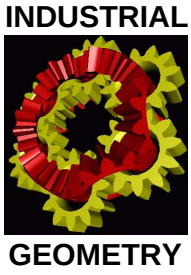
National Research Network S92

Industrial Geometry

http://www.industrial-geometry.at



NRN Report No. 92

Shape Spaces via Medial Axis Transforms for Segmentation of 3D Voxel Data

Jochen Abhau, Oswin Aichholzer, Sebastian Colutto, Bernhard Kornberger and Otmar Scherzer

January 2010

FWF
Der Wissenschaftsfonds.

universität innsbruck

Shape Spaces via Medial Axis Transforms for Segmentation of 3D Voxel Data

Jochen Abhau * Oswin Aichholzer † Sebastian Colutto ‡
Bernhard Kornberger‡ Otmar Scherzer * §

Thursday 7th January, 2010

Abstract

In this paper we construct a shape space of medial ball representations from given shape training data using methods of Computational Geometry and Statistics. The ultimate goal is to employ the shape space as prior information in supervised segmentation algorithms for 3D voxel data. The construction of the shape space, the statistics of the training data, and their subsequent use in segmentation require up-to-date, respectively novel, methods to standardize the medial ball representation of the shape space. Aside from novel methods the whole pipeline of shape space, statistics, and segmentation is novel to the best of our knowledge.

Keywords: medial ball representation,medial axis transform, image segmentation, Procrustes analysis, skin surfaces, EM algorithm

1 Introduction

Segmentation and detection of objects in image data is of major importance in several applications, such as for instance medical imaging, video surveillance and motion tracking.

Segmentation methods based on minimization of energy functionals have proven to be very efficient. In this context it is common to differ between edge based [48] and region based energy functionals [39]. Another distinctive feature of segmentation

*Computational Science Center, University of Vienna, Nordbergstraße 15, A-1090 Wien, Austria, {jochen.abhau,otmar.scherzer}@univie.ac.at

†Institute for Software Technology, Graz University of Technology, Austria, {oaich, bkorn}@ist.tugraz.at

‡Institute of Mathematics, University of Innsbruck, Technikerstraße 21a, A-6020 Innsbruck, Austria, Sebastian.Colutto@uibk.ac.at

§Radon Institute of Computational and Applied Mathematics, Austrian Academy of Sciences, Altenberger Straße 69, A-4040 Linz, Austria

algorithms is whether they perform supervised and unsupervised. Historically, unsupervised segmentation algorithms have been introduced first. However, the segmentation results are poor if the data is noisy, has little contrast, or if the object is partially occluded. To cope with these difficult situations supervised segmentation is performed, which commonly is implemented by incorporating a shape statistics term in the energy functional for unsupervised segmentation.

In the following we review some energy minimization segmentation approaches which incorporate shape prior information:
A compact representation of shapes, called *M-reps* (Medial atom representation) has been proposed in [24] and [41]. An M-rep consists of medial atoms aligned on a regular grid structure. Medial atoms consist of a coordinate position in space, a radius, a local coordinate frame, and an angle. The atoms are samples of the medial axis [9]. The actual surface associated with the M-rep is an enfolding B-spline of the medial atoms. M-reps can be regarded as elements of a Riemannian shape manifold, which defines distances between M-reps by lengths of geodesics amongst them. An M-rep shape space is suitable for modeling shapes of very similar geometry and topology. Image segmentation with M-reps has been considered in [28, 40]. Mumford-Shah energy segmentation on an M-rep space has been presented in [16].

In [18] (this work is actually in 2D) contours are represented as simple, closed B-spline curves. The shape prior and statistic are computed as the mean and the principal components of the spline control points of the training shapes. For segmentation a Mumford-Shah like energy functional is supplemented by the Mahalanobis distance to the shape prior, which is calculated from the statistics. This approach is based on the assumption that the shapes are distributed according to a multivariate Gaussian distribution. Using kernel space techniques, a generalization to non-Gaussian distributed shapes is given in [17]. Applications to medical MR image segmentation have been studied in [14].

A level set approach for segmentation with shape priors has been studied in [36]. There shapes are associated with the associated signed distance functions. On the training distance functions of the shapes the actual statistics, which is used in the segmentation algorithm, is calculated. Shape statistics are implemented in a geodesic active contour evolution as maximum a posteriori estimator of the shape. By this approach geometrically and topologically complex shape priors can be considered. However, this representation requires large computational resources, and efficient implementations with narrow bands tend to create inaccuracies and artifacts in the segmentation results. In [22], the work of [36] has been improved concerning efficiency. The same shape prior strategy as in [36] has been used in combination with the unsupervised Chan-Vese energy minimization [13] in [47]. Further generalizations of [36] are given in [43] and [44].

In [15], shape statistical prior information is incorporated in a variational segmentation functional with an additional regularization term. A generalization of this work is presented in [26]. More recently, the paper [12] proposes a variational level set framework that takes into account shape prior information which combines gradient and region based segmentation.

In this paper, we consider the construction of *medial ball shape spaces* and the computation of statistics of training shapes, which are then used for supervised energy

minimization segmentation. We are given training shapes, which are represented as triangular surface meshes. Such can either be provided by expert segmentation, or have been obtained from unsupervised segmentation (see e.g. [2, 16]). For the latter one can also imagine to use data of other imaging modalities, which have higher contrast or are less affected by noise.

The following scheme, which can as well serve as an *outline* of our work, is studied:

Scheme 1.1
Preprocessing (Shape Statistics)

- *For each mesh a discrete medial axis transform*

$$M = (x_1, \ldots, x_k; r_1, \ldots, r_k) \tag{1}$$

  is computed. We refer to $c(M) = (x_1, \ldots, x_k) \in \mathbb{R}^{3\times k}$ *as the* centers *of the medial balls of* M*. It is important for our application that each discrete medial axis transform* M_j *consists of the same number* k *of medial balls* $(x_i, r_i) \in \mathbb{R}^3 \times \mathbb{R}^+$*. The algorithm developed in here is capable to do so, which is a novel aspect in the literature. For sake of simplicity, the discrete medial axis transform computed in this step will be called* ball representation *according to its definition (Section 2).*

- *A labeling of the ball representation* M_i *is computed by an EM-algorithm and the Kuhn-Munkres algorithm (Section 3).*

- *Using Procrustes analysis, a mean ball representation is computed from the labeled ball representations, and a Mahalanobis distance between ball representations is defined. These are the building blocks of our medial ball shape space. (Section 4).*

Segmentation

- *For segmentation the shape space is implemented as shape prior information in a simplified Mumford-Shah functional. The implied surface of the ball representations is constructed by a skin surface [20] (Section 5).*

The following case examples for segmentation are considered. In Section 6 the full pipeline of statistics and segmentation is applied to medical prostate data and synthetic data, and Section 7 concludes the paper.

2 Medial Axis Transform

In this section we consider the problem of approximating a number of bounded open sets $\Omega_1, \ldots, \Omega_n \subset \mathbb{R}^3$ by sets $B_1, \ldots, B_n$ of approximate medial balls, where all sets B_i have the same cardinality. Stability of the approximation, described below, is of vital importance for the statistical analysis of medial axis representations. Our novel

approach for calculating sets of approximate medial balls of the same cardinality for a class of objects combines three well known methods. First we utilize Voronoi diagrams to approximate objects by sets of (approximate) medial balls [7, 8]. As these sets will usually have a rather huge cardinality we then use set covering methods to obtain sufficiently small and stable subsets [5, 4]. Finally, we use a k-means [29] clustering algorithm to control the cardinality of the approximating sets in order to obtain sets of approximate medial balls of uniform cardinality.

We start by reviewing some basic facts of the medial axis transform.

2.1 Discrete medial axis transform

The following definitions of the medial axis and medial axis transform are standard and can for instance be found in [10].

Definition 2.1 ([10]) *The inner* Medial Axis $\mathrm{M}(\Omega)$ *of a bounded open set* $\Omega \in \mathbb{R}^3$ *is the set of points* $x \in \Omega$ *which have at least two closest points in* $\partial\Omega$*, where* $\partial\Omega$ *denotes the boundary of* Ω*. The inner* Medial Axis Transform $\mathrm{MT}^{\mathrm{in}}(\Omega)$ *is the collection of maximal (with respect to inclusion) open balls centered at* $\mathrm{M}(\Omega)$ *and included in* Ω*.*

Note that the cardinality of the inner Medial Axis Transform is in general infinite. Moreover the medial axis $\mathrm{M}(\Omega)$ behaves unstable with respect to perturbations of high curvature of $\partial\Omega$. That is, the medial axis might contain branches that are far from being intuitive. The corresponding inner medial axis transform $\mathrm{MT}^{\mathrm{in}}(\Omega)$ expresses this unstable behavior by small balls near the surface of Ω. For these reasons we seek for an approximation of $\mathrm{MT}^{\mathrm{in}}(\Omega)$, which we define next.

Definition 2.2 ([10]) *The inner* Discrete Medial Axis Transform $\mathrm{DMT}(\Omega)^{\mathrm{in}}$ *is a finite set of open, approximate medial balls, which approximates the inner Medial Axis Transform* $\mathrm{MT}^{\mathrm{in}}(\Omega)$*.*

In the following we assume that the input object Ω is given by a set S_Ω of sample points of the boundary $\partial\Omega$ of Ω and a triangular mesh $T(S_\Omega)$, representing $\partial\Omega$. We first show how to approximate **one** three-dimensional object Ω by the union of its discrete medial axis transform $\mathrm{DMT}(\Omega)^{\mathrm{in}}$ in a stable way. We start with the well known Voronoi approach [7, 8].

Definition 2.3 *The* Voronoi Diagram *of a set of points* S_Ω *in* $\mathbb{R}^3$ *is a partition of* $\mathbb{R}^3$ *into (possibly unbounded) convex polyhedral regions, called Voronoi cells, such that each point* $s_i \in S_\Omega$ *has an associated Voronoi cell* $v(s_i)$ *with*

$$v(s_i) := \{x \in \mathbb{R}^3 : \|x - s_i\| \leq \|x - s_j\|, \forall i \neq j,\ s_i, s_j \in S_\Omega\}.$$

We first compute the Voronoi diagram of S_Ω. Then we extract for each sample point $s_i \in S_\Omega$ the inner pole point p_i, which is the vertex of the Voronoi cell $v(s_i)$ being farthest away from s_i and inside Ω. Finally, we construct for each inner pole point p_i a so called polar ball B_{p_i,ρ_i} centered at p_i with radius $\rho_i = \|s_i - p_i\|$. The set of all polar balls created in this way is the inner discrete medial axis transform $\mathrm{DMT}(\Omega)^{\mathrm{in}}$,

and S_Ω is contained in the boundary of its union [7, 8]. See Figure 1(a) for an example with more than 10000 balls.

In the original approach a dense sampling S_Ω of a smooth surface $\partial\Omega$ of Ω is required in order to be able to distinguish inner Voronoi vertices from outer ones [7, 8]. To overcome this restriction we use, as mentioned above, the triangular surface mesh $T(S_\Omega)$ as additional input. This allows us to easily distinguish between inner and outer Voronoi vertices. Thus noise - always present in real-world data sets - and poor sampling quality do not affect the correctness of our inner/outer labeling, and correct operation of this step is ensured.

The centers of $\mathrm{DMT}(\Omega)^{\mathrm{in}}$ are close to $\mathrm{M}(\Omega)$, see [6] for a quantitative analysis and a precise statement of that fact. This implies that $\mathrm{DMT}(\Omega)^{\mathrm{in}}$ might include small, surface near balls corresponding to unwanted features of $\mathrm{MT}^{\mathrm{in}}(\Omega)$. Moreover, $\mathrm{DMT}(\Omega)^{\mathrm{in}}$ approximates Ω by up to $|S_\Omega|$ balls. For our purposes the instability and the high cardinality of $\mathrm{DMT}(\Omega)^{\mathrm{in}}$ prevent its direct usage.

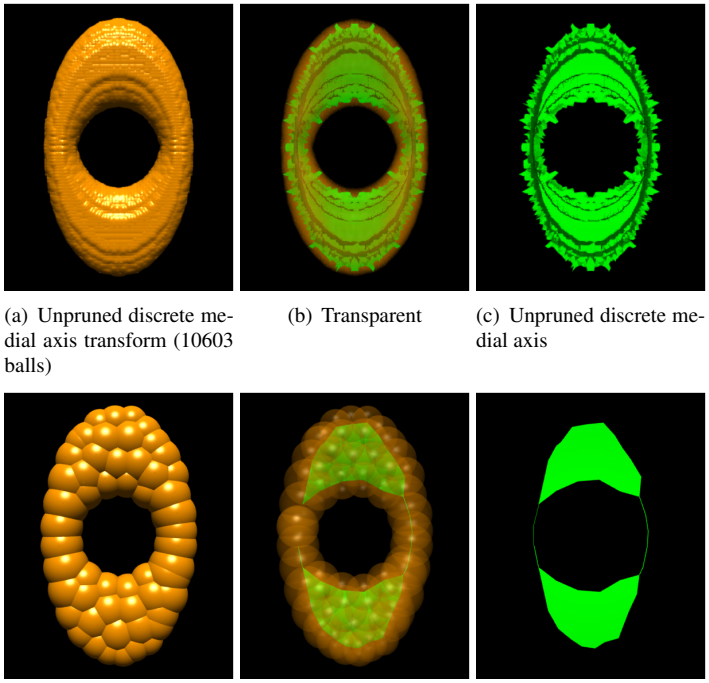
To avoid these disadvantages we apply a pruning algorithm to $\mathrm{DMT}(\Omega)^{\mathrm{in}}$ in order to extract a proper subset of $\mathrm{DMT}(\Omega)^{\mathrm{in}}$. The result will be a stable (we remove surface near balls which result from instability) and compact representation of Ω, cf. Figure 1(d) for an example. Moreover the pruning step will give us control to obtain sets of approximate medial balls of predefined cardinality, see Subsection 2.2 for details. In the following we briefly describe the pruning approach of [4], see there for further details.

By the above construction each ball in $\mathrm{DMT}(\Omega)^{\mathrm{in}}$ has four sample points on its boundary but none in its interior. We enlarge each ball of $\mathrm{DMT}(\Omega)^{\mathrm{in}}$ by a sufficiently small constant $\varepsilon > 0$. Then we use a spatial search structure to find for each enlarged ball b_i all sample points from S_Ω which are now covered by b_i (typically tens or even hundreds of sample points are contained). Finally we use a set-covering algorithm to find an (almost) minimal subset $\mathrm{DMT}(\Omega)^{\mathrm{in}}_*$ of the enlarged balls whose union covers all sample points S_Ω. This set $\mathrm{DMT}(\Omega)^{\mathrm{in}}_*$ is the output of the pruning step.

Let us stress the fact that the applied set-covering algorithm to find a minimal subset of balls will favor balls that cover a large fraction of S_Ω and thus large areas of $\partial\Omega$ in order to make the cardinality of $\mathrm{DMT}(\Omega)^{\mathrm{in}}_*$, $\left|\mathrm{DMT}(\Omega)^{\mathrm{in}}_*\right|$, as small as possible. These balls are centered near stable parts of $\mathrm{M}(\Omega)$. This implies that surface-near balls, which originate from small perturbations of high curvature of $\partial\Omega$ (recall the above discussion) are avoided. Thus stability of our approach is greatly improved by the set-covering algorithm, see also [4]. In addition the one sided Hausdorff distance from the union of the obtained compact discrete medial axis transform $\mathrm{DMT}(\Omega)^{\mathrm{in}}_*$ to the original object Ω is bounded by $O(\varepsilon)$.

2.2 Sets of approximate medial balls of uniform cardinality

As described in Section 1 our goal is to obtain sets of approximate medial balls of uniform cardinality for different input objects, Ω_i, $i = 1, \ldots, n$. So far we have shown how to compute a stable representation $\mathrm{DMT}(\Omega_{\mathrm{i}})^{\mathrm{in}}_*$ of the enlarged discrete medial axis transform for one fixed object Ω_i. We can control $|\mathrm{DMT}(\Omega_{\mathrm{i}})^{\mathrm{in}}_*|$ to some extent by



(a) Unpruned discrete medial axis transform (10603 balls)

(b) Transparent

(c) Unpruned discrete medial axis

(d) Pruned discrete medial axis transform (55 balls)

(e) Transparent

(f) Pruned discrete medial axis

Figure 1: Discrete Medial Axis Transforms and Discrete Medial Axes

the value ε by which we enlarge the polar balls.

For each single object Ω_i, the value ε is chosen such that the resulting set $|\mathrm{DMT}(\Omega_\mathrm{i})^{\mathrm{in}}_*|$ is at least as large as the desired uniform cardinality k, that is, $k \leq \min_i(|\mathrm{DMT}(\Omega_\mathrm{i})^{\mathrm{in}}_*|)$. This approach is justified by the fact, that for piecewise linear $\partial\Omega$, the quantity $|\mathrm{DMT}(\Omega_\mathrm{i})^{\mathrm{in}}_*|$ tends to infinity if $\varepsilon \to 0$. We then apply the k-means clustering algorithm to each $\mathrm{DMT}(\Omega_\mathrm{i})^{\mathrm{in}}_*$ and get for every set $\mathrm{DMT}(\Omega_\mathrm{i})^{\mathrm{in}}_*$ exactly k clusters. For each cluster we choose as an representative ball the one whose center is closest to the center of that cluster. This results in n stable representations of $\{\Omega_1, \ldots, \Omega_n\}$ by sets of balls $\{B_1, \ldots, B_n\}$ with uniform cardinality k.

For the remainder of the paper we will use the term *ball representation* to refer to the pruned inner discrete medial axis transform of uniform cardinality.

3 Labeling of the Ball Representations

In this section we use statistical algorithms for labeling of the ball representations. Given two ball representations $M = (x_1, \ldots, x_k; r_1, \ldots, r_k)$ and $\tilde{M} = (\tilde{x}_1, \ldots, \tilde{x}_k; \tilde{r}_1, \ldots, \tilde{r}_k)$, the labeling problem consists in determining a rearrangement of indices such that the two ball representations optimally match. Every matching involves a distance measure between ball representations. In view of the used definition of an Euclidean invariant shape space of ball representations considered below (see Section 4), we consider the following minimization problem for determining the optimal alinement:

$$\min_{\rho>0, A\in SO(3), b\in\mathbb{R}^3, \pi\in\mathcal{S}^k} \sum_{i=1}^{k} \|\rho A x_{\pi(i)} - b - \tilde{x}_i\|^2 + \alpha(\rho r_{\pi(i)} - \tilde{r}_i)^2 \ . \tag{2}$$

Here $SO(3)$ is the special orthogonal group of degree 3 (that is the group of rotations in $\mathbb{R}^3$, which can be represented as unitary matrices with determinant one), $\mathcal{S}^k$ is the symmetric group of degree k, and $\|\cdot\|$ denotes the Euclidean norm in $\mathbb{R}^3$. The parameter α is a weighting parameter, providing a trade-off between matching of balls and radii. Since it is difficult to compute global minimizers of (2) efficiently, we use a two-step algorithm from [33], to compute approximate minimizers. This algorithm consists of the following two steps:

- An Expectation Maximization (EM) algorithm is used to compute a similarity transformation, i.e. a translation $b \in \mathbb{R}^3$, a rotation $A \in SO(3)$, and a scaling $\rho > 0$, which is close to an optimal one in (2). The algorithm is explained in Subsection 3.1.
- Keeping the similarity transform fixed, the Kuhn-Munkres algorithm is used to compute an optimal labeling $\pi \in \mathcal{S}^k$ in (2). This algorithm is explained in Subsection 3.2.

3.1 EM-algorithm to compute the similarity transformation

EM-algorithms have been used successfully for solving a wide range of labeling problems [32]. In general, EM algorithms [38] consist in maximizing a likelihood function, which depends on parameters and hidden variables. There two successive steps are performed iteratively:

- An **E**xpectation step, which consists in computation of the expected values of the hidden variables and
- a **M**aximization step, which is to optimize the likelihood function for the parameters.

We propose an algorithm which is closely related to [33]. However, the difference is that we take into account information on the radii of the medial balls in addition. The key idea of [33] is to compute the optimal similarity transformation of (2) relaxing the constraint $\pi \in \mathcal{S}^k$. This allows to apply the EM algorithm.

1. *Relaxation step*.

   A mapping $\pi \in \mathcal{S}^n$ can be represented as a matrix $P \in \{0,1\}^{k\times k}$ with entries

   $$P_{j,i} = \begin{cases} 1 & \text{if } \pi(j) = i \\ 0 & \text{otherwise} \end{cases} .$$

   The relaxation consists in assuming instead of $\mathcal{S}^k$ *stochastic matrices*

   $$\{P \in [0,1]^{k\times k} : \sum_{i=1}^{k} P_{j,i} = 1 \text{ for all } j\} .$$

   It is common to interpret $P_{j,i}$ as the probability that j is mapped to i. For medial ball representation this means that the ball (x_j, r_j) is mapped onto $(\tilde{x}_i, \tilde{r}_i)$.

   Furthermore, a distribution $p = (p_1, \ldots, p_k)$ is assumed over $\{1, \ldots, k\}$. The number p_i represents the probability, that an arbitrary element of $\{1, \ldots, k\}$ is mapped to i.

2. *Distances of balls*.

   Under the condition $\pi(j) = i$ and for a given similarity transformation (ρ, A, b), we assume that the transformed ball (x_j, r_j) is normally distributed around $(\tilde{x}_i, \tilde{r}_i)$, i.e. $\rho A x_j + b$ is normally distributed with mean $\tilde{x}_i$ and ρr_j is normally distributed with mean $\tilde{r}_i$. Concerning the radii, this assumption is slightly inaccurate since radii cannot become negative; we overcome this problem by choosing the variance σ sufficiently small, such that the 0.01 quantile is positive. Hence we define (for i, $j = 1, \ldots, k$)

   $$g_i(x_j, r_j) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left(-\frac{1}{2\sigma^2}\left(||\rho A x_j + b - \tilde{x}_i||_2^2 + \alpha(\rho r_j - \tilde{r}_i)^2\right)\right) . \tag{3}$$

   The function $g_i(x_j, r_j)$ is large if (x_j, r_j) is mapped closely to $(\tilde{x}_i, \tilde{r}_i)$ and small otherwise.

The likelihood function for the parameters of the similarity transform (ρ, A, b) and the relaxed labeling (P, p) for given ball representations M and $\tilde{M}$ is then given by the product probability

$$L(\rho, A, b, P, p) = \prod_{i,j=1}^{k} \left(p_i g_i(x_j, r_j)\right)^{P_{j,i}} . \tag{4}$$

Taking the logarithm on both sides defines the log-likelihood function

$$l(\rho, A, b, P, p) = \sum_{i,j=1}^{k} P_{j,i} \left(\log p_i + \log g_i(x_j, r_j)\right) . \tag{5}$$

It is important to note, that if P is a permutation, then $p_i = 1/k$. Therefore in this case, maximizers of (5) are minimizers in (2) and vice versa. Following [33], enlarging

the class of labelings from $\mathcal{S}^k$ to stochastic matrices has little effect on the optimal similarity transform.
The advantage of the formulation (5), compared to (2), is, that (5) can be minimized efficiently with an EM-algorithm, where the variables ρ, A, b, p are regarded as parameters, and the entries of P are regarded as hidden variables of the likelihood function l. The EM algorithm for minimization of (5) reads as follows:

Algorithm 1 EM Maximization of l

Initialize $\rho = 1$, $A = id$ and $b = 0 \in \mathbb{R}^3$. Choose values for σ, α.
while l still increases **do**
E-Step: Compute the expected value of the relaxed permutation matrix P,

$$P_{j,i} \leftarrow \frac{p_i g_i(x_j, r_j)}{\sum_{s=1}^{k} p_s g_s(x_j, r_j)} \tag{6}$$

By this assignment it is guaranteed that $\sum_{i=1}^{k} P_{j,i} = 1$ for all j.

M-Step: Maximize

$$\sum_{i,j=1}^{k} P_{j,i}(\log p_i + \log g_i(x_j, r_j)) \tag{7}$$

over A, ρ, b and $p_1, \ldots, p_k$.
Compute the new value of l.
end while

The formulas to compute the maximum in (7) and their derivation are given in the appendix.

Compared to different algorithms for minimizing (2) such as Markov Chain Monte Carlo Methods (MCMC), the output of the EM Algorithm is more sensitive to initial values of the parameters. This is no drawback in our case, since the ball representations are often already quite well aligned. However, the provided algorithm is assumed to be more efficient than MCMC.

3.2 Labeling

This subsection is concerned with computing an optimal labeling $\pi \in \mathcal{S}^k$ in (2).

It is instructive to reformulate the labeling problem as a combinatorial optimization problem.

1. A weighted graph is constructed of the following data:
   - Nodes are the $2k$ balls $(x_1, r_1) \ldots, (x_k, r_k)$ and $(\tilde{x}_1, \tilde{r}_1), \ldots, (\tilde{x}_k, \tilde{r}_k)$.

- Let α be as in (2). The weights of the edges between the nodes are defined by

$$w_{j,i} = ||\rho A x_j - b - \tilde{x}_i)||^2 + \alpha(\rho r_j - \tilde{r}_i)^2, \quad \text{with } \alpha \in [0,1].$$

  Here, (ρ, A, b) is the similarity transform computed in Subsection 3.1.

2. A permutation $\pi \in \mathcal{S}^k$ minimizing

$$\sum_{i=1}^{k} w_{\pi(i),i} \tag{8}$$

   is computed.

For the fixed similarity transform transformation (ρ, A, b), a minimizer $\pi \in \mathcal{S}^k$ of (2) is a minimizer of (8) and vice versa. For the solution of the matching problem (8) we apply the Kuhn-Munkres algorithm, which is a special case of the primal-dual algorithm in linear programming [3]. According to [46] the complexity of the algorithm is $O(k^3)$. In our applications, the high order of complexity of the algorithm is not an issue since it is only applied for medial ball representations of dimension k, which is in the range of 10 to 50.

4 Medial Ball Shape Space by Procrustes Analysis

In the following we perform a statistics of the ball representations $M_1, \ldots, M_n$ and establish a shape space, which we call *medial ball* shape space. Moreover, a distance between elements of the medial ball shape space is employed to compare them quantitatively.

4.1 Elements of the medial ball shape space

A shape is all the geometrical information that remains when location, scale and rotational effects are filtered out from an object (taken from [19, Chapter 1]). We consider this definition when constructing the medial ball shape space as the factor space of ball representations $M = (x_1, \ldots, x_k; r_1, \ldots, r_k)$ modulo similarity transformations. That is

$$\Sigma = \mathbb{R}^{3\times k} \times (\mathbb{R}^{>0})^k / \sim_{ST} \ . \tag{9}$$

That is, two ball representations $M = (x_1, \ldots, x_k; r_1, \ldots, r_k)$ and $\tilde{M} = (\tilde{x}_1, \ldots, \tilde{x}_k; \tilde{r}_1, \ldots, \tilde{r}_k)$ represent the same shape (or in other words are equivalent with respect to $\sim_{ST}$) if there exists a similarity transformation (ρ, A, b) such that

$$\rho A \tilde{x}_i + b = x_i \text{ and } \rho \tilde{r}_i = r_i \text{ for } i = 1, \ldots, k \tag{10}$$

We emphasize that in the segmentation literature (see e.g. [16, 40, 36, 18]) frequently shape spaces are not defined as factor spaces. As a consequence the according shape space allows for different representations of the same shape.

Below we derive a segmentation model taking into account the original shape concept and coordinates of Kendall [31]. Compared to Bookstein [11], Kendall coordinates do not depend on the choice of a baseline i.e. two medial balls in our case. Therefore statistical analysis like PCA is not distorted. Since it is bulky to define a distance between equivalence classes we systematically choose representatives, for which we define a shape distance. The process is similar to [19, Chapter 4] where mid axis representations $M = (x_1, \ldots, x_k)$ (without radii) have been considered.

1. A *mean* ball representation $\boldsymbol{\mu}$ of the (labeled) ball representations $M_1, \ldots, M_n$ is defined by
$$\boldsymbol{\mu} = \underset{\tilde{\boldsymbol{\mu}}=\{(\tilde{\mu}_1,\tilde{r}_1),\ldots,(\tilde{\mu}_k,\tilde{r}_k)\}}{\arg\min} \sum_{l=1}^{n} d(M_l, \tilde{\boldsymbol{\mu}})^2 \tag{11}$$
where
$$d(M_l, \tilde{\boldsymbol{\mu}})^2 = \min_{\rho>0, A\in SO(3), b\in\mathbb{R}^3} \sum_{i=1}^{k} \|\rho A x_i^{(l)} - b - \tilde{\mu}_i\|^2 + \alpha(\rho r_i^{(l)} - \tilde{r}_i)^2 \tag{12}$$
The parameter α is the same as in Equation (2). The minimization problem (11) can be solved with an iterative algorithm, which is along the lines of [19, Chapter 5].

2. M is normalized with respect to *translation* by multiplying the ball centers of M with the $(k-1)\times k$ ***Helmert submatrix*** H (see [19, p.34] for the definition). The advantage in using the Helmert matrix for this task (compared to for instance centering the coordinates) is that the arising representation only consists of $k-1$ centers and that the Helmert matrices can be computed efficiently inductively. The resulting representation $T(M)$ is given by
$$T(M) = ((H[c(M)^T])^T; r_1, \ldots, r_k) \in \mathbb{R}^{3\times(k-1)} \backslash \{0\} \times (\mathbb{R}^{>0})^k \,. \tag{13}$$

3. $T(M)$ is normalized with respect to *scaling* by dividing the coordinates and radii by the Frobenius norm $\|c(T(M))\|$. The resulting scaled representation is denoted by $S(T(M))$.

4. $S(T(M))$ is normalized with respect to ***Rotation*** by minimizing the Frobenius norm
$$\min_{A\in SO(3)} \|c(S(T(\boldsymbol{\mu})))^T - c(S(T(M)))^T A\| \,. \tag{14}$$
Rotating $S(T(M))$ by the optimal $A \in SO(3)$ gives
$$R_{\boldsymbol{\mu}}(S(T(M))) = \left(c(S(T(M)))A; \frac{r_1}{\|c(T(M))\|}, \ldots, \frac{r_k}{\|c(T(M))\|} \right) , \tag{15}$$
which is the representative of the class.

   As shown in [31], the optimal rotation A in (14) can be calculated as follows: If
$$(c(S(T(\boldsymbol{\mu}))))c(S(T(M)))^T = V\Lambda U^T \tag{16}$$

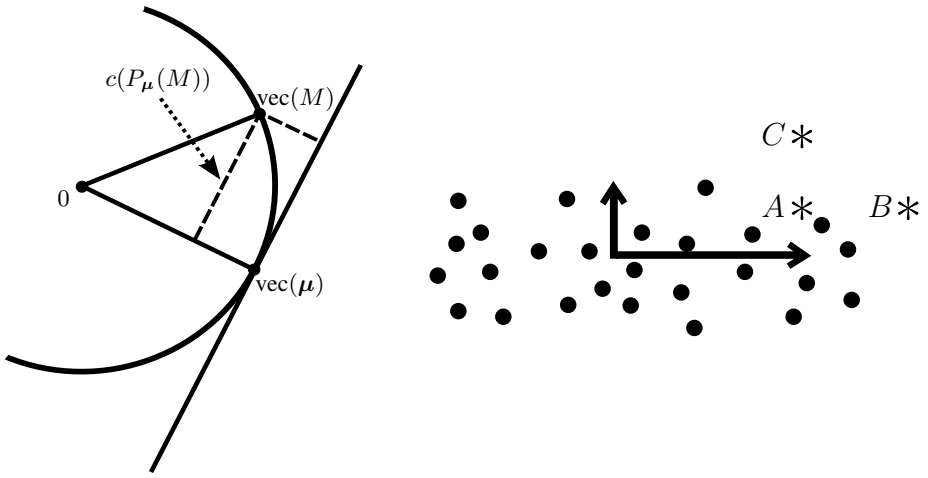


Figure 2: On the left hand side, Procrustes tangent coordinates of the centers of a ball representation are shown. Translated, scaled and rotated ball representations $\boldsymbol{\mu}$, M are on the circle. vec(M) is orthogonally projected to the subspace spanned by vec$(\boldsymbol{\mu})$. The difference between the projection of vec(M) and vec$(\boldsymbol{\mu})$ is $c(P_{\boldsymbol{\mu}}(M))$. PCA and Mahalanobis distance are visualized on the right hand side. The data points vary in horizontal direction much more than in vertical direction. The Mahalanobis distance between A and B is smaller than the distance between A and C, which is the opposite measured in the Euclidean distance.

is the signed singular value decomposition (see Appendix) of $(c(S(T(\boldsymbol{\mu}))))c(S(T(M)))^T$, then

$$A = UV^T. \tag{17}$$

The mean shape μ and the representative are utilized to compute distances in Σ in the next Subsection 4.2.

4.2 Variability in the medial ball shape space

Standard techniques for capturing the variability of data are principal component analysis (PCA) and the Mahalanobis distance. Historically, these techniques have been defined in Euclidean space. A standard way to transfer these concepts to Riemannian manifolds is to project the data to a tangent space (which is Euclidean) and perform an variability analysis there. Recently, techniques have been established, which consider the whole tangent bundle for parallel transport of tangent vectors along geodesics and perform a *principal geodesic analysis*, [23]. In our case, the medial ball shape space Σ is a Riemannian manifold with singularities. As is detailed in [45] for shape spaces of point configurations and carries over to our situation, Σ has singularities at equivalence classes of ball representations under translation and scaling, where the rotation group does not act free. These shapes arise from ball representations consisting of balls of

the same radius lying on a straight line. We can exclude these ball representations from our further considerations since they are not of interest to us, or apply the simulation of simplicity concept [21] to ball representations $M_1, \ldots, M_n$ and avoid this case. For sake of simplicity, we present only details on the projection technique, although the whole shape analysis and segmentation pipeline is generalizable to a nonlinear framework.

The tangent space of Σ at $[\boldsymbol{\mu}]_{\sim_{ST}}$.
For our medial ball shape space Σ, we compute projections of shapes on the tangent space at $[\boldsymbol{\mu}]_{\sim_{ST}}$. Let $M = (x_1, \ldots, x_{k-1}; r_1, \ldots, r_k)$ be a ball representation, in normalized form (15), and $\mathrm{vec}(M) = (x_1^T, \ldots, x_{k-1}^T)^T \in \mathbb{R}^{3k-3}$. Then the orthogonal projection of M to the tangent space of Σ at $\boldsymbol{\mu}$ is given by

$$P_{\boldsymbol{\mu}}(M) = (\mathrm{vec}(M) - \langle \mathrm{vec}(M), \mathrm{vec}(\boldsymbol{\mu}) \rangle \, \mathrm{vec}(\boldsymbol{\mu}); r_1, \ldots, r_k) \in \mathbb{R}^{3k-3} \times (\mathbb{R}^{>0})^k \tag{18}$$

Concerning the center points $c(M)$ of M, these are the Procrustes tangent coordinates as in [19, p.76]. The tangent projection is illustrated in Figure 4.2, left part. Tangent coordinates represent the linear deviation of M relative to $\boldsymbol{\mu}$ and will be used for statistical analysis in the sequel.

PCA and Mahalanobis distance for ball representations in Σ.
First we recall the concepts of *principal component analysis (PCA)* and *Mahalanobis distance* in $\mathbb{R}^d$. For data points $m_1, \ldots, m_n \in \mathbb{R}^d$, a PCA is given by the normed eigenvectors of the covariance matrix

$$\mathrm{Cov}(m_1, \ldots, m_n) = \frac{1}{n} \sum_{i=1}^{n} (m_i - \overline{m})(m_i - \overline{m})^T \qquad \text{with} \qquad \overline{m} = \frac{1}{n} \sum_{i=1}^{n} m_i. \tag{19}$$

The eigenvectors $e_1, \ldots, e_d$ are sorted by the size of the corresponding eigenvalues $\lambda_1, \ldots, \lambda_d$ and give an orthonormal basis of $\mathbb{R}^d$. This basis represents the main directions of variability of the data vectors $m_1, \ldots, m_n$.

A distance which takes into account the variability of the data $m_1, \ldots, m_n$ is given by the Mahalanobis distance. This distance is defined by

$$d_{\mathbb{R}^d}(\tilde{m}_1, \tilde{m}_2) = (\tilde{m}_1 - \tilde{m}_2)^T (\mathrm{Cov}(m_1, \ldots, m_n))^{-1} (\tilde{m}_1 - \tilde{m}_2) \tag{20}$$

for $\tilde{m}_1$ and $\tilde{m}_2 \in \mathbb{R}^d$. Compared to the Euclidean distance, the Mahalanobis distance of two data points is small, if they differ along the main directions of variability (which are the principal components of $m_1, \ldots, m_k$), and it is large, if they differ along other directions. If d is large it might be appropriate to use a PCA first, and taking only the $\tilde{d}$ most significant components $e_1, \ldots, e_{\tilde{d}}$ to eigenvalues $\lambda_1, \ldots, \lambda_{\tilde{d}}$. This results in a simplified Mahalanobis distance

$$\tilde{d}_{\mathbb{R}^d}(\tilde{m}_1, \tilde{m}_2) = (\tilde{m}_1 - \tilde{m}_2)^T \tilde{U} \tilde{D}^{-1} \tilde{U}^T (\tilde{m}_1 - \tilde{m}_2) \tag{21}$$

with $\tilde{U} = (e_1, \ldots, e_{\tilde{d}})$ and $\tilde{D}$ the diagonal matrix with entries $\lambda_1, \ldots, \lambda_{\tilde{d}}$.

Applying these concepts to ball representations, let $P_{\boldsymbol{\mu}}(M_i) = (c^i_{\boldsymbol{\mu}}; r^i_{\boldsymbol{\mu}}) \in \mathbb{R}^{3k-3} \times (\mathbb{R}^{>0})^k$ be the tangent coordinates of M_i, split into centers $c^i_{\boldsymbol{\mu}}$ and radii $r^i_{\boldsymbol{\mu}}$. A PCA of ball representations $M_1, \ldots, M_n$, is given by the eigenvectors and eigenvalues of the covariance matrices

$$\text{Cov}(c^1_{\boldsymbol{\mu}}, \ldots, c^n_{\boldsymbol{\mu}}) \text{ and } \text{Cov}(r^1_{\boldsymbol{\mu}}, \ldots, r^n_{\boldsymbol{\mu}}). \tag{22}$$

The Mahalanobis distance d_Σ between M and $\tilde{M}$ in Σ is then given by (with $P_{\boldsymbol{\mu}}(M) = (c_{\boldsymbol{\mu}}; r_{\boldsymbol{\mu}})$ and $P_{\boldsymbol{\mu}}(\tilde{M}) = (\tilde{c}_{\boldsymbol{\mu}}; \tilde{r}_{\boldsymbol{\mu}})$)

$$\begin{aligned} d_\Sigma(M, \tilde{M}) \quad &= \quad (c_{\boldsymbol{\mu}} - \tilde{c}_{\boldsymbol{\mu}})^T \,\text{Cov}\left(c^1_{\boldsymbol{\mu}}, \ldots, c^n_{\boldsymbol{\mu}}\right)^{-1} (c_{\boldsymbol{\mu}} - \tilde{c}_{\boldsymbol{\mu}}) \\ &\quad + \alpha\, (r_{\boldsymbol{\mu}} - \tilde{r}_{\boldsymbol{\mu}})^T \,\text{Cov}\left(r^1_{\boldsymbol{\mu}}, \ldots, r^n_{\boldsymbol{\mu}}\right)^{-1} (r_{\boldsymbol{\mu}} - \tilde{r}_{\boldsymbol{\mu}}) \ . \end{aligned} \tag{23}$$

In case that the ball representations contain many balls, a PCA is performed first and a Mahalanobis distance as in (21) is computed.
PCA and Mahalanobis distance are illustrated in Figure 4.2, right part.

As a result of this section, a mean shape and distances between shapes have been defined and this statistics can be used as shape prior in segmentation in the sequel.

5 Segmentation

In this section we consider object segmentation in 3D voxel data. Thereby we aim for taking into account shape prior information given in the form that the object to be recovered belongs to medial ball shape space. The distance on the shape space is given by the Mahalanobis distance (see Sections 2-4). The medial ball shape space is computed from boundaries of training objects as discussed in Section 2.

The boundary of a ball representation is defined as the *skin surface* [20]. Below we give a definition and some properties of skin surfaces which motivate this choice.

Finally we describe the whole pipeline of the proposed segmentation process.

5.1 Skin Surfaces

In this subsection, we review the definition and basic properties of skin surfaces. Moreover, we discuss properties of skin surfaces, which predestine them as boundary surfaces of medial ball representations.

Definition of a skin surface.
Basic building blocks of skin surfaces are *weighted points* $p = (x_p, w_p) \in \mathbb{R}^3 \times \mathbb{R}$. Weighted points can for instance represent balls with center x and radius r, setting $x_p = x$ and $w_p = r^2$, and therefore a ball representation can be regarded as a set of weighted points. Addition and scalar multiplication of weighted points p, q are defined by

$$p + q = (x_p + x_q, w_p + w_q + 2\langle x_p, x_q \rangle), \tag{24}$$

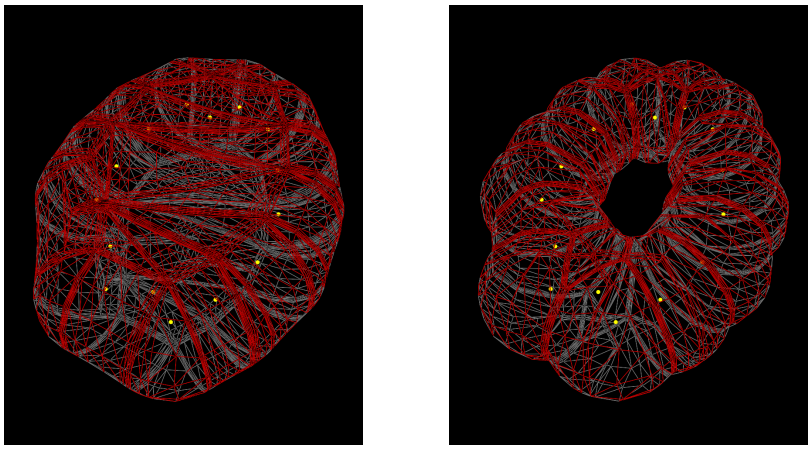


Figure 3: Examples of a skin surface of 15 balls with shrink factor $s = 0.1$ (left) and $s = 0.9$ (right). The dots indicate the centers of the medial balls. Note the topology changing effect of increasing the shrink factor, which can be utilized for constructing shapes of different topology.

and

$$sp = (sx_p, sw_p + (s^2 - s)||x_p||^2) \quad \text{for} \quad s > 0. \tag{25}$$

Here, $\langle \cdot, \cdot \rangle$ and $|| \cdot ||$ denote the Euclidean scalar product and its induced norm. As usual, but with these algebraic operations, the convex hull of a set of weighted points $P = \{p_1, \ldots, p_k\}$ is defined by $\text{conv}(P) = \left\{ \sum_{i=1}^{k} \lambda_i p_i \mid \lambda_i \geq 0, \sum_{i=1}^{k} \lambda_i = 1 \right\}$. The *shrinkage* p^s of a weighted point $p = (x_p, w_p)$ by a factor $s \geq 0$ is defined by $p^s = (x_p, sw_p)$.

Assume now we are given a set P of weighted points $p = (x, w)$ with $w > 0$. Shrinkage applied to P by $s \geq 0$ is defined by $P^s = \{p^s \mid p \in P\}$. Let $B_w(x) = \{y \in \mathbb{R}^3 \mid \|x - y\| \leq w\}$ be a ball with radius w and center x. We denote by ∂ denote the boundary of a set in Euclidean space.

Let $s \geq 0$. Then the *s-skin* of P is defined by

$$\text{skn}^s(P) = \partial \left(\bigcup_{q=(x_q, w_q) \in (\text{conv}P)^s} B_{\sqrt{w_q}}(x_q) \right) \tag{26}$$

Note that the square root of a negative number is imaginary and a ball with an imaginary radius is the empty set.

In the special case $s = 0$, the skin surface of $p_1, \ldots, p_k$ reduces to the convex hull of $x_{p_1}, \ldots, x_{p_k}$, and in case $s = 1$, the skin surface is the boundary of the union of the input weighted points.

Since weighted points are shrunk by a factor of $s \in [0,1]$ in the construction of the skin surface, the skin surface of a surface represented by medial balls is smaller than the original surface. A way to circumvent this problem is by prescaling the input weights $w_{p_1}, \ldots, w_{p_k}$ by $1/s$.

Properties of skin surfaces.
We employ the concept of a skin surface later on for definition of a surface from a ball representation. We believe that skin surfaces are very suitable for this task, since they have several attractive properties which we discuss in the sequel.

- *Information efficiency:* A skin surface only requires center points, weights and a shrinking factor for its complete description. Center points and weights are exactly the information available to us after (mean) ball representation computation in sections (2) and (3). The shrinking factor can be used as a tuning parameter depending on the application. Other popular surface construction methods like NURBS (or other splines), used for instance in [18], or Gregory patches (see [42, 16]) suffer from the drawback, that an additional grid structure of the control points is required for surface generation, which is difficult to obtain in an automatic way.
- *Computational efficiency:* There exists an efficient algorithm for computing skin surfaces (see [35]). For k weighted input points, the algorithm produces a triangular mesh consisting of $\mathcal{O}(k^2)$ vertices.
- *Ball representation fidelity:* If M is a ball representation of some surface in $\mathbb{R}^3$, and $\mathrm{skn}(M)$ is the skin surface defined by M, then M is contained in the ball representation of $\mathrm{skn}(M)$. This follows easily from the definition of the skin surface as a convex hull.
- *Economy:* A small number of input points can generate complicated skin surfaces, [20].
- *Universality:* Every orientable closed surface has a skin representation, [20].

An example of a skin surface, meshed with algorithm [35] is given in Figure 5.1.

The implied skin surface of a shape.
For a fixed ball representation $\boldsymbol{\mu}$, and some given similarity transformation (ρ, A, b), a shape in Σ implies a surface. For its computation, it is necessary to reverse the transformations in Subsection 4.1. To give some details, assume that the shape in Σ is given by its tangent coordinates $P_{\boldsymbol{\mu}} = (c_{\boldsymbol{\mu}}, r_{\boldsymbol{\mu}})$. First note that the tangent projection (18) is injective, with inverse

$$(c_{\boldsymbol{\mu}}; r_{\boldsymbol{\mu}}) \mapsto \left(\mathrm{vec}^{-1}\left(\sqrt{1 - c_{\boldsymbol{\mu}}^T c_{\boldsymbol{\mu}}}\ \mathrm{vec}(c_{\boldsymbol{\mu}}(\boldsymbol{\mu})) + c_{\boldsymbol{\mu}}\right); r_{\boldsymbol{\mu}}\right) \tag{27}$$

mapping tangent coordinates to normalized ball representations. Normalized ball representations are then aligned in $\mathbb{R}^3$ by applying A^{-1} and scaling with $1/\rho$. When reversing translation, note that the Helmert matrix H has a right inverse $\tilde{H} \in \mathbb{R}^{(k-1)\times(k-1)}$.

Applying $\tilde{H}$ and translating the result by b, an (aligned) ball representation M consisting of k balls in $\mathbb{R}^3$ is obtained, to which the skin surface

$$\gamma(M) = \gamma(P_{\boldsymbol{\mu}}, \rho, A, b) \tag{28}$$

is the implied boundary.

5.2 Region and edge based segmentation

For segmentation of voxel images, there are essentially two types of segmentation methods. *Region based* segmentation is applied to images, if the mean intensity of voxels inside the object to segment differs significantly from the mean intensity outside the object. If contrasts are low and objects are only separated by curves, *gradient based* segmentation is used. Here, we briefly discuss both approaches. Let $\Omega \subset \mathbb{R}^3$ be a bounded domain, and $u : \Omega \to \mathbb{R}$ an image intensity function.

Region based segmentation.

For a closed surface $\gamma \subset \Omega$, let $\mathcal{I}(\gamma) \subset \Omega$ be the inner part of γ and $\mathcal{O}(\gamma) \subset \Omega$ the outer part of γ. The mean values of an image intensity function u on $\mathcal{I}(\gamma)$ and $\mathcal{O}(\gamma)$ are then given by

$$\overline{u}_{\mathcal{I}}(\gamma) = \frac{1}{|\mathcal{I}(\gamma)|} \int_{\mathcal{I}(\gamma)} u \, dx \quad \text{and} \quad \overline{u}_{\mathcal{O}}(\gamma) = \frac{1}{|\mathcal{O}(\gamma)|} \int_{\mathcal{O}(\gamma)} u \, dx \tag{29}$$

A simplified version of the Mumford-Shah model introduced in [13] consists in minimization of the functional

$$I_{SMS}(\gamma) = \int_{\mathcal{I}(\gamma)} (\overline{u}_{\mathcal{I}}(\gamma) - u)^2 \, dx + \int_{\mathcal{O}(\gamma)} (\overline{u}_{\mathcal{O}}(\gamma) - u)^2 \, dx \, . \tag{30}$$

If we assume that the mean intensity of an object in image u differs significantly from the mean intensity of the region outside of this object, the functional I_{SMS} is minimized by a submanifold γ which represents the boundary of the object to segment, and $\mathcal{I}(\gamma)$ is the object itself.

Edge based segmentation.

In edge based segmentation, an object and its closed boundary surface γ are implied by high gradients. In this case, it is common to minimize a functional which penalizes small gradients, as for example the Snakes energy introduced in [30]:

$$I_{KWT}(\gamma) = -\int_{\gamma} |\nabla f(\gamma)| dxdy + \alpha \text{Area}(\gamma) \tag{31}$$

Note that edge based energy functionals are usually minimized by active contour methods. They consist of evolving a small initial contour mesh towards the boundary of the object, which is a minimizer of the energy functional. An advantage of active contour methods is that the evolution converges to a minimizer close to the initial contour.

However, since the evolution equations only contain local information, but shape prior is a global information, active contour methods cannot be used efficiently in this case. In order to retain control over local minimizers of (31), we use the term $\beta\text{Area}(\gamma)$ in (31) instead, thus forcing the contour closer to the initial contour.

In the presence of noisy data that e.g. arises in MRI or ultrasound imaging, minimization of both functionals (30) and (31) over γ is not well-posed. Furthermore, if the object to segment is partially occluded, minimizers of (30) and (31) might not only segment the object, but also parts of the background, and knowledge about the true shape of the object to segment has to be considered.
Here we use the mean ball representation $\boldsymbol{\mu}$ of Equation (11) as a priori information and the Mahalanobis distance d_Σ as defined in Subsection 4.2, formula (23) to measure the distance between two ball representations. We recall that with a given fixed ball representation $\boldsymbol{\mu}$, an equivalence class $[M]_{\sim_{ST}}$ of ball representations can be uniquely represented by tangent coordinates $P_{\boldsymbol{\mu}}(M) = (c_{\boldsymbol{\mu}}(M), r_{\boldsymbol{\mu}}(M))$, see (18). Tangent coordinates and a similarity transformation $(\rho, A, b) \in \mathbb{R}^{>0} \times SO(3) + \times\mathbb{R}^3$ uniquely determine a skin surface $\gamma = \gamma(P_{\boldsymbol{\mu}}(M), f)$ as defined in Subsection 5.1, formula (28). We obtain the following functionals as a regularization of (30) resp. (31) for some $\beta > 0$:

$$I_\beta : \Sigma \times \mathbf{SO}(3) \times \mathbb{R}^3 \times \mathbb{R}^{\geq 0} \longrightarrow \mathbb{R},$$

$$(P_{\boldsymbol{\mu}}, f) \mapsto I_{SMS}(\gamma(P_{\boldsymbol{\mu}}, f)) + \beta d_\Sigma(0, P_{\boldsymbol{\mu}}) \tag{32}$$

$$(P_{\boldsymbol{\mu}}, f) \mapsto I_{KWT}(\gamma(P_{\boldsymbol{\mu}}, f)) + \beta d_\Sigma(0, P_{\boldsymbol{\mu}}) \tag{33}$$

Note that the regularization term actually measures the Mahalanobis distance between M and $\boldsymbol{\mu}$, since $P_{\boldsymbol{\mu}}(\boldsymbol{\mu}) = 0$.

6 Results

In the following section we outline the results for the application of our algorithm on two different test images: first, a synthetic dataset of torus voxel images was segmented using functional (32) in Section 6.1. As a second test case we study prostate segmentation in MRI datasets. Thereby we use functional (32), which is described in Section 6.1. Both functionals are minimized using the evolutionary algorithm CMA-ES (see [27, 34]). This choice is mainly motivated by the fact, that the calculation of the gradient for both functionals is tough, since it requires finding the derivative with respect to the skin surface, which is given by the coordinates of the medial ball representation. Thus analytical differentation is numerical differentiation should be used for this purpose. However, the skin surface meshing process we used for surface construction ([35]) does not guarantee that the resulting polyhedra always have the same structure or even the same number of vertices and faces, as it is the case for example by using spline based interpolation surfaces. Those irregularities in the resulting surface meshes complicate the numerical differentiation process and thus we decided to use gradient-free evolutionary minimization algorithms. We chose the evolutionary algorithm CMA-ES for the minimization of 32, because it has been successfully applied to

a similar task in [16], where its superior behavior over two other common evolutionary minimization techniques was demonstrated. For more details on the minimization process, especially also concerning the choice of the minimization parameters, we also refer to [16].

6.1 Synthetic Examples

In this subsection we are concerned with the segmentation of a synthetically created and corrupted torus voxel image. by this example we demonstrate not only the influence of the shape prior on the segmentation result but in particular also the capability of the method to segment surfaces of arbitrary topology, i.e. a torus in this case. We generated 15 binary training voxel images containing elliptic tori with differing radii and positions (see Figure 4). Using the software ITK-SNAP [49] we computed the according input meshes $\Omega_1, \ldots, \Omega_{15}$ which are the actual input data for the shape statistics.

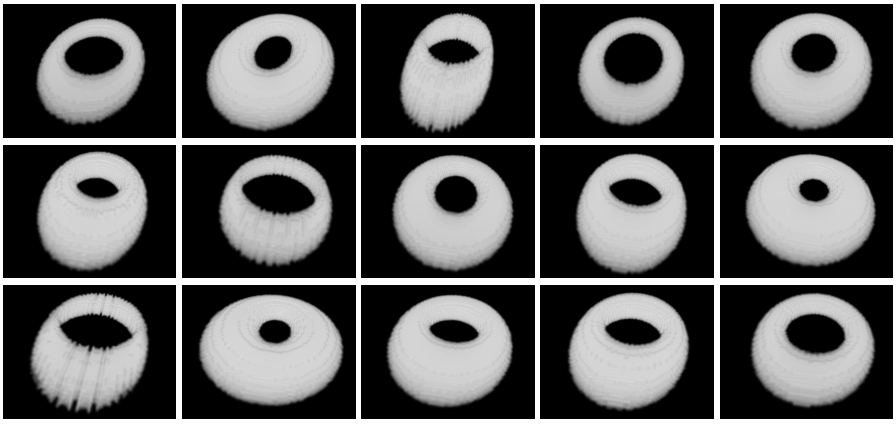


Figure 4: 15 different voxel images of tori used as training data

Using the surface meshes Scheme 1.1 has been used to compute a shape space and a Mahalanobis distance in it. This shape information was then used to segment a torus corrupted by a large black strip as shown in Figure 5(a). Because high contrasts occur in the images, we chose (32) as the segmentation functional, which we minimized using the CMA-ES. The segmentation algorithm was run with $k = 15$ balls. The shrink factor s was included in the minimization process. As initial value for s it proved suitable to take high values if long and thin structures were expected and smaller values in case of more compact objects. The results of the segmentation process with different values for the regularization parameter β are shown in Figure 5(b)-(d). As expected, small values of β, i.e. segmentation with low influence of the regularization, results in improper segmentation results (Figure 5(b)-(c)), while the right choice of β yields a good segmentation, as seen in Figure 5(d). Note that the slight inaccuracies in the

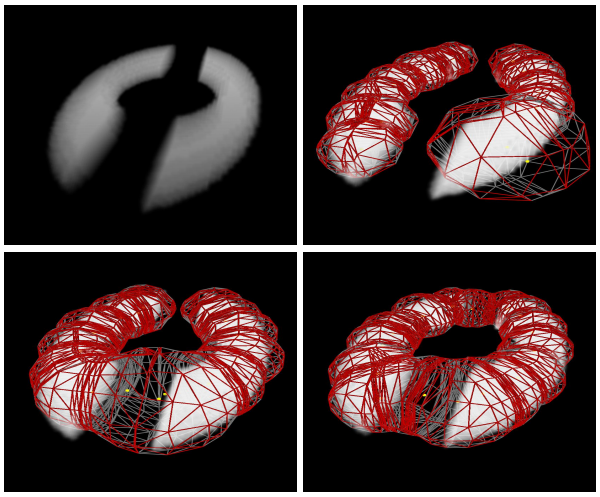


Figure 5: Segmentation results for the corrupted torus image for different values of the regularization parameter β. From upper left to lower right: (a) Voxel image of corrupted torus data. (b) Segmentation result with regularization parameter $\beta = 5 * 10^{-2}$ (c) $\beta = 10^{-2}$ (d) $\beta = 10^{-3}$.

segmentation are a consequence of using skin surfaces as surface generation method, which often results in a surface of ball-like structure by nature.

6.2 Prostate Data

In this example, we are concerned with the segmentation of the prostates in T2-weighted MRI datasets which have been provided for the MICCAI Segmentation Challenge for Clinical Applications [1]. The dataset consisted of 15 training MRI images, which were accompanied by slice-by-slice expert segmentations. Additionally, one dataset was proposed for testing the segmentation algorithm. This dataset is shown in Figure 6(a). As can also be seen there, there are two main challenges for the segmentation algorithm: first, the MRI datasets are very thin, i.e. the number of slices in z-direction is only 28 with the prostate ranging over only about 10 slices. This low resolution poses a general problem for 3D-segmentation algorithms due to rounding errors and leads to the fact, that state- of-the-art techniques for prostate segmentation are often based on the combination of slice-by-slice 2D segmentations.

On the other hand the prostate in the MRI test dataset, which can be seen as the gray colored area in the middle of Figure 6(a), has a very low contrast difference to the other areas of the image. This problem leads to the fact, that a region based functional like (32) is not suitable in this case. Therefore, the edge based segmentation functional (33) was used for the segmentation process. For this purpose, we had to calculate the absolute gradient of the voxel image. To minimize the effect of noise, we first used anisotropic filtering to smooth the voxel image (Figure 6(b)). Afterwards the absolute gradient of the image was calculated using the morphological gradient (Figure 6(c)).

Like in Subsection 6.1, we then again created 15 training meshes $\Omega_1, \ldots, \Omega_{15}$ out of the expert segmentations, which were used to compute shape space and Mahalanobis distance using Scheme 1.1 and the resulting functional combined with the shape prior was minimized using the CMA-ES. The result of this segmentation procedure is shown in Figure 7. As expected from the low input resolution and the complex image contrast structure, the functional needs a high value for the shape prior to yield a good segmentation result. Nevertheless the results are promising.

7 Conclusion

In this paper we presented a new approach for segmentation of 3D voxel images taking into account statistical shape information. The algorithm requires minimal user interactions for segmentation which is achieved by implementing a pipeline of algorithms. The pipeline involves the automatic generation of training shapes represented as medial ball representations that were automatically generated from a given set of input meshes. Afterwards the training shapes are labeled, and a shape space based on Procrustes statistics is established. The resulting shape space and the Mahalanobis distance are used as a prior in region- and edge based segmentation algorithms.

We did a proof-of-concept for our algorithm by first evaluating its performance on a synthetic dataset of 15 randomly generated torus images. By utilizing a segmentation energy which includes statistical regularization using the PCA of the training data, we were able to segment a distorted torus image correctly. Additionally, this example shows the benefit of using skin surfaces as surface generation method for ball representations, which enables us to segment images of arbitrary topology (in contrast to for instance B-splines).

Furthermore, we applied our algorithm to the segmentation of MRI prostate data. Again, we used a statistically regularized segmentation functional supplied with 15 training shapes represented as medial balls which were automatically generated from the available expert segmentations. The automatic generation of medial balls from expert slice-by-slice segmentations is one of the advantages of our algorithm in contrast to the work of Pizer et al. (e.g. [40]), where expert segmentations have to be constructed by using a predefined medial ball model itself. The segmentation result demonstrates that it is in principle possible to automatically reconstruct shapes from MRI images with our method. We want to point out however, that a rigorous evaluation and comparison of the segmentation results for the prostate in MRI images or, more generally spoken, application in the field of medical image segmentation, requires specific fine-tuning of our general formulated algorithm to the specific application field, which is beyond the scope of our paper.

It was shown that the approximation of the medial axis by using only balls with no structure is sufficient in principle but also poses problems when dealing with complex applications. Thus, a future work will be to investigate ways of approximating also the structure of the medial surface. As a first idea, spline surface approximation techniques could be used to approximate the medial surface and interpolate over the resulting medial balls of the discrete medial axis transform. Furthermore, to obtain a better segmentation result for complex medical images one could investigate the use of energy

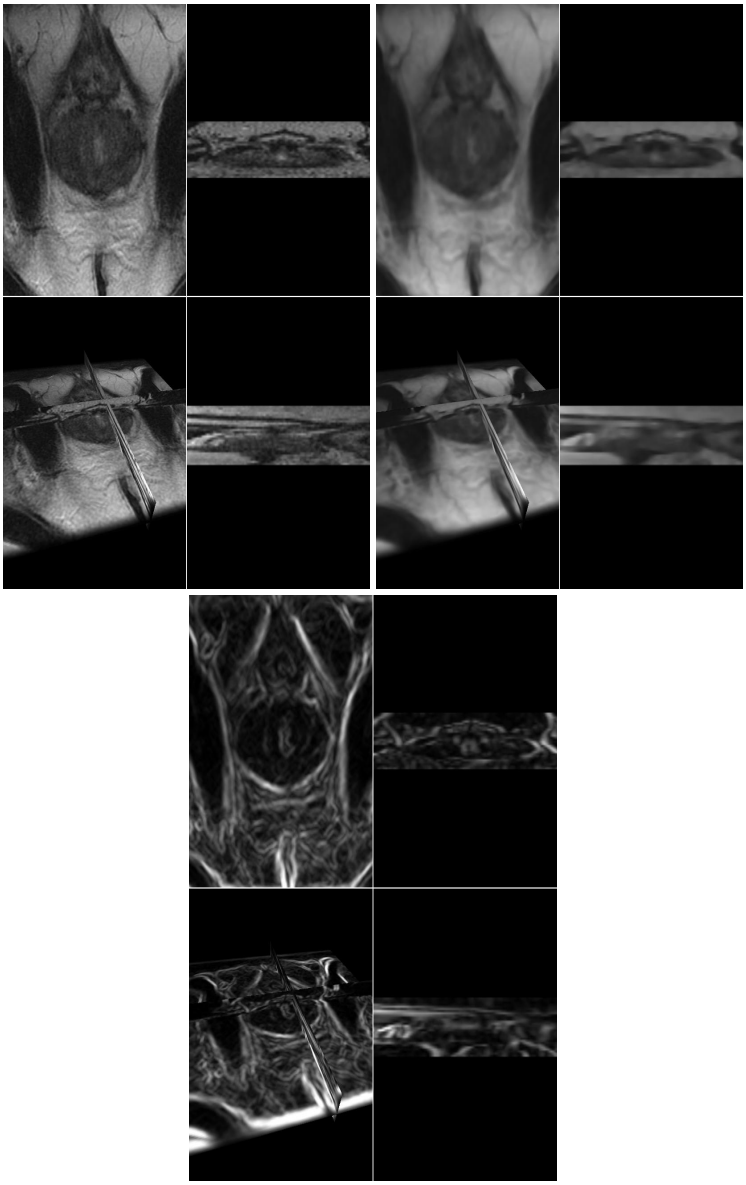


Figure 6: The figure shows (a) the original MRI test image f on the upper left, (b) the MRI test image after smoothing using anisotropic filtering f_σ on the upper right, and (c) the absolute value of the morphological gradient of the smoothed image $|\nabla f_\sigma|$ on the bottom. All images consist of four different views of the voxel image, i.e. from upper left to lower right: transversal, coronal, 3D, sagittal.

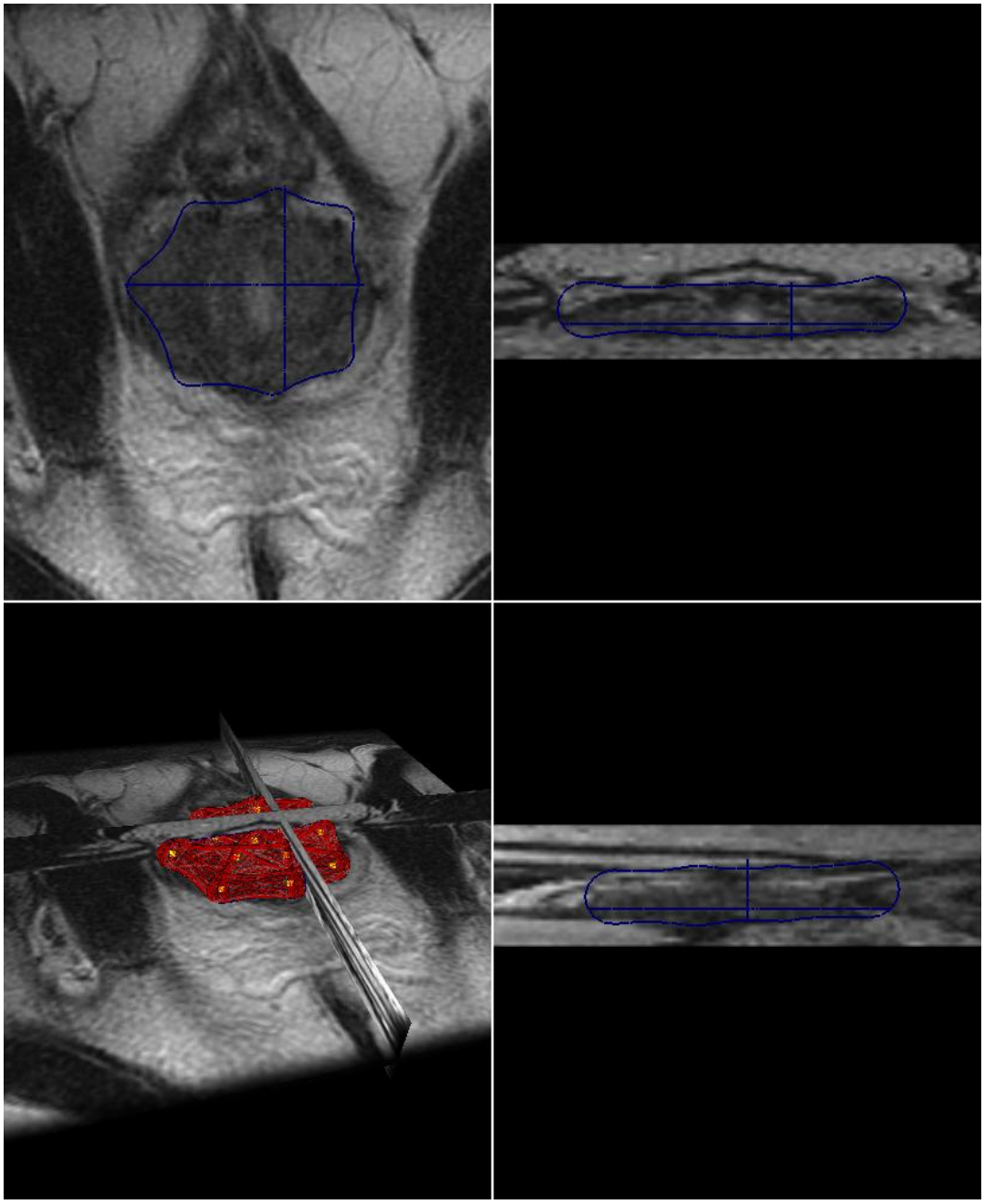


Figure 7: Segmentation result for the prostate in the MRI test image with regularization parameter $\beta = 0.5$. The image shows from upper left to lower right: transversal, coronal, 3D and sagital view of the original image f and the segmentation result. The segmentation result in the 2D-views is visualized by the blue contours, while the resulting mesh surface and its medial ball centers is shown in 3D.

functionals based not only on shape statistics but also image knowledge as e.g. outlined in [25].

Acknowledgement

This work has been supported by the Austrian Science Fund (FWF) within the research networks NFNs *Industrial Geometry*, Projects S09203 and S09205, and Photoacoustic Imaging in Biology and Medicine, Project S10505-N20.

References

[1] Medical image computing and computer assisted intervention. http://www.miccai2009.org.

[2] J. Abhau, W. Hinterberger, and O. Scherzer. Segmenting surfaces of arbitrary topology: A two-step approach. In *Ultrasonic Imaging and Signal Processing*. Proceedings of SPIE – Volume 6437, 2007.

[3] R.K. Ahuja, T.L. Magnanti, and J.B. Orlin. *Network Flows: Theory, Algorithms, and Applications*. Prentice Hall, 1993.

[4] O. Aichholzer, F. Aurenhammer, T. Hackl, M. Kornberger, B. M. Peternell, and H. Pottmann. Approximating boundary-triangulated objects with balls. In *Proc. 23rd European Workshop on Computational Geometry*, pages 130–133, Graz, 2007.

[5] O. Aichholzer, F. Aurenhammer, and B. Kornberger. Stable piecewise-linear approximations of 3d medial axes. Manuscript.

[6] O. Aichholzer, F. Aurenhammer, B. Kornberger, S. Plantinga, G. Rote, A. Sturm, and G. Vegter. Recovering structure from r-sampled objects. In *Proc. Symposium on Geometry Processing*, volume 28, pages 1349–1360, Berlin, 2009.

[7] N. Amenta and M. Bern. Surface reconstruction by "voronoi" filtering. *Discrete & Computational Geometry*, 22:481–504, 1999.

[8] N. Amenta and R. Kolluri. Accurate and efficient unions of balls. In *Proc. 16th Ann. Symp. Computational Geometry*, pages 119–128, Hong Kong, 2000. ACM.

[9] H. Blum and R.N. Nagel. Shape description using weighted symmetric axis features. *Pattern Recognition*, 10:167–180, 1978.

[10] J.-D. Boissonnat and M. Teillaud, editors. *Effective Computational Geometry for Curves and Surfaces*. Springer, 2006.

[11] F. L. Bookstein. *Morphometric tools for landmark data: geometry and biology*. Cambridge University Press, Cambridge, 1997.

[12] X. Bresson, P. Vandergheynst, and J.P. Thiran. A variational model for object segmentation using boundary information and shape prior driven by the mumford-shah functional. *Int. J. Comput. Vision*, 28(2):145–162, 2006.

[13] T. Chan and L. Vese. Active contours without edges. *IEEE Trans. Image Process.*, 10(2):266–277, 2001.

[14] Q. Chen, Z.M. Zhou, M. Tang, P.A. Heng, and D.S. Xia. Shape statistics variational approach for the outer contour segmentation of left ventricle mr images. *IEEE Trans. Inf. Technol. Biomed.*, 10(3):588–597, 2006.

[15] Y. Chen, H.D. Tagare, S. Thiruvenkadam, F. Huang, D. Wilson, K.S. Gopinath, R.W. Briggs, and E.A. Geiser. Using prior shapes in geometric active contours in a variational framework. *Int. J. Comput. Vision*, 50:315–328, 2002.

[16] S. Colutto, F. Frühauf, M Fuchs, and O. Scherzer. The CMA-ES on Riemannian manifolds to reconstruct shapes in 3D voxel images. *IEEE Transactions on Evolutionary Computation*, 2009. to appear.

[17] D. Cremers, T. Kohlberger, and C. Schnoerr. Shape statistics in kernel space for variational image segmentation. *Pattern Recognition*, 36(9):1929–1943, 2003.

[18] D. Cremers, F. Tischhäuser, J. Weickert, and Ch. Schnörr. Diffusion snakes: Introducing statistical shape knowledge into the Mumford-Shah functional. *Int. J. Comput. Vision*, 50:295–313, 2002.

[19] I.L. Dryden and K.V. Mardia. *Statistical Shape Analysis*. John Wiley, Chichester, 1998.

[20] H. Edelsbrunner. Deformable smooth surface design. *Discrete Comp. Geom.*, 21:87–115, 1999.

[21] H. Edelsbrunner and E.P. Mücke. Simulation of simplicity: A technique to cope with degenerate cases in geometric algorithms. *ACM Trans. Graph.*, 9(1):66–104, 1990.

[22] W. Fang and K.L. Chan. Incorporating shape prior into geodesic active contours for detecting partially occluded object. *Pattern Recognition*, 40(7):2163–2172, 2007.

[23] P.T. Fletcher, S. Joshi, C. Ju, and S.M. Pizer. Principal geodesic analysis for the study of nonliner statistics of shape. *IEEE Trans. Med. Imag.*, 23:995–1005, 2004.

[24] D.S. Fritsch, S.M. Pizer, L. Yu, V. Johnson, and E.L. Chaney. Localization and segmentation of medical image objects using deformable shape loci. *Lecture Notes in Computer Science*, 1230:127–140, 1997. IPMI 1997.

[25] M. Fuchs and S. Gerber. Variational shape detection in microscope images based on joint shape and image feature statistics. In *CVPR Workshops 2008. IEEE Computer Society Conference on*, pages 1–8, 2008.

[26] M. Gastaud, M. Barlaut, and G. Aubert. Combining shape prior and statistical features for active contour segmentation. *IEEE Trans. Circuits and Systems*, 14(5):726–734, 2004.

[27] N. Hansen, S. D. Müller, and P. Koumoutsakos. Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES). *Evolutionary Computation*, 11(1):1–18, 2003.

[28] S. Joshi, S. Pizer, P.T. Fletcher, A. Thall, and G. Tracton. Multi-scale 3-d deformable model segmentation based on medical description. In *Proc. International Conference on Information Processing in Medical Imaging (IPMI)*, pages 64–77, 2001.

[29] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu. An efficient k-means clustering algorithm: Analysis and implementation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(7):881–892, 2002.

[30] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *Int. J. Comput. Vision*, 1(4):321–331, 1987.

[31] D.G. Kendall. Shape manifolds, procrustean metrics and complex projective spaces. *Bull. Lond. Math. Soc.*, 16:81–121, 1984.

[32] J. T. Kent, K.V. Mardia, and C.C. Taylor. Matching problems for unlabelled configurations. *Bioinf. Images Wavelets*, pages 33–36, 2004.

[33] J.T. Kent, K.V. Mardia, and C.C. Taylor. Matching unlabelled configurations and protein bioinformatics. *to appear*.

[34] S. Kern and N. Hansen. Evaluating the cma evolution strategy on multimodal test functions. In *Eigth International Conference on Parallel Problem Solving from Nature PPSN VIII*, volume 3242, pages 282–291. Springer, 2004.

[35] N. Kruithof and G. Vegter. Meshing skin surfaces with certified topology. *Comput. Geom.*, 36:166–182, 2007.

[36] M.E. Leventon, W.E.L. Grimson, and O. Faugeras. Statistical shape influence in geodesic active contours. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR-00)*, pages 316–323, Los Alamitos, 2000. IEEE.

[37] K.V. Mardia, J.T. Kent, and J.M. Bibby. *Multivariate Analysis*. Academic Press, London, 1977.

[38] G. McLachlan and T. Krishnan. *The EM algorithm and extensions*, volume 2. Wiley-Interscience, Hoboken, NJ, 2008.

[39] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. Pure Appl. Math.*, 42(5):577–685, 1989.

[40] S.M. Pizer, P.T. Fletcher, S. Joshi, A. Thall, J.Z. Chen, Y. Fridman, D.S. Fritsch, A.G. Gash, J.M. Glotzer, M.R. Jiroutek, C. Lu, K.E. Muller, G. Tracton, P. Yushkevich, and E.L. Chaney. Deformable m-reps for 3d medical image segmentation. *Int. J. Comput. Vision*, 55(2):85–106, 2003.

[41] S.M. Pizer, A.L. Thall, and D.T Chen. M-reps: A new object representation for graphics. Technical Report TR99-030, Department of Computer Science, University of North Carolina — Chapel Hill, 1999.

[42] M. A. Puso and T. A. Laursen. A 3-d contact smoothing algorithm method using gregory patches. *Numer. Meth. in Engineering*, 2002.

[43] M. Rousson and N. Paragios. Shape priors for level set representations. In *European Conference in Computer Vision (ECCV)*, pages 78–93, 2002.

[44] M. Rousson and N. Paragios. Prior knowledge, level set representations & visual grouping. *Int. J. Comput. Vision*, 76(3):231–243, 2007.

[45] C. G. Small. *The Statistical Theory of Shape*. Springer Series in Statistics. Springer, 1996.

[46] S. Suri. Bipartite matching and the hungarian method. *www.cse.ust.hk/ golin/COMP572/Notes/Matching.pdf*.

[47] A. Tsai, A. Yezzi, C. Tempany, D. Tucker, A. Fan, W.E.L. Grimson, and A. Willsky. A shape-based approach to the segmentation of medical imagery using level sets. *IEEE Trans. Med. Imag.*, 22(2):137–154, 2003.

[48] A. Witkin, M. Kass, and D. Terzopoulos. Snakes: Active contour models. *Int. J. Comput. Vision*, 1, No 4:321–331, 1987.

[49] P.A. Yushkevich, J. Piven, C. Hazlett, H.and G. Smith, R. Ho, S., J.C. Gee, and G. Gerig. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage*, 31(3):1116–1128, 2006.

Appendix

THE SIGNED SINGULAR VALUE DECOMPOSITION.

Denote $S_{(i)}$ the i-th row, and $S^{(j)}$ the j-th column of a square matrix S. The signed singular value decomposition of S,

$$S = UDV^T \tag{34}$$

is obtained from the (usual) singular value decomposition $S = UDV^T$ in the following way

- if $\det(U) = \det(V) = 1$, then do nothing
- if $\det(U) = -\det(V) = 1$, replace the last column $V^{(d)}$ of V by $-V^{(d)}$, and replace $D_{d,d}$ by $-D_{d,d}$.
- if $-\det(U) = \det(V) = 1$, replace the last column $U^{(d)}$ of U by $-U^{(d)}$, and replace $D_{d,d}$ by $-D_{d,d}$.
- if $-\det(U) = -\det(V) = 1$, replace the last column $U^{(d)}$ of U by $-U^{(d)}$ and $V^{(d)}$ by $-V^{(d)}$

Performing these manipulations, it is guaranteed that $\det(U) = \det(V) = 1$ and therefore $U, V \in SO(d)$.

FORMULAS FOR THE MAXIMIZATION STEP IN ALGORITHM (1).
Given P, the M-Step of Algorithm (1) consists in maximizing the functional (7),

$$\sum_{i,j=1}^{n} P_{j,i}(\log p_i + \log g_i(x_j, r_j))$$

with respect to f and $p_1, \ldots, p_k, .$ Since $p_1, \ldots, p_k$ and f occur in different summands, maximization can be performed independently for each sum.

(1) Maximizing (7) over $p_1, \ldots, p_k$ subject to the restriction $\sum_{i=1}^{k} p_i = 1$ gives the condition

$$\sum_{j=1}^{k} P_{j,i} \frac{1}{p_i} = \lambda$$

with a Lagrange multiplier $\lambda \in \mathbb{R}$ for all i. Since $\sum_{i,j=1}^{k} P_{j,i} = k$ and $\sum_{i=1}^{k} p_i = 1$ it follows that $\lambda = k$ and therefore

$$p_i = \frac{1}{k} \sum_{j=1}^{k} P_{j,i}, \quad i = 1, \ldots, k.$$

(2) Maximization of (7) with respect to $\rho > 0$, $A \in SO(3)$ and $b \in \mathbb{R}^3$ is equivalent to minimization of

$$\sum_{i,j=1}^{k} P_{j,i} \left(||x_j - \rho A y_i - b||^2 + \alpha (r_j - \rho s_i)^2 \right) \tag{35}$$

over $\rho > 0$, $A \in SO(3)$ and $b \in \mathbb{R}^3$.
We first consider the special case $\alpha = 0$ in (35), and then the general case $\alpha \geq 0$.

(a) The special case $\alpha = 0$.

We denote the weighted means of x and y by

$$\overline{x} = \frac{\sum\limits_{i,j=1}^{k} P_{j,i} x_j}{\sum\limits_{i,j=1}^{k} P_{j,i}} \quad \text{and} \quad \overline{y} = \frac{\sum\limits_{i,j=1}^{k} P_{j,i} y_i}{\sum\limits_{i,j=1}^{k} P_{j,i}} \ .$$

Moreover, let

$$S_{y,x} = \sum_{i,j=1}^{k} P_{j,i} (y_i - \overline{y})(x_j - \overline{x})^T .$$

Following the computations for the unweighted case in [37], we compute the signed singular value decomposition of $S_{y,x}$,

$$S_{y,x} = UDV^T \ .$$

Here, $U, V \in SO(3)$ and D is diagonal with $D_{1,1} \geq D_{2,2} \geq |D_{3,3}|$. Then the solution of minimization problem (35) in case $\alpha = 0$ is given by

$$\rho = \frac{\text{Trace}(D)}{\text{Trace}(S_{x,x})} \tag{36}$$

$$A = UV^T \tag{37}$$

$$b = \overline{x} - A\overline{y} \tag{38}$$

(b) The general case $\alpha \geq 0$.

Since the additional term in case $\alpha \geq 0$ is independent of rotation and scaling, the formulas (37) and (38) for optimal translation and rotation remain the same in the general case. The scaling factor λ, computed by differentiation, computes to

$$\rho = \frac{\text{Trace}(D) + \alpha r^T P s}{\text{Trace}(S_{x,x}) + \alpha r^T P r} \tag{39}$$