

Local Regression based Statistical Model Fitting

Matthias Amberg, Marcel Lüthi, and Thomas Vetter

Computer Science Department, University of Basel, Switzerland
{matthias.amberg, marcel.luethi, thomas.vetter}@unibas.ch

Abstract. Fitting statistical models is a widely employed technique for the segmentation of medical images. While this approach gives impressive results for simple structures, shape models are often not flexible enough to accurately represent complex shapes. We present a fitting approach, which increases the model fitting accuracy without requiring a larger training data-set. Inspired by a local regression approach known from statistics, our method fits the full model to a neighborhood around each point of the domain. This increases the model’s flexibility considerably without the need to introduce an artificial segmentation of the structure. By adapting the size of the neighborhood from small to large, we can smoothly interpolate between localized fits, which accurately map the data but are more prone to noise, and global fits, which are less flexible but constrained to valid shapes only. We applied our method for the segmentation of teeth from 3D cone-beam ct-scans. Our experiments confirm that our method consistently increases the precision of the segmentation result compared to a standard global fitting approach.

1 Introduction

Statistical shape models have become a commonly used tool in computer vision, especially for automated segmentation purposes. Oftentimes statistical models lack the flexibility to accurately represent all valid shapes, as the the number of training data-sets is insufficient to learn the shape-space. Many approaches have been proposed to mitigate that problem. These approaches can be grouped into three categories: The first possibility is to artificially increase the model flexibility by introducing synthetic variations of the training data at model build time [1–3]. Another option is to relax the model restrictions at fitting time [2, 4–7]. This is achieved either by finding other features in an image and then to restrict them with plausible model shapes [2, 6, 7] or by fitting a statistical model which is subsequently relaxed to match other local information [4, 5]. The problem with the above approaches is that they allow for deviations that are not explained by the training shapes. The relaxation of the shape model constraint does not guarantee anymore that the resulting shape is a likely instance of the model. Also it is a delicate matter to define artificial deformations that capture exactly the characteristics of the shape. The third class of methods tries to segment the model either spatially [8–10] or in the frequency domain [11–13]. While this increases the flexibility without above mentioned disadvantages, the main

problem here is how to choose the segments. For natural shapes, neighboring points always strongly correlate and therefore any spatial segmentation must be arbitrary. Furthermore no natural decomposition in frequency bands is known.

The method proposed in this paper focuses on fitting a model locally in the spatial domain and thus increasing its flexibility implicitly. The method is inspired by local linear regression which is commonly used in statistics [14]. We fit a complete shape model to a neighborhood around a given point of the image to segment. The fitting result is used to predict the value at this particular point only. By repeating this procedure for every point, we finally obtain a fit of the whole model. The neighborhood size determines how flexible our model is. In the extreme, considering only a single point allows any arbitrary shape to be matched. In contrast, having the neighborhood correspond to the full image leads to the global fit, which strictly adheres to the model constraint. For any local neighborhood we do adhere to the shape constraint, whereas the global shape constraint is nevertheless relaxed. We thus have the advantages of a segmented model without ever actually having to segment the model. Neither do we have to introduce artificial deformations which are not motivated by the data to increase the model’s flexibility.

This flexibility comes at the cost of an increased computational overhead. Performing a full fitting at every point becomes infeasible for large images. To lessen the computational burden we further present a method to interpolate between such local fittings, which no longer need to be created for every point but rather once for a region of definable size.

Our work was motivated by a project in which we aimed to perform model based segmentation of teeth from 3D Cone Beam CT data. The manual segmentation of such data-sets is tedious and even human experts are not always able to distinguish the tooth from the surrounding bony structure and to clearly separate neighboring teeth. With the more wide-spread use of this 3D imaging technology to plan surgical interventions, the segmentation of teeth becomes a routine task and automated procedures needs to be employed to be able to perform it efficiently. Due to the difficulty of obtaining good reference segmentations, we did not have enough data-sets available to build a shape model that represents the large anatomical variability of these shapes. Our quantitative evaluation, using a leave-one-out procedure on seven manually segmented data-sets, confirms that our localized fitting nevertheless yields accurate segmentation and consistently improves the segmentation accuracy compared to a global fitting approach.

2 Background

2.1 Statistical shape and deformation models

The idea behind statistical models is to learn the normal shape of an anatomical structure from given example shapes. The resulting model is specific to a given structure and thus can be used to perform image processing tasks, which would not be feasible otherwise. In the area of computer vision and medical image analysis, the main application of statistical models is image segmentation [15]. There

are several variants of statistical shape models, such as the Active Shape model [16], the Morphable Model [10] or Statistical Deformation Models [17]. They all apply Principal Component Analysis (PCA) to a number of example data sets S_1, \dots, S_n to model typical deformations. The examples are assumed to be suitably discretized and in correspondence. This means that for each example S_i , there exists a mapping $\phi_i : \Omega \rightarrow \mathbb{R}^d$ from a reference domain $\Omega = \{x_1, \dots, x_N\}$ such that $\phi_i(x)$ and $\phi_j(x)$ denote corresponding points of the examples S_i and S_j . It is well known that a PCA model can be written as a generative model of the form:

$$\mathcal{M}[\boldsymbol{\alpha}] : \Omega \rightarrow \mathbb{R}^d \quad x \mapsto \mu(x) + \sum_{i=1}^n \alpha_i u_i(x) \quad (1)$$

where $\mu : \Omega \rightarrow \mathbb{R}^d$ denotes the sample mean $\mu(x) := \frac{1}{n} \sum_{i=1}^n \phi_i(x)$ and $u_i : \Omega \rightarrow \mathbb{R}^d, i = 1, \dots, n$ are the principal components, computed from the examples [18]. Note that the model is completely defined by the parameter vectors $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n)$. For the case of statistical shape models, the domain Ω is usually a discretized reference surface, or a parametrization domain. The set of points $\{\mathcal{M}[\boldsymbol{\alpha}](x) \mid x \in \Omega\}$ defines the surface induced by the parameter vector $\boldsymbol{\alpha}$. In the case of statistical deformation models, Ω is usually an image domain, and $\mathcal{M}[\boldsymbol{\alpha}]$ defines a deformation field that relates the given reference and target image. To distinguish between these models, we write \mathcal{M}_S for shape models, and \mathcal{M}_D for the deformation model.

2.2 Statistical model fitting

Statistical model fitting aims at finding the set of parameters of the model (1), such that it optimally explains a given target structure. We assume in the following that the target structure has already been rigidly aligned with the model, such that only the non-rigid part of the mapping needs to be found.¹ There are various different methods for shape model fitting (see Heimann et al. [15] for a recent overview). The most simple case of shape model fitting is to directly fit a given target surface $\Gamma_T \subset \mathbb{R}^d$. Such a surface representation can often be obtained from an image using simple, intensity based segmentation methods. This rough first segmentation can then be constrained to a valid shape by fitting the shape model. The problem is formulated as an optimization problem: The goal is to find a shape, which for each point x_i minimizes the distance to the closest point in the target shape Γ_T :

$$\boldsymbol{\alpha}^* := \arg \min_{\boldsymbol{\alpha} \in \mathbb{R}^n} \sum_{x_i \in \Omega} \left(\min_{x' \in \Gamma_T} \|x' - \mathcal{M}_S[\boldsymbol{\alpha}](x_i)\| \right)^2 + \lambda \|\boldsymbol{\alpha}\|^2. \quad (2)$$

The term $\lambda \|\boldsymbol{\alpha}\|^2$ serves as a regularization term, which penalizes shapes that strongly deviate from the modeled mean shape. While this approach is simple, its main problem is that the cost function (2) includes only boundary information of the modeled shapes. An approach that includes the complete image information

¹ This can be achieved, for example, by performing a Procrustes Alignment [19] on a number of landmark points defined on the reference and target shape.

has been proposed by Rückert et al.[17]. The idea is to directly build a statistical model of the deformation fields $\mathcal{M}_D[\alpha]$, which relate the image intensities of a reference image $I_R : \Omega \rightarrow \mathbb{R}$, to given examples images. The goal of model fitting is to find the deformation represented by this model, which best relates the reference image I_R to a new target image $I_T : \Omega \rightarrow \mathbb{R}$:

$$\alpha^* := \arg \min_{\alpha \in \mathbb{R}^n} \sum_{x_i \in \Omega} (I_R(x_i + \mathcal{M}_D[\alpha](x_i)) - I_T(x_i))^2 + \lambda \|\alpha\|^2. \quad (3)$$

Note that (3) establishes point-to-point correspondence between the images. It can therefore be used to map labels defined on the reference image onto the target. In particular, given a segmentation of the reference image $B_R : \Omega \rightarrow \{0, 1\}$, the warped image $B_R(x + \mathcal{M}_D[\alpha^*](x))$ yields a segmentation of the target image. Although the Problems (2) and (3) differ in the way they use the image information, their solution depends only on the model coefficients α . The strategy for increasing the fitting accuracy that we are proposing in the next section holds unchanged for both approaches. For clarity of presentation we will discuss our method on the simpler case of shape model fitting (Equation 2). For the practical application that we present in Section 4 we will, however, apply deformation model fitting (Equation 3).

3 Local shape model fitting

The use of model fitting to obtain a segmentation result, which is restricted to valid anatomical shapes, is appealing. However the models are often not sufficiently flexible to represent complicated structures accurately. It has been shown in the literature that spatial segmentation of the model can increase the fitting accuracy [8–10]. The reason for the improved accuracy is that each segment only has to represent a local part of the shape, which, by itself, is less complex than the complete global shape. These segments can therefore be accurately represented using fewer example shapes. Our approach exploits the same idea, but instead of partitioning the model we consider local areas of the target structure to which we fit the full model. Similar to a local regression approach known from statistics [20], we fit the model to a local neighborhood around each data point of the target. This strategy avoids completely the problem of having to segment the model into independent parts.

We illustrate the method for the case of shape model fitting using a 2D example of hand shapes. Let \mathcal{M}_S be a statistical shape model and $\Gamma_T \subset \mathbb{R}^d$ a target shape. For each point $x_0 \in \Omega$ we fit the entire shape model, but weight the influence of each point $x \in \Omega$ on the total cost with a weight w_{x_0} depending on its distance to the point x_0 :

$$\alpha_{x_0}^* := \arg \min_{\alpha \in \mathbb{R}^n} \sum_{x_i \in \Omega} w_{x_0}(x_i) \left(\min_{x' \in \Gamma_T} \|x' - \mathcal{M}_S[\alpha](x_i)\| \right)^2 + \lambda \|\alpha\|^2, \quad (4)$$

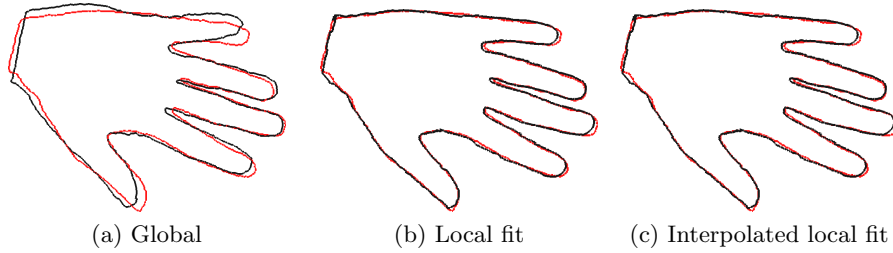


Fig. 1: A comparison between a global fit (a) and local fitting results (b), (c) for a shape model which was built from 15 hand shapes. The target hand shape was not part of the model’s training data. The result shown in (b) has been computed by fitting every point, whereas in (c) only 20 points were fitted and the remaining ones interpolated.

Although the solution $\alpha_{x_0}^*$ defines a full shape, it is only used to determine the result Γ^* at the point x_0 :

$$\Gamma^*(x_0) := \mathcal{M}[\alpha_{x_0}^*](x_0) \quad (5)$$

The fitting is performed at every point $x_0 \in \Omega$, and therefore we eventually obtain a result for the whole shape. Every fit strives to minimize the error around x_0 , and thus the resulting shape explains the target better than any global fit would. As the local neighborhoods of two nearby points greatly overlap, the resulting fits are similar and the resulting target shape $\Gamma^*(x_0)$ is nevertheless smooth. The problem of discontinuous segments, as observed with segmented models, is completely avoided.

A typical choice for w_{x_0} is the *Epanechnikov kernel* [14]:

$$w_{x_0}(x) := \kappa_\sigma(x_0, x) = D\left(\frac{\text{dist}(x, x_0)}{\sigma}\right) \quad (6)$$

with

$$D(t) = \begin{cases} \frac{3}{4}(1 - t^2) & \text{if } |t| \leq 1 \\ 0 & \text{otherwise.} \end{cases} \quad (7)$$

The weight function κ_σ is compact and its support determined by σ . For surface fitting, the distance $\text{dist}(x, x_0)$ is taken to be the geodesic distance on Ω , as we wish to match neighboring points on the surface. Figure 1 shows an example of how our procedure improves the fitting accuracy.

The fitting procedure is only useful if the shape constraints are still enforced to the degree that artifacts and noise in the data are not fitted. This is determined by the parameter σ in our kernel (6), which acts as a regularization parameter. If σ is small the fitting is local and the target shape is accurately represented. In the extreme case when we fit the model to the point x_0 only, the shape statistics is completely ignored. On the other hand, if σ is large, all points of the shape have nearly the same influence, and we arrive at the global fitting. The effect of this parameter is illustrated in Figure 2, where we fitted a hand shape with a number of manually introduced artifacts. We observe that by increasing σ we

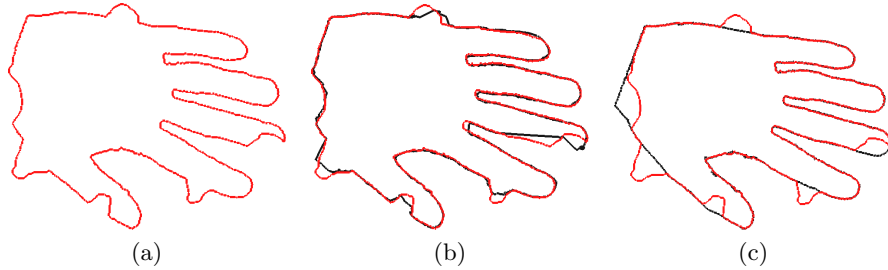


Fig. 2: An example of a local fitting on noisy data. (a) shows the hand with manually introduced artifacts (red line). Fitting only a small neighborhood leads to overfitting (black line). Increasing the size of the neighborhood eliminates the influence of the artifacts almost completely .

can reduce the influence of the artifacts, and still accurately represent the actual target shape.

3.1 Interpolated Local Shape Model Fitting.

Evaluating Equation 4 at every point $x \in \Omega$ quickly results in an excessive computational burden for densely sampled shapes. To address this problem, we perform the fitting only at a subset of the points $\tilde{\Omega} := \{x_0, \dots, x_K\} \subset \Omega$. Observe that to define the fitting result Γ^* at point x_k in Equation 5, we used only the single point $\mathcal{M}[\alpha_{x_k}^*](x_k)$, even though α^* defines a full shape

$$\Gamma_{x_k}^* := \mathcal{M}[\alpha_{x_k}^*].$$

Thus, the idea is to combine these local predictions in a smooth way:

$$\Gamma^*(x) := \frac{\sum_{k=1}^K \hat{w}_k(x) \Gamma_{x_k}^*(x)}{\sum_{k=1}^K \hat{w}_k(x)}, x \in \Omega \quad (8)$$

where we choose \hat{w}_k again to be the Epanechnikov kernel (6). Note that since \hat{w}_k is compactly defined, for (8) to be well defined, the support of the kernel needs to be chosen such that for each point $x \in \Omega$ at least one of the weight functions $\hat{w}_k(x)$ is non-zero.

4 Results

In this section we show how our method can be applied for the automatic segmentation of teeth from Cone Beam CT images. Each image has a resolution of $90 \times 71 \times 31$ voxels. Figure 3a shows a sample slice through such a tooth ct-image. The automatic segmentation of such images is a challenging task. Especially at the roots, the tooth is virtually indistinguishable from the surrounding bony structure. Furthermore, neighboring teeth can be touching and are thus difficult

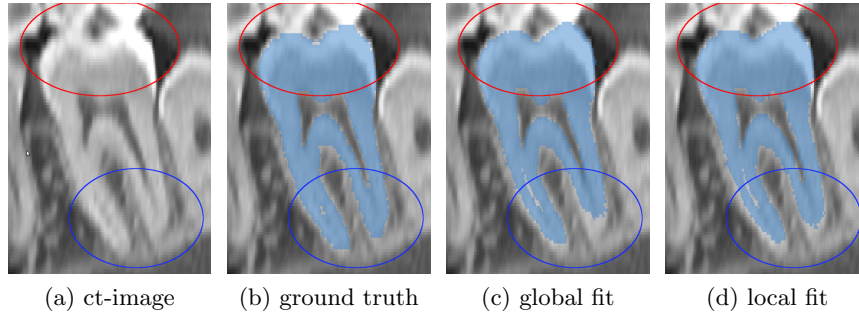


Fig. 3: A typical ct-scan of a tooth is shown in (a) which features typical problems faced in ct-image segmentation. The image is noisy and a neighboring tooth is touching our target tooth (red circle). The local fit (d), while robust enough not to leak into the neighboring tooth at the top, segments the root section better than the global fit (c) (blue circle), compared to the ground truth (b).

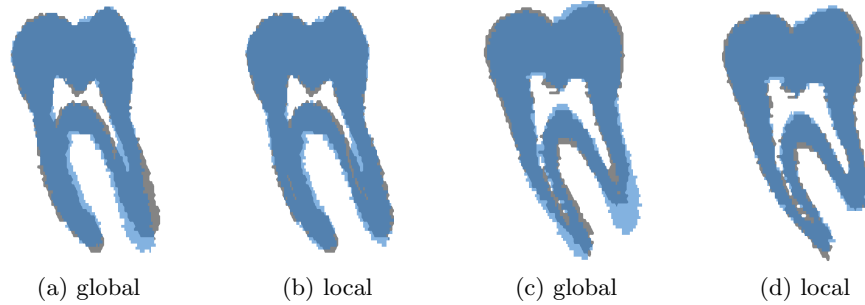


Fig. 4: Fitting results (blue) for two sample teeth are compared to ground truth (grey). Especially at the roots, our local approach (b) and (d) is more precise than the global one (a) and (c).

to separate. To address these problems, we manually segmented seven data-sets to create a deformation model, which we subsequently used for segmentation. We choose one image as the reference, to which we aligned all the images using Procrustes alignment [19] on eight manually defined landmarks. For solving the optimization problem, we used a standard LBFGS-B optimizer [21].

The shape of the tooth can vary greatly among individuals, and a model built from only seven datasets is not sufficient to span this variation. By using our interpolated local shape model fitting, we get more accurate segmentation results, as shown in Figure 4. Especially at the roots of the tooth, which exhibit most variation, the local fitting clearly leads to a large improvement.

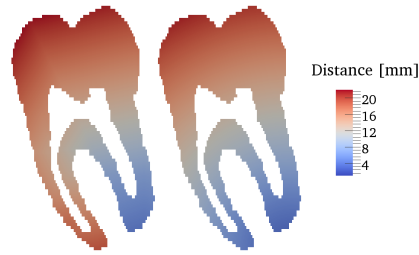


Fig. 5: Comparison of our distance measure (left) to the Euclidean distance (right), measured from a point at the lower right root. Note the difference of the obtained distance in the left root as our distance measure is notably higher than the Euclidean distance in this area.

4.1 Distance Measure

For the shape fitting example discussed in Section 3, we used the geodesic distance to determine the local neighborhood. This is not possible for the deformation model, which is defined on the image domain. However, we would still like to quantify the distance in terms of the shape that we model. We therefore define the distance between two points that lie within the structure as shortest path between those points with the constraint that the path always resides within the structure’s volume. Figure 5 shows a comparison of the Euclidean distance and our distance function. With the Euclidean distance, the tips of the roots are rather close to each other, in our distance measure they are far apart.

4.2 Global fit versus Interpolated Local Fit Comparison

We performed a leave-one-out test to assess the fitting quality. We used six manually segmented data sets, similar to the one shown in Figure 3a, for model building. We then fitted the left-out tooth with the global approach as well as our interpolated local fitting method. We used 35 equidistant points within the tooth as our local region defining $\Gamma_{x_k}^*$ (Cf. Section 3.1). In all cases the interpolated local fitting method provided better fitting results than the global method. Table 1 shows the quantitative results of this leave-one-out test. We used the average Hausdorff Distance to measure the quality of our fittings to the ground truth. Informally the average Hausdorff Distance tells us how far away, on average, the other shape is to be expected, if a random point in a shape is chosen. It measures how well two shapes match each other. More precisely the average Hausdorff Distance between two volumes S_1 and S_2 is defined as the total sum of the shortest path of each point in shape S_1 to the closest point in S_2 and vice versa, divided by the total number of points of both shapes.

Table 1: Measured error based on the ground truth of a global fit and an interpolated local regression based fit.

Average Hausdorff Distance [pixel]							
tooth #	1	2	3	4	5	6	7
global fitting error	0.084	0.089	0.301	0.221	0.203	0.138	0.151
local fitting error	0.071	0.075	0.240	0.143	0.086	0.119	0.113

5 Discussion

We presented an approach for the fitting of statistical models. Our method is based on the local regression method known from statistics. The core idea is to fit the model to a local region around a point for each point of the model. In this way we obtain a fitting result which is solely based on the model information but still allows for more flexibility than a global fitting method. In contrast to previous methods, no segmentation of the structure needs to be performed, nor do artificial deformations need to be included. The size of the neighborhood that is considered for each fit determines the trade-off between obtaining a fit that strictly adheres to the global shape constraint, and one that accurately explains all the data. This parameter should be chosen such that it reflects the noise properties of the images and the quality of the model.

For large images, fitting a full model at every point is computationally too expensive. To reduce the computational burden of such an approach, we proposed a method to interpolate local shape model fits performed only for a subset of all model points. In this way it becomes feasible to apply the method for the segmentation of large 3D images. We presented an application of our method for the segmentation of teeth from Cone Beam CT images. Our tests confirmed that the local fitting method improves the segmentation results consistently, compared to the global fitting results.

The number of points we choose for the interpolated fitting determines how many times we have to perform the fitting. It would therefore be interesting to choose the points, such they most effectively increase the fitting accuracy. The exploration of strategies for choosing these point optimally, is an interesting problem that will be the subject of future work.

Acknowledgements

This work has been supported by the CO-ME/NCCR research network of the Swiss National Science Foundation (<http://co-me.ch>).

References

1. Cootes, T.F., Taylor, C.J.: Combining point distribution models with shape models based on finite element analysis. *Image Vision Comput.* **13**(5) (1995) 403–409

2. Cootes, T.F., Taylor, C.J.: Data driven refinement of active shape model search. In: BMVC, British Machine Vision Association (1996)
3. Loog, M.: Localized maximum entropy shape modelling. In: Information Processing in Medical Imaging, Springer 619–629
4. Pekar, V., Kaus, M., Lorenz, C., Lobregt, S., Truyen, R., Weese, J.: Shape-model-based adaptation of 3D deformable meshes for segmentation of medical images. In: Proceedings of SPIE. Volume 4322. (2001) 281
5. Shang, Y., Dossel, O.: Statistical 3D shape-model guided segmentation of cardiac images. *COMPUTERS IN CARDIOLOGY* **31** (2004) 553
6. Weese, J., Kaus, M., Lorenz, C., Lobregt, S., Truyen, R., Pekar, V.: Shape constrained deformable models for 3D medical image segmentation. *Lecture notes in computer science* (2001) 380–387
7. Shen, D., Herskovits, E.H., Davatzikos, C.: An adaptive-focus statistical shape model for segmentation and shape modeling of 3-d brain structures. *IEEE Trans Med Imaging* **20**(4) (April 2001) 257–70
8. de Bruijne, M., van Ginneken, B., Viergever, M.A., Niessen, W.J.: Adapting active shape models for 3d segmentation of tubular structures in medical images. *Inf Process Med Imaging* **18** (July 2003)
9. Zhao, Z., Aylward, S., Teoh, E.: A novel 3D partitioned active shape model for segmentation of brain MR images. *Lecture Notes in Computer Science* **3749** (2005)
10. Blanz, V., Vetter, T.: A morphable model for the synthesis of 3d faces. In: SIGGRAPH '99: Proceedings of the 26th annual conference on Computer graphics and interactive techniques, ACM Press (1999) 187–194
11. Davatzikos, C., Tao, X., Shen, D.: Hierarchical active shape models, using the wavelet transform. *IEEE Trans Med Imaging* **22**(3) (March 2003) 414–23
12. Nain, D., Haker, S., Bobick, A., Tannenbaum, A.: Multiscale 3-d shape representation and segmentation using spherical wavelets. *IEEE Trans Med Imaging* **26**(4) (April 2007) 598–618
13. Knothe, R.: A Global-to-local model for the representation of human faces. PhD thesis, Computer Science Department, University of Basel (2009)
14. Hastie, T., Tibshirani, R., Friedman, J.: Kernel Smoothing Methods. Springer Series in Statistics. In: The Elements of Statistical Learning. Springer New York Inc., New York, NY, USA (2001)
15. Heimann, T., Meinzer, H.: Statistical shape models for 3D medical image segmentation: A review. *Medical Image Analysis* (2009)
16. Cootes, T., Taylor, C., Cooper, D., Graham, J., et al.: Active shape models-their training and application. *Computer Vision and Image Understanding* **61**(1) (1995) 38–59
17. Rueckert, D., Frangi, A.F., Schnabel, J.A.: Automatic construction of 3d statistical deformation models using non-rigid registration. In: MICCAI '01: Proceedings of the 4th International Conference on Medical Image Computing and Computer-Assisted Intervention, London, UK, Springer-Verlag (2001) 77–84
18. Tipping, M.E., Bishop, C.M.: Probabilistic principal component analysis. *Journal of the Royal Statistical Society* **61** (September 1999) 611–622
19. Dryden, I., Mardia, K.: Statistical shape analysis. Wiley New York (1998)
20. Cleveland, W.S., Devlin, S.J.: Locally-Weighted regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association* **83**(403) (1988) 596–610
21. Zhu, C., Byrd, R., Lu, P., Nocedal, J.: Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. *ACM Transactions on Mathematical Software (TOMS)* **23**(4) (1997) 550–560