Vision, Modeling, and Visualization (2013)
Michael Bronstein, Jean Favre, and Kai Hormann (Eds.)

A New Framework for Fitting Shape Models to Range Scans: Local Statistical Shape Priors Without Correspondences

Carsten Last, Simon Winkelbach, and Friedrich M. Wahl

Institut für Robotik und Prozessinformatik, TU Braunschweig,
Mühlenpfordtstr. 23, 38106 Braunschweig, Germany,
{c.last, s.winkelbach, f.wahl}@tu-bs.de

Abstract

Statistical shape models provide an important means in many applications in computer vision and computer graphics. However, the major problems are that the majority of these shape models require dense point-correspondences along all training shapes and that a large number of training shapes is needed in order to capture the full amount of intra-class shape variation. In this contribution, we focus on a statistical shape model that can be constructed from a set of training shapes without defining any point-correspondences. Additionally, we show how a local statistical shape model can make better use of the available shape information, greatly reducing the number of required training shapes. Finally, we present a new framework to fit this local statistical shape model without correspondences to range scans that represent incomplete parts of the trained shape class. The fitted model is then used to reproduce a natural-looking approximation of the complete shape.

Categories and Subject Descriptors (according to ACM CCS): I.5.1 [Pattern Recognition]: Models—Statistical I.4.8 [Image Processing and Computer Vision]: Scene Analysis—Surface Fitting I.4.10 [Image Processing and Computer Vision]: Image Representation—Volumetric

1. Introduction

Statistical Shape Models (SSMs) can be used for many applications in computer vision and computer graphics, e.g. realtime facial animation [LYYB13] and face recognition [AKV08]. In these tasks they have proven their superior performance compared to other approaches (see e.g. [BCF06]). However, there exist two major drawbacks with statistical shape models that we try to address in this contribution:

1. Dense point-correspondences need to be defined along the training shapes of the model.
2. A large number of training shapes is needed in order to capture the full amount of intra-class shape variation.

Dense point-correspondences may be determined manually or automatically. However, both approaches have their drawbacks. While the manual definition of point-correspondences is tedious and time-consuming, the automatic generation of point-correspondences is a difficult problem that is not easy to solve [HM09]. This contradicts the fact that wrong point-correspondences can cause strong artifacts in the shapes generated by the statistical shape model [PKA*09]. So, we think that it would be better to use a statistical shape model which does not require any point-correspondences at all.

The second major drawback is that for natural shape classes, like e.g. faces, a large number of training shapes is needed in order to capture the full amount of intra-class shape variation. For example, Paysan et al. used 200 training shapes to build a statistical shape model of the human face [PKA*09]. Even this large number of training samples is not enough to capture the full amount of face variation, as four predefined segments have been manually introduced in the shape model in order to enlarge the space of possible shape configurations. However, a large number of training shapes means a lot of work, because all these shapes have to be acquired (which always includes a certain amount of manual editing) or they have to be segmented from volume data (which is often also done manually, e.g. [PTW*09]).

This means, the larger the number of required training shapes, the higher the threshold to start research in the area of statistical shape models. Paysan et al. also recognized this

DOI: 10.2312/PE.VMV.VMV13.153-160

problem and made their model publicly available in order to emphasize the use of statistical shape models in various research areas. We also use their data for our experiments in this contribution and would like to thank them for making it available. However, their model has been built only for the class of face shapes. For other shape-classes, the problem of obtaining the required training data still exists.

To address the two above mentioned problems, we propose a framework that allows the fully-automatic fitting of a local statistical shape model without correspondences to range scans that represent incomplete parts of the trained shape class. The model that we use has no need for point-correspondences and can capture a great amount of shape variation with only a few training shapes at hand. The fitting result of the local model can then be used to reproduce a natural-looking approximation of the complete shape.

Our contribution is structured as follows: In section 2 we will review the basic concepts of statistical shape models based on correspondences, and we will introduce an alternative approach that addresses the above mentioned problems: a local statistical shape model without correspondences. Then, in section 3 we will present our new framework for fitting the local statistical shape model to range scans. In section 4 we will compare a standard statistical shape model with the two alternative approaches and show that the two alternatives perform equal to or even better than the standard model. Finally, in section 5 we will demonstrate the use of our framework by fitting the local statistical shape model to incomplete range scans of faces.

2. Related Work

In this section we review prior work in the field of statistical shape models on which we build up in our contribution. At first, section 2.1 deals with the basic concepts of statistical shape models based on point-correspondences. Then, in section 2.2 we introduce a volumetric statistical shape model which has no need for point-correspondences, and in section 2.3 we recapitulate our recently proposed local statistical shape model that is tailored to limited training data.

2.1. Shape Models Based On Point-Correspondences

In the majority of statistical shape models, the training shapes are represented as ordered lists of vertices that are in correspondence with each other. Given a set of m training shapes $\{\vec{C}_1,\dots,\vec{C}_m\}$, each shape is represented by a vector

$$\vec{C}_i = [x_1,y_1,z_1,x_2,y_2,z_2,\dots,x_n,y_n,z_n] \tag{1}$$

of n vertices, where each triple (x_k,y_k,z_k), with $1 \le k \le n$, represents one vertex of the training shape. Under the assumption of a normal intra-class shape distribution, a mean shape $\vec{C}_{\text{mean}}$ is extracted from the training data by averaging over all shapes. Afterwards, a shape matrix is obtained by stacking all mean-subtracted training shapes on top of each other. Then, a principal component analysis (PCA) is performed on the shape matrix in order to estimate the principal axes $\{\vec{B}_1,\dots,\vec{B}_{m-1}\}$ and the corresponding standard deviations $\{s_1,\dots,s_{m-1}\}$ of the multivariate normal distribution that best describes the set of training shapes. The statistical shape model is finally given by

$$\vec{C}_{\vec{w}} = \vec{C}_{\text{mean}} + \sum_{i=1}^{m-1} w_i \vec{B}_i, \tag{2}$$

where $\vec{w} = [w_1,\dots,w_{m-1}]$ is a vector of shape weights that is used to control the shapes $\vec{C}_{\vec{w}}$ which can be synthesized by the model. In order to comply to the trained shape distribution, the set of possible shape weights is limited to the interior of the hyperellipse which is bounded by three standard-deviations of the multivariate normal distribution:

$$\sqrt{\sum_{i=1}^{m-1} \frac{{w_i}^2}{{s_i}^2}} \le 3. \tag{3}$$

2.2. A Shape Model Without Correspondences

As stated in section 1, we think that the use of correspondences in statistical shape models has many drawbacks. An alternative approach has been proposed by Leventon et al. in the context of level set segmentation of medical image data [LGF00]. They chose an implicit representation of the training shapes $\vec{C}_i$ as the zero level set of a higher-dimensional signed distance function $\Phi_i : \Omega \subset \mathbb{R}^3 \to \mathbb{R}$:

$$\vec{C}_i = \{\vec{x} \,|\, \Phi_i(\vec{x}) = 0\}. \tag{4}$$

Using this volumetric representation, they showed that it is possible to construct a statistical shape model by using linear combinations of the implicit, volumetric training shapes without defining any correspondences.

Similar to the explicit shape representation in eq. (1), one can arrange the implicit shape representation row-wise in a single vector $\vec{\Phi}_i$ by defining a raster scanning pattern on the volume Ω. Then, a mean shape Φ_{mean} is extracted again by averaging over all training shapes, and a PCA is performed on the shape matrix, yielding the principal axes $\{\vec{\tilde{B}}_1,\dots,\vec{\tilde{B}}_{m-1}\}$ and corresponding standard deviations $\{\tilde{s}_1,\dots,\tilde{s}_{m-1}\}$ of the volumetric statistical shape model. Finally, the volumetric statistical shape model is given as:

$$\Phi_{\vec{w}}(\vec{x}) = \Phi_{\text{mean}}(\vec{x}) + \sum_{i=1}^{m-1} w_i \vec{\tilde{B}}_i(\vec{x}), \tag{5}$$

for all $x \in \Omega$. Explicit shapes $\vec{C}_{\vec{w}}$ can be obtained as the zero level set of the implicit shapes $\Phi_{\vec{w}}$ which are synthesized by the volumetric model.

2.3. A Shape Model for Limited Training Data

Another major drawback with statistical shape models is that they require a great amount of training shapes to capture

the full amount of variation which is inherent in a class of shapes. Various approaches have been presented to deal with this problem. The most promising ones consist of partitioning the model (e.g. [dBGVN03] and [ZAT05]) or using hierarchical models (e.g. [DTS03] and [NHBT07]). Nevertheless, it is a nontrivial task to identify the correct segments. So, we recently proposed another approach that also allows for variable adaptations in different regions of the model but has no need for any predefined segments [LWW*11].

The central idea is to replace the vector of shape weights $\vec{w} \in \mathbb{R}^{m-1}$ by a *sufficient smooth* field of weight vectors $\vec{w} : \Omega \to \mathbb{R}^{m-1}$ so that equation (5) changes to

$$\Phi(\vec{w}(\vec{x})) = \Phi_{\text{mean}}(\vec{x}) + \sum_{i=1}^{m-1} w_i(\vec{x})\vec{B}_i(\vec{x}), \qquad (6)$$

for all $x \in \Omega$. This modification allows a different adaptation of the statistical shape model in every location of the underlying data domain Ω. The smoothness constraint of the weight field is important because it controls how much the local statistical model of eq. (6) is restricted by the global statistical model of eq. (5): If $\vec{w}(\vec{x})$ is a constant function, then eq. (6) and eq. (5) are identical. If $\vec{w}(\vec{x})$ is allowed to develop arbitrary steep edges, then nearly any shape can be represented through eq. (6). For a *sufficient smooth* function $\vec{w}(\vec{x})$, eq. (6) gives us the ability to continuously crossfade between different instances of the global statistical model in local regions of the data domain Ω. This is basically the same idea as partitioning the model in various segments, where the degree of smoothness roughly corresponds to the size of the segments. However, the great advantage is that the segments do not need to be predefined. Furthermore, the degree of smoothness can be varied throughout the fitting process.

3. Methods

In the following we present our approach for the rigid alignment of shapes without correspondences, and we present our new framework that allows the fully-automatic fitting of the local volumetric statistical shape model to possibly incomplete range scans.

3.1. Rigid Shape Alignment

The statistical shape models of section 2 capture only non-rigid shape deformations, so all shapes have to be rigidly aligned with regard to pose and scale. When the shapes are in correspondence with each other, as for the statistical shape model of section 2.1, one can keep one shape fixed and calculate pairwise similarity transformations [Dha03] of all other shapes to the reference shape. Each similarity transformation minimizes the mean-square distance between all points on a training shape to the reference shape with regard to rotation, translation, and isotropic scale. However, when we do *not* assume that the shapes are in correspondence with each other, we cannot simply compute a similarity transformation in order to rigidly align the shapes. Therefore, we propose to use a two-step approach for the alignment of two shapes without correspondences: A coarse registration followed by a fine registration.

For the coarse registration we propose to use the RANSAM approach [WMW06]. RANSAM (RANdom SAmple Matching) is a fast and robust approach to coarsely register a pair of 3D scans. It builds up on the well-known probabilistic RANSAC method. However, the RANSAM approach only registers two shapes with regard to rotation and translation. So, in order to obtain an estimate for the initial scale factor, we use the following algorithm:

```
for f=minScale:0.05:maxScale do
  scale shape by factor f
  find optimal RANSAM match
  evaluate matching quality
end
```

This means, we simply iterate over a finite set of scale factors and compute a RANSAM match for each scale factor. Like in the RANSAC method, we calculate the matching quality for each obtained match by counting the inliers in an ε-neighborhood around the matched shape after registration. The final matching result is simply the one with the highest matching quality. This simple approached proved to perform very well in estimating the initial rotation, translation, and scale for the following fine registration step.

For the fine registration we propose to use a modified version of the well-known Iterative Closest Points (ICP) algorithm [BM92]. The ICP algorithm iteratively selects a random set of closest points on a pair of shapes and minimizes the squared distance between these two point sets with regard to rotation and translation. After a few iterations, this usually leads to a fine registration of the two shapes. Our modification consists of including the scale factor in the minimization process by minimizing the distance between the two point sets with the help of the similarity transformation introduced above. Again, this simple modification performed very well in registering the shapes with regard to rotation, translation, and scale.

3.2. Fitting the Global Volumetric Model to 3D Scans

Before we describe the fitting process of the global volumetric statistical shape model of section 2.2, we briefly review the training process: We start with a set of m training shapes $\{\vec{C}_1, \ldots, \vec{C}_m\}$ given in an explicit, mesh-based representation. After rigid alignment, the training shapes are converted to a volumetric representation using the approach presented in [CL96]. This approach approximates a point cloud by a signed distance volume in a small hull around the surface. We chose a diameter of 20 mm for the hull. This signed distance hull is propagated to the rest of the volume using the approach described in [FH04]. From the thus obtained

volumetric representations of our input shapes, we calculate the volumetric statistical shape model as described in section 2.2. Please note that the so obtained PCA subspace is also used by the local volumetric SSM of section 2.3.

Now that we have trained the volumetric statistical shape model, we can describe how to fit the model to a new range scan $\vec{C}$ in order to approximate the scan by the model. At first, we align the range scan to the explicit representation of the mean shape of our volumetric SSM using the approach described in section 3.1. Since the new scan may contain holes, the interior and exterior regions of the shape are not clearly defined. So, no complete signed distance information can be obtained from the aligned range scan, and we only approximate the signed distance hull with the approach in [CL96], but no further propagation of the distance values to the rest of the volume is performed. Having the range scan rigidly registered to our 3D model, we need to find the weights that belong to the best shape approximation which can be synthesized through our volumetric statistical shape model. Let

$$A = \begin{bmatrix} \vec{B}_1 \\ \vdots \\ \vec{B}_{m-1} \end{bmatrix} \tag{7}$$

be the subspace of possible shapes spanned by the volumetric statistical shape model of equation (5). With the help of equation (7), equation (5) can be rewritten as:

$$\Phi_{\vec{w}} = \Phi_{\text{mean}} + \vec{w} \cdot A. \tag{8}$$

From the PCA theory follows that, given a shape Φ, the weights w_{opt} obtained by

$$\vec{w}_{\text{opt}} = (\Phi - \Phi_{\text{mean}}) \cdot A^T \tag{9}$$

are the best approximation of the shape Φ in the subspace A that minimizes the sum of squared errors between the original and reconstructed shape [Jol02]. This means that

$$||\Phi - \Phi_{\vec{w}}||^2 \tag{10}$$

is minimal for $\vec{w} = \vec{w}_{\text{opt}}$, yielding a closed-form solution for the best approximation (in least squares sense) which can be synthesized from the volumetric statistical shape model of equation (5). Please note that the same argumentation applies for the correspondence-based statistical shape model of equation (2). The problem is that we do not have a complete signed distance representation of our acquired 3D scan. So, we cannot calculate the closed form solution of equation (9). Instead we formulate the model fitting problem as an iterative fixed point problem:

Let Φ_{hull} be the signed distance representation of our range scan $\vec{C}$ that contains only distance values in a small hull around the scanned surface (further denoted by hull$(\vec{C})$). We start with initial weights $\vec{w}_0 = \vec{0}$ and calculate a first model approximation as

$$\Phi_{\vec{w}_k} = \Phi_{\text{mean}} + \vec{w}_k \cdot A. \tag{11}$$

In this approximation, the synthesized signed distance values are replaced by the known signed distance values of the hull around the scanned surface:

$$\Phi^{\star}_{\vec{w}_k}(\vec{x}) = \begin{cases} \Phi_{\text{hull}}(x) & \text{, if } \vec{x} \in \text{hull}(\vec{C}) \\ \Phi_{\vec{w}_k}(\vec{x}) & \text{, else} \end{cases}. \tag{12}$$

Then, a new weight vector $\vec{w}_{k+1}$ is obtained by inserting $\Phi^{\star}_{\vec{w}_k}(\vec{x})$ in equation (9):

$$\vec{w}_{k+1} = (\Phi^{\star}_{\vec{w}_k} - \Phi_{\text{mean}}) \cdot A^T \tag{13}$$

By iterating equations (11), (12), and (13) until convergence one obtains an accurate approximation of the range scan by the volumetric statistical shape model. However, we additionally perform one step of our modified ICP after each weight change in order to adjust the rigid transformation parameters of the updated model approximation.

3.3. Fitting the Local Volumetric Model to 3D Scans

Our approach for fitting the local volumetric statistical shape model of equation (6) to range scans is closely related to the approach presented by us in [LWW*11]. Like the fitting of the global volumetric SSM, it is also an iterative process, whereby the fitting result $\vec{w}_{\text{opt}}$ from section 3.2 is used as an initial solution for the local fitting process. So, we initialize the local weight field of equation (6) by setting $\vec{w}_0(\vec{x}) = \vec{w}_{\text{opt}}$ for all $\vec{x} \in \Omega$. Given the signed distance hull representation Φ_{hull} of our scanned shape $\vec{C}$, we perform a weight update for each voxel $\vec{x}$ that resides inside the hull in order to minimize the mean-square error between the local volumetric model approximation and the range scan $\vec{C}$. Thus, for each $\vec{x} \in \text{hull}(\vec{C})$ the weight update is calculated as follows:

$$\vec{w}_{k+1}(\vec{x}) = \vec{w}_k(\vec{x}) + \nabla r(\vec{w}_k(\vec{x})), \tag{14}$$

where the error function is defined as

$$r(\vec{w}_k(\vec{x})) = (\Phi(\vec{w}_k(\vec{x})) - \Phi_{\text{hull}}(\vec{x}))^2, \tag{15}$$

and the gradient of the error function is given as

$$\nabla r(\vec{w}_k(\vec{x})) = \vec{w}_k(\vec{x}) \cdot \vec{\vec{b}}\vec{\vec{b}}^T + (\Phi_{\text{mean}}(\vec{x})) - \Phi_{\text{hull}}(\vec{x})) \cdot \vec{\vec{b}}, \tag{16}$$

with $\vec{\vec{b}} = [\vec{B}_1(\vec{x}), \cdots, \vec{B}_{m-1}(\vec{x})]$.

Using this weight update scheme, the mean-square error between the local volumetric model representation and the signed distance hull of our range scan is minimized without posing any restrictions on the weight function $\vec{w}(\vec{x})$.

Now, in order to satisfy the smoothness constraint discussed in section 2.3, we convolve the resulting weight field $\vec{w}_{k+1}(\vec{x})$ after each weight update step with a cubic smoothing kernel. By doing so, the updated weight values of the signed distance hull are slowly propagated over the whole volume. We start with a big smoothing kernel of about 65 mm x 65 mm x 65 mm in order to emphasize a fast spreading. The size of the smoothing kernel is further reduced during the fitting process in order to being able to adjust to fine

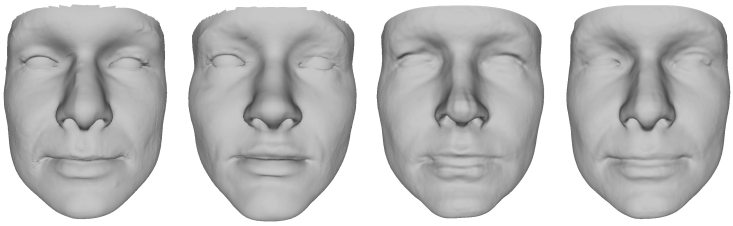

Figure 1: *Examples for the best possible approximation of a given target shape (left), reconstructed using the standard SSM (2nd column), the volumetric SSM (3rd column), and the local volumetric SSM (right). The SSMs have been trained using 30 data sets. It can be seen that the result synthesized by the local volumetric SSM looks most similar to the target shape.*

structures of the range scan. The dimensions of the final kernel are dependent on the amount of noise in the range scan and on the size of the holes in the scan. In order to stick to the trained shape distribution, we additionally project each weight vector $\vec{w}_k(\vec{x})$ that resides outside the hyperellipse of eq. (3) back onto this hyperellipse. Please note that no update of the rigid transformation parameters is performed during the fitting process of the local model as we consider them accurate enough after the global fitting process.

4. Evaluating the Best Possible Shape Approximation

In section 1 it has been argued that a large amount of training data is needed in order to capture the full amount of variation in a given shape class. In this section we want to substantiate these arguments by evaluating the best possible shape approximation which can be synthesized by a statistical shape model depending on the number of training shapes that have been used to build the model. Also, we evaluate how the approximation quality differs between the traditional statistical shape model of section 2.1 and the volumetric statistical shape models of sections 2.2 and 2.3.

The training data for our experiment has been provided by Paysan et al. [PKA*09]. As stated in section 1, they made their model publicly available together with a method for synthesizing random sample shapes. We synthesized up to 90 shapes with the help of their model that were in turn used to train the statistical shape models of section 2. They also provided ten exemplary range scans that are in correspondence to their model but have not been used in the model generation process. So, dense point-correspondences were available that we could use to compute the optimal reconstruction of the traditional statistical shape model, defined in section 2.1.

For the global statistical shape models of sections 2.1 and 2.2, we have presented a closed-form solution for the optimal weights in section 3.2. However, no easy closed form solution exists for the local statistical shape model of section 2.3. So, we used our local fitting framework (section 3.3) also for the determination of the best possible approximation. The dimensions of the final cubic smoothing kernel were chosen to 2.5 mm x 2.5 mm x 2.5 mm.

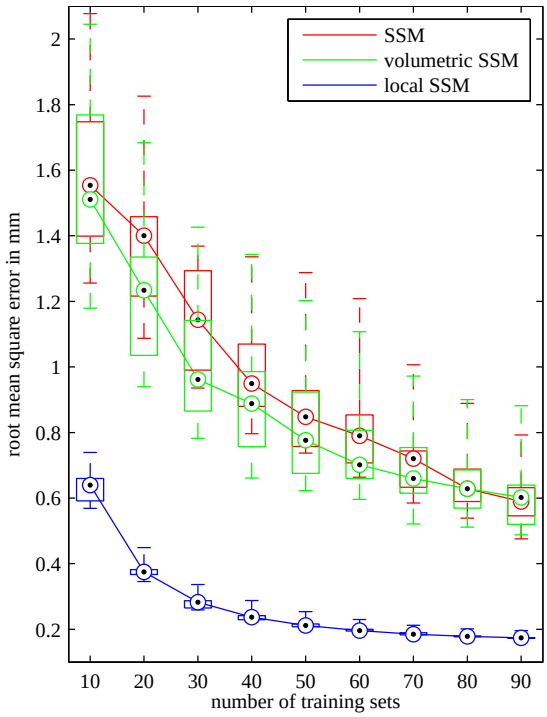

Figure 2: *Boxplots of the root-mean-square deviations from shapes synthesized by various statistical shape models (SSMs) to ten given target shapes, plotted against the number of data sets which were used to train the SSMs. It can be seen that the volumetric SSM (without point correspondences) is as good as a standard SSM (with point correspondences) and that the local volumetric SSM clearly outperforms the other two approaches for limited training data.*

As the statistical shape models capture only the nonrigid variations, prior to the model reconstruction, each of the target shapes has been rigidly aligned to the mean shape of the

statistical shape model of section 2.1 using a similarity transformation as explained in section 3.1. The same rigid alignment was used for the volumetric statistical shape models so that the evaluation of the reconstruction quality is not influenced by the rigid alignment procedure. For the volumetric shape models, each of the target shapes was transformed to an implicit signed-distance representation by the approach presented in section 3.2. After the model fitting process, the resulting volumetric, implicit shape approximations were transformed back to a mesh-based, explicit representation using the marching cubes algorithm [LC87].

The reconstruction quality has been evaluated by calculating the root-mean-square deviation from the model approximation to the target shape:

$$e = \frac{1}{n}\sqrt{\sum_{i=1}^{n}\left\| d\left(\begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix}, \vec{C}\right)\right\|^2}, \qquad (17)$$

where $d((x_i, y_i, z_i)^T, \vec{C})$ is the closest euclidean distance from a point $(x_i, y_i, z_i)^T$ on the approximated shape to the surface of the target shape $\vec{C}$. Please note that this is not necessarily the distance to the nearest vertex of the target shape, it could as well be the closest distance to the nearest triangle of the target shape. The resulting approximation errors are depicted in figure 2. On the horizontal axis one can see the number of training samples that have been used to train the statistical shape models, and on the vertical axis one can see the approximation error as defined in equation (17). The approximation errors for the ten face scans are given as boxplots, where the median error of all ten faces is depicted as a black dot.

It is clearly visible how the approximation error decreases with a growing number of training datasets for all three statistical shape models. The results also show that the volumetric shape model of equation (5), which does not use any point-correspondences, can synthesize a solution that is equal to (or even slightly better than) the solution synthesized by the statistical shape model with correspondences defined in equation (2). Also, the results show that the local volumetric statistical shape model of equation (6) can synthesize a much better approximation with limited training data than the other two approaches. For example at 30 training sets, the median error of the correspondence-based statistical shape model is 1.14 mm, the median error of the volumetric statistical shape model is 0.96 mm, and the median error of the local volumetric statistical shape model is already as low as 0.28 mm. Additionally, it can be seen that the local volumetric statistical shape model shows a nearly exponential decay of the approximation error. After a strong decrease of the approximation error of about 0.36 mm, when increasing the number of training sets from 10 to 30, the decay is strongly reduced for a growing number of training sets. From 30 to 50 datasets the reduction is 0.07 mm, from 50 to 70 datasets the reduction is 0.03 mm, and from 70 to 90 datasets the reduction is 0.01 mm. Hence, it can be argued that a small number of about e.g. 30 datasets (for the class of face shapes) is enough to train the local volumetric statistical shape model.

Some exemplary results of the model approximations are depicted in figure 1. For these examples the statistical shape models have been trained using 30 training data sets. It can be seen that all three approaches are able to synthesize satisfactory result. However, the result which looks most similar to the target shape has been synthesized by the local volumetric statistical shape model.

5. Application: Reconstructing Incomplete Face Scans

In this final section we want to demonstrate the use of our new shape-fitting framework, which we presented in section 3, by reconstructing the missing regions of 3D face scans with the help of our local volumetric statistical shape model. As for the experiment in section 4, we use training shapes for the volumetric statistical shape model that have been synthesized by the model of Paysan et al. [PKA*09]. We use 30 shapes in the training process. The test scans have been acquired by ourselves with the help of a structured light scanner (http://www.david-laserscanner.com/). We evaluate our approach using 6 range scans of 3 different individuals. The scanned shapes are depicted in the leftmost column of figure 3. The shapes in rows 3 and 5 have been acquired from the front and the shapes in rows 4 and 6 from the side. The shapes in rows 1 and 2 have been stitched together with the help of the RANSAM approach [WMW06] using 3 overlapping scans. It can be seen that only the stitched scans contain almost no holes. All the other scans contain non-negligible holes that have to be filled with the help of statistical shape information in order to reproduce a natural-looking face. Most of them are due to occlusions, but also other factors, like e.g. specular reflection, are plausible explanations.

The fitting of the volumetric statistical shape model has been carried out as described in section 3. We first perform a global fitting of the volumetric model (section 3.2) followed by a local weight adaptation (section 3.3). Please note that our approach works fully automatic without any user interaction. For the almost complete shapes in rows 1 and 2 of figure 3 we use a minimum smoothing kernel size of 2.5 mm x 2.5 mm x 2.5 mm and for the other shapes we use a minimum kernel size of 16.5 mm x 16.5 mm x 16.5 mm. The fitting results are depicted in the second column (global fitting result) and in the third column (local fitting result) of figure 3, respectively. The fourth and fifth column show the modified 3D scans, where the holes have been filled using the local fitting results. In column four, the local fitting result is additionally highlighted in orange.

It can be seen that the local model was able to capture the almost complete scans very accurately. But, more important, also for the other shapes the local model approximation

Figure 3: *Results of the 3D face scan completion experiment explained in section 5. In the leftmost column, the original 3D scan is shown, in the 2nd column one can see the initial fit by the global volumetric SSM, and in the third column one can see the final result obtained with the local volumetric SSM. The 4th and 5th column show a combination of the local fitting result and the original scan, with the local result highlighted in the 4th column.*

looks more similar to the original scan than the global fitting result. The existing regions of the shapes were precisely approximated and the missing regions have been nicely interpolated. However, in those cases where only one side of the face was present in the scan, the model fitting results do not look exactly symmetric. This is due to the fact that there exists no similarity criterion in the incorporated statistical shape model. So, the side with no data mostly resembles the global fitting result, whereas the other side is given as a

precise fit of the local model. Nevertheless, also the reconstructed partial faces look all very natural.

6. Conclusion and Discussion

In this contribution we have shown that it is possible to use a local volumetric statistical shape model, which has no need for point-correspondences, for a natural-looking interpolation of the missing regions in a 3D range scan. Therefore, we have presented a new framework that allows the fully-automatic fitting of the model to range scans without any user interaction. Also, we have shown experimental results that a volumetric shape model without point-correspondences can synthesize a solution that is comparable to the solution synthesized by a standard statistical shape model with point-correspondences and that a local adaptation of the model substantially improves the fitting performance.

One may argue that the ICP algorithm, utilized for the rigid alignment of two shapes, still uses dense point-correspondences, as it repeatedly computes the nearest neighbors of all shape vertices. However, these simple correspondences are much easier to obtain and they are much more robust than the correspondences used in other models, since no nonlinear deformations are required. Additionally, the RANSAM approach is very robust against outliers so that the rigid alignment procedure converges to good results even for strongly disturbed shapes. Nevertheless, we plan to investigate other approaches for the rigid alignment in the future (see e.g. [DPTA09]) and to compare them with our approach. We also plan to extend our local volumetric statistical shape model to include symmetries, which we think will further improve the reconstructed 3D face scans.

References

[AKV08] AMBERG B., KNOTHE R., VETTER T.: Expression invariant 3d face recognition with a morphable model. In *Proceedings of the 8th IEEE International Conference on Automatic Face & Gesture Recognition* (2008), IEEE, pp. 1–6. 1

[BCF06] BOWYER K. W., CHANG K., FLYNN P.: A survey of approaches and challenges in 3d and multi-modal 3d + 2d face recognition. *Computer Vision and Image Understanding 101*, 1 (2006), 1–15. 1

[BM92] BESL P. J., MCKAY N. D.: A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence 14*, 2 (1992), 239–256. 3

[CL96] CURLESS B., LEVOY M.: A volumetric method for building complex models from range images. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques* (New Orleans, LA, USA, 1996), ACM, pp. 303–312. 3, 4

[dBGVN03] DE BRUIJNE M., GINNEKEN B., VIERGEVER M., NIESSEN W.: Adapting active shape models for 3d segmentation of tubular structures in medical images. In *Information Processing in Medical Imaging, 18th International Conference* (2003), Taylor C., Noble J., (Eds.), vol. 2732 of *LNCS*, Springer, Heidelberg, pp. 136–147. 3

[Dha03] DHAWAN A.: *Medical Image Analysis*. Wiley-IEEE Press, 2003. 3

[DPTA09] DURRLEMAN S., PENNEC X., TROUVÉ A., AYACHE N.: Statistical models of sets of curves and surfaces based on currents. *Medical Image Analysis 13*, 5 (2009), 793–808. 8

[DTS03] DAVATZIKOS C., TAO X., SHEN D.: Hierarchical active shape models, using the wavelet transform. *IEEE Transactions on Medical Imaging 22*, 3 (2003), 414–423. 3

[FH04] FELZENSZWALB P., HUTTENLOCHER D.: *Distance transforms of sampled functions*. Tech. Rep. TR2004-1963, Cornell University Computing and Information Science, September 2004. 3

[HM09] HEIMANN T., MEINZER H.-P.: Statistical shape models for 3d medical image segmentation: A review. *Medical Image Analysis 13*, 4 (2009), 543–563. 1

[Jol02] JOLLIFFE I. T.: *Principal component analysis*, second ed. Springer, Heidelberg, 2002. 4

[LC87] LORENSEN W. E., CLINE H. E.: Marching cubes: A high resolution 3d surface construction algorithm. In *Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques* (Anaheim, CA, USA, 1987), ACM, pp. 163–169. 6

[LGF00] LEVENTON M. E., GRIMSON W. E. L., FAUGERAS O.: Statistical shape influence in geodesic active contours. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Hilton Head, SC, USA, 2000), IEEE, pp. 316–323. 2

[LWW*11] LAST C., WINKELBACH S., WAHL F., EICHHORN K., BOOTZ F.: A locally deformable statistical shape model. In *Machine Learning in Medical Imaging, Second International Workshop* (2011), Suzuki K., Wang F., Shen D., Yan P., (Eds.), vol. 7009 of *LNCS*, Springer, Heidelberg, pp. 51–58. 3, 4

[LYYB13] LI H., YU J., YE Y., BREGLER C.: Realtime facial animation with on-the-fly correctives. *ACM Transactions on Graphics (TOG) - SIGGRAPH 2013 Conference Proceedings 32*, 4 (2013), 42. 1

[NHBT07] NAIN D., HAKER S., BOBICK A., TANNENBAUM A.: Multiscale 3-d shape representation and segmentation using spherical wavelets. *IEEE Transactions on Medical Imaging 26*, 4 (2007), 598–618. 3

[PKA*09] PAYSAN P., KNOTHE R., AMBERG B., ROMDHANI S., VETTER T.: A 3d face model for pose and illumination invariant face recognition. In *Proceedings of the 6th IEEE International Conference on Advanced Video and Signal based Surveillance for Security, Safety and Monitoring in Smart Environments* (Genova, Italy, 2009), IEEE, pp. 296–301. 1, 5, 6

[PTW*09] PIRNER S., TINGELHOFF K., WAGNER I., WESTPHAL R., RILK M., WAHL F., BOOTZ F., EICHHORN K.: CT-based manual segmentation and evaluation of paranasal sinuses. *European Archives of Oto-Rhino-Laryngology 266*, 4 (2009), 507–518. 1

[WMW06] WINKELBACH S., MOLKENSTRUCK S., WAHL F.: Low-cost laser range scanner and fast surface registration approach. In *Pattern Recognition, 28th DAGM Symposium* (2006), Franke K., Müller K.-R., Nickolay B., Schäfer R., (Eds.), vol. 4174 of *LNCS*, Springer, Heidelberg, pp. 718–728. 3, 6

[ZAT05] ZHAO Z., AYLWARD S., TEOH E.: A novel 3d partitioned active shape model for segmentation of brain mr images. In *Medical Image Computing and Computer-Assisted Intervention, 8th International Conference* (2005), Duncan J., Gerig G., (Eds.), vol. 3749 of *LNCS*, Springer, Heidelberg, pp. 221–228. 3