Isometric Embedding of Facial Surfaces into $\mathbb{S}^3$

Alexander M. Bronstein[1], Michael M. Bronstein[2], and Ron Kimmel[2]

[1] Department of Electrical Engineering,
alexbron@ieee.org
[2] Department of Computer Science,
bronstein@ieee.org, ron@cs.technion.ac.il
Technion - Israel Institute of Technology,
32000 Haifa, Israel

Abstract. The problem of isometry-invariant representation and comparison of surfaces is of cardinal importance in pattern recognition applications dealing with deformable objects. Particularly, in three-dimensional face recognition treating facial expressions as isometries of the facial surface allows to perform robust recognition insensitive to expressions.

Isometry-invariant representation of surfaces can be constructed by isometrically embedding them into some convenient space, and carrying out the comparison in that space. Presented here is a discussion on isometric embedding into $\mathbb{S}^3$, which appears to be superior over the previously used Euclidean space in sense of the representation accuracy.

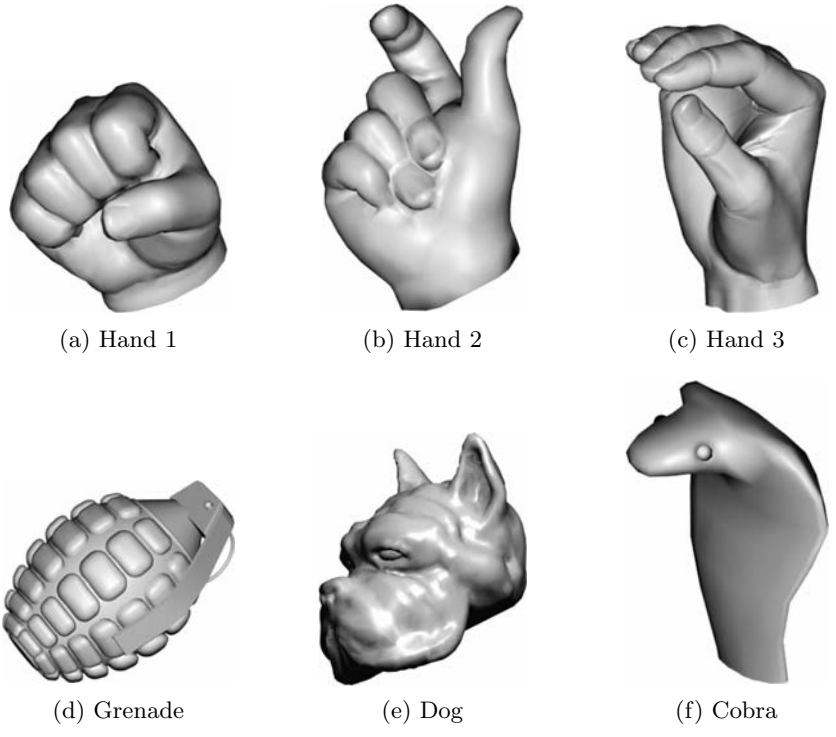
1 Introduction

The problem of isometry-invariant representation of surfaces arises in numerous pattern recognition applications dealing with deformable objects. Particularly, in three-dimensional face recognition, it was shown that facial expressions can be modelled as isometric transformations of the facial surface [1, 2]. Under this assumption, the problem of expression invariant face recognition is reduced to finding similarity between isometric surfaces.

Figure 1 illustrates the problem of isometric surface matching. The first row (a)-(c) shows three isometric transformations of the same hand (assume that the fingers do not touch each other, such that the topology is preserved), which, with a bit of imagination, look like a grenade, a dog and a cobra (d)-(f). In other words, from the point of view of their *extrinsic* geometry[1], isometric surfaces can look completely different, while being just instances of the same surface.

[1] Formally, *intrinsic* geometry refers to all the properties of the manifold expressed in terms of the metric (first fundamental form), and *extrinsic* geometry refers to properties expressed in terms of the second fundamental form. We use these terms in a broader sense, by which extrinsic geometry defines the properties that describe the way the manifold is immersed in the ambient space.

R. Kimmel, N. Sochen, J. Weickert (Eds.): Scale-Space 2005, LNCS 3459, pp. 622–631, 2005.
© Springer-Verlag Berlin Heidelberg 2005



(a) Hand 1 (b) Hand 2 (c) Hand 3

(d) Grenade (e) Dog (f) Cobra

Fig. 1. Illustration of the deformable surface matching problem. (a) - (c): isometries of a hand. (d) - (f): different objects that resemble the hands from extrinsic geometry point of view

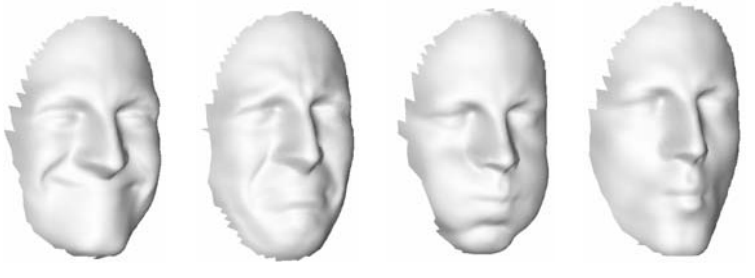
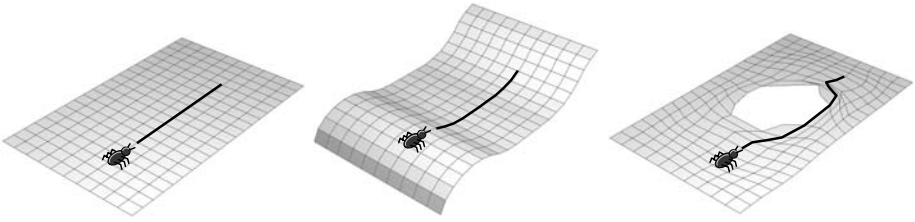


Fig. 2. Several facial expressions of the same person that can be modelled as isometries. Face data shown by courtesy of Eyal Gordon

Formally, given two complete compact smooth Riemannian manifolds $(\mathcal{S}, g)$ and $(\mathcal{Q}, h)$, the diffeomorphism $f : (\mathcal{S}, g) \to (\mathcal{Q}, h)$ is called an *isometry* if $f*h = g$, where $f*h$ denotes pullback of the metric. As the result, all the *intrinsic* geometric properties of the surface are preserved. Equivalently, an isometry can be defined as a diffeomorphism preserving the geodesic distances, that is



(a) Original surface (b) Isometric transformation (c) Non-isometric transformation

Fig. 3. Illustration of isometric (b) and non-isometric (c) transformations of a surface (a). Isometries do not change the intrinsic geometry of the surface, such that an imaginable creature living on the surface does not feel the transformation

$$d_{\mathcal{S}}(\xi_1, \xi_2) = d_{\mathcal{Q}}(\eta_1, \eta_2) \quad \forall \xi_1, \xi_2 \in \mathcal{S},\ \eta_1, \eta_2 \in \mathcal{Q}. \tag{1}$$

where $d_{\mathcal{S}}$ and $d_{\mathcal{Q}}$ denote the geodesic distances induced by g and h, respectively.

2 Bending-Invariant Canonical Forms

Let $(\mathcal{S}, h)$ and $(\mathcal{Q}, h)$ be two-dimensional Riemannian manifolds (surfaces) related by an isometry $f(\mathcal{S}) = \mathcal{Q}$. In the context of face recognition $\mathcal{S}$ and $\mathcal{Q}$ are different expressions of the same face. Since geodesic distances are preserved under an isometry, they are suitable candidates for an isometry-invariant representation of the surface.

However, we should remember that the surfaces $\mathcal{S}$ and $\mathcal{Q}$ are sampled and therefore in practice we have two *finite metric spaces* $(\{\xi_1, ..., \xi_{N_{\mathcal{S}}}\}, \mathbf{D}_{\mathcal{S}})$ and $(\{\eta_1, ..., \eta_{N_{\mathcal{Q}}}\}, \mathbf{D}_{\mathcal{Q}})$, respectively. The matrices $\mathbf{D}_{\mathcal{S}} = (d_{\mathcal{S}}(\xi_i, \xi_j))$ and $\mathbf{D}_{\mathcal{Q}} = (d_{\mathcal{Q}}(\eta_i, \eta_j))$ denote the mutual geodesic distances between the points in $\mathcal{S}$ and $\mathcal{Q}$. There is neither guarantee that $\mathcal{S}$ and $\mathcal{Q}$ are sampled at the same points, nor that the number of samples of the two surfaces is the same ($N_{\mathcal{S}} \neq N_{\mathcal{Q}}$). Moreover, even if the samples are the same, they can be ordered arbitrarily. This ambiguity makes impractical the use of $\mathbf{D}$ itself as an invariant representation.

2.1 Isometric Embedding

An alternative proposed in [3] is to avoid dealing explicitly with the matrix of geodesic distances and represent the Riemannian surface as a subset of some convenient m-dimensional space $\mathcal{S}'^m$, such that the original intrinsic geometry is preserved. We call such a procedure *isometric embedding*. This embedding allows to get rid of the extrinsic geometry, which no more exists in the new space. As a consequence, the resulting representation is identical for all isometric transformations of the surface. Another advantage is related to the fact that a general Riemannian metric is usually inconvenient to work with. The embedding

space, on the other hand, can be chosen completely to our discretion. That is, the embedding replaces a complicate geometric structure by a convenient one.

Isometric embedding is a mapping between two finite metric spaces

$$\varphi : (\{\xi_1, ..., \xi_N\} \subset \mathcal{S}, \mathbf{D}) \rightarrow (\{\xi'_1, ..., \xi'_N\} \subset \mathcal{S}'^m, \mathbf{D}') \quad , \tag{2}$$

such that

$$d'_{ij} = d_{ij} \quad \forall i, j = 1, ..., N. \tag{3}$$

The matrices $\mathbf{D} = (d_{ij}) = (d(\xi_i, \xi_j))$ and $\mathbf{D'} = (d'_{ij}) = (d'(\xi'_i, \xi'_j))$ denote the mutual geodesic distances between the points in the original and the embedding space, respectively. Following Elad and Kimmel, the image of $\{\xi_1, ..., \xi_N\}$ under φ is called the *canonical form* of $(\mathcal{S}, g)$ [3].

In general, such isometric embedding does not exist, and therefore one has to bear in mind that the canonical form is an *approximate* representation of the discrete surface. It is possible to find optimal canonical forms in sense of some metric distortion criterion. Also, the canonical form is not defined uniquely, but up to any transformation in the embedding space that does not alter the distances (e.g. in an Euclidean space, such transformations are translations, rotations and reflections). Yet this ambiguity is much easier to cope with compared to the vast degrees of freedom in the matrix $\mathbf{D}$.

2.2 The Choice of the Embedding Space

An important question is how to choose the embedding space. First, the geometry of the embedding space is important. Popular choices include spaces with flat [4, 3, 5, 6], spherical [7] or hyperbolic [8] geometry. This choice should be dictated by the convenience of a specific space and the resulting embedding error, which, in turn, depends on the embedding error definition.

Secondly, the dimension m of the embedding space must be chosen in such a way that the codimension of $\varphi(\mathcal{S})$ in $\mathcal{S}'^m$ is at least 1. The reason is made clear if we limit our manifolds to be graphs of functions (in our case - functions of two variables). Sampling of a graph $z(x, y)$ produces a set of points, which when embedded into $\mathbb{R}^2$ (or some other two-dimensional manifold), reflect the sampling pattern while the intrinsic geometry is captured mainly by the deformation of the boundaries of that function. Increasing the sampling rate will not enhance the intrinsic geometry captured by the embedded space. It would mainly indicate the x, y parametric shadow, or numerical support in $\mathbb{R}^2$ of the sampled function. On the other hand, when embedded into $\mathbb{R}^3$ (or other higher-dimensional manifolds), the sample points will lie along some two-dimensional submanifold of $\mathbb{R}^3$, and increasing the sampling rate would better capture the geometry of this submanifold.

Embedding with codimension zero (e.g. embedding of a surface in a plane) is useful when the manifold is endowed with some additional property, for example, texture. Such embedding can be thought of as an intrinsic parameterization of the manifold and has been explored in the context of medical visualization [4], texture mapping [6] and registration of facial images [9].

The main focus of present paper is embedding into a three-dimensional sphere $\mathbb{S}^3$. This space appears to be more suitable for embedding of facial surfaces than the Euclidean space used beforehand in [3, 1, 2]. Embedding into a two-dimensional sphere $\mathbb{S}^2$ was employed by Elad and Kimmel for visualization of convoluted brain cortical surface [7]. Here, we present embedding into $\mathbb{S}^3$ as a more accurate representation of facial surfaces.

3 Embedding into $\mathbb{S}^3$

A unit[2] m-dimensional sphere can be represented as the geometric locations of all unit vectors in $\mathbb{R}^{m+1}$

$$\mathbb{S}^m = \left\{ \mathbf{x} \in \mathbb{R}^{m+1} : \|\mathbf{x}\|_2 = 1 \right\}. \tag{4}$$

For every point on $\mathbb{S}^m$, there exists a correspondence between the parameterization coordinates $\xi^1, ..., \xi^m$ and the unit vector in $\mathbb{R}^{m+1}$. Given two points $\boldsymbol{\xi}_i, \boldsymbol{\xi}_j$ on the sphere (corresponding to unit vectors $\mathbf{x}_i, \mathbf{x}_j \in \mathbb{R}^{m+1}$), the geodesic distance between them is the great circle arc length, given by

$$d_{\mathbb{S}^m}(\boldsymbol{\xi}_i, \boldsymbol{\xi}_j) = \cos^{-1}(\langle \mathbf{x}_i, \mathbf{x}_j \rangle). \tag{5}$$

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

$$\begin{aligned} x^1(\boldsymbol{\xi}) &= \cos\xi^1 \cos\xi^2 \cos\xi^3, \\ x^2(\boldsymbol{\xi}) &= \cos\xi^1 \sin\xi^2 \cos\xi^3, \\ x^3(\boldsymbol{\xi}) &= \sin\xi^1 \cos\xi^3, \\ x^4(\boldsymbol{\xi}) &= \sin\xi^3. \end{aligned} \tag{6}$$

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

$$\begin{aligned} d_{\mathbb{S}^3}(\boldsymbol{\xi}_i, \boldsymbol{\xi}_j) = \cos^{-1}(&\cos\xi_i^1 \cos\xi_i^3 \cos\xi_j^1 \cos\xi_j^3 \cos(\xi_i^2 - \xi_j^2) + \\ &\cos\xi_i^3 \cos\xi_j^3 \sin\xi_i^1 \sin\xi_j^1 + \sin\xi_i^1 \sin\xi_j^3). \end{aligned} \tag{7}$$

We use the normalized *weighted stress* [10] as the embedding error criterion

$$\epsilon(\boldsymbol{\Xi}'; \mathbf{D}, \mathbf{W}) = \frac{\sum_{i<j} w_{ij}(d'_{ij}(\boldsymbol{\Xi}') - d_{ij})^2}{\sum_{i<j} d'^2_{ij}(\boldsymbol{\Xi}')} = \frac{A}{B}, \tag{8}$$

where $\boldsymbol{\Xi}' = (\xi_j'^i)$ is a $3 \times N$ matrix representing the parametric coordinates of the canonical form points, and $\mathbf{D} = (d_{ij})$ and $\mathbf{W} = (w_{ij})$ are $N \times N$ matrices of geodesic distances and weights, respectively. The weight are chosen as $w_{ij} = 1$; choosing $w_{ij} \propto d_{ij}^{-2}$ gives the *relative stress* [10].

[2] Without loss of generality, we discuss a unit sphere. The sphere radius is taken into account by scaling the distances.

The stress $\epsilon(\mathbf{\Xi}'; \mathbf{D}, \mathbf{W})$ is minimized w.r.t. $\mathbf{\Xi}$ using the BFGS quasi-Newton algorithm (medium-scale optimization implemented in MATLAB function `fminunc`). The gradient of $\epsilon(\mathbf{\Xi}'; \mathbf{D}, \mathbf{W})$ w.r.t. $\mathbf{\Xi}'$ is given by

$$\frac{\partial}{\partial \xi_k'^l} \epsilon(\mathbf{\Xi}'; \mathbf{D}, \mathbf{W}) = B^{-2} \left(B \frac{\partial}{\partial \xi_k'^l} A - A \frac{\partial}{\partial \xi_k'^l} B \right), \tag{9}$$

where

$$\begin{aligned} \frac{\partial}{\partial \xi_k'^l} A &= 2 \sum_i w_{ik} (d'_{ij} - d_{ij}) \frac{\partial}{\partial \xi_k'^l} d'_{ik}, \\ \frac{\partial}{\partial \xi_k'^l} B &= 2 \sum_i d'_{ik} \frac{\partial}{\partial \xi_k'^l} d'_{ik}, \end{aligned} \tag{10}$$

and

$$\begin{aligned} \frac{\partial}{\partial \xi_k'^l} d'_{ik} &= (1 - C_{ik}^2)^{-1/2} \frac{\partial}{\partial \xi_k'^l} C_{ik}, \\ C_{ik} &= \cos \xi_i'^1 \cos \xi_i'^3 \cos \xi_k'^1 \cos \xi_k'^3 \cos(\xi_i'^2 - \xi_k'^2) + \\ &\quad \cos \xi_i'^3 \cos \xi_k'^3 \sin \xi_i'^1 \sin \xi_k'^1 + \sin \xi_i'^1 \sin \xi_k'^1 \end{aligned} \tag{11}$$

4 Numerical Results

The spherical embedding approach was tested on a set of human faces. Facial surfaces were acquired using a structured light 3D scanner [11] and underwent preprocessing by cropping, smoothing and subsampling (see details in [2]). The final surfaces contained about 750 points. An efficient modification of the parametric Fast Marching algorithm [12, 13, 14] was used to measure geodesic distances on the discrete surfaces.

Comparison of canonical form was performed in $\mathbb{R}^4$ using the moments signatures. A $pqrs$ moment of the canonical form is defined as

$$\mu_{pqrs}^{\mathbf{X}'} = \sum_{i=1}^{N} \left(x_i'^1\right)^p \left(x_i'^2\right)^q \left(x_i'^3\right)^r \left(x_i'^4\right)^s, \tag{12}$$

where $\mathbf{X}' = x_j'^i$ denotes the $4 \times N$ matrix of $\mathbb{R}^4$ coordinates of the canonical form points, corresponding to the parametric coordinates ξ_j^i.

Canonical forms were first aligned using an Euclidean transformation, by eliminating the first-order moments $m_{1000}, m_{0100}, m_{0010}, m_{0001}$ and the mixed second-order moments $m_{1100}, m_{1010}, m_{1001}, m_{0110}, m_{0101}, m_{0011}$. The axes were reordered according to the second order moments, making the projection onto the first axis x^1 have the largest variance, and onto the fourth axis x^4 the smallest variance. This operation resolves the translation and rotation ambiguities for non-trivial objects. Next, reflections were applied to each axis x^k in order to set

$$\sum_{i=1}^{N} \text{sign } \left(x_i'^k\right) \geq 0. \tag{13}$$

This operation resolves the reflection ambiguity.

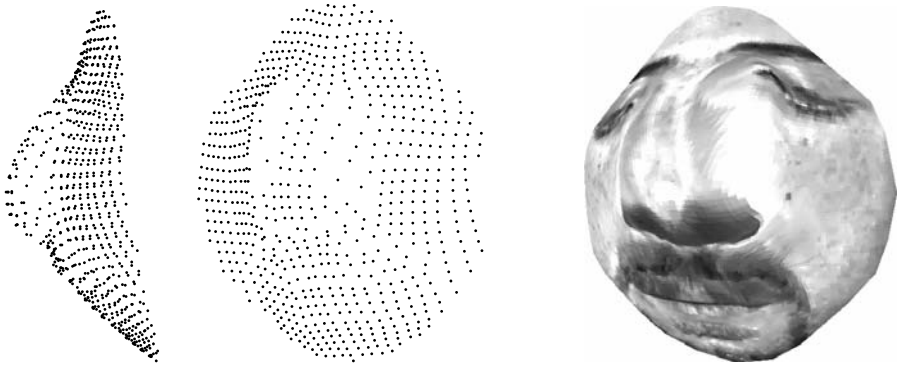
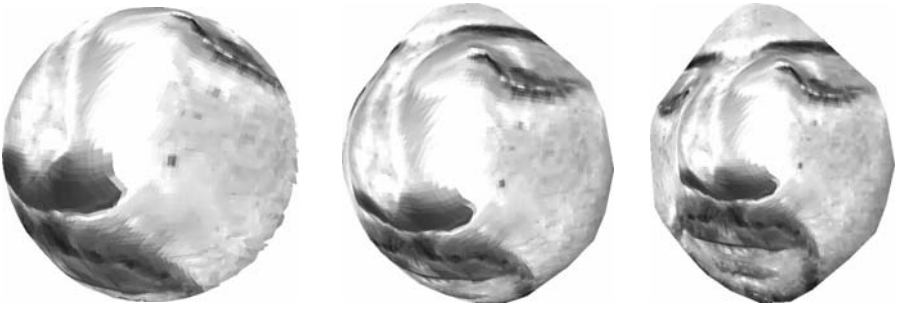


Fig. 4. Maximum-variance projection onto $\mathbb{R}^3$ of a facial surface embedded into $\mathbb{S}^3$ with radius $R = 10cm$



(a) $R = 5cm$ (b) $R = 7.5cm$ (c) $R = 15cm$

Fig. 5. Maximum-variance projections onto $\mathbb{R}^3$ of a facial surface embedded into $\mathbb{S}^3$ with different radii

Next, a signature of moments up to the fifth order was computed. The distance between two canonical forms $\mathbf{X}'$ and $\mathbf{Y}'$ was computed according to the standard Euclidean distance between their moments signatures:

$$d^2_{mom}(\mathbf{X}', \mathbf{Y}') = \sum_{p+q+r+s=2,..,5} \left(\mu^{\mathbf{X}'}_{pqrs} - \mu^{\mathbf{Y}'}_{pqrs}\right)^2. \tag{14}$$

Figure 4 depicts the maximum-variance projection onto $\mathbb{R}^3$ of a canonical form obtained by embedding into a sphere of radius $R = 10cm$. Figure 5 depicts the embedding of the same face into spheres of different radii.

Figure 6 shows the embedding error ϵ as a function of the sphere radius. The minimum error is obtained around $R = 7.5cm$ and then increases asymptotically, as R grows to infinity. The asymptote corresponds to embedding into $\mathbb{R}^3$.

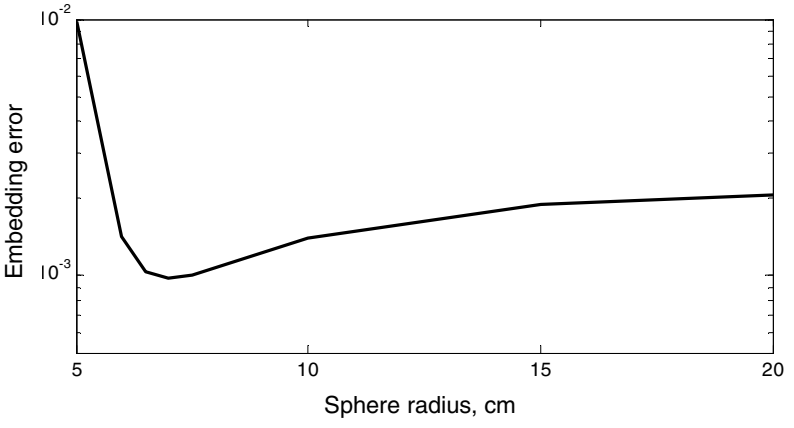
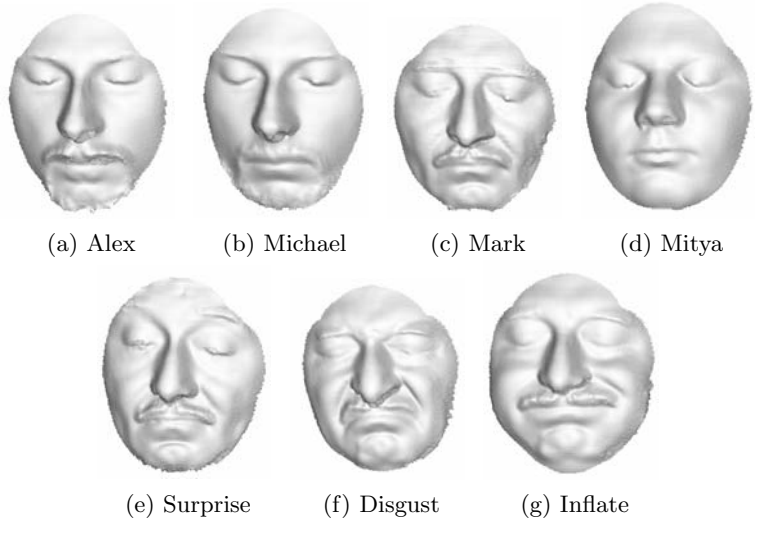


Fig. 6. Embedding error as the function of the sphere radius in cm. The asymptote $R \to \infty$ corresponds to embedding into $\mathbb{R}^3$



(a) Alex (b) Michael (c) Mark (d) Mitya

(e) Surprise (f) Disgust (g) Inflate

Fig. 7. (a) - (d) Faces used in the experiment. Alex and Michael are identical twins. (e) - (g) Representative facial expressions of subject Mark

Therefore, spherical embedding allows to obtain more than twice lower embedding error.

Finally, Figures 2–8 show a toy "face recognition" experiment that was performed on a set of 33 faces of four subjects in the presence of extreme facial expressions (Figure 2). Figure 8 depicts a two-dimensional visualization of the similarities (in sense of d_{mom}) between the faces.

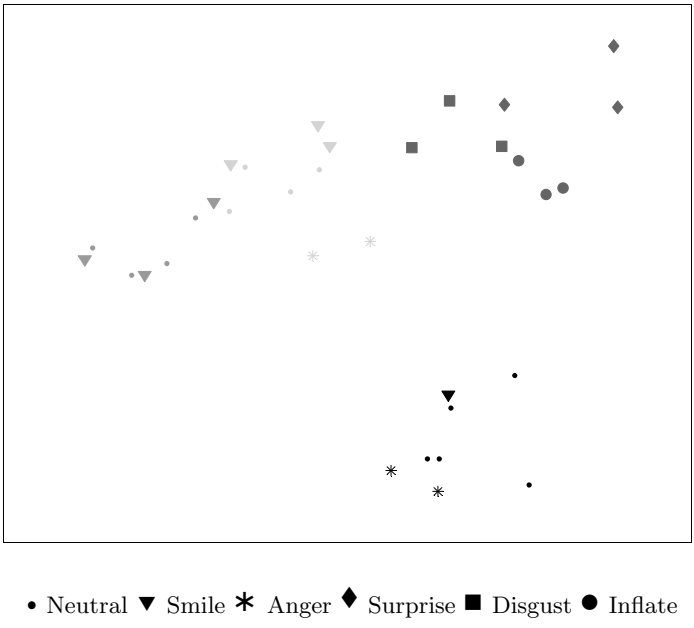


Fig. 8. Two-dimensional visualization of dissimilarities between faces. Colors represent different subjects. Symbols represent different facial expressions

5 Conclusions

We presented the three-dimensional sphere $\mathbb{S}^3$ as an alternative to the Euclidean space used beforehand in [3, 1, 2] to construct expression-invariant representation of human faces. Using $\mathbb{S}^3$ results in smaller embedding error, and leads to a more accurate representation.

References

1. A. Bronstein, M. Bronstein, and R. Kimmel. Expression-invariant 3D face recognition. In Proc. Audio and Video-based Biometric Person Authentication, pages 62–69, 2003.
2. A. Bronstein, M. Bronstein, and R. Kimmel. Three-dimensional face recognition. Technical Report CIS-2004-04, Dept. of Computer Science, Technion, Israel, 2004.
3. A. Elad and R. Kimmel. Bending invariant representations for surfaces. In Proc. CVPR, pages 168–174, 2001.
4. E. L. Schwartz, A. Shaw, and E. Wolfson. A numerical solution to the generalized mapmaker's problem: attening nonconvex polyhedral surfaces. IEEE Trans. PAMI, 11:1005–1008, 1989.
5. R. Grossman, N. Kiryati, and R. Kimmel. Computational surface attening: a voxel-based approach. IEEE Trans. PAMI, 24(4):433–441, 2002.

6. G. Zigelman, R. Kimmel, and N. Kiryati. Texture mapping using surface attening via multi-dimensional scaling. IEEE Trans. Visualization and computer graphics, 9(2):198–207, 2002.
7. A. Elad and R. Kimmel. Spherical attening of the cortex surface. In R. Malladi, editor, Geometric methods in bio-medical image processing, volume 2191, pages 77–89. Springer-Verlag, Berlin Heidelberg New York, 2002.
8. 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.
9. A. Bronstein, M. Bronstein, E. Gordon, and R. Kimmel. Fusion of 3D and 2D information in face recognition. In Proc. ICIP, 2004. to appear.
10. I. Borg and P. Groenen. Modern multidimensional scaling - theory and applications. Springer-Verlag, Berlin Heidelberg New York, 1997.
11. A. Bronstein, M. Bronstein, E. Gordon, and R. Kimmel. High-resolution structured light range scanner with automatic calibration. Technical Report CIS-2003-06, Dept. of Computer Science, Technion, Israel, 2003.
12. J. A. Sethian. A review of the theory, algorithms, and applications of level set method for propagating surfaces. Acta numerica, pages 309–395, 1996.
13. R. Kimmel and J. A. Sethian. Computing geodesic on manifolds. In Proc. US National Academy of Science, volume 95, pages 8431–8435, 1998.
14. A. Spira and R. Kimmel. An efficient solution to the eikonal equation on parametric manifolds. Interfaces and Free Boundaries, 2004. To appear.