Morphological Elastic Graph Matching applied to frontal face authentication under optimal and real conditions

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

Abstract

In this paper, morphological elastic graph matching is applied to frontal face authentication on databases collected either under optimal conditions or during real-world tests (i.e., access-control to buildings or tele-services via INTERNET in a typical office environment etc.). It is demonstrated that the morphological elastic graph matching achieves a very low equal error rate on databases collected under optimal conditions. However, its performance deteriorates in real-world experiments. The compensation for variable recording conditions, such as changes in illumination, scale differences and varying face position prior to the application of morphological elastic matching is proposed. The results obtained indicate that the proposed approach overcomes the image variations and stabilizes the performance of the authentication algorithm.

1. Introduction

The interest and research activities in automatic face recognition have increased significantly over the past five years. The growth is mainly driven by application demands, such as identification for law enforcement and authentication for remote banking and access-control applications. A recent survey on face recognition can be found in [1]. Machine recognition of faces yields problems of the following two categories:

- **Face recognition**. Given a test face and a set of reference faces in a database find the N closest reference faces to the test one.
- **Face authentication**. Given a test face and a reference one, decide if the test face is identical to the reference one.

*This work was carried out within the framework of European ACTS project "Multi-modal Verification Techniques for Tele-services and Security Applications" (M2VTS).

The just mentioned problems are conceptually different. On the one hand, a face recognition system usually assists a human face-recognition expert to determine the identity of the test face by computing all similarity scores between the test face and each human face stored in the system database and by ranking them. On the other hand, a face authentication system should decide *itself* if the test face is a *client* (i.e., he or she claims his/her own identity) or is *an impostor* (i.e., he or she pretends to be someone else). The evaluation criteria for face recognition systems are different from those applied to face authentication systems. The performance of face recognition systems is quantified in terms of the percentage of correctly identified faces within the N best matches [2]. The performance of face authentication systems is measured in terms of the *false rejection rate* (FRR) achieved at a fixed *false acceptance rate* (FAR) or vice versa. By varying FAR, the *Receiver Operating Characteristic* (ROC) curve is obtained. A third difference is in the requirements needed when face recognition/authentication systems are tested. Although the algorithms employed in both face recognition and authentication systems are of common origin (for example, the dynamic link architecture [11]), there is no a priori guarantee that the same algorithm would enjoy the same performance level in both cases.

Face recognition has received more attention than face authentication by the scientific community. There are not many published works on face authentication. Even for face recognition only a handful of algorithms have been tested on databases greater than 150 individuals. Moreover, the conditions under which face databases are collected affect seriously the performance of a face recognition system. Recently, a comparative study has been performed for three well known face recognition techniques, namely, the eigenfaces, the auto-association and classification neural networks, and the elastic graph matching [3]. It has been found that the eigenfaces work well when the face images have relatively small lighting and moderate expression variations. Their performance deteriorates significantly as lighting variation increases. On the contrary, elastic graph matching is found relatively insensitive to variations in lighting, face position and expressions. Recognition errors close to 20% are the best rates reported for elastic graph matching when scale variations are present. Auto-association and classification networks actually break down (i.e., their recognition errors $\approx 60\%$) in the presence of scale variations. In the closely related work [4], Adini et al. presented an empirical study that evaluates the sensitivity of several image representations (e.g. edge maps, directional and non-directional derivatives of Gaussian filters, images convolved by 2-D Gabor-like filters) in changes of the illumination conditions. It has been found that all the aforementioned image representations are insufficient to overcome variations due to changes in illumination direction, view-point and expressions.

Motivated by [2, 3, 4], we test the performance of morphological elastic graph matching on databases ranging from small ones to large ones collected either under optimal or real conditions. Secondly, by evaluating the sensitivity of face authentication systems employing the morphological elastic graph matching to changes in illumination, face size and position, and facial expressions, we propose simple and powerful face normalization techniques that compensate for the varying recording conditions. Accordingly, the paper complements the previously published works [3, 4].

2. Morphological Elastic Graph Matching

In the following, a brief description of morphological dynamic link architecture (MDLA) is presented. In MDLA, we substitute the Gabor-based feature vectors used in elastic graph matching [11] by the *multiscale morphological dilation-erosion*. Let $\mathcal{R}$ and $\mathcal{Z}$ denote the set of real and integer numbers, respectively. Given an image $f(\mathbf{x}) : \mathcal{D} \subseteq \mathcal{Z}^2 \to \mathcal{R}$ and a structuring function $g(\mathbf{x}) : \mathcal{G} \subseteq \mathcal{Z}^2 \to \mathcal{R}$, the dilation of the image $f(\mathbf{x})$ by $g(\mathbf{x})$ is denoted by $(f \oplus g)(\mathbf{x})$.Its complementary operation, the erosion, is denoted by $(f \ominus g)(\mathbf{x})$. The *scaled hemisphere* is employed as structuring function. The multiscale dilation-erosion of the image $f(\mathbf{x})$ by $g_\sigma(\mathbf{x})$ is defined by [12]:

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

For $\sigma = -9, \ldots, 9$, the outputs of multiscale dilation-erosion form the feature vector located at grid node $\mathbf{x}$, i.e., $\mathbf{j}(\mathbf{x}) = ((f \star g_9)(\mathbf{x}), \ldots, f(\mathbf{x}), \ldots, (f \star g_{-9})(\mathbf{x}))$. An 8 $\times$ 8 sparse grid is created by measuring the feature vectors $\mathbf{j}(\mathbf{x})$ at equally spaced nodes over the output of the face detection algorithm described in [15]. The design of a face authentication system based on morphological elastic graph matching spans several issues that have been solved separately. Fast multiscale morphological operations for efficient feature extraction are proposed in [5]. Linear projection algorithms for feature selection have been studied in [6]. Automatic weighting of nodes according to their discriminatory power has been treated in [7].

3 Performance assessment on multimedia databases collected under optimal conditions

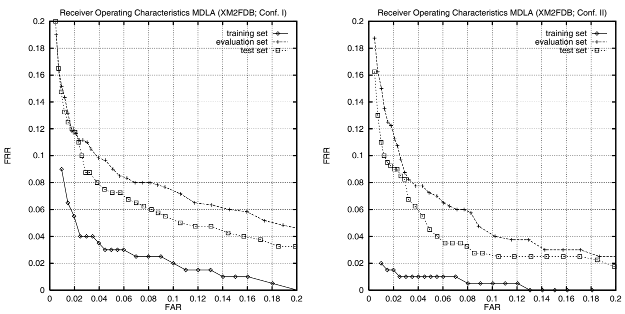
In this section, we assess the performance of morphological elastic graph matching when it is applied for face authentication on two multimedia databases collected under optimal conditions. The databases are the M2VTS database [10], and the extended M2VTS database (XM2VTSDB) [13].

The M2VTS database contains four recordings (i.e., shots) of 37 individuals. Four experimental sessions have been implemented by employing the "leave one out" principle producing 5,328 client claims and another 5,238 impostor claims. Details on the experimental protocol can be found in [5]. Table 1 summarizes the equal error rate achieved by the face authentication algorithm under study and compares it to those achieved by other competitive frontal authentication algorithms developed within M2VTS project. It is seen that the optimized MDLA with discriminatory power coefficients is ranked as the first method with respect to the equal error rate.

Table 1. Best equal error rates achieved by frontal face authentication algorithms within M2VTS project.

Method	EER (%)
Elastic Graph Matching (residual matching) [8]	11.8
Morphological Dynamic Link Architecture (without linear projections) [5]	9.3
Elastic Graph Matching (local discriminants) [8]	6.1
Morphological Dynamic Link Architecture (with linear projections of the feature vectors) [6]	5.4
Optimized robust correlation [9]	4.8
Morphological Dynamic Link Architecture with discriminatory power coefficients [7]	3.7

The extended M2VTS database contains 295 individuals' video data which include speech and image sequences of rotated heads. Eight recordings of 295 persons have been collected under optimal recording conditions. Morphological elastic graph matching has been tested on the XM2VTSDB according to the two configurations defined in [14]. The design of a person authentication system based on MDLA is split into three procedures, namely, the training, evaluation and test procedure. In the first configuration three frontal images have been used for each client. The training procedure results in 200 client claims and 39800 impostor claims. The receiver operating characteristic on the training set is plotted in Figure 1a. It is seen that an EER $\approx$ 3.0 % is achieved. Another 600 client claims and 40,000



(a) (b)

Figure 1. Receiver Operating Characteristics of morphological elastic graph matching on the training, evaluation and test set of the extended M2VTS database (a) in configuration I, and, (b) in configuration II.

impostor ones are produced during the evaluation procedure. Although there is a possibility to adapt the thresholds determined previously, we decided not to do so on purpose to keep the aforementioned claims disjoint to the claims that are employed in the subsequent test procedure. The evaluation claims can be exploited then to train a fusion manager. The ROC on the evaluation set is overlaid in Figure 1a. It can easily be seen that an EER $\approx$ 8.0 % is obtained on the evaluation set. The test procedure produces 400 client claims and 112,000 impostor ones. The test ROC is shown overlaid in Figure 1a. The inspection of Figure 1a reveals that an EER $\approx$ 6.57 % is achieved.

A second configuration is also tested that employs four training images for each client, two images for evaluation and another two images for testing. This configuration results in 400 evaluation client claims whereas the numbers of evaluation impostor claims, test client claims and test impostor claims are left intact. The ROCs are plotted in Figure 1b. It is seen that the EER is approximately 1.5 % on the training set, 6.2 % on the evaluation set and 5% on the test set.

4. Problems occurred in real-world tests

It is well-known that the conditions under which face databases are collected affect drastically the performance of face recognition/authentication algorithms [3, 4]. A key to the successful development of a general face authentication system is to systematically account for the different acquisition parameters, i.e., the varying lighting, background, pose and scale of the face etc. Accordingly, the evaluation of any authentication algorithm on a database collected under optimal conditions is inadequate when the algorithm is to be integrated in a platform for commercial exploitation. In such a case, it is of utmost importance to evaluate the authentication performance of the algorithm on real-world tests. Towards this goal two small galleries are collected under real operating conditions, namely, the MATRANORTEL database and the IBERMATICA database. The former database was collected by MATRANORTEL Communications, France, and the latter one by Ibermática, Spain.

The experiments reported on MATRANORTEL database have been conducted on 21 persons. Several sources of degradation are modeled in the database: (a) varying face size and position, (b) changes in illumination, and, (c) variable facial expressions. In addition to images belonging to the just-mentioned cases, the database contains one set of training images (1 image per person) and one set of test images (2 images per person) recorded under optimal conditions. That is, a uniform white background exists in the images, uniform lighting conditions are used, the face is of neutral expression and it is located at the center of the image. A set of experiments has been implemented so that the performance of the morphological elastic matching is correlated to each degradation source. 2,445 impostor claims and 117 client claims are produced in total. The first column in Table 3 summarizes the performance of morphological dynamic link architecture for each source of degradation separately and in the general case.

A set of similar experiments has been conducted on 11 persons from the IBERMATICA database. For each person, two training sample images are stored in the database. A large number of test images has also been recorded. In particular, more than 20 test client images of each person have been recorded. In addition, a couple of test impostor images per individual have been recorded. In total, 514 client accesses have been tested against 56 impostor ones. The three types of degradation previously described are present in this database as well. The performance of the elastic graph matching [8] and that of the morphological elastic graph matching have been evaluated on the IBERMATICA database. The EERs obtained are shown in Table 2. From the inspection of Tables 2 and 3, it becomes evident that the performance of the algorithms tested depends strongly

Table 2. Equal error rates achieved by elastic graph matching and morphological elastic graph matching when they are applied to IBERMATICA database.

Method	EER (%)
Elastic Graph Matching [8]	25
Morphological Elastic Graph Matching (MDLA)	35
Normalization + Morphological Elastic Graph Matching	20

on the variable recording conditions. To alleviate such a dependence, the compensation for the different conditions is needed. To do so, the use of simple and powerful face normalization techniques prior to the application of any authentication algorithm is proposed in the next Section.

5. Face normalization techniques

The proposed techniques are based on the detection of the facial region in the image and its splitting into two segments, the left segment and the right one. We assume that: (1) image background is uniform, and, (2) one person only appears in a scene. The algorithm is comprised of the following steps:

Step 1: The oval shape of a face can be approximated by an ellipse. Therefore, the detection of the facial area in an image can be performed by detecting an object of elliptical shape.

Step 2: The next step is to model the face-like region by an ellipse using moment-based features [16]. Let us denote the face-like area by C and the best-fit ellipse by $\mathcal{E}$. An ellipse is defined by its center (x_0, y_0), its orientation θ and the length a and b of its semi-major and semi-minor axes. The center of the ellipse is estimated by the center of mass of region C. The orientation of the ellipse is computed by determining the angle between the axis of the least moment of inertia and the horizontal axis of the coordinating system. The length of semi-major axis a and the length of semi-minor axis b can be computed by evaluating the least and the greatest moments of inertia. To find the ellipse which models better the given region, we iteratively maximize the measure:

$$\mathcal{M} = \sum_{(x,y)\in\mathcal{E}\cap C} 1 - \sum_{(x,y)\in\mathcal{E}\cap C^c} 1 \qquad (2)$$

where C^c denotes the complement of region C (i.e., the background). The maximization of (2) corresponds to the maximization of the number of correctly modeled pixels (i.e., $(x,y) \in \mathcal{E} \cap C$) and the minimization of the number of incorrectly modeled pixels (i.e., $(x,y) \in \mathcal{E} \cap C^c$).

Step 3: The first ellipse found is a coarse approximation of the facial area, because the hair, and in some cases, parts of the clothes are included. To overcome this problem the ellipse is subdivided into its left and right segments with respect to the vertical axis.

Step 4: The next step is to apply a clustering algorithm to each segment of the ellipse, separately. By choosing K=2, a K-means algorithm hopefully succeeds to relate the skin-like areas with a single cluster in each segment.

Step 5: The union of clusters in the left and right segment that correspond to skin-like areas is modeled by an ellipse using the algorithm described in **Step 2**. The quality of fit is measured again by (2). By doing so, a finer approximation is obtained. The ellipse found at the last iteration is the best-fit ellipse we searched for.

Lighting may cause unevenly illuminations of the right and the left face segments. To compensate for the aforementioned effect, the mean intensities of both face segments should be equalized. The problem of varying face position can be solved by translating the initial image so that the ellipse center always coincides with the image plane center. The width of the face can be approximated by the length of minor axis $2b$. Size (i.e., scale) normalization can be achieved by resizing the image with a horizontal scale factor $\frac{W_d}{2b}$, where W_d is the desired width of the normalized face. Image resizing is achieved by linear interpolation. The successful compensation for variable recording conditions in images from MATRANORTEL database is demonstrated in Figure 2.

To quantify the success of face normalization techniques, we evaluate the EER when the proposed techniques are incorporated into the morphological elastic graph matching. The gain in EER on MATRANORTEL database is shown in Table 3 for each degradation source separately as well as in the general case. A significant drop of 7.3% in EER (which amounts to a 33 % relative drop) is achieved in the general case. For comparison purposes, the EERs achieved without normalization are also included. The EER drops to 20% when the face normalization techniques are applied prior to MDLA on IBERMATICA database, as can be seen from the inspection of Table 2. That is, a significant drop of 15% in EER (which amounts to a 43 % relative drop) is achieved.

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.

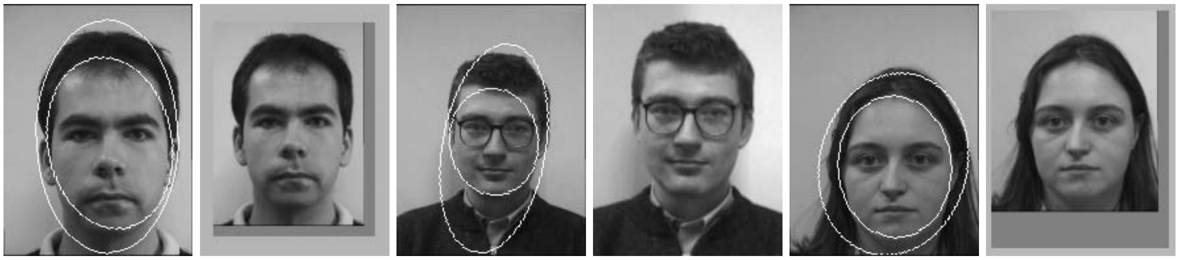


Figure 2. Facial images collected under variable recording conditions. The first, third and fifth images are original ones. The second, fourth and sixth images are normalized ones. The best ellipses determined by the algorithm are drawn on the original images.

Table 3. Equal error rates achieved on MATRANORTEL database by morphological elastic graph matching with and without face normalization.

Conditions	EER (%)	
	without normalization	with normalization
Optimal	12	9.5
lighting	33	15
face size	28	19
face position	23	17
expressions	17	13
general	22	14.7

[2] P. Jonathon Phillips, "Matching pursuit filters applied to face identification," *IEEE Trans. on Image Processing*, vol. 7, no. 8, pp. 1150–1164, August 1998.

[3] 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.

[4] 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.

[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)*, vol. I, pp. 105–108, Santa Barbara, California, U.S.A., 1997.

[6] A. Tefas, C. Kotropoulos and I. Pitas, "Variants of Dynamic Link Architecture based on mathematical morphology for frontal face authentication," in *Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 814–819, Santa Barbara, California, U.S.A., 1998.

[7] C. Kotropoulos, A. Tefas, and I. Pitas, "Frontal face authentication using variants of Dynamic Link Matching based on mathematical morphology," in *Proc. of the IEEE Int. Conf. on Image Processing*, vol. I, pp. 122–126, Chicago, U.S.A., October 1998.

[8] B. Duc, S. Fischer, and J. Bigün, "Face authentication with Gabor information on deformable graphs," *IEEE Trans. on Image Processing*, accepted for publication 1998.

[9] J. Matas, K. Jonsson, and J. Kittler, "Fast face localisation and verification," in *Proc. of British Machine Vision Conference*, 1997.

[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] 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.

[12] 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.

[13] K. Messer, J. Matas, and J. Kittler. "Acquisition of a large database for biometric identity verification," in *Proc. of BioSig 98*, June 1998.

[14] J. Luettin, and G. Maître, "Evaluation protocol for the extended M2VTS database (XM2VTSDB)," in *IDIAP Communication 98-05*, IDIAP, Martigny, Switzerland 1998.

[15] 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.

[16] A.K. Jain, *Fundamentals of Digital Image Processing*. Prentice-Hall, Englewood Cliffs, NJ, 1989.