Expression-invariant representations of faces

Alexander M. Bronstein, *Student Member, IEEE,* Michael M. Bronstein, *Student Member, IEEE,* and Ron Kimmel, *Senior Member, IEEE*

***Abstract*— Addressed here is the problem of constructing and analyzing expression-invariant representations of human faces. We demonstrate and justify experimentally a simple geometric model that allows to describe facial expressions as isometric deformations of the facial surface. The main step in the construction of expression-invariant representation of a face involves embedding of the facial intrinsic geometric structure into some convenient low-dimensional space. We study the influence of the embedding space geometry and dimensionality choice on the representation accuracy and argue that compared to its Euclidean counterpart, spherical embedding leads to notably smaller metric distortions. We experimentally support our claim showing that a smaller embedding error leads to better recognition.**

Keyword: Isometry-invariant representation, multidimensional scaling, isometric embedding, face recognition, spherical harmonics.

I. Introduction

Expression-invariant features of the human face play an important role in various applications. Such are face recognition in computer vision [1], [2], texture mapping for facial animation in computer graphics [3], [4], [5], [6], [7], emotion interpretation in psychology [8], and measurement of geometric parameters of the face in cosmetic surgery. The variability of the face appearance due to its non-rigid structure makes these tasks non-trivial and challenges for a convenient model to analyze the nature of facial expressions.

In face recognition, the problem of expression-invariant representation is of particular importance as most of today's face recognition systems are sensitive to facial expression variations. The development of robust face recognition algorithm insensitive to facial expression is one of the greatest challenges of current research in this field [9], [10], [11], [12]. Several approaches have been proposed for this purpose. One possibility is to use only regions of the face the least susceptible to variations due to facial expression [13]. For example, one can remove the mouth region as it varies greatly with expression. Yet, practice shows that there is no large subset of the face that is perfectly rigid across a broad range of expressions [14]

Another possibility is to add different expressions of each subject into the gallery. The main limitation is the large number of instances of each subject due to the richness of all the possible expression. Even though, the probe expression may still be substantially different from those sampled.

A third approach is a model of facial expression used to generate synthetic facial expressions for any 3D surface. Such models are usually very complicated and therefore, not suitable for the face recognition application, where real- or near-real time performance is required. Moreover, it is practically impossible to capture the subtle individual differences of expressions in different subjects.

The present paper focuses on a simple geometric model of facial expressions, referred to as the *isometric model* and expression-invariant representation of the face based on this model. Our thesis is that facial expressions can be modelled as isometries of the facial surface. A simplified setting of this model assuming all expressions to be with closed mouth was used in our three-dimensional expression-invariant face recognition system [2], [15], [16], [17]. Here, we extend this model to handle both open and closed mouth.

Using the isometric model, expression-invariant representation of the face is essentially equivalent to isometry-invariant representation of a surface. We apply the isometric embedding approach, introduced by Elad and Kimmel [18]. The key idea is to represent the intrinsic metric structure of the facial surface by embedding the surface into a low-dimensional Euclidean space and replacing the geodesic distances by Euclidean ones. The resulting representation, known as *canonical form*, can then be treated by conventional methods for rigid surface matching.

The advantage of this approach over previously proposed methods is that we do not attempt *generating* facial expressions, but rather create a representation which is (approximately) invariant to facial expressions. Secondly, we use the whole facial surface without removing parts sensitive to expressions. Thirdly, the gallery contains only a single instance of a subject, e.g. the neutral expression. Finally, our algorithm is computationally efficient and can be performed in near real-time.

We start with a description of the isometric model and its experimental validation in Section II. In Section III, we present the Elad-Kimmel isometry-invariant surface matching approach in a broader perspective, for both Euclidean and non-Euclidean embedding spaces. We address some criteria for the choice of the embedding space, and present the computational core of the approach, based on multidimensional scaling. Section IV deals with facial geometry representation in three-dimensional spaces with Euclidean and spherical geometry. Embedding into a three-dimensional sphere ($\mathbb{S}^3$) introduces smaller distance distortion, and consequently, yields better recognition. We provide an experimental evidence of this fact. Section V deals with incorporating the texture information using embedding into a two-dimensional space, specifically discussing the two-dimensional sphere $\mathbb{S}^2$. Finally, Section VI concludes the paper.

Manuscript revised March 7, 2006.

A. M. Bronstein, M. M. Bronstein and R. Kimmel are with the Department of Computer Science, Technion – Israel Institute of Technology.

II. Isometric model of facial expressions

We describe a facial surface as a smooth compact connected two-dimensional Riemannian manifold (surface), denoted by $\mathcal{S}$. The geodesic distances (shortest path lengths) on $\mathcal{S}$ are induced by the Riemannian metric and are denoted by $d_{\mathcal{S}}(\xi_1, \xi_2)$ for $\xi_1, \xi_2 \in \mathcal{S}$. A transformation $\psi : \mathcal{S} \to \mathcal{S}'$ is called an *isometry* if

$$d_{\mathcal{S}}(\xi_1, \xi_2) \quad = \quad d_{\mathcal{S}'}(\psi(\xi_1), \psi(\xi_2)), \tag{1}$$

for all $\xi_1, \xi_2 \in \mathcal{S}$. 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 (approximately) *intrinsically* equivalent (i.e. have the same metric structure). 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.

However, the isometric model also requires the *topology* of the surface to be preserved. 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." Fig. 1 demonstrates this particular property by showing a minimal geodesic between two points on the upper and the lower lips. In [2], [15], we ignored this possibility, assuming all the expressions to be with closed mouth. Even in such a limited setting, the isometric model appeared to be able to gracefully represent strong facial expressions. In order to be able to handle expressions with open and closed mouth, we propose fixing the topology of the surface. We could enforce the mouth to be always closed by "gluing" the lips when the mouth is open; alternatively, we can constrain the mouth to be open and disconnecting the lips by means of a cut in the surface when the mouth is closed. Here, the later option is adopted.

A. Validation of the isometric model

In order to validate the isometric model, we tracked a set of feature points on the facial surface and measured how distances between them change due to expressions, while preserving the topology of the surface. In this experiment, 133 white round markers (approximately $2\,mm$ in diameter) were placed on the face of a subject as invariant fiducial points (Fig. 2, left). The subject was asked to demonstrate different facial expressions of varying strength with open and closed mouth (total of 16 expressions; see some examples in Fig. 2, right). The faces in this and all the following experiments were acquired using a coded-light scanner with acquisition time of about $150\,msec$ and depth resolution of $0.5\,mm$ [19]. As the reference surface, we used a "neutral expression." The lips were manually cropped in all the surfaces. Fast Marching Method (FMM) [20] was used to measure the geodesic distances.

The changes of the distances due to facial expressions were quantified using two measures: the absolute and the relative error with respect to the reference distances. Relative and absolute error distributions are plotted in Fig. 3. The standard deviation of the absolute and the relative errors were $5.89\,mm$ and 15.85%, respectively, for geodesic distances, and $12.03\,mm$ and 39.62%, respectively, for the Euclidean ones. We conclude that the changes of the geodesic distances due to facial expressions are insignificant even for extreme expressions, which justifies our model. Moreover, the Euclidean distances were demonstrated to be much more sensitive to changes due to facial expressions compared to the geodesic ones. This observation will be reinforced in Section IV-A, where we compare our approach based on geodesic distances and straightforward rigid matching of facial surfaces.

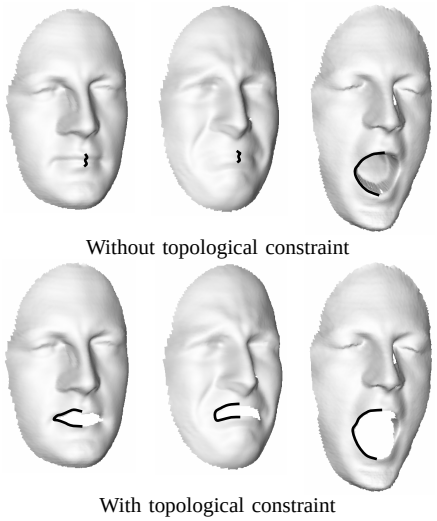

Fig. 1. Illustration of the open mouth problem. First row: Geodesics on the facial surface with three facial expressions. An open mouth changes the topology of the surface, which leads to loosing the consistency of the geodesics. Second row: The same face, in which the topology is constrained by cutting out the lips. Consistency of the geodesics is better preserved.

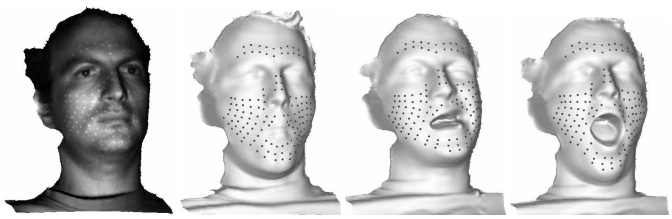

Fig. 2. Isometric model validation experiment. Left: facial image with the markers. Right: example of one moderate and two strong facial expressions with marked reference points.

III. Isometry-invariant representation

Under the assumption of the isometric model, in order to obtain an expression-invariant representation of the face $\mathcal{S}$, we would like to keep only the *intrinsic* metric structure of $\mathcal{S}$. Equivalently, we would like to somehow ignore the *extrinsic geometry*, that is, the way the surface $\mathcal{S}$ is immersed into the ambient space. An obvious invariant to isometries is the set of all geodesic distances on the surface $\mathcal{S}$. In practice, $\mathcal{S}$ is given at a finite set of points $\{\xi_1, ..., \xi_N\}$, and the geodesic distances

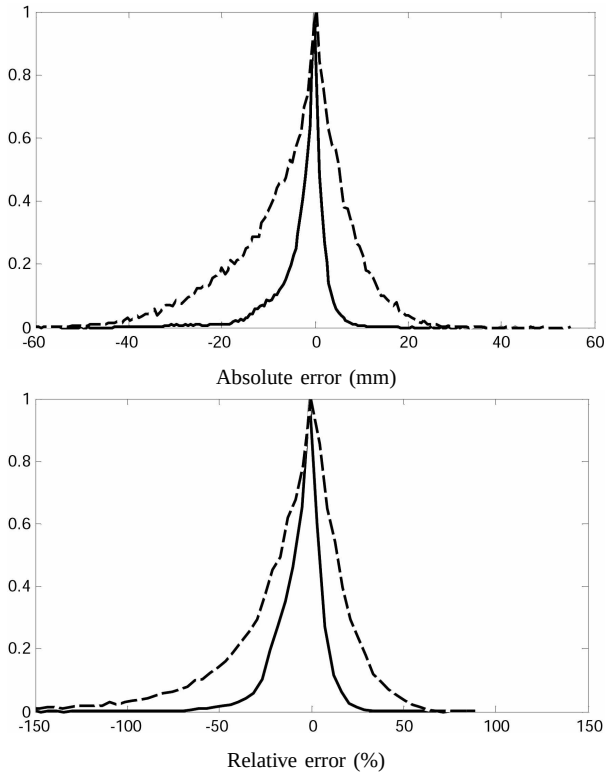

Fig. 3. Normalized histograms of the absolute and the relative error of the geodesic (solid) and the Euclidean (dotted) distances.

$d_{\mathcal{S}}(\xi_i, \xi_j)$ are computed approximately from the samples of $\mathcal{S}$ using FMM and represented as an $N \times N$ matrix $\mathbf{\Delta}$.

Yet, comparing explicitly the geodesic distances is difficult since the correspondence between the samples of two surfaces is unknown. As a solution, Elad and Kimmel [18] proposed to represent $\mathcal{S}$ as a subset of some convenient m-dimensional space $\mathcal{Q}^m$, such that the original intrinsic geometry is approximately preserved. Such a procedure is called an *(almost) isometric embedding*, and allows to get rid of the extrinsic geometry. As the result of isometric embedding, the representations of all the isometries of $\mathcal{S}$ are identical, up to the isometry group in $\mathcal{Q}^m$, which is usually easy to deal with (for example, all possible isometries in $\mathbb{R}^3$ are rotations, translations and reflections.)

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

$$\varphi : (\{\xi_1, ..., \xi_N\} \subset \mathcal{S}, \mathbf{\Delta}) \to (\{x_1, ..., x_N\} \subset \mathcal{Q}^m, \mathbf{D}),$$

such that $d_{\mathcal{S}}(\xi_i, \xi_j) \approx d_{\mathcal{Q}^m}(x_i, x_j)$ for all $i, j = 1, ..., N$. The matrices $\mathbf{\Delta} = (\delta_{ij}) = (d_{\mathcal{S}}(\xi_i, \xi_j))$ and $\mathbf{D} = (d_{ij}) = (d_{\mathcal{Q}^m}(x_i, x_j))$ denote the pair-wise geodesic distances between the points in the original and the embedding space, respectively. Elad and Kimmel [18] called the image $\{x_1, ..., x_N\} = \varphi(\{\xi_1, ..., \xi_N\})$ the *canonical form* of $\mathcal{S}$. We emphasize that in general, such an isometric embedding does not exist since embedding inevitable introduces metric distortion. As the result, the canonical form is just an *approximate* representation of the discrete surface. Yet, it is possible to find optimal canonical forms by minimizing some *embedding error* criterion (see Section III-B).

Originally, Elad and Kimmel transformed the isometry-invariant deformable surface matching problem into a problem of rigid surface matching in $\mathbb{R}^3$. However, for non-trivial surfaces there is usually an unavoidable representation error resulting from the theoretical impossibility to embed a curved surface into a Euclidean space [21]. The canonical representation of a face is therefore degraded by two types of error: one stemming from the deviation from the isometric model of facial expressions, and the other caused by embedding distortion. While the first error is a limitation of our model, the second one can be reduced by choosing a better embedding space. An important question is whether such a choice will also lead to better recognition.

There are two important aspects to be addressed. First, the *geometry* of the embedding space is important. Certain geometries appear to be more suitable for certain types of surfaces, and thus allow smaller embedding errors [22], [6], [23], [24]. For practical reasons, the geodesic distances in the embedding space should be efficiently computable.[1] Secondly, the *codimension* of the embedding is important. In case of non-zero codimension ($m \geq 3$), embedding produces a canonical form that can be plotted as a surface in the corresponding embedding space. Yet, in this way we take into consideration only the geometry of the face and do not make explicit use of the photometric information (*texture*), which may contain important information that can help to discriminate between faces.

A way to incorporate the texture, described as a scalar field $\alpha : \mathcal{S} \to \mathbb{R}$, is to perform a zero-codimension embedding of $\mathcal{S}$ into some two-dimensional space ($m = 2$). Such an embedding can be thought of as *canonical parameterization* of the facial surface. Alternatively, one can think of the embedding as a warping of the texture image, producing a *canonical image* $a = \varphi(\alpha)$ [2]. Since the embedding approximately preserves the intrinsic geometry of the facial surface, the canonical image is expression-invariant. The difference between these two cases is exemplified in Section IV where we show embedding into $\mathbb{R}^3$ and $\mathbb{S}^3$ and in Section V, where embedding into $\mathbb{S}^2$ is shown.

A. Preprocessing

The face canonization procedure consists of three stages: preprocessing, embedding and matching. The aim of the first stage is to crop a region of interest out of the facial surface, which will be later used for the canonical form computation. First, the surface undergoes standard processing for scanning artifact removal (spike removal, hole filling, smoothing, etc). Then, the facial contour is cropped using the *geodesic mask* [15]. The main idea is to compute a geodesically-equidistant contour of a fixed radius from some set of anchor points (e.g. nose tip and the sellion point) and cut all the points outside this contour. If a topological constraint is employed to allow for

[1] A simple analytic expression for geodesic distances is available for spaces with Euclidean, spherical and hyperbolic geometry. In [25], [26] we show an efficient interpolation procedure for geodesic distances, thereby, the geodesic distances need not necessarily to be expressed analytically.

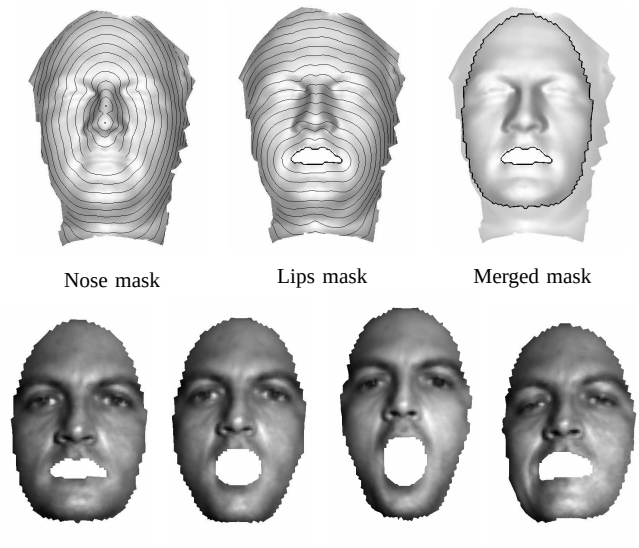

Fig. 4. First row, left to right: geodesic mask computation stages in case of the topologically-constrained isometric model assumption (open and closed mouth): geodesic circles around the nose and the sellion; geodesic circles around the lips; the merged geodesic mask. Second row: examples of the geodesic mask insensitivity to facial expressions.

expressions with both open and closed mouth, two geodesic masks are computed: one around the nose tip and the sellion and another around the lips (Fig. 4, first row). The lips region can be found using standard methods, e.g. according to the texture information. A radius of $30\,mm$ is typically used for the lip mask. The two masks are then merged together. This way, we crop the facial surface in a geometrically consistent manner, which is insensitive to facial expressions (Fig. 4, second row).

B. Multidimensional scaling

The second stage is the embedding. We assume the smoothed and cropped surface to be sampled at N points $\{\xi_1, ..., \xi_N\}$ (typically, N ranges between 1000 and 3000), and assume to be given as input the matrix of geodesic distances $\boldsymbol{\Delta}$ between the surface samples. Every point x in the embedding space $\mathcal{Q}^m$ is assumed to be corresponding to some m-dimensional parametric coordinates $\boldsymbol{\theta} = (\theta^1, ..., \theta^m)$. The geodesic distances are given by some explicit or efficiently approximated function $d_{\mathcal{Q}^m}(x_1, x_2) = d_{\mathcal{Q}^m}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$. The canonical form is computed by the following optimization problem

$$\boldsymbol{\Theta} = \underset{\boldsymbol{\Theta}}{\operatorname{argmin}}\, \sigma(\boldsymbol{\Theta}; \boldsymbol{\Delta}), \tag{2}$$

known as *multidimensional scaling* (MDS) [27]. Here $\boldsymbol{\Theta} = (\boldsymbol{\theta}_1, ..., \boldsymbol{\theta}_N)$ is an $N \times m$ matrix of parametric coordinates in $\mathcal{Q}^m$ corresponding to the points $\{x_1, ..., x_N\}$. In case of Euclidean embedding, the parametric coordinates in $\mathbb{R}^m$ coincide with the Cartesian coordinates in $\mathbb{R}^m$. Therefore, we can refer to $\{x_1, ..., x_N\}$ directly by their Cartesian coordinates, which we denote by $\mathbf{X} \equiv \boldsymbol{\Theta}$. In the non-Euclidean case, if the embedding space $\mathcal{Q}^m$ can be realized as a geometric place in some higher-dimensional Euclidean space $\mathbb{R}^{m'}$ (with $m' > m$; for example, $\mathbb{S}^3$ is represented as a unit sphere in $\mathbb{R}^4$), we will use the $N \times m'$ matrix $\mathbf{X}$ to denote the coordinates in $\mathbb{R}^{m'}$ corresponding to $\boldsymbol{\Theta}$.

The function σ is an embedding error criterion, measuring the discrepancies between the original (geodesic) distances and the embedding space distances. Most common choices of σ include the *raw stress*

$$\sigma_{\mathrm{raw}}(\boldsymbol{\Theta}; \boldsymbol{\Delta}) = \sum_{i>j} \left(d_{ij}(\boldsymbol{\Theta}) - \delta_{ij}\right)^2, \tag{3}$$

and the *normalized stress*

$$\sigma_{\mathrm{norm}}(\boldsymbol{\Theta}; \boldsymbol{\Delta}) = \frac{\sum_{i>j} \left(d_{ij}(\boldsymbol{\Theta}) - \delta_{ij}\right)^2}{\sum_{i>j} d_{ij}^2(\boldsymbol{\Theta})}. \tag{4}$$

Here $d_{ij}(\boldsymbol{\Theta})$ and δ_{ij} are a short notation for $d_{\mathcal{Q}^m}(\boldsymbol{\theta}_i, \boldsymbol{\theta}_j)$ and $d_{\mathcal{S}}(\xi_i, \xi_j)$, respectively. For additional embedding error criteria, see [27].

Note that in fact we compute directly the image $\varphi(\{\xi_1, ..., \xi_N\})$ rather than φ itself. Also note that the optimization problem is non-convex and using convex optimization algorithms is liable to converge to a local rather than global minimum. Nevertheless, convex optimization are commonly used in the MDS literature [27]. Having a good initialization or using multiscale or multigrid optimization [28] reduces the risk of local convergence. Usually, first-order (gradient-descent) optimization algorithms are employed for large scale MDS problems. The computational complexity of such methods grows as $\mathcal{O}(N^2)$.

C. Canonical form matching

The final stage is matching of canonical forms. In case of geometry-only representation (non-zero codimension), we assume to be given two canonical forms $\{x_1, ..., x_N\} = \varphi(\{\xi_1, ..., \xi_N\})$ and $\{x'_1, ..., x'_{N'}\} = \varphi'(\{\xi'_1, ..., \xi'_{N'}\})$, corresponding to two surfaces $\mathcal{S}$ and $\mathcal{S}'$. The canonical forms are represented in the parametric coordinates of $\mathcal{Q}^m$ by the matrices $\boldsymbol{\Theta}$ and $\boldsymbol{\Theta}'$. The goal of matching is to compute some distance $d(\boldsymbol{\Theta}, \boldsymbol{\Theta}')$ between the canonical forms.

First of all, the matching should account for the fact that the canonical forms have some unresolved degrees of freedom and are defined up to an isometry in $\mathcal{Q}^m$. The most straightforward approach is to use an *iterative closest point* (ICP) algorithm [29], [30]. These algorithms minimize a distance between two point clouds over all isometries in $\mathcal{Q}^m$, i.e. find an optimal rigid alignment between $\boldsymbol{\Theta}$ and $\boldsymbol{\Theta}'$. The ICP problem is well-studied in $\mathbb{R}^3$, and practically untouched for non-Euclidean spaces. Moreover, ICP algorithms are usually computationally intensive and thus comparison between a large number of canonical forms, which is typical, for example, in one-to-many face recognition applications with large databases, becomes prohibitive.

As an alternative, the *method of moments* can be used [31]. In case of Euclidean embedding, the distance between two canonical forms represented by the Cartesian coordinates $\mathbf{X}$ and $\mathbf{X}'$ is computed as the difference of their P-th order

moments

$$d_{\mathrm{M}}(\mathbf{X},\mathbf{X}') = \sum_{p_1+...+p_m\leq P} \left(\mu^{\mathbf{X}}_{p_1,...,p_m} - \mu^{\mathbf{X}'}_{p_1,...,p_m}\right)^2, \quad (5)$$

where

$$\mu^{\mathbf{X}}_{p_1,...,p_m} = \sum_{i=1}^{N}\prod_{k=1}^{m} x_{ik}^{p_k}, \quad (6)$$

is the m-dimensional moment of $\mathbf{X}$ for any $p_1,...,p_m \geq 0$. A somewhat better approach is to consider the canonical form as a two-dimensional triangulated mesh rather than just a cloud of points, and compute the moments according to

$$\mu^{\mathbf{X}}_{p_1,...,p_m} = \sum_{i=1}^{T} s_i \prod_{k=1}^{m} \overline{x}_{ik}^{p_k}, \quad (7)$$

where T denotes the number of triangular faces, and $\overline{x}_i$ and s_i stand for the centroid and the area of the i-th triangle, respectively. It is assumed that the canonical forms are first aligned [18]. In the non-Euclidean case, if $\mathcal{Q}^m$ is a subset of some $\mathbb{R}^{m'}$, we can apply the method of moments to the Cartesian m'-dimensional coordinates $\mathbf{X}$ and $\mathbf{X}'$. Alternatively, one can define the moments directly in $\mathcal{Q}^m$, and compute them in the parameterization domain. The exact way of defining such moments depends on the specific choice of the embedding space and is usually non-trivial.

In case of geometry and texture representation (zero codimension), we assume in addition to be given two texture images $\alpha, \alpha' : \mathcal{S} \to \mathbb{R}$. The embeddings φ and φ' map N points of α and N' points of α' to $\mathcal{Q}^m$; the other points are interpolated. As the result, we have two canonical images $a, a' : \mathcal{Q}^2 \to \mathbb{R}$. The matching in this case is performed between the canonical images rather than canonical forms. One can compare the images a and a' directly, e.g. using correlation, or, as an alternative, apply some transform $\mathcal{T} : a \to \hat{a}$ and compare $\hat{a}$ and $\hat{a}'$. For example, in [15], we proposed to project the Euclidean canonical images onto an eigenspace created in a manner similar to eigenfaces [32], [33]. In Section V we show the use of the spherical harmonic transform for comparison of 2D spherical canonical images. An estimate of the albedo (reflection coefficient) can be used to make the canonical image insensitive to illumination [16].

IV. Euclidean and spherical canonical forms

The face recognition method introduced in [15] was based on embedding facial surfaces into $\mathbb{R}^3$, under a tacit assumption of closed mouth. In this section, we show Euclidean canonical forms in $\mathbb{R}^3$ incorporating a topological constraint, which allows handling both open and closed mouth. Then, we show that embedding into a three-dimensional sphere ($\mathbb{S}^3$) can be considered as a generalization of the Euclidean embedding and allows more accurate representation of faces.

A. Topologically-constrained Euclidean canonical forms

The main difference between Euclidean canonical forms used in [15] and the topologically-constrained ones proposed here is the preprocessing stage, in which a more complicated geodesic mask is used (see Section III-A). In order to test the topologically-constrained canonical form, we used a data set containing 102 instances with different facial expressions of 7 subjects, including both open (70 instances) and closed (32 instances) mouth. This is an extended version of the experiment presented in [15], which contained only expressions with closed mouth. In all our experiments, lip contour was segmented manually. The matching of canonical forms was based on 3D moments. For reference, we also show rigid matching of the original surfaces.

Fig. 6 depicts a low-dimensional representation of the dissimilarities between faces in sense of d_{M}. Each symbol on the plot represents a face; colors denote different subjects; the symbol's shape represents the facial expression and its size represents the expression strength. Ideally, clusters corresponding to a specific subject should be as tight as possible (meaning that the representation is insensitive to facial expressions) and as distant as possible from other subjects, which means that the representation allows us to discriminate between different subjects. It can be seen that for rigid surface matching (Fig. 6, left) the clusters overlap, implying that variability due to facial expressions is larger than that stemming from subject's identity. On the other hand, using canonical forms, we obtain tight and distinguishable clusters (Fig. 6, right). This is a practical confirmation of our isometric model, implying that the intrinsic geometry (captured by the geodesic distances) is much better for discriminating the subject's identity than the extrinsic geometry (Euclidean distances).

B. Spherical canonical forms in $\mathbb{S}^3$

In [34], we argued that $\mathbb{R}^3$ is not necessarily the best candidate embedding space for face representation. Despite being the most convenient to work with, the Euclidean space intuitively seems to introduce large distortion, as facial surfaces are curved while the Euclidean space is flat. As an alternative, a three-dimensional sphere $\mathbb{S}^3$ with variable radius R is suggested. The choice of R changes the curvature of the embedding space, thus influencing the embedding error. Embedding into $\mathbb{R}^3$ can be considered as the asymptotic case with $R \to \infty$. Therefore, embedding into $\mathbb{S}^3$ can potentially lead to smaller embedding error.

A three-dimensional sphere of radius R can be represented as the geometric location of all vectors of length R in $\mathbb{R}^4$

$$\mathbb{S}^3 = \left\{\mathbf{x} \in \mathbb{R}^4 : \|\mathbf{x}\|_2 = R\right\}. \quad (8)$$

For convenience, we discuss a unit sphere; the radius R can be taken into account by scaling the geodesic distances. For every point on $\mathbb{S}^3$, there exists a correspondence between a vector of parametric coordinates $\boldsymbol{\theta} = (\theta^1,...,\theta^3)$ and a unit vector in $\mathbb{R}^4$. Given two points $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2$ on the sphere (corresponding to unit vectors $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^4$), the geodesic distance between them is the great circle arc length, given by

$$d_{\mathbb{S}^3}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \cos^{-1}(\langle \mathbf{x}_1, \mathbf{x}_2 \rangle). \quad (9)$$

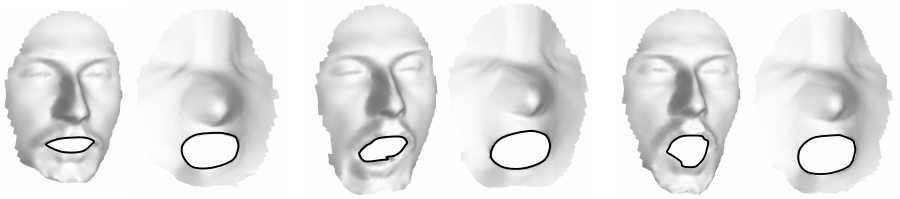

Fig. 5. Examples of canonical forms obtained by embedding into $\mathbb{R}^3$ with topology constraint. Bold line denotes the lip contour.

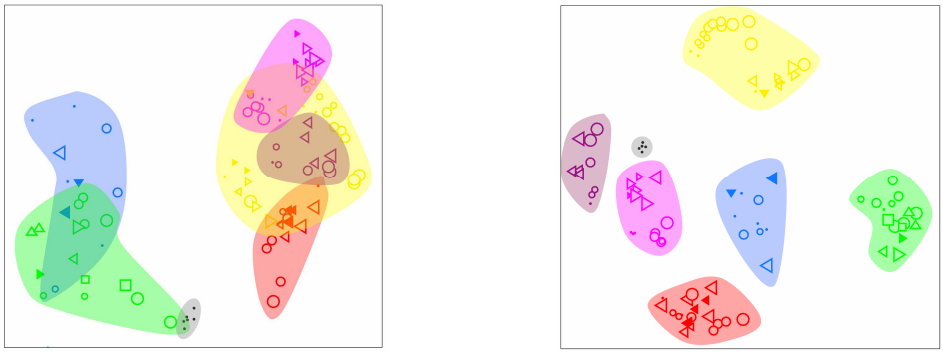

Fig. 6. Low-dimensional visualization of dissimilarities between faces obtained using rigid surface matching (left) and embedding into $\mathbb{R}^3$ with topology constraint (right). Colors represent different subjects. Symbols represent different facial expression with open (empty) and closed mouth (filled).

$\mathbb{S}^3$ can be parameterized as

$$\begin{aligned} x^1(\boldsymbol{\theta}) &= \cos\theta^1 \cos\theta^2 \cos\theta^3, \\ x^2(\boldsymbol{\theta}) &= \cos\theta^1 \sin\theta^2 \cos\theta^3, \\ x^3(\boldsymbol{\theta}) &= \sin\theta^1 \cos\theta^3, \\ x^4(\boldsymbol{\theta}) &= \sin\theta^3. \end{aligned} \tag{10}$$

where $\boldsymbol{\theta} \in [0,\pi] \times [0,2\pi] \times [0,\pi]$. The geodesic distance is explicitly expressed as

$$\begin{aligned} d_{\mathbb{S}^3}(\boldsymbol{\theta}_1,\boldsymbol{\theta}_2) &= \\ &\cos^{-1}\left[\cos\theta_1^1 \cos\theta_1^3 \cos\theta_2^1 \cos\theta_2^3 \cos(\theta_1^2-\theta_2^2) + \right. \\ &\left. \cos\theta_1^3 \cos\theta_2^3 \sin\theta_1^1 \sin\theta_2^1 + \sin\theta_1^1 \sin\theta_2^3\right]. \end{aligned} \tag{11}$$

The normalized stress

$$\sigma_{\mathrm{norm}}(\boldsymbol{\Theta};\boldsymbol{\Delta}) = \frac{\sum_{i<j}(d_{ij}(\boldsymbol{\Theta})-\delta_{ij})^2}{\sum_{i<j} d_{ij}^2(\boldsymbol{\Theta})} \equiv \frac{A}{B}, \tag{12}$$

appears advantageous from the point of view of optimization in case of embedding into $\mathbb{S}^3$ and has been adopted as the embedding error criterion [34]. Here, $\boldsymbol{\Theta}$ is a $3 \times N$ matrix of parametric coordinates and $d_{ij} = d_{\mathbb{S}^3}(\boldsymbol{\theta}_i,\boldsymbol{\theta}_j)$. The gradient of $\nabla_{\boldsymbol{\Theta}}\sigma_{\mathrm{norm}}(\boldsymbol{\Theta};\mathbf{D})$ is given by

$$\frac{\partial}{\partial\theta_k^l}\sigma_{\mathrm{norm}}(\boldsymbol{\Theta};\boldsymbol{\Delta}) = B^{-2}\left(B\frac{\partial}{\partial\theta_k^l}A - A\frac{\partial}{\partial\theta_k^l}B\right), \tag{13}$$

for $k = 1,...,N$ and $l = 1,...,3$, where

$$\begin{aligned} \frac{\partial}{\partial\theta_k^l}A &= 2\sum_i (d_{ij}-\delta_{ij})\frac{\partial}{\partial\theta_k^l}d_{ik}; \\ \frac{\partial}{\partial\theta_k^l}B &= 2\sum_i d_{ik}\frac{\partial}{\partial\theta_k^l}d_{ik}, \end{aligned} \tag{14}$$

and

$$\begin{aligned} \frac{\partial}{\partial\theta_k^l}d_{ik} &= (1-C_{ik}^2)^{-1/2}\frac{\partial}{\partial\theta_k^l}C_{ik}; \\ C_{ik} &= \cos\theta_i^1 \cos\theta_i^3 \cos\theta_k^1 \cos\theta_k^3 \cos(\theta_i^2-\theta_k^2) \\ &\quad + \cos\theta_i^3 \cos\theta_k^3 \sin\theta_i^1 \sin\theta_k^1 + \sin\theta_i^1 \sin\theta_k^1. \end{aligned} \tag{15}$$

The matching of spherical canonical forms is still an open research question. In [34], we treated the spherical canonical forms as point clouds in $\mathbb{R}^4$, and used four-dimensional moments to compare them. However, it appears that moments computed according to (7), where the canonical form is considered as a two-dimensional triangulated mesh in $\mathbb{R}^4$, yield notably better recognition results.

C. Embedding and recognition error

Numerical experiments were performed on a database of four subjects, each appearing with 26 facial expressions of variable strength. The first two subjects in the data set were identical twins. We emphasize that the main purpose of these experiments is not to demonstrate the performance of the isometry-invariant face recognition approach, which has been already demonstrated on larger data sets [15], [35]. Our primary goal is to establish the important connection between

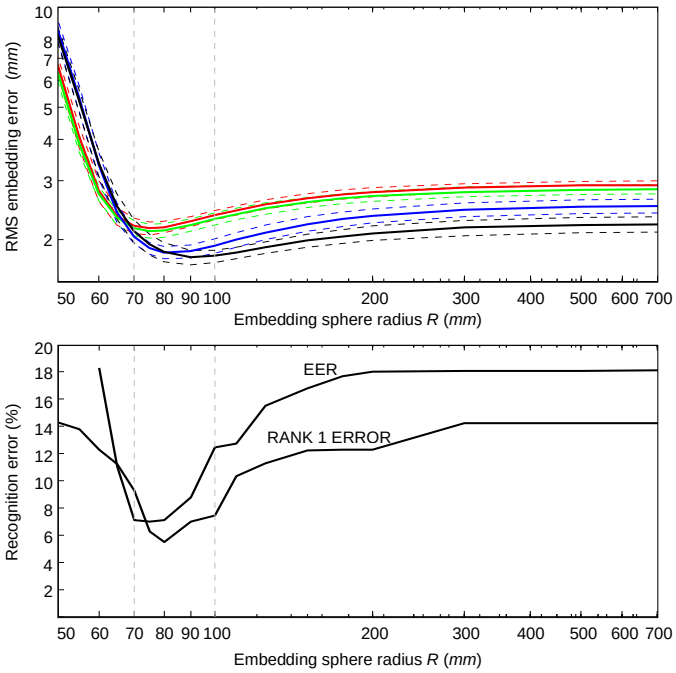

Fig. 7. First row: embedding error versus the embedding sphere radius for four different subjects (colors denote different subjects, dashed lines indicate 95% confidence intervals). Second row: EER and rank-1 error rate versus the embedding sphere radius. The asymptote $R \to \infty$ corresponds to embedding into $\mathbb{R}^3$.

metric distortion introduced by the embedding procedure and recognition accuracy. Our toy set of 104 faces with controlled rich facial expressions seems to be sufficient for this purpose.

Fig. 7 (first row) shows the average embedding error as a function of the sphere radius R. The minimum error is obtained around $R \approx 75 \div 90\,mm$ slightly depending on the subject, and then increases asymptotically as R grows to infinity. The asymptote $R \to \infty$ corresponds to embedding into $\mathbb{R}^3$. We conclude that 3D spherical embedding allows to obtain more than twice lower embedding error compared to 3D Euclidean embedding.

Fig. 7 (second row) presents the equal error rate (EER) and rank-1 recognition error as a function of the embedding sphere radius. The minimum EER of 7.09% is achieved for $R = 70 \div 80\,mm$; the lowest rank-1 error of about 5.48% is obtained for $R = 75\,mm$. Both measures grow more than twice as the embedding sphere radius is decreased ($R = 50\,mm$) or increased ($R \to \infty$). We conclude that both EER and rank-1 error achieve the minimum at embedding sphere radii that yield minimum embedding error. This experimental evidence supports our conjecture that smaller embedding error results in better recognition, and motivates the search for low-distortion representation of surfaces.

V. Spherical canonical images in $\mathbb{S}^2$

In case of zero-codimension embedding (i.e. embedding the facial surface into a two-dimensional space), the embedding produces a canonical parameterization of the surface, which allows comparing texture in an expression-invariant way. The simplest case is embedding into $\mathbb{R}^2$, producing a warped texture, which we called the *canonical image* [2].

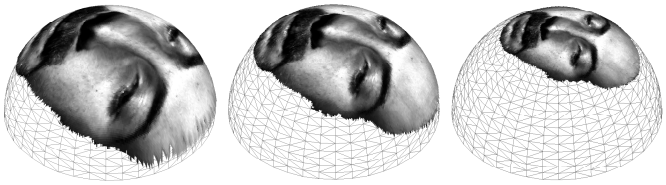

Fig. 8. Face embedded into $\mathbb{S}^2$ with different radii, from left to right: $R = 80mm, 100mm$ and $150mm$.

In this section we propose to embed the facial image into a two-dimensional sphere ($\mathbb{S}^2$) rather than a plane, which will be shown to produce lower embedding distortions. Secondly, we present another experimental evidence supporting our conjecture that lower embedding error directly implies higher recognition accuracy. Finally, we take advantage of the fact that the new canonical image is defined on $\mathbb{S}^2$ and use spherical harmonics to measure dissimilarity between two images. Rotation and reflection invariance of spherical harmonics removes the embedding ambiguity and does not require alignment of the canonical images.

We parameterize $\mathbb{S}^2$ using two angles: the elevation $\theta^1 \in \left[-\frac{\pi}{2}, +\frac{\pi}{2}\right]$ measured from the azimuthal plane in the northern direction, and the azimuthal angle $\theta^2 \in [0, 2\pi)$. The geodesic distance between two points $\boldsymbol{\theta}_1 = (\theta_1^1, \theta_1^2)$ and $\boldsymbol{\theta}_2 = (\theta_2^1, \theta_2^2)$ is given by

$$d_{\mathbb{S}^2}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = R \cos^{-1}\left(\cos\theta_1^1 \cos\theta_2^1 \cos(\theta_1^2 - \theta_2^2) + \sin\theta_1^1 \sin\theta_2^1\right), \tag{16}$$

where R is the sphere radius. As before, we consider a unit sphere; different values of R are achieved by scaling the input distance matrix $\boldsymbol{\Delta}$.

As the embedding error criterion, we use the raw stress (3) where $d_{ij}(\boldsymbol{\Theta}) = d_{\mathbb{S}^2}(\boldsymbol{\theta}_i, \boldsymbol{\theta}_j)$. One of the points, say $\boldsymbol{\theta}_1$, is anchored to the north pole of the sphere ($\theta_1^1 = \frac{\pi}{2}$). This point is chosen to be the nose tip and is determined as a local maximum of the Gaussian curvature on the facial surface.

The mapping $\varphi : \mathcal{S} \to \mathbb{S}^2$ from the original surface to the sphere can be regarded as a warping transformation, which maps the original facial image $\alpha : \mathcal{S} \to \mathbb{R}$ onto a section of the sphere. The resulting image, $a : \varphi(\mathcal{S}) \subset \mathbb{S}^2 \mapsto \mathbb{R}$, can be computed for any $\boldsymbol{\theta}$ by means of linear interpolation (see Figures 8, 9). The image a referred to as the *spherical canonical image* [2] is invariant to isometric deformations of the face and, hence, insensitive to facial expressions. Yet, the spherical canonical image is not a fully invariant signature of the face, since fixing a single fiducial point on the pole still allows one degree of freedom of rotation and reflection about that point.

A. Spherical harmonic signatures

In order to obtain a truly invariant signature of the face, we use the *spherical harmonic transform* (SHT). A function $a \in \mathbb{L}^2\left(\mathbb{S}^2\right)$ can be expanded in the spherical harmonic basis

with coefficients

$$\begin{aligned}\hat{a}_{l,m} &= \langle a, Y_l^m \rangle \\ &= \int_0^{\pi}\int_0^{2\pi} a\left(\theta^1,\theta^2\right)\overline{Y_l^m\left(\theta^1,\theta^2\right)}d\theta^2\cos(\theta^1)d\theta^1,\end{aligned} \quad (17)$$

for $l \in \mathbb{N}\cup\{0\}$ and $|m| \le l$, where

$$Y_l^m\left(\theta^1,\theta^2\right) = \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}}P_l^m\left(\sin\theta^1\right)e^{\imath m\theta^2} \quad (18)$$

is the (l,m)-spherical harmonic, and P_l^m is the associate Legendre function of degree l and order m. A discrete version of the SHT $\hat{a}_{l,m}$ can be carried out efficiently using the FFT [36].

A handy property of spherical harmonics is that for every $\Delta\theta^2$, $a\left(\theta^1,\theta^2\right)$ and $a\left(\theta^1,\theta^2\pm\Delta\theta^2\right)$ are transformed to two sets of coefficients, which differ only in the complex phase. Hence, the set of coefficients $c_{l,m} = |\hat{a}_{l,m}|$ removes the rotation and reflection ambiguities from the canonical image and defines an invariant signature of the face.

As a dissimilarity measure between two such signatures, we use the Euclidean norm

$$d_{\mathrm{SH}}\left(c_{l,m},c'_{l,m}\right) = \sum_{l\ge 0,|m|\le 0}\left(c_{l,m}-c'_{l,m}\right)^2. \quad (19)$$

Using basic properties of spherical harmonics, it is straightforward to show that such a measure is characterized by the *similarity* property, i.e.

$$d_{\mathrm{SH}}\left(c_{l,m},c'_{l,m}\right) \le \sum_{l\ge 0,|m|\le 0}\left|\hat{a}_{l,m}-\hat{a}'_{l,m}\right|^2 = \|a-a'\|, \quad (20)$$

and since $d_{\mathrm{SH}}\left(c_{l,m},c'_{l,m}\right)$ is invariant under any $\mathcal{R}\in\mathcal{G}$, where $\mathcal{G}$ is the group of rotations and reflections about the north pole on $\mathbb{S}^2$,

$$d_{\mathrm{SH}}\left(c_{l,m},c'_{l,m}\right) \le \min_{\mathcal{R}\in\mathcal{G}}\|a-\mathcal{R}a'\|. \quad (21)$$

In other words, similar a and a' (up some $\mathcal{R}\in\mathcal{G}$) will result in small dissimilarities in sense of d_{SH}, whereas dissimilar a and a' will result in large values of d_{SH}. In practice, different "frequencies" in the spherical harmonics domain typically carry different amounts of information useful for discriminating between different subjects. Therefore, a more sophisticated dissimilarity measure could be based on a weighted Euclidean norm. Optimal weights can be found from a training set e.g. by using PCA.

A disadvantage of the proposed representation is that it is invariant only to azimuthal roto-reflection, while being sensitive to general roto-reflections on $\mathbb{S}^2$. This, in turn, requires the use of constrained embedding, which fixes the location of the nose tip, and thus relies on its location. The use of a non-constrained embedding is feasible in combination with a signature invariant under a general roto-reflection group on $\mathbb{S}^2$. Construction of such a signature is based on the fact that the subspace

$$V_l = \mathrm{span}\left\{Y_l^m : |m|\le l\right\} \quad (22)$$

is closed under a roto-reflection group on $\mathbb{S}^2$ [37]. Hence, a signature of the form

$$c_l = \|\hat{a}_l\| = \|(\hat{a}_{l,-l},...,\hat{a}_{l,l})\| \quad (23)$$

is invariant under general rotations and reflections. Dissimilarity between such signatures can be measured as in (19).

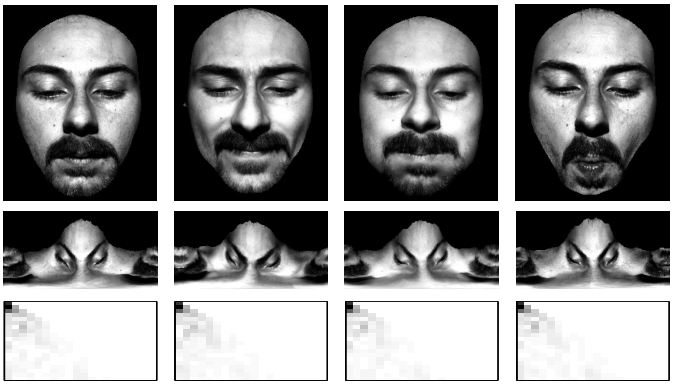

Fig. 9. First row: original facial images; second row: spherical canonical images shown in parametric coordinates; third row: spherical harmonics transform coefficients

B. Embedding and recognition error

The spherical canonical images method was applied to the data set from Section IV-C. Fig. 10 (first row) depicts the average embedding error for each subject plotted as a function of R. It appears that embedding sphere radius yielding the minimum embedding error ranges from 90 to $100\,mm$, slightly depending on the subject.

Fig. 10 (second row) presents the equal error rate (EER) and rank-1 recognition error as a function of the embedding sphere radius. The minimum EER of 12.39% is achieved at $R = 90\,mm$. At this embedding sphere radius, rank-1 error of 1.96% is achieved, which remains nearly constant for $R = 90\div 125\,mm$. Both recognition error measures appear to be in close correspondence with the embedding error. This provides an additional experimental evidence to support our conjecture that smaller embedding error results in better recognition.

VI. Discussion and conclusion

We started with the assumption that facial expressions can be considered as isometries of the facial surface and showed an empirical justification for this model. We demonstrated how this model can be applied to constructing expression-invariant representations of faces using the isometric embedding approach. The resulting representation is useful, for example, in 3D face recognition. We studied the impact of the embedding space choice on the metric distortion introduced by the embedding and concluded that spaces with spherical geometry are more favorable for representation of facial surfaces. We also provided an experimental evidence that spherical embedding leads to better recognition rates compared to Euclidean ones. Furthermore, the use of spherical canonical images in $\mathbb{S}^2$ allows us to perform matching in the spherical harmonic transform domain, which does not require preliminary alignment of the images. The results of Mémoli

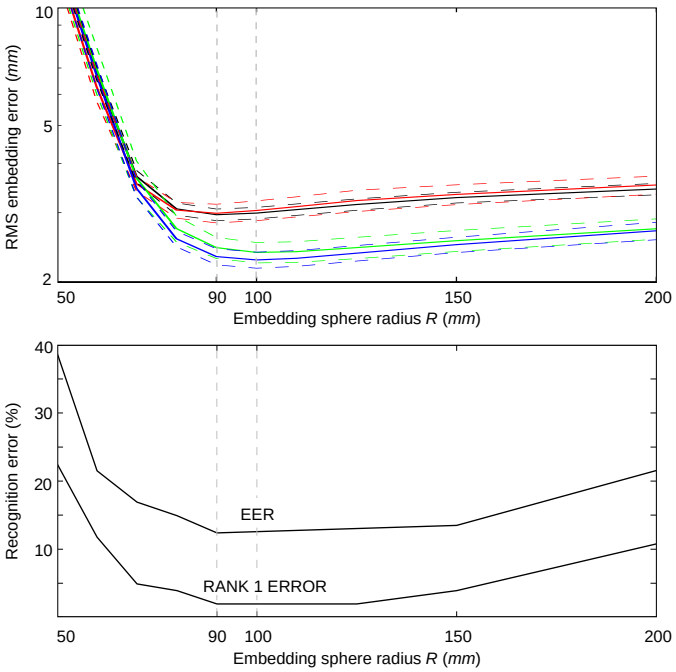

Fig. 10. First row: embedding error versus the embedding sphere radius for four different subjects (colors denote different subjects, dashed lines indicate 95% confidence intervals). Second row: EER and rank-1 error rate versus the embedding sphere radius. The asymptote $R \to \infty$ corresponds to embedding into $\mathbb{R}^2$.

and Sapiro [38] and our follow-up results [25], [26], [35] show that the idea of non-Euclidean embeddings can be generalized to embedding one surface into another. The errors introduced by embedding into a space with some predefined geometry like $\mathbb{R}^3$ or $\mathbb{S}^3$ are avoided in this way.

Our algorithm is computationally efficient and has near real-time performance. A basic version of the algorithm was implemented in a prototype system developed at the Department of Computer Science, Technion [15]. The system is based on a commodity PC platform with AMD Opteron processor supporting SSE extensions. All the software has been tailored for this architecture. In a typical face recognition setup, end-to-end processing comprises the following stages: 3D surface acquisition and reconstruction ($0.90\,sec$), smoothing ($0.33\,sec$), subsampling to approximately 2500 points ($0.08\,sec$), geodesic mask computation ($0.10\,sec$), facial surface cropping ($0.08\,sec$) measurement of geodesic distances between all the points using a parallelized parametric version of FMM ($1.60\,sec$), MDS carried out by 40 iterations of the SMACOF algorithm [27] ($2.08\,sec$) and canonical form comparison ($< 0.005\,sec$). The overall processing time is about $5\,sec$. In light of our recent results [39], [28], we foresee that the MDS stage performance can be further significantly improved using multigrid optimization.

REFERENCES

[1] Y. Tian, T. Kanade, and J. Cohn, *Handbook of face recognition*, chapter Facial expression analysis, Springer-Verlag, Heidelberg, 2003.

[2] A. M. Bronstein, M. M. Bronstein, and R. Kimmel, "Expression-invariant 3D face recognition," in *Proc. Audio and Video-based Biometric Person Authentication*. 2003, Lecture Notes on Computer Science, pp. 62–69, Springer.

[3] F. I. Parke and K. Waters, *Computer Facial Animation*, AK Peters, 1994.

[4] F. Pighin, R. Szeliski, and D. H. Salesin, "Modeling and animating realistic faces from images," *IJCV*, vol. 50, no. 2, pp. 143–169, November 2002.

[5] V. Blanz, C. Basso, T. Poggio, and T. Vetter, "Reanimating faces in images and video," *Computer Graphics Forum*, vol. 22, no. 3, pp. 641–650, 2003.

[6] G. Zigelman, R. Kimmel, and N. Kiryati, "Texture mapping using surface flattening via multi-dimensional scaling," *IEEE Trans. Visualization and computer graphics*, vol. 9, no. 2, pp. 198–207, 2002.

[7] A. M. Bronstein, M. M. Bronstein, and R. Kimmel, "Interpolation and extrapolation of facial surfaces using generalized multidimensional scaling," *IEEE Trans. Vis. Comp. Graphics*, 2006, submitted.

[8] P. Ekman, *Darwin and facial expression; a century of research in review*, Academic Press, New York, 1973.

[9] K. Chang, K. Bowyer, and P. Flynn, "Effects on facial expression in 3D face recognition," in *SPIE Conference on Biometric Technology for Human Identification*, 2005, pp. 132–143.

[10] T. Theoharis G. Toderici G. Passalis, I.A. Kakadiaris and N. Murtuza, "Evaluation of 3D face recognition in the presence of facial expressions: an annotated deformable model approach," in *IEEE Workshop on Face Recognition Grand Challenge Experiments*, 2005.

[11] X. Lu and A. K. Jain, "Deformation analysis for 3D face matching," in *Workshop on Applications of Computer Vision*, 2005, pp. 99–104.

[12] K. Bowyer, K. Chang, and P. Flynn, "A survey of approaches and challenges in 3D and multi-modal 3D+2D face recognition," *Computer Vision and Image Understanding*, vol. 101, no. 1, pp. 1–15, 2006.

[13] C. Chua, F. Han, and Y. Ho, "3D human face recognition using point signature," in *International Conference on Automatic Face and Gesture Recognition*, 2000, pp. 233–238.

[14] K. Chang, K. Bowyer, and P. Flynn, "Multiple nose region matching for 3D face recognition under varying facial expression," preprint.

[15] A. M. Bronstein, M. M. Bronstein, and R. Kimmel, "Three-dimensional face recognition," *IJCV*, vol. 64, no. 1, pp. 5–30, August 2005.

[16] A. M. Bronstein, M. M. Bronstein, E. Gordon, and R. Kimmel, "Fusion of 3D and 2D information in face recognition," in *Proc. IEEE ICIP*, 2004.

[17] A. M. Bronstein, M. M. Bronstein, and R. Kimmel, "Expression-invariant face recognition via spherical embedding," in *Proc. IEEE ICIP*, 2005, vol. 3, pp. 756–759.

[18] A. Elad and R. Kimmel, "On bending invariant signatures for surfaces," *IEEE Trans. PAMI*, vol. 25, no. 10, pp. 1285–1295, 2003.

[19] A. M. Bronstein, M. M. Bronstein, E. Gordon, and R. Kimmel, "High-resolution structured light range scanner with automatic calibration," Tech. Rep. CIS-2003-06, Dept. of Computer Science, Technion, Israel, 2003.

[20] R. Kimmel and J. A. Sethian, "Computing geodesic on manifolds," in *Proc. US National Academy of Science*, 1998, vol. 95, pp. 8431–8435.

[21] N. Linial, E. London, and Y. Rabinovich, "The geometry of graphs and some its algorithmic applications," *Combinatorica*, vol. 15, pp. 333–344, 1995.

[22] E. L. Schwartz, A. Shaw, and E. Wolfson, "A numerical solution to the generalized mapmaker's problem: flattening nonconvex polyhedral surfaces," *IEEE Trans. PAMI*, vol. 11, pp. 1005–1008, 1989.

[23] A. Elad and R. Kimmel, *Geometric methods in bio-medical image processing*, vol. 2191, chapter Spherical flattening of the cortex surface, pp. 77–89, Springer-Verlag, Berlin Heidelberg New York, 2002.

[24] 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*, 2002, pp. 123–131.

[25] A. M. Bronstein, M. M. Bronstein, and R. Kimmel, "Generalized multidimensional scaling: a framework for isometry-invariant partial surface matching," *Proc. National Academy of Sciences*, vol. 103, no. 5, pp. 1168–1172, January 2006.

[26] A. M. Bronstein, M. M. Bronstein, and R. Kimmel, "Efficient computation of the Gromov-Hausdorff distance for smooth surfaces," *SIAM Journal on Scientific Computing*, submitted.

[27] I. Borg and P. Groenen, *Modern multidimensional scaling - theory and applications*, Springer-Verlag, Berlin Heidelberg New York, 1997.

[28] M. M. Bronstein, A. M. Bronstein, R. Kimmel, and I. Yavneh, "Multigrid multidimensional scaling," *Numerical Linear Algebra with Applications (NLAA)*, vol. 13, pp. 149–171, March-April 2006.

[29] P. J. Besl and N. D. McKay, "A method for registration of 3D shapes," *IEEE Trans. PAMI*, vol. 14, pp. 239–256, 1992.
[30] Z. Y. Zhang, "Iterative point matching for registration of free-form curves and surfaces," *IJCV*, vol. 13, pp. 119–152, 1994.
[31] 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.
[32] M. Turk and A. Pentland, "Face recognition using eigenfaces," in *Proc. Computer Vision and Pattern Recognition*, 1991, pp. 586–591.
[33] L. Sirovich and M. Kirby, "Low-dimensional procedure for the characterization of human faces," *JOSA A*, vol. 2, pp. 519–524, 1987.
[34] A. M. Bronstein, M. M. Bronstein, and R. Kimmel, "On isometric embedding of facial surfaces into $\mathbb{S}^3$," in *Intl. Conf. on Scale Space and PDE Methods in Computer Vision*, 2005, pp. 622–631.
[35] A. M. Bronstein, M. M. Bronstein, and R. Kimmel, "Robust expression-invariant face recognition from partially missing data," in *Proc. ECCV*. 2006, Lecture Notes on Computer Science, pp. 396–408, Springer.
[36] P. Kostelec D. Healy Jr., D. Rockmore and S. Moore, "FFTs for the 2-sphere – improvements and variations," *The Journal of Fourier Analysis and Applications*, vol. 9, no. 4, pp. 341–385, 2003.
[37] M. Kazhdan, T. Funkhouser, and S. Rusinkiewicz, "Rotation invariant spherical harmonic representation of 3D shape descriptors," in *Proc. Eurographics Symp. on Geom. Proc.*, 2003.
[38] F. Mémoli and G. Sapiro, "A theoretical and computational framework for isometry invariant recognition of point cloud data," *Foundations of Computational Mathematics*, vol. 5, no. 3, pp. 313–347, 2005.
[39] 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.