Shell PCA: statistical shape modelling in shell space

Chao Zhang[1], Behrend Heeren[2], Martin Rumpf[2], and William A.P. Smith[1]

[1]Department of Computer Science, The University of York, UK
[2]Institute for Numerical Simulation, University of Bonn, Germany

[1]{cz679, william.smith }@york.ac.uk
[2]{heeren, martin.rumpf }@ins.uni-bonn.de

Abstract

In this paper we describe how to perform Principal Components Analysis in "shell space". Thin shells are a physical model for surfaces with non-zero thickness whose deformation dissipates elastic energy. Thin shells, or their discrete counterparts, can be considered to reside in a shell space in which the notion of distance is given by the elastic energy required to deform one shape into another. It is in this setting that we show how to perform statistical analysis of a set of shapes (meshes in dense correspondence), providing a hybrid between physical and statistical shape modelling. The resulting models are better able to capture nonlinear deformations, for example resulting from articulated motion, even when training data is very sparse compared to the dimensionality of the observation space.

1. Introduction

Statistical models of the shape or appearance of a class of objects are widely used in computer vision and graphics to model the variability over the object class. They can be used to constrain synthesis or analysis problems and their parameter space provides a compact representation which can be used for classification or intuitive editing. Human faces and bodies have proven particularly amenable to statistical modelling. In this case, the source of shape variability is usually changes in identity or dynamic deformation.

Such models seek to satisfy a number of competing goals. The first is to capture the variability in the training data as efficiently as possible, measured by the *compactness* of the model. Hence, one purpose of statistical modelling is dimensionality reduction. The second is to approximate unseen data as accurately as possible, measured by the *generalisation* ability of the model. Finally, we require that the model have *specificity*, i.e. can only generate instances that are plausibly members of the object class being modelled.

Often, the dimensionality of the observations is orders of magnitude greater than the number of samples in the training set. With modern shape capture techniques, typical meshes may contain tens of thousands of vertices. In contrast, the number of training samples that can feasibly be collected is typically only in the tens or hundreds. Hence, the underlying shape space is sampled very sparsely.

In this scenario, the quality of the model is dependent on the validity of the assumed or learnt structure of the manifold on which the data is assumed to lie. For example, PCA assumes that the input data lies on (or can be well approximated by) a hyper-planar manifold, the axes of which are those that capture maximal variance. This makes it optimal with respect to compactness in Euclidean space but a poor choice when the data contains highly nonlinear variations.

In this paper, we use a physically-motivated, nonlinear model of surfaces: thin shells that can undergo tangential and bending distortion. Geodesics in the space of shells have already shown great promise in realistically interpolating and extrapolating between sparse samples of shapes undergoing complex deformations [21, 19, 16]. In other words, shell space potentially provides a useful constraint in modelling the nonlinear variability in a sparsely sampled set of shapes. This motivates our idea of performing statistical shape analysis in shell space. We define a notion of covariance based on the Hessian of an elastic energy term. In analogy to PCA, we extract principal components based on an eigendecomposition of the resulting covariance matrix. The resulting principal components are able to capture nonlinear articulations and complex deformations.

We provide results on human face and body data and evaluate the resulting models in terms of compactness, generalisation and specificity.

2. Related Work

Constructing a statistical shape model requires knowledge of the structure of the underlying shape space, along with a prior over that space. The first ingredient is obtained by using manifold learning or assuming a known manifold structure, the second via a statistical analysis in this space. For a recent review of statistical shape modelling see [6].

Statistical shape modelling Statistical models of shape have been used widely in computer vision and graphics. In a 2D setting, PCA-based models such as Active Shape [9] or Appearance Models [8] provide a parametric representation of shape that can be used for segmentation, tracking and recognition. In a 3D setting, they are typically used for fitting to noisy or ambiguous data or for 3D reconstruction via analysis-by-synthesis. Essentially, the model provides a constraint that significantly reduces the search space for many shape processing problems. Such models begin by transforming raw data into a shape space. This requires establishing a dense point-to-point correspondence between the training samples. This problem has previously been addressed using non-rigid ICP [1], template fitting [31], optical flow [3] and minimum description length [11].

Manifold learning Manifold learning refers to the situation in which data is assumed to lie on an unknown, usually Riemannian, manifold and the structure of this manifold is learnt from training data. PCA can be viewed as parametric manifold learning, in that the structure of the manifold is assumed hyperplanar but its parameters (principal axes) are learnt from data. PCA has been widely used to model shape variation [2, 28, 34]. More complex, usually nonlinear, manifold learning approaches have also been applied to shape modelling. For example, Locally Linear Embedding for matching articulated shapes [26] and Isomap for shape clustering [35]. Multilinear models [33, 10, 5] have proved useful for a number of tasks since they provide a natural way to independently model different sources of variation, for example identity and expression. In contrast to manifold learning approaches, in this paper we use a physical model to define the manifold on which we perform statistical learning. Note that some statistical methods incorporate manifold learning. For example, Hauberg *et al*. compute shape statistics on a learned Riemannian manifold [20].

Statistics on manifolds An alternative to manifold learning is to assume that the Riemannian manifold on which the input data lie is known explicitly. On this manifold, statistical analysis can be performed in a manner that respects the Riemannian geometry of the manifold. This requires Riemannian notions of concepts such as distance, mean value and covariance. In this direction, Pennec [29] showed how to compute these measures for a number of geometric primitives that do not form vector spaces. Fletcher *et al*. [13] went further, building statistical models on such manifolds via computation of the principal geodesics of a set of data. This was done via a linear approximation on the tangent plane. Freifeld and Black [14] use this principal geodesic analysis to build nonlinear models of human body shape variation. They define a Lie group characterising deformations of a triangle mesh and perform statistical analysis on the resulting Riemannian manifold. In contrast to the work we present here, their manifold of deformations is not physically motivated. In all these cases, explicit expressions for the log and exponential maps are required and hence knowledge of the underlying manifold. Of close relation to the work we present here, Rumpf and Wirth [30] describe a covariance analysis of shapes represented as boundary contours of elastic objects.

Geometry of shape space The classical treatment of shape space is due to Kendall [23], in which sets of landmarks are considered points on a shape manifold in which the effects of scale, rotation and translation are factored out. The tangent plane to Kendall's shape space enables linear analysis in which Euclidean distance approximates Procrustes distance. Srivastava *et al*. [32] propose a representation for analysing shapes of curves under an elastic metric. This allows them to define geodesics and geodesic distance between curves in a way that is invariant to transformations including reparameterisation. Killian *et al*. [24] model the space of triangulated shapes. They use a Riemannian metric which measures the stretching of triangle edges in order to compute geodesics. While this space requires an ad hoc regularisation term, Heeren *et al*. [21] propose a purely physically-based notion of geodesics by making use of a bending energy term. Furthermore they introduce a notion of shape extrapolation via the exponential map in Riemannian geometry. The physically-based definition of a shell space and the geometric exploration in [21] is of particular relevance to this work.

3. Thin Shells

A thin shell $\mathcal{S}^\delta$ is a physical material with tiny but positive thickness δ. Mathematically this shell is represented by a smooth surface $\mathcal{S}$ embedded in $\mathbb{R}^3$ which is thought of as the middle layer of the physical material, i.e.

$$\mathcal{S}^\delta = \{p + z\, n(p) \,|\, p \in \mathcal{S}\,, z \in (-\delta/2, \delta/2)\}\,,$$

where $n : \mathcal{S} \to S^2$ denotes the unit normal field.

3.1. Shell Deformations

In the following we will consider a reference material $\bar{\mathcal{S}}^\delta \subset \mathbb{R}^3$ which is in a stress-free state and an elastic deformation $\phi^\delta : \bar{\mathcal{S}}^\delta \to \mathbb{R}^3$ thereof. All quantities corresponding to $\bar{\mathcal{S}}^\delta$ and its middle layer $\bar{\mathcal{S}}$, respectively, will be labelled with a bar, e.g. $\bar{g}$ denotes the first fundamental form of $\bar{\mathcal{S}}$. The corresponding elastic deformation energy is given by

$$\mathcal{W}^\delta[\phi^\delta] = \int_{\bar{\mathcal{S}}^\delta} W_{\text{mem}}(D\phi^\delta)\,\mathrm{d}x\,, \tag{1}$$

where $D\phi^\delta \in \mathbb{R}^{3,3}$ and W_{mem} denotes some elastic energy density.

LeDret and Raoult [25] have shown in the context of Γ-convergence that to leading order, the energy (1) scales *linearly* in the thickness parameter δ and after rescaling with $\frac{1}{\delta}$ is given by the *membrane* energy

$$\mathcal{W}_{\text{mem}}[\bar{\mathcal{S}}, \phi] = \int_{\bar{\mathcal{S}}} W_{\text{mem}}(\mathcal{G}[\phi])\, \mathrm{d}a\,, \tag{2}$$

where $\phi : \bar{\mathcal{S}} \to \mathbb{R}^3$ is the deformation of the middle layer and $\mathrm{d}a = \sqrt{\det \bar{g}}\, \mathrm{d}t$ denotes the area element. The *Cauchy-Green* strain tensor $\mathcal{G}[\phi] \in \mathbb{R}^{2,2}$ is given by

$$\mathcal{G}[\phi] = \bar{g}^{-1} g_\phi\,, \tag{3}$$

where g_ϕ is the intrinsic first fundamental form on the deformed shell $\phi(\mathcal{S})$. The membrane energy density in (2) can be chosen e.g. as

$$W_{\text{mem}}(A) = \frac{\mu}{2}\operatorname{tr} A + \frac{\lambda}{4}\det A - (\frac{\mu}{2} + \frac{\lambda}{4})\log\det A - \mu - \frac{\lambda}{4}\,.$$

Here the trace accounts for local length changes while the determinant accounts for local change of area.

Friesecke *et al*. [15] demonstrated that for *isometric* deformations ϕ (for which $\mathcal{W}_{\text{mem}}[\bar{\mathcal{S}}, \phi] = 0$) the leading order term of (1) is *cubic* in the thickness δ and after rescaling with $\frac{1}{\delta^3}$ is given by an energy term that solely depends on the so-called relative shape operator. This *bending* energy is supposed to account for out of plane bending and changes in curvature. One particular choice is the *Willmore* energy

$$\mathcal{W}_{\text{bend}}[\bar{\mathcal{S}}, \phi] = \int_{\bar{\mathcal{S}}} |\bar{H} - H \circ \phi|^2\, \mathrm{d}a\,, \tag{4}$$

which measures changes in the mean curvature H. Note that H can be regarded as the trace of the matrix-valued relative shape operator. The energy (4) is widely used as bending energy e.g. in the computer graphics community, cf. [17, 19]. Combining (2) and (4) we obtain

$$\mathcal{W}_{\bar{\mathcal{S}}}[\phi] = \mathcal{W}_{\text{mem}}[\bar{\mathcal{S}}, \phi] + \gamma\, \mathcal{W}_{\text{bend}}[\bar{\mathcal{S}}, \phi] \tag{5}$$

where the physical parameter γ is a bending weight that is proportional to the squared thickness δ^2 of the thin shell. It can be viewed as a parameter that balances between membrane energy and bending energy.

For two shells $\bar{\mathcal{S}}, \mathcal{S}$ we define

$$\mathcal{W}[\bar{\mathcal{S}}, \mathcal{S}] = \min_{\phi : \phi(\bar{\mathcal{S}}) = \mathcal{S}} \mathcal{W}_{\bar{\mathcal{S}}}[\phi]\,. \tag{6}$$

We might consider this as an approximation of the squared Riemannian distance in the shell space although it is not symmetric. However, for infinitesimal small deformations the Hessian of $\mathcal{W}$ allows us to retrieve a symmetric, positive-definite Riemannian metric on the space of shell [21]. Note that the formulation is invariant with respect to rigid body motions as $\mathcal{W}_{\bar{\mathcal{S}}}[\phi] = 0$ if $\phi(x) = Qx + b$, $Q \in SO(3), b \in \mathbb{R}^3$.

3.2. Discrete Shells

The above description of shell deformation is based on a smooth shell. In practice, we discretize a smooth shell $\mathcal{S}$ by a triangular mesh S. Furthermore we assume a fixed connectivity constraint that means there is a one-to-one correspondence (dense correspondence) between all nodes and all faces of two meshes as e.g. in [24]. Hence we will represent a triangular mesh S by the vector of its nodal positions $\mathbf{X} \in \mathbb{R}^{3n}$, where n is the number of vertices in the mesh.

We can think of each triangle $T \subset \mathbb{R}^3$ being parametrized over the unit triangle $\omega \subset \mathbb{R}^2$ consisting of the nodes $(0,0)$, $(1,0)$ and $(0,1)$. If $q_0, q_1, q_2 \in \mathbb{R}^3$ are the nodes of T we consider the local linear parametrization $X_T : (t_1, t_2) \mapsto t_1 q_1 + t_2 q_2 + (1 - t_1 - t_2) q_0$ with $0 \le t_1, t_2 \le 1$. Hence we have

$$DX_T = [\partial_{t_1} X_T \,|\, \partial_{t_2} X_T] = [q_1 - q_0 \,|\, q_2 - q_0] \in \mathbb{R}^{3,2}$$

and can deduce an elementwise constant discrete first fundamental form via $G_T = (DX_T)^T DX_T \in \mathbb{R}^{2,2}$. Due to the dense correspondence we can define the discretization of the distortion tensor (3) elementwise via

$$\mathcal{G}[\Phi]_{\bar{T}} = \bar{G}_{\bar{T}}^{-1} G_T.$$

Hence the discrete version of membrane energy is

$$\mathcal{W}_{\text{mem}}[\bar{\mathbf{X}}, \mathbf{X}] = \sum_{\bar{T} \in \bar{\mathcal{S}}} |\bar{T}|\, W_{\text{mem}}(\mathcal{G}[\Phi]_{\bar{T}})\,, \tag{7}$$

where we use the same energy density W_{mem} as in the smooth setting.

For the discrete bending energy, we make use of the *discrete shells* energy proposed in [19]:

$$\mathcal{W}_{\text{bend}}[\bar{\mathbf{X}}, \mathbf{X}] = \sum_{\bar{e} \in \bar{\mathcal{S}}} \frac{(\theta_{\bar{e}} - \theta_e)^2 |\bar{e}|^2}{D_{\bar{e}}}\,, \tag{8}$$

where the sum is over all undeformed edges $\bar{e} \in \bar{\mathrm{S}}$. If T and T' share one edge e we have $D_e = \frac{1}{3}(|T| + |T'|)$ and θ_e is the angle between the triangle normals of T and T'. In [7] the relation between (8) and the Willmore energy is analysed rigorously.

Analogously to (5) the discrete deformation energy $\mathcal{W}[\bar{\mathrm{S}}, \mathrm{S}] = \mathcal{W}[\bar{\mathbf{X}}, \mathbf{X}]$ is given as the sum of (7) and (8):

$$\mathcal{W}[\bar{\mathbf{X}}, \mathbf{X}] = \mathcal{W}_{\text{mem}}[\bar{\mathbf{X}}, \mathbf{X}] + \gamma\, \mathcal{W}_{\text{bend}}[\bar{\mathbf{X}}, \mathbf{X}] \tag{9}$$

Due to the dense correspondence there is a natural elementwise linear deformation between two meshes $\bar{\mathcal{S}}$ and $\mathcal{S}$. Hence - different from (6) - we do not need to optimize for deformations in (9). Note that (9) is invariant with respect to rigid body motions as well.

4. Statistics in Shell Space

A principal component analysis (PCA) relies on notions of averaging and covariance and uses an eigendecomposition of the covariance matrix in order to extract linear principal components. Following [30] we will now introduce for a given set of input data (i) a general notion of an average (depending on a distance measure d) and (ii) a covariance operator as a generalization of a covariance matrix (depending on an inner product g). In section 4.1 and 4.4 we will introduce two particular choices for d and g, respectively, namely the standard Euclidean distance/metric and a physically-based distance/metric induced by the shell deformation energy (9). The corresponding PCA will be refered to as Euclidean PCA in the former and as Shell PCA in the latter setup.

4.1. Averaging

Consider a given set of input data $\mathrm{S}_1, \dots, \mathrm{S}_m$ which we now consider as triangular meshes that are in dense correspondence. Hence we can represent each mesh S_i by its vector of nodal positions $\mathbf{X}_i \in \mathbb{R}^{3n}$. For a given (squared) distance measure $d^2 : \mathbb{R}^{3n} \times \mathbb{R}^{3n} \to \mathbb{R}$ the group average $\hat{\mathbf{X}} \in \mathbb{R}^{3n}$ is given by the Fréchet mean

$$\hat{\mathbf{X}} := \arg\min_{\mathbf{X}} \sum_{i=1}^{m} d^2[\mathbf{X}_i, \mathbf{X}] \,. \tag{10}$$

In the Euclidean setup we have $d^2[\mathbf{X}_i, \mathbf{X}] = \|\mathbf{X}_i - \mathbf{X}\|^2$ and hence

$$\hat{\mathbf{X}}_{\text{euc}} = \tfrac{1}{m} \sum_{i=1}^{m} \mathbf{X}_i \,.$$

However, in the Shell PCA setup we have $d^2[\mathbf{X}_i, \mathbf{X}] = \mathcal{W}[\mathbf{X}_i, \mathbf{X}]$ and (10) becomes a nonlinear optimization problem and the average $\hat{\mathbf{X}}_{\text{shell}}$ has to fulfil the necessary condition

$$F[\hat{\mathbf{X}}_{\text{shell}}] := \sum_{i=1}^{m} \partial_2 \mathcal{W}[\mathbf{X}_i, \hat{\mathbf{X}}_{\text{shell}}] = 0 \,. \tag{11}$$

Here ∂_2 denotes differentiation with respect to the second argument of $\mathcal{W}$.

4.2. Covariance operator

Inherently a PCA is defined on a linear space. Hence we will consider the linear space of nodal displacements $\mathbf{U}_i = \mathbf{X}_i - \hat{\mathbf{X}} \in \mathbb{R}^{3n}$, $i = 1, \dots, m$, from the average $\hat{\mathbf{X}}$. Given an inner product g on $\mathbb{R}^{3n}$ we define the covariance operator

$$\mathbf{Cov}\, \mathbf{V} = \frac{1}{m} \sum_{i=1}^{m} g(\mathbf{V}, \mathbf{U}_i) \mathbf{U}_i \tag{12}$$

and a matrix $C = (C_{ij})_{ij} \in \mathbb{R}^{m,m}$ via

$$C_{ij} := g(\mathbf{U}_i, \mathbf{U}_j) \,. \tag{13}$$

Obviously, C is symmetric and positive semi-definite, that means there is a spectral decomposition

$$C = O \Lambda O^T \,, \quad \Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_m) \,,$$

where $\lambda_1 \geq \dots \geq \lambda_m \geq 0$ are eigenvalues of C and O is an orthogonal matrix, i.e. $OO^T = O^T O = \mathrm{id}_m$. We define $\mathbf{W}_1, \dots, \mathbf{W}_m$ via

$$\mathbf{W}_k := \frac{1}{\sqrt{\lambda_k}} \sum_{i=1}^{m} O_{ik} \mathbf{U}_i \tag{14}$$

if $\lambda_k > 0$ and $\mathbf{W}_k = 0$ else. A straight forward calculation reveals that

$$\mathbf{Cov}\, \mathbf{W}_k = \frac{\lambda_k}{m} \mathbf{W}_k$$

and $g(\mathbf{W}_k, \mathbf{W}_l) = \delta_{kl}$, i.e. $\mathbf{W}_1, \dots, \mathbf{W}_m$ are in fact eigenvectors of $\mathbf{Cov}$. Formally, we can extend $\mathbf{W}_1, \dots, \mathbf{W}_m$ to an orthonormal basis of $\mathbb{R}^{3n}$ with $\mathbf{Cov}\, \mathbf{W}_k = 0$ for $k > m$.

Remark 1: Usually eigenvectors of $\mathbf{Cov}$ are found by a spectral decomposition of the $(3n)$-by-$(3n)$ covariance matrix. However, as in most applications $m \ll 3n$ it is more efficient to decompose $C \in \mathbb{R}^{m,m}$ as defined in (13) and obtain eigenvectors via (14).

Remark 2: Due to the rigid body motion invariance the representation of a discrete shell S by its nodal positions $\mathbf{X}$ is not unique. In fact, S is represented by an equivalence class of nodal position vectors induced by rigid body motions. This issue becomes crucial when defining nodal displacements $\mathbf{U} = \mathbf{X} - \hat{\mathbf{X}}$, as we can construct an arbitrary large displacement by a simple translation. However, this obstacle is overcome by taking $\mathbf{X}$ such that $\|\mathbf{X} - \hat{\mathbf{X}}\|^2 \leq \|\mathbf{Y} - \hat{\mathbf{X}}\|^2$ for all $\mathbf{Y}$ in the equivalence class.

4.3. Principal component analysis

For a data set $\mathbf{U}_1, \dots, \mathbf{U}_m$ the first component $\mathbf{V}_{(1)}$ of a principal component analysis (PCA) is defined as

$$\begin{aligned} \mathbf{V}_{(1)} &= \arg\max_{\|\mathbf{V}\|=1} \sum_{i=1}^{m} g(\mathbf{V}, \mathbf{U}_i)^2 \\ &= \arg\max_{\|\mathbf{V}\|=1} g(\mathbf{Cov}\, \mathbf{V}, \mathbf{V}) \end{aligned} \tag{15}$$

where we used the definition (12) of $\mathbf{Cov}$ in the second equality. If we now write $\mathbf{V} = \sum_k \alpha_k \mathbf{W}_k$, i.e. represent $\mathbf{V}$ in the orthonormal basis as defined in (14), we get $g(\mathbf{Cov}\, \mathbf{V}, \mathbf{V}) = \sum_k \alpha_k^2 \lambda_k$. Hence (15) is equivalent to solving

$$\bar{\alpha} = \arg\max_{\|\alpha\|=1} \sum_{k=1}^{m} \alpha_k^2 \lambda_k \,.$$

As λ_1 is the largest eigenvalue we have $\bar{\alpha} = (1, 0, 0, \dots)$ and hence $\mathbf{V}_{(1)} = \mathbf{W}_1$. Similarly we obtain further components $\mathbf{V}_{(k)}$ as $\mathbf{V}_{(k)} = \mathbf{W}_k$ for $k = 2, \dots, m$. Hence the

principal components are given by the eigendisplacements of **Cov** as defined in (14).

4.4. Choice of metric

We have not specified an inner product g on the space of discrete shells yet. For the Euclidean setup we can define g as the standard scalar product on $\mathbb{R}^{3n}$, i.e.

$$g_{\text{euc}}(\mathbf{U}, \mathbf{V}) = \mathbf{U}^T \mathbf{V} \,.$$

It has been shown in [21] that for a deformation energy as defined in (9) the bilinear form

$$g_{\text{shell}}(\mathbf{U}, \mathbf{V}) = \frac{1}{2}\, \mathbf{U}^T (\text{Hess}\, \mathcal{W})[\hat{\mathbf{X}}_{\text{shell}}, \hat{\mathbf{X}}_{\text{shell}}]\mathbf{V} \qquad (16)$$

in fact defines a metric on the space of discrete shells modulo rigid body motions. Here $\text{Hess}\, \mathcal{W} = \partial_2^2 \mathcal{W} \in \mathbb{R}^{3n,3n}$ denotes the Hessian matrix w.r.t. the second argument of $\mathcal{W}$ which is a positive semi-definite symmetric matrix whose nullspace corresponds to displacements induced by rigid body motions.

4.5. Visualization of principal modes

A PCA model amounts to an average shape $\hat{\text{S}}$ represented by its nodal positions $\hat{\mathbf{X}} \in \mathbb{R}^{3n}$ and its eigendisplacements $(\mathbf{W}_k)_k$ obtained by the PCA. A simple linear visualization of these dominant modes is to compute meshes via $\hat{\mathbf{X}} + t\,\mathbf{W}_k$, where $t \in [-T, T]$.

However, a more reasonable way to express pure nonlinear variations is to use the nonlinear shooting method via the time-discrete exponential map $\text{EXP}_{\mathbf{X}}$ as proposed in [21]. On a smooth manifold $\mathcal{M}$ the exponential map $\exp_{\mathbf{X}}$ maps a tangent vector $\mathbf{V}$ at some point $\mathbf{X}$ to the endpoint $\mathbf{X}(1)$ of the geodesic $t \mapsto \mathbf{X}(t)$ with $\mathbf{X}(0) = \mathbf{X}$ and $\dot{\mathbf{X}}(0) = \mathbf{V}$. Here, we use the exponential map on shell space as a natural extrapolation of shell variations and pick up the corresponding time-discrete definition from [21]. In fact, for the visualizing of (nonlinear) principal modes of variation we consider the elastic average $\hat{\mathbf{X}}$ as start point $\mathbf{X}$ and a (possibly scaled) mode $\alpha \mathbf{W}_k$, $\alpha \in \mathbb{R}$, as the initial velocity $\mathbf{V}$.

4.6. Projection and reconstruction

In an analogous fashion to Euclidean PCA, Shell PCA can also be used for reconstructing shapes from a set of PCA coefficients. Given the elastic average $\hat{\mathbf{X}}$, the eigenvectors $(\mathbf{W}_k)_k$ and some (possibly unseen) shape $\mathbf{X} \in \mathbb{R}^{3n}$, we first compute the nodal displacement $\mathbf{U} = \mathbf{X} - \hat{\mathbf{X}}$. We then project $\mathbf{U}$ onto the Shell PCA space $\mathcal{W}_m$, where $\mathcal{W}_m$ is the linear subspace spanned the principal modes $\mathbf{W}_k, k = 1, \ldots, m$. The projection is given by

$$P_{\mathcal{W}_m} : \mathbf{U} \to P_{\mathcal{W}_m}\mathbf{U} := \sum_{k=1}^{m} g_{\text{shell}}(\mathbf{W}_k, \mathbf{U})\mathbf{W}_k \,,$$

where g_{shell} is the Shell (16). Finally we are able to get the reconstruction $\bar{\mathbf{X}}$ via nonlinear shooting:

$$\bar{\mathbf{X}} := \text{EXP}_{\hat{\mathbf{X}}}(P_{\mathcal{W}_m}\mathbf{U}) \,.$$

4.7. Elastic vs. Riemannian PCA

We consider a linear space of (possibly large) nodal displacements and *not* the tangent space in a Riemannian setup of infinitesimal displacements. This is computationally more efficient (no higher resolution of geodesic paths from the average $\hat{\mathbf{X}}$ to each input shape $\mathbf{X}_i$ is required) and sufficient also for large displacements due to the involved nonlinear elastic deformation energy and its invariance w.r.t. rigid body motions. However, the input displacements $\mathbf{U}_i = \mathbf{X}_i - \hat{\mathbf{X}}$ as well as the resulting principal modes $\mathbf{W}_k$ can be considered as approximate tangent vectors in the Riemannian tangent space at the average shape $\hat{\mathbf{X}}$ whereas this approximation is only valid in a neighborhood of $\hat{\mathbf{X}}$. Furthermore, this motivates the nonlinear shooting of principal modes using the discrete exponential map (cf. sec. 4.5).

5. Implementation and optimization

The implemention is realized as an extension of the open source C++ library QuocMesh[1]. Computationally, the most demanding part is evaluating the shell mean $\hat{\mathbf{X}}_{\text{shell}} \in \mathbb{R}^{3n}$ in (11). This is done by means of Newton's method, i.e. for $F[.]$ as defined in (11) and an initial guess $\mathbf{X}_0$ we compute iteratively

$$DF[\mathbf{X}_k]\,\mathbf{D}_k = -F[\mathbf{X}_k]\,, \quad \mathbf{X}_{k+1} = \mathbf{X}_k + \tau_k \mathbf{D}_k \,,$$

until $\|DF[\mathbf{X}_k]\| < \epsilon$ for some $k < K_{\max}$ and set $\hat{\mathbf{X}}_{\text{shell}} = \mathbf{X}_k$. The stepsize τ_k is determined by Armijo's backtracking line search [27]. Note that each iteration step requires (i) an evaluation of F and DF, i.e. the assembling and addition of m Hessian matrices $\partial_2^2 \mathcal{W}[.,.] \in \mathbb{R}^{3n,3n}$, and (ii) solving a linear system in $3n$ dimensions. To improve the robustness and efficiency of the optimization we make us of an hierarchical scheme based on progressive meshes [22, 18] as it was used in [24]. The computation of the nonlinear mean for the input data shown in Figure 1 with $m = 5$ and $n \approx 6000$ takes 5 minutes on an Intel Core 3.40GHz. However, the running time can be improved substantially by picking up the two-level hierarchical method proposed in [16]. Note that the spectral decomposition of (13) by means of a standard QR algorithm is fast as m is usually small.

6. Experiments

In this section we provide an experimental evaluation of our model. We begin with a qualitative evaluation of

[1] http://numod.ins.uni-bonn.de/software/quocmesh/index.html

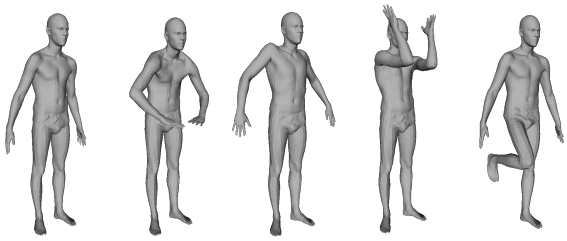


Figure 1: Human bodies data with large nonlinear, articulated deformations from FAUST dataset [4]. These 5 samples are used in our experiment for evaluating the principal modes in section 6.1.

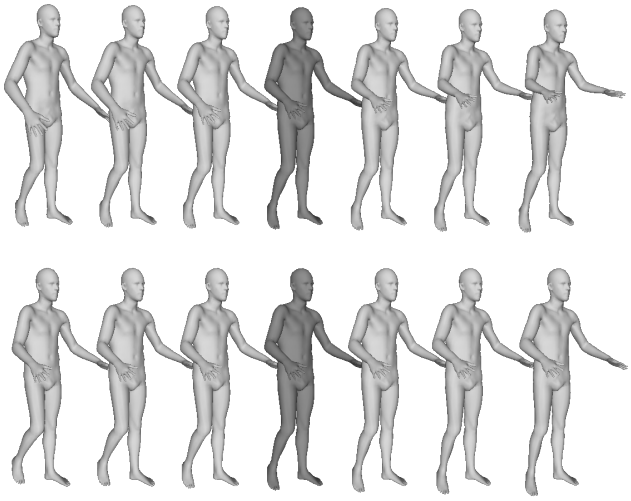
the Shell PCA model by visualising the principal components. Next we compare reconstructions using Euclidean and Shell PCA . Finally, we provide a quantitative comparison in terms of compactness, generalisation and specificity.

We use two datasets in our experiments. The first contains scans of human bodies drawn from the FAUST dataset [4]. The 5 training shapes are shown in Figure 1. The meshes are watertight, genus zero and we apply groupwise simplification to reduce their resolution to 6,000 vertices. For this data, Euclidean PCA fails to obtain a meaningful average, let alone principal components due to the articulated motion. The second is the B3D(AC)^2 [12] dataset containing facial motion sequences in dense correspondence. From this dataset, we extract a subset containing 40 expressions of a single subject. The meshes are genus three (holes for the mouth and eyes) with a boundary. Once again, we apply groupwise simplification to reduce the mesh resolution to 3,000 vertices. Note that in both cases, the training data is extremely sparse and that there are large, nonlinear deformations between shapes.

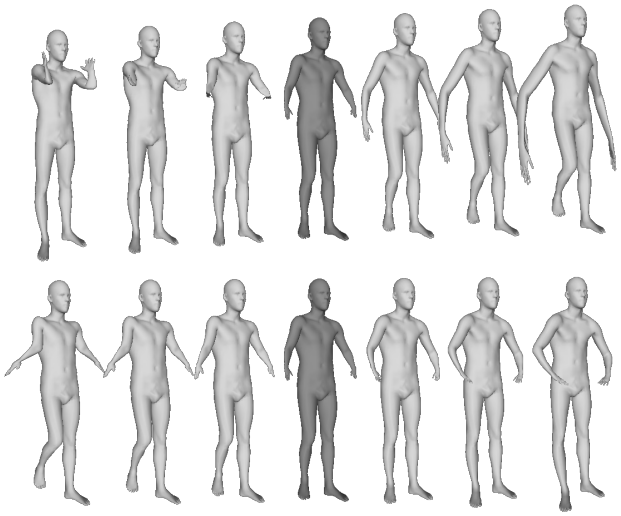
6.1. Qualitative evaluation of PCA

In Figure 2, we show the first two principal components for the body data. The elastic average for Shell PCA and the linear average for Euclidean PCA are depicted in the middle column (shapes in dark grey), respectively. In each row, we show a geodesic path traversing from the average in each direction along the principal component. Note that the shell PCA modes successfully capture the nonlinear, articulated motion. The first mode appears to capture the raising and lowering of the arms and the second the bending of the leg, while Euclidean PCA fails to capture meaningful deformations and leads to degenerated surfaces.

To emphasise the nonlinear nature of the Shell PCA modes, we show the vertex trajectories for the first principal component in Figure 3. This is done by generating a sequence of shapes by nonlinear shooting and plotting the resulting trajectories in red. For comparison, the Euclidean PCA trajectories are shown in blue. Shell PCA clearly leads to nonlinear trajectories.



(a) Shell PCA modes (elastic average in dark grey)



(b) Euclidean PCA modes (linear average in dark grey)

Figure 2: Top 2 modes of body data with (a) Shell PCA and (b) Euclidean PCA. The mean shape in dark color is placed in the middle, and negative and positive shooting (linear combination for Euclidean PCA) results are shown on the left and the right, respectively.

In Figure 4, we show the first three principal components for the facial expression data. We show Euclidean PCA in the first row, Shell PCA with linear combination in the second row, and Shell PCA with nonlinear shooting in the third row. Note that, while both Euclidean and Shell PCA capture similar characteristics in their principal three modes, Shell PCA with nonlinear shooting prevents the surface from folding over itself and retains a more plausible face shape (for example, mode 1 of Euclidean PCA appears to correspond approximately to mode 3 of Shell PCA but Shell PCA preserves a more plausible chin shape in the positive direction).

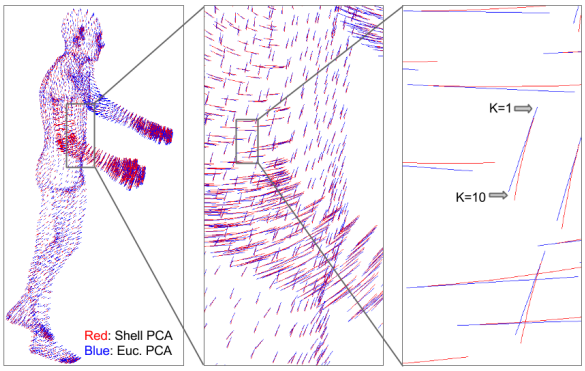


Figure 3: Vertex trajectories of Shell PCA and Euclidean PCA. Note that the vertices follow a curve with the Shell PCA (red) and a straight line with the Euclidean PCA (blue).

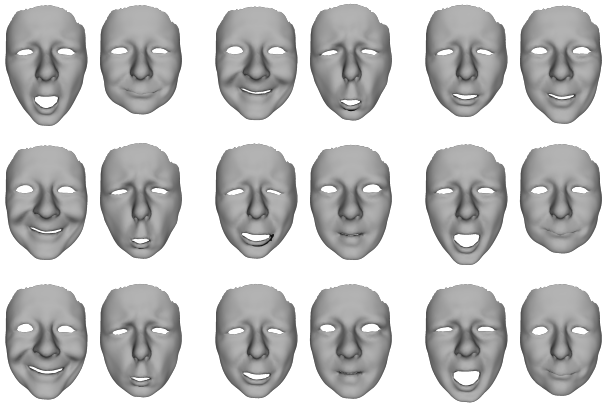


Figure 4: Top 3 modes of the face expression data. (Top: Euclidean PCA; middle, Shell PCA with linear combination; bottom: Shell PCA with nonlinear shooting. Col. 1 and 2 show mode 1, col. 3 and 4 show mode 2, col. 5 and 6 show mode 3.)

6.2. Results of reconstruction

In Figure 5 we show the result of reconstructing a face using an increasing number of model dimensions. The results in the top row are for Euclidean PCA and in the bottom row for Shell PCA. While PCA is optimal in terms of minimising Euclidean distance, it is clear that there is a perceptual improvement in the reconstruction results using Shell PCA. Using only 5 dimensions (column 2), Shell PCA has successfully reconstructed the strong elastic deformation of the smile while the Euclidean PCA reconstruction does not successfully convey the smiling expression.

6.3. Quantitative evaluation of PCA

We now provide a quantitative comparison between a Euclidean PCA and Shell PCA model constructed using the facial expression data.

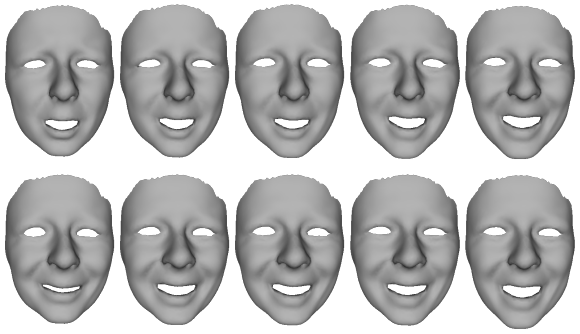


Figure 5: Reconstruction results of using increasing number of parameters. (Top: Euclidean PCA reconstruction results, bottom: Shell PCA reconstruction results. From left to right: dimension = (1,5,10,39), and the last column for ground-truth.)

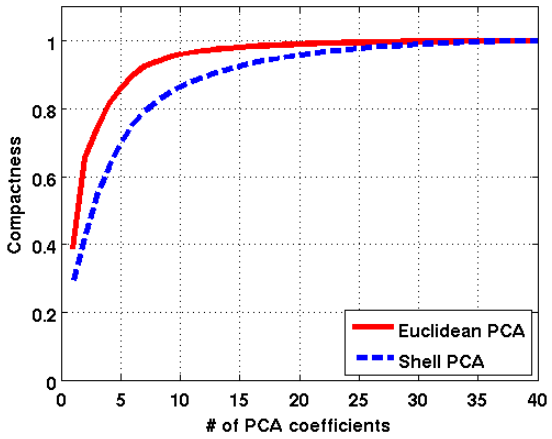
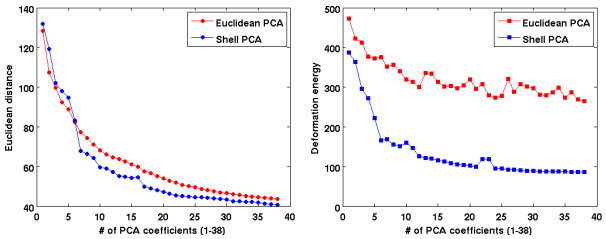


Figure 6: Compactness of Shell PCA and Euclidean PCA.

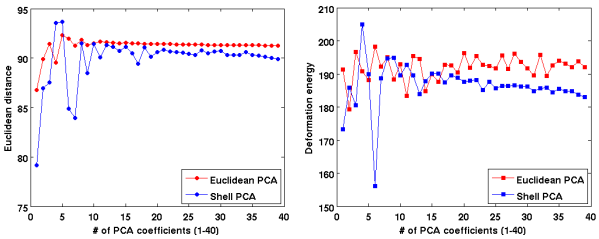
Compactness measures how efficiently a model captures the variability in the training data. Specifically, it is the cumulative variance captured by the top K principal components as a proportion of the total variance within the training data. Hence, we define the compactness as $\alpha(K) = \frac{\sum_{i=1}^{K} \lambda_i}{\sum_{i=1}^{n} \lambda_i}$. We show compactness as a function of K for the expression dataset for Shell and Euclidean PCA in Figure 6. Euclidean PCA is optimal in the sense of least squares, i.e. measuring in the (squared) Euclidean norm. Shell PCA is optimal in the sense of (squared) Riemannian distance in shell space. It is clear that Euclidean PCA is superior in terms of compactness, implying that Euclidean variance in the expression data is more easily captured than elastic variance. However, we should be cautious in how we interpret these plots since the variances in the two models are computed under different distance measures.

Generalisation evaluates the ability of a shape space to



(a) Error in Euclidean distance (b) Error in deformation energy

Figure 7: Generalisation error of Shell PCA and Euclidean PCA.



(a) Error in Euclidean distance (b) Error in deformation energy

Figure 8: Specificity error of Shell PCA and Euclidean PCA.

represent unseen examples of the class. Given a training set of m shapes, generalisation is measured using leave-one-out cross reconstruction of training samples, i.e. the model is learned using $m-1$ samples and the model is then fitted to the excluded sample. The fitting error is measured using two distances: the mean vertex-to-vertex Euclidean distance, and the elastic deformation energy. Generalisation is then reported as mean fitting error averaged over all trials and is a function of the number of model parameters. It is expected that the mean error decreases until convergence as the number of shape parameters increases. In the following experiment, the training data consists of $m = 40$ shapes.

The generalisation error results are shown in Figures 7a (measured in terms of Euclidean distance) and 7b (measured in terms of elastic deformation energy). The errors are computed between the reconstructed and ground truth shape.

It is unsurprising that Shell PCA outperforms Euclidean PCA in terms of minimising elastic deformation energy. However, perhaps more surprising it that Shell PCA seems to perform better in terms of Euclidean distance once > 7 model dimensions are used. This implies that the Shell principal components are more successfully modelling the underlying shape space, enabling better reconstructions of unseen data.

Specificity measures how well a model is able to generate instances that are similar to real data. Firstly, a set of instances are sampled from the learned shape space. Then, for each sample, the error to closest real sample is computed and an average taken over all samples. It is expected that the mean distance increases until convergence with increasing number of parameters. This is because increasing numbers of model dimensions gives the model more flexibility to create more variable shapes, increasing the likelihood that they lie a long way from real samples. In practical applications, generalisation ability and specificity are traded off against each other.

To measure specificity, we consider a PCA model as a probabilistic model by assuming that the data forms a Gaussian cloud in either Euclidean or shell space. Hence, we consider the shape parameter vector to be drawn from a multivariate normal distribution zero mean and standard deviations given by the eigenvalues. We generate 10,000 random samples for both the Euclidean and Shell PCA models. We show the specificity error as a function of the number of parameters in terms of Euclidean distance (Figure 8a) and elastic deformation energy (Figure 8b). In both cases, Shell PCA converges more slowly but gives better specificity for > 20 dimensions. Again, this suggests that the shell principal components are a better model of the underlying shape space.

7. Conclusions

In this paper we have presented an extension of Principal Components Analysis to operate in shell space, providing a hybrid between physical and statistical modelling of shape variation. Principal components are obtained via an eigendecomposition of a covariance matrix arising from an inner product based on the Hessian of an elastic energy. We have shown that such a model is better able to capture the nonlinear variations present in articulated body pose data and face expression data with complex deformations.

In future work we would like to explore applications of the model in areas such as 3D reconstruction, motion tracking and recognition. In addition, we would like to explore an alternative formulation of PCA in shell space, namely computing principal *g*eodesics via statistical analysis in the tangent space computed using the discrete log map. We believe such a model may perform even better when the nonlinear deformations are very large.

Acknowledgements

The work presented here was initiated by a discussion between M. Rumpf and W.A.P. Smith at the Dagstuhl Seminar New Perspectives in Shape Analysis in 2014.

References

[1] B. Amberg, S. Romdhani, and T. Vetter. Optimal step nonrigid icp algorithms for surface registration. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1–8. IEEE, 2007. 2

[2] V. Blanz and T. Vetter. A morphable model for the synthesis of 3d faces. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 187–194. ACM Press/Addison-Wesley Publishing Co., 1999. 2

[3] V. Blanz and T. Vetter. Face recognition based on fitting a 3d morphable model. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 25(9):1063–1074, 2003. 2

[4] F. Bogo, J. Romero, M. Loper, and M. J. Black. FAUST: Dataset and evaluation for 3D mesh registration. In *Proc. CVPR*, pages 3794–3801, 2014. 6

[5] T. Bolkart and S. Wuhrer. Statistical analysis of 3d faces in motion. In *3D Vision-3DV 2013, 2013 International Conference on*, pages 103–110. IEEE, 2013. 2

[6] A. Brunton, A. Salazar, T. Bolkart, and S. Wuhrer. Review of statistical shape spaces for 3d data with comparative analysis for human faces. *Computer Vision and Image Understanding*, 128:1–17, 2014. 2

[7] D. Cohen-Steiner and J.-M. Morvan. Restricted Delaunay triangulations and normal cycle. *Sympos. Comput. Geometry*, pages 312–321, 2003. 3

[8] T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. *IEEE Transactions on pattern analysis and machine intelligence*, 23(6):681–685, 2001. 2

[9] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham. Active shape models-their training and application. *Computer vision and image understanding*, 61(1):38–59, 1995. 2

[10] K. Dale, K. Sunkavalli, M. K. Johnson, D. Vlasic, W. Matusik, and H. Pfister. Video face replacement. *ACM Transactions on Graphics (TOG)*, 30(6):130, 2011. 2

[11] R. H. Davies, C. J. Twining, T. F. Cootes, J. C. Waterton, and C. J. Taylor. 3d statistical shape models using direct optimisation of description length. In *Computer Vision—ECCV 2002*, pages 3–20. Springer, 2002. 2

[12] G. Fanelli, J. Gall, H. Romsdorfer, T. Weise, and L. Van Gool. A 3-d audio-visual corpus of affective communication. *Multimedia, IEEE Transactions on*, 12(6):591–598, 2010. 6

[13] P. T. Fletcher, C. Lu, S. M. Pizer, and S. Joshi. Principal geodesic analysis for the study of nonlinear statistics of shape. *Medical Imaging, IEEE Transactions on*, 23(8):995–1005, 2004. 2

[14] O. Freifeld and M. J. Black. Lie bodies: A manifold representation of 3D human shape. In *Proc. ECCV*, pages 1–14, 2012. 2

[15] G. Friesecke, R. James, M. G. Mora, and S. Müller. Derivation of nonlinear bending theory for shells from three-dimensional nonlinear elasticity by Gamma-convergence. *C. R. Math. Acad. Sci. Paris 336*, 8:697–702, 2003. 3

[16] S. Fröhlich and M. Botsch. Example-driven deformations based on discrete shells. *Computer Graphics Forum 30*, 8:2246–2257, 2011. 1, 5

[17] A. Garg, E. Grinspun, M. Wardetzky, and D. Zorin. Cubic shells. *Symposium on Computer Animation*, pages 91–98, 2007. 3

[18] M. Garland and P. Heckbert. Surface simplification using quadric error metrics. *ACM SIGGRAPH*, pages 209–216, 1997. 5

[19] E. Grinspun, A. N. Hirani, M. Desbrun, and P. Schröder. Discrete shells. In *Eurographics/SIGGRAPH Symposium on Computer Animation*, 2003. 1, 3

[20] S. Hauberg, O. Freifeld, and M. J. Black. A geometric take on metric learning. In *Advances in Neural Information Processing Systems 25*, pages 2024–2032, 2012. 2

[21] B. Heeren, M. Rumpf, P. Schröder, M. Wardetzky, and B. Wirth. Exploring the geometry of the space of shells. *Computer Graphics Forum (Proceedings of SGP)*, 33(5):247–256, 2014. 1, 2, 3, 5

[22] H. Hoppe. Progressive meshes. *Computer Graphics Proceedings*, 1996. 5

[23] D. G. Kendall. Shape manifolds, procrustean metrics, and complex projective spaces. *Bulletin of the London Mathematical Society*, 16(2):81–121, 1984. 2

[24] M. Kilian, N. J. Mitra, and H. Pottmann. Geometric modeling in shape space. *ACM Transactions on Graphics (TOG)*, 26(3):64, 2007. 2, 3, 5

[25] H. LeDret and A. Raoult. The nonlinear membrane model as a variational limit of nonlinear three-dimensional elasticity. *J. Math. Pures Appl.*, 73:549–578, 1995. 3

[26] D. Mateus, F. Cuzzolin, R. Horaud, and E. Boyer. Articulated shape matching using locally linear embedding and orthogonal alignment. In *Proc. ICCV*, 2007. 2

[27] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer, New York / Berlin, 1999. 5

[28] A. Patel and W. A. Smith. 3D morphable face models revisited. In *Proc. CVPR*, pages 1327–1334, 2009. 2

[29] X. Pennec. Intrinsic statistics on riemannian manifolds: Basic tools for geometric measurements. *Journal of Mathematical Imaging and Vision*, 25(1):127–154, 2006. 2

[30] M. Rumpf and B. Wirth. An elasticity-based covariance analysis of shapes. *International Journal of Computer Vision*, 92(3):281–295, 2011. 2, 4

[31] A. Salazar, S. Wuhrer, C. Shu, and F. Prieto. Fully automatic expression-invariant face correspondence. *Machine Vision and Applications*, 25(4):859–879, 2014. 2

[32] A. Srivastava, E. Klassen, S. Joshi, and I. Jermyn. Shape analysis of elastic curves in Euclidean spaces. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 33(7):1415–1428, 2011. 2

[33] D. Vlasic, M. Brand, H. Pfister, and J. Popović. Face transfer with multilinear models. *ACM Transactions on Graphics (TOG)*, 24(3):426–433, 2005. 2

[34] F. Yang, J. Wang, E. Shechtman, L. Bourdev, and D. Metaxas. Expression flow for 3d-aware face component transfer. *ACM Transactions on Graphics (Proc. SIGGRAPH)*, 30(4):60, 2011. 2

[35] D. Yankov and E. Keogh. Manifold clustering of shapes. In *Data Mining, 2006. ICDM'06. Sixth International Conference on*, pages 1167–1171. IEEE, 2006. 2