Available online at www.sciencedirect.com
ScienceDirect
Medical Image Analysis 12 (2008) 136–151
MEDICAL IMAGE ANALYSIS
www.elsevier.com/locate/media

Automatic shape model building based on principal geodesic analysis bootstrapping

Erik B. Dam [a,*], P. Thomas Fletcher [b], Stephen M. Pizer [b]

[a] *Nordic Bioscience, Imaging, Herlev Hoved 207, Herlev, Denmark*
[b] *MIDAG, University of North Carolina, Chapel Hill, NC, United States*

Received 16 May 2006; received in revised form 11 June 2007; accepted 31 August 2007
Available online 25 September 2007

Abstract

We present a novel method for automatic shape model building from a collection of training shapes. The result is a shape model consisting of the mean model and the major modes of variation with a dense correspondence map between individual shapes. The framework consists of iterations where a medial shape representation is deformed into the training shapes followed by computation of the shape mean and modes of shape variation.

In the first iteration, a generic shape model is used as starting point – in the following iterations in the bootstrap method, the resulting mean and modes from the previous iteration are used. Thereby, we gradually capture the shape variation in the training collection better and better. Convergence of the method is explicitly enforced.

The method is evaluated on collections of artificial training shapes where the expected shape mean and modes of variation are known by design. Furthermore, collections of real prostates and cartilage sheets are used in the evaluation.

The evaluation shows that the method is able to capture the training shapes close to the attainable accuracy already in the first iteration. Furthermore, the correspondence properties measured by generality, specificity, and compactness are improved during the shape model building iterations.

Keywords: Shape modeling; Correspondence; Shape variation; Shape mean; Principal geodesic analysis

1. Introduction

Methods based on analysis of shape variation are becoming widespread in medical imaging. These methods allow a statistical modeling of prior shape knowledge in tasks where the image information itself often is not strong enough to solve the task automatically. The obvious example is the use of deformable models in segmentation, where the preferred deformations are determined by a statistical shape model. Another important task is shape analysis and classification, where a statistical shape model provides distributions of healthy and diseased organs for diagnostic methods.

The most common statistical shape models consist of a mean shape with deformations. The mean and the corresponding deformations are constructed through statistical analysis of shapes from a collection of training data. Each shape in the training set is represented by the chosen shape representation, and analysis of the parameters for the representation gives the mean and variations.

The best known shape model of this type is the *Active Shape Model* (Cootes et al., 1995). Here, the shapes are represented by a point distribution model (PDM) with given point-wise correspondence. The mean model is achieved after Procrustes alignment of the shapes followed by simple mean position computation of each point in the model.

* Corresponding author. Tel.: +45 4454 7777; fax: +45 4454 8888.
E-mail address: erikdam@nordicbioscience.dk (E.B. Dam).
URL: http://www.diku.dk/~erikdam (E.B. Dam).

1361-8415/$ - see front matter © 2007 Published by Elsevier B.V.
doi:10.1016/j.media.2007.08.004

Principal component analysis (PCA) is used to provide the variations.

PCA is only valid for Euclidean parameter spaces and is restricted to producing linear shape variations. Other approaches for shape models based on PDM's allow more complex, non-linear shape variations through the use of kernel PCA (Schölkopf et al., 1998). Alternatively, the shape model can be based on other shape representations such as the implicit Level Set formulation as in Cremers et al. (2002).

This paper pursues a medial shape representation in the form of the *m-rep* (Pizer et al., 1996). This representation offers an intuitive visualization of the shape by means of the sheet of sampled medial atoms. Compared to PDM's this representation is less simple since the parameter space is not Euclidean but consists of a combination of position, scaling, and orientation parameters. Standard PCA is therefore not applicable. However, the analogue of PCA has been developed for a more general space of shape representations. This is the principal geodesic analysis (PGA) that applies to shape representations that form Lie groups (Fletcher et al., 2003b).

1.1. Establishing correspondence

A key element in constructing shape models is the representation of the shapes in the training collection. This must be done in a manner that defines/preserves correspondence across the population. For PDM's the simplest method is by manual selection of the boundary points by an expert of the specific anatomical structure. In 2D this is a time-consuming and tedious process – in 3D it is not feasible. However, this process can be automated and several approaches for this have been published.

Morphable Surface Models (Shelton, 2000) create correspondences and a shape model from surface meshes that are deformed onto first one of the shapes and then the resulting mean shape in a bootstrap optimization method. The deformations are guided by regularization terms penalizing surface deformation and correspondence distortion in a multi-resolution optimization.

The Iterative Closest Point algorithm (Besl and McKay, 1992) is used as basis in the approach by Brett and Taylor (2000) where corresponding points are generated from sparse surface meshes and then propagated in a binary tree.

The approach presented by Davies et al. (2002) starts by generating roughly corresponding PDM boundary points from a spherical harmonics shape representation. This set of boundary points and their correspondences are then optimized through a minimum description length (MDL) approach. The method is inspired by the method in Kotcheff and Taylor (1998) and both aim at optimizing an expression for correspondence optimality (model compactness) directly rather than the previous methods that rely on optimizing properties that are intuitively geometrically sensible. Recently, a faster, gradient descent based improvement of the MDL optimization scheme in 3D has been developed (Heimann et al., 2005).

The correspondence problem can also be attacked as a registration problem as done in both (Rueckert et al., 2003; Twining et al., 2005). Finally, for shape models based on a Level Set representation, there is also a mathematically well-founded framework for automated shape model building based on statistics of warpings between example shapes (Charpiat et al., 2005). For more references to methods for constructing shape model correspondence, see Cootes' time-line of developments (Cootes, 2005).

1.2. The contributions of this paper

The contribution of this paper is a novel, automated shape modeling method. The core is a bootstrap process that iteratively optimizes the shape representation on a training collection and then derives the PGA mean and modes of deformation. Through the bootstrap iterations, the PGA mean converges to an increasingly accurate shape mean for the collection. The PGA modes of variation converge and give the modes of variation for the shape model.

The flavor of this work resembles the MDL method in Davies et al. (2002). The main difference – apart from the shape representation – is that the MDL approach starts the optimization process from shape representations derived directly from the training shapes but with relatively poor correspondence. The MDL optimization then hopes to retain the representation/shape fit while optimizing the correspondence. Conversely, the PGA bootstrap starts from a generative model with explicit correspondence but with poor fit to the individual training shapes. The bootstrap process then keeps the correspondence while optimizing the individual shapes to fit to the training shapes.

We evaluate the PGA bootstrap method for constructing shape models for three shape collections. The first is a set of simple, artificial shapes where the expected mean and shape variations are intuitively clear.

Second, we use a population of prostates. The collection consists of 46 cases where the prostates were segmented in the course of radiation treatment of prostate cancer. In the CT scans routinely used in radiation treatment planning, the prostate boundaries are quite fuzzy – therefore, the prior knowledge in a statistical shape model is essential for segmenting the prostate.

Finally, we build a shape model for articular cartilage from knee MR scans. Analysis of articular cartilage is central in disease progression quantification (for osteoarthritis) and the shape model is useful for automatic segmentation methods and crucial for defining the correspondence needed for analysis of focal measures such as thickness.

We extend previous work (Dam et al., 2004) but this paper has several key contributions:

- The shape match that drives the model deformation is novel.

- The model regularization is novel.
- Convergence of the shape model building iterations is now enforced.
- The method is here evaluated for modeling of artificial shapes and knee cartilage sheets as well as the prostate collection from our previous work.
- Correspondence properties of the framework are evaluated.

These contributions result in far better shape representations and more intuitive shape models than in our previous work.

We first introduce the basic problem of shape modeling in Section 2. After this we present our framework in Sections 3–6. The collections of real prostate and knee cartilage sheets are introduced in Sections 7 and 8.

The evaluation in Section 9 shows that our framework manages to capture the shapes in the training collection near the attainable accuracy and that the method iterations improve the correspondence properties. Thereby, we conclude that our method is a promising automatic shape model building method.

2. Problem statement

When building a shape model, the task is to construct a compact description of a distribution of shapes observed through a training collection. The task at hand determines which aspects of shape to model. In this work, we focus on the "overall shape". Thereby, we ignore fine scale details and seek a compact representation of smooth shapes.

Here we introduce a collection of artificial shapes where the desired shape model is intuitively clear (see Fig. 1). We also use a noisy version of the basic shapes with pixel-wise, normal distributed variation in the local thickness (with a standard deviation of 20% of the basic thickness).

Since we seek a model of smooth shapes, we aim for a smooth representation of the training shapes. The basic version could be seen as a smooth approximation of the noisy version. For pairs of basic and noisy shapes, the mean dice similarity coefficient (DSC) is 0.94. DSC defines volume overlap between sets A and B as $2|A \cap B|/(|A| + |B|)$ (Dice, 1945). A shape model representation of the noisy shapes should aim for an equally good overlap.

For a given, real-world data collection, only the "noisy" version is available. It is therefore not possible to determine the desired overlap between the data collection shapes and the smooth shape model representation directly. In order to estimate how accurately a smooth model can at best represent the shapes, such a collection can be compared to a version where each shape is smoothed. This is done by smoothing the corresponding binary image with a Gaussian at the scale of the data voxels followed by thresholding. Other smoothing kernels are possible, but this seems to be the natural, least committed choice defined by the data collection. We denote the mean volume overlap between the original collection shapes and the smoothed versions as the *attainable accuracy*. This attainable accuracy is determined by the training collection alone. Furthermore, it is not a strict, theoretical upper limit for the overlap between the collection shapes and the smooth shape model but rather a pragmatic golden standard. Whether a smooth shape model will be able to approach this attainable accuracy when representing the shapes will be determined by the model rigidity/flexibility, sampling/compactness, and other model properties.

For the basic version shapes, the DSC between shapes and smoothed shapes is 1.00. For the noisy version, this attainable accuracy is 0.95. Thereby, we have estimates for how well the overlap between the shapes and the shape representations should be.

A shape model constructed from shape representations of these training collections should result in a distribution of flat, rectangular shapes with varying degrees of thickness, bending and width. Since we focus on modeling

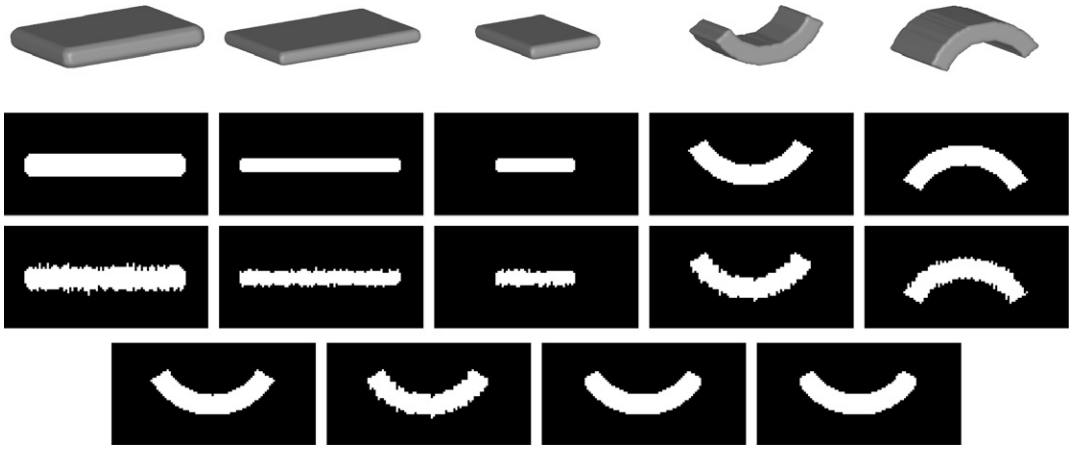

Fig. 1. Collection of artificial, flat shapes. Top row shows the five shapes with varying thickness, width/length ratio, and bending. The two center rows show center slices of the basic and noisy shapes, respectively. Bottom row shows the slice from the fourth shape as basic, noisy, basic smoothed and noisy smoothed.

"overall shape" the basic and the noisy collections should result in virtually identical shape models.

2.1. Properties of a statistical shape model

In general, when building shape models for anatomical shapes, there are a number of desirable properties to strive for. First, the model should represent the shapes accurately.

Second, the model should approximate an assumed underlying statistical distribution of shapes. This distribution is obviously unknown for anatomical shapes, but a collection of training shapes can be used to model it from. The success of this modeling is often evaluated indirectly in terms of the generality and specificity of the model. Generality is the ability to recognize/model unknown shapes from the underlying distribution. Specificity is the ability to accept only plausible shapes as belonging to the distribution. These two desirable properties are to some degree opposing: generality suggests a flexible model whereas specificity suggests a more constrained model.

In addition to modeling the anatomical shapes, the shape model should also provide a coordinate system that allows focal analysis of corresponding locations in the shapes. The correspondence given by the coordinate system should respect anatomically corresponding landmarks. This requirement suggests that the shape model should restrict deformations that destroy correspondence – making the model less flexible and thereby opposing both the desires for accuracy and generality.

Finally, for robustness and computation reasons, the model should ideally be a sparse representation of the anatomical shapes. This desire for compactness in the model opposes the desire to represent the shapes accurately.

It is non-trivial to propose a unifying theory that reveals how these desirable properties should be balanced against each other. For specific tasks in medical image analysis, the evaluation should be focused on the clinical end goal instead. In our presentation of a general framework, we will discuss how to evaluate *accuracy*, *generality*, *specificity*, *correspondence*, and *compactness* individually in the sections below.

3. Medial shape representation: m-rep

We use a medial shape representation, m-rep, as our smooth shape representation of choice. Below, we briefly review the geometry and the framework for image segmentation (Pizer et al., 2003; Joshi et al., 2002) before the image match and the regularization terms essential for the shape model building framework are introduced in the next section.

The m-rep is based on the medial axis of Blum and Nagel (1978). In this framework, a 3D geometric object is represented as a set of connected continuous medial sheets, which are formed by the centers of all spheres that are interior to the object and tangent to the object's boundary at two or more points. Here we focus on 3D objects that can be represented by a single medial sheet.

We sample the medial sheet $\mathscr{M}$ over a spatially regular lattice of *medial atoms* (see Fig. 2) defined as a 4-tuple $\mathbf{m} = \{\mathbf{x}, r, \mathbf{F}, \theta\}$, consisting of: $\mathbf{x} \in \mathbb{R}^3$ and $r \in \mathbb{R}^+$, the center and radius of the sphere, $\mathbf{F} \in \mathbf{SO}(3)$ an orthonormal local frame parameterized by $(\mathbf{b}, \mathbf{b}^{\perp}, \mathbf{n})$, where $\mathbf{n}$ is the normal to the medial manifold, $\mathbf{b}$ is the direction in the tangent plane of the fastest narrowing of the implied boundary sections, and $\theta \in [0, \pi)$ the object angle determining the angulation of the two implied opposing boundary points to the local frame. Given an m-rep figure, we fit a smooth boundary surface to the model. We use a subdivision surface method (Thall, 2002) that interpolates the boundary positions and normals implied by each atom.

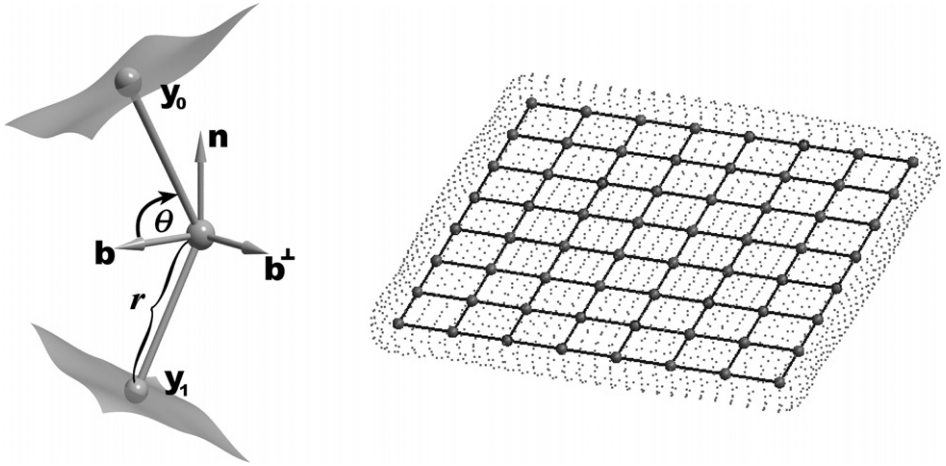

Fig. 2. Left: medial atom. Right: grid of 8×8 medial atoms with the interpolated, implied boundary illustrated by the dotted surface. The online color version illustrates the **b** vectors and implied boundary point vectors as well.

4. From shapes to m-reps

The previous section describes how to derive a shape from an m-rep model. For shape model building, it is required to go the other way: from the shape to an m-rep. Since the m-rep is a generative shape model, representing a shape is equivalent to segmenting the binary image corresponding to the shape.

Following the deformable models paradigm, an m-rep model $\mathbf{M}$ is deformed into an image I by optimizing an objective function:

$$F(\mathbf{M}, I) = L(\mathbf{M}, I) + \alpha G(\mathbf{M}) + R(\mathbf{M}) \quad (1)$$

The function L, the *image match*, measures how well the model matches the image information, while G, the *geometric typicality*, gives a prior on the possible variation of the geometry of the model weighted by $\alpha \geqslant 0$. The last term R, the *model regularization*, is an addition to the standard m-rep framework. In the basic m-rep framework, there is no unique model corresponding to a given shape – the model regularization term is designed to pick a "nice" representative among the manifold of possible representations. These terms are described in detail below.

4.1. Objective function: image match

In this shape model building framework, m-rep models are fitted to the desired shapes. In the previous work (Dam et al., 2004), the image match term of the objective was computed as the correlation between the image and a Gaussian derivative kernel in the normal direction of the object at the implied boundary points. This image match term proved to be too local. In a segmentation setting where the model can be initialized close to the desired shape, this type of image match is often sufficient – but while building a shape model there is no prior shape model and therefore the image match needs to be able to attract the model boundary to the shape boundary without requiring accurate initialization.

Instead we here introduce the distance from the model boundary to the shape boundary integrated over the model boundary as the basic image match function to minimize. This has some clear advantages. First, a distance transform from the shape boundary can be computed prior to model optimization and thereby ensure fast optimization. Second, the model can be attracted to the shape from infinitely far away.

The computation of the distance from the model boundary to the shape boundary has two special cases in order to help the optimization.

First, we use the shape boundary distance gradient and the model boundary normal to determine whether the nearest shape boundary is the correct one (as illustrated in Fig. 3). This is done to ensure convergence to the correct optimum. Otherwise the model could shrink to a single point arbitrarily located at the shape boundary and reach minimal shape match penalty. Also, it avoids faulty local minima where a part of the model boundary gets stuck at the wrong shape boundary. This is checked at the model boundary point given by the parameter s with the transition function ws (for wrong side):

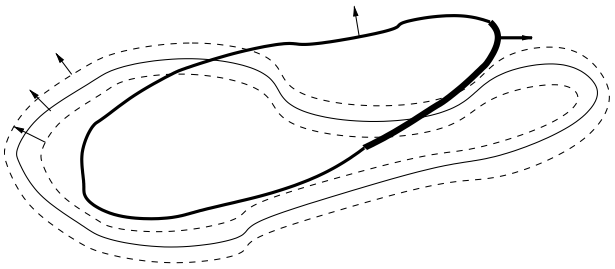

Fig. 3. The image match at a model boundary point is computed as the distance to the shape boundary. The desired shape is shown in thin black with dotted iso-distance curves. The model boundary is shown in thick black. The gradient of the shape boundary distance function and the model boundary normals are used to ensure that the boundary point does not get attracted to the wrong boundary. Shape distance is negative inside the object, so shape distance gradient arrows point away from the center of the shape. When the model normal is in opposite direction to the shape distance gradient (determined by the sign of the dot product) the boundary distance is not just the value of the shape boundary distance function but approximated by adding the maximal shape radius. This is the case for the thicker part of the model boundary. The principles are identical in 3D.

$$ws(s) = \min(1, \max(0, -\mathbf{mbn}(s) \cdot \mathbf{g}(s)))$$

where **mbn** is the model boundary normal and **g** is the shape distance function gradient. By not normalizing the gradient, we avoid arbitrary directions given by gradients close to zero. The function gives a soft transition between 0 and 1 for model boundary point on the right vs. wrong side.

Second, for thin shapes the central part of the model optimization is the phase where the model ends need to crawl into the shape ends (illustrated in Fig. 4). This is problematic since the shape boundary distance function offers no help. The ends of the model are equally likely to move slightly inwards or outwards depending on the local narrowing of the thin shape and then get stuck in a local minimum.

One solution to this problem could be to let the end parts of the model be optimized by the distance to the shape boundary in a direction normal to the model (as opposed to normal to the shape boundary as described above). This will fail for special shapes (for example a constantly curving shape) and experiments on both artificial and real data also show that the ends get stuck in local minima.

Another solution is to add the distance from the model boundary to the shape boundary in a direction normal to the model boundary – integrated over the shape boundary instead of over the model boundary. This would solve the problem. However, computing distances from the model boundary is computationally costly since the model is repeatedly changed during optimization.

We propose to add a constant expanding force to the end parts of the model boundary that are inside the shape

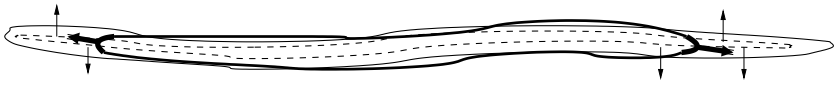

Fig. 4. For boundary points at the ends of the model, the shape boundary distance function offers no help for the deformation of the model. The thick black model boundary points with a normal approximately perpendicular to the black shape boundary distance gradient are assigned an outward force in the optimization that allows the model to crawl into the ends of the shape. The principles are identical in 3D.

boundary. In the optimization scheme, this simply means that the change of the image match is defined as the signed displacement in a direction normal to the model boundary. This is applied at model boundary points where the normal is approximately perpendicular to the shape boundary distance gradient. As the model crawls into the end of the shape during optimization, the boundary distance gradient becomes approximately parallel to the model boundary normal and this expanding force is canceled.

Analogously, for model boundary points outside the shape boundary also with normal approximately perpendicular to the shape boundary distance gradient, an imploding force is added. Thereby, the model boundary can also crawl into thin indentations in the shape.

The model boundary places where crawling is relevant are detected by the transition function *crawl*:

$$crawl(s) = \min\left(1, \left(\max\left(0, 2 - \frac{1}{\theta_{\text{crawl}}}\left|angle(\mathbf{mbn}(s), \mathbf{g}(s)) - \frac{\pi}{2}\right|\right)\right)\right)$$

The free parameter θ_{crawl} determines how close to perpendicular the model boundary and the shape boundary distance function gradient must be before crawling. The term inside $|\cdot|$ is an angle being zero when the model boundary normal and the shape boundary distance function gradient are perfectly perpendicular. The remaining expression ensures *crawl* equal 1 for this angle smaller then θ_{crawl} with a linear transition to 0 for the angle increasing to $2\theta_{\text{crawl}}$.

The image match term of the objective function, L, is now computed as the integral of the boundary distance, modified as described above, at each model boundary point:

$$L(\mathbf{M}, I) = \int_{\mathscr{B}(\mathbf{M})} (1 - ws(s)) Dist(s) + ws(s)(S_{\text{dist}}(s) + shapeDiam)\, \mathbf{ds}$$

$$Dist(s) = (1 - crawl(s))|S_{\text{dist}}(\mathbf{mb}(\mathbf{s}))| + crawl(s)\, Displ(s)$$

$$Displ(s) = -\text{sign}(S_{\text{dist}}(\mathbf{mb}(\mathbf{s}))) \frac{\partial \mathbf{mb}(\mathbf{s})}{\partial t} \cdot \mathbf{mbn}(\mathbf{s})$$

where **mb** is a model boundary point using some parameterization s of the model boundary $\mathscr{B}(\mathbf{M})$. S_{dist} is the shape boundary distance transformation function – computed using an anisotropic version of the Danielsson distance transform (Danielsson, 1980) and tri-linearly interpolated to provide values at arbitrary model boundary point positions.

As explained above, the transition function *ws* determines whether to use the shape boundary distance directly or to add the diameter of the shape *shapeDiam* (computed as twice the maximal internal absolute shape boundary distance) and thereby go across the shape when computing the distance to the shape boundary.

And *crawl* determines whether crawling is relevant at the model boundary and thereby invokes the displacement given by *Displ*. The parameter t in the partial derivative on **mb** is a fiducial optimization evolution parameter – the term simply means that a movement in the direction of the model boundary normal is interpreted as a step towards the shape boundary. The sign of the shape distance function ensures that this term applies both for model boundary points inside the shape crawling out into extrusions and for points outside crawling into indentations.

The details make the image match function appear complicated. It is simply the distance from the model boundary to the shape boundary – just with two intuitively simple special cases that ensure that the model boundary does not get stuck on the outside shape boundary and that the model boundary is allowed to crawl into thin extrusions/indentations.

These two special cases would not be necessary if the image match was, for instance, the sum of the distances from the shape to the model and the distances from the model to the shape. However, since the shape boundary distance transformation can be computed once before optimization of the model the method presented here is much more computationally efficient.

4.2. Objective function: geometric typicality

Intuitively, the geometric typicality term is a geometry prior that ensures that the model optimization favors models with a probable geometry. However, when building a shape model, no shape prior is available. Therefore, the geometric typicality term, G, is defined as the change in the boundary from the previous optimization stage (where $\mathbf{mb}_0$ is the previous boundary position):

$$G(\mathbf{M}) = -\int_{\mathscr{B}(\mathbf{M})} \frac{\|\mathbf{mb}(s) - \mathbf{mb}_0(s)\|^2}{r^2}\, \mathbf{ds}$$

With a statistical shape model that defines global deformations by a linear combination of major modes of variation, such as the PCA model in the active shape model or the PGA modes presented in Section 5, the term above can be supplemented by the standard Mahalanobis distance (Mahalanobis, 1936) as is done in the bootstrap framework presented in Section 6.

4.3. Objective function: model regularization

The model regularization term, R, corresponds in spirit to the curvature and neighbor distance terms that are present in the original active contour model (Kass et al., 1988) as well as contemporary implementations of shape model building frameworks based on the active shape model (Thodberg and Olafsdottir, 2003; Thodberg, 2003; Heimann et al., 2006). The term is added in order to keep the model "nice" and thereby also ensures correspondence.

Model regularity is essential for ensuring proper correspondence across model instances. Ideally, the model regularization does not affect the geometry of the model – it just dictates the choice of model parameterization among the manifold of possible models that can represent a given shape. However, due to sampling and implementation details, the model regularization will affect the resulting model geometry marginally.

For a medial atom $\mathbf{m}$ with neighbor atoms $\mathbf{m}_i$ (see Fig. 5), we define the model regularization term:

$$\begin{aligned} R(\mathbf{m}) = {} & \gamma_{\text{dist}} \frac{1}{8} \sum_{j=1}^{8} rd\left[d(\mathbf{m}, \mathbf{m}_j), \overline{d(\cdot,\cdot)}\right]^2 \\ & + \gamma_{\text{curv}} \frac{1}{4} \sum_{i=1}^{4} angle(\mathbf{m} - \mathbf{m}_i, \mathbf{m} - \mathbf{m}_{-i})^2 \\ & + \gamma_{\text{orient}} \left(angle(\mathbf{b}, \mathbf{b}_{\text{ideal}})^2 + \frac{1}{4} \sum_{i=1}^{4} |angle(\mathbf{m}_i - \mathbf{m}_{-i}, \mathbf{n}) \right. \\ & \left. -\pi/2|^2 \right) + \gamma_{\text{boundary}} \int_{\mathscr{B}(\mathbf{M})} \frac{1}{8} \sum_{j=1}^{8} 1 - \mathbf{mbn}(\mathbf{s}) \cdot \mathbf{mbn}_j(\mathbf{s}) \end{aligned}$$

The first term penalizes unevenly spaced atoms weighted by the constant γ_{dist}. Here $d(\cdot,\cdot)$ is the distance between atoms, rd is the relative difference, and $\overline{d(\cdot,\cdot)}$ is the mean vertical/horizontal/diagonal distance between neighboring atoms for the model.

The second term penalizes high curvature and non-regular angles in the atom grid weighted by γ_{curv}. The atom $\mathbf{m}_{-i}$ is the neighbor opposing $\mathbf{m}_i$ (as illustrated in Fig. 5).

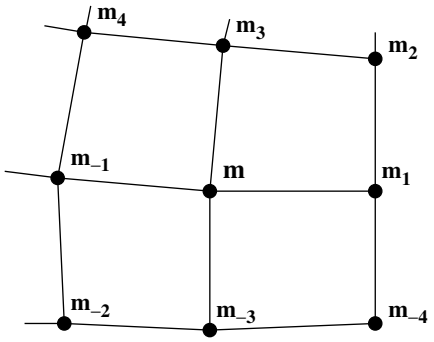

Fig. 5. The central medial atom, $\mathbf{m}$, has eight neighbors numbered $i = 1 \ldots 4$ with opposing neighbors $i = -1 \ldots -4$. When the relative neighbor positions are not important, we number the neighboring atoms $j = 1 \ldots 8$.

The term weighted by γ_{orient} penalizes non-regular orientation of the atom coordinate frame. The first part compares the $\mathbf{b}$ direction with $\mathbf{b}_{\text{ideal}}$ which is a weighted sum of the horizontal and vertical vectors between opposing neighbors across the atom. The weights are given by a linear combination of the atom coordinates scaled to between 0 and 1 such that orientation should preferably be straight away from the coordinate center (illustrated in the online color version in Fig. 2). The second half of the term penalizes if the atom orientation normal vector $\mathbf{n}$ is not perpendicular to the atom grid.

The final term penalizes high curvature of the resulting model boundary weighted by γ_{boundary}. The curvature is approximated by the dot products between the boundary point normal and the neighboring boundary point normals. The equations above demand some special cases for medial atoms with fewer than eight neighbors at the edge of the medial sheet. We leave these special cases as implementation details.

Each of the regularization terms above are reasonable as they penalize undesirable shape model behavior. However, the specific formulation and combination of the terms could be rephrased in a more theoretical framework. The formulation above has the advantage that each term is simple, fast to compute, and as the results section shows, it does actually ensure model regularity.

4.4. Automatic fitting of shape model

Optimization of the objective function above now allows deformation of an m-rep model such that it represents a given shape. The optimization function is optimized using a multi-scale conjugate gradient method. The multi-scale element consists of optimizing with the gradient computed at successively smaller scales in order to avoid getting stuck in local minima.

We start from a suitable m-rep model and then perform the following steps:

- The model is translated and scaled to the center of mass and volume of the shape (computed using the corresponding binary volume).
- The model is optimized at the figural level. The allowed transformations are a global similarity transform plus an elongation of the entire figure.[1]
- The model is optimized at the atom level. Each medial atom is independently transformed by a similarity transform plus a rotation of the object angle.

We apply this model fitting method using the starting model shown in Fig. 2 to represent the shapes in the collec-

[1] This choice of allowed transformation is natural in the m-rep figural framework – however, in a general shape representation it would be natural to allow anisotropic scaling of the shape in all directions. This loss of a degree of freedom will then has to be compensated at the deformations at the atom level.

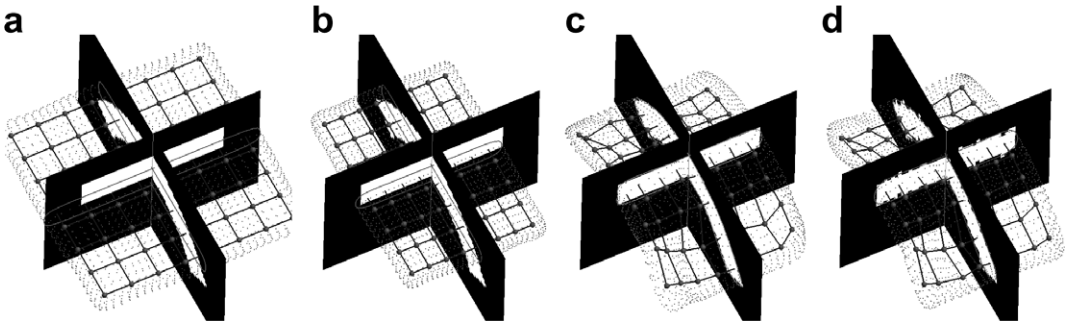

Fig. 6. The result of the automatic fitting of the initial model from Fig. 2 to the basic and noisy versions of the fourth shape from the artificial collection in Fig. 1 (it is here rotated compared to Fig. 1). (a) The starting model after translation and scaling. (b) After optimization at the figural level. (c) After optimization at the atom level against the basic version. (d) After optimization at the atom level against the noisy version.

tion of artificial, flat shapes. For this we use the following parameters: $\alpha = 100$ in the figure stage and $\alpha = 0$ in the atom stage, $\theta_{\text{crawl}} = 0.4$, $\gamma_{\text{dist}} = 0.000015$, $\gamma_{\text{curv}} = \gamma_{\text{orient}} = 0.0003$, and $\gamma_{\text{boundary}} = 0.003$.

The choice of parameters is somewhat arbitrary and the result of experiments showing the suitable order of magnitude rather than specific values. The exception is the choice of $\alpha = 100$ in the figure stage. For building a shape model (as opposed to performing segmentation using an existing shape model) geometric typicality does not really make sense so $\alpha = 0$ would be the natural choice. However, for shapes with no natural orientation, such as the potato-shaped prostates introduced below, the figure stage optimization can result in an arbitrarily oriented model. For correspondence considerations a stable model orientation is desirable and therefore $\alpha = 100$ is used in the figure stage even though it is not significant for the collection of flat shapes.

The resulting deformed models are illustrated in Fig. 6. The mean DSC between pairs of shapes and fitted models is 0.97 for the basic collection and 0.93 for the noisy collection. This is slightly lower than the overlap between the shapes and the smoothed versions of them (1.00 and 0.95 as stated in Section 2). There are a few causes for this. First, the choice of smoothing scale is somewhat fiducial – chosen such that the smoothing of the noisy shapes estimates the smooth basic shapes well. Second, Gaussian smoothing is not explicitly modeling the implicit smoothing performed by the shape representation. However, in general the shape model samples the medial axis sparsely and thereby trades attainable precision for compactness in representation.

Thereby, a collection of shapes can be represented automatically by the model optimization method. In order to build a statistical shape model, the distribution of shapes needs to be determined from the resulting collection of models. This is done below.

5. Principal geodesic analysis

Principal geodesic analysis (PGA) (Fletcher et al., 2003b) is a generalization of principal component analysis (PCA) to curved manifolds. We briefly review the results here.

As shown in Fletcher et al. (2003b), the set of all medial atoms forms a Lie group $M = \mathbb{R}^3 \times \mathbb{R}^+ \times \mathbf{SO}(3) \times \mathbf{SO}(2)$, which we call the *medial group*. Likewise, the set of all m-rep models containing n medial atoms forms a Lie group M^n, i.e., the direct product of n copies of M. This allows the definition of the exponential and logarithmic maps, $\exp(\cdot)$ and $\log(\cdot)$, that define the geodesics of the medial group.

5.1. M-rep means and PGA

The Riemannian distance between m-rep models $\mathbf{M}_1$, $\mathbf{M}_2 \in M^n$ is given by $d(\mathbf{M}_1, \mathbf{M}_2) = \| \log(\mathbf{M}_1^{-1}\mathbf{M}_2) \|$. Thus, the intrinsic mean of a set of m-rep models $\mathbf{M}_1, \ldots, \mathbf{M}_N$ is the minimizer of the sum-of-squared geodesic distances: $\mu = \arg\min_{\mathbf{M} \in M^n} \sum_{i=1}^{n} \| \log(\mathbf{M}_i^{-1}\mathbf{M}) \|^2$.

Principal components of Gaussian data in $\mathbb{R}^n$ are defined as the projection onto the linear subspace through the mean spanned by the eigenvectors of the covariance matrix. If we consider a general manifold, the counterpart of a line is a geodesic curve.

As shown in Fletcher et al. (2003a), the covariance structure of a Gaussian distribution on M^n may be approximated by a covariance matrix $\mathbf{\Sigma}$ in the Lie algebra $\mathfrak{m}^n$. The eigenvectors of this covariance matrix correspond via the exponential map to geodesics on M^n, called *principal geodesics*.

Analogous to linear PCA models, we may choose a subset of the principal directions $\mathbf{u}^{(k)} \in \mathfrak{m}^n$ with corresponding variations λ_k that is sufficient to describe the variability of the m-rep shape space. New m-rep models may be generated within this subspace of typical objects. Given a set of coefficients $\{\alpha_1, \ldots, \alpha_l\}$, we generate a new m-rep model by $\mathbf{M} = \mu \exp\left(\sum_{k=1}^{l} \alpha_k \mathbf{u}^{(k)}\right)$, where α_k is chosen to be within $[-3\sqrt{\lambda_k}, 3\sqrt{\lambda_k}]$.[2]

[2] Alternatively, the limits on the coefficients α_k could be defined in terms of the Mahalanobis distance which would be particularly appropriate for higher dimensions.

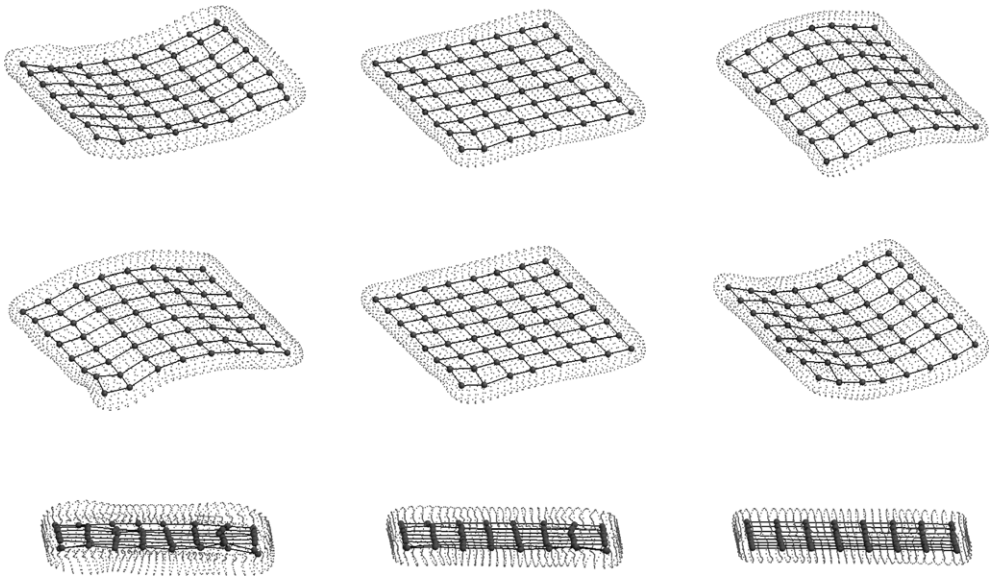

Fig. 7. Montage with the first three modes of variation for the noisy collection. The center column shows the mean shape. The first row shows first mode of variation that captures both bending and the width/length ratio (note that the correlation between bending and width/length ratio is present in the training data – similar to the variation between fourth and fifth shape in Fig. 1). Second mode captures mainly bending. The third mode captures thickness almost exclusively.

Algorithms for computing the m-rep mean and the principal geodesic analysis on a population of m-rep figures are given in Fletcher et al. (2003a) and revised in Fletcher and Joshi (2004).

The resulting mean shape and the first three modes of variation for the noisy version of the artificial collection are illustrated in Fig. 7. The first two modes capture bending and width/length ratio. The third mode captures thickness. These three modes capture 99% of the total variation.

6. Shape model bootstrap

The building blocks for the shape model building framework are now ready. Using a fiducial starting model, we use the model optimization method from Section 4 to create representations of each shape in the desired training collection. Principal geodesic analysis is used to extract a mean and modes of deformation from these shape representations as described in Section 5.

In principle, this could be it. However, if the starting model requires much deformation in order to fit the shapes, the correspondence from shape to shape is potentially ruined to some degree. And since the fiducial starting model is by definition not taylor-made to fit the shapes, this is a real risk. Furthermore, the resulting shape mean could have a bias towards the starting model.

Therefore, it makes sense to redo the shape representation optimization step again – now starting from the derived shape mean instead of the fiducial starting model. And in the natural generalization, to iterate this process until the derived shape mean and modes of variation no longer change.

During these iterations, the shape model modes of variation are added to the shape transformations. This is done during the figure level deformation where the allowed transformations are then a linear combination of a similarity transformation, a scaling, and the PGA modes. Thereby, the Mahalanobis distance is added to the geometric typicality term in Section 4.2 (using the notation from Section 5.1, the Mahalanobis distance is $\sum_{k=1}^{l} \frac{\alpha_k}{\gamma_k}$).

Possibly, the variation in the atom level deformations could also be trained during the bootstrap iterations and included in the allowed deformations. The reasons for not doing this are firstly that we wish to model the overlap shape and its variation which is naturally captured in global modes of deformations. Second, anatomical correspondence is likely better preserved with global deformations than independent, local deformations. Finally, focusing the majority of the deformation in the global, figural level makes the optimization faster.

The overall idea is that during the bootstrap iterations, the shape mean will converge, and that the global modes of variation will capture most of the necessary deformation such that only little deformation is needed at the atom level. This ensures both fast optimization and good correspondence in the optimized models.

6.1. Convergence

The PGA shape model building bootstrap iterations should run until the resulting shape mean and modes of deformation converge. However, a mathematical proof of convergence of the process seems unattainable – or at the very least highly complicated giving requirements on the

regularization parameters based on properties of the collection of training shapes.

In order to ensure well-behaved evolution of the mean and modes of variation we therefore propose a simple scheme inspired by simulated annealing (Kirkpatrick et al., 1983). By adding an increasing penalty for change of the model representing each individual shape between bootstrap iterations, we explicitly dampen the evolution and enforce convergence.

From the second iteration, we add the following annealing term, A, to the objective function in Eq. (1) for the model M in bootstrap iteration k:

$$A(M^k) = \psi 2^{(k-2)} \left(\sum_{n=1}^{N} \|\mathbf{x}_n^k - \mathbf{x}_n^{k-1}\| + \int_{\mathscr{B}(\mathbf{M})} \|\mathbf{mb}(\mathbf{s})^k - \mathbf{mb}(\mathbf{s})^{k-1}\| \, \mathrm{d}s \right)$$

The first term is the sum of squared distances between atom positions compared to the model for the same shape in the previous bootstrap iteration. Analogously, the second term is the sum of squared distances between corresponding boundary positions compared to the previous bootstrap iteration. The parameter ψ determines the overall magnitude of the penalty – which is simply doubled each iteration in order to ensure relatively quick convergence.

6.2. Resulting shape model properties

We demonstrate the full shape model building framework on the collection of basic artificial shapes from Section 2. The method is run using the starting model shown in Fig. 2, the parameters given in Section 4.4, and $\psi = 0.001$. The effects of the iterations in the bootstrap method are illustrated by the graphs in Figs. 8 and 9.

Fig. 8 focuses on the accuracy of the shape representation by evaluating how well the models fit the shapes in the training collection. In our framework it is obvious that representation accuracy needs to be evaluated. It is often assumed unnecessary to evaluate (as in Styner and Gerig, 2001), however, optimization of correspondence properties alone can easily lead to flawed representation accuracy as demonstrated in Heimann et al. (2006). We measure the fit both as the boundary distance from the image match term L in Eq. (1) and as DSC.

Fig. 9 (left) illustrates the convergence of the bootstrap iterations by measuring mean distances between the corresponding points in each shape from the previous iteration. Fig. 9 (right) illustrates the improved correspondence during the bootstrap iterations as measured by the model com-

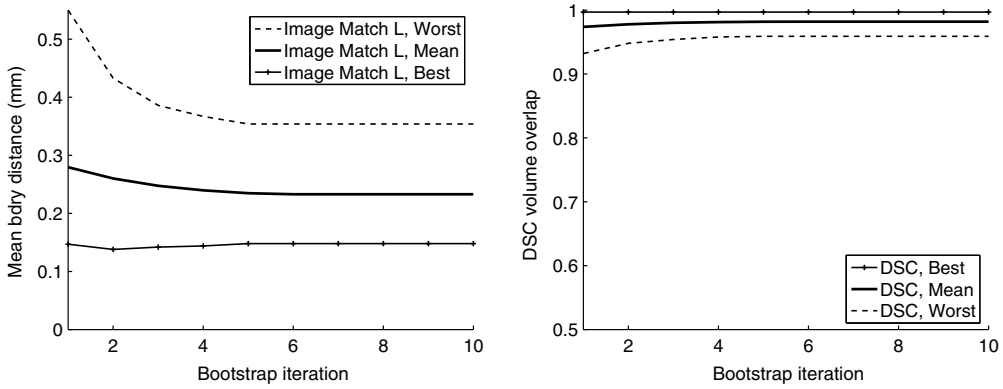

Fig. 8. Accuracy of the shape representation. Left: the fit of the models to the shapes as a function of bootstrap iteration measured as the mean boundary distance between the shapes and the fitted models. Right: the fit measured as DSC volume overlap.

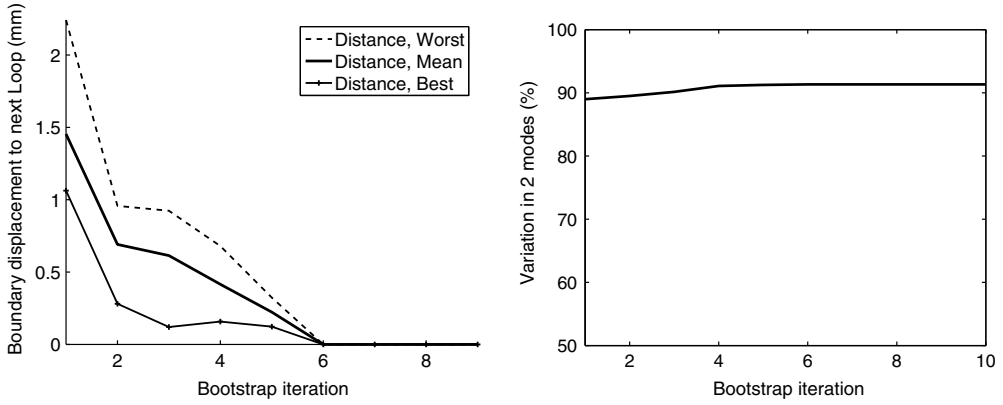

Fig. 9. Convergence and correspondence properties of the method. Left: the mean distance between corresponding boundary points are measured between iterations of the bootstrap. Right: the compactness of the model.

pactness (defined as the accumulated variation captured by a fixed number of modes of variation).

These figures illustrate that as the model building iterations progress, the resulting shape model represents the training shapes slightly more accurately while improving the correspondence at the same time. The correspondence properties of the resulting shape models are evaluated more thoroughly for the prostate and cartilage collections in Section 9.2.

The effects of the model building iterations are not spectacular on this collection of shapes where the model performance only improves little after the first iteration. This is simply because the artificial shapes are quite simple and therefore relatively easy to capture without need for optimizing the shape model through the iterations. Accordingly, the resulting mean model and modes of variation after the 10 iterations (not illustrated) are similar to those illustrated in Fig. 7.

6.3. Basic vs. noisy model

As stated in Section 2, due to the aim of capturing "overall shape", the models resulting from the basic and noisy versions of the artificial collection should be virtually identical. Fig. 10 illustrates that this is actually the case. We measure the difference between the basic and noisy model with varying degree of deformation along the main modes of variation (up to 3 standard deviations in each mode) as captured by the shape model building method. The difference between pairs of deformed basic and noisy models is measured as the mean Euclidean distance between corresponding boundary and medial points in the two shape representations. By measuring distance between corresponding points rather than just between nearest boundary/medial points, the distance also captures the degree of correspondence between the two models.

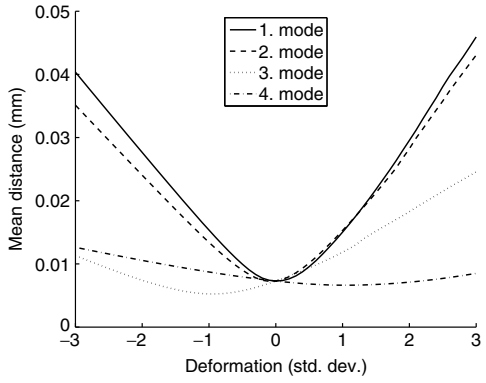

Fig. 10. Comparison of the statistical shape models captured from the noisy and basic collections of artificial shapes. For each mode of variation, the two resulting shapes for the two models are compared by computing the mean distance between corresponding points. Even for large deformations along the main modes, the shapes produced by the two models differ minimally compared to the voxel size (1 mm).

In the following sections, we introduce the collections of real prostate and cartilage sheet shapes. These structures have less obvious shape characteristics and are more challenging with respect to shape model building.

7. The UNC pelvis CT collection

Prostate cancer is a common type of cancer among males from age 50. In order to perform radiation treatment, the prostate needs to be segmented. This is particularly challenging when done from CT due to low contrast with neighboring tissue – a statistical shape model is therefore useful for the segmentation. Furthermore, a shape model can be used for analyzing the optimal radiation beam configuration.

The pelvis scan collection was retrospectively selected from the archives of UNC Health-care (Chapel Hill, NC, USA) and Western Wake Radiology (Cary, NC, USA). The scans were acquired from Siemens Somotom 4+ CT scanners without contrast agents of resolution $0.01 \times 0.01 \times 0.4\,\text{mm}^3$ or $0.015 \times 0.015 \times 0.3\,\text{mm}^3$. The collection has 46 scans in total – all manually segmented by trained radiologists by slice-wise outlining in the axial plane using *MASK* (Tracton et al., 1994) and *anastruct_editor*, from the PLan-UNC suite developed at UNC–CH Radiation Oncology. For more detail on the pelvis scan collection, see Dam et al. (2004).

The collection has manual segmentations for prostate, bladder, and rectum (see Fig. 11). All cases are diagnosed

Fig. 11. Sagittal slices of the manual segmentations of rectum, prostate (brightest), and bladder (darkest) from two cases in the UNC pelvis collection. Notice the variation in the shape of the segmented prostates as well as the jagged profiles typical of manual slice-wise segmentations when viewing them from another orientation than the one they were segmentation from. The volumes vary from 12 cm^3 to 144 cm^3.

with prostate cancer so the resulting shape model will not necessarily model prostates in general. For instance, an increase of the size of the prostate is common for prostate cancer patients. Since the shape model is to be used for analysis of patients diagnosed with prostate cancer, this bias towards cancerous prostates is desirable.

Using the estimate from Section 2 of the attainable volume overlap between the prostate shapes and a smooth shape representation, we get a highest attainable mean DSC overlap of 0.92. The lowering of the attainable overlap is caused by the jagged profiles in the shapes due to slice-wise manual outlining.

8. The CCBR knee MRI collection

Osteoarthritis is a common degenerative joint disease that affects the majority of the elderly population resulting in pain and reduced range of motion. The central process defining disease progress is breakdown of articular cartilage. Quantitative analysis of articular cartilage is crucial for clinical studies aimed at developing treatments beyond pain relief. A cartilage shape model is useful in cartilage segmentation and needed to provide the coordinate system for analysis of focal measures such as thickness.

We prospectively acquired the knee scans on an Esaote C-Span low-field 0.18 T scanner dedicated to imaging of extremities using a Turbo 3D T1 sequence (flip angle 40°, TR 50 ms, TE 16 ms). The scans are made through the sagittal plane with a voxel size in the range $0.7031 \times 0.7031 \times (0.7031/0.7813/0.8594)$ mm^3. Approximate scan time is 10 min. From a larger collection, 25 scans were randomly selected for training the shape model. The articular cartilage was manually segmented by a radiologist by slice-wise-outlining (see Fig. 12) using a dedicated annotation tool.

The scans include both left and right knees – right knee scans were reflected in order to apply the same methodology to all scans. The test subjects were males and females of ages between 21 and 72 years with varying degrees of osteoarthritis (scores 0–3 on the Kellgren and Lawrence scale (Kellgren and Lawrence, 1957)).

Using the estimate from Section 2 of the attainable volume overlap between the cartilage sheets and a smooth shape representation, we get a highest attainable mean DSC overlap of 0.86. The relatively low attainable overlap is due to firstly slice-wise manual outlining, but primarily the fact that the shapes are quite thin, and therefore small boundary displacements account for a relatively large change in volume (overlap).

9. Evaluation

The shape model building framework that we present here is almost entirely automatic. The exception is the choice of starting model. For all the shape collections we model here, we start from a generic, flat, rectangular model – however, the grid dimensions differ. The collection of basic/noisy flat shapes are modeled using an 8×8 model, the knee cartilage by an 8×4 model more obvious for elongated shapes, and the prostates by a 4×4 model. The prostates could also be modeled by an 8×8 model – the specific choice is an informal choice of representation compactness against accuracy. Alternatively, the choice could be automated like in Styner and Gerig (2001) simply by trying different dimensions and picking the one with best performance.

The parameters for the optimization process are the same for all collections (listed in Sections 4.4 and 6.1).

In the evaluation, we first focus on the evolution during the model building bootstrap iterations. Then we illustrate the resulting shape models for the collections of real prostate and cartilage shapes.

9.1. Evolution of accuracy

We first evaluate the evolution of the resulting shape representations as the shape model building iterations progress.

Like in Section 6.2 we track the accuracy of the shape representations in terms of mean boundary distance (L in Eq. (1)) (Fig. 13, left) as well as DSC volume overlap (Fig. 13, right) between shapes in the collection and the optimized model. As the graphs show, there are slight improvements in representation accuracy as the bootstrap iterations evolve.

The convergence of the shape model building process is also illustrated (Fig. 13, bottom) and for three of the collections, the method converges within the 10 iterations. For the prostate collection, there is still some minor evolution.

Fig. 12. Left: sagittal slice from an MR scan. Center: zoom of the same slice. Right: the corresponding golden standard segmentation resulting from the radiologist outlining in that slice.

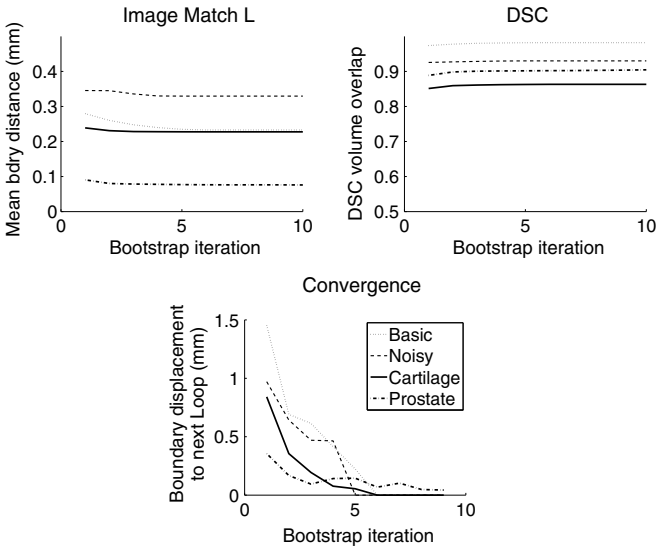

Fig. 13. Evolution of the mean boundary distance (left) and the DSC volume overlap (center) as function of the bootstrap iteration. For all shape collections, the accuracy of the representations improve slightly. The process converges for three collections, while the prostate model is still evolving slightly (right).

9.2. Evolution of correspondence

Evaluation of the evolution of correspondence properties is illustrated in Fig. 14. Inspired by Styner et al. (2003), we measure generality, specificity and compactness.

Generality measures the ability to recognize unknown examples from the same class of shapes. Specificity is the ability not to allow shapes that are far from the shapes in the training collection – meaning that only plausible shapes are statistically probable in the resulting shape distribution. To some degree, generality and specificity are opposing requirements. Finally, compactness is the ability to model the shape distribution with few, powerful parameters.

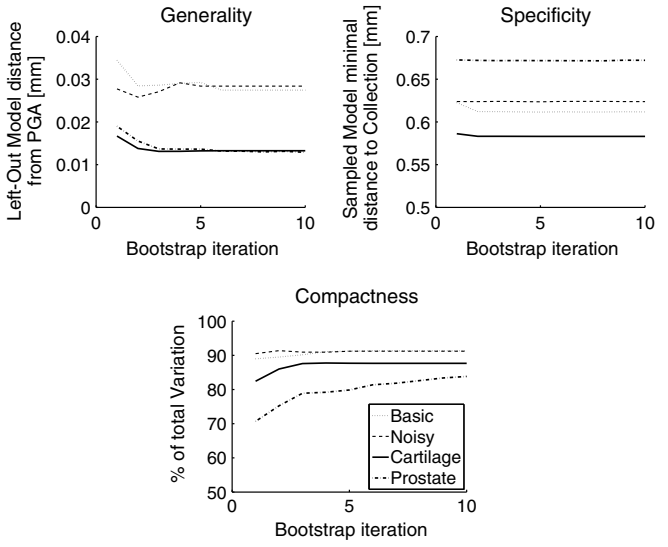

Fig. 14. Evolution of correspondence properties during the shape model building iterations. We evaluate generality, specificity, and compactness for the four collections of shapes. In general, generality and compactness improve during the iterations while specificity remains constant.

Generality is approximated in a leave-one-out fashion by measuring the minimal distance from each shape instance to the PGA shape space spanned by the remaining shapes in the collection. Specificity is evaluated by sampling 1000 random shapes from the PGA shape space (we sample the shape parameters from a multi-variate Gaussian since for relatively small deformations this is a close approximation of the true shape space) and measuring the minimal distance to one of the shapes from the training collection.

For both generality and specificity, distance between two shapes is measured as the mean Euclidean distance between corresponding boundary and medial points in the two shape representations generated following Section 4.4. By measuring distance between corresponding points rather than just between nearest boundary points, the evaluation of generality and specificity also indirectly evaluates the correspondence ability of the model.

Compactness is measured as the relative variance covered by a fixed number of shape variation modes. For the artificial shapes, we use two modes, and for the real shapes we use 10 modes.

The overall picture is that generality and compactness are improved during the shape model building iterations while specificity is virtually constant. Thereby, the framework manages to improve the overall correspondence of the shape models.

9.3. The resulting shape models

We illustrate the resulting shape models for the prostate and cartilage sheet collections with the standard displays of the mean shape accompanied with shapes at a few standard deviations away along the major modes of variation.

As is typically the case, the modes capture both shape variation which is difficult to describe in simple terms as well as modes with nice, intuitive anatomical foundation. Examples of the latter is the third mode for the prostate model (in Fig. 15) which is most likely related to deformation due to pressure from the bladder.

The third mode in the model for the tibial medial cartilage sheet (Fig. 16) also has an interesting anatomical relation – it captures the overall bending of the sheet which in previous work has been shown to have a close relation to the presence of osteoarthritis (Folkesson et al., 2006).

9.4. Focal statistics

The coordinate system of the resulting shape model allows focal statistics across a population of shapes. We illustrate this by the map in Fig. 17 of the focal variation in cartilage thickness for the collection of tibial cartilage

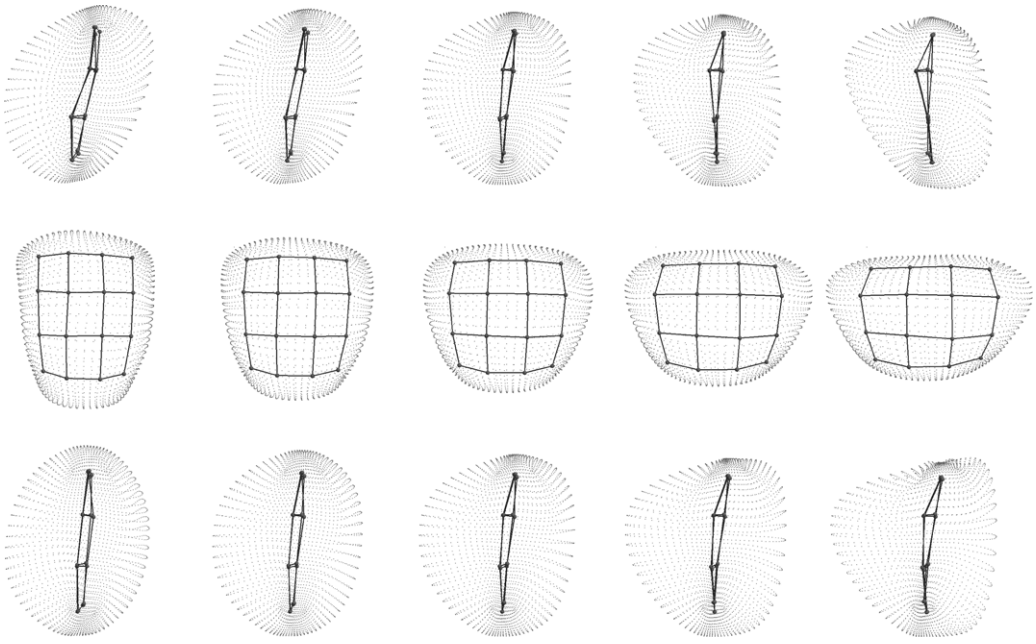

Fig. 15. Prostate shape model resulting from the shape model building method. The rows illustrate first, second, and third mode of variation. Each shape mode is shown with the mean shape in the center and then ± one and two standard deviations away from the mean. The first mode (shown from sagittal view) can be characterized as skewness of the prostate. The second mode (from coronal view) is a Laurel/Hardy variation. The third mode (from sagittal view) could be related to the location of the bladder which pushes down on the prostate from top left.

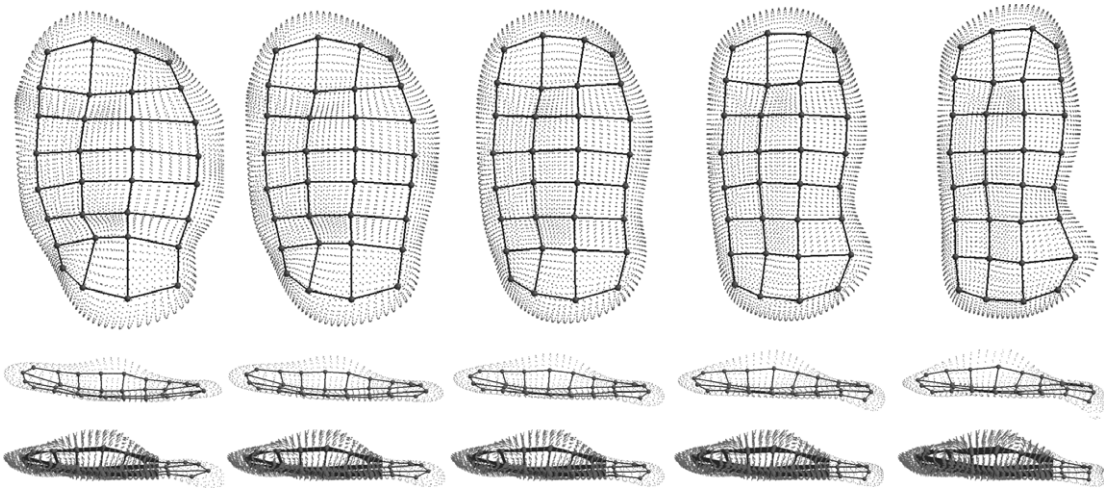

Fig. 16. Cartilage sheet shape model resulting from the shape model building method. The first mode (shown from axial view) is again a Laurel/Hardy variation. The second mode (from sagittal view) mainly captures curving of the posterior part of the sheet. The third mode (from sagittal view) captures the overall bending of the cartilage sheet.

sheets. The thickness variation is focused in the central valley of the cartilage sheet and up towards the intercondylar spine – anatomically corresponding to the load-bearing region of the cartilage sheet.

9.5. Results summary

In Table 1, we summarize the key performance measures for the shape models generated above for the four collections of training shapes. The number of training shapes is given parenthesized for each collection.

First, we show the accuracy of the optimized shape representations as the resulting fit vs. the training collection shapes. This is measured both in terms of mean boundary distance and as mean DSC overlap. For DSC, both values for the estimate for the attainable value from Section 2 and the performance of the optimized shape models are listed.

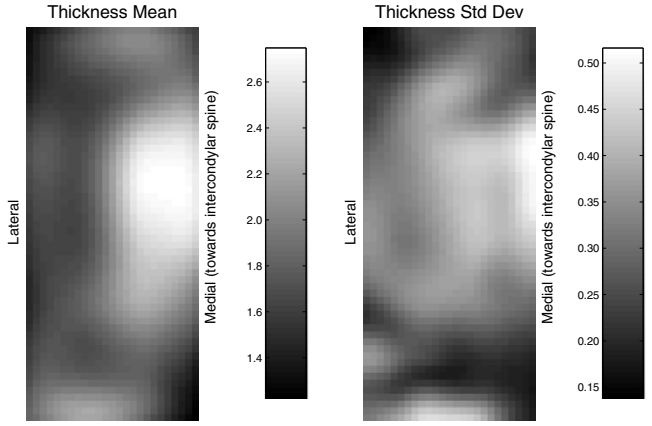

Fig. 17. Focal variation for the tibial cartilage sheets. The maps of the local mean (left) and standard deviation (right) across the collection of cartilage sheets show that the largest thickness as well as variation is primarily located in the load-bearing region of the cartilage sheet.

Table 1
Summary of results

Collection Size	Basic 5	Noisy 5	Prostate 46	Cartilage 25
Accuracy				
Voxel size (mm)	1.0	1.0	0.13	0.8
Boundary distance (mm)	0.23	0.33	0.08	0.23
DSC, attainable	1.00	0.95	0.92	0.86
DSC, evaluation	0.98	0.93	0.90	0.86
Correspondence				
Generality (mm)	0.027	0.028	0.013	0.013
Specificity (mm)	0.61	0.62	0.67	0.58
Compactness (%)	91	91	84	88

First, for each collection the representation accuracy is given as mean boundary distance as well as volume overlap (DSC) between shapes and models. The boundary distance can be compared by the mean voxel side length. The DSC values compare the "attainable accuracy" for the shape collection to performance of the models. Second, the correspondence properties for each collection are given.

Second, we show the correspondence property evaluation values for generality, specificity, and compactness.

10. Conclusion

We present a framework for automatic shape model building from a collection of training shapes. The method consists of alternating steps of deforming a shape representation into the training shapes following by calculation of the shape mean and the modes of shape variation. The mean and modes are then used as the starting point in the next iteration of the shape model building method.

The basic framework is heavily inspired by our previous work (Dam et al., 2004). However, in particular the optimization based on boundary distance ensures that the new method represents the training shapes close to the attainable accuracy. For the prostate collection, this means a significant improvement from a mean DSC volume overlap of 0.80–0.90 (compared to the attainable accuracy of 0.92).

We evaluate our shape model building method on four collections of training shapes, among these prostates and cartilage sheets. The evaluation shows that the method is able to represent the shapes close to the attainable accuracy already after the first iteration. Through the iterations the resulting shape mean and modes of variation are optimized such that the correspondence properties are improved.

The resulting shape means and modes of deformation look intuitively pleasing and correspond well with the expectations. This is particularly clear for the collections of artificial shapes where the expected mean and modes are known by design.

Thereby, we argue that our overall framework is suitable for automatic construction of statistical shape models.

10.1. Future work

The framework presented above allows construction of a statistical shape model from the a collection of training shapes. The results show that the method is good for representing the shapes and for extracting reasonable shape means and modes of variation.

The evaluation of the correspondence properties gives intuitively pleasing results. The collection that proves to be the most challenging is the collection of prostates. This is revealed by the slower rate of convergence combined with the slightly worse values for correspondence properties. This is to be expected since the prostate shapes have little natural correspondence defined by their geometry. For such shapes the correspondence should probably rather be defined by the surrounding organs. Since our shape representation of choice has been used extensively for multi-object models, this is a natural, future extension of our method. Thereby, our method will also be extended to allow more complicated shapes.

The intuitive look of the resulting shape mean and modes of variations indicates that the framework is reasonably well behaved. Also, the evaluation of both the representation accuracy and as well as the correspondence properties shows good performance of the framework. However, since the method has been evaluated on different data sets than the sets leading, alternative methods were evaluated against in Styner et al. (2003) or Heimann et al. (2005), it is very hard to conclude how our method performs compared to these methods.

The devil's advocate could argue that our method is based on a naive confidence in the robustness of the correspondence properties of a medial model and that we therefore focus too much on geometric regularization and representation accuracy instead of directly optimizing the correspondence properties as done in Davies et al. (2002). Therefore, we look forward to performing a proper evaluation against the main competing methods – inspired by the excellent study in Styner et al. (2003).

This future will be focused on evaluation the ability to model actual anatomical correspondence rather than mathematical properties related to correspondence. Only such a direct evaluation will reveal whether the ideas underlying our framework are sound and robust.

Acknowledgments

We sincerely thank Per Halverson at Western Wake Radiology (Cary, NC, USA) for supplying and annotating pelvis scans. Julian Rosenman and Gregg Tracton at Radiation Oncology, University of North Carolina (Chapel Hill, NC, USA) for supplying and annotating pelvis scans. Paola Pettersen and Claus Christiansen at the Center for Clinical and Basic Research (Ballerup, Denmark) for supplying and annotating knee MRI. Our colleagues at UNC and University of Copenhagen for discussions and valuable input – in particular Sarang Joshi, Gregg Tracton, Mads Nielsen, Ole Fogh Olsen, and Marleen de Bruijne. Corné Hoogendoorn, Pompeu Fabra University, Barcelona, for proofreading. Parts of the work was carried out while Erik B. Dam was at the IT University of Copenhagen.

References

Besl, P.J., McKay, N., 1992. A method for registration of 3-d shapes. IEEE PAMI 14 (2).

Blum, H., Nagel, R., 1978. Shape description using weighted symmetric axis features. Pattern Recognition 10 (3), 167–180.

Brett, A., Taylor, C., 2000. A method of automated landmark generation for automated 3d pdm construction. Image and Vision Computing 18.

Charpiat, G., Faugeras, O., Keriven, R., 2005. Approximations of shape metrics and application to shape warping and empirical shape statistics. Foundations of Computational Mathematics 58 (1).

Cootes, T., 2005. Timeline of developments in algorithms for finding correspondences across sets of shapes and images. Available from: <http://www.isbe.man.ac.uk/~bim/Papers/correspondence_timeline.pdf>.

Cootes, T., Taylor, C., Cooper, D., Graham, J., 1995. Active shape models: their training and application. CVIU (1), 38–59.

Cremers, D., Kohlberger, T., Schnörr, C., 2002. Nonlinear shape statistics in mumford-shah based segmentation. In: 7th European Conference on Computer Vision, vol. 2351 of LNCS, Springer, Berlin.

Dam, E.B., Fletcher, P.T., Pizer, D.S.M., Tracton, G., Rosenman, D.J., 2004. Prostate shape modeling based on principal geodesic analysis bootstrapping. In: Proceedings of MICCAI 2004, vols. 3216–3217 of LNCS, Springer, Berlin.

Danielsson, P.E., 1980. Euclidean distance mapping. Computer Graphics and Image Processing 14.

Davies, R., Twining, C., Cootes, T., Waterton, J., Taylor, C., 2002. A minimum description length approach to statistical shape modeling. IEEE Transactions on Medical Imaging 21 (5).

Dice, L., 1945. Measures of the amount of ecologic association between species. Ecology 26.

Fletcher, P., Joshi, S., 2004. Principal geodesic analysis on symmetric spaces: statistics of diffusion tensors. In: MMBIA, vol. 3117 of LNCS, Springer, Berlin.

Fletcher, P.T., Joshi, S., Lu, C., Pizer, S.M., 2003a. Gaussian distributions on Lie groups and their application to statistical shape analysis. In: Information Processing in Medical Imaging, vol. 2732 of LNCS, Springer, Berlin.

Fletcher, P.T., Lu, C., Joshi, S., 2003b. Statistics of shape via principal geodesic analysis on Lie groups. In: Proceedings of the Computer Vision and Pattern Recognition.

Folkesson, J., Dam, E.B., Olsen, O.F., Pettersen, P.C., Christiansen, C., 2006. Automatic curvature analysis of the articular cartilage surface. In: MICCAI Joint Disease Workshop, pp. 17–24. Available from: <http://www.itu.dk/image/joint>.

Heimann, T., Wolf, I., Williams, T., Meinzer, H.-P., 2005. 3D active shape models using gradient descent optimization of description length, vol. 3565 of LNCS, Springer, Berlin.

Heimann, T., Wolf, I., Meinzer, H., 2006. Optimal landmark distributions for statistical shape model construction. In: Proceedings of the SPIE Medical.

Joshi, S., Pizer, S., Fletcher, P.T., Yushkevich, P., Thall, A., Marron, J.S., 2002. Multiscale deformable model segmentation and statistical shape analysis using medial descriptions. Transactions on Medical Imaging 21 (5), 538–550.

Kass, M., Witkin, A., Terzopoulos, D., 1988. Snakes: active contour models. International Journal of Computer Vision, 321–331.

Kellgren, J., Lawrence, J., 1957. Radiological assessment of osteo-arthrosis. Annals of the Rheumatic Diseases 16 (4), 494–501.

Kirkpatrick, S., Gelatt, C., Vecchi, M., 1983. Optimization by simulated annealing. Science 220 (4598).

Kotcheff, A., Taylor, C., 1998. Automatic construction of eigenshape models by direct optimization. Medical Image Analysis 2 (4).

Mahalanobis, P., 1936. On the generalized distance in statistics. In: Proceedings of the National Institute of Science, vol. 12.

Pizer, S., Fritsch, D., Yushkevich, P., Johnson, V., Chaney, E., 1996. Segmentation, registration, and measurement of shape variation via image object shape. IEEE Transactions on Medical Imaging 18 (10).

Pizer, Chen, Fletcher, Fridman, Fritsch, Gash, Glotzer, Jiroutek, Joshi, Muller, Thall, Tracton, Yushkevich, Chaney, 2003. Deformable *m*-reps for 3d medical image segmentation. IJCV.

Rueckert, D., Frangi, A., Schnabel, J., 2003. Automatic construction of 3-d statistical deformation models of the brain using nonrigid registration. IEEE Transactions on Medical Imaging 22 (8), 1014–1025.

Schölkopf, B., Smola, A., Müller, K.-R., 1998. Nonlinear component analysis as a kernel eigenvalue problem. Neural Computation 10.

Shelton, C.R., 2000. Morphable surface models. International Journal of Computer Vision 38 (1).

Styner, M., Gerig, G., 2001. Medial models incorporating object variability for 3d shape analysis. In: Proceedings of the Information Processing in Medical Imaging.

Styner, M., Rajamani, K., Nolte, L.-P., Zsemlye, G., Szekely, G., Taylor, C., Davies, R., 2003. Evaluation of 3d correspondence methods for model building. In: IPMI, vol. 2732 in LNCS, Springer, Berlin.

Thall, A., 2002. Fast C^2 interpolating subdivision surfaces using iterative inversion of stationary subdivision rules. Technical report, University of North Carolina. Available from: <http://midag.cs.unc.edu/pubs/papers/Thall_TR02-001.pdf>.

Thodberg, H.H., 2003. Minimum description length shape and appearance models. In: Information Processing in Medical Imaging, vol. 2732 in LNCS, Springer, Berlin.

Thodberg, H.H., Olafsdottir, H., 2003. Adding curvature to minimum description length shape models. In: Proceedings of the British Machine Vision Conference.

Tracton, G., Chaney, E., Rosenman, J., Pizer, S., 1994. Mask: combining 2d and 3d segmentation methods to enhance functionality. In: Proceedings of Mathematical Methods in Medical Imaging III. SPIE, vol. 2299.

Twining, C., Cootes, T., Marsland, S., Petrovic, V., Schestowitz, R., Taylor, C., 2005. A unified information-theoretic approach to group-wise non-rigid registration and model building. In: Proceedings of the Information Processing in Medical Imaging, vol. 3565 of LNCS, Springer, Berlin.