, we use a discriminant function μ_r^l that is a person and node specific combination of the measure of similarity between jets and the measure of local deformation. Its usefulness is to assign proper weights to each of the measures. For instance, for an extremely large interocular distance the deformation component should be emphasized while for a scar the jet similarity should receive a higher weight. The new local similarity measure is

$$c_t^l(r) = \mu_r^l(\widehat{\mathbf{d}}_t^l(r)). \tag{18}$$

Such transforms can be constructed by using linear or nonlinear methods for building discriminant function [30]. We have used LDA in order to find the discriminant transform μ_r^l.

Let $\mathscr{L}_C^l(r)$ and $\mathscr{L}_I^l(r)$ be the sets of normalized local similarity vectors $\widehat{\mathbf{d}}_t^l(r)$ that correspond to genuine and impostor claims, respectively. In order to form the optimization criterion, the between-class scatter matrix and the within-class scatter matrix of the normalized local similarity vectors $\widehat{\mathbf{d}}_t^l(r)$ are employed. Let $\mathbf{D}_W^l(r)$ and $\mathbf{D}_B^l(r)$ be the within scatter and the between scatter matrices, of the vectors $\widehat{\mathbf{d}}_t^l(r)$, respectively. The $\mathbf{D}_W^l(r)$ and $\mathbf{D}_B^l(r)$ can be calculated using Eqs. (7) and (8) for the vectors $\widehat{\mathbf{d}}_t^l(r)$, respectively.

When using the criterion (9) for the vectors $\widehat{\mathbf{d}}_t^l(r)$ the optimal weighting coefficients are given by Ref. [23]

$$\acute{\mathbf{q}}^l(r) \triangleq \frac{\mathbf{D}_W^l(r)^{-1}(\mathbf{m}(\mathscr{L}_I^l(r)) - \mathbf{m}(\mathscr{L}_C^l(r)))}{\|\mathbf{D}_W^l(r)^{-1}(\mathbf{m}(\mathscr{L}_I^l(r)) - \mathbf{m}(\mathscr{L}_C^l(r)))\|}. \tag{19}$$

The new similarity value between the lth node of the reference graph and the same node of the test graph is now:

$$c_t^l(r) = \mu_r^l(\widehat{\mathbf{d}}_t^l(r)) = \acute{\mathbf{q}}^l(r)^{\mathrm{T}}\widehat{\mathbf{d}}_t^l(r). \tag{20}$$

Fig. 1 depicts the sequence of steps that should be followed in order to form the new local similarity measure, defined in Eq. (20), for the lth node between the reference person r and the test person t.

3.3. Discriminant node weighting

The final step of the DEGM algorithm is to find a person specific discriminant function β_r of the new local similarity values and create the total similarity measure between a reference face r and a test face t. The idea here is to weight the similarity measures of nodes that correspond to different landmarks with weights that correspond to their discriminant power. The weights should be person specific due to the fact that different persons have different discriminant landmarks. Let $\mathbf{c}_t(r) \in \Re^L$ be a column vector comprised by the new local similarity values at every node:

$$\mathbf{c}_t(r) = [c_t^1(r)\ c_t^2(r)\ \cdots\ c_t^L(r)]^{\mathrm{T}}, \tag{21}$$

where L is the number of graph nodes. The vector $\mathbf{c}_t(r)$ is the total similarity vector between the reference face r and a test face t. The standard EGM algorithm [1,6] treats uniformly all the similarity values $c_t^l(r)$. That is, the total similarity measure between a reference person r and a test person t is simply the sum of all node similarity measures:

$$D_t(r) = \sum_{i=1}^{L} c_t^i(r) = \mathbf{1}^{\mathrm{T}}\mathbf{c}_t(r), \tag{22}$$

where $\mathbf{1}$ is an $L \times 1$ vector of ones. The proposed algorithm should learn a discriminant function β_r that is person specific and form the total similarity measure between faces:

$$\acute{D}_t(r) = \beta_r(\mathbf{c}_t(r)). \tag{23}$$

The transform β_r could be just a weighting vector or a more complicated nonlinear support vector machine. We will use LDA to create a total similarity measure between the reference person r and a test person t. A modified LDA algorithm that can cope with the small sample size problem and can be applied in this step is the one presented in Ref. [31]. Let $\mathscr{T}_C(r)$ and $\mathscr{T}_I(r)$ be the sets of the total similarity vectors for the genuine and impostor claims of the reference person r, respectively. Let $\mathbf{V}_W(r)$ and $\mathbf{V}_B^l(r)$ be the within scatter and the between scatter matrices, of the vectors $\mathbf{c}_t^l(r)$, respectively. The $\mathbf{V}_W(r)$ and $\mathbf{V}_B(r)$ can be calculated using Eqs. (7) and (8) for the vectors $\mathbf{c}_t^l(r)$, respectively. The optimal weighting coefficients, that are derived from the maximization of criterion (9), are the elements of the vector $\acute{\mathbf{w}}(r)$ [23]:

$$\acute{\mathbf{w}}(r) \triangleq \frac{\mathbf{V}_W(r)^{-1}(\mathbf{m}(\mathscr{T}_I(r)) - \mathbf{m}(\mathscr{T}_C(r)))}{\|\mathbf{V}_W(r)^{-1}(\mathbf{m}(\mathscr{T}_I(r)) - \mathbf{m}(\mathscr{T}_C(r)))\|}. \tag{24}$$

The similarity measure between the reference person r and the test person t, after all the successively discriminant steps, is given by

$$\acute{D}_t(r) = \beta_r(\mathbf{c}_t(r)) = \acute{\mathbf{w}}(r)^{\mathrm{T}}\mathbf{c}_t(r). \tag{25}$$

4. Normalized morphological facial analysis

In this section, we briefly outline the new multiscale morphological analysis that is proposed in this paper for feature extraction. The interested reader can refer to Refs. [7,9] for motivations and details concerning the MEGM.

As was already mentioned, the first step of all EGM algorithms is to build an information pyramid in the reference facial image. In the morphological EGM this information pyramid is built using multiscale morphological dilation–erosions [12]. Given an image $f(\mathbf{x}) : \mathscr{D} \subseteq \mathscr{Z}^2 \to \Re$ and a structuring function $g(\mathbf{x}) : \mathscr{G} \subseteq \mathscr{Z}^2 \to \Re$, the dilation of the image $f(\mathbf{x})$ by $g(\mathbf{x})$ is denoted by $(f \oplus g)(\mathbf{x})$. Its complementary operation, the erosion, is denoted by $(f \ominus g)(\mathbf{x})$ [7]. The multiscale dilation–erosion pyramid of the image $f(\mathbf{x})$ by $g_\sigma(\mathbf{x})$ is defined in Ref. [12], where σ denotes the scale parameter of the structuring function. In Ref. [7] it was shown that the choice of structuring function does not lead to statistically significant changes in the verification performance. However, it affects the computational complexity of feature calculation.

Such morphological operations can highlight and capture important information for key facial features such as eyebrows, eyes, nose tip, nostrils, lips, face contour, etc. but can be affected by different illumination conditions and noise [7]. To compensate for these conditions, the normalized multiscale dilation–erosion is proposed for facial image analysis. It is well known that the different illumination conditions affect the facial region in a nonuniform manner. However, it can safely be assumed that the illumination changes are locally uniform inside the area of the structuring element used for multiscale analysis. The proposed morphological features are calculated by subtracting the mean value of the intensity of the image f inside the area of structuring element from the corresponding

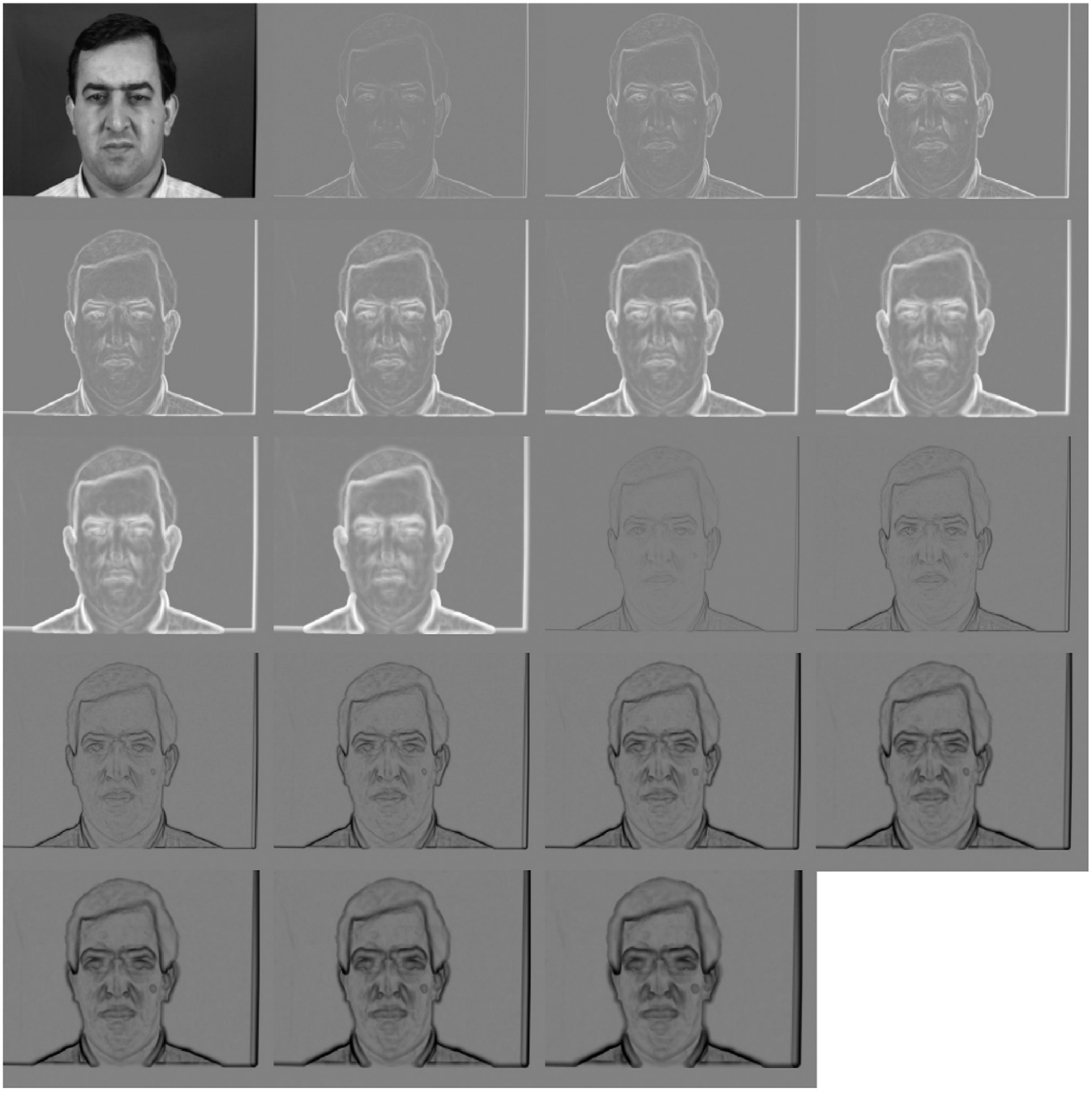

Fig. 2.

maximum (dilation) or minimum (erosion) of the area. Formally, the normalized multiscale morphological analysis is given by

$$(f \star g_\sigma)(\mathbf{x}) = \begin{cases} (f \oplus g_\sigma)(\mathbf{x}) - m_-(f, \mathbf{x}, G_\sigma) & \text{if } \sigma > 0, \\ f(\mathbf{x}) & \text{if } \sigma = 0, \\ (f \ominus g_{|\sigma|})(\mathbf{x}) - m_+(f, \mathbf{x}, G_\sigma) & \text{if } \sigma < 0, \end{cases} \quad (26)$$

where $m_-(f, \mathbf{x}, \mathscr{G}_\sigma)$ and $m_+(f, \mathbf{x}, \mathscr{G}_\sigma)$ are the mean values of the image $f(\mathbf{x} - \mathbf{z})$, $\mathbf{x} - \mathbf{z} \in \mathscr{D}$ and $f(\mathbf{x} + \mathbf{z})$, $\mathbf{x} + \mathbf{z} \in \mathscr{D}$ inside the support area of the structuring element $\mathscr{G}_\sigma = \{\mathbf{z} \in \mathscr{G} : \|\mathbf{z}\| < \sigma\}$, respectively. Another implementation for the operators $m_+(f, \mathbf{x}, \mathscr{G}_\sigma)$ and $m_-(f, \mathbf{x}, \mathscr{G}_\sigma)$ would be the median of the values of the image inside the support area of the structuring element. The output of these morphological operations forms the jet $\mathbf{j}(\mathbf{x}^l)$, at the graph node l that is located in image coordinates $\mathbf{x}^l$:

$$\mathbf{j}(\mathbf{x}^l) = ((f \star g_{\sigma_\Lambda})(\mathbf{x}^l), \ldots, (f \star g_{\sigma_1})(\mathbf{x}^l), f(\mathbf{x}^l), (f \star g_{\sigma_{-1}})(\mathbf{x}^l), \ldots, (f \star g_{\sigma_{-\Lambda}})(\mathbf{x}^l)). \quad (27)$$

where Λ is the number of different scales used. Fig. 2 depicts the output of normalized dilation–erosion for various scales used. The upper left image is the original image extracted from the XM2VTS database. The first nine images starting from the left corner, apart from the upper left one, are the normalized dilated images and the remaining nine are the normalized eroded images. The *Normalized Morphological Elastic Graph Matching* (NMEGM) will be used in experiments that will be presented in the experimental results section. When applying the DEGM using the normalized morphological multiscale analysis we will use the abbreviation DNMEGM.

5. Experimental results

5.1. Database and protocol

The experiments were conducted in the XM2VTS database using the protocol described in Ref. [32]. There is another popular database for face recognition tasks, the so-called FERET database, and is accompanied with an evaluation methodology [33–35]. Unfortunately, in the standard FERET evaluation protocol only one facial image is allowed for training. Thus, discriminant techniques can be difficultly used in the FERET evaluation protocol unless virtual samples are created in the training set (the use of synthetic samples arises many implementation issues). As can be seen in many recent published face recognition methods that are based on the exploitation of discriminant information, the researchers do not follow the FERET evaluation methodology and they choose a subset of FERET (in many cases the subset contains no more than 200 persons) to apply their own protocol [36–39]. Thus, their results cannot be compared in a straightforward manner. Moreover, the FERET evaluation methodology do not define strictly the size of the train and test sets and the set of people that should be considered as clients and impostors.

On the other hand XM2VTS offers a strict verification protocol with more than one sample per person. Thus, XM2VTS database is chosen in more cases instead of FERET database for testing face verification technologies. We have verified this fact by reviewing the recent bibliography of face verification algorithms and we have seen that the XM2VTS database is preferred more than the FERET database [40–46]. This is also verified by the fact that many competitions have been conducted in the XM2VTS database [47,48] over the past few years.

For the experiments a typical graph setup has been used [6,7]. More precisely, the graph was selected to be a 8×8 sparse graph and for the multiscale analysis Λ was set to 9. Thus, the jet dimension has been set to 19. The structuring element used in all experiments was cylindrical for computational reasons [7]. Only the luminance information at a resolution of 720×576 has been considered in our experiments. The appropriateness of the proposed normalized morphological features used in our experiments has been verified in Ref. [18] where it has been shown that they outperform the morphological features proposed in Ref. [7].

Unlike the most of subspace techniques that require a perfect manual alignment in order to perform well [18], DNMEGM has been combined with a fully automatic alignment method according to the eyes position of each facial image using the eye coordinates that have been derived from the method reported in Ref. [49]. No other image preprocessing technique has been used. In order to simplify the approach, graphs of the same size were considered for all persons. As an alternative, the face normalization technique reported in Ref. [8] could be used in order to find the width and the height of the face and create person specific graphs. The way the choice of the graph affects the verification procedure is out of the scope of this work.

The XM2VTS database contains 295 subjects, four recording sessions and two shots (repetitions) per recording session. The XM2VTS database provides two experimental setups namely, Configuration I and Configuration II [32]. Each configuration is divided in three different sets the training set, the evaluation set and the test set. The training set is used to create genuine and impostor models for each person. The evaluation set is used to learn the verification decision thresholds. In case of multimodal systems, the evaluation set is also used for training the fusion manager. In our case, due to the lack of sufficient data, we have used the evaluation sets in order to train the weights of the local similarity vector, the weighting coefficients of the total similarity vector and to find the person specific thresholds. This strategy has been followed in other cases as well, like Ref. [41], where the evaluation set has been used for fusing the scores of different face classifiers for frontal face verification.

For both configurations the training set has 200 clients, 25 evaluation impostors and 70 test impostors. The two configurations differ in the distribution of client training and client evaluation data. For additional details concerning XM2VTS database the interested reader can refer to Ref. [32]. Recently, frontal face verification competitions using the XM2VTS [47,48] have been conducted. The interested reader can refer to Refs. [47,48] and to the references within for the tested face verification algorithms.

The performance of face verification systems is measured in terms of the *false rejection rate* (FRR) achieved at a fixed *false acceptance rate* (FAR). There is a trade-off between FAR and FRR. That is, it is possible to reduce either of them with the risk of increasing the other one. This trade-off between the FAR and FRR can create a curve where FRR is plotted as a function of FAR. This curve is called *receiver operating characteristic* (ROC) curve [8,14]. The performance of a verification system is often quoted by a particular operating point of the ROC curve where FAR = FRR. This operating point is called *equal error rate* (EER). When a verification technique is to be evaluated for a real application then the thresholds should be set a priori. The evaluation set is used for setting the thresholds. The same thresholds will then be used on the test set. Let FAE and FRE be the corresponding FAR and FRR obtained on the evaluation set. Since application requirements might constrain the FAR or FRR to stay within certain limits, the system is evaluated for three different vectors of thresholds that correspond to the operating points where FAE = 0, FRE = 0 and FAE = FRE. For each given threshold, the *total error rate* (TER) can be obtained as the sum of FAR and FRR.

5.2. Experimental results in Configuration I

The training set of the Configuration I contains 200 persons with three images per person. The evaluation set contains three images per client for genuine claims and 25 evaluation impostors with eight images per impostor. Thus, evaluation set gives a total of $3\times200=600$ client claims and $25\times8\times200=40.000$ impostor claims. The test set has two images per client and 70 impostors with eight images per impostor and gives $2\times200=400$ genuine claims and $70\times8\times200=112.000$ impostor claims. The training set is used for calculating for each reference person r and for each node l a matrix for feature selection. In the training set three reference graphs per person are created. The $3\times2=6$ graphs that comprise the genuine class are created by applying EGM having one image as reference (i.e., in order to create the graph) and the other two images are used as test images. The impostor class contains $3\times3\times199=1797$ graphs. The jet dimension is 19 and thus, for a reference person r and a node l the matrix $\mathbf{B}^l(r)$, defined in Eq. (13), is singular. We have used *principal component analysis* [21] in order to satisfy the invertibility of the matrix $\mathbf{B}^l(r)$. The evaluation set is used for learning the discriminant vector for weighting the local similarity vector and the vector that weights the total similarity vector of the graph nodes. The invertibility of the within-class scatter matrix $\mathbf{D}_W^l(r)$ of the local similarity vector in Section 3.2 is satisfied due to the fact that the local similarity vectors contain only two dimensions. The invertibility of $\mathbf{V}_W(r)$ is also satisfied since we have used sparse graphs with 64 nodes.

In the proposed approach linear techniques have been used for training the different discriminant steps. The selection of linear techniques has been done due to the fact that the risk of overtraining in comparison with the nonlinear discriminant transforms is smaller and that they are less computational complex than the nonlinear. The problem of overtraining is of much greater intensity when nonlinear techniques are applied in training sets containing small and nonrepresentative data and, therefore in many cases very poor generalization is observed [30]. We have experienced great difficulties when applying nonlinear techniques in the various discriminant steps using publicly available databases used for testing face verification technologies like M2VTS [50] and XM2VTS [32] since they provide very few samples for the genuine class of each person (three or four images per person).

A possible solution in order to avoid overtraining is to insert virtual samples or noisy observations in the training set as proposed in Ref. [14]. However, in these cases many practical issues, like the number of the virtual samples to be included or the amount of noise to be inserted, cannot be easily assessed. Another implementation issue concerning the nonlinear methods is the choice of the appropriate nonlinear function (or kernel) that would represent the data, which in many cases are difficult to define [14,30].

For threshold calculation we have used the method proposed in Ref. [7]. That is, the similarity measures for every person calculated in the training set form the distance vector $\mathbf{o}(r)$. The elements of the vector $\mathbf{o}(r)$ are sorted in ascending order and are used for the person specific thresholds on the distance measure. Let $T_Q(r)$ denote the Qth order statistic of the vector of distances, $\mathbf{o}(r)$. The threshold of the person r is chosen to be equal to $T_Q(r)$. Let r_1, r_2 and r_3 be the three instances of the person r in the training set. A claim of a person t is considered valid if $\min_j\{\acute{D}_t(r_j)\}<T_Q(r)$ where $\acute{D}_t(r_j)$ is the distance between the graph of test person t and the reference graph r_j.

Table 1
The tested elastic graph matching approaches

Algorithm	Abbreviation
Normalized morphological elastic graph matching	NMEGM
NMEGM applying only node discriminant feature selection	NMEGM-FD
NMEGM applying only local discriminant weighting	NMEGM-LD
NMEGM applying only discriminant node weighting	NMEGM-ND
NMEGM applying all the proposed discriminant steps	DNMEGM

In order to illustrate the contribution of each discriminant step and also show the performance of the combined discriminant approach, we have conducted the following experiments:

- NMEGM without discriminant analysis.
- NMEGM applying only discriminant feature selection, as described in Section 3.1 (abbreviated as NMEGM-FD).
- NMEGM applying only local discriminant weighting using LDA, as described in Section 3.2, without using feature vector discriminant analysis or discriminant node weighting (abbreviated as NMEGM-LD).
- NMEGM applying only discriminant node weighting using LDA, as described in Section 3.3, without any other discriminant step (abbreviated as NMEGM-ND).

The abbreviations of the tested approaches are summarized in Table 1.

The NMEGM without any discriminant step has given a TER = 12.9% in the test set of Configuration I. For NMEGM-FD the best TER has been achieved by keeping the first three discriminant projections of the solution of the maximization of criterion (14) and has been estimated about 5.7%. The evaluation set has been used in order to estimate how many discriminant dimensions we should keep. We have, as well, tested feature vector discriminant analysis using the Fisher's criterion (9) and we have seen that the projection to the one-dimensional space using Eq. (11) has not lead to so significant improvement in the performance, giving an TER = 10%, as has been achieved by using criterion (14). The NMEGM-LD gave a TER = 9.2%. The TER that has been obtained with node weighting using LDA has been estimated about 10.7% (NMEGM-ND). The best TER achieved was 2.8% using successively all the discriminant steps described in this paper. As can be seen the feature vector discriminant analysis step is very important since it reduces the TER from 12.9% to 5.7% (that is a 50% reduction in terms of TER). But the other two steps are very significant, as well, and reduce the TER from 5.7% to 2.8% (another 50% reduction in terms of TER).

Table 2
Error rates according to XM2VTS protocol for Configuration I

Algorithm	Configuration I											
	Evaluation set			Test set								
	FAE = FRE	FAE(FRE = 0)	FRE(FAE = 0)	FAE = FRE		FRE = 0		FAE = 0		Total error rate (TER)		
				FA	FR	FA	FR	FA	FR	FAE = FRE	FRE = 0	FAE = 0
NMEGM	9.2	98.2	65.0	7.9	5.0	98.8	0.0	0.0	61.0	12.9	98.8	61.0
NMEGM-ND	6.3	62.8	56.3	6.7	4.2	63.8	0.0	0.0	61.0	10.7	63.8	61.0
NMEGM-LD	5.2	45.5	20.0	5.2	4.0	45.0	0.5	0.0	17.0	9.2	45.5	17.0
NMEGM-FD	2.5	29.9	55.3	2.5	3.2	11.2	0.2	0.2	14.7	5.7	11.4	14.9
DNMEGM	0.2	0.7	6.5	1.6	1.2	10.2	0.0	0.0	13.1	2.8	10.2	13.1

Table 3
A comparison of TER for Configuration I using fully automatic registration

Algorithm	TER
IDIAP-Cardinaux [47]	4.7
UPV [47]	3.98
UNIS-NC [47]	3.86
DNMEGM	2.8

The error rates according to the XM2VTS protocol are illustrated in Table 2. Table 3 shows a comparison of DNMEGM with other methods that use fully automatic alignment. The results have been acquired by the most recent competition in XM2VTS database [48]. Obviously, our method outperforms all the approaches tested in Ref. [48] using fully automatic alignment.

The ROC curves are depicted in Fig. 3 where the EER and the operating point using the thresholds of the operating point that corresponds to EER in the evaluation set are also shown (abbreviated as EER-E in Fig. 3). The operating points that corresponds to the thresholds of the EER measured in the evaluation set are shown in columns 5 and 6 of Table 2.

5.3. Experimental results in Configuration II

The Configuration II differs from the Configuration I in the distribution of client training and client evaluation data. The training set of the Configuration II contains 200 persons with four images per person. The evaluation set contains two images per client for genuine claims. Thus, evaluation set gives a total of $2 \times 200 = 400$ genuine claims. The training set contains four references images for each client. The same approach as

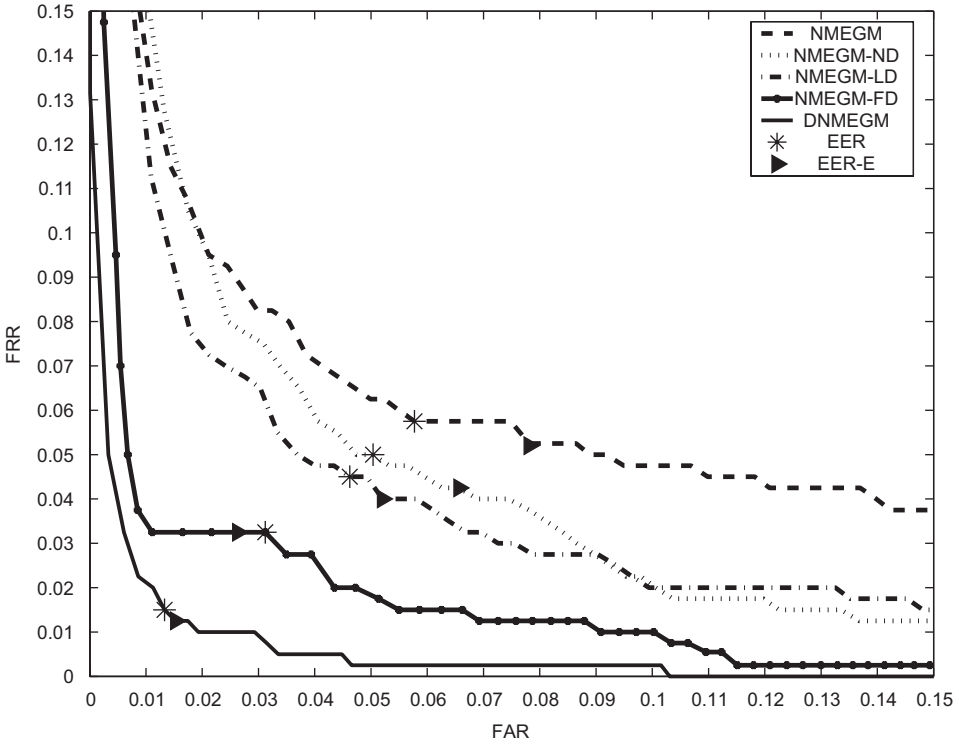

Fig. 3.

Table 4
Error rates according to XM2VTS protocol for Configuration II

Algorithm	Configuration II											
	Evaluation set			Test set								
	FAE = FRE	FAE(FRE = 0)	FRE(FAE = 0)	FAE = FRE		FRE = 0		FAE = 0		Total error rate (TER)		
				FA	FR	FA	FR	FA	FR	FAE = FRE	FRE = 0	FAE = 0
NMEGM	6.0	75.0	65.5	5.2	4.0	76.9	0.0	0.0	43.2	9.2	76.9	43.2
NMEGM-ND	5.2	46.5	69.5	4.9	4.2	71.2	0.0	0.1	38.1	9.1	71.2	38.2
NMEGM-LD	4.9	44.5	68.3	4.2	2.7	57.2	0.0	0.1	34.2	6.9	57.2	34.3
NMEGM-FD	2.4	28.5	51.0	2.8	2.7	26.9	0.0	0.1	18.7	5.5	26.9	18.8
DNMEGM	0.1	0.3	5.5	1.0	0.7	9.7	0.0	0.0	12.2	1.7	9.7	12.2

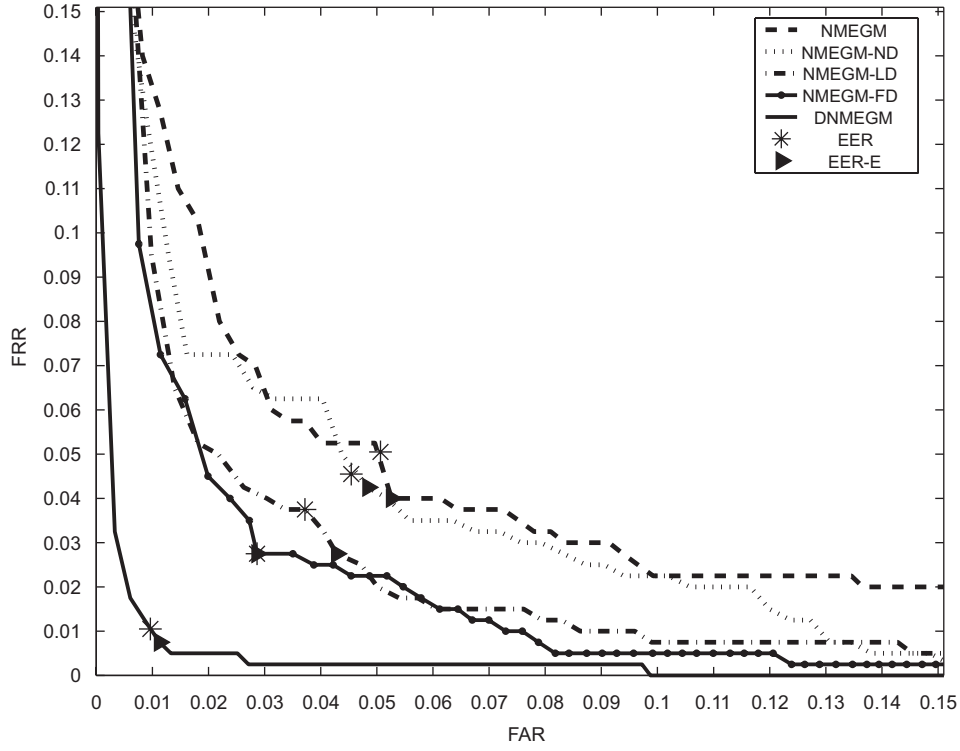

Fig. 4.

Table 5
A comparison of TER for Configuration II using fully automatic registration

Algorithm	TER
IDIAP-Cardinaux [47]	2.1
UPV [47]	2.3
UNIS-NC [47]	3.3
DNMEGM	1.7

in Configuration I has been used for training, for accepting a claim as valid and for threshold calculation. In Table 4 the error rates according to the XM2VTS Configuration II protocol are illustrated for the tested approaches. As can be seen the DNMEGM achieves a very low TER = 1.7% in this configuration. The corresponding ROC curves are depicted in Fig. 4. Table 5 shows a comparison of DNMEGM with other methods that use fully automatic alignment in Configuration II. Obviously, our method outperforms as well all the approaches tested in Ref. [48] using fully automatic alignment in Configuration II.

6. Conclusion

A general method for enhancing the performance of the EGM algorithm by employing discriminant analysis techniques in all phases of EGM has been proposed. The first step of the proposed method is to use discriminant techniques at the node feature vectors for feature selection. Afterwards, the two local similarity values, i.e., the similarity measure for the projected node feature vector and the node deformation, are combined in

a discriminant manner in order to form the new local similarity measure. Moreover, the new local similarity values at the nodes of the elastic graph are weighted by coefficients that are also derived from discriminant analysis in order to form a total similarity measure between objects. The methodology has been successfully applied to a modified MEGM algorithm in the frontal face verification problem. Further research on the topic includes the exploration of other pattern recognition algorithms, like support vector machines or relevance vector machines, in the various discriminant steps of the proposed algorithm in order to boost further the verification performance of DEGM.

Acknowledgments

This work has been partially funded by the integrated project BioSec IST-2002-001766 (Biometric Security, http://www.hrefhttp://www.biosec.orgbiosec.org) and by the network of excellence BioSecure IST-2002-507634 (Biometrics for Secure Authentication, http://www.biosecure.info), both under Information Society Technologies (IST) priority of the Sixth Framework Programme of the European Community.

References

[1] M. Lades, J.C. Vorbrüggen, J. Buhmann, J. Lange, C.v.d. Malsburg, R.P. Würtz, W. Konen, Distortion invariant object recognition in the dynamic link architecture, IEEE Trans. Comput. 42 (3) (1993) 300–311.

[2] J. Zhang, Y. Yan, M. Lades, Face recognition: eigenface, elastic matching, and neural nets, Proc. IEEE 85 (9) (1997) 1423–1435.

[3] L. Wiskott, J. Fellous, N. Krüger, C.v.d. Malsburg, Face recognition by elastic bunch graph matching, IEEE Trans. Pattern Anal. Mach. Intell. 19 (7) (1997) 775–779.

[4] L. Wiskott, Phantom faces for face analysis, Pattern Recognition 30 (6) (1997) 837–846.

[5] R.P. Würtz, Object recognition robust under translations, deformations, and changes in background, IEEE Trans. Pattern Anal. Mach. Intell. 19 (7) (1997) 769–775.

[6] B. Duc, S. Fischer, J. Bigün, Face authentication with Gabor information on deformable graphs, IEEE Trans. Image Process. 8 (4) (1999) 504–516.

[7] C. Kotropoulos, A. Tefas, I. Pitas, Frontal face authentication using discriminating grids with morphological feature vectors, IEEE Trans. Multimedia 2 (1) (2000) 14–26.

[8] C. Kotropoulos, A. Tefas, I. Pitas, Morphological elastic graph matching applied to frontal face authentication under well-controlled and real conditions, Pattern Recognition 33 (12) (2000) 31–43.

[9] C. Kotropoulos, A. Tefas, I. Pitas, Frontal face authentication using morphological elastic graph matching, IEEE Trans. Image Process. 9 (4) (2000) 555–560.

[10] A. Tefas, C. Kotropoulos, I. Pitas, Face verification using elastic graph matching based on morphological signal decomposition, Signal Processing 82 (6) (2002) 833–851.

[11] N. Krüger, An algorithm for the learning of weights in discrimination functions using a priori constraints, IEEE Trans. Pattern Anal. Mach. Intell. 19 (7) (1997) 764–768.

[12] P.T. Jackway, M. Deriche, Scale-space properties of the multiscale morphological dilation–erosion, IEEE Trans. Pattern Anal. Mach. Intell. 18 (1) (1996) 38–51.

[13] J. Triesch, C.v.d. Malsburg, A system for person-independent hand posture recognition against complex backgrounds, IEEE Trans. Pattern Anal. Mach. Intell. 23 (12) (2001) 1449–1453.

[14] A. Tefas, C. Kotropoulos, I. Pitas, Using support vector machines to enhance the performance of elastic graph matching for frontal face authentication, IEEE Trans. Pattern Anal. Mach. Intell. 23 (7) (2001) 735–746.

[15] V. Vapnik, The Nature of Statistical Learning Theory, Springer, New York, 1995.

[16] K.-A. Toh, J. Xudong, W.-Y. Yau, Exploiting global and local decisions for multimodal biometrics verification, IEEE Trans. Signal Process. 52 (10) (2004) 3059–3072.

[17] J. Kittler, Y.P. Li, J. Matas, Face verification using client specific Fisherfaces, in: J.T. Kent, R.G. Aykroyd (Eds.), The Statistics of Directions, Shapes and Images, 2000, pp. 63–66.

[18] S. Zafeiriou, A. Tefas, I. Pitas, Elastic graph matching versus linear subspace methods for frontal face verification, in: International Workshop on Nonlinear Signal and Image Processing, Sapporo, Japan, 2005.

[19] L. Chengjun, Gabor-based kernel PCA with fractional power polynomial models for face recognition, IEEE Trans. Pattern Anal. Mach. Intell. 26 (5) (2004) 572–581.

[20] L. Juwei, K. Plataniotis, A. Venetsanopoulos, Face recognition using kernel direct discriminant analysis algorithms, IEEE Trans. Neural Networks 14 (1) (2003) 117–126.

[21] M. Turk, A.P. Pentland, Eigenfaces for recognition, J. Cognitive Neurosc. 3 (1) (1991) 71–86.

[22] B. Scholkopf, A. Smola, Learning with Kernels, MIT Press, Cambridge, MA, 2002.

[23] K. Fukunaga, Statistical Pattern Recognition, Academic, San Diego, CA, 1990.

[24] C. Songcan, Y. Xubing, Alternative linear discriminant classifier, Pattern Recognition 37 (7) (2004) 1545–1547.

[25] P.N. Belhumeur, J.P. Hespanha, D.J. Kriegman, Eigenfaces vs. Fisherfaces: recognition using class specific linear projection, IEEE Trans. Pattern Anal. Mach. Intell. 19 (7) (1997) 711–720.

[26] H. Yu, J. Yang, A direct LDA algorithm for high-dimensional data with application to face recognition, Pattern Recognition 34 (2001) 2067–2070.

[27] D.L. Swets, J. Weng, Using discriminant eigenfeatures for image retrieval, IEEE Trans. Pattern Anal. Mach. Intell. 18 (8) (1996) 831–836.

[28] Y. Li, J. Kittler, J. Matas, On matching scores of LDA-based face verification, in: Proceedings of the British Machine Vision Conference BMVC2000, 2000.

[29] R. Snelick, U. Uludag, A. Mink, M. Indovina, A. Jain, Large-scale evaluation of multimodal biometric authentication using state-of-the-art systems, IEEE Trans. Pattern Anal. Mach. Intell. 27 (3) (2005) 450–455.

[30] K.-R. Muller, S. Mika, G. Ratsch, K. Tsuda, B. Scholkopf, An introduction to kernel-based learning algorithms, IEEE Trans. Neural Networks 12 (2) (2001) 181–201.

[31] M. Kyperountas, A. Tefas, I. Pitas, Weighted piecewise lda for solving the small sample size problem in face verification, IEEE Trans. Neural Networks, accepted for publication.

[32] K. Messer, J. Matas, J. Kittler, J. Luettin, G. Maitre, XM2VTSDB: the extended M2VTS database, in: Proceedings of the Second International Conference on Audio- and Video-based Biometric Person Authentication (AVBPA'99), 1999, pp. 72–77.

[33] S. Rizvi, P. Phillips, H. Moon, The FERET verification testing protocol for face recognition algorithms, in: Third IEEE International Conference on Automatic Face and Gesture Recognition, Nara, Japan, 1998, pp. 48–53.

[34] P. Phillips, H. Wechsler, J. Huang, P. Rauss, The FERET database and evaluation procedure for face recognition algorithms, Image Vision Comput. 16 (5) (1998) 295–306.

[35] P.J. Phillips, H. Moon, P.J. Rauss, S. Rizvi, The FERET evaluation methodology for face recognition algorithms, IEEE Trans. Pattern Anal. Mach. Intell. 22 (10) (2000) 1090–1104.

[36] L. Chengjun, H. Wechsler, A shape- and texture-based enhanced Fisher classifier for face recognition, IEEE Trans. Image Process. 10 (4) (2001) 598–608.

[37] L. Chengjun, H. Wechsler, Gabor feature based classification using the enhanced Fisher linear discriminant model for face recognition, IEEE Trans. Image Process. 11 (4) (2002) 467–476.

[38] L. Chengjun, H. Wechsler, Independent component analysis of gabor features for face recognition, IEEE Trans. Neural Networks 14 (4) (2003) 919–928.

[39] J. Yang, A. Frangi, J. Yang, D. Zhang, Z. Jin, KPCA plus LDA: a complete kernel fisher discriminant framework for feature extraction and recognition, IEEE Trans. Pattern Anal. Mach. Intell. 27 (2) (2005) 230–244.

[40] K. Jonsson, J. Kittler, Y.P. Li, J. Matas, Support vector machines for face authentication, Image Vision Comput. 20 (5–6) (2002) 369–375.

[41] J. Czyz, J. Kittler, L. Vandendorpe, Multiple classifier combination for face-based identity verification, Pattern Recognition 37 (7) (2004) 1459–1469.

[42] J. Kittler, R. Ghaderi, T. Windeatt, J. Matas, Face verification via error correcting output codes, Image Vision Comput. 21 (13–14) (2003) 1163–1169.

[43] S. Bengio, C. Marcel, S. Marcel, J. Mariethoz, Confidence measures for multimodal identity verification, Inf. Fusion 2 (2002) 267–276.

[44] N. Poh, S. Bengio, Database, protocols and tools for evaluating score-level fusion algorithms in biometric authentication, Pattern Recognition 39 (2006) 223–233.

[45] S. Srisuk, M. Petrou, W. Kurutach, A. Kadyrov, A face authentication system using the trace transform, Pattern Anal. Appl. 8 (2005) 50–61.

[46] S. Zafeiriou, A. Tefas, I. Buciu, I. Pitas, Exploiting discriminant information in nonnegative matrix factorization with application to frontal face verification, IEEE Trans. Neural Networks 17 (3) (2006) 683–695.

[47] J. Matas, M. Hamou, K. Jonsson, J. Kittler, Y. Li, C. Kotropoulos, A. Tefas, I. Pitas, T. Tan, H. Yan, F. Smeraldi, J. Bigun, N. Capdevielle, W. Gerstner, S. Ben-Yacouba, Y. Abdelaoued, E. Mayoraz, Comparison of face verification results on the XM2VTS database, in: Proceedings of 2000 International Conference on Pattern Recognition (ICPR'00), 2000, pp. 858–863.

[48] K. Messer, J. Kittler, M. Sadeghi, S. Marcel, C. Marcel, S. Bengio, F. Cardinaux, C. Sanderson, J. Czyz, L. Vandendorpe, S. Srisuk, M. Petrou, W. Kurutach, A. Kadyrov, R. Paredes, B. Kepenekci, F. Tek, G. Akar, F. Deravi, N. Mavity, Face verification competition on the XM2VTS database, in: AVBPA03, 2003, pp. 964–974.

[49] K. Jonsson, J. Matas, J. Kittler, Learning salient features for real-time face verification, in: Proceedings of the Second International Conference on Audio- and Video-based Biometric Person Authentication (AVBPA'99), 1999, pp. 60–65.

[50] S. Pigeon, L. Vandendorpe, The M2VTS multimodal face database, in: J. Bigun, G. Chollet, G. Borgefors (Eds.), Audio and Video-based Biometric Person Authentication, Lecture Notes in Computer Science, vol. 1206, 1997, pp. 403–409.

About the Author—STEFANOS ZAFEIRIOU was born in Thessaloniki, Greece, in 1981. He received the B.Sc. degree in Informatics with highest honors in 2003 and the Ph.D. degree in Informatics in 2007, both from the Aristotle University of Thessaloniki, Thessaloniki, Greece. He has received various scholarships and awards during his undergraduate and Ph.D. studies. He has co-authored over than 20 journal and conference publications. He is currently a researcher and teaching assistant at the Department of Informatics at the University of Thessaloniki. His current research interests lie in the areas of signal and image processing, computational intelligence, pattern recognition and computer vision as well as in the area of watermarking for copyright protection and authentication of digital media.

About the Author—ANASTASIOS TEFAS received the B.Sc. in Informatics in 1997 and the Ph.D. degree in Informatics in 2002, both from the Aristotle University of Thessaloniki, Greece. Since 2006, he has been an assistant professor at the Department of Information Management, Technological Educational Institute of Kavala. From 1997 to 2002, he was a researcher and teaching assistant in the Department of Informatics, University of Thessaloniki. From 2003 to 2004, he was a temporary lecturer in the Department of Informatics, University of Thessaloniki where he is currently, a senior researcher. He has co-authored over 50 journal and conference papers. His current research interests include computational intelligence, pattern recognition, digital signal and image processing, detection and estimation theory, and computer vision.

About the Author—IOANNIS PITAS received the Diploma of Electrical Engineering in 1980 and the Ph.D. degree in Electrical Engineering in 1985 both from the Aristotle University of Thessaloniki, Greece. Since 1994, he has been a professor at the Department of Informatics, Aristotle University of Thessaloniki. From 1980 to 1993 he served as scientific assistant, lecturer, assistant professor, and associate professor in the Department of Electrical and Computer Engineering at the same University. He served as a visiting research associate at the University of Toronto, Canada, University of Erlangen-Nuernberg, Germany, Tampere University of Technology, Finland, as visiting assistant professor at the University of Toronto and as visiting professor at the University of British Columbia, Vancouver, Canada. He was lecturer in short courses for continuing education. He has published over 145 journal papers, 380 conference papers and contributed in 18 books in his areas of interest. He is the co-author of the books Nonlinear Digital Filters: Principles and Applications (Kluwer, 1990), 3-D Image Processing Algorithms (Wiley, 2000), Nonlinear Model-Based Image/Video Processing and Analysis (Wiley, 2001) and author of Digital Image Processing Algorithms and Applications (Wiley, 2000). He is the editor of the book Parallel Algorithms and Architectures for Digital Image Processing, Computer Vision and Neural Networks (Wiley, 1993). He has also been an invited speaker and/or member of the program committee of several scientific conferences and workshops. In the past he served as associate editor of the IEEE Transactions on Circuits and Systems, IEEE Transactions on Neural Networks, IEEE Transactions on Image Processing, EURASIP Journal on Applied Signal Processing and co-editor of Multidimensional Systems and Signal Processing. He was general chair of the 1995 IEEE Workshop on Nonlinear Signal and Image Processing (NSIP95), technical chair of the 1998 European Signal Processing Conference and general chair of IEEE ICIP 2001. His current interests are in the areas of digital image and video processing and analysis, multidimensional signal processing, watermarking and computer vision.