



**HAL**  
open science

## Learning and Transfer of Deformations

C. Samir, T. Deregnacourt, H. Laga

► **To cite this version:**

C. Samir, T. Deregnacourt, H. Laga. Learning and Transfer of Deformations. Université Clermont Auvergne. 2019. hal-04518231

**HAL Id: hal-04518231**

**<https://hal.science/hal-04518231>**

Submitted on 9 Apr 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Research Highlights

- Spatio-temporal registration
- Dynamic surface analysis
- Transfer and prediction of facial expressions

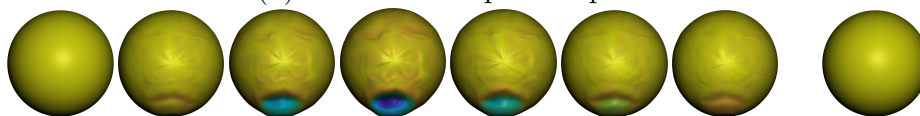
## Learning and Transfer of Deformations

July 2019

T. Deregnaucourt, C. Samir, H. Laga, and A.-F. Yao



(a) Mean 4D surprise expressions.



(b) Surprise expression synthesized on a input (most left) hemisphere.



(c) Surprise expression synthesized on a input (most left) statue head.

**Figure 1:** Learns the mean expression from an input (a) and transfer to (b) or predict (c).

## Abstract

We propose a new framework for modeling surface dynamics, analyzing their elastic shapes, and predicting future dynamics from a given surface. We represent a facial expression, hereinafter referred to as a 4D facial surface, using its initial surface and a sequence of deformation vector fields. We then treat facial expressions as sampling from a mixture of spatio-temporal stochastic processes. This fact, in connection with expression's patterns learning, has made the incorporation of meaningful prior knowledge possible. The proposed framework can be applied to a wide range of applications where the analysis and the prediction of temporal shape dynamics are important. In this paper, we focus on 4D facial expressions and demonstrate the utility of the proposed framework in performing temporal alignment between facial expressions, deforming one 4D surface (*i.e.* a facial expression) onto another, and simulating and synthesizing 4D surfaces.

## 1 Introduction

Automated analysis of dynamic surfaces, *i.e.* surfaces that deform and evolve over time, hereinafter referred to as 4D surfaces, is an important problem in computer graphics and geometry processing. For instance, anatomical objects such as human body shapes, and human body parts such as faces, bones, and subcortical structures in the brain, deform as a result of growth and/or disease progression. In general, one is interested in understanding and modeling the typical deformation patterns of these objects and the variability of these deformations within and across subjects. This is referred to as statistical modeling and analysis of 4D surfaces, where the fourth dimension corresponds to the temporal component.

In this paper, we focus on facial surfaces that deform due to facial expressions, hereinafter referred to as *4D expressions*, and propose a new framework for their statistical analysis. Formally, we are given a finite set of 4D expressions where each expression is represented as a sequence of deforming 3D faces. The goal is to:

- Compute the mean (or average) 4D expression. For example, the same person can smile in different ways. Similarly, different person smile differently. The goal is to learn a typical smile from a data set. This is equivalent to computing a statistical mean from a set of smile observations either on the same person or on different persons.
- Learn a statistical model from a given set of observations
- Simulate new expressions as predictions or sampling from the statistical model.

Developing such framework requires solving important fundamental challenges. Faces exhibit large elastic deformations within the same subject, across different subjects, within the same expression and across different expressions. Also, expressions exhibit large variability in execution rate (*i.e.* speed) within and across subjects. The former requires spatial elastic registration, *i.e.* finding a one-to-one correspondence between two facial surfaces of the same or different individuals. The latter requires accurate temporal registration, which adds another complexity to the problem.

We propose in this paper a new framework that treats the shape of a 4D surface in the setting of Riemannian geometry and stochastic processes. This allows solving three main issues: rate invariant spatio-temporal alignment of facial trajectories, elastic geodesics (*i.e.* optimal deformations) between them, and learning a predictive model from observations. A stochastic process model is then used to learn and simulate new facial expressions as predicted observations. In this context, the proposed formulation is well posed, guaranteeing an optimal solution.

We illustrate this framework on various examples of facial expression analysis including the synthesis of natural 4D expressions. We also show that the proposed approach is able to learn long-term dynamic patterns, a problem that was unsolved by previous state-of-the-art techniques [33, 23, 3]. This allows us to predict an entire facial expression from just its initial surface.

The remainder of the paper is organized as follows. Section 2 reviews the related work. Section 3 overviews the proposed method and summarizes the contributions of this paper. Section 4 presents the different statistical tools needed for the prediction. In Section 5, we describe efficient numerical methods to this end. Section 6 presents the proposed prediction method. Section 7 demonstrates the results and discusses the performance of the proposed framework.

## 2 Related work

The work presented in this paper can be seen as a generalization of the recent advances in statistical analysis of the shape of 3D objects to the statistical analysis of the shape of 4D objects where the fourth dimension refers to time. Thus, in our literature survey, we first review the techniques that have been used for statistical shape analysis and then focus on the analysis of dynamic objects such as facial expressions.

**Statistical analysis of the shape of 3D objects and surfaces.** Statistical analysis and modeling of shapes of objects take their origin in the work of [13], where the shape of an object in Euclidean spaces is defined as all the geometrical information that remains when location, scale and rotation are factored out. Since then,

various methods for computing deformations between two shapes as geodesics in the Kendall shape space have been proposed. [14] investigated a method for computing minimal geodesics in Kendall shape space based on minimizing sums of squares of the Procrustes distance. In a regression context, [15] proposed a method for estimating smoothing splines in the general Kendall shape space using unrolling and unwrapping procedures [12] on spherical data.

While Kendall’s definition of shape space took major strides in shape analysis, it has some limitations, especially when dealing with objects that exhibit large elastic deformations, due to the use of landmarks to define the shape space. Much work has been done in order to find a convenient shape representation that enables simple physical interpretations of deformations (*i.e.* bending and stretching) and efficient numerical methods. Srivastava *et al.* [27, 11] propose a new geometric representation of surfaces based on computational differential geometry. They introduced a Square-Root Normal Field (SRNF) representation for analyzing shapes of surfaces in Euclidean spaces under an elastic metric and compute geodesics between hemispherical surfaces using path-straightening approach [27] and between genus-0 surfaces using SRNF inversion procedure [18]. Other authors have also presented other alternative methods in order to find optimal deformations [21, 17, 25].

**Analysis and simulation of 4D facial expressions.** To the best of our knowledge, statistical modeling and simulation of the trajectories of facial surfaces has not been addressed in the past. There is however an active research for studying regression and fitting data on Riemannian manifolds, see [29, 20, 24, 26].

One of the straightforward techniques for simulating facial expressions would consist of transporting the mean linear deformation to another candidate in order to predict its dynamics [33, 1]. These methods, however, are only successful when the starting surface is close to training data. The results are less convincing when the observations include nonlinear deformations and thus are not similar to the training. Furthermore, we note that the state-of-the-art methods were widely developed based on *short-term* or *one-step* prediction modeling. In contrast to this generic deformation model, we propose an adaptive and more flexible way to capture the *long-term* deformation vector field.

Expression transfer has been studied for prediction inside the same 4D object [22, 30, 16]. [8] use geodesic segments between successive frames in order to construct a piecewise-geodesic path. Due to abrupt velocity change, *i.e.* deformations are not smooth at intermediate points, they extended their work in order to create a smooth path using splines [9]. Other alternative techniques like Bezier curves can be applied for smoothing paths [2]. Although these works achieve good results, they are not applicable to predicting a new 4D object. [7] overcome this issue by first learning the deformation of known 4D triangulated objects. A weighted sum of these defor-

mations is then applied to a new surface in order to predict a new 4D object. This work, however, does not consider the spatial and temporal registration issue, which makes it not applicable in general. [19] use a statistical model in order to transmit an expression. From a reference frame, they compute the shape, pose and expression variation for generating a new face. Despite this process allows to realistically transmit an expression, this works does not allow to analyse expressions due to the lack of temporal registration. Moreover, the large uses of facial specific features make this work not usable for other kind of dynamic surfaces.

### 3 Overview and contributions

Motivated by the need for accurate statistical modeling and simulation of nonlinear elasticity of 4D facial shapes, we propose a new framework for modeling surface dynamics, analyzing their elastic shapes, and predicting future dynamics from a given surface. The framework takes a set of facial expressions  $\alpha_i$  and a new neutral surface  $\alpha_*(0)$ , and produces an animated facial expression  $\alpha_*(t), t \in [0, 1]$ . This is achieved in six steps:

1. Preprocess the parametrized surfaces and paths.
2. Perform spatial registration of the surfaces.
3. Perform temporal registration with geodesics between paths
4. Compute the Fréchet mean per expression.
5. Learn the regression model from the observations. Here, we consider that the trajectory of a surface (4D expressions for example) have a stochastic characteristic. We assume that each trajectory is a realization of a random function sampled from a stochastic process.
6. Generate  $\alpha_*(t), t \in [0, 1]$  from  $\alpha_*(0)$  using the regression model.

The proposed prediction model plays a central role in predicting the dynamic patterns for an unobserved static surface, even far in terms of deformations and parametrization from a common 3D face. It takes different conditions into account to improve the prediction accuracy. This will be illustrated on real as well as on synthetic examples.

### 4 Formulation

We are given a finite set  $\alpha_i, i \in \{1, \dots, n\}$  of independent trajectories on the space of hemispherical surfaces. A hemispherical surface is defined to be a surface that is

homeomorphic to a disc. Then, such surfaces can be represented as:

$$\begin{aligned} f : \Omega = [0, 1] \times [0, 2\pi] &\rightarrow \mathbb{R}^3 \\ s = (r, \theta) &\mapsto (x(r, \theta), y(r, \theta), z(r, \theta)), \end{aligned} \quad (1)$$

where  $r$  is the parameter along the radial direction and  $\theta$  is the parameter along circular direction. We denote by  $\mathcal{F}$  the space of such surfaces, which we equip with the  $\mathbb{L}^2$  metric:

$$\langle f_1, f_2 \rangle = \int_{\Omega} f_1(s) f_2(s) ds. \quad (2)$$

The distance between two surfaces  $f_1, f_2 \in \mathcal{F}$  is then given by:

$$d_{\mathcal{F}}(f_1, f_2) = \|f_1 - f_2\|_2. \quad (3)$$

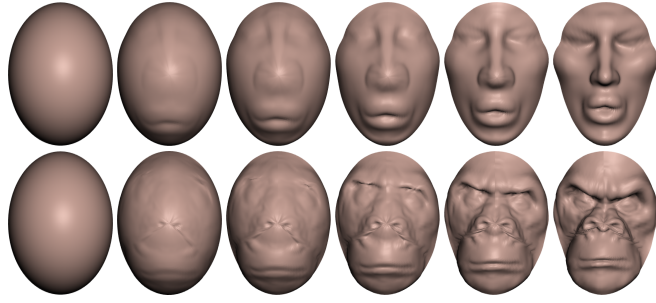
Let  $\alpha_i$  an arbitrary trajectory from  $I = [0, 1]$  to  $\mathcal{F}$ . Equivalently, each  $\alpha_i$  can be seen as a temporal evolution of a unique surface  $\alpha_i(0)$  on  $\mathcal{F}$  whereas all the  $\alpha_i$ 's are realizations of a spatio-temporal process  $A$ . We assume that  $T$ , the space of the  $\alpha_i$ 's is endowed with the  $\mathbb{L}^2$  metric such that  $\langle \alpha_1, \alpha_2 \rangle = \int_I \langle \alpha_1(t), \alpha_2(t) \rangle dt$ .

Now, let us assume that  $\alpha(t)$  only depends on  $\alpha(0)$  and the time instant  $t \in I$ , *i.e.*  $\alpha(t) = g(\alpha(0))(t) + \delta(t)$  where  $\delta$  is a random error and  $g$  is the link function. In other terms, our objective to learning a predictive model becomes a problem of estimating the link function  $g$  from the observations  $(\alpha_1, \dots, \alpha_n)$ . We assume that the function  $g$  is a regression function such that  $g(\alpha(0)) = \mathbb{E}(A|A(0) = \alpha(0))$ .

To solve this problem, we need a set of tools for analyzing facial shapes and trajectories. We propose to solve two main challenging issues. The first one is due to the parametrization of facial surfaces since each surface may have a different curvilinear system. For example, let us suppose that we want to compute the optimal deformation from a surface  $f_1$  to a surface  $f_2$ . In addition to translation and pose, we have to fix their parameterizations (correspondences) before computing the optimal vector field that deforms one surface onto the other. In this example, assuming that we have a numerical method to solve the first issue as a cost, we need to define a good representation and an elastic metric that will allow us to capture diffeomorphic deformations. The shape of facial surfaces has been extensively studied in the past. In the next section, we partially review some of the previous methods and summarize the steps and tools that are required for our study.

## 5 Analyzing Trajectories of Facial Shapes

The notion of deformation between any two surfaces  $f_1$  and  $f_2 \in \mathcal{F}$  must be measured with respect to the deformation between their shapes. To reach such goal one needs



**Figure 2:** Examples of geodesic paths between the most left and the most right surfaces in each row. The intermediate surfaces show the evolving deformation. These deformations are clearly nonlinear as illustrated by the intermediate surfaces along the geodesic paths.

to consider  $\mathcal{F}$  as a Riemannian manifold by putting a Riemannian structure on it in order to define a geodesic distance for our statistical study. In this section, we adapt the ideas from [11] and [27] to study the geometry of facial shapes.

## 5.1 Statistical Analysis of Facial Shapes

We assume that spatial registration problems due to translation, rotation, and scaling between surfaces have been already removed using different optimization methods [5, 32, 6]. We are then interested in finding the best spatial correspondences between two surfaces  $f_1, f_2 \in \mathcal{F}$ . This problem is equivalent to looking for the best re-parameterization of these surfaces, *i.e.* the reparameterization that minimizes the  $\mathbb{L}^2$  distance between them. Recall that the re-parameterization group  $\Gamma$  is defined by:

$$\Gamma = \{ \gamma : \Omega \rightarrow \Omega \mid \gamma(\partial\Omega) = \partial\Omega, |\dot{\gamma}| > 0 \}, \quad (4)$$

where  $\Gamma$  is the set of orientation-preserving diffeomorphisms. This optimal re-parameterization problem can be solved by minimizing the following cost function:

$$E_{\Gamma}(\gamma_1, \gamma_2) = d_{\mathcal{F}}(f_1 \circ \gamma_1, f_2 \circ \gamma_2). \quad (5)$$

Since the re-parameterization group  $\Gamma$  does not act on  $\mathcal{F}$  by isometries, it is hard to find the optimal solution of  $E_{\Gamma}$ . Indeed, from a practical point of view, one has to search for a couple of solutions  $(\gamma_1, \gamma_2)$ , simultaneously. Even if we assume that good estimates may be computed for each candidate separately, we can deduce that finding the couple is challenging. Thus, we use another representation in order to find the optimal solution: the Square-Root Normal Field (SRNF) representation [10]. Here, we summarize the main ideas for clarity and completeness. We refer the reader to the



paper [10] for more details. The SRNF representation of a surface  $f$  is defined as:

$$h : \mathcal{F} \rightarrow \mathcal{Q}$$

$$f \mapsto q = \frac{\dot{f}}{\sqrt{\|\dot{f}\|}} \text{ where } \dot{f} = \frac{\partial f}{\partial r} \times \frac{\partial f}{\partial \theta} \neq 0_{\mathbb{R}^3}.$$

It is essentially the normal scaled by the square-root of its normal. The space  $\mathcal{Q}$  of SRNFs has a natural  $\mathbb{L}^2$  metric. Moreover,  $\Gamma$  acts on  $\mathcal{Q}$  by isometries [10]. So, by minimizing the  $\mathbb{L}^2$  cost between  $q_1$  and  $q_2$ , the SRNF representation of surfaces  $f_1$  and  $f_2$  respectively, the initial re-parameterization problem becomes tractable. Indeed, we rewrite the cost  $E_\gamma$  with  $\tilde{E}_\gamma$  as follows:

$$\tilde{E}_\Gamma(\gamma) = \|q_1 - (q_2, \gamma)\|_2. \quad (6)$$

We denote the optimal solution by  $\hat{\gamma}$  and we further assume that the surfaces  $f_1$  and  $\bar{f}_2 = f_2 \circ \bar{\gamma}$  are spatially registered, accordingly. Once spatially registered, we can compute the geodesic path between them using Euclidean geometry. Two examples of such geodesics are illustrated in Figure 2. For each example, the geodesic starts from the surface at the left and ends at the right. The intermediate points illustrate the evolving deformation field along the geodesic path.

## 5.2 Statistical Analysis of Trajectories

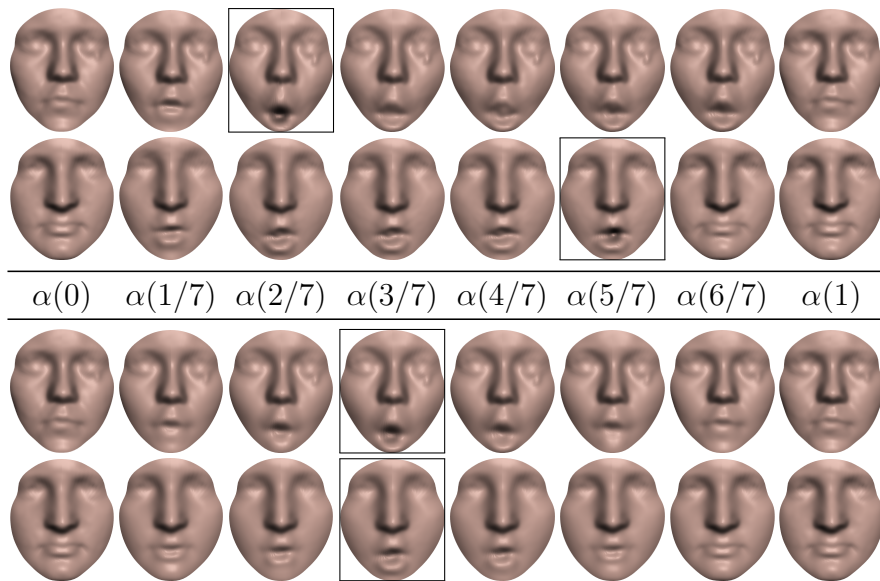
We extend the geometric tools for facial shape analysis of the previous section to study variations between trajectories of facial shapes. Let  $\alpha_1, \alpha_2$  be two different trajectories on  $\mathcal{F}$  with the correct spatial registration estimated using the approach presented in the previous section. Another issue that needs to be solved is how to find a good temporal alignment between  $\alpha_1$  and  $\alpha_2$ ? The problem of temporal alignment is equivalent to finding the best time-warping function that minimizes the  $\mathbb{L}^2$  distance between  $\alpha_1$  and  $\alpha_2$ . To this end, we first define the time-warping group by:

$$\mathcal{E} = \{ \epsilon : [0, 1] \rightarrow [0, 1] \mid \epsilon(0) = 0, \epsilon(1) = 1, \dot{\epsilon} > 0 \}, \quad (7)$$

Using the fact that a spatio-temporal process is a trajectory of surfaces, we estimate the appropriate temporal correspondences by optimizing the cost function:

$$E_{\mathcal{E}}(\epsilon_1, \epsilon_2) = \left( \int_0^1 d_{\mathcal{F}}(\alpha_1(\epsilon_1(t)), \alpha_2(\epsilon_2(t)))^2 dt \right)^{1/2}. \quad (8)$$

This is a complex problem to optimize directly. Since an expression is a path on  $\mathcal{F}$ , which is an Euclidean space, we estimate the temporal alignment using the Square-Root Vector Field (SRVF) representation [28] and dynamic programming. The trajectories  $\alpha_1$  and  $\bar{\alpha}_2 = \alpha_2 \circ \bar{\epsilon}$  are then temporally registered.



**Figure 3:** Examples of surprise expressions. The first two rows show two expressions before temporal registration. The last two rows show the same two expressions but after temporal alignment. For each case, we highlight, using a bounding box, the face where the intensity of the expression is maximal.



**Figure 4:** Estimated Fréchet mean for three prototypical expressions: (a) Angry, (b) Happy, (c) Surprise.

Figure 3 shows an example of temporal alignment between two different expressions. This example illustrates the importance of temporal registration in the proposed framework. The first two rows display a surprise expression from two different persons before temporal alignment. The last two rows show the same expressions after temporal alignment. We can easily check that the original expressions undergo a time-shift that has to be corrected before analysis. In particular, the maximum deformations, highlighted with boxes, occur at different time instants ( $t = 3/7$  for the

first row and  $t = 6/7$  for the second row). Note that this nuisance has been correctly estimated and corrected using the proposed temporal alignment (time instant  $t = 4/7$  for both).

Once the expressions are spatio-temporally registered, we can compute the Fréchet mean of each prototypical expression, see Figure 4.



**Figure 5:** Comparison of the predictive results on two examples: (top) the empirical mean estimator, and (bottom) the Nadaraya-Watson estimator. Examples of surprise deformations as  $\|D(t)\|_2$  from  $t = 0$  to  $t = 0.5$ : training (black), predicted (blue), and mean (red) in the right.

## 6 Parameter Estimation and 4D Expression Simulation

A dynamic surface  $\alpha : [0, 1] \rightarrow \mathcal{F}$  can be interpreted as a trajectory of surfaces. In what follows we assume that all the surfaces, within a trajectory and across trajectories, are correctly registered, *i.e.* they have been normalized for translation and scale, and optimally rotated and re-parameterized. Moreover, we assume that these trajectories have been also temporally registered using the previous section.

With this representation, one can formulate the problem of statistical analysis of 4D expressions as the problem of statistical analysis of trajectories in  $\mathcal{F}$ . However, since we seek to transmit the deformation due to an expression, we focus on deformation vector fields. That is, if we assume that the expressions are smooth enough then  $\alpha$  can be interpreted as a continuous sequence of infinitesimal deformations  $D$  such

that:

$$\alpha(t) = \alpha(0) + D(t), \forall t \in [0, 1]. \quad (9)$$

where  $D$  belongs to  $\mathbb{L}^2([0, 1], \mathcal{F})$ .

Knowing the first surface  $\alpha_*(0)$ , predicting a trajectory  $\alpha_*$  is equivalent to predicting its successive deformations  $D_*$ . Then, instead of solving the original regression estimation  $g$  on  $\alpha$ 's, we reformulate our problem as a regression function  $R$  on deformations such that:

$$g(\alpha_*(0))(t) = \alpha_*(0) + R(\alpha_*(0))(t). \quad (10)$$

Note that the space of  $D_i$  is Euclidean. Thus, one can exploit the state-of-the-art processes for Euclidean spaces. In this paper, we will consider two estimators. The first one is given by:

$$\hat{R}(\alpha_*(0)) = \frac{1}{N} \sum_{i=1}^N D_i$$

We will refer to the resulting deformation as the mean deformation or the generic expression. As shown in Figure 4, this estimator, or equivalently the Fréchet mean when it comes to expressions, may successfully learn the common deformations. However, predicting an expression for a given neutral surface may fail if the input surface is far from the observed ones. To make the prediction step flexible and more adaptive to the initial surface, we propose a second estimator based on the Nadaraya-Watson type kernel estimator [4]:

$$\hat{R}(\alpha_*(0)) = \sum_{i=1}^N \frac{K_h(d_{\mathcal{F}}(\alpha_i(0), \alpha_*(0)))}{\sum_{j=1}^N K_h(d_{\mathcal{F}}(\alpha_j(0), \alpha_*(0)))} D_i, \quad (11)$$

where  $K$  is a Gaussian kernel and  $h$  is a bandwidth. We have chosen here

$$h^2 = \frac{1}{N} \sum_{i=1}^N \left( d_{\mathcal{F}}(\alpha_i(0), \alpha_*(0)) - \frac{1}{N} \sum_{i=1}^N d_{\mathcal{F}}(\alpha_i(0), \alpha_*(0)) \right)^2 \quad (12)$$

## 7 Results

In this section, we apply the proposed framework to learn a 4D expression model from a collection of 4D facial expressions and then use it to simulate the same expression from a given neutral surface. All the experiments are conducted on the BP4D database [31], which  $N = 6$  expressions for each class of expression.

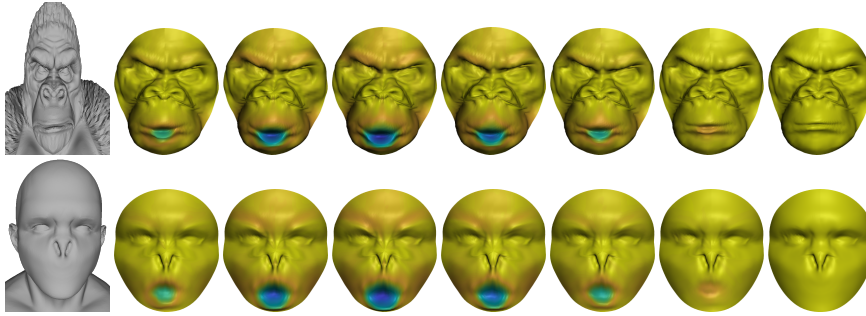
**Spatial and temporal registrations.** In the following, we assume that translation, rotation and scaling variabilities are removed using Generalized Procrustes Analysis [5]. Before analyzing facial shapes, we need a parameterization for each face  $f$  inside a centred geodesic ball: We computed the radial curves of  $f$  at an angle  $\theta$ , noted  $c_\theta$ , which is the intersection between  $f$  and a rotated half-plane  $P_\theta$ . Finally, we parameterized each curve  $c_\theta$  in a way it lies in  $[0, 1]$  using the arc-length. This process allows a full automatic parameterization of  $f$ .

After setting the initial parameterizations for facial shapes (spatial) and trajectories (temporal), we use our method as detailed in Section 5.1 to register two facial surfaces, and then our method detailed in Section 5.2 for temporal registration of trajectories. Now that each trajectory is spatio-temporally registered, we can learn  $\alpha_*$  from available observations using the estimator  $\hat{R}$ . The first model is an estimation by the empirical mean. As shown in Figure 5, this model correctly predicts a surprise expression on the first example but resulted in a less convincing simulation for the second example. This limitation is due to the uniform weights in the model where close and far faces have the same influence. This is not realistic since the input neutral faces are more likely to behave similarly to "closer" faces. In order to overcome this limitation, we use a Nadaraya-Watson estimator as defined in Equation. (11). As shown by the two examples of Figure 5, this model outperforms the estimation by empirical mean. This result is also confirmed with Figure 5 right where we display examples of surprise deformations as  $\|D(t)\|$  from  $t = 0$  to  $t = 0.5$ : training (green), predicted (blue). The same figure displays the mean without and with the scaled standard deviation to give a confident interval of level 95% of the mean. We point out that the predictions are in a small area around the mean.

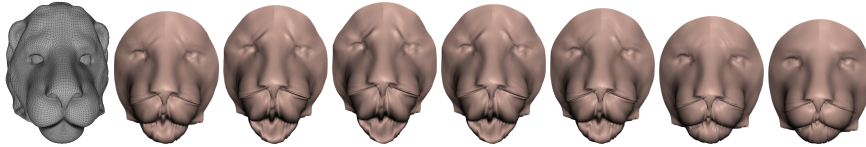
**Extension to more general surfaces.** Now we test if the proposed model can be extended to simulate expressions on input surfaces that are relatively far from being a human face. To partially answer this question, we simulate expressions from non-human faces. Recall that the nose tip has been manually extracted but the remaining tasks remain fully automatic. We display two examples as animated surfaces in Figure 6. One can note that the predicted expressions have a very good quality. Note that, in general, learning models require large and complete datasets in order to capture all variabilities.

## 8 Summary

We have proposed a novel framework to learn dynamic patterns from 4D expressions and predict a trajectory from a neutral surface. We have demonstrated the efficiency of the proposed estimator by applying it to learn six mean deformations from



**Figure 6:** Examples of surprise expressions simulated on non-human faces: (top)) King Kong, (bottom) face without mouth. The input surface is the most left one.



**Figure 7:** Example of simulating the surprise expression on an input (most left) lion head.

a dataset of 4D expressions. From our experiments, we conclude that the proposed approach can be useful for learning and automatically synthesizing new 4D expressions. The same framework can be adapted to transfer a learned animation from available spatio-temporal observations to a newly-observed surface. The framework is sufficiently flexible to be useful in other applications.

We believe that the proposed framework is a promising direction for realistic simulation and synthesis of long-term facial motion and animation. We expect to see more animation methods using the analysis of trajectories in shape space. For instance, fitting Bézier and cubic spline curves could be adapted to simulate trajectories on the shape space from a finite set of indexed control surfaces. The proposed framework includes a spatio-temporal registration component, which can be extended to more general surfaces where the notion of expression makes sense. If the registration fails, the quality of the prediction will drop down. We illustrate this idea in Figure 7 with an example of a surface representing a lion. As expected, the method fails to estimate an accurate registration on the input surface due to the open mouth and the mane. We think that an improvement of the registration step can make the proposed framework more general. Finally, accurate and fast numerical approximation methods could be investigated for the computational aspect when processing larger data sets.

## References

- [1] Volker Blanz and Thomas Vetter. A morphable model for the synthesis of 3d faces. In *Proceedings of the 26th Annual Conference on Computer Graphics and*

- Interactive Techniques*, SIGGRAPH '99, pages 187–194, New York, NY, USA, 1999. ACM Press/Addison-Wesley Publishing Co.
- [2] Christopher Brandt, Christoph von Tycowicz, and Klaus Hildebrandt. Geometric flows of curves in shape space for processing motion of deformable objects. In *Computer Graphics Forum*, volume 35, pages 295–305. Wiley Online Library, 2016.
  - [3] Byungkuk Choi, Roger Blanco i Ribera, J. P. Lewis, Yeongho Seol, Seokpyo Hong, Haegwang Eom, Sunjin Jung, and Junyong Noh. Sketchimo: Sketch-based motion editing for articulated characters. *ACM Trans. Graph.*, 35(4):146:1–146:12, 2016.
  - [4] Luc P. Devroye. The uniform convergence of the Nadaraya-Watson regression function estimate. *Canadian Journal of Statistics*, 6(2):179–191, 1978.
  - [5] L. Dryden and K. V. Mardia. *Statistical Shape Analysis*. 1998.
  - [6] Shaoyi Du, Nanning Zheng, Shihui Ying, Qubo You, and Yang Wu. An extension of the icp algorithm considering scale factor. In *Image Processing, 2007. ICIP 2007. IEEE International Conference on*, volume 5, pages V–193. IEEE, 2007.
  - [7] Lin Gao, Yu-Kun Lai, Dun Liang, Shu-Yu Chen, and Shihong Xia. Efficient and flexible deformation representation for data-driven surface modeling. *ACM Transactions on Graphics (TOG)*, 35(5):158, 2016.
  - [8] Behrend Heeren, Martin Rumpf, Peter Schröder, Max Wardetzky, and Benedikt Wirth. Exploring the geometry of the space of shells. In *Computer Graphics Forum*, volume 33, pages 247–256. Wiley Online Library, 2014.
  - [9] Behrend Heeren, Martin Rumpf, Peter Schröder, Max Wardetzky, and Benedikt Wirth. Splines in the space of shells. In *Computer Graphics Forum*, volume 35, pages 111–120. Wiley Online Library, 2016.
  - [10] Ian H Jermyn, Sebastian Kurtek, Eric Klassen, and Anuj Srivastava. Elastic shape matching of parameterized surfaces using square root normal fields. In *European conference on computer vision*, pages 804–817. Springer, 2012.
  - [11] S.H. Joshi, E. Klassen, A. Srivastava, and I. Jermyn. Removing shape-preserving transformations in square-root elastic (sre) framework for shape analysis of curves. In *EMMCVPR*, pages 387–398, 2007.
  - [12] P. E. Jupp and J. T. Kent. Fitting smooth path to spherical data. *Journal of Applied Statistics.*, 36:34–46, 1987.

- [13] D.G Kendall. Shape manifolds, procrustean metrics and complex projective spaces. *Bulletin of the London Mathematical Society*, 16:81–121, 1984.
- [14] K. Kenobi, I. L. Dryden, and H. Le. Shape curves and geodesic modelling. *Biometrika*, 97:567–584, 2010.
- [15] K.R Kim, I.L Dryden, and H. Le. Smoothing splines on riemannian manifolds, with applications to 3d shape space. 2018.
- [16] Yeara Kozlov, Derek Bradley, Moritz Bächer, Bernhard Thomaszewski, Thabo Beeler, and Markus H. Gross. Enriching facial blendshape rigs with physical simulation. *Comput. Graph. Forum*, 36(2):75–84, 2017.
- [17] Sebastian Kurtek and Hassen Drira. A comprehensive statistical framework for elastic shape analysis of 3d faces. *Computers & Graphics*, 51:52–59, 2015.
- [18] Hamid Laga, Qian Xie, Ian H Jermyn, and Anuj Srivastava. Numerical inversion of srnf maps for elastic shape analysis of genus-zero surfaces. *IEEE transactions on pattern analysis and machine intelligence*, 39(12):2451–2464, 2017.
- [19] Tianye Li, Timo Bolkart, Michael J Black, Hao Li, and Javier Romero. Learning a model of facial shape and expression from 4d scans. *ACM Transactions on Graphics (TOG)*, 36(6):194, 2017.
- [20] X. Li and Y. Chang. Non-uniform interpolatory subdivision surface. *Applied Mathematics and Computation*, 324:239–253, 2018.
- [21] P. W. Michor and D. Mumford. Riemannian geometries on spaces of plane curves. *Journal of the European Mathematical Society*, 8:1–48, 2006.
- [22] Igor S. Pandzic. Facial motion cloning. *Graph. Models*, 65(6):385–404, November 2003.
- [23] Chandan Pawaskar, Wan-Chun Ma, Kieran Carnegie, John P. Lewis, and Tae-hyun Rhee. Expression transfer: A system to build 3d blend shapes for facial animation. In *IVCNZ*, pages 154–159. IEEE, 2013.
- [24] C. Samir, P.-A. Absil, A. Srivastava, and E Klassen. A gradient-descent method for curve fitting on Riemannian manifolds. *Foundations of Computational Mathematics*, 12:49–73, 2012.
- [25] C. Samir, A. Srivastava, M. Daoudi, and E. Klassen. An intrinsic framework for analysis of facial surfaces. *International Journal of Computer Vision*, 82:80–95, April 2009.



- [26] O. Sander. Geodesic finite elements of higher order. *IMA Journal of Numerical Analysis*, 2015.
- [27] A. Srivastava, E. Klassen, S. H. Joshi, and I. H. Jermyn. Shape analysis of elastic curves in Euclidean spaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 33:1415–1428, 2011.
- [28] Anuj Srivastava, Eric Klassen, Shantanu H Joshi, and Ian H Jermyn. Shape analysis of elastic curves in euclidean spaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(7):1415–1428, 2011.
- [29] J. Wllner, E. Nava Yazdani, and P. Grohs. Smoothness properties of lie group subdivision schemes. *Multiscale Model. Simul.*, 6:493–505, 2007.
- [30] Jian Zhang, Jun Yu, Jane You, Dapeng Tao, Na Li, and Jun Cheng. Data-driven facial animation via semi-supervised local patch alignment. *Pattern Recognition*, 57:1 – 20, 2016.
- [31] Xing Zhang, Lijun Yin, Jeffrey F. Cohn, Shaun J. Canavan, Michael Reale, Andy Horowitz, Peng Liu, and Jeffrey M. Girard. Bp4d-spontaneous: a high-resolution spontaneous 3d dynamic facial expression database. *Image Vision Comput.*, 32(10):692–706, 2014.
- [32] Timo Zinßer, Jochen Schmidt, and Heinrich Niemann. Point set registration with integrated scale estimation. In *International conference on pattern recognition and image processing*, pages 116–119, 2005.
- [33] Michael Zollhöfer, Justus Thies, Pablo Garrido, Derek Bradley, Thabo Beeler, Patrick Pérez, Marc Stamminger, Matthias Nießner, and Christian Theobalt. State of the art on monocular 3d face reconstruction, tracking, and applications. *Comput. Graph. Forum*, 37(2):523–550, 2018.