Robust expression-invariant face recognition from partially missing data

Alexander M. Bronstein[1], Michael M. Bronstein[1], and Ron Kimmel[1]

Dept. of Computer Science, Technion – Israel Institute of Technology, Haifa 32000, Israel
{alexbron, bronstein}@ieee.com, ron@cs.technion.ac.il

Abstract. Recent studies on three-dimensional face recognition proposed to model facial expressions as isometries of the facial surface. Based on this model, expression-invariant signatures of the face were constructed by means of approximate isometric embedding into flat spaces. Here, we apply a new method for measuring isometry-invariant similarity between faces by embedding one facial surface into another. We demonstrate that our approach has several significant advantages, one of which is the ability to handle partially missing data. Promising face recognition results are obtained in numerical experiments even when the facial surfaces are severely occluded.

1 Introduction

Face recognition deals with the problem of identifying a human subject using information describing his or her face. A description of a subject to be identified (*probe*), is compared to those stored in the database of subjects with known identity (usually referred to as *gallery*). The probe is accepted if identified with one of the gallery subject, or otherwise rejected, based on some distance function. Ideally, there should be no false acceptances or false rejections.

Recently, three-dimensional (3D) face recognition has become an emerging modality, trying to use 3D geometry of the face for accurate identification of the subject. While traditional two-dimensional (2D) face recognition methods suffer from sensitivity to factors such as illumination, head pose and the use of cosmetics, 3D methods appear to be more robust to these factors. Yet, the problem of *facial expressions* is a major issue in 3D face recognition, since the geometry of the face significantly changes as a result of facial expressions. One of the focuses of the recent Face Recognition Grand Challenge (FRGC) competition is robustness to facial expressions [1, 2].

In [3], we introduced an expression-invariant 3D face recognition method. Our main thesis is the *isometric model*, according to which facial expressions are modelled as isometries of the facial surface. The subject's identity is associated with the intrinsic geometry of the surface (i.e. its metric structure), which appears to be nearly expression-invariant [3]. Getting rid of the extrinsic geometry of the surface and using its intrinsic geometry only, an expression-invariant representation of the face is constructed. We used the approach presented by Elad and Kimmel [4]. Mapping the surface in an isometric way into $\mathbb{R}^3$, where the original geodesic distances are replaced with the Euclidean ones, one creates a representation of the intrinsic geometry, which can be simply handled as a rigid surface. Such a mapping is termed *isometric embedding*. Elad and Kimmel used a numerical procedure called *multidimensional scaling* (MDS) [5] to compute the embedding.

Face recognition based on the isometric embedding approach is simple and computationally-efficient (in [3], a real-time system that acquires and matches two surfaces in less than 5 *sec* was obtained). The disadvantage is the fact that in general, a surface cannot be isometrically embedded into $\mathbb{R}^m$, and therefore, such a mapping introduces an inevitable distortion of the distances (*embedding error* or *metric distortion*), which reduces the recognition accuracy. Attempts to reduce the embedding error were made in [6–8] by resorting to non-Euclidean embedding spaces. In [8], it was conjectured that smaller embedding error results in better face recognition accuracy. This conjecture was proved experimentally using two-dimensional spheres with different radii as the embedding space.

The main limitation of the different embedding spaces used beforehand was the demand that the geodesic distances are expressed *analytically*. This practically limits the possibilities to spheres and flat domains. It is obvious, however, that the embedding of one surface into another results in *zero* metric distortion if the surfaces are isometric. If the surfaces are not isometric, the embedding error could be a measure of how different their intrinsic geometry is.

Unfortunately, a facial surface has a complicated metric structure and the geodesic distances can not be expressed analytically. The price we have to pay in order to perform embedding into such spaces is that the geodesic distances must be computed numerically. However, the apparent advantages seem to justify it. In addition to higher accuracy, this kind of embedding allows to compare *partially missing* surfaces. This is especially important in practical 3D face recognition, where due to acquisition imperfections the facial surfaces can be occluded. The ability to handle partially missing data also frees us from the need to perform sophisticated cropping identical for all faces, which is required in [3].

This paper consists of five sections. In Section 2, we review the expression-invariant face recognition method based on the isometric model. In Section 3, we introduce the notion of partial embedding and outline a recent generalization of MDS as a way to compute it [9]. Section 4 outlines a hierarchical matching scheme for one-to-many face recognition with very large databases. Section 5 is devoted to experimental results. We show that our approach works accurately even in a setting where the facial surface is severely occluded. Section 6 concludes the paper.

2 Expression-invariant face recognition

Our starting point is the isometric model of facial expressions, introduced in [3]. The facial surface is described as a smooth compact connected two-dimensional Riemannian manifold (surface), denoted by $\mathcal{S}$. The *minimal geodesics* between

$s_1, s_2 \in \mathcal{S}$ are curves of minimum length on $\mathcal{S}$ connecting s_1 and s_2. The geodesics are denoted by $C^*_{\mathcal{S}}(s_1, s_2)$. The *geodesic distances* refer to the lengths of the minimum geodesics and are denoted by $d_{\mathcal{S}}(s_1, s_2) = \text{length}(C^*_{\mathcal{S}}(s_1, s_2))$. A transformation $\psi : \mathcal{S} \to \mathcal{Q}$ is called an *isometry* if $d_{\mathcal{S}}(s_1, s_2) = d_{\mathcal{Q}}(\psi(s_1), \psi(s_2))$ for all $s_1, s_2 \in \mathcal{S}$. In other words, an isometry preserves the intrinsic metric structure of the surface.

The isometric model, assuming facial expressions to be isometries of some "neutral facial expression", is based on the intuitive observation that the facial skin stretches only slightly. All expressions of a face are assumed to be *intrinsically* equivalent (i.e. have the same metric structure), and *extrinsically* different. Broadly speaking, the intrinsic geometry of the facial surface can be attributed to the subject's identity, while the extrinsic geometry is attributed to the facial expression. The isometric model tacitly assumes that the expressions preserve the *topology* of the surface. This assumption is valid for most regions of the face except the mouth. Opening the mouth changes the topology of the surface by virtually creating a "hole", which was treated in [10] by imposing topological constraints.

The goal of expression-invariant face recognition, under the assumption of the isometric model, is to perform an isometry-invariant matching of facial surfaces. In other words, we are looking for some distance function $d(\mathcal{S}, \mathcal{Q})$ to compare between two facial surfaces $\mathcal{S}$ and $\mathcal{Q}$, such that $d(\mathcal{S}, f(\mathcal{S})) = 0$ for all isometries f of $\mathcal{S}$. Since the geodesic distances are an obvious isometry-invariant, one could think of $d(\mathcal{S}, \mathcal{Q})$ comparing the geodesic distances on $\mathcal{S}$ and $\mathcal{Q}$. However, in practice only *sampled* versions of the surfaces are available, and therefore we have the intrinsic geometry of $\mathcal{S}$ and $\mathcal{Q}$ represented as *finite metric spaces* $(\{s_1, ..., s_N\}, \mathbf{\Delta}_{\mathcal{S}})$ and $(\{q_1, ..., q_M\}, \mathbf{\Delta}_{\mathcal{Q}})$, where the $N \times N$ matrix $\mathbf{\Delta}_{\mathcal{S}} = (d_{\mathcal{S}}(s_i, s_j))$ and the $M \times M$ matrix $\mathbf{\Delta}_{\mathcal{Q}} = (d_{\mathcal{Q}}(q_i, q_j))$ denote the pair-wise geodesic distances (which, in practice, must be computed numerically) between the samples of $\mathcal{S}$ and $\mathcal{Q}$. There is no guarantee that different instances of the same facial surface are sampled at the same points, nor that the number of samples is the same. Moreover, even if the samples are the same, they can be ordered arbitrarily. This ambiguity, which theoretically requires examining all the permutations between the points on the two surfaces, makes impractical the use of the geodesic distances *per se* for isometry-invariant surface matching.

A recent fundamental paper by Mémoli and Sapiro [11] relates the permutation-based distance between surfaces represented as point clouds to the Gromov-Hausdorff distance and shows a probabilistic framework allowing to approximate it without computing all the permutations.

2.1 Isometric embedding

An alternative proposed by Elad and Kimmel [4] and adopted in [3] for face recognition is to avoid dealing explicitly with the matrix of geodesic distances and represent $\mathcal{S}$ as a subset of $\mathbb{R}^m$, such that the original intrinsic geometry is approximately preserved. Such a procedure is called an *isometric embedding*. The image of $\mathcal{S}$ under the embedding is referred to as the *canonical form* of $\mathcal{S}$, and

$\mathbb{R}^m$ as the *embedding space*. As a result of isometric embedding, the canonical forms of all the isometries of $\mathcal{S}$ are identical, up to the isometry group in $\mathbb{R}^m$ (rotations, translations and reflections), which is easy to deal with. The distance $d(\mathcal{S}, \mathcal{Q})$ is computed by comparing the canonical forms of $\mathcal{S}$ and $\mathcal{Q}$ in a rigid way.

In the discrete setting, isometric embedding is a mapping between two finite metric spaces

$$\varphi : (\{s_1, ..., s_N\} \subset \mathcal{S}, \mathbf{\Delta}_{\mathcal{S}}) \rightarrow (\{x_1, ..., x_N\} \subset \mathbb{R}^m, \mathbf{D}), \tag{1}$$

such that $d_{\mathcal{S}}(s_i, s_j) = d_{\mathbb{R}^m}(x_i, x_j)$ for all $i, j = 1, ..., N$. Here $d_{\mathbb{R}^m}$ denotes the Euclidean metric, and $\mathbf{D} = (d_{\mathbb{R}^m}(x_i, x_j))$ is the matrix of pair-wise geodesic distances between the points in the embedding space. In practice, this matrix is computed approximately using the *fast marching method* (FMM) [12].

2.2 Multidimensional scaling

Unfortunately, it appears that a general surface like the human face usually cannot be isometrically embedded into $\mathbb{R}^m$ of any finite dimension [13], and therefore, such an embedding introduces a distortion of the geodesic distances, referred to as the *embedding error*. Yet, though an exact isometric embedding of $\mathcal{S}$ into $\mathbb{R}^m$ does not exist, it is possible to compute an approximately isometric embedding, which minimizes the embedding error. In [4], the *raw stress* [5] was used

$$\sigma_{\text{raw}}(\mathbf{X}; \mathbf{\Delta}_{\mathcal{S}}) = \sum_{i>j} \left(d_{\mathbb{R}^m}(\mathbf{x}_i, \mathbf{x}_j) - d_{\mathcal{S}}(s_i, s_j) \right)^2. \tag{2}$$

as the embedding error criterion. Here $\mathbf{\Delta}_{\mathcal{S}}$ denotes the geodesic distances matrix of the surface $\mathcal{S}$, and $\mathbf{X}$ is a $N \times m$ matrix of coordinates in $\mathbb{R}^m$. The solution

$$\mathbf{X}_{\mathcal{S}} = \underset{\mathbf{X}}{\text{argmin}}\, \sigma_{\text{raw}}(\mathbf{X}; \mathbf{\Delta}_{\mathcal{S}}) \tag{3}$$

obtained by gradient descent minimization of the stress is the discrete canonical form of $\mathcal{S}$, and the whole process is called *multidimensional scaling* (MDS). The optimization problem (3) is non-convex and therefore convex optimization algorithms cannot guarantee global convergence. Yet, this problem can be usually resolved using good intialization or employing multiscale or multigrid optimization [14].

2.3 Canonical form matching

The similarity function between two surfaces $\mathcal{S}$ and $\mathcal{Q}$ in the canonical forms (CF) algorithm is computed as the Euclidean distance between the vector of P-th order high-dimensional moments of the corresponding canonical forms $\mathbf{X}_{\mathcal{S}}$ and $\mathbf{X}_{\mathcal{Q}}$ after alignment [15]

$$d_{\text{CF}}(\mathcal{S}, \mathcal{Q}) = \sum_{p_1 + \dots p_m \leq P} \left(\mu^{\mathbf{X}_{\mathcal{S}}}_{p_1, \dots, p_m} - \mu^{\mathbf{X}_{\mathcal{Q}}\mathbf{R} + \mathbf{b}}_{p_1, \dots, p_m} \right)^2, \tag{4}$$

where

$$\mu^{\mathbf{X}}_{p_1,...,p_m} = \sum_{i=1}^{N} \prod_{j=1}^{m} x_{ij}^{p_j}, \quad (5)$$

is the $(p_1, ..., p_m)$-th moment of $\mathbf{X}$, and $\mathbf{X}_{\mathcal{Q}}\mathbf{R}+\mathbf{b}$ is an m-dimensional Euclidean transformation (rotation, reflection and translation) aligning the canonical forms $\mathbf{X}_{\mathcal{S}}$ and $\mathbf{X}_{\mathcal{Q}}$. In [4], the alignment transformation is obtained by centering the canonical forms, diagonalizing their matrices of second-order moments, and re-ordering the axes such that the variance values in each axis are decreasing. Alternatively, the alignment can be performed using three fiducial points [3].

2.4 Remarks

The CF approach does not allow exact isometry-invariant surface matching, as the embedding is not exactly isometric and inevitably introduces an error. In other words, generally $d_{\mathrm{CF}}(f(\mathcal{S}), \mathcal{S}) > 0$ for a surface $\mathcal{S}$ and its isometry f. The algorithm is sensitive to the definition of the boundaries of the surfaces, and does not allow for matching of surfaces with different topologies, or more generally, partial matching.

The alignment ambiguity and the use of rigid matching algorithms for the canonical forms poses a restriction on the number of the surface samples. It must be sufficiently large ($N \sim 1000$) in order for the alignment and matching to work accurately. In [3], we found that 2500 samples were required for face recognition with a reasonable recognition rate. The number of points is a major issue in terms of computational complexity, as the cost of the stress and its gradient computation is $\mathcal{O}(N^2)$, while the computation of the geodesic distance matrix is at least $\mathcal{O}(N^2)$.

Another major issue in face recognition with large databases is precomputation of distances between faces. Using the method of moments, in the CF approach it is possible to precompute the moments signatures for all faces in the gallery. When a new probe face has to be matched, its moment signature is computed and efficiently matched to the gallery signatures.

3 Generalized multidimensional scaling

The main thesis of this paper is measuring the intrinsic similarity of two facial surfaces by embedding them into each other, based on [9]. Embedding two isometric surfaces into each other results in *zero* embedding error. In the general case when two surfaces are not isometric, the embedding error is a measure of their similarity. Conceptually, the difference between Euclidean and partial embedding is presented in Figure 1.

We assume to be given the model surface $\mathcal{S}$ from the gallery, sampled at N points, and the probe surface $\mathcal{Q}$ sampled at M points (typically, $M \ll N$). Possibly, $\mathcal{Q}$ is partially missing. We are looking for a mapping

$$\varphi : (\{q_1, ..., q_N\} \subset \mathcal{Q}, \mathbf{\Delta}_{\mathcal{Q}}) \rightarrow (\{s_1, ..., s_N\} \subset \mathcal{S}, \mathbf{\Delta}_{\mathcal{S}}),$$

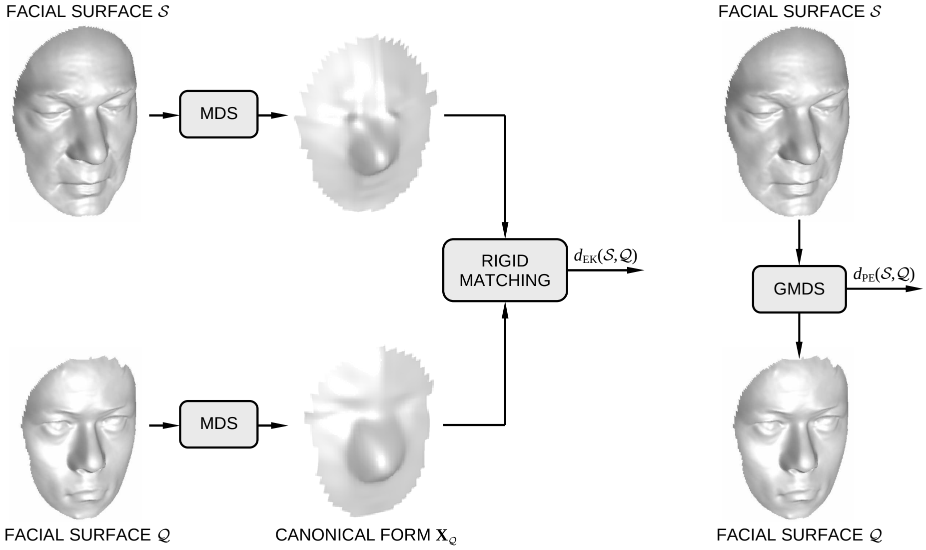

Fig. 1. Schematic representation of face recognition using expression-invariant canonical forms obtained using MDS (left), and the proposed method of embedding one facial surface into another using GMDS (right).

such that $d_{\mathcal{Q}}(q_i, q_j)$ is as close as possible to $d_{\mathcal{S}}(\varphi(q_i), \varphi(q_j))$ for all $i, j = 1, ..., N$. Note that $d_{\mathcal{S}}$ is assumed continuous here, as $s_i = \varphi(q_i)$ can be an arbitrary point. In practice, the values of $d_{\mathcal{S}}$ must be approximated numerically. We refer to such φ as *partial embedding* of $\mathcal{Q}$ into $\mathcal{S}$.

In order to compute the partial embedding, we use a procedure similar to MDS, which we call the *generalized* MDS or GMDS [9]. Since our new embedding space $\mathcal{S}$ is a general 2D manifold, we have to represent the points on $\mathcal{S}$ in their parametric coordinates. Let us assume that $\mathcal{Q}$ is given in parametric from by the mapping $\mathbf{u} \in I \subset \mathbb{R}^2 \to \mathcal{S}$, where I is the parametrization domain, which can be assumed to be $[0, 1] \times [0, 1]$. Similarly to the Euclidean case, the *generalized stress* is defined as [9]

$$\sigma_{\mathrm{gen}}(\mathbf{U}; \mathbf{\Delta}_{\mathcal{Q}}, \mathbf{W}, d_{\mathcal{S}}) = \sum_{i>j} w_{ij} \left(d_{\mathcal{S}}(\mathbf{u}_i, \mathbf{u}_j) - d_{\mathcal{Q}}(q_i, q_j) \right)^2 . \qquad (6)$$

Here $\mathbf{u}_i$, $i = 1, ..., M$ denote the vectors of parametric coordinates of s_i, and $\mathbf{W} = (w_{ij})$ is a symmetric matrix of non-negative weights. In case of full matching, $w_{ij} = 1$ are used. When the probe is partially missing, the weights must be chosen differently [9].

Minimization of the stress is performed iteratively, like in the former case, using gradient-descent type methods or more sophisticated optimization algorithms [16]. Note that, in practice, $d_{\mathcal{S}}$ is available only between N samples of $\mathcal{S}$. Hence, it must be approximated for all the rest of the points. This computation is critical for the GMDS. For that goal we developed the *three-point geodesic distance approximation.* The idea of this numerical procedure is to produce a computationally efficient $\mathcal{C}^1$-approximation for $d_{\mathcal{S}}$ and its derivatives, interpolating their values from the matrix $\mathbf{\Delta}_{\mathcal{S}}$ of pairwise geodesic distances on $\mathcal{S}$. For further details, the reader is referred to [17].

The partial embedding distance function between the probe surface $\mathcal{Q}$ and the model surface $\mathcal{S}$ is defined as

$$d_{\mathrm{PE}}(\mathcal{S},\mathcal{Q}) = \sqrt{\frac{\mathrm{argmin}_{\mathbf{U}}\, \sigma_{\mathrm{gen}}(\mathbf{U};\mathbf{\Delta}_{\mathcal{Q}},\mathbf{W},d_{\mathcal{S}})}{\sum_{i>j} w_{ij}}}, \tag{7}$$

which is independent of the number of points. d_{PE} has length units and can be interpreted as an RMS metric distortion.

3.1 Partial matching

One of the most important properties of d_{PE} is that it allows to perform *partial matching* of surfaces (note that, indeed, $d_{\mathrm{PE}}(\mathcal{S},\mathcal{Q})$ is not symmetric, which allows to embed a patch of $\mathcal{Q}$ into $\mathcal{S}$). Partial matching is important in practical face recognition applications, where imperfections of the acquisition devices and occlusions of the face (e.g. when the subject is wearing glasses) result in a partially missing probe surfaces.

Let us assume that we wish to compare two facial surfaces: a model $\mathcal{S}$ and a probe $\mathcal{Q}$, which is acquired with occlusions such that only a patch $\mathcal{Q}' \subset \mathcal{Q}$ is available. If $\mathcal{Q}'$ is sufficiently large, $d_{\mathrm{PE}}(\mathcal{S},\mathcal{Q}') \approx d_{\mathrm{PE}}(\mathcal{S},\mathcal{Q})$; the difference can be bounded by the diameter of $\mathcal{Q}\setminus\mathcal{Q}'$ [9]. Yet, it is tacitly assumed that the geodesic distances on $\mathcal{Q}'$ are given by $d_{\mathcal{Q}'}(q_1,q_2) = d_{\mathcal{Q}}|_{\mathcal{Q}'}(q_1,q_2)$ (this notation implies that $d_{\mathcal{Q}'}(q_1,q_2) = d_{\mathcal{Q}}(q_1,q_2)$ for all $q_1,q_2 \in \mathcal{Q}'$). However, $d_{\mathcal{Q}'}$ is computed numerically on $\mathcal{Q}'$ and can be inconsistent with $d_{\mathcal{Q}}|_{\mathcal{Q}'}$. The problem potentially arises for example with geodesics that touch the boundary $\partial\mathcal{Q}'$; such geodesics can be different on $\mathcal{Q}$ and $\mathcal{Q}'$ (see Figure 2), and the corresponding distance is therefore inconsistent. To resolve this problem, we assign zero weight $w_{ij} = 0$ to every pair of points (q_i,q_j) such that $d_{\mathcal{Q}'}(q_i,\partial\mathcal{Q}') + d_{\mathcal{Q}'}(q_j,\partial\mathcal{Q}') < d_{\mathcal{Q}'}(q_i,q_j)$ For more details, the reader is referred to [17].

3.2 Comparison to the canonical forms approach

A major difference of the CF and the PE algorithms is that in the former, isometric embedding is used only as an intermediate stage to obtain an isometry-invariant representation of the surfaces, whereas in our approach isometric embedding is used directly to compute the similarity between surfaces. The consequences of this difference are several. First, the *codimension* of the canonical form in the embedding space is at least one. In PE, the codimension is always zero. Secondly, embedding into Euclidean space still leaves the degrees of freedom of an Euclidean isometry (rotation, translation and reflection). In embedding into a general surface, if it is rich enough, such ambiguity usually does not exist.

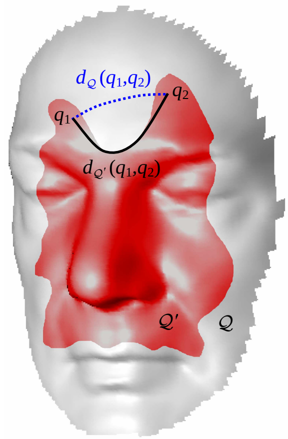

Fig. 2. Partial matching problem. Shown in blue dotted is a geodesic between the points $q_1, q_2 \in \mathcal{Q}$; the corresponding inconsistent geodesic on $\mathcal{Q}'$ is shown in black.

Thirdly, this ambiguity requires alignment of the CF canonical forms, which is avoided in PE. Due to the fact that the metric distortion serves as a dissimilarity measure in PE rather than a side effect (as in CF), a small number of surface samples suffices for accurate matching, and practically, as few as tens of points were enough in all of our face recognition experiments.

Another major issue is preprocessing. The performance of the CF approach depends heavily on the facial surface preprocessing, since it is important that the probe and the model surfaces contain the same region of the face. In [3], a *geodesic mask* was used to crop the facial surfaces. The problem is especially acute if one wishes to handle expressions with open mouth and uses a topological constraint by cutting off the lips [10]. The PE approach, on the other hand, is insensitive to preprocessing, since it allows partial matching. Practically, the probe can be an arbitrary patch of the model surface.

Since d_{PE} between any two faces is computed iteratively, it is impossible to precompute it as in the CF approach. However, it is still possible to speed-up the matching significantly using a hierarchial comparison. We address this issue in Section 4. The comparison of the PE and the CF approaches is summarized in Table 1.

4 Hierarchial matching

An apparent limitation of the proposed face recognition method stems from the fact that, unlike the CF approach that allows to match canonical forms using moment signatures, the partial embedding distance cannot be precomputed. ¿From this point of view, our approach is similar to methods proposing the use of the iterative closest point algorithm (ICP). Taking, for example, about 1 *sec* per

Table 1. Comparison of partial embedding and the canonical forms algorithm

	Canonical forms	*Partial Embedding*
Accuracy	up to minimum distortion caused by the embedding space	up to numerical errors
Distance function	moments or ICP	embedding error is used as distance
Alignment	required to resolve rotation, translation and reflection ambiguity	no alignment ambiguity
Partial matching	difficult	natural
Precomputation	possible using moment signatures	possible to some extent using hierarchial matching
Samples	~ 1000	$10 \sim 100$
Preprocessing	requires geometrically-consistent cropping of the facial surfaces using geodesic mask; particularly sensitive to lip cropping in case of open-mouth expressions	the probe can be an arbitrary patch of the facial surface
Numerical core	FMM, MDS	FMM, GMDS

comparison, matching a probe to a gallery of $100,000$ faces would take about 30 hours on a single CPU. Such computational complexity makes the one-to-many face recognition scenario infeasible. However, our method can still be used for one-to-many face recognition with very large databases by taking advantage of a hierarchical matching scheme, which is briefly outlined here.

Let the gallery database consist of K_0 faces, $\left\{\mathcal{S}_1^0, ..., \mathcal{S}_{K_0}^0\right\}$. We aggregate groups of faces close in the sense of d_{PE}, replacing them with a single representative, as usually done in vector quantization [18]. The number of faces forming an aggregate can be either constant or adaptive, and depends on the specific aggregation algorithm used. As a result, a smaller set $\left\{\mathcal{S}_1^1, ..., \mathcal{S}_{K_1}^1\right\}$ is obtained. Repeating the procedure iteratively, a tree-like structure is created, where at the top level there is a relatively small set $\left\{\mathcal{S}_1^L, ..., \mathcal{S}_{K_L}^L\right\}$ of representative faces. Here L denotes the number of levels in the tree. Such hierarchial representation can be computed off-line once for a given database. Adding new faces to it can be made very efficient using techniques from heap and sorting trees, requiring $\mathcal{O}(\log K)$ comparisons.

Hierarchical comparison of a probe face to the entire database is performed in a top-down manner: first, the probe $\mathcal{Q}$ is compared to K_L top-level faces $\left\{\mathcal{S}_i^L\right\}$; $\mathcal{S}_i^L$ minimizing $d_{\mathrm{PE}}\left(\mathcal{S}_i^L, \mathcal{Q}\right)$ is selected, and the probe is compared to its subtree. The process is repeated until the lowest level is reached. Such a scheme allows to perform face recognition with only $\mathcal{O}(\log K)$ matches and is suitable for one-to-many comparison scenarios with very large database of faces.

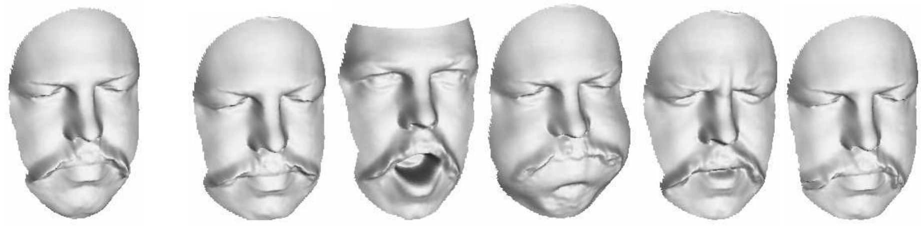

Fig. 3. Gallery (leftmost) and four probe faces of a representative subject in the database.

5 Results

The presented face recognition algorithm was tested on a set of 30 subjects from the Notre Dame 3D face database used in the FRGC competition [1, 2]. The gallery consisted of one neutral expression per subject; five instances with moderate facial expressions were used as probes for each subject, yielding the total of 180 faces (see Figure 3). Gallery faces were cropped with a wide rectangular mask, which included most of the facial surface. These surfaces were subsequently sampled on a regular Cartesian grid consisting of approximately 2500 points. Pairwise geodesic distances between these points were measured using an efficient modification of parametric FMM [12, 12, 19] requiring about 1 *sec* for computing a 2500×2500 distance matrix. Two sets of probes were created: in the first experiment, the probe faces were cropped using a narrow geodesic mask, which excluded hair and other unrelated details, covering most of the relevant parts of the face (Figure 4, left). In the second experiment, random parts of the surface were intentionally removed, resulting in severe occlusions of the facial surface (see example in Figure 4, right). In both experiments, the surfaces were sampled at 53 points using *furthest point sampling* strategy [20].

Face recognition was carried out by embedding the probe surface into the gallery surface using GMDS[1]; d_{PE} served as a dissimilarity measure between the faces. Figure 5 depicts the receiver operator characteristic (ROC) curves obtained in the two experiments. Comparison of mildly and severely occluded faces resulted in about 3.1% and 5.5% equal-error rate (EER), respectively. In both experiments 100% rank-1 recognition rate was achieved. Our non-optimized C code required about $1 \div 5$ *sec* per surface comparison.

[1] The GMDS MATLAB implementation will be available for download from `http://tosca.cs.technion.ac.il` as a part of the TOSCA (Toolbox for Surface Comparison and Analysis) Project.

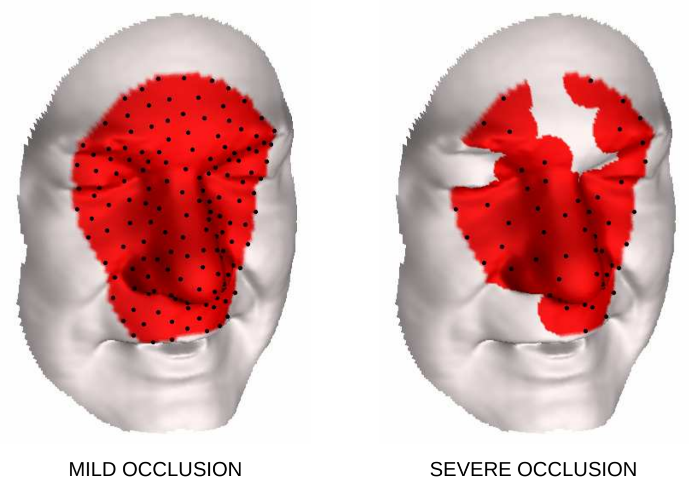

Fig. 4. Left: probe face with mild occlusions; right: the same probe face with severe occlusions. Surface samples are denoted by black dots.

6 Conclusions

Following the isometric model of facial expression introduced in [3], we proposed a novel expression-invariant face recognition algorithm. The main idea of our approach is to embed the probe facial surface into that of the model. Faces belonging to the same subject are nearly isometric and thus result in low embedding error, whereas different subjects are expected to have different intrinsic geometry, and thus produce higher embedding error. Unlike the previous approaches, our method does not introduce unnecessary metric distortions due to the embedding itself. Moreover, the probe and the model are not required to contain the same amount of information; matching a probe with partially missing data is natural to our approach. To the best of our knowledge, it is the first method to allow partial isometry-invariant matching of surfaces in general and of facial geometry in particular.

The numerical core of our face recognition method is the GMDS procedure, which has the same computational complexity as that of the standard MDS procedure. Although our algorithm does not permit pre-computation of simple efficiently comparable signatures, we outlined a hierarchical matching strategy that enables the use of our approach for one-to-many face recognition in large databases.

Promising face recognition results were obtained on a small database of 30 subjects even when the facial surfaces were severely occluded. In sequel studies, we intend to demonstrate the performance of our approach on larger databases with extreme facial expression. Noting that GMDS is capable of finding intrin-

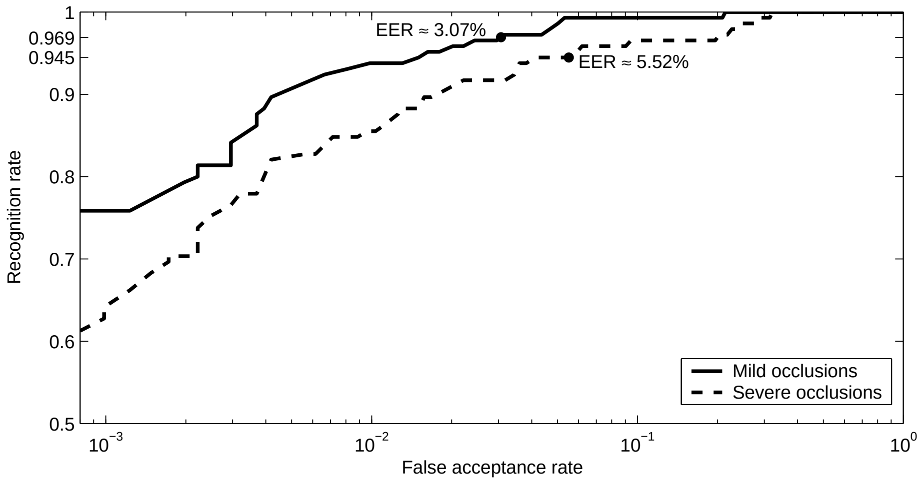

Fig. 5. ROC curves obtained in the face recognition experiment with mild (solid) and severe (dashed) occlusions.

sic correspondence between two facial surfaces, our approach can be readily extended to handle texture as well, similarly to [21].

References

1. K. Chang, K. W. Bowyer, and P. J. Flynn. Face recognition using 2D and 3D facial data. In *ACM Workshop on Multimodal User Authentication*, pages 25–32, 2003.
2. P. J. Flynn, K. W. Bowyer, and P. J. Phillips. Assessment of time dependency in face recognition: an initial study. In *Audio- and Video-Based Biometric Person Authentication*, pages 44–51, 2003.
3. A. M. Bronstein, M. M. Bronstein, and R. Kimmel. Three-dimensional face recognition. *Intl. J. Computer Vision*, 64(1):5–30, August 2005.
4. A. Elad and R. Kimmel. On bending invariant signatures for surfaces. *IEEE Trans. PAMI*, 25(10):1285–1295, 2003.
5. I. Borg and P. Groenen. *Modern multidimensional scaling - theory and applications*. Springer-Verlag, Berlin Heidelberg New York, 1997.
6. J. Walter and H. Ritter. On interactive visualization of high-dimensional data using the hyperbolic plane. In *Proc. ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, pages 123–131, 2002.
7. A. M. Bronstein, M. M. Bronstein, and R. Kimmel. On isometric embedding of facial surfaces into $\mathbb{S}^3$. In *Proc. Int'l Conf. Scale Space and PDE Methods in Computer Vision*, number 3459 in Lecture Notes on Computer Science, pages 622–631. Springer, 2005.
8. A. M. Bronstein, M. M. Bronstein, and R. Kimmel. Expression-invariant face recognition via spherical embedding. In *Proc. ICIP*, 2005.

9. A. M. Bronstein, M. M. Bronstein, and R. Kimmel. Generalized multidimensional scaling: a framework for isometry-invariant partial surface matching. *Proc. Nat. Acad. Sci.*, 103(5):1168–1172, January 2006.
10. A. M. Bronstein, M. M. Bronstein, and R. Kimmel. Expression-invariant representations for human faces. Technical Report CIS-2005-01, Dept. of Computer Science, Technion, Israel, June 2005.
11. F. Mémoli and G. Sapiro. A theoretical and computational framework for isometry invariant recognition of point cloud data. *Foundations of Computational Mathematics*, 2005. to appear.
12. R. Kimmel and J. A. Sethian. Computing geodesic on manifolds. In *Proc. US National Academy of Science*, volume 95, pages 8431–8435, 1998.
13. N. Linial, E. London, and Y. Rabinovich. The geometry of graphs and some its algorithmic applications. *Combinatorica*, 15:333–344, 1995.
14. M. M. Bronstein, A. M. Bronstein, R. Kimmel, and I. Yavneh. A multigrid approach for multidimensional scaling. In *Proc. Copper Mountain Conf. Multigrid Methods*, 2005.
15. A. Tal, M. Elad, and S. Ar. Content based retrieval of VRML objects - an iterative and interactive approach. In *Proc. Eurographics Workshop on Multimedia*, 2001.
16. D. Bertsekas. *Nonlinear programming.* Atlanta Scientific, 2 edition, 1999.
17. A. M. Bronstein, M. M. Bronstein, and R. Kimmel. Effcient computation of the Gromov-Hausdorff distance for smooth surfaces. Technical Report CIS-2006-01, Dept. of Computer Science, Technion, Israel, January 2006.
18. A. Gersho and R. M. Gray. *Vector quantization and signal compression.* Kluwer Academic Publishers, Boston, 1992.
19. A. Spira and R. Kimmel. An efficient solution to the eikonal equation on parametric manifolds. *Interfaces and Free Boundaries*, 6(3):315–327, 2004.
20. Y. Eldar, M. Lindenbaum, M. Porat, and Y. Y. Zeevi. The farthest point strategy for progressive image sampling. *IEEE Trans. Image Processing*, 6(9):1305–1315, 1997.
21. G. Zigelman, R. Kimmel, and N. Kiryati. Texture mapping using surface flattening via multi-dimensional scaling. *IEEE Trans. Visualization and computer graphics*, 9(2):198–207, 2002.