Variants of Dynamic Link Architecture based on mathematical morphology for frontal face authentication

A. Tefas C. Kotropoulos I. Pitas
Department of Informatics
Aristotle University of Thessaloniki
Box 451, Thessaloniki 540 06, GREECE
email:{tefas,costas,pitas}@zeus.csd.auth.gr

Abstract

Two novel variants of Dynamic Link Architecture that are based on mathematical morphology and incorporate coefficients which weigh the contribution of each node in elastic graph matching according to its discriminatory power are developed. They are the so called Morphological Dynamic Link Architecture and the Morphological Signal Decomposition-Dynamic Link Architecture. The proposed variants are tested for face authentication in a cooperative scenario where the candidates claim an identity to be checked. Their performance is evaluated in terms of their Receiver Operating Characteristic and the Equal Error Rate achieved in M2VTS database. An Equal Error Rate in the range 3.7-6.8% is reported.

1 Introduction

Face recognition has exhibited a tremendous growth for more than two decades. A critical survey of the literature on human and machine face recognition can be found in [1]. An approach that exploits both sources of information, that is, the grey-level information and shape information, is the so-called *Dynamic Link Architecture* (DLA) [8]. The principles of this pattern recognition scheme can be traced back to the origins of self-organization in neural networks. The algorithm is split in two phases, i.e., the training and the recall phase. In the training phase, the objective is to build a sparse grid for each person included in the reference set. Towards this goal a sparse grid is overlaid on the facial region of a person's digital image and the response of a set of 2-D Gabor filters tuned to different orientations and scales is measured at the grid nodes. The responses of Gabor filters form a *feature vector* at each node. In the recall phase, the reference grid of each person is overlaid on the face image of a test person and is deformed so that a cost function is minimized. The research on elastic graph matching and its applications has been an active research topic. A different topology cost for a particular pair of nodes has been proposed in [9]. It is based on the radius of the Apollonius sphere defined by the Euclidean distances between the nodes being matched. This scheme is further used in [16].

A problem in elastic graph matching that has received much attention is the weighting of graph nodes according to their discriminatory power. Several methods have been proposed in the literature. For example, a Bayesian approach yields the more reliable nodes for gender identification, beard and glass detection in bunch graphs [15]. An automatic weighting of the nodes according to their significance by employing local discriminants is proposed in [2]. 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 [7].

In this paper, we propose variants of DLA that are based on mathematical morphology and incorporate local coefficients that weigh the contribution of each grid node according to its discriminatory power. More specifically, we develop first a variant of DLA that is based on multiscale morphological dilation-erosion, the so-called Morphological Dynamic Link Architecture (MDLA), and we incorporate linear projections of the feature vectors (i.e., Principal Component Analysis and Linear Discriminant Analysis) in this variant. Second, we propose a DLA variant that is based on morphological shape decomposition, the so-called Morphological Signal Decomposition-DLA (MSD-DLA), which employs local discriminatory power coefficients aiming at separating more efficiently the intra-class distances from the inter-class ones as well. In both cases, a two-class classification problem is considered. That is, we are seeking methods to separate more efficiently feature vectors extracted from frontal facial images of the same person (i.e., the clients) and the ones extracted from frontal facial images of the remaining persons in a database (i.e., the impostors). A comparative study of the verification capability of the proposed methods in M2VTS database [10] is also undertaken. The performance of the algorithms is evaluated in terms of their Receiver Operating Characteristic (ROC) and the equal error rate (EER) achieved in the M2VTS database. It is demonstrated that by using linear projections of feature vectors in MDLA an EER of 5.4% is achieved improving a previously reported EER [6] by 4%. Moreover, by weighting the matching errors at each node of MDLA using the proposed discriminatory power coefficients an EER of 3.7% is obtained which is 5.65% less than the EER reported in [6]. The latter technique applied within MSD-DLA yields an EER of 6.6%-6.8% reducing the initially achieved EER (without node weighting) by 5%.

2 Linear projections in morphological dynamic link matching

An alternative to linear techniques for generating an information pyramid is the scale-space morphological techniques. In the following, a brief description of MDLA is given and the incorporation of linear projections is explained. In MDLA, we substitute the Gabor-based feature vectors by the *multiscale morphological dilation-erosion* [4]. The multiscale morphological dilation-erosion is based on the two fundamental operations of the grayscale morphology, namely, the *dilation* and the *erosion*. Let $\mathcal{R}$ and $\mathcal{Z}$ denote the set of real and integer numbers, respectively. Given an image $f(\mathbf{x}) : \mathcal{D} \subseteq \mathcal{Z}^2 \rightarrow \mathcal{R}$ and a structuring function $g(\mathbf{x}) : \mathcal{G} \subseteq \mathcal{Z}^2 \rightarrow \mathcal{R}$, the dilation of the image $f(\mathbf{x})$ by $g(\mathbf{x})$ is denoted by $(f \oplus g)(\mathbf{x})$. Its complementary the erosion is denoted by $(f \ominus g)(\mathbf{x})$. Their definitions can be found in any book on Digital Image Processing. The *scaled hemisphere* is employed as a structuring function [4]. The multiscale dilation-erosion of the image $f(\mathbf{x})$ by $g_\sigma(\mathbf{x})$ is defined by [4]:

$$(f \star g_\sigma)(\mathbf{x}) = \begin{cases} (f \oplus g_\sigma)(\mathbf{x}) & \text{if } \sigma > 0 \\ f(\mathbf{x}) & \text{if } \sigma = 0 \\ (f \ominus g_\sigma)(\mathbf{x}) & \text{if } \sigma < 0. \end{cases} \quad (1)$$

The outputs of multiscale dilation-erosion for $\sigma = -9, \ldots, 9$ form the feature vector located at the grid node $\mathbf{x}$:

$$\mathbf{j}(\mathbf{x}) = ((f \star g_9)(\mathbf{x}), \ldots, f(\mathbf{x}), \ldots, (f \star g_{-9})(\mathbf{x})). \quad (2)$$

An 8×8 sparse grid has been created by measuring the feature vectors $\mathbf{j}(\mathbf{x})$ at equally spaced nodes over the output of the face detection algorithm described in [5]. $\mathbf{j}(\mathbf{x})$ has been demonstrated that captures important information for the key facial features [6].

Frequently, linear projection algorithms are used to reduce the dimensionality of the feature vectors. Two are the most popular linear projection algorithms: The *Karhunen-Loeve or Principal Component Analysis* (PCA) that does not employ category information and the *linear discriminant analysis* (LDA) that exploits category labels.

Representations based on PCA have been studied and extensively used for various applications. Among others PCA has been applied to face recognition e.g. [14]. For a detailed list of applications the interested reader is referred to [13, 3]. It is well known that PCA aims at reducing the dimensionality of the original feature vectors so that the new vectors after this projection approximate in the sense of minimum mean squared error the original ones. PCA methods have shown good performance in image reconstruction/compression tasks. Accordingly, the feature vectors produced are called *most expressive features* (MEFs) [13]. In addition to dimensionality reduction PCA decorrelates the feature vectors and facilitates the LDA that is applied next in eigenvalue/eigenvector computations as well as in matrix inversion. Although MDLA has originally been applied to 4 sets of 37 frontal images, one for each person in M2VTS database [10], we now need much more frontal images. We have either extracted other frontal images for each person or we have augmented the original frontal images with others produced by slightly adding Gaussian noise as is proposed in [3]. Let $\mathbf{j}'(\mathbf{x}_l) = \mathbf{j}(\mathbf{x}_l) - \mathbf{m}(\mathbf{x}_l)$ be the normalized feature vector at node $\mathbf{x}_l$ where $\mathbf{j}(\mathbf{x}_l) = (j_1(\mathbf{x}_l), \ldots, j_{19}(\mathbf{x}_l))^T$ and $\mathbf{m}(\mathbf{x}_l)$ is the mean feature vector at $\mathbf{x}_l$. Let N denote the total number of frontal images extracted from a database for all persons. Let also $\mathbf{\Gamma}(\mathbf{x}_l)$ be the covariance matrix of the feature vectors $\mathbf{j}'(\mathbf{x}_l)$ at node $\mathbf{x}_l$. In PCA we compute the eigenvectors that correspond to the p largest eigenvalues of $\mathbf{\Gamma}(\mathbf{x}_l)$, say $\mathbf{e}_1(\mathbf{x}_l), \ldots, \mathbf{e}_p(\mathbf{x}_l)$. The PCA projected feature vector is given by:

$$\tilde{\mathbf{j}}(\mathbf{x}_l) = \begin{bmatrix} \mathbf{e}_1^T(\mathbf{x}_l) \\ \vdots \\ \mathbf{e}_p^T(\mathbf{x}_l) \end{bmatrix} \mathbf{j}'(\mathbf{x}_l) = \mathbf{P}(\mathbf{x}_l)\mathbf{j}'(\mathbf{x}_l) \quad (3)$$

where T denotes the transposition operator. $\tilde{\mathbf{j}}(\mathbf{x}_l)$ is of dimensions $p \times 1$, $p \leq 19$.

Next LDA is applied to feature vectors produced by PCA. It is well known that there is no guarantee that the MEFs are necessarily good for discriminating among classes defined by a set of samples

[13, 3]. It is well known that optimality in discrimination among all possible linear combinations of features can be achieved by employing Linear Discriminant Analysis (LDA). The feature vectors produced after the LDA projection are called *most discriminating features* (MDFs) [13]. We are interested in applying the LDA at each grid node locally. In the following, the explicit dependence on $\mathbf{x}$ is omitted for notation simplicity. Let $\mathcal{S}$ be the entire set of feature vectors at a grid node and $\mathcal{S}_k$ be the corresponding set of features vectors at the same node extracted from the frontal facial images of the k-th person in the database. Our local LDA scheme determines a weighting matrix $(d \times p)$ $\mathbf{V}_k$ for the k-th person defined by:

$$\mathbf{V}_k = \begin{bmatrix} \mathbf{v}_{1k} \\ \vdots \\ \mathbf{v}_{dk} \end{bmatrix} \tag{4}$$

such that the ratio:

$$\begin{aligned} \mathcal{M}_k &= \frac{\operatorname{tr}\left[\mathbf{V}_k \left\{\sum_{\tilde{\mathbf{j}} \in \mathcal{S}_k} (\tilde{\mathbf{j}} - \tilde{\mathbf{m}}_k)(\tilde{\mathbf{j}} - \tilde{\mathbf{m}}_k)^T\right\} \mathbf{V}_k^T\right]}{\operatorname{tr}\left[\mathbf{V}_k \left\{\sum_{\tilde{\mathbf{j}} \in (\mathcal{S}-\mathcal{S}_k)} (\tilde{\mathbf{j}} - \tilde{\mathbf{m}}_k)(\tilde{\mathbf{j}} - \tilde{\mathbf{m}}_k)^T\right\} \mathbf{V}_k^T\right]} \\ &= \frac{\operatorname{tr}\left[\mathbf{V}_k \mathbf{W}_k \mathbf{V}_k^T\right]}{\operatorname{tr}\left[\mathbf{V}_k \mathbf{B}_k \mathbf{V}_k^T\right]} \end{aligned} \tag{5}$$

is minimized where $\tilde{\mathbf{m}}_k$ is the class-dependent mean vector of the feature vectors which result after PCA. This is a generalized eigenvalue problem. Its solution, (i.e., the vectors $\mathbf{v}_{ik}$, $i = 1, \ldots, d$) are given by the eigenvectors that correspond to the d smallest in magnitude eigenvalues of $\mathbf{B}_k^{-1}\mathbf{W}_k$ or equivalently by the eigenvectors that correspond to the d largest in magnitude eigenvalues of $\mathbf{W}_k^{-1}\mathbf{B}_k$ provided that both $\mathbf{W}_k$ and $\mathbf{B}_k$ are invertible. We shall confine ourselves to the case $d = 2$, where only two MDFs are used. Because the matrix $\mathbf{W}_k^{-1}\mathbf{B}_k$ is not symmetric in general, the eigenvalue problem could be computationally unstable. A very elegant method that diagonalises the two symmetric matrices $\mathbf{W}_k$ and $\mathbf{B}_k$ and yields a stable computation procedure for the solution of the generalized eigenvalue problem has been proposed in [12]. This method has been used to solve the generalized eigenvalue problem.

Let the superscripts t and r denote a test and a reference person (or grid), respectively. Let us also denote by $\mathbf{x}_l$ the l-th grid node. Having found the weighting matrix $\mathbf{V}_k(\mathbf{x}_l)$, for the l-th node of the k-th person in the database, we project the reference feature vector after PCA at this node onto the plane defined by $\mathbf{v}_{1k}(\mathbf{x}_l)$ and $\mathbf{v}_{2k}(\mathbf{x}_l)$ as follows:

$$\breve{j}(\mathbf{x}_l^r) = \mathbf{V}_k \left[\mathbf{P}(\mathbf{x}_l^r)\left(\mathbf{j}(\mathbf{x}_l^r) - \mathbf{m}_l\right) - \tilde{\mathbf{m}}_{kl}\right]. \tag{6}$$

Let us suppose that a test person claims the identity of the k-th person. Then the test MDF vector at the l-th node can be derived as in (6). The L_2 norm of the difference between the MDF vectors at the l-th node has been used as a (signal) similarity measure, i.e.:

$$C_v(\breve{j}(\mathbf{x}_l^t), \breve{j}(\mathbf{x}_l^r)) = \|\breve{j}(\mathbf{x}_l^t) - \breve{j}(\mathbf{x}_l^r)\|. \tag{7}$$

Let us denote by $\mathcal{V}$ the set of grid nodes. The grid nodes are simply the vertices of a graph. Let also $\mathcal{N}(l)$ denote the four-connected neighbourhood of vertex l. The objective is to find the set of test grid node coordinates $\{\mathbf{x}_l^t,\ l \in \mathcal{V}\}$ that yields the best matching. As in DLA [8], the quality of the match is evaluated by taking into account the grid deformations as well. Grid deformations can be penalized using the additional cost function:

$$C_e(l,\xi) = C_e(\mathbf{d}_{l\,\xi}^t, \mathbf{d}_{l\,\xi}^r) = \|\mathbf{d}_{l\,\xi}^t - \mathbf{d}_{l\,\xi}^r\| \quad \xi \in \mathcal{N}(l) \tag{8}$$

with $\mathbf{d}_{l\,\xi} = (\mathbf{x}_l - \mathbf{x}_\xi)$. The penalty (8) can be incorporated to a cost function:

$$C(\{\mathbf{x}_l^t\}) = \sum_{l \in \mathcal{V}} \left\{ C_v(\breve{j}(\mathbf{x}_l^t), \breve{j}(\mathbf{x}_l^r)) + \lambda \sum_{\xi \in \mathcal{N}(l)} C_e(l,\xi) \right\}. \tag{9}$$

One may interpret (9) as a simulated annealing with an additional penalty (i.e., a constraint on the objective function). Since the cost function (8) does not penalize translations of the whole graph the random configuration $\mathbf{x}_l$ can be of the form of a random translation $\mathbf{s}$ of the (undeformed) reference grid and a bounded local perturbation $\mathbf{q}_l$, i.e.:

$$\mathbf{x}_l^t = \mathbf{x}_l^r + \mathbf{s} + \mathbf{q}_l \quad ; \quad \|\mathbf{q}_l\| \le q_{\max} \tag{10}$$

where the choice of $q_{\max}$ controls the rigidity/plasticity of the graph. It is evident that the proposed approach differs from the two stage coarse-to-fine optimization procedure proposed in [8]. In our approach we replace the two stage optimization procedure with a probabilistic hill climbing algorithm which attempts to find the best configuration $\{\mathbf{s}, \{\mathbf{q}_l\}\}$ at each step.

3 Incorporating discriminatory power coefficients in Morphological Signal Decomposition-Dynamic Link Architecture

The modeling of a gray-scale facial image region by employing morphological shape decomposition (MSD) is described in this section. Let us denote by $f(\mathbf{x}) : \mathcal{D} \subseteq \mathcal{Z}^2 \rightarrow \mathcal{Z}$ the facial image region that can be extracted by using a face detection module such as the one proposed in [5]. Without any loss of generality it is assumed that the image pixel values are non-negative, i.e., $f(\mathbf{x}) \geq 0$. Let $g(\mathbf{x}) = 1$, $\forall \mathbf{x} : \|\mathbf{x}\| \leq \sigma$ denote the structuring function. The value $\sigma = 2$ has been used in all experiments. Symmetric operators will not explicitly denoted hereafter. Given $f(\mathbf{x})$ and $g(\mathbf{x})$, the objective of shape decomposition is to decompose $f(\mathbf{x})$ into a sum of components, i.e.:

$$f(\mathbf{x}) = \sum_{i=1}^{K} f_i(\mathbf{x}) \tag{11}$$

where $f_i(\mathbf{x})$ denotes the i-th component. The i-th component can be expressed as:

$$f_i(\mathbf{x}) = [h_i \oplus n_i\, g]\,(\mathbf{x}) \tag{12}$$

where $h_i(\mathbf{x})$ is the so-called *spine* and

$$n_i\, g(\mathbf{x}) = \underbrace{[g \oplus g \oplus \cdots \oplus g]}_{n_i \text{ times}}(\mathbf{x}). \tag{13}$$

An intuitively sound choice for $n_1\, g(\mathbf{x})$ is the maximal function in $f(\mathbf{x})$, that is, to choose n_1 such that:

$$[f \ominus (n_1 + 1)g]\,(\mathbf{x}) \leq 0 \quad \forall x \in \mathcal{D}. \tag{14}$$

Accordingly, the first spine is given by:

$$h_1(\mathbf{x}) = [f \ominus n_1\, g]\,(\mathbf{x}). \tag{15}$$

Morphological shape decomposition can then be implemented recursively as follows.

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

Step 2. i-th level of decomposition: Starting with $n_i = 1$ increment n_i until

$$\left[(f - \hat{f}_{i-1}) \ominus (n_i + 1)g\right](\mathbf{x}) \leq 0. \tag{16}$$

Step 3. Calculate the i-th component by:

$$f_i(\mathbf{x}) = \left\{ \underbrace{\left[(f - \hat{f}_{i-1}) \ominus n_i\, g\right]}_{h_i(\mathbf{x})} \oplus n_i\, g \right\}(\mathbf{x}). \tag{17}$$

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

$$\hat{f}_i(\mathbf{x}) = \hat{f}_{i-1}(\mathbf{x}) + f_i(\mathbf{x}). \tag{18}$$

Step 5. Let $\mathcal{A}(f - \hat{f}_i)$ 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. Increment i and go to Step 2 until $i > K$ or $\mathcal{A}(f - \hat{f}_{i-1})$ is sufficiently small.

We propose a dynamic link matching with feature vectors that are extracted from the reconstructed images $\hat{f}_i(\mathbf{x})$ at the last K successive levels of decomposition $i = L - K, \ldots, L$ where L denotes the maximal number of decomposition levels. A successful choice of K found by experiments is $K = 15$. That is, the grey level information $\hat{f}_i$ at the node $\mathbf{x}$ of the sparse grid for the levels of decomposition $i = L - 15, \ldots, L$ along with the grey level information f is concatenated to form the feature vector $\mathbf{j}(\mathbf{x})$, the so-called jet [8]:

$$\mathbf{j}(\mathbf{x}) = \left(f(\mathbf{x}), \hat{f}_{L-15}(\mathbf{x}), \ldots, \hat{f}_L(\mathbf{x})\right). \tag{19}$$

Accordingly, the variant of DLA that results is termed Morphological Shape Decomposition-Dynamic Link Architecture.

It is well known that some facial features (e.g. the eyes, the nose) are more distinctive for certain persons. Therefore they play a more crucial role in the verification procedure than other features. It would be helpful if we may calculate a weighting coefficient that enables discriminating among feature vectors extracted from frontal facial images of the same person (i.e., the clients) and the ones extracted from frontal facial images of the remaining persons in a database (i.e., the impostors). To do so, we would like to weigh the signal similarity measure at node l given by:

$$C_v(\mathbf{j}(\mathbf{x}_l^t), \mathbf{j}(\mathbf{x}_l^r)) = \|\mathbf{j}(\mathbf{x}_l^t) - \mathbf{j}(\mathbf{x}_l^r)\| \tag{20}$$

using class-dependent discriminatory power coefficients (DPCs) $DP_l(\mathcal{S}_r)$ so that when person t claims the identity of person r a distance measure between them is computed by:

$$D(t, r) = \sum_{l \in \mathcal{V}} \frac{DP_l(\mathcal{S}_r)\; C_v(\mathbf{j}(\mathbf{x}_l^t), \mathbf{j}(\mathbf{x}_l^r))}{\sum_{n \in \mathcal{V}} DP_n(\mathcal{S}_r)} \tag{21}$$

with $\mathcal{S}_r$ denoting the class of the reference person r. Let $m_{\text{intra}}(\mathcal{S}_r, l)$ be the mean intra-class distance for class $\mathcal{S}_R$ 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 grid node l:

$$\begin{aligned} m_{\text{intra}} &= E\left\{C_v(\mathbf{j}(\mathbf{x}_l^t), \mathbf{j}(\mathbf{x}_l^r))\right\} \;\forall t,\, r \in \mathcal{S}_r \qquad (22) \\ m_{\text{inter}} &= E\left\{C_v(\mathbf{j}(\mathbf{x}_l^t), \mathbf{j}(\mathbf{x}_l^r))\right\} \;\forall r \in \mathcal{S}_r,\, t \in (\mathcal{S} - \mathcal{S}_r) \end{aligned}$$

where $\mathcal{S}$ denotes the set of all classes in the database. Let $\text{var}_{\text{intra}}(\mathcal{S}_r, l)$ and $\text{var}_{\text{inter}}(\mathcal{S}_r, l)$ be the variances of the intra-class distances and the inter-class distances, given by (20), respectively. A plausible measure of the discriminatory power of the grid node l for the class $\mathcal{S}_r$ is the *Fisher's Linear Discriminant* (FLD) [11] function or first canonical variate that takes under consideration both the difference between the two class-dependent mean distances and the distance variances in order to yield a DPC for the grid node l:

$$DP_l(\mathcal{S}_r) = \frac{(m_{\text{inter}}(\mathcal{S}_r, l) - m_{\text{intra}}(\mathcal{S}_r, l))^2}{\text{var}_{\text{inter}}(\mathcal{S}_r, l) + \text{var}_{\text{intra}}(\mathcal{S}_r, l)}. \quad (23)$$

We can see that in (23) the $DP_l(\mathcal{S}_r)$ is maximized when the denominator $\text{var}_{\text{inter}}(\mathcal{S}_r, l) + \text{var}_{\text{intra}}(\mathcal{S}_r, l)$ is minimized. This can be interpreted as an **AND** rule for the variances of the clusters. Alternatively, one can use a more relaxed criterion of the form:

$$DP_l(\mathcal{S}_r) = \frac{(m_{\text{inter}}(\mathcal{S}_r, l) - m_{\text{intra}}(\mathcal{S}_r, l))^2}{\sqrt{\text{var}_{\text{inter}}(\mathcal{S}_r, l)\text{var}_{\text{intra}}(\mathcal{S}_r, l)}}. \quad (24)$$

The denominator of (24) is interpreted as an **OR** rule for the variances of the clusters. Figure 1 depicts the grids formed in the procedure of matching when $DP_l(\mathcal{S}_r)$ given by (24) are employed.

4 Performance evaluation of the combined schemes

The combined schemes of MDLA/MSD-DLA with linear projections and with discriminatory power coefficients have been tested on the M2VTS database [10]. The database contains both sound and image information. Four recordings (i.e., shots) of the 37 persons have been collected. 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. In the authentication experiments we use only one frontal image from the image sequence of each person that has been chosen based on symmetry considerations. Four experimental sessions have been implemented by employing the "leave one out" principle. Details on the experimental protocol used in the performance evaluation as well as on the computation of thresholds that discriminate each person from the remaining persons in the database can be found in [6]. We may create a plot of False Rejection Rate (FRR) versus the False Acceptance Rate (FAR) with the varying thresholds as an implicit parameters. This plot is the *Receiver Operating Characteristic* (ROC) of the verification technique. The ROCs of the MDLA with and without linear projections or discriminatory power coefficients are plotted in Figure 2. In the same plot the ROCs of MSD-DLA with and without discriminatory power coefficients are also depicted. The Equal Error Rate (EER) of a technique (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 EER of MDLA with one MDF is 6.8% and with two MDF is 5.4% whereas the EER of MDLA without any linear projections is 9.35 % [6]. It is seen that the incorporation of linear projections improves the EER by 2.55-4%. Moreover, the EER of MDLA using the discriminatory power coefficients (23) is found 3.7%. It is worth noting that the EER of MSD-DLA without local discriminatory power coefficients is 11.89 %. By using this discrimination criterion (23), we achieve an EER of 6.73% following the same experimental protocol. The same figure of merit using the discrimination criterion (24) is found to be 6.58%. That is, a significant drop of 5.3% in EER is reported. The comparison of EERs achieved by the proposed schemes are identical or better than the ones reported in [2], i.e., EER between 5.4 % and 9.2 %.

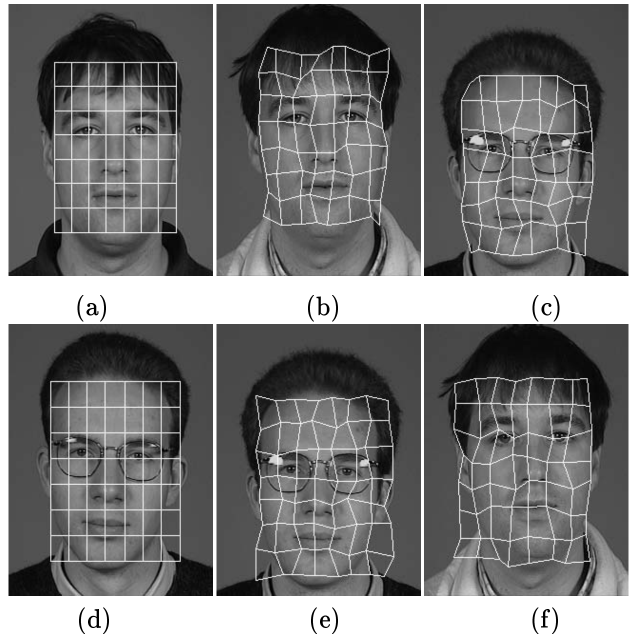

Figure 1: The graph matching procedure in MSD-DLA: model grid, best grid for the test person after translation and deformation of the grid. Figures (a),(d): Reference person. Figures (b),(e): The test person is identical to the reference one (After Linear Discriminant Distance b-a=1665, Distance e-d=1406). Figures (c),(f): The test person is different from the reference one (After Linear Discriminant Distance c-a=5347, Distance f-d=4868).

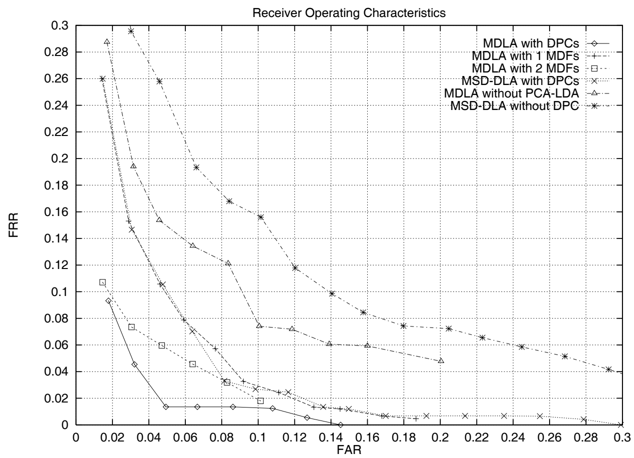

Figure 2: Receiver Operating Characteristics of MDLA and MSD-DLA with/without linear projections or local discriminatory power coefficients.

Acknowledgments

The research reported in this paper has been carried out within the framework of the European ACTS-M2VTS project.

References

[1] R. Chellapa, 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] B. Duc, S. Fischer, and J. Bigün, "Face authentication with Gabor information on deformable graphs," *IEEE Trans. on Image Processing*, submitted 1997.

[3] K. Etemad, and R. Chellappa, "Discriminant Analysis for Recognition of Human Face Images," in *Lecture Notes in Computer Science: Audio- and Video- based Biometric Person Authentication* (J. Bigün, G. Chollet and G. Borgefors, Eds.), vol. 1206, pp. 127–142, 1997.

[4] P.T. Jackway, and M. Deriche, "Scale-space properties of the multiscale morphological dilation-erosion," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 18, no. 1, pp. 38–51, January 1996.

[5] C. Kotropoulos, and I. Pitas, "Rule-based face detection in frontal views," in *Proc. of IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP 97)*, vol. IV, pp. 2537–2540, Munich, Germany, April 21-24, 1997.

[6] C. Kotropoulos, and I. Pitas, "Face authentication based on morphological grid matching," in *Proc. of the IEEE Int. Conf. on Image Processing (ICIP 97)*, pp. I-105–I-108, Santa Barbara, California, U.S.A., 1997.

[7] 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.

[8] 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.

[9] B.S. Manjunath, R. Chellapa, 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*, pp. 373–378, 1992.

[10] S. Pigeon, and L. Vandendorpe, "The M2VTS multimodal face database," in *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.

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

[12] D.L. Swets, and J. Weng, "Using discriminant eigenfeatures for image retrieval," *IEEE Trans. on Pattern Analysis and Machine Intelligence* vol. 18, no. 8, pp. 831–837, August 1996.

[13] D.L. Swets, and J. Weng, "Discriminant analysis and eigenspace partition tree for face and object recognition from views," in *Proc. of the IEEE Int. Conf. on Automatic Face and Gesture Recognition*, pp.192–197, Killington, Vermont, U.S.A., October 14-16, 1996.

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

[15] L. Wiskott, "Phantom faces for face analysis," in *Lecture Notes in Computer Science:Computer Analysis of Images and Patterns* (G.Sommer, K. Daniilidis, J. Pauli, Eds.), vol. 1296, pp. 480–487, 1997.

[16] X. Wu, and B. Bhanu, "Gabor wavelet representation for 3-D object recognition," *IEEE Trans. on Image Processing*, vol. 6, no. 1, pp. 47–64, January 1997.