A comprehensive statistical framework for elastic shape analysis of 3D faces

Sebastian Kurtek, Hassen Drira

A Comprehensive Statistical Framework for Elastic Shape Analysis of 3D Faces

Sebastian Kurtek[a], Hassen Drira[b],

[a]*Department of Statistics, The Ohio State University, Columbus, OH, USA*
[b]*Institut Mines-Télécom/Télécom Lille, CRIStAL (UMR CNRS 9189) Lille, France*

Abstract

We develop a comprehensive statistical framework for analyzing shapes of 3D faces. In particular, we adapt a recent elastic shape analysis framework to the case of hemispherical surfaces, and explore its use in a number of processing applications. This framework provides a parameterization-invariant, elastic Riemannian metric, which allows the development of mathematically rigorous tools for statistical analysis. Specifically, this paper describes methods for registration, comparison and deformation, averaging, computation of covariance and summarization of variability using principal component analysis, random sampling from generative shape models, symmetry analysis, and expression and identity classification. An important aspect of this work is that all tasks are preformed under a unified metric, which has a natural interpretation in terms of bending and stretching of one 3D face to align it with another. We use a subset of the BU-3DFE face dataset, which contains varying magnitudes of expression.

Keywords: 3D face, statistical framework, elastic Riemannian metric, generative face model

1. Introduction

In recent years, there has been an exponential growth of accessible 3D face datasets due to increasing technological progress in development of acquisition and storage sensors. The 3D face represents important cues for many applications such as human-machine interaction, medical surgery, surveillance, etc., and thus, studying the shape of facial surfaces has become a fundamental problem in computer vision and graphics [1, 2]. Any appropriate shape analysis framework applied to the face problem should be able to automatically find optimal correspondences between facial surfaces (one-to-one nonlinear matching of points across surfaces), produce natural deformations that align one 3D face to another, and provide tools for statistical analysis such as computation of an average or template face, exploration of variability in different expression classes, random sampling of 3D faces from statistical models, and even reflection symmetry analysis. These tools, if developed properly, allow for principled and efficient modeling of complex 3D face data. The 3D face registration, deformation and statistical modeling problems are closely related, and thus, should be solved simultaneously under a unified Riemannian shape analysis framework. The 3D facial surfaces are assumed to be genus-0 and are allowed to undergo complex isometric and elastic deformations, and may contain missing parts. Below, we summarize some of the state-of-the-art methods for 3D face modeling that are relevant to our paper; most of these methods focus on face recognition rather than the general statistical analysis task.

Many approaches are based on markers to model the 3D face. Marker-based systems are widely used for face animation [3, 1]. Explicit face markers significantly simplify tracking, but also limit the amount of spatial detail that can be captured. There have been several approaches in recent years that rely on deforming facial surfaces into one another, under some chosen criterion, and use quantifications of these deformations as metrics for face recognition. Among these, the ones using nonlinear deformations facilitate local stretching, compression, and bending of surfaces to match each other and are referred to as elastic methods. For instance, Kakadiaris et al. [4] utilize an annotated face model to study geometrical variability across faces. The annotated face model is deformed elastically to fit each face, thus matching different anatomical areas such as the nose and eyes. In affective computing, the markers correspond to action units and allow one to model the 3D face for expression understanding [5]. A strong limitation of all marker-based approaches is the need for manual segmentation and/or annotation of a 3D face. In other approaches, the 3D face is represented by a markerless morphable model, which can be used for identity recognition [6] and face animation [7, 8]. In [6], a hierarchical geodesic-based resampling approach is applied to extract landmarks for modeling facial surface deformations. The deformations learned from a small group of subjects (control group) are then synthesized onto a 3D neutral model (not in the control group), resulting in a deformed template. The proposed approach is able to handle expressions and pose changes simultaneously by fitting a generative deformable model. In [8], facial expressions are represented as a weighted sum of blendshape meshes and the non-rigid iterative closest point (ICP) algorithm is applied together with face tracking to generate 3D face animations. This class of approaches is automatic and can be performed in real time. However, in all of these methods there is no definition of a proper metric, which is needed for statistical analysis. On the other hand, the proposed method provides a proper metric in the shape space of 3D faces allowing the definition of statistics such as an average and covariance.

Majority of previous approaches to 3D face analysis are based on extracting local cues leading to discriminant features used for many applications such as identity, expression and gender classification [9, 10]. The advantage of these approaches is high classification accuracy along with low computational cost for computer vision applications. However, these approaches are less significant in the computer graphics context. This is due to the fact that statistical analysis of facial surfaces in the feature space is generally not easily mapped back to the original surface space. Thus, the obtained results, while computationally inexpensive, are very difficult to interpret and use in practice.

In several approaches, the 3D face is embedded into a particular space of interest, and the faces are compared in that space. Tsalakanidou et al. [11] apply principal component analysis to build eigenfaces, where each face image in the database can be represented as a vector of weights; the weights of an image are obtained by its projection onto the subspace spanned by the eigenface directions. Then, identification of the test image is done by locating the image in the database whose weights have the smallest Euclidean distance from the weights of the test image. The main limitation of this method is that it is not invariant to pose changes. Furthermore, the model is image-based where, in addition to the face of interest, one must account for the image background. Bronstein et al. [12] construct a computationally efficient, invariant representation of surfaces undergoing isometric deformations by embedding them into a low-dimensional space with a convenient geometry. These approaches allow deformation-robust metrics that are useful for several applications including biometrics. However, computation of statistics is not possible under this model.

Drira et al. [13] represent the 3D face as a collection of radial curves that are analyzed under a Riemannian framework for elastic shape analysis of curves [14]. This framework provides tools for computation of deformations between facial surfaces, mean calculation of 3D faces via the curve representation, and 3D face recognition. Along similar lines, [15, 16] used facial curves to model facial surfaces for several other applications. The main limitation of these works is that they utilize a curve representation of 3D faces. Thus, registrations between the surfaces are curve-based, and the correspondence between the radial curves must be known a priori (very difficult in practice). As a result, the computed correspondences and any subsequent computations tend to be suboptimal. Furthermore, to the best of our knowledge, these approaches did not thoroughly investigate the use of the Riemannian framework for more complex statistical modeling such as random sampling of facial surfaces from a generative model.

There is also a number of methods in the graphics literature, which provide tools for various shape modeling tasks [17, 18, 19]. While these methods are very general and provide good results on complex shapes, they require the surface registration problem to be solved either manually or via some other unrelated method. Thus, these methods do not provide proper metrics for shape comparison and statistical modeling in the presence of different surface parameterizations. The main benefit of the proposed approach is that the registration and comparison/modeling problems are solved simultaneously under a unified Riemannian metric.

In this paper, we adapt a recent elastic shape analysis framework [20, 21] to the case of hemispherical surfaces, and explore its use in a number of 3D face processing applications. This framework was previously defined for quadrilateral, spherical and cylindrical surfaces. All of the considered tasks are performed under an elastic Riemannian metric allowing principled definition of various tools including registration via surface re-parameterization, deformation and symmetry analysis using geodesic paths, intrinsic shape averaging, principal component analysis, and definition of generative shape models. Thus, the main contributions of this work are:

(1) We extend the framework of Jermyn et al. [20] for statistical shape analysis of quadrilateral and spherical surfaces to the case of hemispherical surfaces.

(2) We consider the task of 3D face morphing using a parameterized surface representation and a proper, parameterization-invariant elastic Riemannian metric. This provides the formalism for defining optimal correspondences and deformations between facial surfaces via geodesic paths.

(3) We define a comprehensive statistical framework for modeling of 3D faces. The definition of a proper Riemannian metric allows us to compute intrinsic facial shape averages as well as covariances to study facial shape variability in different expression classes. Using these estimates one can form a generative 3D face model that can be used for random sampling.

(4) We provide tools for symmetry analysis of 3D faces, which allows quantification of asymmetry of a given face and identification of the nearest (approximately) symmetric face.

(5) We study expression and identity classification under this framework using the defined metric. We compare our performance to the state-of-the-art method in [13]. The main idea behind presenting this application is to showcase the benefits of an elastic framework in the recognition task. We leave a more thorough study of classification performance and comparisons to other state-of-the-art methods as future work.

The rest of this paper is organized as follows. Section 2 defines the mathematical framework. Section 3 presents the applicability of the proposed method to various 3D face processing tasks. We close the paper with a brief summary in Section 4.

2. Mathematical Framework

In this section, we describe the main ingredients in defining a comprehensive, elastic shape analysis framework for facial surfaces. We note that these methods have been previously described for the case of quadrilateral, spherical and cylindrical surfaces in [20, 21]. We extend these methods to hemispherical surfaces and apply them to statistical shape analysis of 3D faces. Let $\mathcal{F}$ be the space of all smooth embeddings of a closed unit disk in $\mathbb{R}^3$, where each such embedding defines a parameterized surface $f : \bar{\mathbb{D}} \to \mathbb{R}^3$. Let Γ be the set of all boundary-preserving diffeomorphisms of $\bar{\mathbb{D}}$. For a facial surface $f \in \mathcal{F}$, $f \circ \gamma$ represents its re-parameterization. In other words, γ is a warping of the coordinate system on f. As previously shown in [20], it is inappropriate to use the $\mathbb{L}^2$ metric for analyzing shapes of parameterized surfaces, because Γ does not act on $\mathcal{F}$ by isometries. Thus, we utilize the square-root normal field (SRNF) representation of surfaces and the corresponding Riemannian metric proposed in [20]. We summarize these methods next and refer the reader to those papers for more details.

Let $s = (u, v) \in \bar{\mathbb{D}}$ define a polar coordinate system on the closed unit disk. The SRNF representation of facial surfaces is then defined using a mapping $Q : \mathcal{F} \to \mathbb{L}^2$ as $Q(f)(s) = \frac{n(s)}{|n(s)|^{1/2}}$. Here, $n(s) = \frac{\partial f}{\partial u}(s) \times \frac{\partial f}{\partial v}(s)$ denotes a normal vector to the surface f at the point $f(s)$. The space of all SRNFs is a subset of $\mathbb{L}^2(\bar{\mathbb{D}}, \mathbb{R}^3)$, henceforth referred to simply as $\mathbb{L}^2$, and it is endowed with the natural $\mathbb{L}^2$ metric. The differential of Q is a smooth mapping between tangent spaces, $Q_{*,f} : T_f(\mathcal{F}) \to T_{Q(f)}(\mathbb{L}^2)$, and is used to define the corresponding Riemannian metric on $\mathcal{F}$ as $\langle\langle w_1, w_2\rangle\rangle_f = \langle Q_{*,f}(w_1), Q_{*,f}(w_2)\rangle_{\mathbb{L}^2}$, where $w_1,\ w_2 \in T_f(\mathcal{F})$, $n_w(s) = \frac{\partial f}{\partial u}(s) \times \frac{\partial w}{\partial v}(s) + \frac{\partial w}{\partial u}(s) \times \frac{\partial f}{\partial v}(s)$, $|\cdot|$ denotes the 2-norm in $\mathbb{R}^3$, and ds is the Lebesgue measure on $\bar{\mathbb{D}}$ [21]. Using this expression, one can verify that the re-parameterization group Γ acts on $\mathcal{F}$ by isometries, i.e. $\langle\langle w_1 \circ \gamma, w_2 \circ \gamma\rangle\rangle_{f\circ\gamma} = \langle\langle w_1, w_2\rangle\rangle_f$. Another advantage of this metric is that it has a natural interpretation in terms of the amount of stretching and bending needed to deform one surface into another. For this reason, it has been referred to as the partial elastic metric [20]. Furthermore, this metric is automatically invariant to translation. Scaling variability can be removed by rescaling all surfaces to have unit area. We let C denote the space of all unit area surfaces. This defines the pre-shape space in our analysis.

Rotation and re-parameterization variability is removed from the representation space using equivalence classes. Let $q = Q(f)$ denote the SRNF of a facial surface f. A rotation of f by $O \in SO(3)$, Of, results in a rotation of its SRNF representation, Oq. A re-parameterization of f by $\gamma \in \Gamma$, $f \circ \gamma$, results in the following transformation of its SRNF: $(q, \gamma) = (q \circ \gamma)\sqrt{J_\gamma}$, where J_γ is the determinant of the Jacobian of γ. Now, one can define two types of equivalence classes, $[f] = \{O(f \circ \gamma) | O \in SO(3), \gamma \in \Gamma\}$ in C endowed with the metric $\langle\langle\cdot,\cdot\rangle\rangle$ or $[q] = \{O(q, \gamma) | O \in SO(3), \gamma \in \Gamma\}$ in $\mathbb{L}^2$ endowed with the $\mathbb{L}^2$ metric; each equivalence class represents a shape uniquely in its respective representation space. This results in two strategies to account for the rotation and re-parameterization variabilities in 3D face data. Given two surfaces $f_1, f_2 \in C$, the exact solution comes from the following optimization problem: $(O^*, \gamma^*) = \text{arginf}_{(O,\gamma)\in SO(3)\times\Gamma}\, d_C(f_1, O(f_2 \circ \gamma))$. Unfortunately, there is no closed form expression for the geodesic distance d_C because of the complex structure of the Riemannian metric $\langle\langle\cdot,\cdot\rangle\rangle$. There is a numerical approach, termed path-straightening, which can be used to compute this geodesic distance, but it is computationally expensive. Thus, we use an approximate solution to the registration problem in our analysis, which can be computed using the SRNF representation as $(O^*, \gamma^*) = \text{arginf}_{(O,\gamma)\in SO(3)\times\Gamma} \|q_1 - (Oq_2, \gamma)\|$. This problem is much easier to solve and provides a very close approximation to the original problem, because the partial elastic metric on C is the pullback of the $\mathbb{L}^2$ metric from the SRNF space.

The optimization problem over $SO(3) \times \Gamma$ is solved iteratively using the general procedure presented in [20, 21]. First, one fixes γ and searches for an optimal rotation over $SO(3)$ using Procrustes analysis; this is performed in one step using singular value decomposition. Then, given the computed rotation, one searches for an optimal re-parameterization in Γ using a gradient descent algorithm, which requires the specification of an orthonormal basis for $T_{\gamma_{id}}(\Gamma)$. The definition of this basis depends on the domain of the surface. In the present case, we seek a basis of smooth vector fields that map the closed unit disk to itself. In order to define this basis, we make a small simplification. Because all of the initial, facial surface parameterizations were obtained by defining the point $s = (0, 0)$ at the tip of the nose, we treat this point as a landmark, i.e. it is fixed throughout the registration process. Given this simplification, we first construct a basis for $[0, 1]$ as $B_{[0,1]} = \{\sin(2\pi n_1 u), 1 - \cos(2\pi n_1 u), u, 1 - u | n_1 = 1, \ldots, N_1,\ u \in [0, 1]\}$ and a basis for $\mathbb{S}^1$ as $B_{\mathbb{S}^1} = \{\sin(n_2 v), 1 - \cos(n_2 v), v, 2\pi - v | n_2 = 1, \ldots, N_2,\ v \in [0, 2\pi]\}$. We take all products of these two bases while ensuring that the boundary of the unit disk is preserved. Then, to define an orthonormal basis of $T_{\gamma_{id}}(\Gamma)$ we use the Gram-Schmidt procedure. This results in a finite, orthonormal basis $B_{\bar{\mathbb{D}}} = \{b_1, \ldots, b_N\}$ for $T_{\gamma_{id}}(\Gamma)$. In the following sections, we let $f_2^* = O^*(f_2 \circ \gamma^*)$, where $O^* \in SO(3)$ is the optimal rotation and $\gamma^* \in \Gamma$ is the optimal re-parameterization. Then, the geodesic distance in the shape space $\mathcal{S} = C/(SO(3)\times\Gamma)$ is computed using $d([f_1], [f_2]) = \inf_{(O,\gamma)\in SO(3)\times\Gamma} d_C(f_1, O(f_2 \circ \gamma)) \approx d_C(f_1, O^*(f_2 \circ \gamma^*))$. This allows us to compute the geodesic only once, after the two facial surfaces have been optimally registered.

As a next step, we are interested in comparing facial surface shapes using geodesic paths and distances. As mentioned earlier, there is no closed form expression for the geodesic in C, and thus, we utilize a numerical technique termed path-straightening. In short, this approach first initializes a path between the two given surfaces, and then "straightens" it according to an appropriate path energy gradient until it becomes a geodesic. We refer the reader to [22, 21] for more details. In the following sections, we use $F^{*,pre}$ to denote the geodesic path between two facial surfaces f_1 and f_2 in the pre-shape space (no optimization over $SO(3) \times \Gamma$) and $F^{*,sh}$ to denote the geodesic path in the shape space between f_1 and f_2^*. The length of the geodesic path is given by $L(F^*) = \int_0^1 \sqrt{\langle\langle F_t^*, F_t^*\rangle\rangle_F}\,dt$, where $F_t^* = \frac{dF^*}{dt}$. All derivatives and integrals in our framework are computed numerically. The computational cost of the proposed method is similar to that reported in [22].

3. Applications

In this section, we describe the utility of the presented mathematical tools in various 3D face processing tasks including deformation, template estimation, summarization of variability, random sampling and symmetry analysis. We also present two classification tasks concerned with (1) classifying expressions, and (2) classifying person identities. The 3D faces used in this paper are a subset of the BU-3DFE dataset. BU-3DFE is a database of annotated 3D facial expressions, collected by

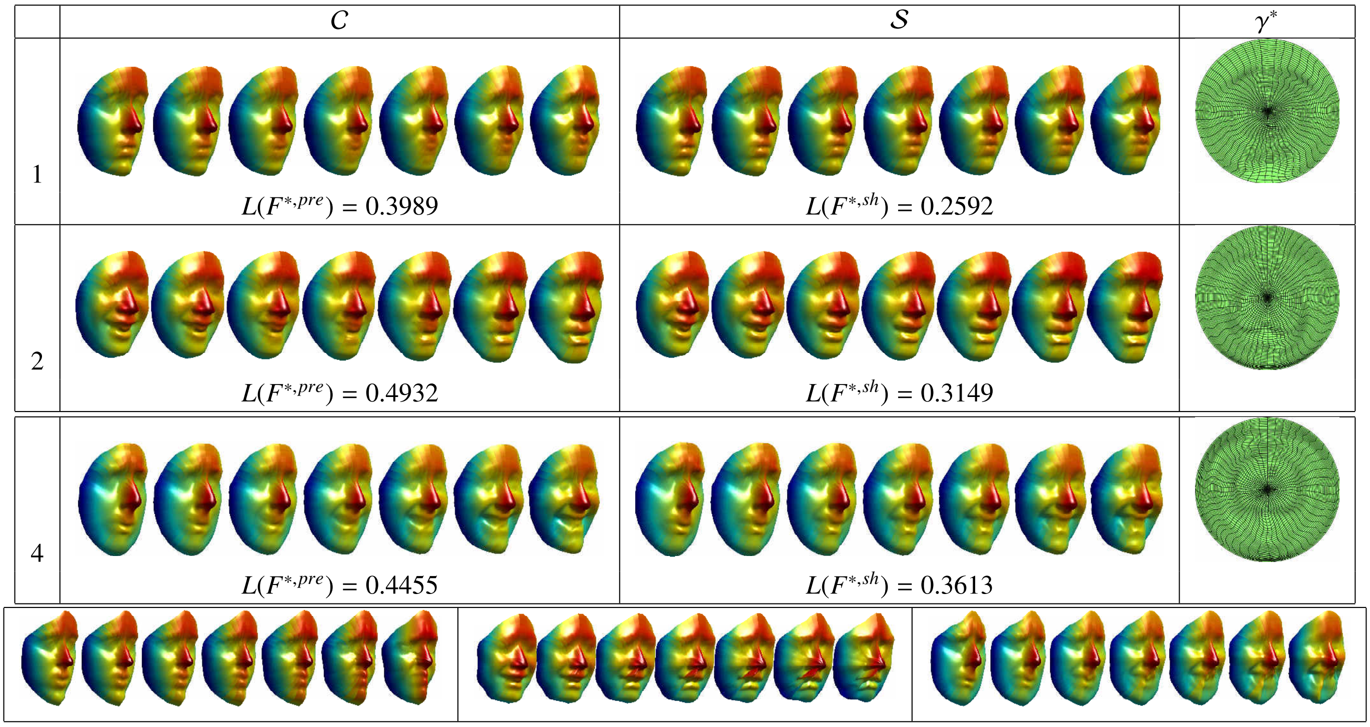

Figure 1: Top: Comparison of geodesic paths and distances in $\mathcal{C}$ and $\mathcal{S}$ for different persons and expressions (1 neutral to anger, 2 happiness to disgust, and 3 sadness to happiness) as well as optimal re-parameterizations (allow elastic deformations between 3D faces). Bottom: Geodesics (1)-(3) computed using [13].

Yin et al. [23] at Binghamton University in Binghamton, NY, USA, which was designed for research on 3D human faces and expressions and to develop a general understanding of human behavior. There are a total of 100 subjects in the database, 56 females and 44 males. A neutral scan was first captured for each subject. Then, each person was asked to perform six expressions reflecting the following emotions: anger, happiness, fear, disgust, sadness and surprise. The expressions varied according to four levels of intensity (low, middle, high and highest). Thus, there were 25 3D facial expression models per subject in the entire database. We use a subset of this data with highest expression intensities (most challenging case) to assess the proposed method.

Each facial surface is represented by an indexed collection of radial curves that are defined and extracted as follows. The reference curve on a facial surface f is chosen to be the vertical curve after the face has been rotated to the upright position. Then, each radial curve β_α is obtained by slicing the facial surface by a plane P_α that has the nose tip as its origin and makes an angle α with the plane containing the reference curve. We repeat this step to extract radial curves at equally-separated angles, resulting in a set of curves that are indexed by the angle α. Thus, the facial surface is represented in a polar (radius-angle) coordinate system. We use 50 radial curves sampled with 50 points in our surface representation (50×50 grid).

Face Deformation: We generate facial shape deformations using geodesic paths. While linear interpolations could also be used here, the geodesic provides the optimal deformation under the defined Riemannian metric. Since we only have to compute the geodesic once per deformation, after the surfaces have been optimally registered, this does not result in a prohibitive computational cost. We compare the results obtained in $\mathcal{C}$ to those in $\mathcal{S}$ in Figure 1. We consider three different examples for various persons and expressions. There is a large decrease in the geodesic distance in each case due to the additional optimization over $SO(3) \times \Gamma$. It is clear from this figure that elastic matching of 3D faces is very important when the main goal is to generate natural deformations between them. This is especially evident in the areas of the lips and eyes. Take, for instance, Example 1. In the pre-shape space, the lips are averaged out along the geodesic path and are pretty much non-existent close to the midpoint. But, due to a better matching of geometric features along the geodesic path in the shape space, the lips are clearly defined. The same can be observed in the eye region. As will be seen in the next section, these distortions become even more severe when one considers computing averages and variability within a set of 3D faces. In the right panel of the figure we display the optimal re-parameterizations that achieve the correspondence between these surfaces; these are clearly nonlinear and depict natural transformations. We also generated geodesics for the same examples using the curve-based method in [13] (bottom panel of Figure 1). These results suggest that considering the radial curves independently can generate severe distortions in the geodesic paths and produce unnatural deformations between 3D faces.

Face Template: We generate 3D face templates using the notion of the Karcher mean. Tools and results for computing shape statistics for cylindrical surfaces under the SRNF rep-

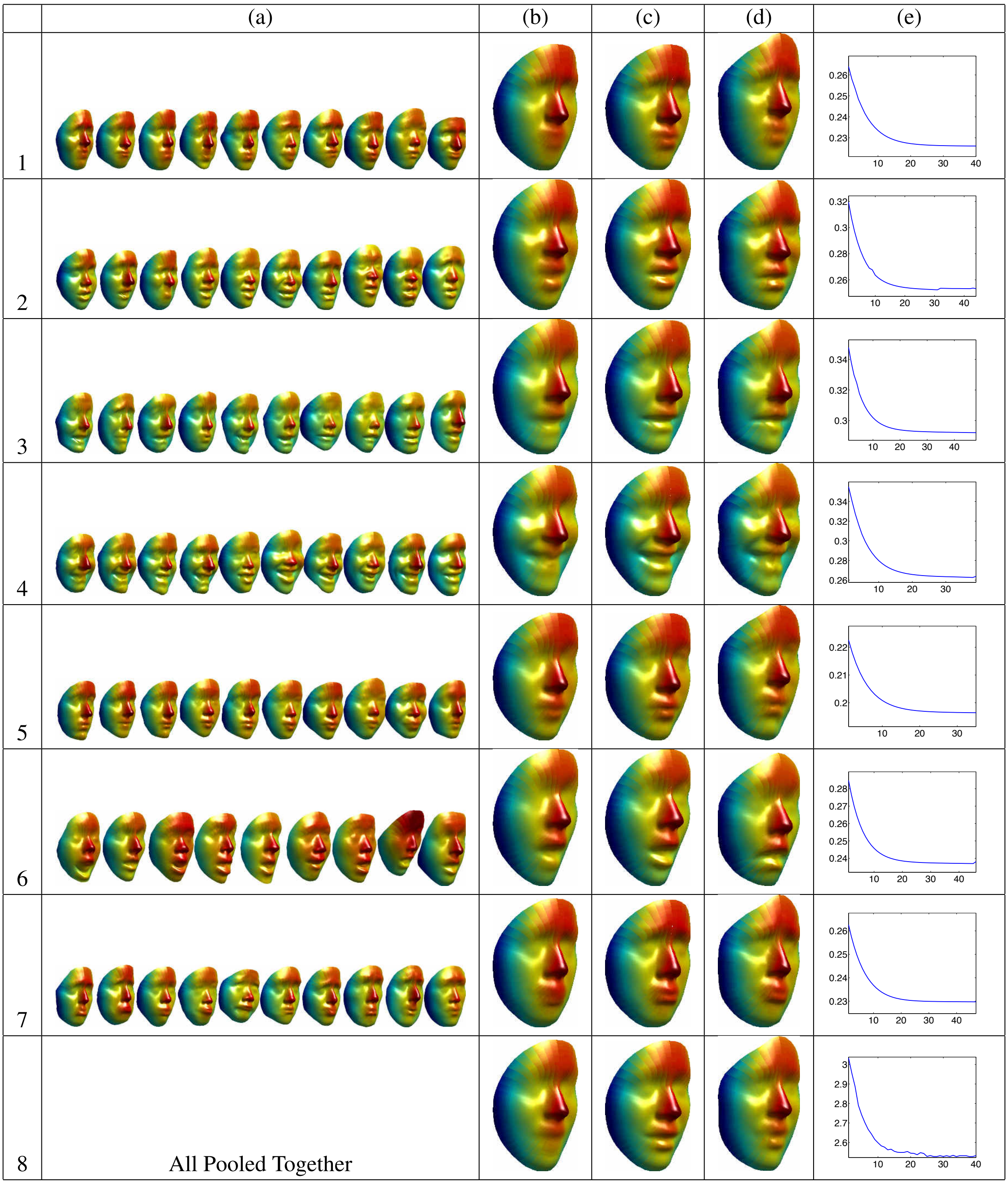

Figure 2: (a) Sample of surfaces used to compute the face template for each expression: (1) anger, (2) disgust, (3) fear, (4) happiness, (5) neutral, (6) surprise, (7) sadness, and (8) all samples pooled together. (b) Sample average computed in $\mathcal{C}$. (c) Karcher mean computed in $\mathcal{S}$. (d) Karcher mean computed using [13]. (e) Optimization energy in $\mathcal{S}$ (sum of squared distances of each shape from the current average) at each iteration.

resentation have been previously described in [24]; we review some of the concepts relevant to current analysis in the following sections. Let $\{f_1, \ldots, f_n\} \in \mathcal{C}$ denote a sample of facial surfaces. Then, the Karcher mean is defined as $[\bar{f}] = \mathrm{argmin}_{[f]\in\mathcal{S}} \sum_{i=1}^{n} L(F_i^{*,sh})^2$, where $F_i^{*,sh}$ is a geodesic path between a surface $F_i^{*,sh}(0) = f$ and a surface in the given sample $F_i^{*,sh}(1) = f_i^*$ that was optimally registered to f. A gradient-based approach for finding the Karcher mean is given in [24]. The Karcher mean is actually an equivalence class of surfaces and we select one element as a representative $\bar{f} \in [\bar{f}]$. As one can see from this formulation, the computation of the Karcher mean requires n geodesic calculations per iteration. This can be very computationally expensive, and thus, we approximate the geodesic using a linear interpolation when computing the facial surface templates. We present all results in Figure 2. We

Figure 3: The first two principal directions of variation (PD1 and PD2) computed in the pre-shape ($\mathcal{C}$) and shape ($\mathcal{S}$) spaces for expressions (1)-(8) in Figure 2.

compare the facial template computed in $\mathcal{S}$ to a standard sample average computed in $\mathcal{C}$ and the curve-based Karcher mean [13]. First, we note from panel (e) that there is a large decrease in energy in each example. The qualitative results also suggest that the 3D face templates computed in $\mathcal{S}$ are much better representatives of the given data than those computed in $\mathcal{C}$ or using the curve-based method. Again, the biggest differences are noticeable around the mouth and eyes. In fact, when looking at panels (b) and (d), it is fairly difficult to recognize the expression; this distinction is much clearer in panel (c).

Summary of Variability and Random Sampling: Once the sample Karcher mean has been computed, the evaluation of the Karcher covariance is performed as follows. First, we optimally register all surfaces in the sample to the Karcher mean $\bar{f}$, resulting in $\{f_1^*, \ldots, f_n^*\}$, and find the shooting vectors $\{v_1, \ldots, v_n\}$ from the mean to each of the registered surfaces. The covariance matrix K is computed using $\{v_i\}$, and principal directions of variation in the given data can be found using standard principal component analysis (singular value decomposition). Note that due to computational complexity, we do not use the Riemannian metric $\langle\langle \cdot, \cdot \rangle\rangle$ to perform PCA; thus, we sacrifice some mathematical rigor in order to improve computational efficiency. The principal singular vectors of K can then be mapped to a surface f using the exponential map, which we approximate using a linear path; this approximation is reasonable in a neighborhood of the Karcher mean. The results for all eight samples displayed in Figure 2 are presented in Figure 3. For each example, we display the two principal directions of variation in $\mathcal{C}$ and $\mathcal{S}$. These paths are sampled at $-2, -1, 0, 1, 2$ standard deviations around the mean. The summary of variability in the shape space more closely resembles deformations present in the original data. This leads to more parsimonious shape models. In contrast to the principal directions seen in $\mathcal{C}$, the ones in $\mathcal{S}$ contain faces with clear facial features.

Given a principal component basis for the tangent space $T_{[\bar{f}]}(\mathcal{S})$, one can sample random facial shapes from an approximate Gaussian model. A random tangent vector is generated using $v = \sum_{j=1}^{k} z_j \sqrt{S_{jj}} u_j$, where $z_j \overset{iid}{\sim} N(0,1)$, S_{jj} is the variance of the jth principal component, and u_j is the corresponding principal singular vector of K. A sample from the approximate Gaussian is then obtained using the exponential map $f_{rand} = \exp_{\bar{f}}(v)$, which again is approximated using a linear path. The results are presented in Figure 4. As expected, the facial surfaces sampled in the shape space are visually preferred to those sampled in the pre-shape space; this is due to better matching of similar geometric features across 3D faces such as the lips, eyes and cheeks.

Symmetry Analysis: To analyze the level of symmetry of a facial surface f we first obtain its reflection $\tilde{f} = H(v)f$, where $H(v) = (I - 2\frac{vv^T}{v^T v})$ for a $v \in \mathbb{R}^3$. Let $F^{*,sh}$ be the geodesic path between f and $\tilde{f}^* = O^*(\tilde{f} \circ \gamma^*)$. We define the length of the path $F^{*,sh}$ as a measure of symmetry of f, $\rho(f) = L(F^{*,sh})$. If $\rho(f) = 0$ then f is perfectly symmetric. Furthermore, the halfway point along the geodesic, i.e. $F^{*,sh}(0.5)$, is approximately symmetric (up to numerical errors in the registration and geodesic computation). If the geodesic path is unique, then

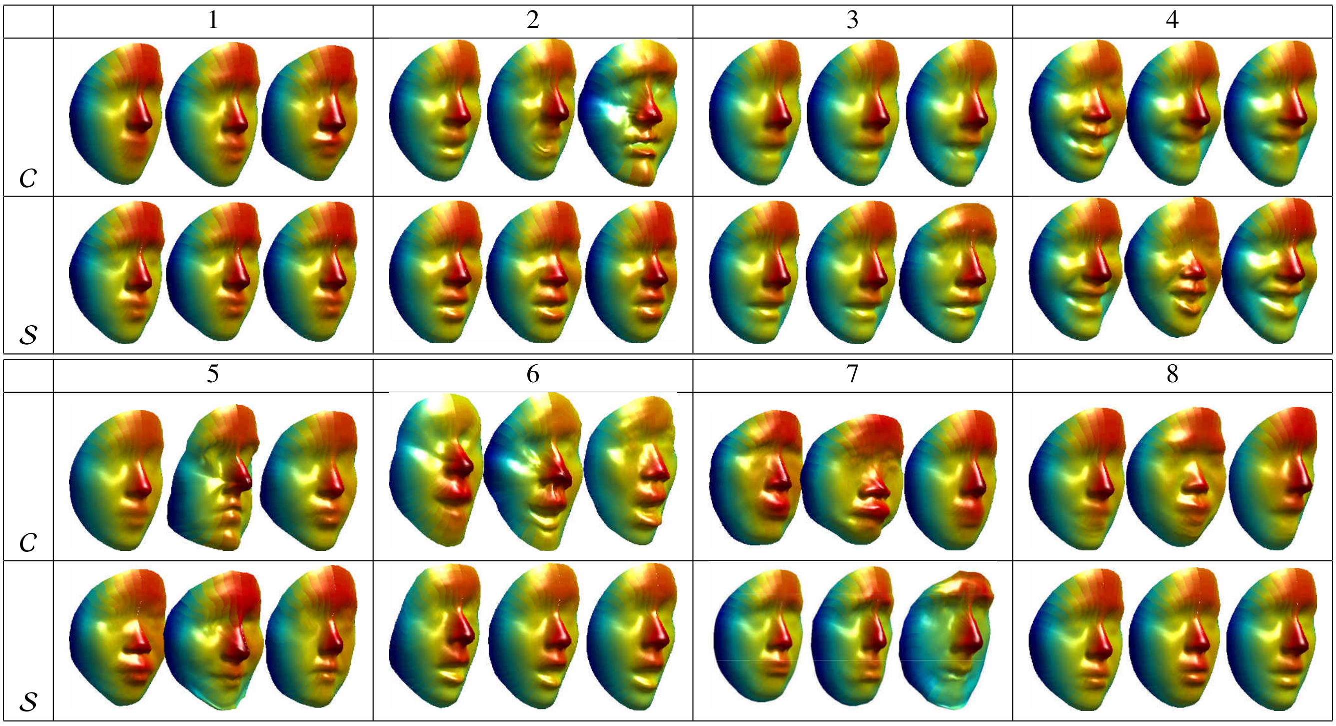

Figure 4: Random samples generated from the approximate Gaussian distribution in the pre-shape ($\mathcal{C}$) and shape ($\mathcal{S}$) spaces for expressions (1)-(8) in Figure 2.

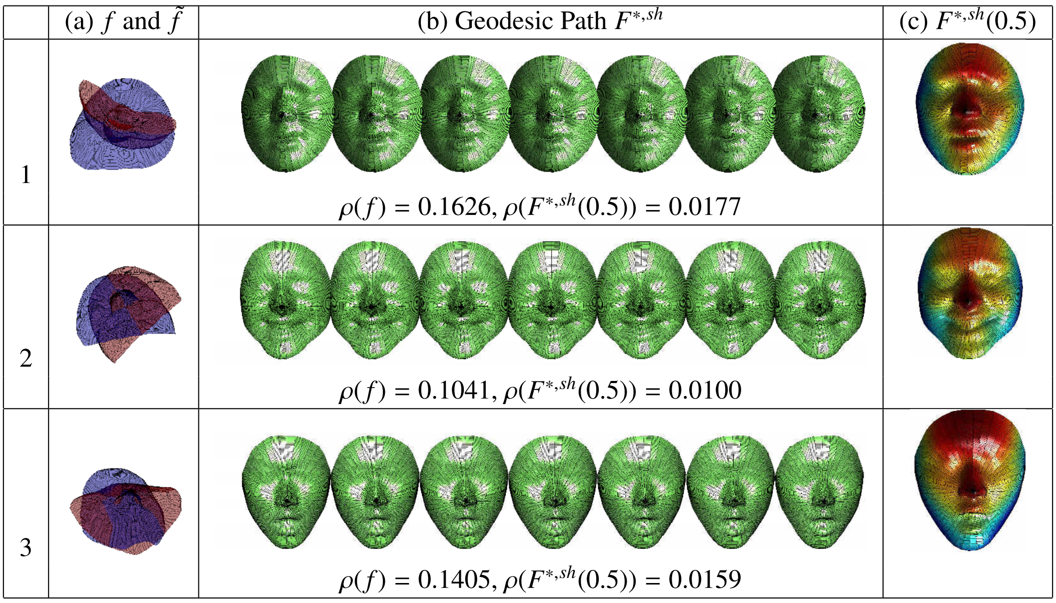

Figure 5: (a) Facial surface f in blue and its reflection $\tilde{f}$ in red. (b) Geodesic path in $\mathcal{S}$ between f and $\tilde{f}$ and the measure of symmetry $\rho(f)$. We also compute the measure of symmetry for the midpoint of the geodesic $\rho(F^{*,sh}(0.5))$, which is expected to be 0 for perfectly symmetric faces. (c) Midpoint of the geodesic.

amongst all symmetric shapes, $F^{*,sh}(0.5)$ is the closest to f in $\mathcal{S}$. Three different examples are presented in Figure 5. The average measure of symmetry for the geodesic midpoints (averaged over all of the presented examples) is 0.0145, which is very close to 0 (perfect symmetry). In the presented examples, the faces are already fairly symmetric. Nonetheless, the symmetrized faces (right panel) have a natural appearance with clearly defined facial features.

Identity and Expression Classification: In the final application, we explore the use of the proposed framework in two different classification tasks. We compare our results to the method presented in [13], which reported state-of-the-art recognition performance in the presence of expressions. We do not compare our performance to any other state-of-the-art methods because many of them are specifically designed for classification experiments (feature based). Our framework is more general as it also allows deformation and statistical modeling of faces. The proposed framework can be tuned to maximize classification performance by extracting relevant elastic features from the computed statistical models, but we believe that this is beyond the scope of the current paper.

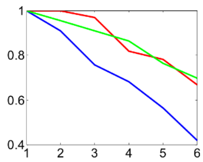

Figure 6: Identity recognition in $\mathcal{C}$ (blue), $\mathcal{S}$ (red), and using [13] (green).

The first task we consider is concerned with classifying expressions. We selected 66 total surfaces divided into six expression groups (11 persons per group): anger, disgust, fear, happiness, surprise and sadness. We computed the pairwise distance matrices in $\mathcal{C}$, $\mathcal{S}$, and using [13]. We calculated the classification performance in a leave-one-out manner by leaving out all six expressions of the test person from the training set. The classification accuracy in $\mathcal{C}$ was 62.12% while that in $\mathcal{S}$ was 74.24%. The classification accuracy of [13] was 68.18%. This result highlights the benefits of elastic shape analysis of hemispherical surfaces applied to this recognition task. It also suggests that considering the radial curves independently, as done in [13], deteriorates the recognition performance. The second task we considered was identity classification irrespective of the facial expression. Here, we added 11 neutral expression facial surfaces (one per person) to the previously used 66 and computed 11×66 distance matrices in $\mathcal{C}$, $\mathcal{S}$, and using the method in [13]. We performed classification by first checking the identity of the nearest neighbor. This resulted in a 100% classification rate for all methods. Figure 6 shows the classification results when accumulating over more and more nearest neighbors (up to six since there are six total expressions for each person). It is clear from this figure that identity classification in the shape space is far superior to that in the pre-shape space. The additional search over Γ allows for the expressed faces to be much better matched to the neutral faces, and in a way provides "invariance" to facial expressions in this classification task. The performance of the proposed method is comparable to [13].

4. Summary and Future Work

We defined a Riemannian framework for statistical shape analysis of hemispherical surfaces and applied it to various 3D face modeling tasks including morphing, averaging, exploring variability, defining generative models for random sampling, and symmetry analysis. We considered two classification experiments, one on expressions and one on person identities, to showcase the benefits of elastic shape analysis in this application. This leads us to several directions for future work. First, we will investigate the use elastic facial shape features, which can further improve the reported classification accuracy. Second, we will utilize the proposed 3D face shape models as priors in processing corrupted or incomplete raw data obtained from 3D scanners. Third, we want to study expression transfer via parallel transport. These tools have not yet been developed for hemispherical surfaces, and to the best of our knowledge, there exist very few automatic methods for this task. Finally, we want to move toward the difficult problem of modeling 3D dynamic faces.

References

[1] Deng Z, Chiang P, Fox P, Neumann U. Animating blendshape faces by cross-mapping motion capture data. In: Interactive 3D Graphics. 2006, p. 43–8.

[2] Huang H, Chai J, Tong X, Wu H. Leveraging motion capture and 3D scanning for high-fidelity facial performance acquisition. ACM Trans Graphics 2011;30(4):74:1–74:10.

[3] Lin I, Ouhyoung M. Mirror MoCap: Automatic and efficient capture of dense 3D facial motion parameters from video. The Visual Computer 2005;21(6):355–72.

[4] Kakadiaris IA, Passalis G, Toderici G, Murtuza MN, Lu Y, Karampatziakis N, et al. Three-dimensional face recognition in the presence of facial expressions: An annotated deformable model approach. IEEE Trans Pattern Analysis and Machine Intelligence 2007;29(4):640–9.

[5] Sandbach G, Zafeiriou S, Pantic M. Local normal binary patterns for 3D facial action unit detection. In: International Conference on Image Processing. 2012, p. 1813–6.

[6] Lu X, Jain AK. Deformation modeling for robust 3D face matching. In: Computer Vision and Pattern Recognition. 2006, p. 1377–83.

[7] Bouaziz S, Wang Y, Pauly M. Online modeling for realtime facial animation. ACM Trans Graphics 2013;32(4):40.

[8] Weise T, Bouaziz S, Li H, Pauly M. Realtime performance-based facial animation. ACM Trans Graphics 2011;30(4):77.

[9] Gupta S, Aggarwal JK, Markey MK, Bovik AC. 3D face recognition founded on the structural diversity of human faces. In: Computer Vision and Pattern Recognition. 2007,.

[10] Wang Y, Liu J, Tang X. Robust 3D face recognition by local shape difference boosting. IEEE Trans Pattern Analysis and Machine Intelligence 2010;32(10):1858–70.

[11] Tsalakanidou F, Tzovaras D, Strintzis MG. Use of depth and colour eigenfaces for face recognition. Pattern Recognition Letters 2003;24(9-10):1427–35.

[12] Bronstein AM, Bronstein MM, Kimmel R. Three-dimensional face recognition. International Journal of Computer Vision 2005;64(1):5–30.

[13] Drira H, Ben Amor B, Srivastava A, Daoudi D, Slama R. 3D face recognition under expressions, occlusions, and pose variations. IEEE Trans Pattern Analysis and Machine Intelligence 2013;35(9):2270–83.

[14] Srivastava A, Klassen E, Joshi SH, Jermyn IH. Shape analysis of elastic curves in Euclidean spaces. IEEE Trans Pattern Analysis and Machine Intelligence 2011;33(7):1415–28.

[15] Ben Amor B, Drira H, Berretti S, Daoudi M, Srivastava A. 4-D facial expression recognition by learning geometric deformations. IEEE Trans Cybernetics 2014;44(12):2443–57.

[16] Samir C, Srivastava A, Daoudi M, Kurtek S. On analyzing symmetry of objects using elastic deformations. In: International Conference on Computer Vision Theory and Applications. 2009, p. 194–200.

[17] Sorkine O, Alexa M. As-rigid-as-possible surface modeling. In: Symposium on Geometry Processing. 2007, p. 109–16.

[18] Kilian M, Mitra NJ, Pottmann H. Geometric modeling in shape space. ACM Trans Graphics 2007;26(3):64.

[19] Zhang Z, Li G, Lu H, Ouyang Y, Yin M, Xian C. Fast as-isometric-as-possible shape interpolation. Computers & Graphics 2015;46(0):244 –56.

[20] Jermyn IH, Kurtek S, Klassen E, Srivastava A. Elastic shape matching of parameterized surfaces using square root normal fields. In: European Conference on Computer Vision. 2012, p. 804–17.

[21] Samir C, Kurtek S, Srivastava A, Canis M. Elastic shape analysis of cylindrical surfaces for 3D/2D registration in endometrial tissue characterization. IEEE Trans Medical Imaging 2014;33(5):1035–43.

[22] Kurtek S, Klassen E, Gore J, Ding Z, Srivastava A. Elastic geodesic paths in shape space of parameterized surfaces. IEEE Trans Pattern Analysis and Machine Intelligence 2012;34(9):1717–30.

[23] Yin L, Wei X, Sun Y, Wang J, Rosato MJ. A 3D facial expression database for facial behavior research. In: Automatic Face and Gesture Recognition. 2006, p. 211–6.

[24] Kurtek S, Samir C, Ouchchane L. Statistical shape model for simulation of realistic endometrial tissue. In: International Conference on Pattern Recognition Applications and Methods. 2014,.