FACE AUTHENTICATION USING VARIANTS OF ELASTIC GRAPH MATCHING BASED ON MATHEMATICAL MORPHOLOGY THAT INCORPORATE LOCAL DISCRIMINANT COEFFICIENTS

Constantine Kotropoulos and Anastasios Tefas and Ioannis Pitas

Dept. of Informatics, Aristotle University of Thessaloniki
Box 451, Thessaloniki 540 06, GREECE
e-mail: {costas, tefas, pitas }@zeus.csd.auth.gr

ABSTRACT

Two novel variants of Dynamic Link Architecture that are based on mathematical morphology and incorporate local coefficients which weigh the contribution of each node according to its discriminatory power in elastic graph matching are proposed, namely, the Morphological Dynamic Link Architecture and the Morphological Signal Decomposition-Dynamic Link Architecture. They 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 of 6.6% - 6.8 % is reported.

1. INTRODUCTION

Face recognition has exhibited a tremendous growth for more than two decades. A critical survey of the literature related to human and machine face recognition are 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) [7]. The principles of this pattern recognition scheme can be traced back to the origins of self-organisation 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 2D 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 minimised.

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 [11]. 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 [6].

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 propose first a variant of DLA that is based on multi-scale morphological dilation-erosion, the so-called Morphological Dynamic Link Architecture (MDLA), and we incorporate in this variant linear projections of the feature vectors (i.e., Principal Component Analysis and Linear Discriminant Analysis). Second, we develop a DLA variant that is based on morphological shape decomposition, the so-called Morphological Signal Decomposition-DLA (MSD-DLA), which also employs local discriminatory power coefficients aiming at separating more efficiently the intra-class distances and the inter-class ones. 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 the database (i.e., the impostors). A comparative study of the verification capability of the proposed methods in M2VTS database [8] is also undertaken. The performance of the algorithms is evaluated in terms of their receiver operating characteristic and the equal error rate (EER) achieved in the M2VTS database. It is demonstrated that the combined use of discriminatory power coefficients with DLA variants improves their EER by 2.55% to 5.3%.

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 used in dynamic link matching by the multi-scale morphological dilation-erosion [3]. The multi-scale morphological dilation-erosion is based on the two fundamental operations of the gray-scale morphology, namely the dilation and the

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

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 [3]. The multi-scale dilation-erosion of the image $f(\mathbf{x})$ by $g_\sigma(\mathbf{x})$ is defined by [3]:

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

The outputs of multi-scale 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 [4]. $\mathbf{j}(\mathbf{x})$ has been demonstrated that captures important information for the key facial features [5].

Subsequently, a feature vector dimensionality reduction is pursued by employing PCA. 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. Let $\mathbf{j}_l^{'}(\mathbf{x}) = \mathbf{j}_l(\mathbf{x}) - \mathbf{m}(\mathbf{x})$ be the normalised feature vector at node $\mathbf{x}$ where $\mathbf{j}_l(\mathbf{x}) = (j_{l,1}(\mathbf{x}), \ldots, j_{l,19}(\mathbf{x}))^T$ and $\mathbf{m}(\mathbf{x})$ is the mean feature vector at $\mathbf{x}$. Let N denote the total number of frontal images extracted for all persons. Let also $\mathbf{R}(\mathbf{x})$ be the covariance matrix of the feature vectors $\mathbf{j}^{'}(\mathbf{x})$ at node $\mathbf{x}$. In PCA we compute the eigenvectors that correspond to the p largest eigenvalues of $\mathbf{R}(\mathbf{x})$, say $\mathbf{e}_1(\mathbf{x}), \ldots, \mathbf{e}_p(\mathbf{x})$. The PCA projected feature vector is given by:

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

and is of dimension $p \times 1$, $p \leq N$.

Next LDA is applied to feature vectors produced by PCA. It is well known that optimality in discrimination among all possible linear combinations of features can be achieved by employing LDA [10]. 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 vector $\mathbf{v}_k$ for the k-th person such that the ratio:

$$\mathcal{M}_k = \frac{\mathbf{v}_k^T \left[\sum_{\hat{\mathbf{j}} \in \mathcal{S}_k} (\hat{\mathbf{j}} - \hat{\mathbf{m}}_k)(\hat{\mathbf{j}} - \hat{\mathbf{m}}_k)^T\right] \mathbf{v}_k}{\mathbf{v}_k^T \left[\sum_{\hat{\mathbf{j}} \in (\mathcal{S} - \mathcal{S}_k)} (\hat{\mathbf{j}} - \hat{\mathbf{m}}_k)(\hat{\mathbf{j}} - \hat{\mathbf{m}}_k)^T\right] \mathbf{v}_k} = \frac{\mathbf{v}_k^T \mathbf{W}_k \mathbf{v}_k}{\mathbf{v}_k^T \mathbf{B}_k \mathbf{v}_k} \quad (4)$$

is minimised where $\hat{\mathbf{m}}_k$ is the class-dependent mean vector of the feature vectors which result after PCA. This is a generalised eigenvalue problem. Its solution is given by the eigenvector that corresponds to the minimal eigenvalue o $\mathbf{B}_k^{-1}\mathbf{W}_k$ or equivalently by the eigenvector that corresponds to the maximal eigenvalue of $\mathbf{W}_k^{-1}\mathbf{B}_k$ provided that both $\mathbf{W}_k$ and $\mathbf{B}_k$ are invertible. 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 generalised eigenvalue problem has been proposed in [10]. This method has been used to solve the generalised 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 vector $\mathbf{v}_{kl}$ 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 $\mathbf{v}_{kl}$ as follows:

$$\breve{j}(\mathbf{x}_l^r) = \mathbf{v}_{kl}^T \left[\mathbf{P}_l \left(\mathbf{j}(\mathbf{x}_l^r) - \mathbf{m}_l\right) - \hat{\mathbf{m}}_{kl}\right]. \quad (5)$$

It is seen that a scalar reference feature value results after the linear discriminant projection. Let us suppose that a test person claims the identity of the k-th person. Then the test scalar feature value at the l-th node can be derived as in (5). The absolute value of the difference between the scalar feature values at the l-th node has been used as a (signal) similarity measure, i.e.:

$$C_v(\breve{j}(\mathbf{x}_l^t), \breve{j}(\mathbf{x}_l^r)) = |\breve{j}(\mathbf{x}_l^t) - \breve{j}(\mathbf{x}_l^r)| \quad (6)$$

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 [7], the quality of the match is evaluated by taking into account the grid deformations as well. Grid deformations can be penalised using the additional cost function:

$$C_e(i,j) = C_e(\mathbf{d}_{l\,\xi}^t, \mathbf{d}_{l\,\xi}^r) = \|\mathbf{d}_{l\,\xi}^t - \mathbf{d}_{l\,\xi}^r\| \quad \xi \in \mathcal{N}(l) \quad (7)$$

with $\mathbf{d}_{l\,\xi} = (\mathbf{x}_l - \mathbf{x}_\xi)$. The penalty (7) can be incorporated to a cost function:

$$C(\{\mathbf{x}_l^t\}) = \sum_{l \in \mathcal{V}} \left\{ C_v(\breve{j}(\mathbf{x}_l^t), \breve{j}(\mathbf{x}_l^r)) + \lambda \sum_{\xi \in \mathcal{N}(l)} C_e(\mathbf{d}_{l\xi}^t, \mathbf{d}_{l\xi}^r) \right\}. \quad (8)$$

One may interpret (8) as a simulated annealing with an additional penalty (i.e., a constraint on the objective function). Since the cost function (7) does not penalise translations of the whole graph the random configuration $\mathbf{x}_l$ can be of the form of a random translation $\mathbf{d}$ of the (undeformed) reference grid and a bounded local perturbation $\mathbf{n}_l$, i.e.:

$$\mathbf{x}_l^t = \mathbf{x}_l^r + \mathbf{d} + \mathbf{n}_l \quad ; \quad \|\mathbf{n}_l\| \leq \eta\text{max} \quad (9)$$

where the choice of ηmax controls the rigidity/plasticity of the graph. It is evident that the proposed approach differs from the two stage coarse-to-fine optimisation procedure proposed in [7]. In our approach we replace the two stage optimisation procedure with a probabilistic hill climbing algorithm which attempts to find the best configuration $\{\mathbf{d}, \{\mathbf{n}_l\}\}$ 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 [4]. 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.:

$$f(\mathbf{x}) = \sum_{i=1}^{K} f_i(\mathbf{x}) \tag{10}$$

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

$$f_i(\mathbf{x}) = [l_i \oplus n_i\, g]\,(\mathbf{x}) \tag{11}$$

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

$$n_i\, g(\mathbf{x}) = \underbrace{[g \oplus g \oplus \cdots \oplus g]}_{n_i \text{ times}}(\mathbf{x}). \tag{12}$$

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:

$$[f \ominus (n_1 + 1)g]\,(\mathbf{x}) \leq 0 \quad \forall x \in \mathcal{D}. \tag{13}$$

Accordingly, the first spine is given by:

$$l_1(\mathbf{x}) = [f \ominus n_1\, g]\,(\mathbf{x}). \tag{14}$$

Morphological shape decomposition can then be implemented recursively as follows.

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

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

$$\left[(f - \hat{f}_{i-1}) \ominus (n_i + 1)g\right](\mathbf{x}) \leq 0. \tag{15}$$

Step 3. Calculate the i-th component by:

$$f_i(\mathbf{x}) = \left\{ \underbrace{\left[(f - \hat{f}_{i-1}) \ominus n_i\, g\right]}_{l_i(\mathbf{x})} \oplus n_i\, g \right\}(\mathbf{x}). \tag{16}$$

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

$$\hat{f}_i(\mathbf{x}) = \hat{f}_{i-1}(\mathbf{x}) + f_i(\mathbf{x}). \tag{17}$$

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 $i = L-K, \ldots, L$ for K=15, where L denotes the maximal number of decomposition levels. 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 [7]:

$$\mathbf{j}(\mathbf{x}) = \left(f(\mathbf{x}), \hat{f}_{L-K}(\mathbf{x}), \ldots, \hat{f}_L(\mathbf{x})\right) \tag{18}$$

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 crucial in the verification procedure than others. Thus, 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 the database (i.e., the impostors). To do so, we would like to weigh the signal similarity measure at node i given by:

$$C_v(\mathbf{j}(\mathbf{x}_i^t), \mathbf{j}(\mathbf{x}_i^r)) = ||\mathbf{j}(\mathbf{x}_i^t) - \mathbf{j}(\mathbf{x}_i^r)|| \tag{19}$$

using class-dependent $DP_i(\mathcal{R})$ so that when person t claims the identity of person r a distance measure between them is computed by:

$$D(t,r) = \sum_{i \in \mathcal{V}} DP_i(\mathcal{R})\, C_v(\mathbf{j}(\mathbf{x}_i^t), \mathbf{j}(\mathbf{x}_i^r)) \tag{20}$$

where $\mathcal{R}$ denotes the class of the reference person r. Let $m_{\text{intra}}(\mathcal{X}, i)$ be the mean intra-class distance and $m_{\text{inter}}(\mathcal{X}, i)$ bet the mean inter-class distance for the class $\mathcal{X}$ at grid node i:

$$\begin{aligned} m_{\text{intra}} &= E\left\{C_v(\mathbf{j}(\mathbf{x}_i^t), \mathbf{j}(\mathbf{x}_i^r))\right\} \ \forall t, r \in \mathcal{X} \\ m_{\text{inter}} &= E\left\{C_v(\mathbf{j}(\mathbf{x}_i^t), \mathbf{j}(\mathbf{x}_i^r))\right\} \ \forall r \in \mathcal{X},\, t \in (\mathcal{S} - \mathcal{X}) \end{aligned} \tag{21}$$

where $\mathcal{S}$ denotes the set of all classes in the database. Let $\sigma^2_{\text{intra}}(\mathcal{X}, i)$ and $\sigma^2_{\text{inter}}(\mathcal{X}, i)$ be the variances of the intra-class distances and the inter-class distances, respectively. A plausible measure of the discriminatory power of the grid node i for the class $\mathcal{X}$ is the Fisher's Linear Discriminant (FLD) 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 Discriminatory Power Coefficient (DPC) for the grid node i, [9]:

$$DP_i(\mathcal{X}) = \frac{(m_{\text{inter}}(\mathcal{X}, i) - m_{\text{intra}}(\mathcal{X}, i))^2}{\sigma^2_{\text{inter}}(\mathcal{X}, i) + \sigma^2_{\text{intra}}(\mathcal{X}, i)}. \tag{22}$$

We can see that in (22) the $DP_i(\mathcal{X})$ is maximised when the denominator $\sigma^2_{\text{inter}}(\mathcal{X}, i) + \sigma^2_{\text{intra}}(\mathcal{X}, i)$ is minimised. 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:

$$DP_i(\mathcal{X}) = \frac{(m_{\text{inter}}(\mathcal{X}, i) - m_{\text{intra}}(\mathcal{X}, i))^2}{\sigma_{\text{inter}}(\mathcal{X}, i)\sigma_{\text{intra}}(\mathcal{X}, i)}. \qquad (23)$$

The denominator of (23) is interpreted as an OR rule for the variances of the clusters.

4. PERFORMANCE EVALUATION OF THE COMBINED SCHEMES

The combined schemes of MDLA with linear projections and MSD-DLA with discriminatory power coefficients have been tested on the M2VTS database [8]. 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 × 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 [5]. 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 are plotted in Figure 1. 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.,

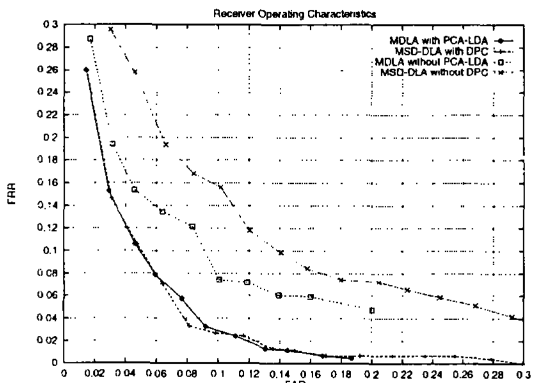

Figure 1: Receiver Operating Characteristics of MDLA with/without linear projections and MSD-DLA with/without local discriminatory power coefficients.

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 linear projections is 6.8% whereas the EER of MDLA is 9.35 % [5]. It is seen that the incorporation of linear projections improves the EER by 2.55 %. It is worth noting that the EER of MSD-DLA without local discriminatory power coefficients is 11.89 %. By using this discrimination criterion (22), we achieve an EER of 6.73% following the same experimental protocol. The same figure of merit using the discrimination criterion (23) 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 is very close to the one reported in [2], i.e., EER between 6.0 % and 9.2 %. However, it should be noted that only the most discriminant feature value per node has been used in our approach whereas the three most discriminative feature values have been employed in [2].

5. 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] 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.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

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

[10] 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.

[11] 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.