

Expression-Compensated 3D Face Recognition with Geodesically Aligned Bilinear Models

Iordanis Mpiperis^{1,2}, Sotiris Malassiotis¹ and Michael G. Strintzis^{1,2}

Abstract—In this paper, we present a technique for addressing 3-dimensional face recognition in presence of facial expressions using a bilinear model. The bilinear model allows decoupling the impact of identity and expression on face appearance and encoding their contribution in separate control parameters. This is achieved by first representing faces as parametric surface models described by a fixed length parameter vector. A generic face model is fitted to each face based on a novel technique that relies on geodesic distances to find implicitly corresponding facial landmarks between the model and the face in hand. Model parameters are then used for bilinear decomposition. The experimental results on the publicly available BU-3DFE face database demonstrate the effectiveness of our technique.

I. INTRODUCTION

This paper describes a novel technique for expression-invariant 3D face recognition based on a bilinear model. Bilinear models were introduced by Tenenbaum and Freeman [1] to describe two-factor observations, where concepts like the “content” and “style” of observations should be analyzed and manipulated separately from each other. In this paper, we use a symmetric bilinear model [1] to decouple the impact of identity and expression on the appearance of the face surface and encode quantitatively their contribution in independent control parameters. These parameters are then used for expression-invariant face recognition.

The bilinear decomposition may be applied to vector representations of faces with the same dimensions. Thus a deformable face model is fitted to each face and then the parameters of this model are used for the bilinear decomposition. Model fitting must be carried so that corresponding anatomical features between the model and the face be aligned. Therefore, we follow an iterative optimization scheme that minimizes the model-surface Euclidean distance and also leads to similar mappings onto the geodesic polar coordinate plane [2].

In the following, we present a brief review of some representative works related to our own and discuss their limitations that motivated our technique. In Section II, we describe the technique for fitting the deformable face model to a face surface using their geodesic polar representations, while in Section III, we present the construction of the bilinear model and a novel technique for training it, which can handle unevenly distributed training data. Its use in face

recognition and experiments performed on the BU-3DFE [3] database are presented in Section IV, while final conclusions are drawn in Section V.

A. Related work

Over the past years, a great number of different 3D face recognition techniques have been reported in literature (see [4] for an extensive survey). Recently though, it became evident that the presence of facial expressions limits the performance of 3D face recognition and thus several researchers sought to address this problem. The majority of works detect the deformable regions of the face and then try to suppress their contribution during matching. [5] and [6] are representative works falling into this category. In [5], Kakadiaris *et al.* recognize faces based on the wavelet coefficients of rigid facial parts, which are found by fitting an Annotated Face Model to the face. In [6] on the other hand, Chang *et al.* follow a multi-region technique in which multiple overlapping regions around the nose are matched using the Iterative Closest Point algorithm (ICP).

The disadvantage of the previous approaches is that they reject deformable parts of the face that still encompass discriminative information. To avoid this, other researchers use features and face representations that are invariant to deformations of the face surface due to expressions. In this framework, Li and Zhang [7] classify faces using descriptors based on surface curvature, geodesic distance and attributes of a 3D mesh fitted to the face. Bronstein *et al.* [8] and Mpiperis *et al.* [2] form expression-invariant representations of faces assuming an isometric model for surface deformations. Bronstein *et al.* use multi-dimensional scaling on pair-wise geodesic distances to embed the surface to a 4D sphere, where classification is performed on the basis of normalized moments. Instead, Mpiperis *et al.* use a geodesic polar parameterization of the surface to construct expression-invariant attribute images that are classified following standard 2D techniques. These approaches depend on the practical observation that facial skin does not stretch significantly during expressions, which, however, is not true in case of intense expressions.

As it can be seen from above, most face recognition techniques work along the following lines: First, they build a discriminative feature vector that is robust to expressions, and then, they classify this vector using typical pattern recognition techniques. Instead of trying to devise appropriate invariant representations, we alternatively propose to encode the contribution of identity and expression in independent parameters. An advantage of such an approach is that it

¹Informatics and Telematics Institute, Centre for Research and Technology Hellas, 1st Km Thermi-Panorama Rd, Thessaloniki, Greece 57001.

²Information Processing Laboratory, Electrical and Computer Engineering Department, Aristotle University of Thessaloniki, Thessaloniki, Greece 54124.

Email: iordanis@iti.gr, malasiot@iti.gr, strintzi@iti.gr

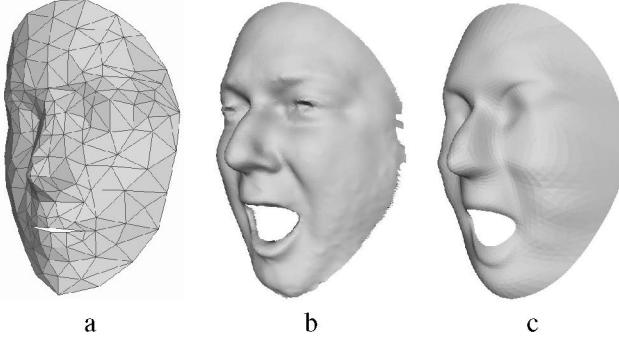


Fig. 1. Fitting the geodesically aligned deformable model to a surface. a: base-mesh, b: original surface, c: subdivision-mesh fitted to the surface.

can be used unaltered for facial expression recognition as well, which could be useful if, apart from the identity of the individual, an estimation of his/her psychological state is also important. To our knowledge, there are no published studies about 3D face recognition following this approach. Only few researchers have investigated it and they have used 2D imagery for this purpose. In general, they decouple appearance components by applying Singular Value Decomposition (SVD) to matrices formed from flattening data tensors. For instance, Vasilescu and Terzopoulos [9] use the N-mode SVD tensor decomposition to separate the influence of identity, pose, illumination and expression, while Wang and Ahuja [10] use Higher-Order SVD to recognize and synthesize facial expressions.

Our work is motivated by the work of Tenenbaum and Freeman in [1] and related publications, who used bilinear models to describe two-factor observations such as the font and the letters of a text or the accent and the meaning of a spoken word. Bilinear models are linear in either factor when the other is held constant and as such they share almost all the advantages of linear models: they are simple in structure, computation and implementation; they can be trained with well known algorithms; and their complexity can be easily adjusted by their dimensionality. Despite their simplicity however, they can model subtle interactions by allowing factors to modulate each other's contribution multiplicatively [1].

II. ESTABLISHMENT OF ANATOMICAL CORRESPONDENCE AND VECTOR REPRESENTATION OF FACES

The bilinear modelling of the face manifold requires face surfaces be given in a vector form. This is achieved by fitting a parameterized deformable 3D face model to each sample surface. Its parameters define bijectively its configuration and therefore they can be used for the vector representation of the face. However, this is true only if the fit of the model satisfies the constraint of anatomical correspondence.

Fitting begins by defining a 3D mesh M_0 with N vertices \mathbf{v}_i (see Fig. 1). This base-mesh is subdivided using the Loop subdivision scheme [11] to give a more smooth and dense mesh. At each subdivision step, each triangle is subdivided into 4 sub-triangles with the introduction of new vertices on its edges. Each vertex of the resulting 3D mesh may

be written as a linear combination of the vertices of the previous level mesh and eventually of the initial mesh M_0 . Theoretically, after an infinite number of subdivisions the mesh converges to a continuous smooth surface, which is a function of the initial mesh and the subdivision rule. In practice however, we do not make infinite subdivisions, since after a few levels (e.g. 3 in our experiments) the mesh becomes dense enough to approximate the subdivision surface. This final mesh, called subdivision-mesh, serves as the deformable model, while the vertices of the base-mesh that define its form comprise the vector representation of the face.

Let \tilde{M} denote the subdivision-mesh and $\tilde{\mathbf{v}}_i$, $i = 1 \dots S$ its vertices. If also h_{ij} are the coefficients of the linear combinations between the vertices of the base-mesh and the subdivision-mesh, then altogether we have

$$\tilde{\mathbf{v}}_i = \sum_{j=1}^N h_{ij} \mathbf{v}_j. \quad (1)$$

A common approach to model fitting (e.g. [12]–[14]) is formulating it as an energy minimization problem. The energy is comprised of the Euclidean distances between the points of the model and their nearest counterparts on the surface. Considering (1) and defining the function $mc(i)$ that returns the index k of the facial point \mathbf{p}_k nearest to the \tilde{M} vertex i , this energy term can be written as

$$E_{mc} = \sum_{i=1}^S \left(\sum_{j=1}^N h_{ij} \mathbf{v}_j - \mathbf{p}_{mc(i)} \right)^2. \quad (2)$$

However, this energy term alone leads to an under-constrained problem whose solution may not represent a plausible human face (e.g. vertices may be set to disparate points and fold the triangles of the mesh). Therefore, a regularization or smoothness term that tries to prevent the model from distorting is also added to the energy formulation. Smoothness is defined as a measure of the elastic energy of the base-mesh that penalizes non-parallel displacements of the edges and is given by

$$E_e = \sum_{i=1}^N \frac{1}{N_i} \sum_{j \in \mathcal{N}_i} (\mathbf{v}_i - \mathbf{v}_j - \mathbf{v}_i^0 + \mathbf{v}_j^0)^2, \quad (3)$$

where \mathcal{N}_i is the set of \mathbf{v}_i 's neighbors, N_i is its cardinality and \mathbf{v}_i^0 , \mathbf{v}_j^0 are the initial positions of the vertices.

E_{mc} gives rise to forces that attract model vertices towards their nearest points on the surface instead of the anatomically corresponding points. This is not a problem if the model is relatively close to the surface, since nearest points and anatomically corresponding points almost coincide. But if the model is relatively far from the surface, vertices may be displaced so that anatomically erroneous correspondence is established. To overcome this problem, first we define a dense correspondence field between the model and the face taking into account their geodesic polar parameterizations, and then, we perform fitting in an iterative way.

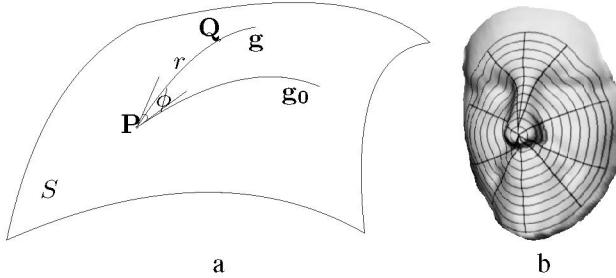


Fig. 2. Geodesic polar parameterization of a surface. a: Definition of geodesic paths and circles. b: Geodesic paths and circles over a face surface.

Given a surface S and two points \mathbf{P} , \mathbf{Q} on it, as shown in Fig. 2, there is an infinite number of curves that belong to S and connect \mathbf{P} with \mathbf{Q} . The curve g with the minimum length is called *geodesic path* and its length *geodesic distance* of the two points. We also define the *geodesic circle* as the geometric locus of the surface points that are at constant geodesic distance r from a given reference point called *geodesic pole*. Now, let \mathbf{P} be the geodesic pole. Let also G be the set of all geodesic paths that pass through \mathbf{P} and g_0 a geodesic path selected arbitrarily from G and used as reference. If the surface is smooth and without holes, then there exists a region around \mathbf{P} , where \mathbf{P} is the only common intersection point of any two geodesic paths of G . \mathbf{Q} may then be defined bijectively by the geodesic distance r from \mathbf{P} and the polar angle ϕ between g and g_0 . (r, ϕ) form the geodesic polar coordinates of \mathbf{Q} .

Mpiperis *et al.* [2] showed that geodesic polar coordinates are preserved during expressions and used them to build expression-invariant attribute images. They also proposed a modification of geodesic polar coordinates that can handle the mouth hole. Here, we use them to define a correspondence field between the model and the face. That is, we add an extra energy term

$$E_c = \sum_{i=1}^S (\tilde{\mathbf{v}}_i - \mathbf{p}_{c(i)})^2, \quad (4)$$

where we assume that surface points \mathbf{p}_i correspond to model vertices $\tilde{\mathbf{v}}_i$ (correspondence is given by function $c(\cdot)$), if they have the same geodesic polar coordinates. Again, this may result to poor anatomical correspondence if the model is too different from the face. To handle both this problem and the similar one regarding nearest points above, we adopt an iterative approach.

In each iteration, we compute the geodesic polar coordinates of the surface points and the subdivision-mesh vertices and we form the couples with the same coordinates. We also form the couples of vertices and their nearest points on the surface, *i.e.* we build functions $c(\cdot)$ and $mc(\cdot)$ participating in (4) and (2) respectively. Model parameters are then estimated by minimizing the energy function

$$E_{def} = \lambda_1 E_c + \lambda_2 E_{mc} + \lambda_4 E_e. \quad (5)$$

E_{def} is quadratic with respect to the unknown model parameters and therefore its minimization can be achieved

easily by solving a simple linear system. Let $\hat{\mathbf{v}}[k] = [\hat{\mathbf{v}}_1[k]^T \dots \hat{\mathbf{v}}_N[k]^T]^T$ stand for the base-mesh vertices that minimize (5) in the k -th iteration and η be a step chosen in $(0, 1)$. Then, the base-mesh vertices $\mathbf{v} = [\mathbf{v}_1^T \dots \mathbf{v}_N^T]^T$ that are used for the vector representation of the surface may be found using the following update rule

$$\mathbf{v}[k] = (1 - \eta)\mathbf{v}[k - 1] + \eta\hat{\mathbf{v}}[k] \quad (6)$$

upon convergence to a final position.

III. MODELLING EXPRESSION AND IDENTITY VARIATION

Using the vector representation described above, we may now model the face manifold by means of a bilinear model.

Let \mathbf{v}^{xp} be the K -dimensional stacked column vector of the N base-mesh vertices of the facial surface of person p with expression x ($K = 3N$). Then each component v_k^{xp} is given by the general bilinear form [1]

$$v_k^{xp} = \sum_{i=1}^I \sum_{j=1}^J w_{ijk} a_i^x b_j^p, \quad (7)$$

where a_i^x and b_j^p are the control parameters that control expression and identity respectively¹, while w_{ijk} are the coefficients that model the interaction of the factors². Using matrix notation (7) is simplified to

$$\mathbf{v}_k^{xp} = \mathbf{a}^{xT} \mathbf{W}_k \mathbf{b}^p, \quad (8)$$

where now $\mathbf{a}^x = [a_1^x \dots a_I^x]^T$, $\mathbf{b}^p = [b_1^p \dots b_J^p]^T$ and $\mathbf{W}_k(i, j) = w_{ijk}$

Let us assume that there exist T faces in our database belonging to T_p individuals and depicting one of T_x possible expressions. Let also $h_{xp}[t]$ be a zero-one function that is one if the t -th face $\mathbf{v}[t]$ belongs to individual p with expression x . Unknown coefficients arise from the minimization of the total squared error [1]

$$E_s = \sum_{t=1}^T \sum_{x=1}^{T_x} \sum_{p=1}^{T_p} \sum_{k=1}^K h_{xp}[t] (\mathbf{v}_k[t] - \mathbf{a}^{xT} \mathbf{W}_k \mathbf{b}^p)^2. \quad (9)$$

By differentiating with respect to \mathbf{a}^x , \mathbf{b}^p and \mathbf{W}_k and setting partial derivatives equal to zero, we end up with the system of equations

$$\mathbf{a}^x = \left(\sum_{p=1}^{T_p} \sum_{k=1}^K n_{xp} \mathbf{W}_k \mathbf{b}^p \mathbf{b}^{pT} \mathbf{W}_k^T \right)^{-1} \left(\sum_{p=1}^{T_p} \sum_{k=1}^K m_k^{xp} \mathbf{W}_k \mathbf{b}^p \right) \quad (10)$$

¹Dimensionalities I and J of a_i^x and b_j^p respectively will be discussed later.

²Equation (7) shows that coefficients w_{ijk} are weighted symmetrically by a_i^x and b_j^p and thus this model is called *symmetric* in the literature.

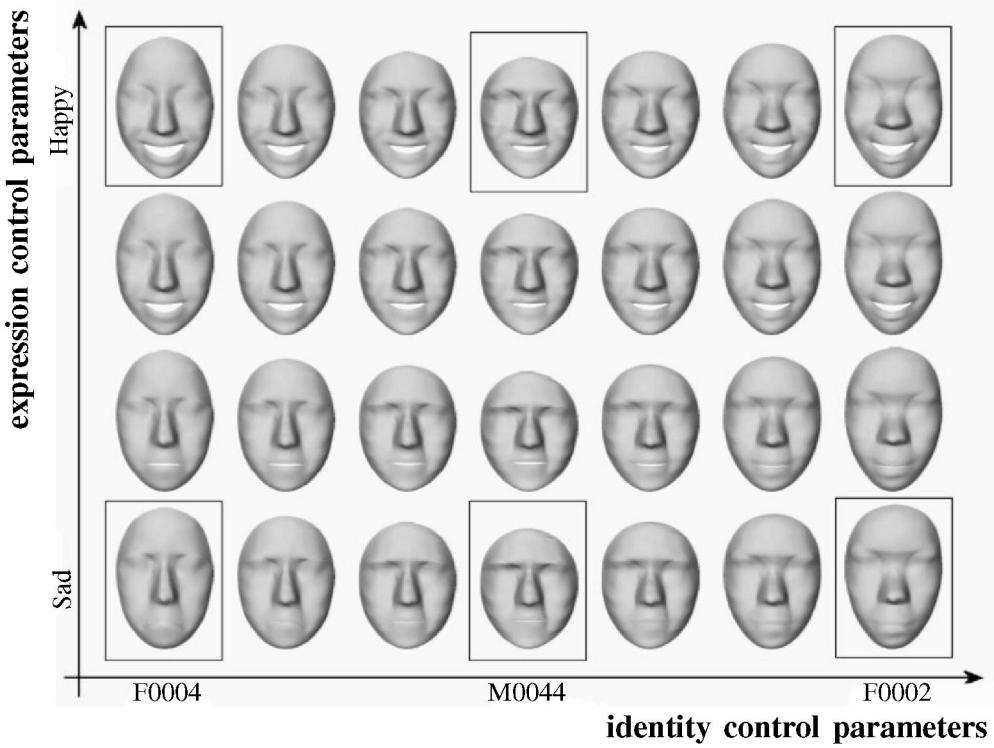


Fig. 3. Deformation of facial surface across expression and identity control parameters. Inside boxes are surfaces which are stored in the BU-3DFE database. Specifically, it is subjects F0004, M0044 and F0002 displaying the *sad* and *happy* expressions. The rest surfaces have been generated by linear interpolation of expression and identity control parameters of the former surfaces.

$$\mathbf{b}^p = \left(\sum_{x=1}^{T_x} \sum_{k=1}^K n_{xp} \mathbf{W}_k^T \mathbf{a}^x \mathbf{a}^{xT} \mathbf{W}_k \right)^{-1} \left(\sum_{x=1}^{T_x} \sum_{k=1}^K m_k^{xp} \mathbf{W}_k^T \mathbf{a}^x \right) \quad (11)$$

$$\text{vec}(\mathbf{W}_k) = \left(\sum_{x=1}^{T_x} \sum_{p=1}^{T_p} n_{xp} \mathbf{b}^p \mathbf{b}^{pT} \otimes \mathbf{a}^x \mathbf{a}^{xT} \right)^{-1} \text{vec} \left(\sum_{x=1}^{T_x} \sum_{p=1}^{T_p} m_k^{xp} \mathbf{a}^x \mathbf{b}^{pT} \right). \quad (12)$$

$n_{xp} = \sum_{t=1}^T h_{xp}[t]$ is the number of training faces that belong to subject p and depict expression x , \mathbf{m}^{xp} is their sum, $\mathbf{m}^{xp} = [m_1^{xp} \dots m_k^{xp} \dots m_K^{xp}]^T = \sum_{t=1}^T h_{xp}[t] \mathbf{v}[t]$, \otimes is the Kronecker product operator and $\text{vec}(\cdot)$ is the matrix vectorization operator that stacks the columns of the matrix.

Interaction matrices \mathbf{W}_k and control vectors \mathbf{a}^x and \mathbf{b}^p may now be found by iterating equations (12), (10) and (11) respectively. To ensure the stability of the solution, updating is performed progressively according to the rule

$$\mathbf{X}[n] = (1 - \eta) \mathbf{X}[n-1] + \eta \hat{\mathbf{X}}[n], \quad (14)$$

where η is the step size usually chosen in $[0.2, 0.8]$, $\mathbf{X}[n]$ stands for the final value of \mathbf{a}^x , \mathbf{b}^p or \mathbf{W}_k in the n -th

iteration, and $\hat{\mathbf{X}}[n]$ stands for the value resulting from (10), (11) or (12) respectively.

Convergence depends on the dimensionalities I of \mathbf{a}^x , and J of \mathbf{b}^p , which control the exactness of the training data reproduction. Convergence is guaranteed if I and J are less than or equal to T_x and T_p , the number of expressions and the number of individuals respectively. If I is equal to T_x and J is equal to T_p , training data are reproduced exactly, while more coarse but also more compact representations result if these dimensionalities are decreased.

The minimization of the total squared error through equations (10), (11) and (12) presented here differs from the optimization scheme proposed by Tenenbaum and Freeman in [1] and their previous work. There, minimization is achieved by means of SVD applied to a mean observation block matrix. That technique relies on evenly distributed data across expression and identity, which may not be valid in practice. Even worse, there may not be available data for a particular expression-identity combination. In such cases the mean observation matrix has indeterminate entries and SVD is not directly applicable. In [1] the authors propose filling the missing entries with the mean values of the appropriate expression and identity, but this substitution does not guarantee the global minimization of the error. In contrast, our method may be applied directly without any assumptions for the data distribution, but at the expense of computational complexity.

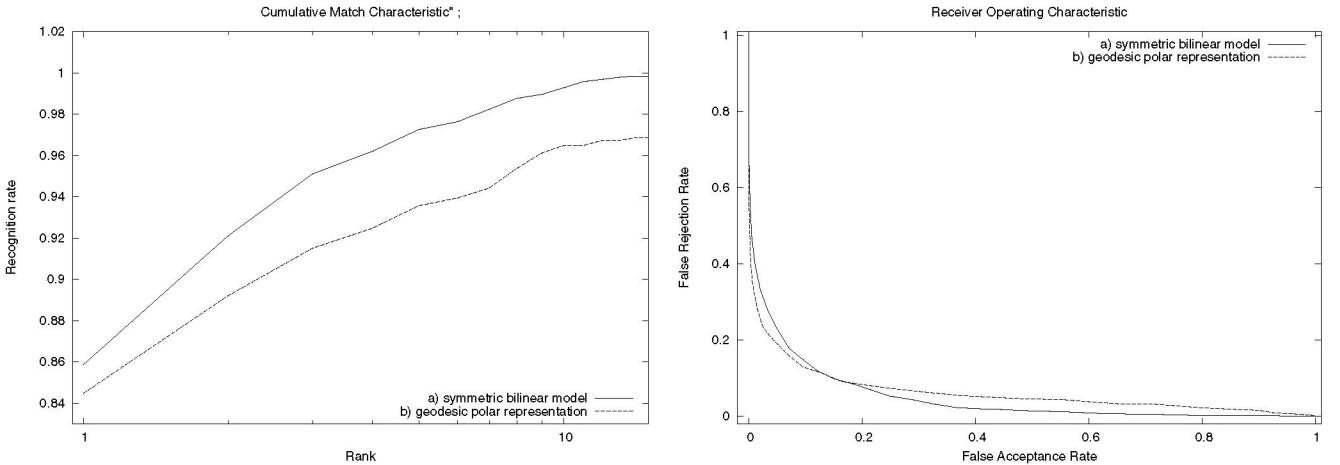


Fig. 4. Cumulative Match Characteristic (CMC) and Receiver Operating Characteristic (ROC) for face recognition based on: a) the symmetric bilinear model, b) the geodesic polar representation presented in [2].

IV. FACE RECOGNITION

In this section, we present an application of the deformable and bilinear model to expression-invariant face recognition using the BU-3DFE face database [3]. BU-3DFE contains 2,500 face scans of 100 subjects, 56 female and 44 male with a variety of ethnic/racial ancestries, who display the six universal expressions of *anger*, *fear*, *disgust*, *happiness*, *sadness* and *surprise*. To demonstrate the effectiveness of our technique, we adopt a procedure that simulates a realistic application scenario and we compare it with Mpiperis *et al.*'s work [2].

First, we split the database into two parts based on subject identity. The first part serves as a bootstrap set and is used for training the bilinear model. The rest of the data, the test set, is further split into the gallery set that contains a single 3D image per subject (neutral or non-neutral), and the probe set that contains the images to be classified (various facial expressions).

The bootstrap set is comprised of 50 subjects chosen randomly from the database while the gallery and probe set are comprised of the rest 50 subjects. Faces in the bootstrap set are set into correspondence and then they are used to train the symmetric bilinear model. The dimension of vectors \mathbf{b}^p and \mathbf{a}^x are set to 45 and 5 respectively. \mathbf{a}^x , \mathbf{b}^p and \mathbf{W}_k are initialized randomly and then they are optimized by iterating equations (10), (11) and (12). Optimization terminates when the relative change in their Frobenius norm gets below a threshold (0.01) or a maximum number of iterations (150–180) is reached.

Then we proceed with gallery image processing and probe image classification. Gallery and probe images are first set into correspondence and then their expression and identity control vectors are extracted using the optimal mixing matrices \mathbf{W}_k found from the bootstrap set. Control vectors \mathbf{a}^x and \mathbf{b}^p are obtained by minimizing the reconstruction squared

error

$$E = \sum_{k=1}^K \left(\mathbf{v}_k - \mathbf{a}^{xT} \mathbf{W}_k \mathbf{b}^p \right)^2. \quad (15)$$

Minimization is achieved similarly to the minimization of the total squared error in (9) during the training of the symmetric bilinear model. By differentiating (15) with respect to \mathbf{a}^x and \mathbf{b}^p and setting partial derivatives equal to zero, control vectors are found by iterating the system of equations

$$\mathbf{a}^x = \left(\sum_{k=1}^K \mathbf{W}_k \mathbf{b}^p \mathbf{b}^{pT} \mathbf{W}_k^T \right)^{-1} \left(\sum_{k=1}^K \mathbf{v}_k \mathbf{W}_k \mathbf{b}^p \right) \quad (16)$$

$$\mathbf{b}^p = \left(\sum_{k=1}^K \mathbf{W}_k^T \mathbf{a}^x \mathbf{a}^{xT} \mathbf{W}_k \right)^{-1} \left(\sum_{k=1}^K \mathbf{v}_k \mathbf{W}_k^T \mathbf{a}^x \right). \quad (17)$$

Now that \mathbf{a}^x and \mathbf{b}^p are known, we may proceed to matching probe images against each gallery image. To account for expression, we force the probe and gallery image to display a common expression, that defined by gallery image's expression control vector \mathbf{a}^x . Classification can then be easily performed based on the Euclidean distance between probe and gallery face vector representations in a nearest neighbor framework.

The above experiments are repeated on several randomly chosen subdivisions of bootstrap and test sets, under the constraint that all subjects are included at least once in the test set. The recognition results are averaged and presented in the form of a cumulative match characteristic and a receiver operating characteristic in Fig. 4. Fig. 4 includes also the results obtained by [2], where the authors use geodesic polar coordinates to build expression-invariant attribute images that are used for recognition. The increase in the recognition rate shows that our deformable and bilinear model may deal well with expressions, which are one of the main limitations of current 3D face recognition systems.

V. CONCLUSION

In this paper, we have demonstrated the application of bilinear models to expression-invariant face recognition, obtaining improved classification results. One of the advantages of our techniques is that it can be used also for identity-invariant facial expression recognition simply by interchanging the roles of identity and expression. However, there are still some issues that limit performance and provide room for further investigation.

In our experiments, we observed that poor correspondence affects substantially recognition performance. The problem is twofold: During training, the bilinear model cannot learn the true identity-expression manifold implying errors in coefficient estimation. During testing, expression manipulation is actually applied on a slightly (or quite) different face. This error is further amplified by inaccurate bilinear parameters and leads to a distorted facial surface.

Another limitation is the need of a large bootstrap set, which should also be annotated with respect to facial expressions. The more different expressions are present in the bootstrap set, the better is the estimation of the interaction matrices \mathbf{W}_k and the better is the fit to novel faces. Training with a few expressions leads to unbalanced generalization ability in favour of identity, which in turn leads to better surface approximation but poorer expression control.

Considering the above remarks, the improvement of the point correspondence is the most challenging problem and this is where we will focus our future efforts.

VI. ACKNOWLEDGMENTS

This work was supported by: a) the European Commission under the FP6 IST Project “PASION - Psychologically Augmented Social Interaction over Networks” (contract No FP6-027654); and b) the Greek Ministry of Development-General Secretariat of Research and Technology under Project “PENED 2003 Ontomedia” (03EΔ475).

REFERENCES

- [1] J. Tenenbaum and W. Freeman, “Separating style and content with bilinear models,” *Neural Computation*, vol. 12, pp. 1247–1283, 2000.
- [2] I. Mpiperis, S. Malassiotis, and M. G. Strintzis, “3D face recognition with the geodesic polar representation,” *IEEE Trans. Information Forensics and Security*, vol. 2, no. 3, pp. 537–547, September 2007.
- [3] L. Yin, X. Wei, Y. Sun, J. Wang, and M. J. Rosato, “A 3D facial expression database for facial behavior research,” in *7th International Conference on Automatic Face and Gesture Recognition (FGR06)*, April 2007, pp. 211–216.
- [4] K. Bowyer, K. Chang, and P. Flynn, “A survey of approaches and challenges in 3D and multi-modal 3D+2D face recognition,” *Comp. Vision and Image Understanding*, vol. 101, pp. 1–15, 2006.
- [5] I. Kakadiaris, G. Passalis, G. Toderici, M. Murtuza, Y. Lu, N. Karampatziakis, and T. Theoharis, “Three-dimensional face recognition in the presence of facial expressions: An annotated deformable model approach,” *IEEE Trans. Pattern Anal. and Mach. Intell.*, vol. 29, no. 4, pp. 640–649, April 2007.
- [6] K. Chang, K. Bowyer, and P. Flynn, “Multiple nose region matching for 3D face recognition under varying facial expression,” *IEEE Trans. Pattern Anal. and Mach. Intell.*, vol. 28, no. 10, pp. 1695–1700, October 2006.
- [7] X. Li and H. Zhang, “Adapting geometric attributes for expression-invariant 3D face recognition,” in *IEEE Int. Conf. on Shape Modeling and Applications*, 2007, pp. 21–32.
- [8] A. M. Bronstein, M. M. Bronstein, and R. Kimmel, “Expression-invariant representations of faces,” *IEEE Trans. Image Processing*, vol. 16, no. 1, January 2007.
- [9] M. A. O. Vasilescu and D. Terzopoulos, “Multilinear subspace analysis of image ensembles,” in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, June 2003, pp. 93–99.
- [10] H. Wang and N. Ahuja, “Facial expression decomposition,” in *Proc. Ninth IEEE Int. Conf. on Computer Vision*, vol. 2, October 2003, pp. 958–965.
- [11] C. Mandal, H. Qin, and B. C. Vemuri, “Novel FEM-based dynamic framework for subdivision surfaces,” *Computer-Aided Design*, vol. 32, no. 8, pp. 479–497, 2000.
- [12] C. Shelton, “Morphable surface models,” *International Journal of Computer Vision*, vol. 38, no. 1, pp. 75–91, 2000.
- [13] D. Metaxas and I. Kakadiaris, “Elastically adaptive deformable models,” *IEEE Trans. Pattern Anal. and Mach. Intell.*, vol. 24, no. 10, pp. 1310–1321, October 2002.
- [14] B. Allen, B. Curless, and Z. Popović, “The space of human body shapes: reconstruction and parameterization from range scans,” in *SIGGRAPH '03: ACM SIGGRAPH 2003 Papers*. New York, NY, USA: ACM Press, 2003, pp. 587–594.