Statistically Motivated 3D Faces Reconstruction

C. Basso and T. Vetter
Curzio.Basso@unibas.ch Thomas.Vetter@unibas.ch
University of Basel, Computer Science Department

Abstract

Facial surgeons require a vast amount of knowledge about the human face, in order to decide how to reconstruct an injured or traumatized part. Does this knowledge consist only of explicit notions about face anatomy, or is there also an implicit knowledge of the normal appearance of a face? And if the answer to the latter question is affirmative, is there a way for a computer to automatically learn this implicit knowledge and exploit it to predict the optimal reconstruction of part of a face? A set of examples of human faces, acquired as 3D surfaces, can be used to build a statistical model, which can generate synthetic faces or analyze novel ones. Such a model is also able to reconstruct the missing part of a face in a statistically meaningful way: any face which can be generated by such a model has a certain probability, and the optimal reconstruction can be defined as the one which maximizes it, given the available data. However, since the model is built from a finite set of examples, it cannot generate any possible face, and therefore the purely statistical reconstruction will not perfectly fit the available data, leading to discontinuities at the boundary between the missing and the available data. In order to avoid this, we look for the surface that explicitly satisfies the continuity constraints at the boundary, and at the same time approximates the first derivatives of the statistical prediction. This leads to a partial differential equation, which can be solved numerically as a linear system. As an example, we show how this method performs reconstructing the noses of a set of test faces.

1. Introduction

We define the problem of surface reconstruction as the one of estimating the optimal shape of a three-dimensional surface's region, which for some reason is not considered trustworthy, or is altogether missing. In Computer Graphics, this problem often arises in the context of post-processing three-dimensional surfaces acquired with range scanning devices, in order to remove artifacts due to errors or limitations of the acquisition process. However, the problem has a more general application scope than just post-processing raw acquisition data, and is especially interesting for medical applications, e.g. to reconstruct the shape of a traumatized region.

The optimality of a given reconstruction is set by two requirements:

- Continuity (or smoothness) at the boundaries of the reconstruction;
- Minimal distance from the true surface (known as reconstruction error).

Variational approaches typically satisfy the first requirement: the optimal reconstruction is defined as the one which satisfies the given boundary conditions and minimizes a certain functional (e.g. membrane or thin-plate energy) of the surface over the invalid region. With respect to the second requirement, these methods work well if the energy to be minimized captures some of the properties of the surface, for example when modeling a surface resulting from a well-understood physical phenomenon. But in case we want to reconstruct something with a more complex structure, like the nose of a face, a variational approach, by itself, is not sufficient, as shown in fig.1(a,b).

The minimization of the reconstruction error, on the other hand, is the explicit goal of statistical reconstruction methods. In the case of faces, since the surface belongs to a specific

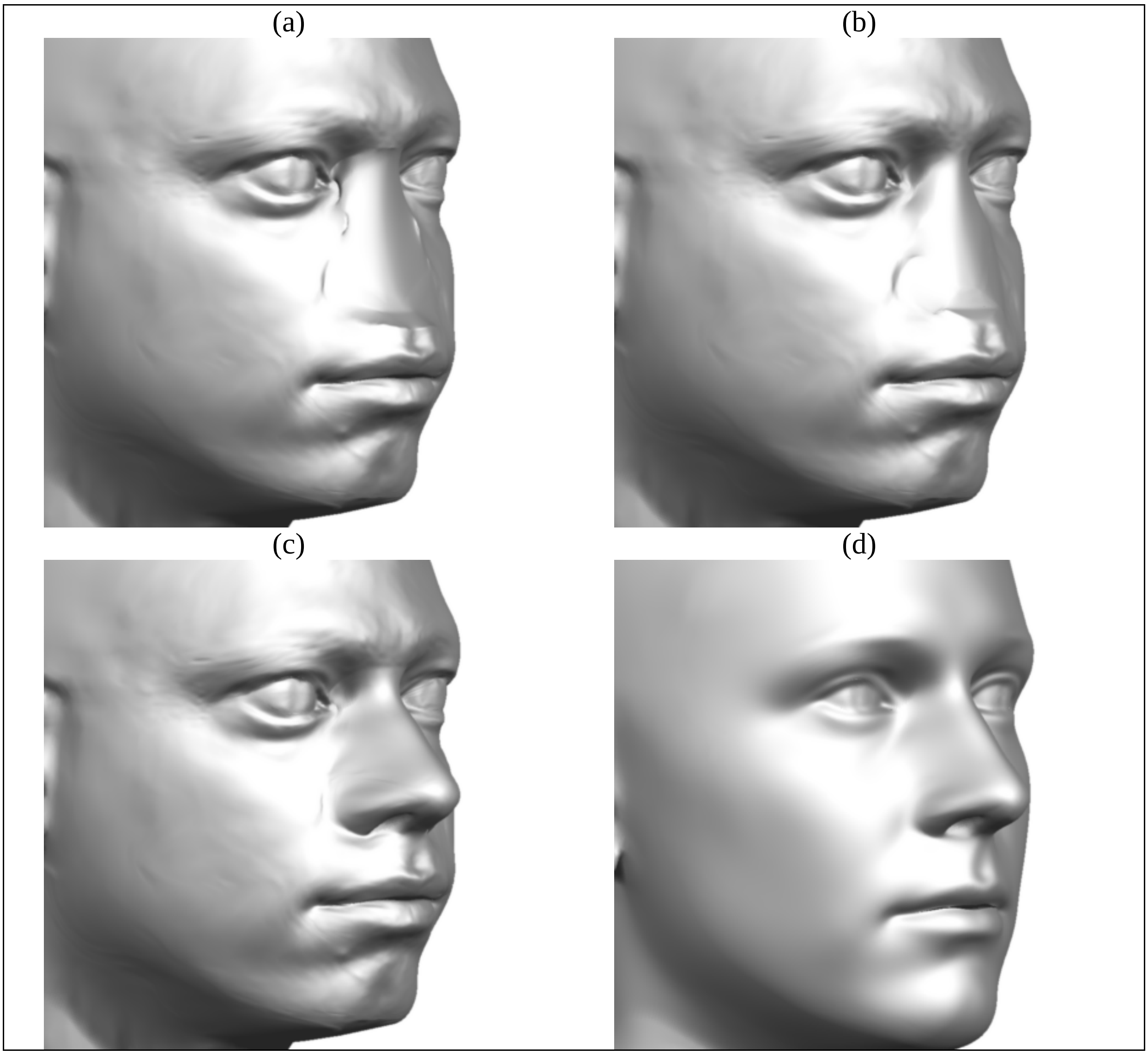


Figure 1. Examples of variational reconstruction applied to the nose. (a)-(b) Reconstructions obtained by minimization of the membrane and the thin-plate energy, respectively. Note that the former results in continuity, while the latter in smoothness at the boundaries. In both cases the shape of the nose is not recovered. (c) Result of a Poisson reconstruction using (d) as a guidance surface.

class of objects, we can build an example-based statistical model in order to infer a realistic nose from the known portion of the face. The disadvantage of this approach is that it works globally, reconstructing a full face that is generally different from the input. Trying to cut the nose from the reconstruction and paste it into the input face will result in discontinuities at the boundaries.

To summarize, each of the two approaches satisfies one of the two requirements: our goal is to combine them into a method that satisfies both. As we will see, the statistical reconstruction can be further refined solving a variational problem, so that it also achieves continuity at the boundaries.

1.1 Previous Works

We formalize the problem by explicitly defining the three-dimensional surface as a function from a parameterization domain S to R^3, where S is a closed subset of R^2. Part of this surface is invalid or unknown, and we denote by $\Omega \subset S$ the corresponding of the parameterization domain. The subset of S outside Ω is denoted by $\partial\Omega$, and on this subset the surface is known. We denote by f^* the valid surface defined only on $\partial\Omega$, and by f the unknown, full surface defined on the whole S. The problem consists in finding the function f satisfying the two requirements mentioned in the previous section: (1) continuity, which we can write as $f|_{\partial\Omega} = f^*$, and (2) minimal reconstruction error. However, we will first consider a class of methods that address the problem by minimizing various kinds of surface energies rather than the reconstruction error.

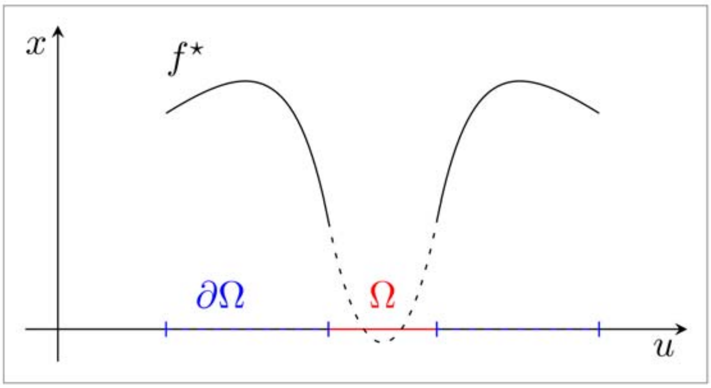


Figure 2. A schematic representation of the continuous formulation for a one-dimensional surface embedded in R^2. We denote by $\partial\Omega$ not just the boundaries of Ω, but also the whole subset of S that is outside $\partial\Omega$, so that $S = \Omega \cup \partial\Omega$. The solid line is the known part of the surface f^*, while the dashed line is the unknown part which has to be reconstructed.

Variational methods define the optimal f as the one minimizing a given functional as well as satisfying the boundary constraints. Typical choices for the functional of the surface are the membrane and the thin-plate energy, and depending on the choice more strict constraints can be imposed on the boundary, requiring for instance C^1-continuity (smoothness). The variational problem is transformed to a partial differential equation (PDE), which is discretized over a polygonal mesh and solved as a sparse (typically non-symmetric) linear system. Examples are the methods of [1], which uses a simple discretization of the thin-plate energy, and of [2], using the Willmore energy. Both methods ensure C^1-continuity at the boundaries.

Note that the discretization of the PDE requires the topology of the polygonal mesh in Ω to be defined, which is not true if Ω corresponds to an actual hole in the data; in such a case, a necessary preprocessing step is the identification and triangulation of the hole. However, this is not necessary if the surface f has been registered against a template surface g as in [3], where the author makes use of the difference $(f - g)$ to obtain f as a deformation of g, using volume splines.

The problem of the missing topology can also be avoided by defining the surface implicitly rather than explicitly. In this case S and Ω are subset of R^3, and the surface is defined by a level set of the function f from S to R. The invalid regions have still to be identified, but since the topology of the surface is implicitly defined in f, the reconstruction of the latter yields both the shape and the topology of the surface. In [4] the authors use as f a clamped signed distance function, and define a function discriminating the valid regions from the invalid ones. The reconstruction is obtained by a diffusion of both functions in the invalid regions. A more sophisticated approach uses anisotropic diffusion ([5], an extension to surfaces of the image inpainting method presented in [6]). Also using implicit functions is the method of [7], which minimizes the L^1-norm, measured on the implicit surface, of a distance function from the valid data. The method is in fact intended for surface reconstruction, but hole filling descends as a side effect.

All the methods discussed so far do not try to explicitly minimize the reconstruction error in the invalid region; in fact, the solution depends only on a small region surrounding the hole, if not simply on its boundary. Since all other information about *f* is discarded, it is difficult for such a method to obtain convincing results if the invalid surface has a complex structure. Sharf et al. ([8]) propose a method that uses the valid surface to predict the structure of the invalid region: using an implicit representation for the surface, the invalid voxels are filled in a multi-resolution approach with examples extracted from the available surface. The choice depends on the context of the voxel, defined by its valid neighboring voxels. This is an interesting approach, in that it makes use of the information provided by the whole valid surface, but it would still be unable to solve the nose problem we mentioned previously.

Our method is based on [9] and [10]. The first work shows how, given a surface registered with a 3D morphable model ([11]), the reconstruction problem can be formulated in a statistical framework. As previously mentioned, since the surface is registered, the topology of the mesh in the invalid region is known even in the case of a hole. This method however does not ensure a solution continuous at the boundaries. In the second work a variational method for image editing is presented, by which the gradient of an image can be used as a guidance field for reconstructing part of another image: as a result the reconstruction has the appearance of the guiding image, but is continuous at the boundaries. This method has also been applied to polygonal meshes in [12], but only to the aim of editing the shape of the meshes, rather than to the reconstruct their surface.

We will show how the statistical reconstruction of [9] can be used as guidance in a variational problem that yields a reconstruction close to the statistical one but continuous at the boundaries. We will first describe in detail the statistical reconstruction method (section 2); then, in section 3, we will describe the variational method of [10] applied to a polygonal 3D mesh, and show how to use it for refining the results of the statistical method.

2. Statistical Reconstruction

In order to approach the problem from a statistical point of view, we assume that the surfaces we want to reconstruct belong to a specific object class (e.g. human faces, teeth, cars), which can be modeled statistically. That is, we assume there exist a parameterized class of functions $f_{\boldsymbol{\alpha}}$, where $\boldsymbol{\alpha}$ is a vector holding the parameters of the surface, such as for any surface *f* of this class there is a choice of $\boldsymbol{\alpha}$ for which $f = f_{\boldsymbol{\alpha}}$. Moreover, not all choices of the parameters $\boldsymbol{\alpha}$ are equally probable, but rather $\boldsymbol{\alpha}$ follows a certain probability distribution $P(\boldsymbol{\alpha})$, which depends on the type of model used. We will use a linear model

$$f_{\boldsymbol{\alpha}} = \sum_{i=0}^{k} \alpha_i f_i \ ,$$

where the surfaces are a linear combination of $k+1$ registered examples f_i. By registered we mean that: (1) all the exemplar surfaces share a common parameterization domain *S*, and (2) that every point (u,v) of the common domain corresponds to a unique physical point in all examples (e.g. the tip of the nose). Of course, when 3D models of faces are acquired, they are not registered, and they have to be processed in order to be used for building such a linear model. We refer to [11] and [13] for methods on how to register 3D models of faces, and we will assume in the following that we have access to such a database of registered examples.

Given such a model and a partial measurement f^* of an unknown complete surface $\hat{f}$, our goal is to obtain a reconstruction $f_{\boldsymbol{\alpha}}$ as close as possible to $\hat{f}$ (for the moment we will not consider the continuity at the boundaries). More precisely, we want to minimize the generalization error defined as the mean reconstruction error over the whole objects class:

$$E = \left\langle \iint \left\| f_\alpha - \hat{f} \right\|^2 \right\rangle$$

given a certain probability distribution $P(\hat{f})$ of the surfaces. Note that this equation sets the goal, but that it cannot be explicitly used for estimating the reconstruction since $\hat{f}$ is unknown; what can be done is to minimize the error over the known surface $\left\| f_\alpha|_{\partial\Omega} - f^* \right\|^2$, but it is well known that this method may lead to overfitting. In the following section we describe a statistical approach that reduces this problem.

2.1 3D Morphable Models

We will now proceed to define the linear model in discrete terms, in order to apply it to the case where the surfaces are actually described by triangular (or in general polygonal) meshes. Remember that such a mesh consists of an ordered list of vertices in R^3 connected by edges: we denote by $\mathbf{p}_i$ the i-th vertex of the mesh, and by i^* the set of indices of its neighboring vertices. In the discrete formulation we denote by S, Ω and $\partial\Omega$ the sets of indices of, respectively, all the vertices of the mesh, the invalid vertices, and the valid vertices. If $i \in \partial\Omega$, the i-th vertex is valid, and we denote its known position by $\mathbf{p}_i^*$, in analogy with f^*.

If the mesh has N vertices, we can represent its full shape (the positions of all the vertices) in two different ways, which we will use depending on the context. One way is to put all vertices positions in a $3 \times N$ matrix $\mathbf{P}$, where the i-th column of $\mathbf{P}$ corresponds to the i-th vertex $\mathbf{p}_i$:

$$\mathbf{P} = (\mathbf{p}_1, \ldots, \mathbf{p}_N) \in R^{3\times N}$$

Alternatively, we can use a $3N$-dimensional vector $\mathbf{v}$, where the vertices positions are stacked one after the other:

$$\mathbf{v} = (\mathbf{p}_1^T, \ldots, \mathbf{p}_N^T)^T \in R^{3N\times 1}$$

The two forms are obviously related, and when necessary we can convert one to the other by rearranging their elements. In particular, we denote the conversion from $\mathbf{P}$ to $\mathbf{v}$ as $\mathbf{v} = vec(\mathbf{P})$, where the *vec* operator stacks the columns of $\mathbf{P}$ one after the other into a single column vector. The opposite conversion is denoted by $\mathbf{P} = \mathbf{v}^{(3)}$: given an n-dimensional vector $\mathbf{x}$, we denote by $\mathbf{x}^{(m)}$ the $m \times (n/m)$ matrix obtained by splitting $\mathbf{x}$ in contiguous blocks of m elements, and rearranging them in (n/m) columns.

The term 3D Morphable Model was introduced by [11] to denote a statistical linear model for 3D meshes, learned from a set of registered examples: a generic surface shape $\mathbf{v}$ is modeled as

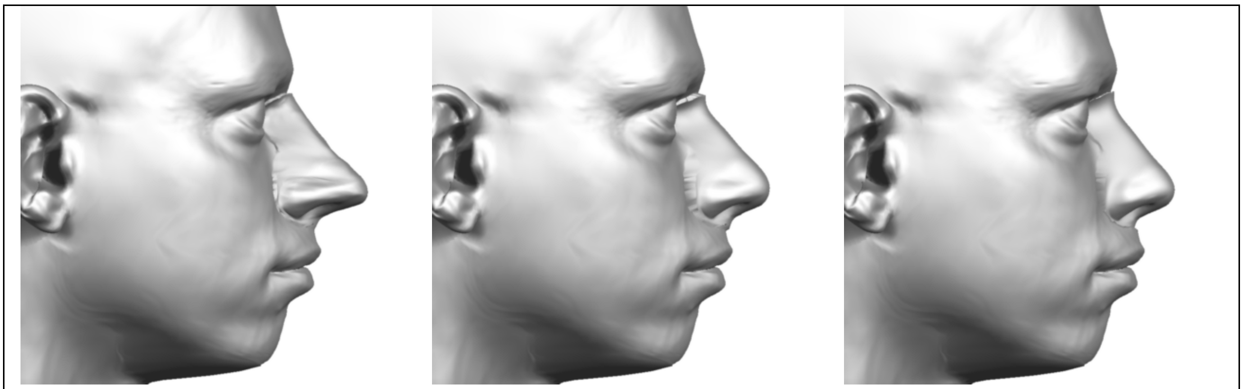
$$\begin{aligned} \mathbf{v} &= \bar{\mathbf{v}} + \mathbf{C}\cdot\boldsymbol{\alpha} + \boldsymbol{\varepsilon} \\ \boldsymbol{\alpha} &\propto N(0,\mathbf{I}) \\ \boldsymbol{\varepsilon} &\propto N(0,\sigma^2\mathbf{I}) \end{aligned}$$

where $\bar{\mathbf{v}}$ is the mean of the model, $\mathbf{C}$ is the $3N \times k$ generative matrix, $\boldsymbol{\alpha} \in R^k$ is the vector holding the coefficients of the model, and $\boldsymbol{\varepsilon}$ is a gaussian noise (this last term expresses the fact that in general our model will not be able to generate any possible shape $\mathbf{v}$). Both $\boldsymbol{\alpha}$ and $\boldsymbol{\varepsilon}$ are normally distributed, and the model parameters ($\bar{\mathbf{v}}, \mathbf{C}, \sigma$) are estimated from the set of examples (see subsection 4.1).

Assume now an incomplete vector $\mathbf{v}^*$ is given, built with the known vertices positions $\mathbf{p}_i^*$: the reconstruction problem consists of inferring, given a model defined by the parameters ($\bar{\mathbf{v}}, \mathbf{C}, \sigma$), the optimal model coefficients $\boldsymbol{\alpha}$ for $\mathbf{v}^*$, where the optimality is defined in terms of the posterior probability:

$$P(\boldsymbol{\alpha}|\mathbf{v}^*) = \frac{P(\boldsymbol{\alpha})P(\boldsymbol{\alpha}|\mathbf{v}^*)}{P(\mathbf{v}^*)} \propto P(\boldsymbol{\alpha})P(\mathbf{v}^*|\boldsymbol{\alpha})$$

That is, the optimal model coefficients will be the one with maximal probability, conditioned to $\mathbf{v}^*$. The two terms on the right hand side of the equation can be derived from the definition



of the model. We know that the probability distribution of the model coefficients is normal with covariance equal to the identity, and therefore we have the following expression:

$$P(\boldsymbol{\alpha}) = (2\pi)^{-k/2} \exp\left\{ -\frac{\|\boldsymbol{\alpha}\|^2}{2} \right\}$$

The second component of the posterior probability, $P(\mathbf{v}^* | \boldsymbol{\alpha})$, is the likelihood of the observed data given the estimate, and we can derive it again from the definition of the model, taking into account the invalid dimensions by a matrix **L** which maps them to zero:

Figure 3. The nose is statistically reconstructed with different values of σ. Without noise the reconstruction correspond to an orthogonal projection into the subspace of the linear model; increasing the value of the noise regularizes the reconstruction until, for high values, the reconstruction is equal to the model mean. Note also the evident discontinuities at the boundary of the reconstructed region.

$$P(\mathbf{v}^* | \boldsymbol{\alpha}) = (2\pi\sigma^2)^{-3l/2} \exp\left\{ -\frac{\|\mathbf{L}(\mathbf{v}^* - \overline{\mathbf{v}} - \mathbf{C}\cdot\boldsymbol{\alpha})\|^2}{2\sigma^2} \right\}$$

The matrix **L** is a diagonal $3N \times 3N$ matrix, with element equal to one if it correspond to the *x*, *y* or *z* coordinate of a known vertex, and zero otherwise. It is convenient to maximize the posterior probability $P(\boldsymbol{\alpha} | \mathbf{v}^*)$ by minimizing its log-inverse. Defining the matrix **Q** = **LC**, the log-inverse is

$$E = -\log P(\boldsymbol{\alpha} | \mathbf{v}^*) \propto -\log P(\boldsymbol{\alpha}) - \log P(\mathbf{v}^* | \boldsymbol{\alpha}) =$$

$$= \|\boldsymbol{\alpha}\|^2 / 2 + \|\mathbf{L}\cdot(\mathbf{v}^* - \overline{\mathbf{v}}) - \mathbf{Q}\cdot\boldsymbol{\alpha}\|^2 / 2\sigma^2$$

The effect of the first term of *E* is to penalize choices of $\boldsymbol{\alpha}$ which, according to the model, have low probabilities; the second term is a cost that depends on how well **v** fits the available data. Observe that the parameter σ is practically weighting the importance of a good fit to the data with respect to the likelihood of it: for very low values of σ, the likelihood will have no influence on the result, and the model will be very flexible; on the contrary, for very high values, the original data will have no importance and the likelihood will be maximized by setting all the model coefficients to zero.

It can be shown (see appendix A for details) that decomposing the matrix **Q** via Singular Value Decomposition (SVD) as

$$\mathbf{Q} = \mathbf{U}\,\mathbf{W}\,\mathbf{V}^{\mathrm{T}}$$

the global minimum of *E* is given by

$$\boldsymbol{\alpha} = \mathbf{V}(\mathbf{W}^2 + \sigma^2\mathbf{I})^{-1}\mathbf{W}\mathbf{U}^T\mathbf{L}(\mathbf{v}^* - \overline{\mathbf{v}})$$

A complete reconstruction of $\mathbf{v}^*$ is therefore

$$\mathbf{v} = \overline{\mathbf{v}} + \mathbf{C}\cdot\mathbf{V}(\mathbf{W}^2 + \sigma^2\mathbf{I})^{-1}\mathbf{W}\mathbf{U}^T\mathbf{L}(\mathbf{v}^* - \overline{\mathbf{v}})$$

Here again we can note how the noise parameter acts as a regularization factor, smoothly constraining the degrees of freedom of the model. If the noise of the model were null, the equation above would reduce to the familiar orthogonal projection onto the subspace spanned by the columns of **C**:

$$\mathbf{v} = \overline{\mathbf{v}} + \mathbf{C}\cdot\mathbf{Q}^+\cdot\mathbf{L}(\mathbf{v}^* - \overline{\mathbf{v}}), \text{ with } \mathbf{Q}^+ = \mathbf{V}\cdot\mathbf{W}^{-1}\cdot\mathbf{U}^T$$

where $\mathbf{Q}^+$ is called the pseudoinverse of **Q**. But as the noise value increases the reconstruction gets closer to the mean of the model, as shown in fig. 3.

3. Refinement via Variational Methods

The statistical reconstruction described in the previous section is global, in the sense that provides an estimate not only for the positions of the invalid vertices, but also for the valid ones. Since in general the statistical model will not be able to achieve a perfect reconstruction of the valid vertices, the constraint $f|_{\partial\Omega} = f^*$ used throughout the introduction is not respected, and therefore the method cannot guarantee continuity at the boundaries of Ω. We will now examine a variational method by which the problem can be overcome.

3.1 Laplacian Reconstruction

As noted in the introduction, variational methods reconstruct the surface f in the invalid region Ω by minimizing a functional, and an often used one is the membrane energy, in which case the variational problem is

$$\min_f \iint_\Omega \|\nabla f\|^2 \text{ with } f|_{\partial\Omega} = f^*.$$

Thinking to the surface in terms of an elastic membrane, the minimization of the integral on the left ensures a surface with minimal tension, while the conditions on the right ensure the continuity of f at the boundaries of Ω. The solution of the above problem is defined by its Euler-Lagrange equation, a PDE of simple form (a Laplace equation with Dirichlet boundary conditions):

$$\Delta f|_\Omega = 0 \text{ with } f|_{\partial\Omega} = f^*.$$

To solve the problem numerically we define a discrete approximation of the Laplacian operator and rewrite the above equation as a system of linear equations. With the notation introduced in the previous section, we define a discretization of Δf known in Computer Graphics as the umbrella-operator ([14], [15]):

$$U(\mathbf{p}_i) = \frac{1}{|i^*|}\sum_{j\in i^*}(\mathbf{p}_j - \mathbf{p}_i)$$

where $|i^*|$ is the valence (the number of neighbors) of $\mathbf{p}_i$. $U(\mathbf{p}_i)$ is the average of the distance vectors between $\mathbf{p}_i$ and its neighbors. By representing the shape of the mesh with the matrix **P** introduced in the previous section, the action of U on the whole mesh can be compactly represented in matrix form; the discrete approximation of the Laplacian for all the vertices of the mesh can be written as

$$U(\mathbf{P}) = \mathbf{K}\mathbf{P}^T$$

where $\mathbf{K} \in R^{N\times N}$ is a sparse matrix, with

$$(\mathbf{K})_{ii} = -1 \text{ and } (\mathbf{K})_{ij} = 1/|i^*| \forall j \in i^*$$

which depends on the connectivity of the mesh. As an example, let us consider a tetrahedron. Each vertex has three neighbors, so that $|i^*| = 3$ for any vertex, and

$$\mathbf{K} = \frac{1}{3}\begin{pmatrix} -3 & 1 & 1 & 1 \\ 1 & -3 & 1 & 1 \\ 1 & 1 & -3 & 1 \\ 1 & 1 & 1 & -3 \end{pmatrix}$$

Note that some of the rows of **K** would be typically discarded, since they operate on vertices outside Ω. Without changing the shape of the matrix **K**, the same effect can be obtained defining a diagonal matrix $\mathbf{\Lambda} \in R^{N \times N}$, with elements $\lambda_i = \{0,1\}$ whose values depend on whether $i \in \Omega$ or not. Using $\mathbf{\Lambda}$ the Euler-Lagrange equation can be written in matrix form, together with the boundary conditions, as

$$(\mathbf{\Lambda} + (\mathbf{I} - \mathbf{\Lambda})\mathbf{K})\mathbf{P}^T = \mathbf{\Lambda}\mathbf{P}^{*T},$$

where $\mathbf{P}^*$ is as $\mathbf{P}$ a $3 \times N$ matrix, where the *i*-th column is either $\mathbf{p}_i^*$ if the *i*-th vertex was valid, or a zero vector otherwise. If the vertex is outside, then $\lambda_i = 1$, and its linear equation will reduce to

$$\mathbf{p}_i = \mathbf{p}_i^*$$

On the other hand, if the vertex is inside, then $\lambda_i = 0$ and the corresponding equation will be

$$\frac{1}{n}\sum_{j \in i^*}(\mathbf{p}_j - \mathbf{p}_i) = 0$$

The linear system is sparse, separable in the *x,y,z* coordinates, and efficiently solvable by LU decomposition (e.g. using the UMFPACK library, see [16]).

3.2 Poisson Reconstruction

We show in fig. 1(a) the results of solving the linear system for one of our test cases: as we would have expected the minimization of the membrane energy only cares about continuity at the boundaries. Trying to use a different functional does not improve the situation either: see the result obtained with a thin-plate energy, in fig. 1(b).

The method for image editing presented in [10] offers a simple way to improve the results: if we know something about the gradient of *f* in Ω, we can use this information along with the boundary conditions. Let $\mathbf{w}=(u,v)$ be an R^2 vector field, called guidance field, defined on Ω. If we modify the membrane energy to take **w** into account, the variational problem becomes

$$\min_f \iint_\Omega \|\nabla f - \mathbf{w}\|^2 \text{ with } f|_{\partial\Omega} = f^*,$$

and the Euler-Lagrange equation becomes a Poisson equation with Dirichlet boundary conditions

$$\Delta f|_\Omega = \nabla \cdot \mathbf{w} = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \text{ with } f|_{\partial\Omega} = f^*.$$

where $\nabla \cdot \mathbf{w}$ is the divergence of the field.

This equation can be further simplified in the case where the guidance field is itself the gradient of a known function *g*, and we have therefore $\nabla \cdot \mathbf{w} = \nabla \cdot \nabla g = \Delta g$. The above Poisson equation reduces again to a Laplace equation:

$$\Delta(f - g)|_\Omega = 0 \text{ with } f|_{\partial\Omega} = f^*,$$

which can be solved as shown above. It is useful to reformulate the problem by defining the displacement function $h = f - g$:

$$\Delta h\big|_{\Omega} = 0 \text{ with } h\big|_{\partial\Omega} = f^* - g\big|_{\partial\Omega},$$

(a) (b)

Figure 4. The nose is first reconstructed statistically, shown in (a). To remove the evident discontinuities the statistical reconstruction can be used as guidance surface by the Poisson reconstruction, yielding the result in (b).

from which is evident that the reconstruction will be the sum of the function *g* (which provides the structure) and of *h* (which provides continuity at the boundary of Ω). In a discrete formulation the reconstruction will given by **P**=**G**+**H**, where **H** is the solution to

$$(\mathbf{\Lambda} + (\mathbf{I} - \mathbf{\Lambda})\mathbf{K})\mathbf{H}^T = \mathbf{\Lambda}\mathbf{H}^{*T} \text{ with } \mathbf{H}^* = \mathbf{P}^* - \mathbf{G}.$$

Fig. 1(c) shows how the solution follows the shape of the guidance function, while at the same time respecting the boundary constraints.

As already mentioned, in [10] the method was used for image editing rather than reconstruction, since for this latter use a good guess of the guidance function *g* from the known surface would be needed. In our case, this can be provided by the result of the statistical reconstruction.

The combination of the two methods yields the following scheme:

• given a model defined by the parameters ($\overline{\mathbf{v}}$,**C**,σ), the statistical reconstruction **v** of $\mathbf{v}^*$ is computed, as in shown in section 2.1;

• setting **G**=$\mathbf{v}^{(3)}$, the solution **H** of the above sparse linear system is computed;

• the final reconstruction is given by **P**=$\mathbf{v}^{(3)}$+**H**

This method has all the advantages of the statistical reconstruction, while at the same time removing the discontinuities at the boundary of Ω. Moreover, results show an improvement of the performance in terms of generalization error (see section 5).

4. Technical Aspects

Before examining the results, we want to describe more in detail two technical issues:

- In section 4.1 we explain how the three parameters ($\overline{\mathbf{v}}$,**C**,σ) of the model are estimated, in particular the noise level;
- In section 4.2 we examine a more general definition for a discrete Laplacian operator than the umbrella operator, and consider if it can yield any advantage.

4.1 Model Estimation

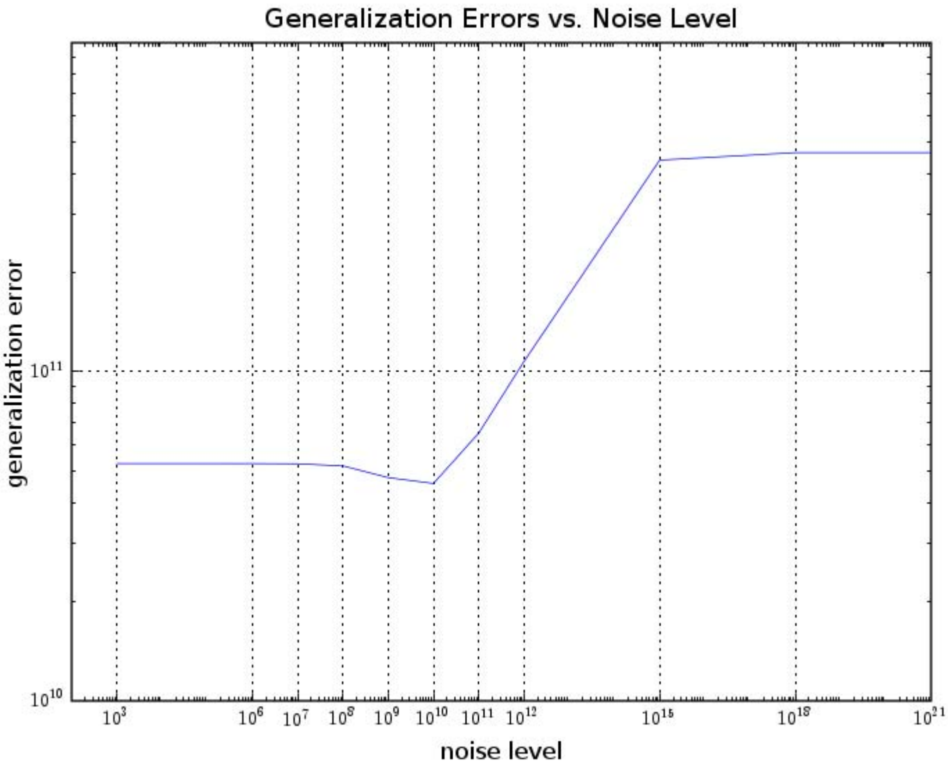


Figure 5. Variation of the generalization error with respect to different values of the noise variance. The error is computed by 5-fold cross validation on the reconstruction of the invalid vertices, given the valid ones. The global minimum is at about 10^{10}.

The optimal model parameters are estimated from a set of training examples $\{\mathbf{v}_i | i=1\ldots m\}$, arranged into a data matrix $\mathbf{A}=(\mathbf{v}_1,\ldots,\mathbf{v}_m)$. The optimal estimate of $\overline{\mathbf{v}}$ is always the sample mean:

$$\overline{\mathbf{v}} = \frac{1}{m}\sum_i \mathbf{v}_i \ .$$

The other two parameters $\mathbf{C}$ and σ can be estimated from the centered data matrix

$$\tilde{\mathbf{A}} = (\mathbf{v}_1 - \overline{\mathbf{v}}, \ldots, \mathbf{v}_m - \overline{\mathbf{v}})$$

using an EM-algorithm as shown in [17] and [18]. However, in the case of $m<3N$ the estimation is simplified by the fact that the estimated value for σ will always be zero: in this particular case the matrix $\mathbf{C}$ can be estimated by Principal Component Analysis (PCA). In practice, given an SVD of the centered data matrix $\tilde{\mathbf{A}} = \mathbf{UWV}^T$, the optimal estimate of $\mathbf{C}$ is:

$$\mathbf{C} = \mathbf{UW}.$$

Our training set has $m=200$ examples, but $3N$ is of the order of 10^6, and therefore we estimated $\mathbf{C}$ as above. It must be clear however that, although the optimal estimate of the noise variance is zero, this is due to the relatively small number of training vectors with respect to their dimensionality, and not because the model can generate any possible shape. Therefore we estimated the optimal value for σ by performing a 5-fold cross-validation (see [19]) of the model with different values of σ and choosing as optimal the value for which the reconstruction error of the invalid vertices is minimal. As shown in fig. 5, the reconstruction error has a minimum when the noise variance is equal to 10^{10}, and we will use this value for the experiments detailed in section 5.

4.2 Discrete Laplacian Design

As shown in [15], there is a more general form of the umbrella operator by which the Laplacian operator can be discretized over a polygonal mesh:

$$\Delta(\mathbf{p}_i) = \sum_{j \in i^*} w_{ij} (\mathbf{p}_j - \mathbf{p}_i)$$

where the coefficients w_{ij} weight the contribution from the different neighbors to the Laplacian of each vertex, and are constrained to sum up to one for a given vertex, that is $\sum_{j \in i^*} w_{ij} = 1$. The weighting coefficients can be derived from a set of positive values $\phi_{ij} = \phi_{ji}$ defined over the edges of the mesh as

$$w_{ij} = \frac{\phi_{ij}}{\sum_{h \in j^*} \phi_{ih}}$$

It is easy to see that by setting $\phi_{ij} = 1$ for every edge we obtain the same discretization of the umbrella operator; another sensible choice is $\phi_{ij} = \left\| \mathbf{p}_i - \mathbf{p}_j \right\|^{-1}$, which weights the contribution of a given neighbor with the inverse of its distance from the vertex. Note in passing that in general **K** is not symmetric since we could have

$$\sum_{h \in i^*} \phi_{ih} = \sum_{h \in j^*} \phi_{jh}$$

even for simple choices of ϕ (like for the umbrella operator if the vertices have different valence).

The question we want to address now is the following: given that our goal is not to approximate the Laplacian of the continuous surface, but rather to minimize the generalization error, is there a better choice for the discrete Laplacian than the umbrella-operator?

In order to answer to this question it helps to note that the condition on the invalid vertices $\mathbf{KP}^T$=0 defines the global minimum of the energy

$$E = \mathbf{P} \cdot \mathbf{K} \cdot \mathbf{P}^T = \sum w_{ij} \left\| \mathbf{p}_i - \mathbf{p}_j \right\|^2 ,$$

which can be seen as a discrete formulation of the membrane energy in terms of edges elongations. From the above equation we can observe that the values w_{ij} determine the stiffness of each edge: edges with a low value will deform more easily than others with a higher one, since their contribution to the overall energy will also be lower.

We argue that this property can be used, by learning the values w_{ij} from the set of training examples that we used to estimate the statistical model. In fact, the most sensible choice is to set

$$\phi_{ij} = \frac{1}{\sigma_{ij}^2}$$

where σ_{ij} is the standard deviation of the edge length in our example set. We will investigate in the next section if this choice yields any improvement to the reconstruction performance.

5. Results

In order to prove the validity of our method we have to verify the following points:

- Let alone the issue of the continuity at the boundaries, does our method perform better than the statistical reconstruction alone?
- Does it perform better than a variational scheme using as guidance the mean shape $\overline{\mathbf{v}}$, rather than the statistical reconstruction?
- Does the particular choice of the discrete Laplacian operator, described in the previous section, improve the results?

We answer to these questions by comparing the generalization errors of the different reconstruction methods, not only in terms of vertex positions, but also of normals directions. The tests are carried on by 5-fold cross validation on a set of 200 examples: for each method the examples set is split in 5 groups, and each group is used in turn as test set, while the other groups are used as training set. With the training set we estimate the model parameters, apart from the noise variance, which is set to 10^{10}; then, for the examples in the test set, the vertices of the nose are reconstructed. For each method, the error on the vertices positions is computed as

$$e_p = \left(\frac{1}{200 \cdot N} \sum \left\| \mathbf{p}_i - \hat{\mathbf{p}}_i \right\|^2 \right)^{1/2},$$

while the error on the normals directions is computed as

$$e_n = \left(\frac{1}{200 \cdot N} \sum \arccos^2 (\mathbf{n}_i \cdot \hat{\mathbf{n}}_i) \right)^{1/2}$$

The results are shown in the histograms of fig. 6-7. With respect to the first question, we can observe that our method improves the performance of the statistical reconstruction both in terms of positions and normals errors. It is also clear that using the mean shape as guidance for the variational method yields the worse results. However, with respect to the last question, the advantage of the operator described in section 4.2 with respect to the umbrella operator is less evident: especially for the position error, the improvement is marginal. Finally, we have to qualify the results by saying that we did these tests only on the reconstruction of the nose, and it may be that on other regions the results would be different.

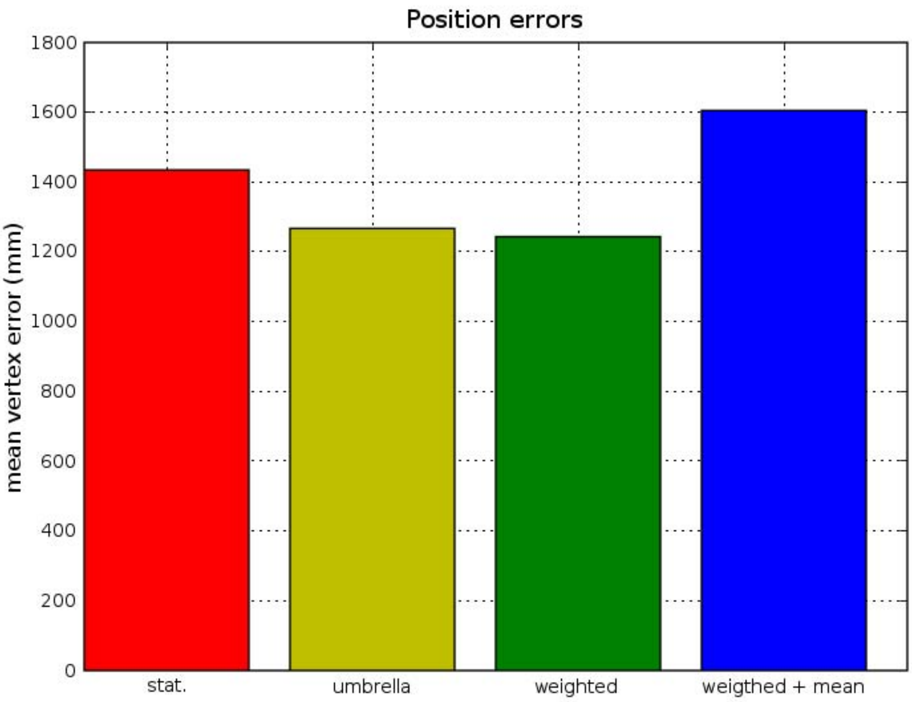


Figure 6. Comparison of the error in generalizing the vertices positions, for the different reconstruction methods: statistical ($1.43 \cdot 10^3$), Poisson using the umbrella operator ($1.27 \cdot 10^3$), Poisson using the operator described in section 4.2 ($1.24 \cdot 10^3$), and the Poisson using the latter operator but the mean as guidance ($1.61 \cdot 10^3$).

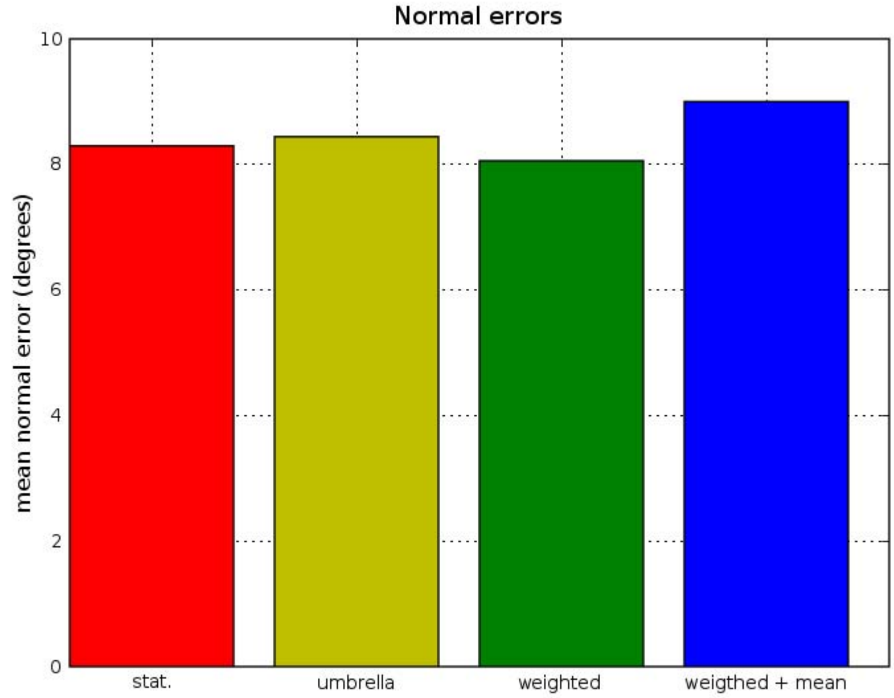


Figure 7. Comparison of the error in generalizing the normals directions, for the different reconstruction methods: statistical (8.30 degrees), Poisson using the umbrella operator (8.44), Poisson using the operator described in section 4.2 (8.06), and the Poisson using the latter operator but the mean as guidance (9.00).

A sample of the results obtained during the tests are in fig. 8, where we selected the four examples with maximum and minimum error, both in terms of vertex positions and normal directions. Visually, the reconstructions with minimum error are very close to the original shape of the nose. The reconstructions with maximum error are in fact different from the originals; note however how they still fits very well with the overall shape of the face.

6. Conclusions

We presented a method to reconstruct invalid or missing areas of a surface by first computing an approximation via a statistical approach, and then refining it by solving a variational problem. The results satisfy the two requirements we set at the beginning: they are smooth at the boundaries and minimize the generalization error. The performance in terms of two different measures of the error between the reconstructed surface and the ground truth, prove that the method does not only remove the discontinuities from the statistical reconstruction, but also that it improves the results of the latter.

The method could be improved from two sides: first, by substituting the membrane energy with a functional that would yield a smooth solution at the boundaries (rather than only a continuous one), and second, by reducing the generalization error of the statistical reconstruction. Regarding the first point, we are currently investigating the possibility of using a guidance field in conjunction with a thin-plate energy; this should directly provide a smooth solution. With respect to the second point, we assume that choosing a subset of the known vertices to drive the statistical reconstruction could reduce its generalization error; in particular, it seems reasonable to select only the vertices that are highly correlated with the missing ones.

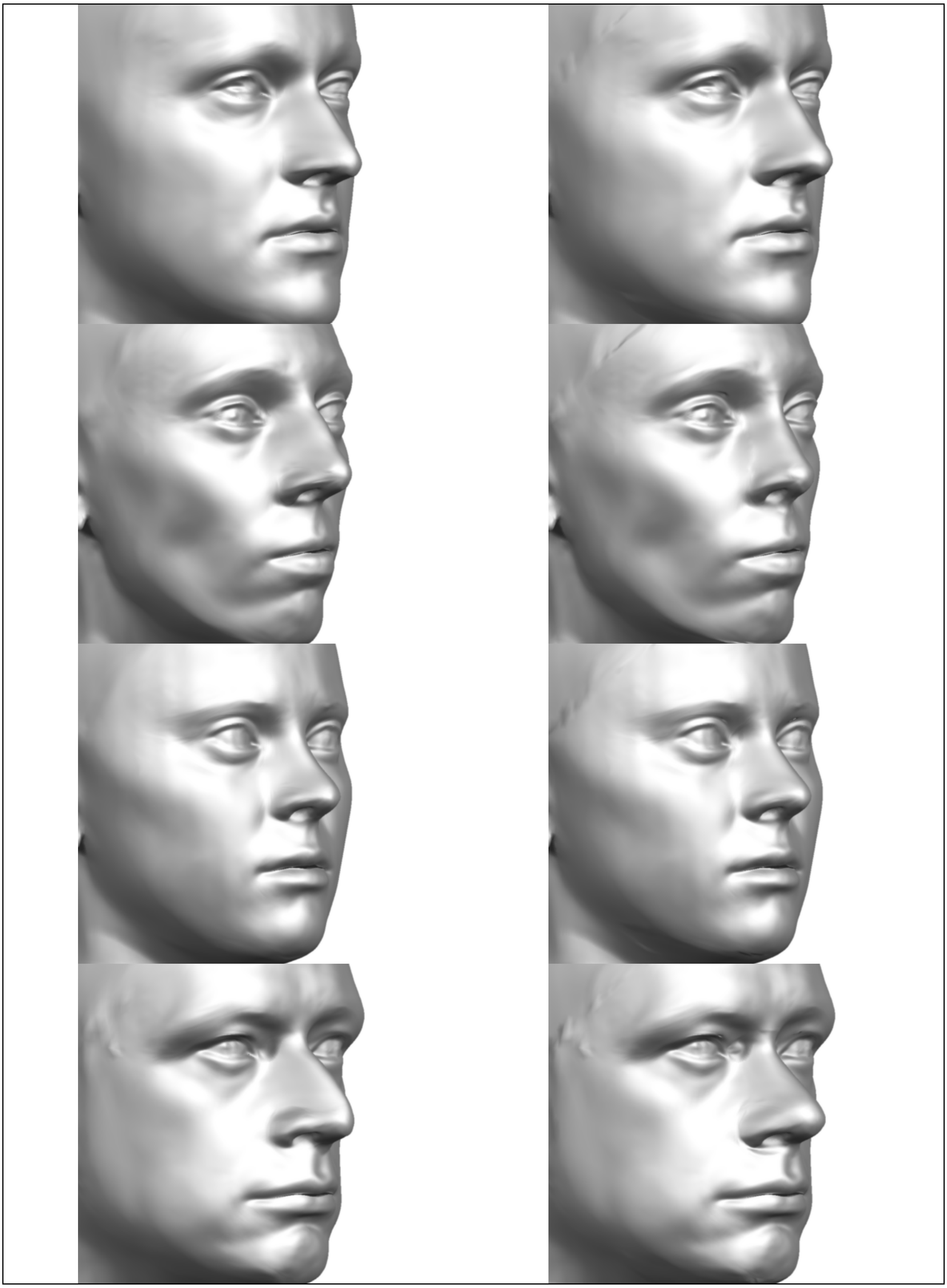


Figure 8. Some results of the experiments: in the left column are the originals, in the right column the reconstructions. The first two rows shows the examples with minimum and maximum average error on the vertices positions ($6.49 \cdot 10^2$ and $2.84 \cdot 10^3$ micrometers respectively), the last two rows are the examples with minimum and maximum average error on the normals directions (5.21 and 11.3 degrees respectively).

Acknowledgements

This work was funded by the Swiss National Science Foundation in the scope of the NCCR Co-Me project 5005-66380

References

[1] P. Liepa, “Filling holes in meshes”, in *Symposium on Geometry Processing*, 2003, pp. 200–205.

[2] U. Clarenz, U. Diewald, G. Dziuk, M. Rumpf, and R. Rusu, “A finite element method for surface restoration with smooth boundary conditions”, *Computer Aided Geometric Design*, vol. 21, pp.427–445, May 2004.

[3] V. Savchenko and N. Kojekine, “An approach to blend surfaces”, in *Proceedings of CGI‘2002 Conference, Advances in Modeling, Animation and Rendering*, 2002, pp. 139–150.

[4] J. Davis, S.R. Marschner, M. Garr, and M. Levoy, “Filling holes in complex surfaces using volumetric diffusion”, in *Proceedings of the 1st International Symposium on 3D Data Processing Visualization and Transmission (3DPVT-02)*, 2002, pp. 428–438.

[5] J. Verdera, V. Caselles, M. Bertalmio, and G. Sapiro, “Inpainting surface holes”, in *Proceedings of the IEEE International Conference on Image Processing (ICIP 2003)*, 2003.

[6] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, “Image inpainting”, in *Proceedings of SIGGRAPH’00*, 2000.

[7] H.-K. Zhao, S. Osher, and R. Fedkiw, “Fast surface reconstruction using the level set method”, in *Proceedings of the IEEE Workshop on Variational and Level Set Methods (VLSM’01)*, 2001.

[8] A. Sharf, M. Alexa, and D. Cohen-Or, “Context-based surface completion”, in *Proceedings of SIGGRAPH’04*, 2004.

[9] V. Blanz, A. Mehl, T. Vetter, and H. P. Seidel, “A statistical method for robust 3d surface reconstruction from sparse data”, in *Int. Symp. on 3D Data Processing, Visualization and Transmission*, Thessaloniki, Greece, 2004.

[10] P. Perez, M. Gangnet, and A. Blake, “Poisson image editing”, in *Proceedings of SIGGRAPH’03*, 2003.

[11] V. Blanz and T. Vetter, “A morphable model for the synthesis of 3d faces”, in *Proceedings of SIGGRAPH’99*, 1999, pp. 187–194.

[12] Y.Yu et al., “Mesh editing with Poisson-based gradient field manipulation”, in *Proceedings of SIGGRAPH’04*, 2004.

[13] C. Basso, T. Vetter, and V. Blanz, “Regularized 3D Morphable Models”, in *Proceedings of the 1st IEEE International Workshop on Higher-Level Knowledge in 3D Modeling and Motion Analysis (HLK2003)*, 2003.

[14] L. Kobbelt, S. Campagna, J. Vorsatz, and H.-P. Seidel, “Interactive multi-resolution modeling on arbitrary meshes”, in *Proceedings of SIGGRAPH ’98*, 1998.

[15] G. Taubin, “A signal processing approach to fair surface design”, in *Proceedings of SIGGRAPH’95*, 1995.

[16] T. A. Davis, UMFPACK version 4.3, 2004, http://www.cise.ufl.edu/research/sparse/umfpack/.

[17] S. Roweis, “EM algorithms for PCA and SPCA”, in *Advances in Neural Information Processing Systems (NIPS)*, 1997.

[18] M. E. Tipping and C. M. Bishop, “Probabilistic principal component analysis”, *Journal of the Royal Statistical Society, Series B*, vol. 21, pp. 611–622, 1999.

[19] M. Stone, “Cross-validatory choice and assessment of statistical predictions”, *Journal of the Royal Statistical Society*, vol. 36, pp. 111-147, 1974.

A. Optimal PPCA reconstruction

In section 2.1 we have shown how the optimal model coefficients $\boldsymbol{\alpha}$ given an incomplete vector $\mathbf{v}^*$ are obtained from the minimization of the following energy:

$$E = \|\boldsymbol{\alpha}\|^2 / 2 + \|\mathbf{L} \cdot (\mathbf{v}^* - \overline{\mathbf{v}}) - \mathbf{Q} \cdot \boldsymbol{\alpha}\|^2 / 2\sigma^2 .$$

Deriving E with respect to $\boldsymbol{\alpha}$, we can find the global minimum as the solution of the following linear system:

$$(\mathbf{Q}^T \mathbf{Q} + \sigma^2 \mathbf{I}) \cdot \boldsymbol{\alpha} = \mathbf{Q}^T \cdot \mathbf{L}(\mathbf{v}^* - \overline{\mathbf{v}}) ,$$

Let $\mathbf{Q}=\mathbf{XWV}^T$ be the singular value decomposition of the matrix $\mathbf{Q}$, where W is a diagonal matrix. The equation above can be then rewritten as

$$(\mathbf{VWX}^T \mathbf{XWV}^T + \sigma^2 \mathbf{I}) \cdot \boldsymbol{\alpha} = \mathbf{VWX}^T \cdot \mathbf{L}(\mathbf{v}^* - \overline{\mathbf{v}})$$

Remembering that $\mathbf{X}^T\mathbf{X}=\mathbf{I}$ and $\mathbf{V}^T\mathbf{V}=\mathbf{VV}^T=\mathbf{I}$, we can manipulate the equation further:

$$(\mathbf{VW}^2\mathbf{V}^T + \sigma^2 \mathbf{VV}^T) \cdot \boldsymbol{\alpha} = \mathbf{VWX}^T \cdot \mathbf{L}(\mathbf{v}^* - \overline{\mathbf{v}}) ,$$

and finally:

$$\mathbf{V}(\mathbf{W}^2 + \sigma^2 \mathbf{I})\mathbf{V}^T \cdot \boldsymbol{\alpha} = \mathbf{VWX}^T \cdot \mathbf{L}(\mathbf{v}^* - \overline{\mathbf{v}}) .$$

The matrix on the left hand side can be inverted and brought to the right hand side, yielding the equation for the optimal coefficients $\boldsymbol{\alpha}$:

$$\boldsymbol{\alpha} = \mathbf{V}(\mathbf{W}^2 + \sigma^2 \mathbf{I})^{-1} \mathbf{WX}^T \cdot \mathbf{L}(\mathbf{v}^* - \overline{\mathbf{v}})$$

Curricula Vitae.

Curzio Basso received the diploma-degree in Physics from the University of Genova, Italy, in 1999 with a thesis on Support Vector Machines applied to regression. Since 2000 he is a PhD student in Computer Science, first at the University of Freiburg, Germany, and currently at the University of Basel, Switzerland. His research activity focuses on the application of statistical learning to the modeling of human faces and expressions.

Thomas Vetter studied mathematics and physics and received the PhD degree in biophysics from the University of Ulm, Germany. As a postdoctoral researcher at the Center for Biological and Computational Learning at MIT, he started his research on computer vision. In 1993, he moved to the Max-Planck-Institut in Tübingen and, in 1999, he became professor of computer graphics at the University of Freiburg. Since 2002, he has been a professor of applied computer science at the University of Basel in Switzerland. His current research is on image understanding, graphics, and automated model building. He is a member of the IEEE and the IEEE Computer Society.