Face Verification using Elastic Graph Matching based on Morphological Signal Decomposition

Anastasios Tefas, Constantine Kotropoulos and Ioannis Pitas

Department of Informatics
Aristotle University of Thessaloniki
Box 451, Thessaloniki 540 06, GREECE
{tefas,costas,pitas}@zeus.csd.auth.gr

Abstract

In this paper a novel method for frontal face verification is proposed. It is based on the *morphological signal decomposition*, a procedure that is used to model a facial image region as a sum of components. During the procedure, partial sums of components create a sequence of reconstructed images which starts with a crude approximation of the facial region that recursively becomes finer. More specifically, a feature vector is created at each node of a sparse grid superimposed on the facial area by concatenating the gray level values of the reconstructed images at the grid node position. When a candidate person claims the identity of a reference person, a variant of dynamic link matching, the so called *Morphological Signal Decomposition - Dynamic Link Architecture*, is applied to yield a matching error between the reference grid and a variable grid that is built over the facial image region of the candidate person. Local coefficients are derived to weigh the contribution of each node to the total matching error according to the node discriminatory power. Moreover, an analysis of the discriminatory power of each level in morphological signal decomposition is undertaken to assess better the behavior of the proposed method. Experimental results are reported on the M2VTS facial image database yielding a very low equal error rate.

Keywords : Face verification; Elastic graph matching; Dynamic link architecture; Morphological signal decomposition; Discriminatory power coefficients.

I. Introduction

Many techniques for face recognition have been developed in the last two decades whose principles span several disciplines, such as image processing, pattern recognition, computer vision and neural networks. The increasing interest in face recognition is attributed to the needs of numerous commercial and law enforcement applications requiring the automated person verification and recognition. Although humans recognize faces relatively easily, the robust machine recognition of faces remains an unsolved problem yet.

Machine perception of faces is applied to the problems of *face recognition* where the objective is to find the most similar face that corresponds to a reference face, from a given database, and *face verification* whose objective is to either accept or reject the identity claim of a test person. In the first case the output of the system is an identity, whereas in the second case the output of the system is a decision about the claim. In this paper, we deal with face verification. For a verification system, there is a trade-off between the *false acceptance rate* (FAR) and the *false rejection rate* (FRR). The choice of the factor, i.e., FAR or FRR, that should be low depends on the nature of the application.

A detailed survey of face recognition algorithms and their applications in real systems can be found in [1]. Two main categories for face recognition techniques can be identified: those employing geometrical features (for example [2]) and those using gray-level information (e.g.,

the eigenface approach [3]). A different approach that uses both gray-level information and shape information has been proposed in [4]. More specifically, the response of a set of 2D Gabor filters tuned to different orientations and scales is measured at the nodes of a sparse grid superimposed on the facial image of a person from a reference set. The responses of Gabor filters form a *feature vector* at each grid node. In the recall phase, the reference grid of each person is overlaid on the facial image of a test person and is deformed so that a criterion based on both the feature vectors and the grid distortion (i.e., the geometry) is minimized. This pattern matching algorithm is called *Dynamic Link Architecture* (DLA). An implementation of DLA based on Gabor wavelets is described in [5].

A comparative study of three algorithms for face recognition, namely, the eigenfaces [3,6], the auto-association and classification neural networks [7], and the elastic graph matching [4] can be found in [8]. The outcome of this study reveals that the elastic graph matching achieves a better performance than the other methods, because it is more robust to lighting, face-position and expression variations. The eigenfaces and the neural network algorithms require the images to be of the same scale and viewing angle. Moreover, the aforementioned methods are very sensitive to lighting variations. The problem of compensating for changes in illumination conditions is crucial for face recognition algorithms [9,10]. The interested reader may refer to [11] for the treatment of varying recording conditions.

DLA and its variants has been an active research topic since its invention. A different topology cost for a particular pair of nodes has been proposed in [12]. It is based on the radius of the Apollonius sphere defined by the Euclidean distances between the nodes being matched. Three major extensions to the DLA have been introduced in order to handle large galleries and large variations in pose and to increase the matching accuracy [13]. Another variant of elastic graph matching that aims at increasing the robustness of the method under translations, deformations and changes in background has been proposed [14]. Recently, a novel variant of DLA based on multi-scale morphological dilation-erosion, the so-called Morphological Dynamic Link Architecture (MDLA), has been proposed and tested for face authentication [15].

It is well known that mathematical morphology is very rich in providing means for the representation and analysis of binary and grayscale images [16,17,18]. The morphological representation of images is well suited to the description of the geometrical properties of image objects. The morphological skeleton and the morphological shape decomposition are two popular approaches for morphological shape representation [19]. Algorithms for shape representation are generally divided into two classes, namely, the *external* and the *internal* algorithms. For example, contour description algorithms belong to the former class, whereas region-based algorithms belong to

the latter. In this paper, we deal with an internal algorithm, namely, the *morphological signal decomposition* (MSD). MSD is the decomposition of an image object (in our case of the facial region) into a union of simple components by using the morphological operations of erosion and dilation [20]. MSD has been successfully applied to the decomposition of a binary shape into a union of simple binary shapes using maximal inscribable disks [20]. A flexible search-based shape representation scheme that typically gives more efficient representations than the morphological skeleton and MSD was developed in [21]. MSD can be considered as an alternative of matching pursuit filters, a nonparametric technique for finding the differences among faces [22]. Instead of wavelets that are directional edge detectors (e.g., second partial derivatives of Gaussian densities and their Hilbert transforms) that are used as basis functions used in [22], we employ a simple structuring function (a cylinder of unit height with a circular cross-section of radius 2) and instead of inner products we employ mathematical morphology operations.

A novel dynamic link architecture is proposed in this paper. The method combines the morphological signal decomposition and the elastic graph matching and is tested for face verification. More specifically, we propose the substitution of the responses of a set of Gabor filters by the gray level values in the MSD reconstructed images. The reconstructed images are created by the partial sums of the components during the morphological signal decomposition. The reconstructed images start from a crude approximation of the facial region that recursively becomes finer. Moreover, we derive local coefficients that weigh the contribution of each node as well as of each level in MSD in the total matching error according to their discriminatory power. Experimental results are reported on the M2VTS database [23]. In order to facilitate the understanding of the proposed method, two block diagrams describing the face verification system are sketched that can be used for future reference. The first block diagram illustrates the modules and techniques used in the training phase of the face verification system developed. The output of this phase is the creation of a database for the clients that includes the reference grids, discriminating information for each client (e.g. discriminatory power coefficients), as well as the person-specific thresholds that will be used in the test phase. This block diagram is shown in Figure 1. The second block diagram explains the use of the face verification system in the test phase. A test person claims the identity of a client, and the information stored in the database for the claimed identity is used in order to decide if the test person has indeed the claimed identity. This block diagram is shown in Figure 2.

The proposed method introduces the following new points:

- The use of morphological signal decomposition techniques for analyzing the facial image region in order to form the feature vectors in elastic graph matching.

- The use of a very simple preprocessing step for compensation of illumination variations.
- The use of discriminative information of the different facial features by weighting the nodes of the elastic graph in elastic graph matching.
- The use of discriminative information of the different levels of facial image region representation by weighting the morphological signal decomposition levels in elastic graph matching.

The outline of this paper is as follows. Facial region modeling by using MSD is outlined in Section II. The proposed Morphological Signal Decomposition - Dynamic Link Architecture (MSD-DLA) is described in Section III. Several local discriminatory power coefficients are derived and tested in MSD-DLA framework in Section IV. The experimental protocol used for the evaluation of the verification efficiency of MSD-DLA and experimental results are presented in Section V. Conclusions are drawn in Section VI.

II. Facial Region Modeling using Morphological Signal Decomposition

In this section, the feature extraction algorithm used in the proposed pattern matching scheme is described. The feature extraction is split into two steps:
1. Preprocessing procedure aiming at the compensation of the varying illumination conditions (Section II-A).
2. Facial region modeling (Section II-B).

A. Preprocessing Procedure

Common problems in automated face recognition are the varying illumination conditions and the differences in face position, scale and pose. These problems are crucial in the performance of most techniques. Prior to feature extraction, it is necessary to localize a facial region in the image, that is to perform face detection. A very attractive approach for face detection is based on multiresolution images, also known as *mosaic images*. The face detection algorithm attempts to detect a facial region at a coarse resolution of the image and subsequently to validate the outcome by detecting facial features at the next resolution level [24]. Towards this goal, the method employs a hierarchical knowledge-based pattern recognition system. Recently, a variant of this method that has very good performance for images with a uniform background has been proposed in [25]. By using this method, we may: (i) roughly define a region where the face is included, and (ii) control the placement of a sparse grid over a face for storing a model of each person in dynamic link matching, as described later.

In order to compensate for differences in luminance, the segmentation of the facial area is needed. By using a clustering algorithm, i.e., a K-means algorithm [26], we can distinguish the skin like area from the hair area and the uniform background, because the skin like area

possesses a greater mean intensity than that of the hair region and the background. The mean intensity value of the skin area can differ considerably from one recording to another. Indeed, in Figure 3, the center, i.e., the mean intensity value of the skin region is plotted for the 37 persons in the M2VTS database shot by shot, where each shot comprises a recording of the 37 persons at a specific time instant. Therefore, a normalization procedure is needed to equalize the mean intensity value of the skin area between the test and the reference person prior to the application of the face verification algorithm. The segmentation of the skin-like image region is needed for illumination compensation. That is, the two procedures are sequentially applied and the output of the region segmentation is used as input to the illumination compensation procedure. Let $\mathbb{Z}$ denotes the set of integer numbers. We define the facial image region that can be extracted by using a face detection module, such as the one proposed in [25], as $f(\mathbf{x}) : \mathcal{D} \subseteq \mathbb{Z}^2 \rightarrow \mathbb{Z}$, where $\mathbf{x}$ denotes the pixel coordinates, $f(\cdot)$ is the gray level value assigned to a pixel, and $\mathcal{D}$ is the domain of $f(\mathbf{x})$ that corresponds to a facial region. If M_p is the mean intensity value of the skin area of person p, a simple normalization can be obtained by:

$$f_n(\mathbf{x}) = f(\mathbf{x})\frac{CT}{M_p} \tag{1}$$

where $f_n(\mathbf{x})$ is the normalized image and CT a predefined constant intensity value. By using (1), we can make both the test and the reference image to have the same mean intensity value in the facial area. We did not try to perform any normalization for face scaling and position, because DLA is proven to be robust to small scale and position changes between the reference and the test person facial images. The normalization of the facial image region luminance between two images of the same person is illustrated in Figure 4.

B. Facial Region Modeling

In this section the modeling of a grayscale facial image region by employing the MSD is described. Let us denote by $f_n(\mathbf{x})$ the normalized image produced according to the procedure described in Section II-A. In the following, we shall omit the subscript n for notation simplicity. Given $f(\mathbf{x}) : \mathcal{D} \subseteq \mathbb{Z}^2 \rightarrow \mathbb{Z}$ and a structuring function $g(\mathbf{x}) : \mathcal{G} \subseteq \mathbb{Z}^2 \rightarrow \mathbb{Z}$, the *grayscale dilation* of the image $f(\mathbf{x})$ by the structuring function $g(\mathbf{x})$ is noted by $(f \oplus g)(\mathbf{x})$ and its dual operation, the *grayscale erosion*, is defined as $(f \ominus g)(\mathbf{x})$, [27,18]. Let us choose a structuring function of the form $g(\mathbf{x}) = 1$, $\forall \mathbf{x} : \|\mathbf{x}\| \leq 2$. By definition, $g(\mathbf{x})$ is symmetric. Moreover, it can be seen easily that the structuring function is a cylinder of unit height with a circular cross-section of radius 2. Let $f(\mathbf{x})$ be approximated by:

$$\hat{f}_K(\mathbf{x}) = \sum_{i=1}^{K} f_i(\mathbf{x}) \tag{2}$$

where $f_i(\mathbf{x})$ denotes the i-th component and K is the total number of components. MSD provides a simple method to determine the components. Indeed, if each component is a simple function that can be expressed as:

$$f_i(\mathbf{x}) = [l_i \oplus n_i\, g]\,(\mathbf{x}) \tag{3}$$

where $l_i(\mathbf{x})$ is the so-called *spine* [18] and

$$n_i\, g(\mathbf{x}) = \underbrace{[g \oplus g \oplus \cdots \oplus g]}_{n_i \text{ times}}(\mathbf{x}), \tag{4}$$

an intuitive choice for $n_1\, g(\mathbf{x})$ is the maximal function in $f(\mathbf{x})$. That is, to choose n_1 such that:

$$[f \ominus (n_1 + 1)g]\,(\mathbf{x}) \leq 0 \quad \forall \mathbf{x} \in \mathcal{D}. \tag{5}$$

Accordingly, the first spine is given by:

$$l_1(\mathbf{x}) = [f \ominus n_1\, g]\,(\mathbf{x}). \tag{6}$$

MSD can then be implemented recursively as follows.

Step 1. Initialization: $\hat{f}_0(\mathbf{x}) = 0$.

Step 2. Find the i-th level of decomposition: starting with $n_i = 1$, increment n_i until

$$\left[(f - \hat{f}_{i-1}) \ominus (n_i + 1)g\right](\mathbf{x}) \leq 0. \tag{7}$$

Step 3. Calculate the i-th component by:

$$f_i(\mathbf{x}) = \left\{ \underbrace{\left[(f - \hat{f}_{i-1}) \ominus n_i\, g\right]}_{l_i(\mathbf{x})} \oplus n_i\, g \right\}(\mathbf{x}). \tag{8}$$

Step 4. Calculate the reconstructed image at the i-th level of decomposition:

$$\hat{f}_i(\mathbf{x}) = \hat{f}_{i-1}(\mathbf{x}) + f_i(\mathbf{x}). \tag{9}$$

Step 5. Let $\mathcal{M}(f - \hat{f}_i)$ be a measure of the approximation of the image $f(\mathbf{x})$ by its reconstruction $\hat{f}_i(\mathbf{x})$ at the i-th level of decomposition. Let also L be the maximum number of image components used for reconstruction. Increment i and go to Step 2 until $i > L$ or alternatively $\mathcal{M}(f - \hat{f}_i) < T$ where T is a predefined threshold.

Figure 5 shows the block diagram of the MSD. In Figure 5, the component extraction module (CE) implements the Step 2 and 3 of the algorithm outlined above. There are several reasons supporting the use of MSD as a feature extraction algorithm, namely:

1. The decomposition of a complex object yields simple components that conform with our intuition. In our case, the component is the maximal inscribable cylinder. In addition, the method is object-independent, in the sense that it employs generic structuring functions that do not depend on the object that is approximated [20].
2. It allows arbitrary amounts of detail to be computed and also allows the abstraction from the detail [20].
3. The representation is unique.
4. MSD employs grayscale erosions and dilations with a flat structuring function, namely, a cylinder of unit height having a circular cross-section of radius 2. Grayscale erosions and dilations with a flat structuring function can be computed very efficiently using running min/max selection algorithms [18].

In the case of noisy images, the sensitivity of the crisp mathematical morphology can be reduced by using soft or fuzzy mathematical morphology [28,29].

III. Combined use of Morphological Signal Decomposition and Dynamic Link Architecture

Traditionally, linear methods like the Fourier transform, the Walsh-Hadamard transform, the Gaussian filter banks, the wavelets and Gabor elementary functions have been used in image pyramids. An alternative way to linear techniques is to use morphological signal decomposition techniques. In this paper, we propose the substitution of Gabor-based feature vectors used in dynamic link matching by feature vectors that are extracted from the reconstructed images $\hat{f}_i(\mathbf{x})$ at the last K consecutive decomposition levels $i = L-K-1, \ldots, L$, where L denotes the maximal number of decomposition levels. The reasoning for this decision is to omit the very first reconstructed images that are common to all individuals in the database. The value K=15 was chosen because it yields good results in our experiments. That is, the gray level information $\hat{f}_i$ at the grid node $\mathbf{x}$ for the decomposition levels $i = L-14, \ldots, L$ along with the gray level information f is concatenated to form the feature vector $\mathbf{j}(\mathbf{x})$ (also called *jet*) [4]:

$$\mathbf{j}(\mathbf{x}) = \left(\hat{f}_{L-K-1}(\mathbf{x}), \ldots, \hat{f}_L(\mathbf{x}), f(\mathbf{x})\right). \tag{10}$$

Alternatively, one may also use the feature vector $(\hat{f}_{L-K-1}(\mathbf{x}), \ldots, \hat{f}_L(\mathbf{x}), f(\mathbf{x}) - \hat{f}_L(\mathbf{x}))$. Figure 6 depicts a series of reconstructed images for the facial image region of a sample person from the database at nineteen decomposition levels. The 20th image at the bottom right is the original facial image region that is decomposed. Only the last fifteen reconstructed images have been employed in the DLA that employs the feature vectors given by (10). In Figure 7 the residual of the morphological signal decomposition is illustrated.

Let the superscripts t and r denote a test and a reference person (or grid), respectively. The L_2 norm between the feature vectors at the same grid node has been used as a (signal) similarity measure, i.e.:

$$C_v(\mathbf{j}(\mathbf{x}_l^t), \mathbf{j}(\mathbf{x}_l^r)) = \|\mathbf{j}(\mathbf{x}_l^t) - \mathbf{j}(\mathbf{x}_l^r)\|. \tag{11}$$

As in DLA [4], the quality of a match is evaluated by taking into account the grid deformation as well. Let us denote by $\mathcal{V}$ the set of grid nodes. Then, an additional cost function is used:

$$C_e(l, j) = C_e(\mathbf{d}_{lj}^t, \mathbf{d}_{lj}^r) = \|\mathbf{d}_{lj}^t - \mathbf{d}_{lj}^r\| \quad \forall l \in \mathcal{V}; j \in \mathcal{N}(l) \tag{12}$$

where $\mathcal{N}(l)$ denotes the neighborhood of a vertex l (e.g., a four-connected neighborhood in our case) and $\mathbf{d}_{lj} = \mathbf{x}_l - \mathbf{x}_j$. The objective in DLA is to find the test grid node coordinates $\{\mathbf{x}_l^t,\ l \in \mathcal{V}\}$ that minimize:

$$C(\{\mathbf{x}_l^t\}) = \sum_{l \in \mathcal{V}} \{C_v(\mathbf{j}(\mathbf{x}_l^t), \mathbf{j}(\mathbf{x}_l^r)) + \lambda \sum_{j \in \mathcal{N}(l)} C_e(\mathbf{d}_{lj}^t, \mathbf{d}_{lj}^r)\}. \tag{13}$$

One may interpret the optimization of (13) as a simulated annealing [30] with an additional penalty, i.e., the cost for grid deformations in the objective function. Since the cost function (12) does not penalize translations of the entire graph, the random configuration $\mathbf{x}_l$ can take the form of a random translation $\mathbf{s}$ of the (undeformed) reference grid and a bounded local perturbation $\mathbf{q}_l$, i.e.:

$$\mathbf{x}_l^t = \mathbf{x}_l^r + \mathbf{s} + \mathbf{q}_l\ ; \quad \|\mathbf{q}_l\| \leq q_{\max} \tag{14}$$

where the choice of $q_{\max}$ controls the rigidity/plasticity of the graph. It is evident that the aforementioned grid matching procedure, i.e., (13) and (14), differs from the two stage coarse-to-fine optimization procedure proposed in [4]. In our approach, we replace the two stage optimization procedure with a probabilistic hill climbing algorithm [30, pp. 9-13] which attempts to find the best configuration $\{\mathbf{s}, \{\mathbf{q}_l\}\}$ at each step. A sparse grid of 8×8 equally spaced nodes has been employed. Figure 8 depicts the grids formed during the matching procedure.

IV. Incorporation of Discriminatory Power Coefficients in MSD-DLA

Having described the facial modeling that provides the feature vectors to the elastic graph matching procedure outlined in Section 3, we now proceed in assessing the discriminatory power of each decomposition level and grid node. Intuitively we expect that all feature vectors will not be equally powerful in discriminating one person from others. Thus, the weighting of the graph nodes according to their discriminatory power in elastic graph matching can improve the performance of the verification algorithm. Several methods that address this issue have been proposed in the literature. A Bayesian approach yields the more reliable nodes for gender

identification, beard and glass detection in bunch graphs [31]. An automatic weighting of the nodes according to their significance by employing a local discriminant is proposed in [5]. A weighted average of the feature vector similarities by a set of coefficients that take into account the importance of each feature in assigning a test person to a specific class is investigated in [32].

Some facial features (e.g., the eyes, the nose) are more crucial in the verification procedure than others. For example, the system described in [22] employs regions that contain the noise tip, the left eye, the right eye and the noise bridge. It is well known that both global and local features are used for face recognition in a hierarchical manner, where the local features provide a finer classification [33]. However, the question which features humans use for face recognition has been subject to much debate [34].

Thus, it will be helpful if we calculate weighting coefficients for the grid nodes that correspond to the significance of the facial features in the verification procedure. To do so, we formulate a two-class problem whose objective is to weigh the signal similarity measure at node l given by (11) using class-dependent *discriminatory power coefficients* (DPCs), $DP_l(\mathcal{S}_r)$, so that when person t claims the identity of person r the distance between them is computed by:

$$D(t,r) = \sum_{l \in \mathcal{V}} \frac{DP_l(\mathcal{S}_r) C_v(\mathbf{j}(\mathbf{x}_l^t), \mathbf{j}(\mathbf{x}_l^r))}{\sum_{k \in \mathcal{V}} DP_k(\mathcal{S}_r)} \tag{15}$$

where $\mathcal{S}_r$ denotes the class of the reference person r. $DP_l(\mathcal{S}_r)$ of the l-th grid node for the class $\mathcal{S}_r$ is a factor that shows how well the intra-class distances are separated from the inter-class distances at this node. Accordingly, at node l we consider the distances measured between frontal facial images of the person r (i.e., the distances $\forall t, r \in \mathcal{S}_r$) and the distances measured between frontal facial images of the person r and all the remaining persons in the database, -i.e., the distances $\forall r \in \mathcal{S}_r$ and $\forall t \in (\mathcal{S} - \mathcal{S}_r)$, where $\mathcal{S}$ denotes the set of all classes in the database. Let $m_{\text{intra}}(\mathcal{S}_r, l)$ be the mean intra-class distance for the class $\mathcal{S}_r$ at grid node l:

$$m_{\text{intra}}(\mathcal{S}_r, l) = E\{C_v(\mathbf{j}(\mathbf{x}_l^t), \mathbf{j}(\mathbf{x}_l^r))\} \;\; \forall\, t, r \in \mathcal{S}_r \tag{16}$$

and $m_{\text{inter}}(\mathcal{S}_r, l)$ be the mean inter-class distance between the class $\mathcal{S}_r$ and $(\mathcal{S} - \mathcal{S}_r)$ at the same node:

$$m_{\text{inter}}(\mathcal{S}_r, l) = E\{C_v(\mathbf{j}(\mathbf{x}_l^t), \mathbf{j}(\mathbf{x}_l^r))\} \;\; \forall\, r \in \mathcal{S}_r, \;\; \forall t \in (\mathcal{S} - \mathcal{S}_r). \tag{17}$$

Let $\sigma^2_{\text{intra}}(\mathcal{S}_r, l)$ and $\sigma^2_{\text{inter}}(\mathcal{S}_r, l)$ be the variances of the intra-class node distances and the inter-class node distances, respectively, i.e.:

$$\sigma^2_{\text{intra}}(\mathcal{S}_r, l) = E\{C_v^2(\mathbf{j}(\mathbf{x}_l^t), \mathbf{j}(\mathbf{x}_l^r))\} - m^2_{\text{intra}}(\mathcal{S}_r, l) \;\; \forall\, t, r \in \mathcal{S}_r \tag{18}$$

$$\sigma^2_{\text{inter}}(\mathcal{S}_r, l) = E\{C_v^2(\mathbf{j}(\mathbf{x}_l^t), \mathbf{j}(\mathbf{x}_l^r))\} - m^2_{\text{inter}}(\mathcal{S}_r, l) \;\; \forall\, r \in \mathcal{S}_r, \;\; \forall t \in (\mathcal{S} - \mathcal{S}_r). \tag{19}$$

Several measures for the discriminatory power of each node were tested. Obviously, the grid nodes that do not possess any discriminatory power should be discarded in the face verification procedure. That is,

$$\text{if} \quad m_{\text{inter}}(\mathcal{S}_r, l) < m_{\text{intra}}(\mathcal{S}_r, l) \quad \text{then} \quad DP_l(\mathcal{S}_r) = 0. \tag{20}$$

Rule (20) is applicable when the two classes cannot be discriminated by the feature vectors at node l. A plausible choice for $DP_l(\mathcal{S}_r)$ is the distance between the $m_{\text{intra}}(\mathcal{S}_r, l)$ and $m_{\text{inter}}(\mathcal{S}_r, l)$, i.e., the distance between the center of client distances and the center of the impostor distances:

$$DP_l(\mathcal{S}_r) = m_{\text{inter}}(\mathcal{S}_r, l) - m_{\text{intra}}(\mathcal{S}_r, l). \tag{21}$$

Another measure of the discriminatory power of grid node l for class $\mathcal{S}_r$ is *Fisher's Linear Discriminant* (FLD) function (or first canonical variate) [26,35]. The latter takes into consideration both the distance between the two cluster centers as well as the compactness of the two clusters in order to yield a DPC of the form:

$$DP_l(\mathcal{S}_r) = \frac{(m_{\text{inter}}(\mathcal{S}_r, l) - m_{\text{intra}}(\mathcal{S}_r, l))^2}{\sigma^2_{\text{inter}}(\mathcal{S}_r, l) + \sigma^2_{\text{intra}}(\mathcal{S}_r, l)}. \tag{22}$$

It can be easily proven that the DPCs given by (22) minimize the ratio:

$$J = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} DP_i(\mathcal{S}_r)DP_j(\mathcal{S}_r)(m_{\text{inter}}(\mathcal{S}_r, i) - m_{\text{intra}}(\mathcal{S}_r, i))(m_{\text{inter}}(\mathcal{S}_r, j) - m_{\text{intra}}(\mathcal{S}_r, j))}{\sum_{i=1}^{N} DP_i^2(\mathcal{S}_r)\left(\frac{\sigma^2_{\text{inter}}(\mathcal{S}_r,i)+\sigma^2_{\text{intra}}(\mathcal{S}_r,i)}{|(m_{\text{inter}}(\mathcal{S}_r,i)-m_{\text{intra}}(\mathcal{S}_r,i)|}\right)} \tag{23}$$

where N is the cardinality of $\mathcal{V}$, i.e., 64 for 8×8 sparse grids. Provided that the numerator in (22) remains constant, $DP_l(\mathcal{S}_r)$ is maximized when the denominator $\{\sigma^2_{\text{inter}}(\mathcal{S}_r, l) + \sigma^2_{\text{intra}}(\mathcal{S}_r, l)\}$ is minimized. That is, both variances should be small. Accordingly, we interpret (22) as an **AND** rule for the cluster variances. Alternatively, one can use a more relaxed criterion of the form:

$$DP_l(\mathcal{S}_r) = \frac{(m_{\text{inter}}(\mathcal{S}_r, l) - m_{\text{intra}}(\mathcal{S}_r, l))^2}{\sigma_{\text{inter}}(\mathcal{S}_r, l)\sigma_{\text{intra}}(\mathcal{S}_r, l)}. \tag{24}$$

The denominator of (24) is interpreted then as an **OR** rule for the cluster variances. The DPCs (22) are shown for several persons in the database in Figure 9. As one can see, the nodes that correspond to the key facial features (e.g. the beard, the bald head), are weighted with larger coefficients, i.e., they are shown as white disks, than other nodes.

In addition to the discriminatory power of each grid node, we can assign discriminatory coefficients to the morphological signal decomposition levels as well. Thus, another factor that can be studied is the discriminatory power of each reconstruction level. To do so, we define the following feature vector:

$$\mathbf{j}'(i) = (\hat{f}_i(\mathbf{x}_1), \hat{f}_i(\mathbf{x}_2), \ldots, \hat{f}_i(\mathbf{x}_N)) \quad i = L - K - 1, \ldots, L + 1 \tag{25}$$

where N denotes, as before, the number of grid nodes and i is the current decomposition level. For $i = L + 1$, we have a jet that comprises the gray levels of the initial image. According to this approach, the face is modeled by $K + 1$ feature vectors, $\mathbf{j}'(1), \ldots, \mathbf{j}'(K + 1)$. The signal similarity measure (11) is now redefined as:

$$C'_v(\mathbf{j}'^t(i), \mathbf{j}'^r(i)) = \| \mathbf{j}'^t(i) - \mathbf{j}'^r(i) \| . \tag{26}$$

Our objective is to calculate the coefficients $w_i(\mathcal{S}_r)$, $i = 1, \ldots, K + 1$, that weigh each decomposition level of MSD and the initial image $f(\mathbf{x})$. It is straightforward to repeat the procedure described in this section in order to calculate the desired coefficients for the MSD levels. Having found the weighting coefficients $w_i(\mathcal{S}_r)$ for the MSD levels for each class $\mathcal{S}_r$, the signal similarity measure (11) can be rewritten as follows:

$$\mathbf{C}_v(\mathbf{j}(\mathbf{x}_l^t), \mathbf{j}(\mathbf{x}_l^r)) = \sqrt{w_{L+1}(\mathcal{S}_r)[f(\mathbf{x}_l^t) - f(\mathbf{x}_l^r)]^2 + \sum_{i=L-K+1}^{L} w_i(\mathcal{S}_r)[\hat{f}_i(\mathbf{x}_l^t) - \hat{f}_i(\mathbf{x}_l^r)]^2}. \tag{27}$$

In Figure 10, the DPCs $w_i(\mathcal{S}_r)$ for all decomposition levels are depicted for four persons (i.e., BP,BS,CC and CM) in the database. For example, the fifteenth decomposition level is found to play a more crucial role for several people (e.g., BS, CC) which is manifested by the greater value that the corresponding DPC attains for these individuals. This is not the case for the individuals BP and CM. Similar observations can be made for the other decomposition levels. By determining first the weighting coefficients $w_i(\mathcal{S}_r)$, $i = 1, 2, \ldots, L + 1$, and substituting (27) into (15), we allow the weighting of both the MSD levels and the grid nodes. It will be shown in the next section that this approach enhances the discrimination among the classes. The discriminatory coefficients proposed in this section can easily be computed during the application of the verification algorithm to any database following any experimental protocol. They can easily be modified when persons are added to, or, deleted from the database, because the computation (16)-(22) can be made incrementally. The cost of updating the DPCs given by (22) is only 7 multiplications and 5 additions per distance measure. This number should be compared to the cost of updating the singular values and the left singular vectors of the data matrix of N feature vectors in an incremental fashion [36]:

$$FL(N, M) = \mathcal{O}\left((N + M)\ \min(N, M)\ \log_2^2 \epsilon\right) \quad \text{floating point operations} \tag{28}$$

where M is the feature vector dimensionality (i.e., $M = K + 1$ in our case) and ϵ is the machine precision. A cost $FL(N_t, M)$, where N_t is the total number of training frontal face images, would be the paid if principal component analysis (PCA) were only employed, i.e., for an eigenface approach. If Linear Discriminant Analysis (LDA) were employed then one would roughly pay

$FL(N_c, M) + \mathcal{O}(M^3)$ where we assume that an additional frontal face image for one client is appended to the training database, each client is represented in the database by N_c frontal face images, and the simultaneous diagonalization of the within-class and between-class scatter matrices is performed [37, p. 32]. The sum of the aforementioned costs would be paid if a combination of PCA and LDA were applied, as in [10,15,34].

V. Description of Experiments and Results

MSD-DLA has been tested on the M2VTS database [23]. The database contains 37 persons' video data, which include speech consisting of uttering digits and color image sequences of rotated heads. Four recordings (i.e., shots) of the 37 persons have been collected at different time instants. Let BP, BS, CC, ..., XM be the identity codes of the persons included in the database. In our experiments, the sequences of rotated heads have been considered by using only the luminance information at a resolution of 286 $\times$ 350 pixels. From each image sequence, one frontal image has been chosen based on symmetry considerations. The details of face detection are out of the scope of this paper. Four experimental sessions have been implemented to yield client and impostor claims that are quantified by a distance - e.g., (15) - by employing a combination of "leave-one-out" and rotation estimates. Each session consists of a training and a test procedure that are applied to their training set and test set, respectively. Figure 11 depicts the experimental protocol, when person BP and shot 04 are excluded from the training set. The 4th shot is the test set in that case. The comparisons shown for person BP are repeated for all other persons in the database. Obviously, similar comparisons are made by rotating among the available shots.

A. Training procedure

The training set is built of 3 out of 4 available shots comprising 36 out of 37 available persons. This amounts to $3 \times 36 = 108$ images. All images of the same person build a class. By using these images one may compute: (i) 6 distance measures for all pairwise combinations between the different sample images in the same class, and, (ii) another 6 distance measures for each pairwise combination between the sample images of any two different classes. It is worth noticing that sample images that originate from different shots are used in all pairwise combinations. In total, 6 intra-class distance measures and $35 \times 6 = 210$ inter-class distance measures are computed for each of the 36 trained classes. MSD-DLA has been used to yield all the aforementioned distance measures. Having computed all the 216 distance measures for each trained class, the objective in the training procedure is to determine a threshold on the distance measures that should ideally enable the distinction between the sample images that belong to

the trained class under study, and the sample images that belong to any other class. These person-specific thresholds are to be used in the test procedure. To elucidate the derivation of person-specific thresholds, let us consider the case we leave out the fourth shot and exclude all frontal face images of person BP from the training set. 36 thresholds are determined, namely, $T_{BS}(4, BP)$, $T_{CC}(4, BP)$, ..., $T_{XM}(4, BP)$. The threshold $T_{BS}(4, BP)$ is used to discriminate sample images of person BS that originate from shots 1–3 against all the sample images of the remaining 35 persons (e.g., $CC-XM$) which originate from any of these shots. The threshold is computed by employing the order statistics of the sequence of the minimum inter-class distances for each of the training impostors, e.g., for person BS the training impostors are CC-XM. The vector of 35 minimum distances is ordered in ascending order of their magnitude. An obvious choice for $T_{BS}(4, BP)$ is to choose the minimum impostor distance, denoted by $D_{(1)}$. In the more general case $T_{BS}(4, BP)$ could be chosen as:

$$T_{BS}(4, BP) = D_{(1+Q)}, \quad Q = 0, 1, 2, \ldots . \tag{29}$$

B. Test Procedure

In the test procedure, three shots create the training set while the fourth one is used as the test set. Each person in the test set has been considered in turn as an impostor, while the 36 remaining persons have been treated as clients. Each client tries to access under its own identity, while the impostor tries to access under the identity of each of the 36 clients in turn. This is tantamount to 36 authentic tests and 36 impostor tests. By repeating the procedure four times, $4 \times 37 \times 36 = 5328$ authentic claims and 5328 impostor claims are realized in total. In each authentic or impostor test, the reference grids built for each class are matched and adapted to the feature vectors computed at every image pixel of the frontal face image of a test person, that can be either a client or an impostor, using MSD-DLA. Again, we have used the minimum intra-class or inter-class distances in the comparisons, i.e., for person BP:

$$\overline{D}(BP_4, \{BS\}) = \min\{D(BP_4, BS_1), D(BP_4, BS_2), D(BP_4, BS_3)\} \tag{30}$$

where the first ordinate in distance computations denotes the frontal face image of a test person and the second ordinate denotes the reference grid for a trained class. Then, the resulting distance measure is compared against the threshold that has been derived during the training procedure. For example, if we consider the client and impostor tests produced when person BP is excluded, a false acceptance occurs when:

$$\overline{D}(BP_4, \{\mathcal{X}\}) \leq T_{\mathcal{X}}(4, BP) \quad \mathcal{X} = BS, CC, \ldots, XM \tag{31}$$

and a false rejection occurs when:

$$\overline{D}(\mathcal{X}_4, \{\mathcal{X}\}) > T_{\mathcal{X}}(4, BP) \quad \mathcal{X} = BS, CC, \ldots, XM. \tag{32}$$

C. Receiver Operating Characteristics

For a particular choice of parameter Q, a collection of thresholds is determined that defines an *operating state* of the test procedure. For such an operating state, a false acceptance rate (FAR) and a false rejection rate (FRR) can be computed. By varying the parameter Q several operating states result. Accordingly, we may create a plot of FRR versus FAR with a varying operating state as an implicit set of parameters. Equivalently, the scalar Q can be used as a varying parameter. This plot is the *Receiver Operating Characteristic* (ROC) of the verification technique. ROC curves for MSD-DLA are plotted in Figure 12. The Equal Error Rate (EER) of MSD-DLA (i.e., the operating state of the method when FAR equals FRR) is another common figure of merit used in the comparison of verification techniques.

The ensemble {test images, verification algorithm} is a source of binary events:

- 1 for false rejection (or false acceptance) with probability p_{FR} (or p_{FA}), and,
- 0 for no error with probability $(1 - p_{FR})$ (or $(1 - p_{FA})$),

respectively. These events can be described by Bernoulli trials. Let us denote by $\hat{p}_{FR}$ and $\hat{p}_{FA}$ the estimates of FRR and FAR, respectively, that are measured according to the experimental protocol outlined above. The exact γ confidence interval of p_{FR} and p_{FA} is the segment between the two roots of the quadratic equation [38]:

$$(p - \hat{p})^2 = \frac{z^2_{(1+\gamma)/2}}{N}\, p\,(1 - p), \quad p = p_{FR}, p_{FA} \text{ and } N = 5328 \tag{33}$$

where z_u is the u-percentile of the standard Gaussian distribution having zero mean and unit variance. The $\gamma = 95\%$ confidence interval of the FAR and FRR is indicated with a horizontal and a vertical error bar, respectively, for all ROCs subsequently.

The EER of the MSD-DLA without weighting the grid nodes with DPCs according to the experimental protocol, described in this section, is found to be 11.89%. Figure 12 depicts the Receiver Operating Characteristics of the MSD-DLA by using DPCs given by (21), (22), and (24), respectively. When the ROC curve is used as a figure of merit for the performance of an authentication algorithm, the smaller the area under the ROC curve for a certain method the better the methods performance is. The plot indicates that the verification capability of MSD-DLA is improved by weighting the grid node contributions by the DPCs given by (22) or (24). The superimposed 95% confidence intervals indicate that the reduction in FAR and FRR is statistically significant. Table I summarizes the EERs achieved by MSD-DLA using the

several DPCs proposed in the paper. It is seen that the use of DPCs given by (22) yields the best EER which corresponds to an improvement of 5.5%.

The ROC, when DPCs are applied on the MSD levels only, is plotted in Figure 13. In the same figure, the plot of the MSD-DLA without DPCs is included for comparison purposes. It can be observed that, without weighting the grid nodes, the discriminant analysis on the MSD levels does not offer a significant reduction in the area under the ROC of the raw MSD-DLA. Moreover, it can be seen that the achievements in FAR and FRR are rather statistically insignificant due to the overlap between the 95% confidence intervals. On the contrary, a statistically significant improvement is attained by combining DPCs for both grid nodes and MSD levels, as can be seen in the same figure. Indeed, in this case the EER is found to be 5.7% following the same experimental protocol. Overall, an improvement of 6.2% has been achieved. To facilitate the comparison, the best ROC, when DPCs (22) are used on the grid nodes, is repeated in Figure 13 as well.

D. Comparison with other methods

Table II compares the EER achieved by the proposed MSD-DLA, the classical GDLA [5], and the gray level frontal face matching [39] in the same database according to the described experimental protocol. The EER achieved when DPCs are used in MSD-DLA and the EER achieved when local discriminant analysis is applied for feature selection in GDLA [5] is also quoted. It is obvious that the proposed combination of morphological signal decomposition and elastic graph matching with discriminatory power coefficients offers a higher verification efficiency than the standard dynamic link architecture with Gabor wavelets.

VI. Conclusions

A novel dynamic link architecture that employs morphological signal decomposition as a feature extraction mechanism has been developed and tested for frontal face verification. The resulting feature vectors are proven to be at least equally powerful as the feature vectors resulting by convolving the frontal face image with Gabor wavelets. Local discriminatory power coefficients have been derived that weigh the contribution of each node and each morphological signal decomposition level in the distance measure according to their discriminatory power. The proposed weighting methods do not need more frontal images than those delivered by the experimental protocol described in Section V. That is, three training frontal face images per person are needed. This is not the case with [5,15] where 15-20 frontal face images per person are needed to derive the weighting coefficients. The equal error rate of the morphological signal decomposition dynamic link architecture on the M2VTS database is found to be 11.89%. By

weighting the contribution of each grid node and each level of the morphological signal decomposition, a significant improvement in EER has been achieved, yielding an equal error rate of 5.7%. All figures of merit tabulated in Table II have been obtained using the same experimental protocol and, therefore, the results are absolutely comparable. The present paper has not discussed at all issues related to face localization and the treatment of varying illumination conditions or pose. This is because face detection is a difficult, open problem in the literature. For the treatment of varying illumination conditions and pose, the interested reader may refer to [11]. Potential applications of the proposed method include person verification in tele-services and video surveillance.

Future research will address a recursive calculation of discriminatory power coefficients $DP_l(\mathcal{S}_r)$ and $w_i(\mathcal{S}_r)$ for maximizing a total quality of match.

References

[1] R. Chellappa, C.L. Wilson, and S. Sirohey, "Human and machine recognition of faces: A survey," *Proceedings of the IEEE*, Vol. 83, No. 5, pp. 705–740, May 1995.

[2] R. Brunelli and T. Poggio, "Face recognition: Features versus templates," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 10, pp. 1042–1052, 1993.

[3] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, Vol. 3, No. 1, pp. 71–86, 1991.

[4] M. Lades, J.C. Vorbrüggen, J. Buhmann, J. Lange, C.v.d. Malsburg, R.P. Würtz, and W. Konen, "Distortion invariant object recognition in the dynamic link architecture," *IEEE Trans. on Computers*, Vol. 42, No. 3, pp. 300–311, March 1993.

[5] B. Duc, S. Fischer, and J. Bigün, "Face authentication with gabor information on deformable graphs," *IEEE Transactions on Image Processing*, Vol. 8, No. 4, pp. 504–516, April 1999.

[6] M. Kirby and L. Sirovich, "Application of the Karhunen-Loeve procedure for the characterization of faces," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 12, No. 1, pp. 103–108, January 1990.

[7] G.W. Cottrell and M. Fleming, "Face recognition using unsupervised feature extraction," in *Int. Neural Network Conf.*, Paris, July 1990, Vol. 1, pp. 322–325.

[8] J. Zhang, Y. Yan, and M. Lades, "Face recognition: Eigenface, elastic matching and neural nets," *Proceedings of the IEEE*, Vol. 85, No. 9, pp. 1423–1435, September 1997.

[9] Y. Adini, Y. Moses, and S. Ullman, "Face recognition: The problem of compensating for changes in illumination direction," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp. 721–732, July 1997.

[10] P.N. Belhumer, J.P. Hespanha, and D.J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp. 711–720, July 1997.

[11] C. Kotropoulos, A. Tefas, and I. Pitas, "Morphological elastic graph matching applied to frontal face authentication under well-controlled and real conditions," *Pattern Recognition*, Vol. 33, No. 12, pp. 31–43, October 2000.

[12] B.S. Manjunath, R. Chellappa, and C. v.d. Malsburg, "A feature based approach to face recognition," in *Proc. of the IEEE Int. Conf. on Computer Vision and Pattern Recognition (CVPR-92)*, 1992, pp. 373–378.

[13] L. Wiskott, J.-M. Fellous, N. Krüger, and C. v.d. Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp. 775–779, July 1997.

[14] R.P. Würtz, "Object recognition robust under translations, deformations, and changes in background," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp. 769–775, July 1997.

[15] C. Kotropoulos, A. Tefas, and I. Pitas, "Frontal face authentication using discriminating grids with morphological feature vectors," *IEEE Trans. on Multimedia*, Vol. 2, No. 1, pp. 14–26, March 2000.

[16] J. Serra, *Image Analysis and Mathematical Morphology*, Academic Press, London, 1982.

[17] J. Serra, *Image Analysis and Mathematical Morphology, Part II: Theoretical Advances*, Academic Press, London, 1988.

[18] I. Pitas and A.N. Venetsanopoulos, *Nonlinear Digital Filters: Principles and Applications*, Kluwer Academic Publ., Boston, MA, 1990.

[19] J.M. Reinhardt and W.E. Higgins, "Comparison between the morphological skeleton and morphological shape decomposition," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 18, No. 9, pp. 951–957, September 1996.

[20] I. Pitas and A.N. Venetsanopoulos, "Morphological shape decomposition," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 12, No. 1, pp. 38–45, January 1990.

[21] J.M. Reinhardt and W.E. Higgins, "Efficient morphological shape representation," *IEEE Trans. on Image Processing*, Vol. 5, No. 1, pp. 89–101, January 1996.

[22] P. J. Phillips, "Matching pursuit filters applied to face identification," *IEEE Trans. on Image Processing*, Vol. 7, No. 8, pp. 1150–1164, August 1998.

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

[24] G. Yang and T.S. Huang, "Human face detection in a complex background," *Pattern Recognition*, Vol. 27, No. 1, pp. 53–63, 1994.

[25] C. Kotropoulos and I. Pitas, "Rule-based face detection in frontal views," in *Proc. of the IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP-97)*, Munich, Germany, April 1997, Vol. IV, pp. 2537–2540.

[26] R.J. Schalkoff, *Pattern Recognition: Statistical, Structural and Neural Approaches*, John Wiley and Sons, New York, 1992.

[27] R.M. Haralick, "Image analysis using mathematical morphology," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 9, No. 4, pp. 532–550, July 1987.

[28] L. Koskinen, J. Astola, and Y. Neuvo, "Soft morphological filters," in *Proc. SPIE Symp. on Image Algebra and Morphological Image Processing*, 1991, Vol. 1568, pp. 262–270.

[29] I. Bloch and H. Maitre, "Fuzzy mathematical morphologies: A comparative study," *Pattern Recognition*, Vol. 28, pp. 1341–1387, September 1995.

[30] R.H.J.M. Otten and L.P.P.P. van Ginneken, *The Annealing Algorithm*, Kluwer Academic Publishers, Norwell, MA, 1989.

[31] L. Wiskott, "Phantom faces for face analysis," *Lecture Notes in Computer Science:Computer Analysis of Images and Patterns*, Vol. 1296, pp. 480–487, 1997.

[32] N. Krüger, "An algorithm for the learning of weights in discrimination functions using a priori constraints," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp. 764–768, July 1997.

[33] D.C. Hay and A.W. Young, "The human face," in *Normality and Pathology in Cognitive Function*, H.D. Ellis, Ed., pp. 173–202. Academic Press, London, 1982.

[34] J.J. Weng and D.L. Swets, "Face recognition," in *Biometrics: Personal Identification in Networked Society*, A. Jain, R. Bolle, and S. Pankanti, Eds., pp. 65–86. Kluwer Academic Publishers, Norwell, MA, 1999.

[35] R.A. Fisher, "The use of multiple measures in taxonomic problems," *Ann. Eugenics*, Vol. 7, pp. 179–188, 1936.

[36] M. Gu and S.C. Eisenstat, "A stable and fast algorithm for updating the singular value decomposition," Tech. Rep. TR-966, Yale University, 1994.

[37] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic Press, San Diego, CA, 1990.

[38] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, McGraw-Hill, New York, 1991.

[39] S. Pigeon and L. Vandendorpe, "Image-based multimodal face authentication," *Signal Processing*, Vol. 69, pp. 59–79, August 1998.

List of Tables

I Equal error rates (EERs) achieved by using several discriminatory power coefficients. 22

II Comparison of Equal Error Rates for several verification techniques in the M2VTS database. 22

List of Figures

1 Block diagram of the training phase of the face verification system. 23
2 Block diagram of the test phase of the face verification system. 23
3 Mean intensity value of the skin area for each person in the database. 23
4 Compensation for luminance variations: (a),(d): Initial images of the same person from different shots. (b),(e): Segmentation of the original image in three regions using K-means clustering for K=3. (c),(f): Normalized images having the same mean intensity value in the skin region. 24
5 Block diagram of MSD. 24
6 Reconstructed images at nineteen decomposition levels. The image at the bottom right is the original one. 25
7 Morphological Signal Decomposition residual. 25
8 The graph matching procedure in MSD-DLA: model grid, best grid for the test person after translation and deformation of the grid. Figures (a),(d): Reference person. Figures (b),(e): The test person is identical to the reference one. Figures (c),(f): The test person is different from the reference one. 26
9 Discriminatory power coefficients for the grid nodes in Dynamic Link Architecture. The brighter a node is the bigger discriminatory power possesses. The intensity of the nodes is normalized for visualization reasons. 27
10 Discriminatory power coefficients of the morphological signal decomposition levels for several persons in the database. 28
11 Experimental Protocol. 28
12 Receiver Operating Characteristic curves of MSD-DLA where the contribution of each grid node is weighted by several choices of discriminatory power coefficients, (DPCs). 29
13 Receiver Operating Characteristic curves of raw MSD-DLA, MSD-DLA when Discriminatory Power Coefficients are employed for Morphological Signal Decomposition levels only, MSD-DLA when Discriminatory Power Coefficients are employed for grid nodes only, MSD-DLA when both the contributions of Morphological Signal Decomposition levels and grid nodes are weighted. 29

TABLE I

EQUAL ERROR RATES (EERS) ACHIEVED BY USING SEVERAL DISCRIMINATORY POWER COEFFICIENTS.

Discriminatory Power Coefficient for Node l		EER (%)
1		11.9
$m_{\text{inter}}(\mathcal{S}_r, l) - m_{\text{intra}}(\mathcal{S}_r, l)$	(21)	8.0
$\frac{(m_{\text{inter}}(\mathcal{S}_r,l)-m_{\text{intra}}(\mathcal{S}_r,l))^2}{\sigma^2_{\text{inter}}(\mathcal{S}_r,l)+\sigma^2_{\text{intra}}(\mathcal{S}_r,l)}$	(22)	**6.44**
$\frac{(m_{\text{inter}}(\mathcal{S}_r,l)-m_{\text{intra}}(\mathcal{S}_r,l))^2}{\sigma_{\text{inter}}(\mathcal{S}_r,l)\sigma_{\text{intra}}(\mathcal{S}_r,l)}$	(24)	6.63

TABLE II

COMPARISON OF EQUAL ERROR RATES FOR SEVERAL VERIFICATION TECHNIQUES IN THE M2VTS DATABASE.

Verification Technique	EER (%)
MSD-DLA	11.89
MSD-DLA with DPCs (22) & (27)	**5.7**
Gray level frontal face matching [39]	8.5
Discriminant GDLA [5]	6.0-9.2
GDLA [5]	10.8-14.4

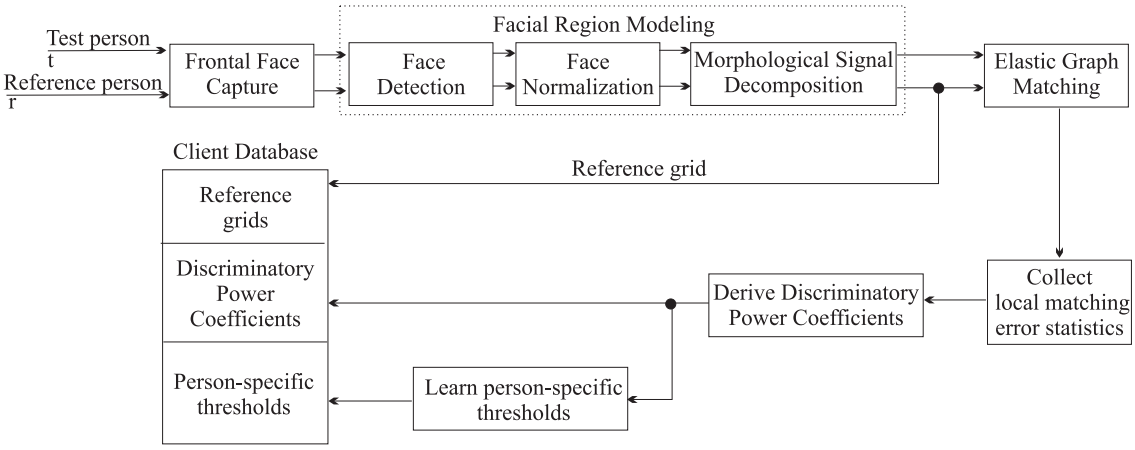

Fig. 1. Block diagram of the training phase of the face verification system.

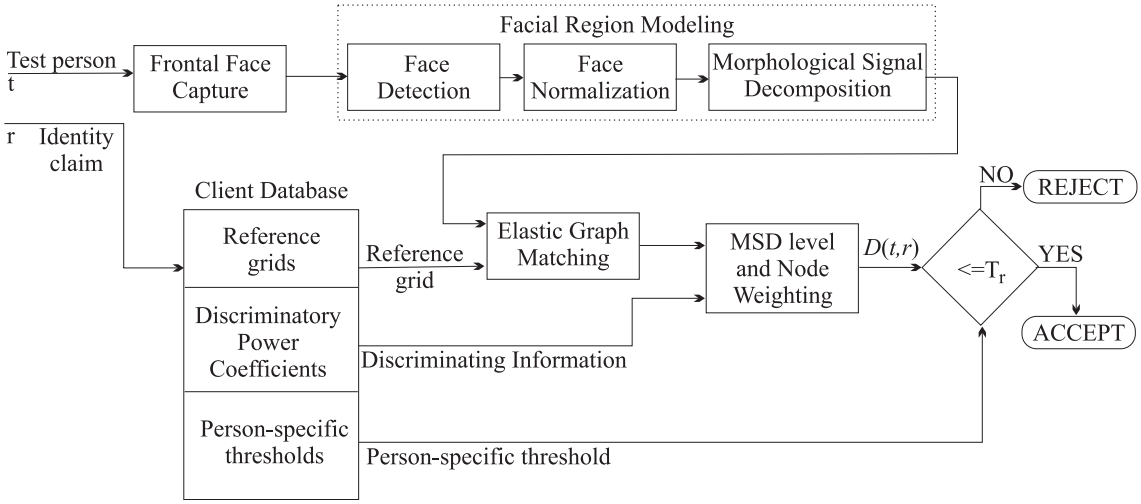

Fig. 2. Block diagram of the test phase of the face verification system.

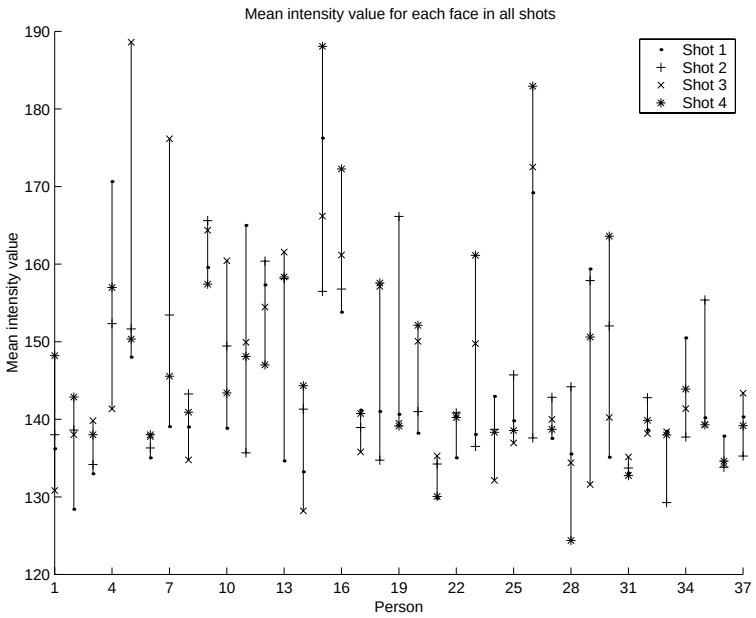

Fig. 3. Mean intensity value of the skin area for each person in the database.

Fig. 4. Compensation for luminance variations: (a),(d): Initial images of the same person from different shots. (b),(e): Segmentation of the original image in three regions using K-means clustering for K=3. (c),(f): Normalized images having the same mean intensity value in the skin region.

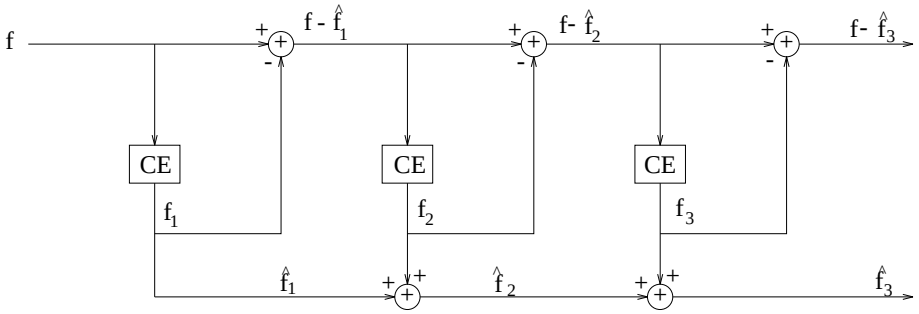

Fig. 5. Block diagram of MSD.

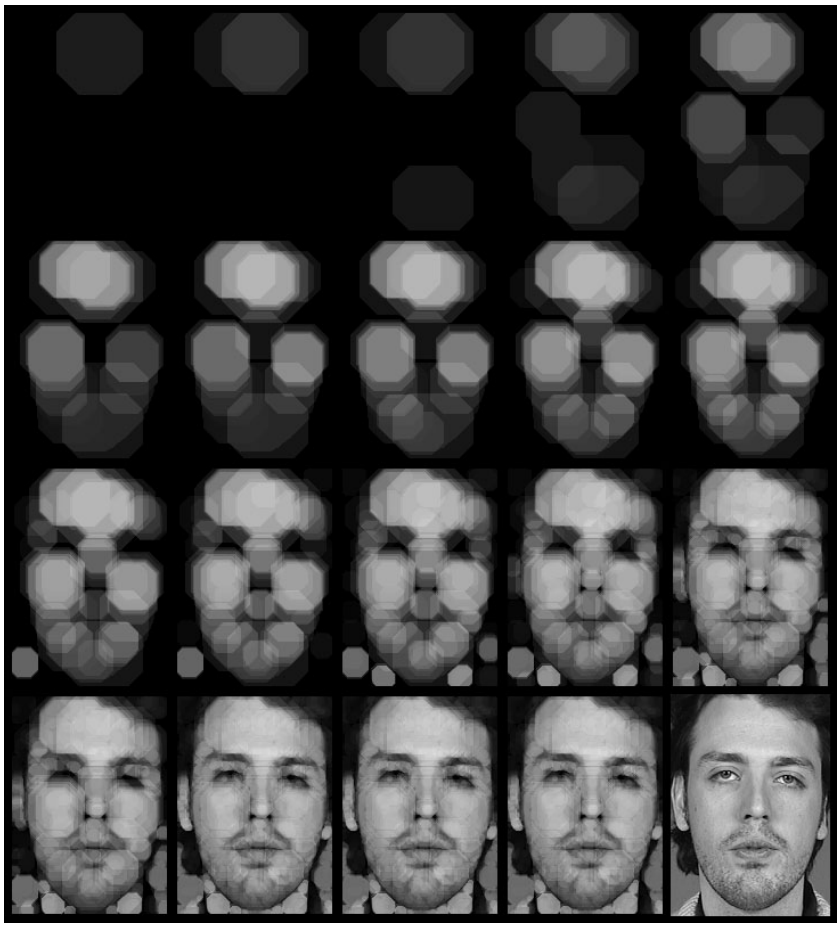

Fig. 6. Reconstructed images at nineteen decomposition levels. The image at the bottom right is the original one.

Fig. 7. Morphological Signal Decomposition residual.

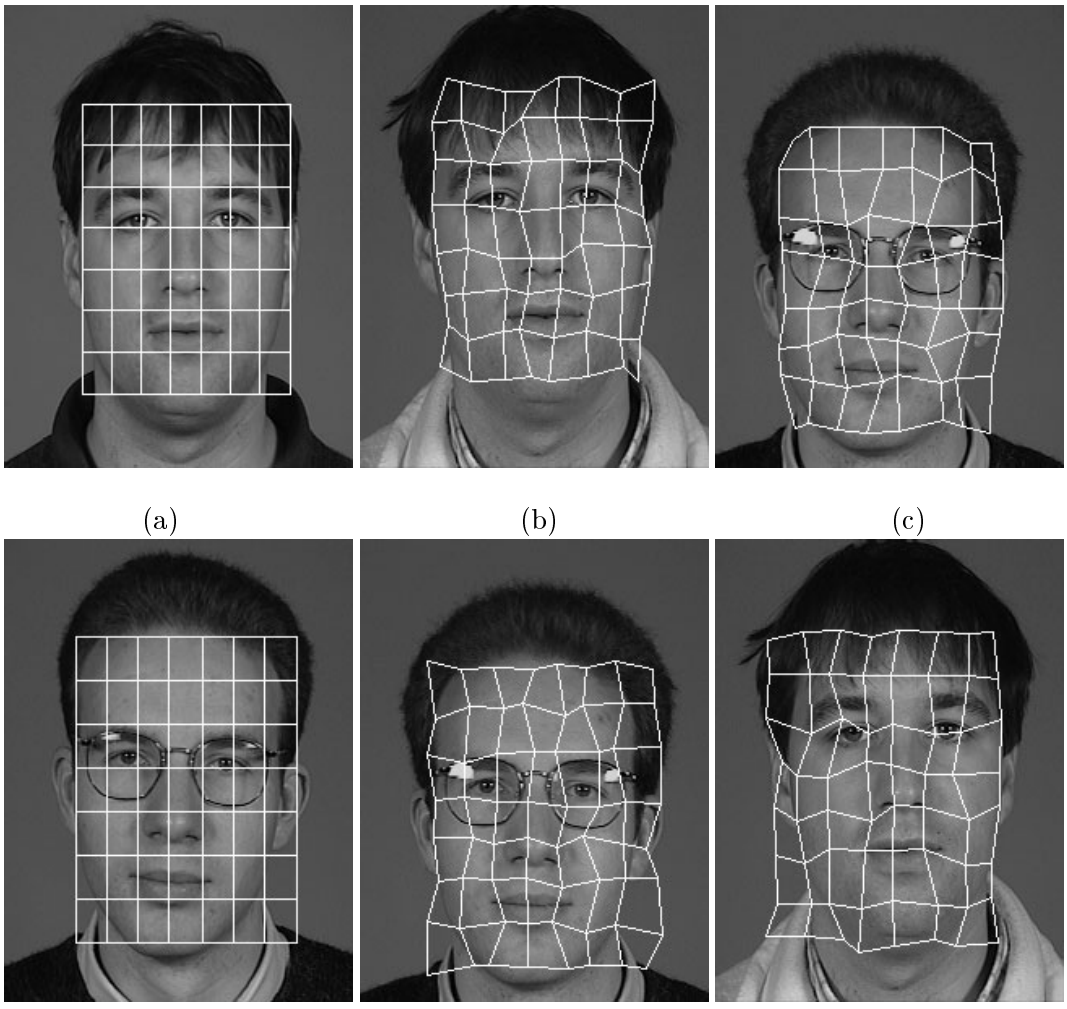

Fig. 8. The graph matching procedure in MSD-DLA: model grid, best grid for the test person after translation and deformation of the grid. Figures (a),(d): Reference person. Figures (b),(e): The test person is identical to the reference one. Figures (c),(f): The test person is different from the reference one.

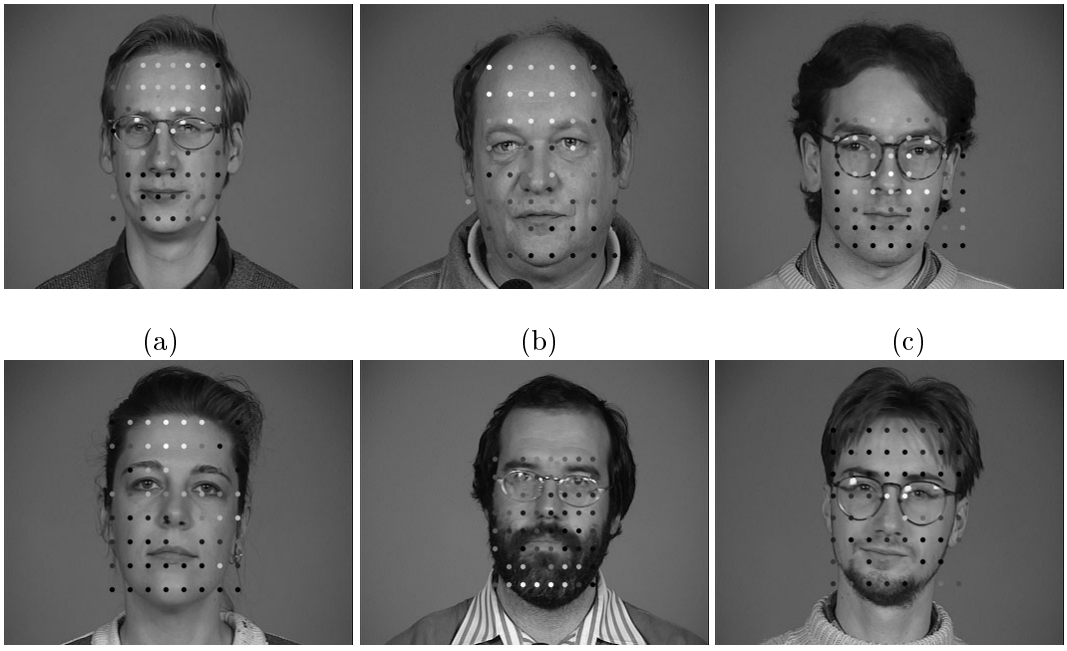

Fig. 9. Discriminatory power coefficients for the grid nodes in Dynamic Link Architecture. The brighter a node is the bigger discriminatory power possesses. The intensity of the nodes is normalized for visualization reasons.

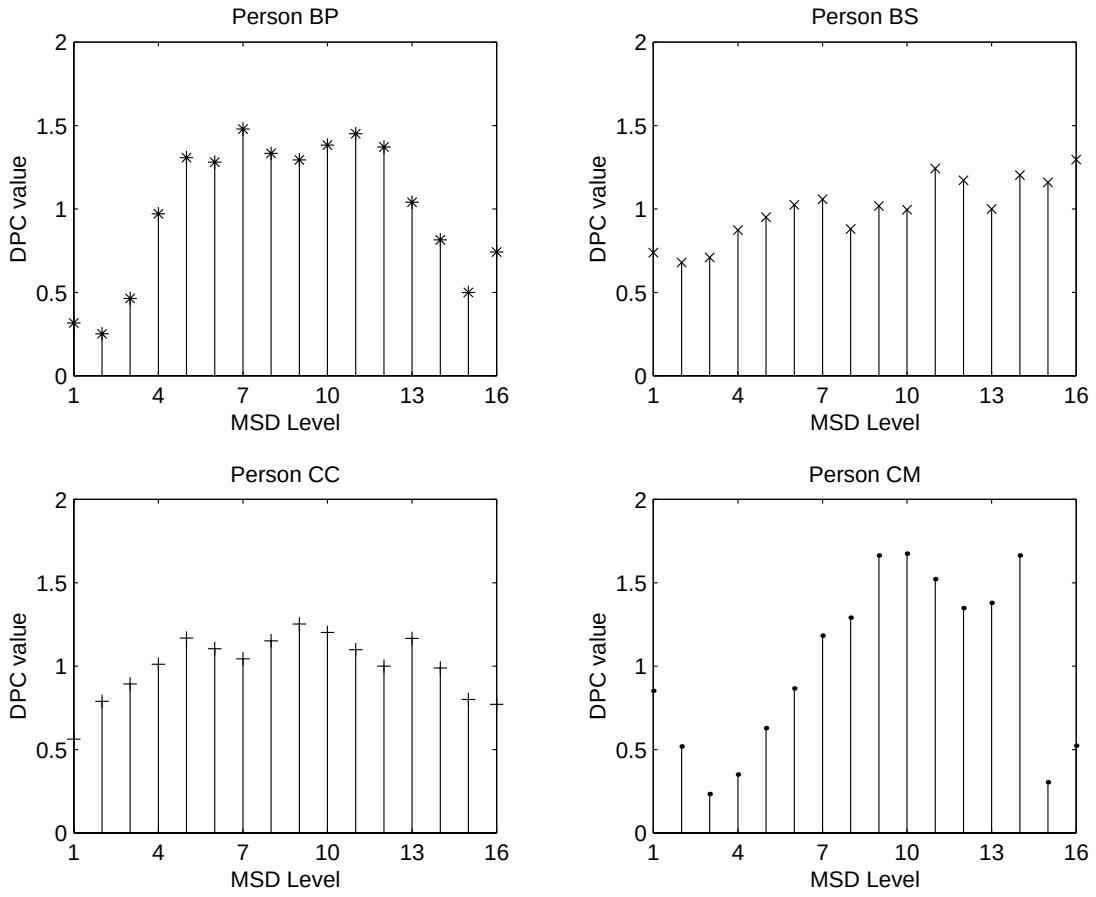

Fig. 10. Discriminatory power coefficients of the morphological signal decomposition levels for several persons in the database.

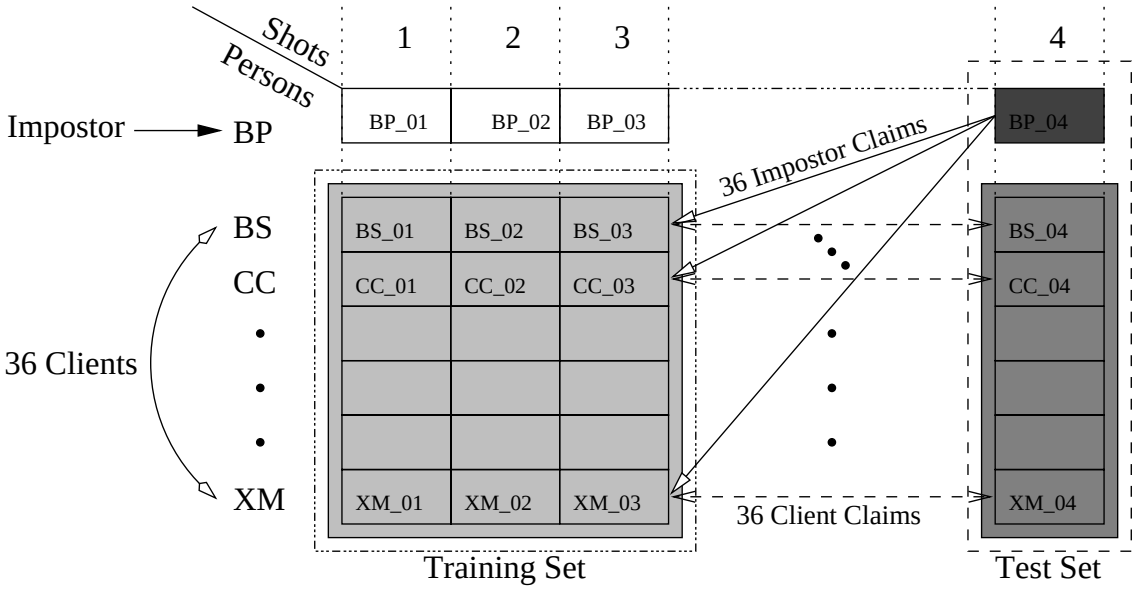

Fig. 11. Experimental Protocol.

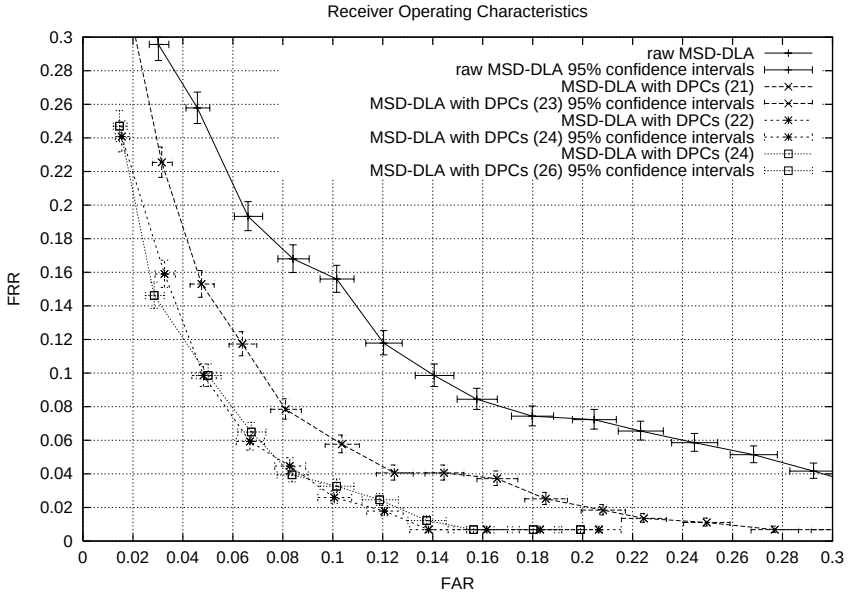

Fig. 12. Receiver Operating Characteristic curves of MSD-DLA where the contribution of each grid node is weighted by several choices of discriminatory power coefficients, (DPCs).

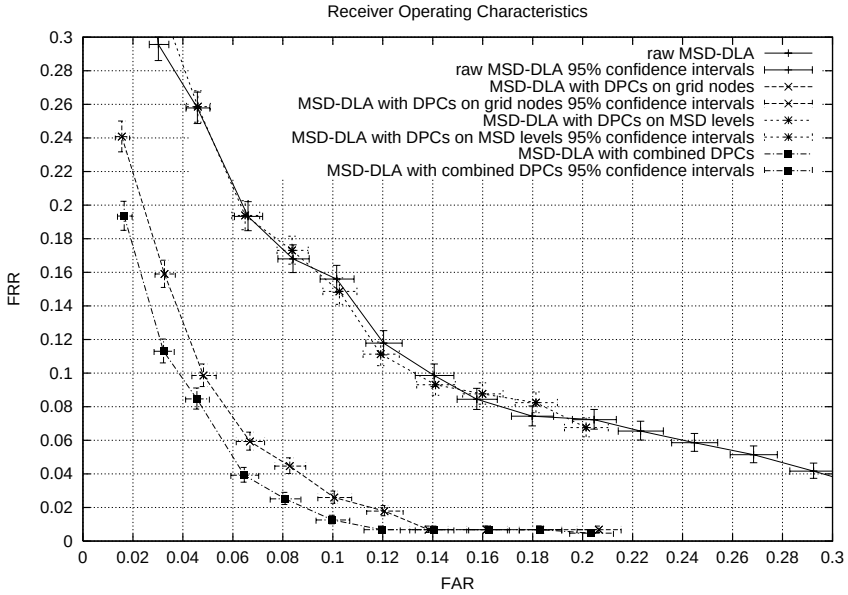

Fig. 13. Receiver Operating Characteristic curves of raw MSD-DLA, MSD-DLA when Discriminatory Power Coefficients are employed for Morphological Signal Decomposition levels only, MSD-DLA when Discriminatory Power Coefficients are employed for grid nodes only, MSD-DLA when both the contributions of Morphological Signal Decomposition levels and grid nodes are weighted.