Medial structure generation for registration of anatomical structures

Sergio Vera (sergio.vera.h@gmail.com)[1,2], Debora Gil*(debora@cvc.uab.es)[2], H. Martin Kjer (hmkj@dtu.dk)[3], Jens Fagertun[3], Rasmus R. Paulsen (rapa@dtu.dk)[3], and Miguel Ángel González Ballester (ma.gonzalez@upf.edu)[4,5]

[1], Alma IT Systems, Barcelona, Spain
[2]Computer Vision Center, Universitat Autònoma de Barcelona, Barcelona, Spain
[3]Department of Applied Mathematics and Computer Science, Technical University of Denmark, Copenhagen, Denmark
[4]Department of Information and Communication Technologies, Universitat Pompeu Fabra, Barcelona, Spain
[5]ICREA, Barcelona, Spain

Abstract

Medial structures (skeletons and medial manifolds) have shown capacity to describe shape in a compact way. In the field of medical imaging they have been employed to enrich the description of organ anatomy, to improve segmentation or to describe organ position in relation to surrounding structures. Methods for generation of medial structures, however, are prone to the generation of medial artifacts (spurious branches) that traditionally need to be pruned before the medial structure can be used for further computations. The act of pruning, can affect main sections of the medial surface, hindering its performance as shape descriptor. In this work we present a method for the computation of medial structures that generates smooth medial surfaces that do not need to be explicitly pruned. Additionally, a validation framework for medial surface evaluation is presented. Finally, we apply this method to create a parametric model of the cochlea shape that yields better registration results between cochleae.

Keywords: medial axis, medial manifold, medial maps, shape modeling, cochlea registration, validation framework.

*Serra Hunter Fellow

1 Medial Maps for Reliable Extraction of Anatomical Medial Surfaces

Many organs in the human anatomy have shapes that can be represented by genus zero surfaces (surfaces that are homeomorphic to a sphere). Nevertheless, such generalization cannot hide the complexity of human anatomical shapes. Even with this broad common characteristics, anatomical shapes display a large amount of variability between different organs but also variability between the same organ over different patients. This large variability makes anatomical shapes challenging to process by computer techniques. Medial structures have demonstrated to be a compact shape descriptor and have shown potential to capture shape variations. Medial structures [Blum, 1967] such as the medial axis and the medial manifold (medial surface), completely determine the geometry of the boundary volume [Gray, 2004]. In the field of medical imaging, they can be used in many applications: information provided by medial surfaces has shown to improve segmentation results [Pizer et al., 2005; Sun et al., 2010], medial structures have been used to characterize pathological abnormalities [Styner and Gerig, 2003; Styner and Lieberman, 2004], and provide detailed representations of complex organs [Yao and Summers, 2009]. They also provide more intuitive and easily interpretable representations of complex organs [Yao and Summers, 2009] and their relative positions [Liu et al., 2010]. Medial information enhanced modelling has been used in a variety of medical imaging analysis applications, including computational neuroanatomy [Yushkevich et al., 2008; Styner and Lieberman, 2004], 3D cardiac modelling [Sun et al., 2008] or cancer treatment planning [Stough et al., 2007; Crouch et al., 2007]. In shape analysis, medial representations can provide better information than Point Distribution Models since they can model not only the shape but also the interior variations too [Yushkevich et al., 2006a].

In any task where a medial surface is to be used to help the modelization, one must ensure that the medial surfaces have certain properties that do not hinder further computations. Ideally the medial surface should be complete enough to capture the key anatomical shapes of an organ but simple enough so that the topology of the medial structure respond to true changes in shape and not to artifacts (spurious branches). In order to provide accurate meshes of anatomical geometry, the extraction of medial manifolds should satisfy three main conditions [Pudney, 1998]:

- Homotopy: The medial manifold should maintain the same topology (number of holes and components) of the original shape.
- Medialness: The medial structure has to lie as close as possible to the center of the original object.
- Thinness: The resulting medial shape should be as thin as possible without breaking the homotopy rule. The ideal case is to have one pixel-wide structures. However, this concept is too generic, and it heavily depends on the selected connectivity.


(a)

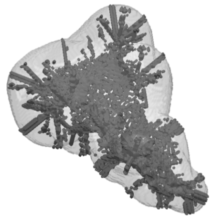

(b)

Figure 1.1: Medial surfaces obtained using a 6-connected neighbourhood, (a), and a 26-connected neighbourhood, (b).

The stability of medial manifold properties depends on the domain on which the medial manifold is computed. Existing methods compute medial structures on either the volumetric voxel domain or a tetrahedral mesh of the volume boundary.

Volumetric approaches can be classified into two big types: morphological thinning and energy-based methods. Morphological methods compute medial manifolds by iterative thinning of the exterior layers of the volumetric object until more thinning breaks surface topology [Bouix et al., 2005; Pudney, 1998; Siddiqi et al., 2002; Palágyi and Kuba, 1999; Ju et al., 2007; Svensson et al., 2002]. Such methods require the definition of a neighbourhood set and conditions for the removal of *simple voxels*, i.e. voxels that can be removed without changing the topology of the object. Furthermore, simplicity tests alone produce (1D) medial axis. Computation of medial manifolds (medial sheets) requires additional tests to know if a voxel lies in a surface and thus cannot be deleted even if it is simple [Pudney, 1998]. Moreover, surface tests might introduce medial axis segments in the medial surface, which is against the mathematical definition of manifold and that may require further pruning [Pudney, 1998; Amenta et al., 2001]. There are many definitions possible for the simplicity and the presence or not of a medial surface voxel, also depending on the local connectivity considered when doing the tests. This means that the skeleton of a shape is not unique and that different methods will generate different medial structures (see Fig. 1.1).

Alternative methods rely on an energy map to ensure medialness on the manifold. Often, this energy image is the distance map of the object [Pudney, 1998] or another energy derived from it, like the average outward flux [Siddiqi et al., 2002; Bouix et al., 2005], level set [Sabry and Farag, 2005; Telea and van Wijk, 2002] or ridges of the distance map [Chang, 2007]. However, to obtain a manifold from the energy image, most methods rely on morphological thinning, in a two step process [Bouix et al., 2005; Pudney, 1998; Siddiqi et al., 2002], thus inheriting the weak points of pure morphological methods.

An alternative to volumetric methods is using the Voronoi diagram tetrahedral mesh of a set of points sampled on the object boundary [Dey and Zhao, 2002; Sheehy et al., 1996; Amenta and Bern, 1998; Amenta et al., 2001; Giesen et al., 2009]. Voronoi methods work on a continuous domain and can naturally resolve branching medial surfaces. However, they still introduce one dimensional spikes associated to boundary irregularities that have to be further pruned [Amenta et al., 2001; Giesen et al., 2009]. Also their computational cost and quality depend on the number of vertices defining the volume boundary mesh and, thus, on the volume resolution [Dey and Zhao, 2002]. Although some recent methods [Giesen et al., 2009] are capable of efficiently dealing with surface perturbations, they are prone to introduce medial loops that distort the medial topology [N.Faraj et al., 2013]. Finally, in the context of medical applications, the voxel discrete domain is the format in which medical data are acquired from medical imaging devices and, thus, it is the natural domain for the implementation of image processing [Khalifa et al., 2010; LL.Dinguraru and et al, 2010] and shape modelling [Peters and Cleary, 2008; Park et al., 2003] algorithms.

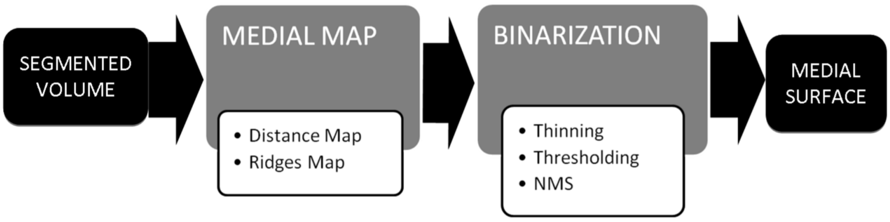


Figure 2.1: Schema of medial surface generation methods.

We have seen that medial surface generation methods are prone to produce noisy medial surfaces with spurial branches. Unwanted branching patterns on the medial surface makes them sub-optimal for the purpose of being used in these medical imaging applications. Pruning operations will decrease these artifacts at the cost of increasing the risk of removing a relevant part of the medial structure. With all this considered, obtaining a good medial structure for usage in medical imaging means to find a compromise between the shape representation and simplicity of medial structure.

We present a method to compute medial structures that is well suited to applications of medical imaging:

1. A novel energy-based method for medial surface computation in images of arbitrary dimensions based on the combination of Gaussian and normalized operators as medialness map followed by non iterative thinning binarization based binarization step is free of topology rules, as it is based on Non-Maxima Suppression (NMS) [Canny, 1986] (Section 2).

2. A validation framework for fair comparison of the quality of medial surfaces: the variability in existing methods for medial surface generation makes comparisons with other methods difficult (Sections 3 and 4).

3. An application of the computation of medial axis for improved registration between human cochleas (Section 5).

2 Extracting Anatomical Medial Surfaces Using Medialness Maps

The computation of medial manifolds from a segmented volume may be split into two main steps: computation of a medial map from the original volume and binarization of such map (Fig. 2.1). Medial maps should achieve a discriminant value on the shape central voxels, while the binarization step should ensure that the resulting medial structures fulfill the three quality conditions [Pudney, 1998] that ensure fair representation of volumes geometry: medialness, thinness and homotopy.

Distance transforms are the basis for obtaining medial manifolds from volumes in any dimension [Pudney, 1998]. The distance transform, also called distance map, is an operator that given a binary volume of a closed domain B, computes for each volume voxel its distance to the domain boundary, ∂B. By definition, maximum values are achieved at the center of B. These voxels correspond to the volume's medial structure and their values depend on the local thickess of the shape.

By the maximality property of distance maps, the medial surface can be obtained using several methods such as iterative thinning [Pudney, 1998], but it can lead to generation of spikes and other discretization artifacts due to the different neighbourhood definitions available. An alternative to iterative thinning is applying a threshold th to $D(B)$. The selection of a good thresholding value of th ensuring homotopy and thinness is problematic. The maximum value of the distance map represents the minimum distance from the medial manifold to the object's boundary. Its value is, therefore, related with the local thickness of the object, and cannot be considered as a global constant value through all the object. On one hand, a too high threshold value is prone to generate unconnected manifold structures violating the homotopy property. On the other hand, a too small value for th means that the medial structure may not be thin. Albeit useful, the distance map is a less than an optimal medialness energy map because it is not selective enough and hinders the binarization step. Further examination of the distance map shows that its central maximal voxels are connected and constitute a ridge surface of the distance map. That is why we claim that the ridges of the distance map provide a better tool to describe the medialness of a set of shape pixels.

2.1 Gaussian Steerable Medial Maps

Ridges/valleys in a digital N-Dimensional image are defined as the set of points that are extrema (minima for ridges and maxima for valleys) in the direction of greatest magnitude of the second order directional derivative [Haralick, 1983]. In image processing, ridge detectors are based either on image intensity profiles [Freeman and Adelson, 1991] or level sets geometry [Lopez et al., 1999]. From the available operators for ridge detection, we have chosen the creaseness measure described in [Lopez et al., 1999] because it provides (normalized) values in the range $[-N, N]$. The ridgeness operator is computed by the structure tensor of the distance map as follows.

Let D denote the distance map to the shape and let its gradient, ∇D, be computed by convolution with partial derivatives of a Gaussian kernel g_σ of variance σ.

The structure tensor or second order matrix [Bigun and Granlund, 1987] is given by averaging the projection matrices onto the distance map gradient:

$$ST_{\rho,\sigma}(D) = \begin{pmatrix} g_\rho * \partial_x D_\sigma^2 & g_\rho * \partial_x D_\sigma \partial_y D_\sigma & g_\rho * \partial_x D_\sigma \partial_z D_\sigma \\ g_\rho * \partial_x D_\sigma \partial_y D_\sigma & g_\rho * \partial_y D_\sigma^2 & g_\rho * \partial_y D_\sigma \partial_z D_\sigma \\ g_\rho * \partial_x D_\sigma \partial_z D_\sigma & g_\rho * \partial_y D_\sigma \partial_z D_\sigma & g_\rho * \partial_z D_\sigma^2 \end{pmatrix} \quad (1)$$

for g_ρ a Gaussian kernel of variance ρ and ∂_x, ∂_y and ∂_z partial derivative operators. Let V be the eigenvector of principal eigenvalue of $ST_{\rho,\sigma}(D)$ and consider its reorientation along the distance gradient, $\tilde{V} = (P, Q, R)$, given as:

$$\tilde{V} = \text{sign}(< \tilde{V} \cdot \nabla D >) \cdot \tilde{V} \quad (2)$$

where $< \cdot >$ the scalar product. The ridgeness measure or NRM (Normalized Ridge Map) [Lopez et al., 1999] is given by the divergence:

$$\mathrm{NRM} := \mathrm{div}(\tilde{V}) = \partial_x P + \partial_y Q + \partial_z R \tag{3}$$

The above operator assigns positive values to ridge pixels and negative values to valley ones. The more positive the value is, the stronger the ridge patterns are. A main advantage over other operators (such as second order oriented Gaussian derivatives) is that NRM $\in [-N, N]$ for N the dimension of the volume. In this way, it is possible to set a threshold, τ, common to any volume for detecting significant ridges and, thus, points highly likely to belong to the medial surface. However, by its geometric nature, NRM has two main limitations. In order to be properly defined, NRM requires that the vector $\tilde{V}$ uniquely defines the tangent space to image level sets. Therefore, the operator achieves strong responses in the case of one-fold medial manifolds, but significantly drops anywhere two or more medial surfaces intersect each other. Additionally, NRM responses are not continuous maps but step-wise almost binary images (Fig.2.2). Such discrete nature of the map is prone to hinder the performance of the NMS binarization step that removes some internal voxels of the medial structure and, thus, introduces holes in the final medial surface.

Ridge maps based on image intensity are computed by convolution with a bank of steerable filters [Freeman and Adelson, 1991]. Steerable filters are given by derivatives of oriented anisotropic 3D Gaussian kernels. Let $\sigma = (\sigma_x, \sigma_y, \sigma_z)$ be the scale of the filter and $\Theta = (\theta, \phi)$ its orientation given by the unitary vector $\eta = (cos(\phi)cos(\theta), cos(\phi)sin(\theta), sin(\phi))$, then the oriented anisotropic 3D Gaussian kernel, g_σ^Θ, is given by:

$$g_\sigma^\Theta = g_{(\sigma_x,\sigma_y,\sigma_z)}^{(\theta,\phi)} = \frac{1}{(2\pi)^{3/2}\sigma_x\sigma_y\sigma_z} e^{-\left(\frac{\tilde{x}^2}{2\sigma_x^2}+\frac{\tilde{y}^2}{2\sigma_y^2}+\frac{\tilde{z}^2}{2\sigma_z^2}\right)} \tag{4}$$

for $(\tilde{x}, \tilde{y}, \tilde{z})$ the change of coordinates given by the rotations of angles θ and ϕ that transform the z-axis into the unitary vector η:

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \end{pmatrix} = R_x(\theta)R_y(\phi)R_x(-\theta) \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{5}$$

with $R_x(\theta)$, $R_y(\phi)$ the following rotation matrices:

$$R_x(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix} \quad R_y(\phi) = \begin{pmatrix} \cos(\phi) & 0 & \sin(\phi) \\ 0 & 1 & 0 \\ -\sin(\phi) & 0 & \cos(\phi) \end{pmatrix} \tag{6}$$

The second partial derivative of g_σ^Θ along the $\tilde{z}$ axis constitutes the principal kernel for computing ridge maps:

$$\partial_z^2 g_\sigma^\Theta = (\tilde{z}^2/\sigma_z^4 - 1/\sigma_z^2) g_\sigma^\Theta \tag{7}$$

We note that by tuning the anisotropy of the Gaussian, we can detect independently medial surfaces and medial axes. For detecting sheet-like ridges, the scales should be set to $\sigma_z > \sigma_x = \sigma_y$, while for medial axes they should fulfill $\sigma_z < \sigma_x < \sigma_y$.

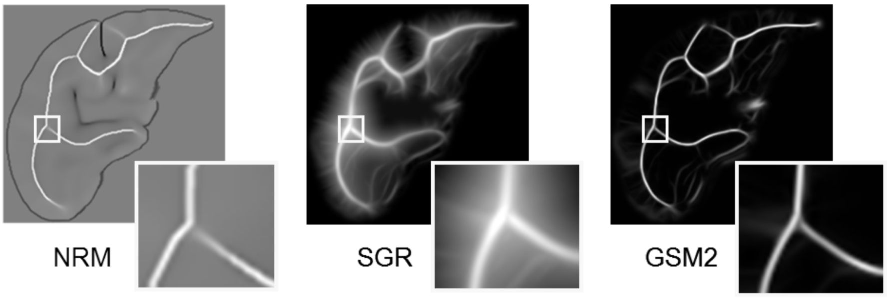


Figure 2.2: Performance of different ridge operators. From left to right: NRM, SGR and GSM2.

The maximum response across Gaussian kernel orientations and the scales gives the Standard Gaussian Ridge (SGR) medial map:

$$\mathrm{SGR} := \max_{\Theta,\sigma}\left(\partial_z^2 g_\sigma^\Theta * D\right) \tag{8}$$

for Θ expressing different orientations of the Gaussian kernel, and σ the scales.

A main advantage of using steerable filters is that their response does not decrease at self-intersections. Their main counterpart is that their response is not normalized, so setting the threshold for binarization becomes a delicate issue [Bouix et al., 2005; Malandain and Fernández-Vidal, 1998].

Given that geometric and intensity methods have complementary properties, we propose combining them into a Geometric Steerable Medial Map (GSM2):

$$\mathrm{GSM2} := \mathrm{SGR}(\mathrm{NRM}) \tag{9}$$

GSM2 generates medial maps with good combination of specificity in detecting medial voxels while having good characteristics for NMS binarization, which does not introduce internal holes.

The three images of Figure 2.2 show the performance of different ridge operators at a 2 dimensional branch (highlighted in the square close up). The geometric NRM (left) produces highly discriminant ridge values. However, they depend on the uniqueness of the direction surface normal, and thus its response significantly decreases at surface branches or self intersections. Steerable Gaussian filters (center) are less sensitive to strong ridges while having increased sensitivity to small, secondary noisy ridges. Finally, the combined approach GSM2 (right) inherits the strong features of each approach. It follows that it achieves a homogeneous response along ridges (induced by NRM normalization) which does not decrease at branches (thanks to the orientations provided by SGR).

2.1.1 Non-Maxima Suppression Binarization

Converting the medialness energy map into a binary set of voxels can be achieved in several ways. As previously stated, thresholding the intensity values of the medialness map yields a reduced set of voxels that are likely to belong to the medial manifold. However, the subset of voxels obtained using thresholding does not necessarily fulfill the property of thinness, and the homotopy heavily depends on the threshold value and the performance of the medial operator at intersections. The usage of iterative thinning schemes after thresholding can generate a thin structure [Bouix et al., 2005], but at the risk of introducing spikes and different surfaces depending on the definition of simple or medial voxels and the order in which voxels are processed.

As an alternative we propose to use Non-Maxima Suppression (NMS) to obtain a thin, one voxel wide medial surface. Non-Maxima Suppression is a well known technique for getting the local maxima of an energy map [Canny, 1986]. For each voxel, NMS consists in checking that the value of its neighbours in a specific direction, V are lower than the actual voxel value. If this condition is not met, the voxel is discarded. In this manner, only voxels that are local maxima along the direction V are preserved. Neighbours of a pixel in a specific direction, V, and delete pixels if their value is not the maximum one:

$$\mathrm{NMS}(x,y,z) = \begin{cases} \mathcal{M}(x,y,z) & \text{if } \mathcal{M}(x,y,z) > \max(\mathcal{M}_{V+}(x,y,z),\ \mathcal{M}_{V-}(x,y,z)) \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

for $\mathcal{M}_{V+} = \mathcal{M}(x+V_x, y+V_y, z+V_z)$ and $\mathcal{M}_{V-} = \mathcal{M}(x-V_x, y-V_y, z-V_z)$.

A main requirement to apply NMS is identifying the local-maxima direction from the medial map derivatives. The search direction for local maxima is given by the eigenvector with highest eigenvalue of the structure tensor of the ridge map, $ST_{\rho,\sigma}(\mathcal{M})$ given by Eq. (1), since it indicates the direction of highest variation of the ridge image. In order to overcome small glitches due to discretization of the direction, NMS is computed using trilinear interpolation.

2.1.2 Parameter Setting

Unlike most of existing parametric methods, the theoretical properties of GSM2 provide a natural way of setting parametric values regardless of the volume size and shape. This new method depends on the parameters involved in the definition of the map GSM2 and in the NMS binarization step.

The parameters arising in the definition of GSM2 are the derivation, σ, and integration, ρ, scales of the structure tensor $ST_{\rho,\sigma}(\mathcal{M})$ used to compute NRM. The derivation scale σ is used to obtain regular gradients in the case of noisy images. The larger it is the more regular the gradient will be at the cost of losing contrast. The integration scale ρ used to average the projection matrices corresponds to time in a solution to the heat equation with initial condition the projection matrix. Therefore large values provide a regular extension of the level sets normal vector, which can be used for contour closing [Gil and Radeva, 2005]. Since in our case we apply NRM to a regular distance map with well defined completed ridges, σ and ρ can be set to their minimum values, $\sigma = 0.5$ and $\rho = 1$.

Concerning steerable filters, the parameters, are the scales, $\sigma = (\sigma_x, \sigma_y, \sigma_z)$, and orientations Θ, defining the steerable filter bank in Eq. (9). These last parameters are usually sampled on a discrete grid, so that Eq. (8) becomes

$$\mathrm{SGR} := \max_{i,j,k} \left(\partial_z^2 g_{\sigma_k}^{\Theta_{i,j}} * D \right) \tag{11}$$

for $\Theta_{i,j}$ given by $\theta_i = \{i\frac{\pi}{N}, \forall i = 1,\ldots,N\}$ and $\phi_j = \{j\frac{\pi}{M}, \forall j = 1,\ldots,M\}$ and $\sigma_k = (\sigma_x^k, \sigma_y^k, \sigma_z^k) = (2^{k+1}, 2^{k+1}, 2^k), k = [0, K]$. Scale depends on the thickness of the ridge and orientations on the complexity of the ridge geometry. The selection of the scale might be critical in the general setting of natural scenes [Lindeberg, 1998]. However in our case, SGR is applied to a normalized ridge map that defines step-wise almost binary images of ridges (see Fig. 2.2, left). Therefore, the choice of scale is not critical anymore. In order to get medial maps as accurate as possible, we recommend using a minimum anisotropic setting: $\sigma_z = 1$, $\sigma_x = \sigma_y = 2$. Finally, orientation sampling should be dense enough in order to capture any local geometry of medial surfaces. In the case of using the minimum scale, eight orientations, $N = M = 8$, are enough.

Therefore, GSM2 is given by:

$$\mathrm{GSM2} = \max_{i,j}\left(\partial_z^2 g_{(2,2,1)}^{\Theta_{i,j}} * NRM\right) \tag{12}$$

for NRM computed over $ST_{1,0.5}(\mathcal{M})$ and $\Theta_{i,j}$ computed setting $N = M = 8$.

The parameters involved in NMS binarization step are the scales of the structure tensor $ST_{\rho,\sigma}(GSM2)$ and the binarizing threshold, τ. Like in the case of NRM, $GSM2$ is a regular function which maximums define closed medial manifolds, so we set the structure tensor scales to their minimum values $\sigma = 0.5$ and $\rho = 1$. Concerning τ, it can be obtained using any histogram threshold calculation, since $GSM2$ inherits the uniform discriminative response along ridges of NRM.

3 Validation Framework for Medial Anatomy Asessment

In order to address the representation of organs for medical use, medial representations should achieve a good reconstruction of the full anatomy and guarantee that the boundaries of the organ are reached from the medial surface. Given that small differences in algorithm criteria can generate different surfaces, we are interested in evaluating the quality of the generated manifold as a tool to recover the original shape.

Validation in the medical imaging field is a delicate issue due to the difficulties for generating ground truth data and quantitative scores valid for reliable application to clinical practice. In this section, we describe our validation framework for evaluating medial surface quality in the context of medical applications. In particular we will generate a synthetic database with ground truth (GT) and two quality tests for assessing the quality of the medial anatomy for data with and without GT. The database can be used to benchmark algorithms using two tests.

3.1 Synthetic Database

The test set of synthetic volumes / surfaces aims to cover different key aspects of medial surface generation (see first row in Fig.4.1). The first batch of surfaces (labelled 'Simple') includes objects generated with a single medial surface. A second batch of surfaces is generated using two intersecting medial surfaces (labelled 'Multiple'), while a last batch of objects (labelled 'Homotopy') covers shapes with different number of holes. Each family of medial topology has 20 samples. The volumetric object obtained from a surface can be generated by using spheres of uniform radii (identified as 'UnifDist') or with spheres of varying radii (identified as 'VarDist').

Volumes are constructed by assigning a radial coordinate to each medial point. In the case of UnifDist, all medial points have the same radial value, while for VarDist they are assigned a value in the range $[5, 10]$ using a polynomial. The values of the radial coordinate must be in a range ensuring that volumes will not present self intersections. Therefore, the maximum range and procedure this radius is assigned depends on the medial topology:

- *Simple*. In this case, there are no restrictions on the radial range.
- *Multiple*. For branching medial surfaces, special care must be taken at surface self-intersecting points. At these locations, radii have to be below the maximum value that ensures that the medial representation defines a local coordinate change [Gray, 2004]. This maximum value depends on the principal curvatures of the intersecting surfaces [Gray, 2004] and it is computed for each surface. Let $\mathcal{M}$ be the medial surface, Z denote the self-intersection points and $D(Z)$ the distance map to Z. The radial coordinate is assigned as follows:

  $$R(X) = min(R(X), max(r_Z, D(Z))) \tag{13}$$

  being $R(X)$ the value of the polynomial function and r_Z the maximum value allowed at self-intersections. In this manner, we obtain a smooth distribution of the radii ensuring volume integrity.
- *Homotopy*. In order to be consistent with the third main property of medial surfaces [Pudney, 1998], volumes must preserve all holes of medial surfaces. In order to do so, the maximum radius r_2 is set to be under the minimum of all surface holes radii.

3.2 Medial Surface Quality Metrics

The database can be used to benchmark algorithms using two tests. The first test evaluates the quality of the medial surface generated, while the second one explores the capabilities of the generated surfaces to recover the original volume and describing anatomical structures. Surface quality tests start from known medial surfaces, that will be considered as ground truth. From this surfaces, volumetric objects can be generated by placing spheres of different radii at each point of the surface. The newly created object is then used as input to several medial surface algorithms and the resulting medial surfaces, compared with the ground truth.

The quality of medial surfaces has been assessed by comparing them to ground truth surfaces in terms of surface distance [Heimann and van Ginneken et al, 2009]. The distance of a voxel y to a surface X is given by: $D_X(y) = \min_{x\in X} \|y - x\|$, for $\|\cdot\|$ the Euclidean norm. If we denote by X the reference surface and Y the computed one, the scores considered are:

1. *Standard Surface Distances*:

   $$AvD = \frac{1}{|Y|}\sum_{y\in Y} D_X(y) \tag{14}$$

$$MxD = \max_{y \in Y}(D_X(y)) \tag{15}$$

2. *Symmetric Surface Distances*:

$$AvSD = \frac{1}{|X| + |Y|} \left(\sum_{x \in X} D_Y(x) + \sum_{y \in Y} D_X(y) \right) \tag{16}$$

$$MxSD = \max \left(\max_{x \in X}(D_Y(x)), \max_{y \in Y}(D_X(y)) \right) \tag{17}$$

Standard distances measure deviation from medialness, while differences between standard and symmetric distances indicate the presence of homotopy artifacts and presence of unnecessary medial segments.

For each family and method, we have computed quality scores statistical ranges as $\mu \pm \sigma$, for μ and σ the average and standard deviation computed over the 20 samples of each group of shapes. The Wilcoxon signed rank test [Wilcoxon, 1945] has been used to detect significant differences across performances.

In medical imaging applications the aim is to generate the simplest medial surface that allows recovering the original volume without losing significant voxels. Volumes recovered from surfaces generated with the different methods are compared with ground truth volumes. Volumes are reconstructed by computing the medial representation [Blum, 1967] with radius given by the values of the distance map on the computed medial surfaces. Ground truth volumes are given by anatomical meshes extracted from original medical scans.

Comparisons with the original anatomical volumes are based on the average and maximum symmetric surface distances (AvSD and MxSD given in (16) and (17)) respectively, computed using the anatomic boundary surface and reconstructed volume boundaries, as well as the following volumetric measures:

1. *Volume Overlap Error:*

$$VOE(A, B) = 100 \times \left(1 - 2\frac{|A \cap B|}{|A| + |B|} \right) \tag{18}$$

2. *Relative Volume Difference:*

$$RVD(A, B) = 100 \times \frac{|A| - |B|}{|B|} \tag{19}$$

3. *Dice coefficient:*

$$Dice(A, B) = \frac{2|A \cap B|}{|A| + |B|} \tag{20}$$

for A, B, being respectively the original and reconstructed volumes. Aside from dice coefficient, lower metric values indicate better reconstruction capability. Like in the case of the synthetic surfaces, for each medial surface method we have computed quality scores statistical ranges as $\mu \pm \sigma$, for μ, σ computed on the medical data set, and Wilcoxon signed rank tests.

4 Validation Experiments

Our validation protocol has been applied to the method described in Section 2.1. To provide a real scenario for the reconstruction tests we have used 14 livers from the SLIVER07 challenge [Heimann and van Ginneken et al, 2009] as a source of anatomical volumes. In order to compare to morphological methods, we have also applied it to an ordered thinning using a 6-connected neighbourhood criterion for defining medial surfaces (labelled Th_6) described in [Bouix and Siddiqi, 2000], a 26-connected neighbourhood surface test (labelled Th_{26}) following [Pudney, 1998]. The consistency of surface pruning is tested on a pruned version of the 26-connected neighbourhood method (labelled ThP_{26}) that does not allow degenerated medial axis segments and the scheme (labelled Tao_6) described in [Ju et al., 2007] that alternates 6-connected curve and surface thinning with more sophisticated pruning stages.

4.1 Medial Surface Quality

Figure 4.1 shows an example of the synthetic volumes in the first row and the computed medial surfaces in the remaining rows. Columns exemplify the different families of volumes generated: one (Simple in 1st and 2nd columns) and two (Multiple in 3rd and 4th columns) foil surfaces, as well as, surfaces with holes (Homotopy in 5th and 6th columns). For each kind of topology we show a volume generated with constant (1st, 3rd and 5th columns) and variable distance (2nd, 4th and last columns). We show medial surfaces in solid meshes and the synthetic volume in semi-transparent color. The shape of surfaces produced using morphological thinning strongly depends on the connectivity rule used. In the absence of pruning, surfaces, in addition, have either extra medial axes attached or extra surface branches in the case pruning is included as part of the thinning surface tests (Tao_6). On the contrary, GSM2 medial surfaces have a well defined shape matching the original synthetic surface.

Table 1 reports error ranges for the four methods and the different types of synthetic volumes, as well as total errors in the last column. For all methods, there are not significant differences between standard and symmetric distances for a given volume. This indicates a good preservation of homotopy. Even with pruning, thinning has significant geometric artifacts (maximum distances increase) and might drop its performance for variable distance volumes due to a different ordering for pixel removal and type of surface preserved.

According to a Wilcoxon signed rank test, strategies alternating curve and surface thinning with pruning stages have worse average distances than other morphological strategies ($p < 0.0001$ for AvD and $p < 0.0001$ for $AvSD$). Given that maximum distances do not significantly differ ($p = 0.4717$, $p = 0.6932$, $p = 0.7752$ for MxD and $p = 0.9144$, $p = 0.7463$, $p = 0.6669$ for $MxSD$), this indicates the introduction of extra structures of larger size (extra surface branches in Tao_6 for the variable volumes shown in Fig. 4.1).

The performance of GSM2 is significantly better than other methods (Wilcoxon signed rank test with $p < 0.0001$), presents high stability across volume geometries and produces accurate surfaces matching synthetic shapes. The small increase in errors for multiple self-crossing surfaces is explained by the presence of holes at intersections between medial manifolds. Still its overall performance clearly surpasses performance of morphological approaches.

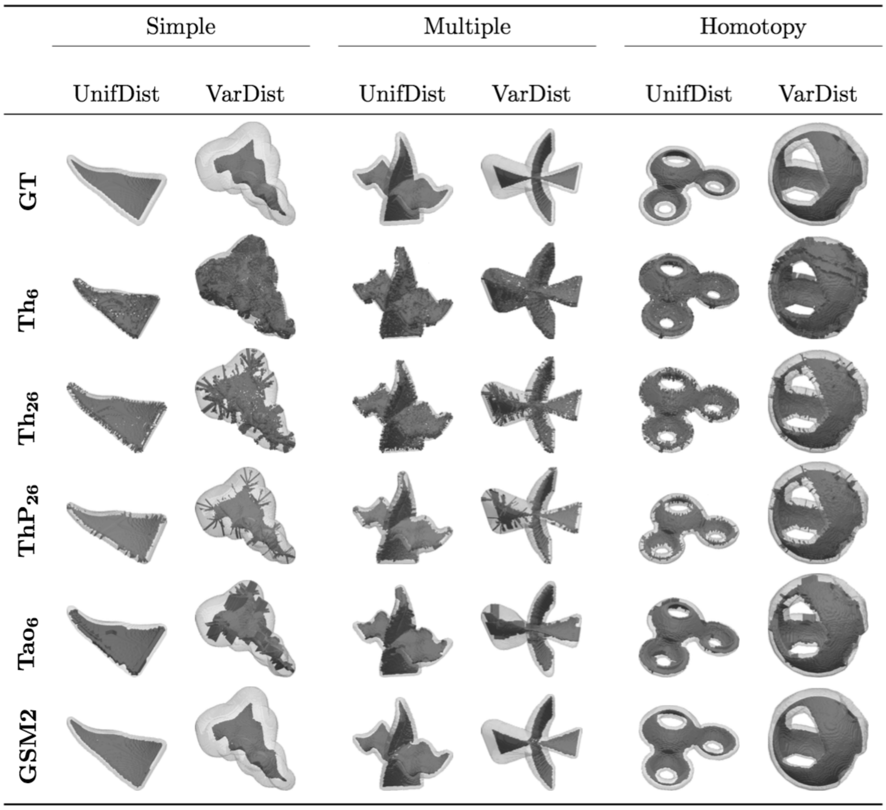


Figure 4.1: Medial surfaces. Examples of the compared methods for each synthetic volume family.

4.2 Reconstruction Power for Clinical Applications

Table 2 reports the statistical ranges for all methods and measures computed for the 14 livers. There are no significant differences among methods and best performers vary depending on the quality measure. However, our approach and the two thinnings, ThP_{26} and Tao_6, have an overall better reconstruction power.

Medial surface of a healthy liver obtained with the thinning methods can be seen in Fig. 4.2 shows the and the GSM2 medial surface in Fig. 4.3 left. In the case of thinning based methods, medial manifolds have a more complex geometry than GSM2 and might include extra structures and self intersections (Fig. 4.2). In medical applications such extra structures might hinder the identification of abnormal or pathological structures. This is not the case for GSM2 surfaces as exemplified in Fig. 4.3. The oversized superior lobe on the right liver is captured by the presence of an unusual medial manifold configuration.

5 Application to Cochlea Registration

The cochlea is the inner ear structure that controls the sensation of hearing and balance, and an understanding of the anatomy and anatomical variability plays

	Simple		Multiple		Homotopy		Total
	UnifDist	VarDist	UnifDist	VarDist	UnifDist	VarDist	
GSM2							
AvD	0.28 ± 0.09	0.28 ± 0.07	0.38 ± 0.09	0.43 ± 0.18	0.37 ± 0.18	0.34 ± 0.14	0.34 ± 0.14
MxD	2.99 ± 0.50	3.50 ± 1.53	3.56 ± 0.53	4.76 ± 1.51	3.39 ± 0.48	3.70 ± 0.84	3.65 ± 1.13
AvSD	0.24 ± 0.05	0.25 ± 0.05	0.37 ± 0.32	0.37 ± 0.18	0.29 ± 0.10	0.28 ± 0.08	0.30 ± 0.17
MxSD	3.02 ± 0.46	3.66 ± 1.52	4.10 ± 2.61	4.76 ± 1.51	3.39 ± 0.48	3.70 ± 0.84	3.78 ± 1.52
$\mathbf{Th_6}$							
AvD	1.52 ± 0.27	5.63 ± 2.19	1.66 ± 0.30	3.05 ± 0.75	1.56 ± 0.35	2.96 ± 1.17	2.73 ± 1.80
MxD	5.55 ± 0.26	16.21 ± 4.76	5.82 ± 0.27	10.75 ± 3.40	5.54 ± 0.20	10.17 ± 3.20	9.01 ± 4.72
AvSD	1.04 ± 0.21	4.34 ± 1.94	1.16 ± 0.24	2.24 ± 0.56	1.09 ± 0.28	2.13 ± 0.98	2.00 ± 1.48
MxSD	5.55 ± 0.26	16.21 ± 4.76	5.82 ± 0.27	10.75 ± 3.40	5.54 ± 0.20	10.17 ± 3.20	9.01 ± 4.72
$\mathbf{Th_{26}}$							
AvD	0.85 ± 0.25	3.15 ± 1.34	1.00 ± 0.19	1.89 ± 0.52	0.86 ± 0.37	1.63 ± 0.84	1.56 ± 1.07
MxD	5.51 ± 0.25	16.17 ± 4.78	5.58 ± 0.19	10.64 ± 3.43	5.46 ± 0.25	10.09 ± 3.21	8.91 ± 4.75
AvSD	0.56 ± 0.14	2.02 ± 0.92	0.67 ± 0.12	1.24 ± 0.35	0.59 ± 0.22	1.05 ± 0.56	1.02 ± 0.69
MxSD	5.51 ± 0.25	16.17 ± 4.78	5.58 ± 0.19	10.64 ± 3.43	5.46 ± 0.25	10.09 ± 3.21	8.91 ± 4.75
$\mathbf{ThP_{26}}$							
AvD	0.57 ± 0.20	2.24 ± 1.00	0.70 ± 0.17	1.38 ± 0.37	0.54 ± 0.24	1.11 ± 0.62	1.09 ± 0.79
MxD	5.49 ± 0.27	16.16 ± 4.78	5.58 ± 0.19	10.61 ± 3.43	5.41 ± 0.27	10.08 ± 3.23	8.89 ± 4.76
AvSD	0.41 ± 0.11	1.38 ± 0.61	0.50 ± 0.11	0.92 ± 0.24	0.41 ± 0.12	0.72 ± 0.37	0.72 ± 0.47
MxSD	5.49 ± 0.27	16.16 ± 4.78	5.58 ± 0.19	10.61 ± 3.43	5.41 ± 0.27	10.08 ± 3.23	8.89 ± 4.76
$\mathbf{Tao_6}$							
AvD	0.79 ± 0.21	4.82 ± 2.05	0.86 ± 0.17	2.46 ± 1.09	0.85 ± 0.29	2.48 ± 1.20	2.04 ± 1.79
MxD	4.87 ± 0.20	17.55 ± 5.19	4.92 ± 0.17	11.10 ± 3.71	4.79 ± 0.21	11.64 ± 4.33	9.14 ± 5.68
AvSD	0.51 ± 0.14	3.92 ± 1.73	0.59 ± 0.13	2.00 ± 0.96	0.59 ± 0.27	1.99 ± 1.03	1.60 ± 1.52
MxSD	4.89 ± 0.18	17.55 ± 5.19	5.32 ± 1.42	11.10 ± 3.71	5.53 ± 3.26	11.87 ± 4.25	9.38 ± 5.73

Table 1: Error ranges (mean and standard deviation) for the Synthetic Volumes.

	$GSM2$	Th_6	Th_{26}	ThP_{26}	Tao_6
Volume Error					
VOE	7.96 ± 1.70	8.84 ± 1.73	8.25 ± 1.72	7.84 ± 1.68	8.49 ± 1.77
RVD	8.49 ± 2.03	9.10 ± 2.10	8.96 ± 2.08	7.86 ± 2.23	5.91 ± 1.99
Dice	$.959 \pm .009$	$.954 \pm .009$	$.957 \pm .009$	$.963 \pm .005$	$.955 \pm .010$
Surface Dist.					
AvSD	0.80 ± 0.06	0.89 ± 0.06	0.83 ± 0.05	0.70 ± 0.11	0.83 ± 0.06
MxSD	5.61 ± 2.68	6.00 ± 2.58	5.52 ± 2.56	5.94 ± 1.45	6.42 ± 2.33

Table 2: Mean and standard deviation of errors in volume reconstruction for each metric.

an important part in utilizing the full potential of Cochlear Implants [Wilson and Dorman, 2008]. Detailed anatomical models have interesting patient-specific applications as they can provide information about the type of electrode design that best suits the anatomy of the user [Vera et al., 2014], or by allowing improvements to the implant programming based on simulations mimicking the actual anatomical and physiological situation [Ceresa et al., 2014].

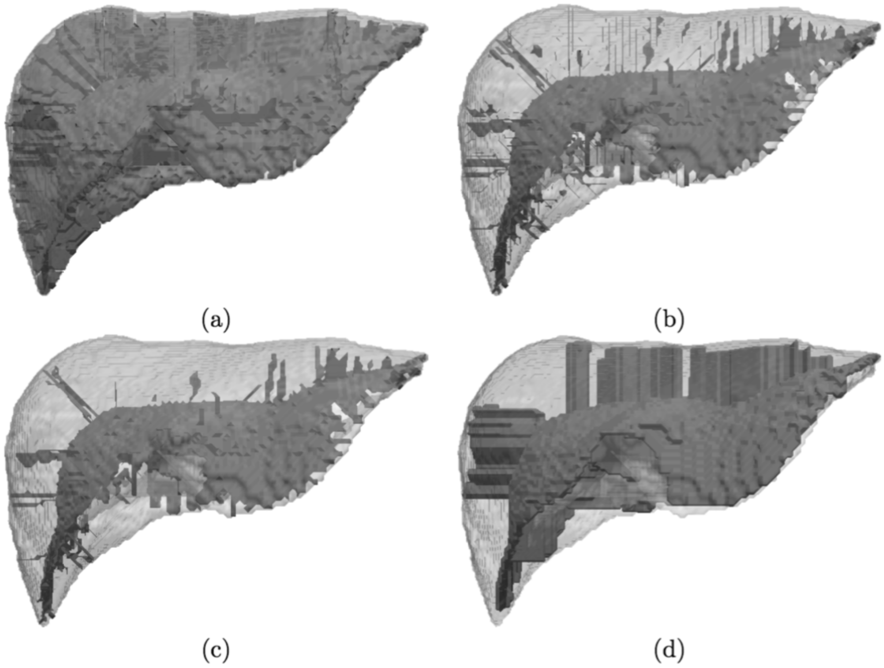


Figure 4.2: Medial manifolds of a healthy liver generated with morphological methods. Th_6 (a), Th_{26} (b), ThP_{26} (c) and Tao_6 (d).

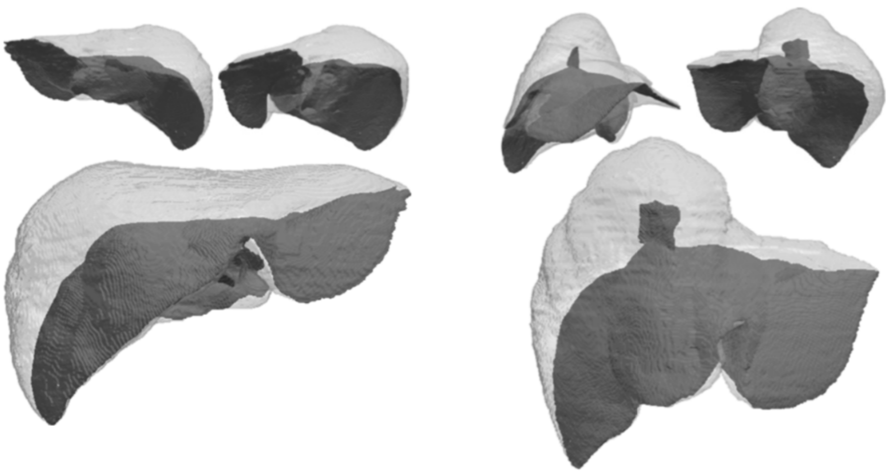


Figure 4.3: GSM2 medial manifolds of a healthy liver (left) and a liver with an unusual lobe (right).

Image registration of the cochlea is challenging for a couple of reasons. The human cochlea is a spiral structure with outer dimensions of approximately 10x8x4 mm. The size and the shape of the spiral can vary extensively. On average, the cochlea winds 2.6 turns [Erixon et al., 2009] but can approach up

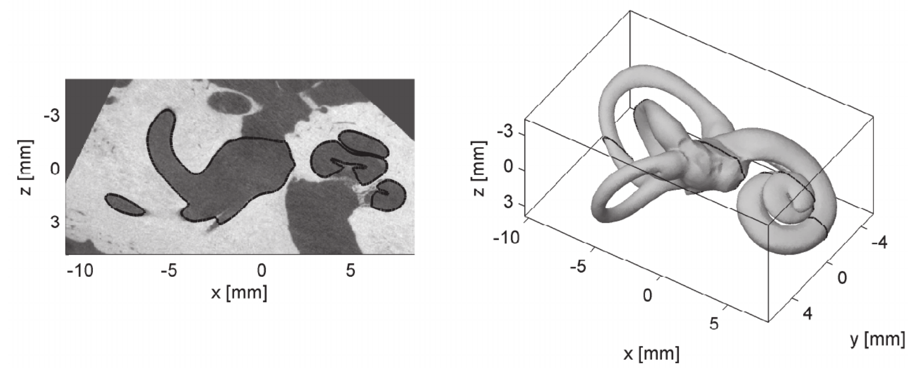


Figure 5.1: Left: Impression of the μCT data and segmentation. Notice the small spacing separating the cochlear turns (right side of CT image), the weak contrast towards internal cochlea borders, and the opening into the middle ear cavity (middle of the image). Right: The corresponding surface model provides an overview of the inner ear topology.

to three full turns - corresponding to a difference in the order of 1-2 mm following the path of the spiral. The separation between the cochlear turns is typically one order of magnitude smaller. Deformations to properly align the most apical region of spiral have been difficult to model to our experience. Further, the whole spiral is a tube-like structure (see Figure5.1, right) with a large degree of self-similarity in the cross-sections. This lack of distinct features makes it difficult to identify corresponding anatomical positions across samples.

The desired registration model should not just expand or compress the apical part of the spiral to align two samples, but rather model a change along the entire spiral. Essentially the model should be able to handle very local deformations while still adhering to the global structure of the samples. This type of behavior is usually not native to non-rigid registration models without some kind of prior or regularization included.

Modifications to a registration model to include such prior knowledge have been studied previously. A way of introducing anatomical shape priors is the use of a statistical shape model [Berendsen et al., 2013; Heimann and Meinzer, 2009]. However, building statistical shape models is in itself a labor intensive task rivaling if not surpassing the task of the registration, as the prerequisite for building the model is data that is already registered to have correspondences.

A multitude of physical constraints have also been proposed as regularizations. For example, local tissue rigidity can be enforced in specified areas [Staring et al., 2007], or conditions of incompressibility or volume-preservation can be applied [Rohlfing et al., 2003]. However, finding the suitable physical constraint for a registration task is not straightforward, as this is case- and application dependent.

In the work of [Baiker et al., 2011] an articulated skeleton model was preregistered to intra-mouse data studies in order to recover large pose-differences between data acquisitions. The presented application is narrow in its scope, but the registration methodology of using landmark correspondences as regularization is more generally applicable, thus we adopt this approach for this work.

In this section we explore the potential of using the skeleton of the cochlea as anatomical prior in free-form registrations using a B-spline transformation model. The skeleton provides a global description of shape in a simplified and structured form. Matching based on skeleton similarity could provide a global anatomical guidance or regularization to a locally defined free-form image registration procedure with a high resistance to noise compared to using only the image intensity similarity.

The use of skeleton similarity in image registrations should be applicable to many different problems and there are many published methods and approaches for finding and matching the skeletons for differing types of data and geometries [Sundar et al., 2003; Tangelder and Veltkamp, 2004]. Skeleton correspondence has been seen in image registration tasks before, relating to for instance 2D/3D multi-modal registration [Liu et al., 1998] and matching of vessels in time-series angiography data [Tom et al., 1994]. More related to our approach is the work of [Tang and Hamarneh, 2008], where multiple different shape features were calculated from surface objects and transformed into vector-valued 2D feature images, which were aligned with a classic image registration formulation. Skeleton features were used for global alignment in the coarser levels of the registration. Our strategy is similar although the prior will be included into the registration model differently.

5.1 Material and methods

A collection of 17 dried temporal bones from the University of Bern were prepared and scanned with a Scanco Medical μCT100 system. The data was reconstructed and processed to obtain image volumes of 24 micron isotropic voxel-sizes containing the inner ear (Figure 5.1, left).

Image segmentation: The border of the inner ear was segmented in all datasets semi-automatically using ITK-SNAP [Yushkevich et al., 2006b].

The semi-automatic tool in the segmentation software was critical for achieving smooth and rounded segmentations in data with that kind of resolution, and for reducing the amount of manual work. A surface model was generated for each dataset using Marching Cubes [Lorensen and Cline, 1987] followed by a surface reconstruction [Paulsen et al., 2010] to obtain a well-formed triangular mesh (Figure 5.1, right).

5.2 Skeletonization

To avoid working with a genus 3 surface, we exclude the vestibular system and focus only on a skeleton of the spiral shaped cochlea. We propose to use a set of corresponding pseudo-landmarks, $\mathcal{Z}^{\mathrm{LM}}$, of the cochleae obtained from a parametric 'curved skeleton' that will improve registration results.

We compute the medial axis of the skeleton by using the $GSM2$ method on the binary segmentations of each cochlea. The computed centerline of the cochlea runs close to the *spiral lamina ossea*, an internal feature of the cochlea that is tied to the perception of different frequencies of sound. We manually define the cochlear apex landmark (A_i), at the extreme of the coclea in each dataset.

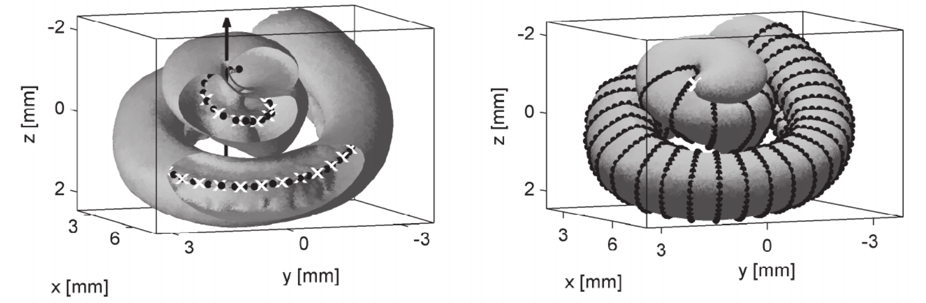


Figure 5.2: The cochlear skeletonization. Left: White 'x' annotations are sampled skeleton information. Right: cochlear apex (A_i) in white. Points on the surface represent parametric pseudo-landmarks.

We generate a naive parametric model of the cochlea. First, we create a parametric description of the cochlea skeleton by sampling 37 corresponding positions on the skeleton with equal arc-length ($\mathcal{Z}_i^{\mathrm{S}}$). Secondly, we extract planar surface cross-section at each of the points, p, in $\mathcal{Z}_i^{\mathrm{S}}$. The cross-section plane is determined by the tangent of the skeleton at p. Each cross-section of the surface mesh is then parameterized using 40 points. These cross-sectional points together with the apex landmark (A_i) provides a set, $\mathcal{Z}_i^{\mathrm{LM}}$, of 1481 corresponding surface pseudo-landmarks (Figure 5.2, right) to be included in a registration model. Finding the cochlea cross-section in the apical region of the cochlear can potentially lead to some ambiguity, as they could intersect with themselves. To avoid this the skeleton cross-sections in the apical turn were not included.

5.3 Image Registration

The registration procedure follows a common work-flow. One dataset was chosen as the reference, to which the remaining moving datasets were registered in two steps - rigid initialization followed by the deformable registration, both detailed in the following subsections.

5.3.1 Initial rigid alignment

There are many approaches for finding rigid transformations. The chosen procedure is independent from the skeleton information and is the same no matter the chosen deformable registration model. In that way, later comparisons of registration results are not affected by the initialization. The whole initialization procedure relies solely upon the extracted surface meshes, but the calculated rigid transformations were also applied to the gray-scale volumes and their segmentations.

Translation: Let $p_{(i,j)}$ be the j-th vertex position of dataset i. A translation was applied so that the center of mass is placed in position (0,0,0), i.e. the mean vertex position, $\bar{p}_i$, was subtracted from all vertices. This places all datasets in a coordinate system where the inner ear center of mass of each dataset is in the origin.

Rotation: Let Σ_i be the 3-x-3 covariance matrix of the mesh vertex positions of dataset i (after the translation). The eigenvectors, W_i, of Σ_i provides a rotation matrix, which when applied transforms the data to the principal component directions. This essentially corresponds to fitting an ellipsoid to the point cloud and aligning the axes.

Check directions: This alignment procedure is robust due to the asymmetry of the inner ear shape (Figure 5.1, right). However, the sign of a principal direction in the i-th dataset could potentially be opposite compared to that of the reference. To handle this we make a simple check. The bounding box of the reference and of the moving point cloud is divided into a coarse grid. We use the sum of squared grid vertex-density difference between the two as a check metric. If the axis-flip would result in a lower metric, then the flip is made to the moving dataset. While there is no guarantee for this to work in all cases, it has worked well for our data. In principle, any kind of rigid alignment could be used instead of the one suggested here.

5.3.2 Deformable registration

The non-rigid image registration follows the formulation and framework of `elastix` [Klein et al., 2010].

The registration is done between the segmentations rather than the gray-scale volumes for two reasons. First, the μCT data contain smaller artifacts and certain weakly contrasted edges, that were handled during the segmentation. Secondly, the registration should not be influenced by the anatomical differences in the surrounding bone structure.

The registration of the moving dataset, I_M, towards the reference, I_F, is formulated as a (parametric) transformation, T_μ, where the vector μ containing the p-parameters of the transformation model are found as an optimization of a cost function, $\mathcal{C}$.

$$\hat{\mu} = \arg\min_{\mu} \mathcal{C}(T_\mu, I_F, I_M) \tag{21}$$

The transformation model used in this paper is the cubic B-spline in a multi-resolution setting. We apply image smoothing with a Gaussian kernel to both the fixed and moving image. For each level of resolution the spacing between grid points and the width of the smoothing kernel follows a decreasing scheme, starting with a coarse registration that is gradually refined. The following scheme was chosen by experimentation:
Control point grid spacing (isotropic, voxels):

$$[144, 72, 48, 48, 36, 24, 18, 12, 6]$$

Width of Guassian kernel (isotropic, voxels):

$$[10, 10, 1, 1, 1, 1, 1, 1, 1]$$

The width of the kernel was deliberately kept narrow in most levels to avoid that small and sharp features would be blurred out (for instance the separation of the cochlear turns). Two of the levels have the same values to overcome limitations of maximum deformation per step.

The cost function used in this 'basic' registration set-up:

$$\mathcal{C}_1 = \alpha \cdot \mathcal{S}_{\text{Sim}}(\mu, I_F, I_M) + (1 - \alpha) \cdot \mathcal{P}_{\text{BE}}(\mu) \tag{22}$$

where α is a weight parameter in the interval [0,1]. The similarity term, $\mathcal{S}_{\text{Sim}}$, is chosen as the sum of squared differences (SSD). The term $\mathcal{P}_{\text{BE}}$ is the energy bending regularization used to penalize strong changes and foldings in the transformation [Rueckert et al., 1999]. The weighting of the similarity term was chosen to 0.9 by experimentation. Increasing α would provide more freedom for deformation of the shapes, but also increase the risk of having non-plausible anatomical results.

The optimization is solved using Adaptive Stochastic Gradient Descent [Klein et al., 2009]. The maximum number iterations was set to 2500. To reduce the computational burden of the optimization only a subset voxels are sampled for the evaluation. For each iteration 2^{14} random coordinate points were sampled. These settings were fixed for all resolutions.

5.3.3 Deformable registration with guidance from skeleton

The free-form registration set-up remains largely the same when a skeleton is included in order to make comparisons fair. The cost function is modified to include a landmark similarity term [Baiker et al., 2011]:

$$\mathcal{C}_2 = \alpha \cdot \mathcal{S}_{\text{Sim}}\left(\mu, I_F, I_M\right) + \beta \cdot \mathcal{S}_{\text{CP}}\left(\mu, \mathcal{Z}_F, \mathcal{Z}_M\right) + (1 - \alpha - \beta) \cdot \mathcal{P}_{\text{BE}}(\mu) \quad (23)$$

where α and β are weightings in the interval [0,1] and fulfilling $\alpha + \beta \leq 1$. The landmark similarity term, $\mathcal{S}_{\text{CP}}\left(\mu, \mathcal{Z}_F, \mathcal{Z}_M\right)$, uses the Euclidean distance between the set of corresponding landmarks, $\mathcal{Z}_F$ and $\mathcal{Z}_M$. In this way intensity-based image registration is guided with features extracted from the anatomical skeleton (i.e. using $\mathcal{Z}_i^{\text{LM}}$ from Section 5.2). By experimentation the weightings were set to $\alpha = 0.8$ and $\beta = 0.11$. The landmark similarity is kept small in order not to force the alignment, and the ratio between image similarity and bending energy regularization is kept similar to the previous set-up $\mathcal{C}_1$ (Eq. 22). Settings for the transformation model and optimizer were unchanged from the previous registration model.

5.4 Evaluation

We are interested in comparing the 16 registration results of model 1 (Eq. 22) and model 2 (Eq. 23) using a number of different image and mesh based metrics.

Image based evaluation: Let $I_i(\mu)$ be the moving segmentation volume after application of the resulting transformation. We compare the Dice Score [Dice, 1945] to the segmentation of the reference dataset, I_{Ref}.

$$\text{DSC} = \frac{2 \cdot |I_{\text{Ref}} \bigcap I_i(\mu)|}{|I_{\text{Ref}}| + |I_i(\mu)|} \quad (24)$$

Mesh based evaluation: We define the surface based scores as follows. Let $\mathbb{S}_{\text{Ref}}(\mu)$ be the reference surface mesh after application of the resulting transformation. There is no direct point correspondence between the reference and the ground truth surfaces, $\mathbb{S}_i$, and they each contain a varying number of vertices. Metrics are therefore based on the closest points, i.e. the minimum

Euclidean distance from a point, p, to any of the points, q, in the other surface, $\mathbb{S}$:

$$d(p, \mathbb{S}) = \min_{\forall q \in \mathbb{S}} (||p - q||_2) \tag{25}$$

The mean surface error, $d_{\bar{s}}$, of each sample is defined as the average of all the closest point distances:

$$d_{\bar{s}} = \frac{1}{N_{\text{Ref}} + N_i} \left(\sum_{\forall p \in \mathbb{S}_{\text{Ref}}(\mu)} d(p, \mathbb{S}_i) + \sum_{\forall p \in \mathbb{S}_i} d(p, \mathbb{S}_{\text{Ref}}(\mu)) \right) \tag{26}$$

where N_{Ref} and N_i are the total number of points in the reference and the moving surface respectively.

The Hausdorff distance, d_H, is the maximum of all the closest point distances:

$$d_H = \max \left\{ \max_{\forall p \in \mathbb{S}_{\text{Ref}}(\mu)} d(p, \mathbb{S}_i), \max_{\forall p \in \mathbb{S}_i} d(p, \mathbb{S}_{\text{Ref}}(\mu)) \right\} \tag{27}$$

The above mentioned metrics are very generic and will hardly be able to reflect and evaluate the change in the registration model that we intend to explore. We therefore include two additional scores, apex error and torque.

First, we calculate the euclidean distance between apexes of the target data and of the reference.

$$d_A = ||A'_{\text{Ref}}(\mu) - A_i||_2 \tag{28}$$

The apex is one of the few locations on the cochlea that can be placed relatively precisely. Even though an arc-length distance might be more correct to use, the euclidean apex error should be indicative of the registration model behavior in the apical region, even though this point is also included in the registration model.

Secondly, we look at the differences in the vector deformation fields obtained by the registration models. The cochlear samples have a different number of turns, and we wish to evaluate the registration models ability to capture this rotational behavior of the anatomy. Our postulation and assumption is that this ability of the registration model should correlate with the 'torque', τ, on the central axis of the cochlear exerted by the deformation field.

Let the force vector, $\vec{F}_p$, on the vertex, p, in the reference mesh be defined simply as the vector between the vertex position before and after application of the registration transformation:

$$\vec{F}_p = p(\mu) - p$$

Further, we can calculate the perpendicular arm from the central axis to the mesh vertex, $\hat{v}_p$. This vector is normalized to unit length, so that the vertices farther from the axis will not contribute with a greater torque.

The scalar projection of the force vector, F_p, onto the unit arm that is perpendicular to both the central axis and $\hat{v}_p$ is then the acting force contributing to the torque:

$$F_p = \vec{F}_p \cdot (\vec{n} \times \hat{v}_p)$$

Using this local vertex torque force leads to our definition of the torque of the registration:

$$\tau = \frac{1}{N_{\text{Ref}}} \sum_{\forall p \in \mathbb{S}_{\text{Ref}}} F_p = \frac{1}{N_{\text{Ref}}} \sum_{\forall p \in \mathbb{S}_{\text{Ref}}} (p(\mu) - p) \cdot (\vec{n} \times \hat{v}_p) \tag{29}$$

	Dice Score	Surf. Error	Hausdorff	Apex Error	Avg. Torque
M1	0.96 ± 0.01	0.040 ± 0.01	0.69 ± 0.24	1.01 ± 0.59	-0.04 ± 0.09
M2	0.95 ± 0.01	0.045 ± 0.01	0.73 ± 0.35	0.69 ± 0.52	-0.53 ± 0.28

Table 3: Statistics of registration evaluation metrics, reported as the mean +/- 1 std. Model 1 is the non-rigid image registration model and Model 2 the non-rigid image registration model incorporating a skeleton prior. Suf Error, Hausdorff and Apex Error metrix are expressed in mm. Average Torque in mm^2

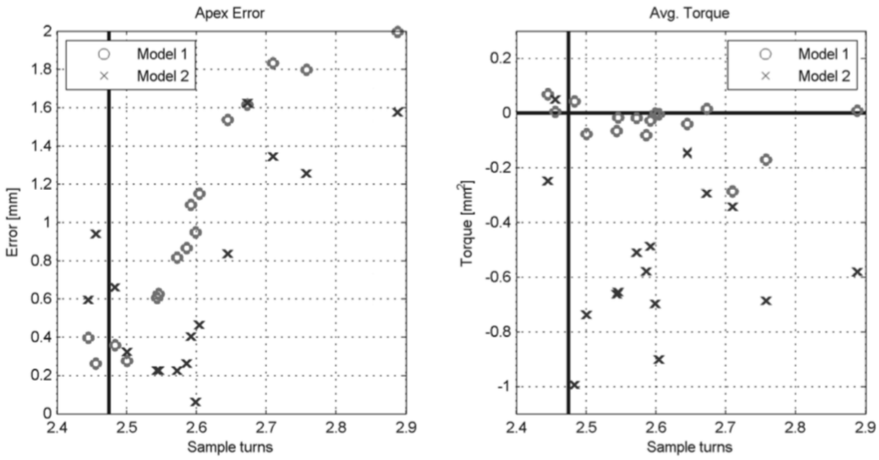


Figure 5.3: Sample-wise apex error (Left) and average torque (Right) plotted against the number of cochlear turns of the target samples. Vertical black line indicate the number of turns in the reference sample.

5.5 Results

The registrations were done on a desktop with a quad-core 3.6 GHz processor, 64 GB RAM, running elastix v4.7. The average time per registration was approximately 0.8 hours and we observed no notable difference in run times or convergence speed between the two registration models.

The statistics of the different metric scores are presented in Table 3. Figure 5.3 elaborates on the sample-wise apex error and torque metric, and Figure 5.4 and 5.5 show the qualitative difference between the registration models.

The general metrics (DSC, $d_{\bar{s}}$, d_H) show a small decrease in performance accuracy for model 2.

From Figure 5.3 it is observed that the apex errors of model 1 grow more or less proportionally to the discrepancy in cochlear turns. The torque is close to zero on average. These observations reflect that model 1 only adapts very locally and behaves indifferently with regards to the turning of the target shape. I.e. the resulting cochlear shapes after registration have little variation in the turns.

The apex errors are seen to be generally lower for model 2. Note, that the apex landmark used to calculate this error was a part of the optimization

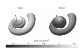


Figure 5.4: Qualitative difference in the local torque acting on the cochlea central axis (black vector). The target sample has 2.60 turns, compared to the 2.46 of the reference (the shown surface). Positive direction of the central axis is defined from the cochlea base towards the apex.

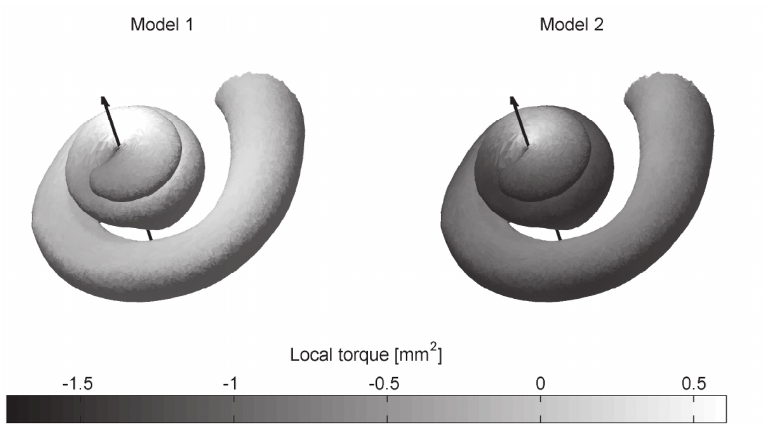


Figure 5.5: The visual difference between registration models. The reference surface is deformed using either model 1 (right, light-grey) or model 2 (left, dark-grey) to align with the target sample (middle, grey). The surfaces have been moved apart to avoid overlap between shapes.

procedure. That the error is reduced is therefore no surprise and it is a biased metric for considering the model accuracy and precision. However, it provides a summarizing pseudo-measure of how much more turning registration model 2 on average is able to capture, which is further illustrated in Figure 5.5. For very large differences in cochlear turns it would seem that both of the registration models have trouble with aligning the apexes.

The torque of model 2 is in most of the cases negative. This indicate vector fields pointing more tangentially in the direction of the spiral towards to the apical region. This would be the expectation as most of the target samples have more turns than the reference. The torque is not a measure of accuracy nor precision. The torque merely provides a simple quantification of the overall rotation of the cochlear shape. Further is gives a good way of illustrating the differences between the registration models as demonstrated in Figure 5.4.

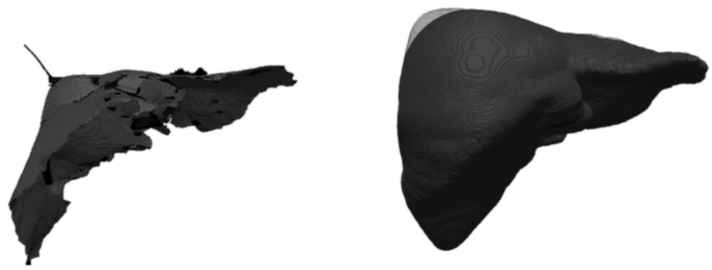
6 Discussion

Medial manifolds are powerful descriptors of shapes. The method presented in this chapter allows the computation of medial manifolds without relying in morphological methods nor neighbourhood or surface tests. Additionally, it can be seamlessly implemented regardless of the dimension of the embedding space.

The performance of our method has been compared to current morphological thinning methods in terms of the quality of medial manifolds and their capability to recover the original volume. For the first experiment a battery of synthetic shapes covering different medial topologies and volume thickness has been generated. For the second one, we have used a public database of CT volumes of livers, including pathological cases with unusual deformations.

The proposed method has several advantages over thinning strategies. It performs equally across medial topologies and volume thickness. The resulting medial surfaces are of greater simplicity than the generated by thinning methods. Although having this minimalistic property, the resulting medial manifolds are suitable for locating unusual pathological shapes and properly restore original volumes. We conclude that our methodology reaches the best compromise between simplicity in geometry and capability for restoring the original volumetric shape.

Any simplification of a medial surface results in a drop in reconstruction quality as illustrated in the images of Fig. 6.1. The images show a medial surface of a liver with a pruned version removing the top branch on the top. Fig. 6.1 (b) shows the volumes reconstructed using the pruned surface (dark color), as well as, the complete one (light color). In this case, the pruned surface cannot reconstruct the external part of the superior lobe of the liver. This drop in accuracy is hard to relate to the simplification process because the branching topology of thinning-based medial manifolds is not always related to the anatomy curvature (concavity-convexity pattern). A main advantage of GSM2 medial surfaces is that their branches are linked to the shape concavities due to the geometrical and normalized nature of the operator. In this context, GSM2 manifolds can be simplified (pruned) ensuring that the loss of reconstruction power will be minimum [Vera and González, 2012].



(a) (b)

Figure 6.1: Impact of pruning in reconstructed volumes: medial manifolds (a) and reconstructed volumes (b).

Regarding computational efficiency, our method is up to 5 times faster than thinning strategies. Unlike parallelization of topological strategies which require special treatment of topological constrains [B., 1995; Palágyi and Kuba, 1999], our code is straightforward to parallelize, even on GPU.

We have shown also how medial information can be used to improve registration of complex anatomy, in this case the cochlea.The Dice Score, surface error and Hausdorff distance serve as very general metrics for evaluating the local adaptability of the registration models. Further, they indicate the general accuracy and precision that we are achieving with the data. The performance with model 2 was decreased on these scores. It would seem that we are trading some local adaption for guiding the model with the landmarks. The determination of the skeleton inherently carries some uncertainties. By introducing the landmarks into the registration model extra noise is added to the procedure. It may happen that a poor skeleton estimate is drawing the spiral in the wrong direction. By providing a more robustly determined skeleton that additionally could fully reach the most apical turn, we expect that the performance of model of 2 could be increased.

Aspects of the skeletonization and its influence can be studied furthermore. For instance the number of landmarks used to represent the skeleton. By experimentation we found an amount of cross-sections that works, but the number of landmarks per cross-section could potentially be reduced. However, the primary concern is the current lack of information in the most apical cochlear turn. For this to be included it would be interesting to look into other skeletal representations. That would in turn potentially require a different way of measuring the similarity of skeletons and possibly an extension to the registration framework to accommodate this. It holds an interesting research potential as both the field of skeletonization and image registration are well-researched areas, but so far joining the two have received little focus. A reason might be the challenge in automatically obtaining consistent skeletons from volumetric data. In this work the skeletons were based on the surface models (i.e. the data segmentation), which in many cases are difficult and/or time-consuming to obtain. Ideally the skeletons should be extracted from volumetric gray-scale data similar to the work of [Abeysinghe et al., 2008; Antúnez and Guibas, 2008].

Using the B-spline grid as the transformation model in the registration has limitations. Choosing a fluid- or optical flow-based model [Oliveira and Tavares, 2014] could potentially be more suited for this kind of spiral anatomy. Alternatively, the performance of the B-spline approach could perhaps be improved with some data preprocessing. If the cochlea was unfolded, possibly based on the skeleton cross-sections, it would be in a space more suited for a B-spline grid transformation. Along the same line of thinking, the deformation control points could be placed in a non-cubic grid structure favoring the spiral nature of the data. However, these suggestions may be difficult to realize and involves adapting the registration method to one very specific task or anatomy. In this and potentially other cases finding a skeleton and including it into the a registration model may be an easier or more feasible approach. The results reflect that it is possible to modify and regularize the registration by using skeleton similarity as a prior, even though there is room for improvements in our methodology.

The registration parameters used in this work were manually determined. A set of parameters that works well on all data samples while running within a reasonable time frame can be difficult to find. Regarding the choice of metric weights, an interval of $\alpha = 0.7 - 0.9$, would seem to be the most appropriate for model 1. Higher α increases the flexibility of the model, which is needed for capturing the cochlear turning. However, increasing beyond 0.9 made some cases fail. In particular the behavior of the deformations in the semi-circular canals performed poorly. The same holds true for model 2. For having a fair comparison between the registration models, the same relative weight of the image similarity and bending energy metric was kept. Having $\beta < 0.15$ was found to be reasonable. Forcing more weight on the landmarks could result in too strong deformations in some cases, and going much lower counters the idea of having the landmarks. Variable metric weights throughout the resolutions were also tested for model 2. I.e. a scheme where a strong weighting was placed on the landmarks in the initial resolutions and then gradually reduced. It worked well in some cases only, so to keep the registration models comparable the fixed weightings scheme was used. Regarding the optimization only the default optimizer and automatically determined settings were used. A number of samples in the range of $2^{14} - 2^{17}$ and a maximum number of iterations between 1000-2500 seemed to produce stable results. Tweaking of registration parameters could result in minor changes of the performance scores, but the same tendencies of the registration models would be observed.

The local torque forces (Figure 5.4) provides the most qualitative view of the differences between the registration models. There is no ground truth torque, but it illustrates that the normal registration model is very local in its adaption, whereas model 2 provides more turning in the region where the skeleton is defined. Ideally we could have shown a more convincingly stronger negative correlation (Figure 5.3) between the differences in the cochlear turns and the average size of the torque. However, we have a low number of samples and the registration also has to deal with general differences in the size and orientation of the samples apart from the turning. In future work the torque could perhaps even be used as a regularization in the registration model, where it could favor a constant torque in the B-spline grid points near the spiral.

Acknowledgments

This research has been supported by the European Commission (GA304857, HEAR-EU), the Spanish Ministry of Economy and Competitiveness (TIN2013-47913-C3-1-R, DEFENSE and DPI2015-65286-R) and the Generalitat de Catalunya (2014-SGR-1470 and Serra Húnter Program).

References

Sasakthi S. Abeysinghe, Matthew Baker, Wah Chiu, and Tao Ju. Segmentation-free skeletonization of grayscale volumes for shape understanding. *IEEE International Conference on Shape Modeling and Applications 2008, Proceedings*, pages 63–71, 2008.

Nina Amenta and Marshall Bern. Surface reconstruction by voronoi filtering. *Discrete and Computational Geometry*, 22:481–504, 1998.

Nina Amenta, Sunghee Choi, and Ravi Krishna Kolluri. The power crust, unions of balls, and the medial axis transform. *Computational Geometry: Theory and Applications*, 19(2-3):127–153, 2001.

Emilio Antúnez and Leonidas Guibas. Robust extraction of 1D skeletons from grayscale 3D images. *Proceedings - International Conference on Pattern Recognition*, 2008. ISSN 10514651.

Gilles B. A parallel thinning algorithm for medial surfaces. *PR*, 16(9):979–986, 1995.

Martin Baiker, Marius Staring, Clemens W. G. M. Löwik, Johan H. C. Reiber, and Boudewijn P. F. Lelieveldt. Automated registration of whole-body follow-up microCT data of mice. *Medical Image Computing and Computer Assisted Intervention*, 14(Pt 2):516–523, 2011. ISSN 03029743, 16113349. doi: 10.1007/978-3-642-23629-7_63.

Floris F. Berendsen, Uulke A. van der Heide, Thomas R. Langerak, Alexis N. T. J. Kotte, and Josien P. W. Pluim. Free-form image registration regularized by a statistical shape model: application to organ segmentation in cervical MR. *Computer Vision and Image Understanding*, 117(9):1119–1127, 2013. ISSN 10773142, 1090235x, 1090235X. doi: 10.1016/j.cviu.2012.12.006.

J. Bigun and G. Granlund. Optimal orientation detection of linear symmetry. In *Proc. of the IEEE 1st International Conference on Computer Vision*, pages 433–438, 1987.

H. Blum. *A transformation for extracting descriptors of shape.* MIT Press, 1967.

S. Bouix and K. Siddiqi. Divergence-based medial surfaces. In *ECCV*, pages 603–618, 2000.

Sylvain Bouix, Kaleem Siddiqi, and Allen Tannenbaum. Flux driven automatic centerline extraction. *Med. Image Anal.*, 9(3):209–21, 2005.

J. Canny. A computational approach to edge detection. *PAMI*, 8:679–698, 1986.

Mario Ceresa, Nerea Mangado Lopez, Hector Dejea Velardo, Noemi Carranza Herrezuelo, Pavel Mistrik, Hans Martin Kjer, Sergio Vera, Rasmus Reinhold Paulsen, and Miguel Ángel González Ballester. Patient-specific simulation of implant placement and function for cochlear implantation surgery planning. *Proc. MICCAI*, pages 49–56, 2014. doi: 10.1007/978-3-319-10470-6_7.

S. Chang. Extracting skeletons from distance maps. *IJCSNS*, 7(7), 2007.

Jessica R. Crouch, Stephen M. Pizer, Senior Member, Edward L. Chaney, Yu chi Hu, Gig S. Mageras, and Marco Zaider. Automated Finite-Element analysis for deformable registration of prostate images. *IEEE Trans. on Medical Imaging*, 26(10):1379–1390, oct 2007.

Tamal K. Dey and Wulue Zhao. Approximate medial axis as a voronoi subcomplex. In *Proc. of the seventh ACM symposium on Solid modeling and applications*, pages 356–366, 2002. ISBN 1-58113-506-8.

Lee R. Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3):pp. 297–302, 1945. ISSN 00129658. URL `http://www.jstor.org/stable/1932409`.

Elsa Erixon, Herman Högstorp, Karin Wadin, and Helge Rask-Andersen. Variational Anatomy of the Human Cochlea: Implications for Cochlear Implantation. *Otology and Neurotology*, 30(1):14–22, 2009. doi: 10.1097/MAO.0b013e31818a08e8.

W.T. Freeman and E.H. Adelson. The design and use of steerable filters. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 13(9):891–906, 1991.

J. Giesen, B. Miklos, M. Pauly, and C. Wormser. The scale axis transform. In *SCG*, pages 106–115, 2009.

Debora Gil and Petia Radeva. Extending anisotropic operators to recover smooth shapes. *Comput. Vis. Image Underst.*, 99(1):110–125, July 2005. ISSN 1077-3142.

A. Gray. *Tubes*. Birkhäuser, 2004.

R. Haralick. Ridges and valleys on digital images. *CVGIP*, 22(10):28–38, 1983.

T. Heimann and B. van Ginneken et al. Comparison and evaluation of methods for liver segmentation from CT datasets. *IEEE Trans Med Imaging*, 28(8): 1251–1265, 2009.

Tobias Heimann and Hans-Peter Meinzer. Statistical shape models for 3D medical image segmentation: A review. *Medical Image Analysis*, 13(4):543–563, 2009. ISSN 13618415, 13618423. doi: 10.1016/j.media.2009.05.004.

Tao Ju, Matthew L. Baker, and Wah Chiu. Computing a family of skeletons of volumetric models for shape description. In *Geometric Modeling and Processing*, pages 235–247, 2007.

F. Khalifa, A. El-Baz, G. Gimel'farb, and R. Ouseph. Shape-appearance guided level-set deformable model for image segmentation. In *icpr*, 2010.

Stefan Klein, Josien P. W. Pluim, Marius Staring, and Max A. Viergever. Adaptive stochastic gradient descent optimisation for image registration. *International Journal of Computer Vision*, 81(3):227–239, 2009. ISSN 09205691, 15731405. doi: 10.1007/s11263-008-0168-y.

Stefan Klein, Marius Staring, Keelin Murphy, Max A. Viergever, and Josien P.W. Pluim. elastix: a toolbox for intensity-based medical image registration. *IEEE Transactions on Medical Imaging*, 29(1):196 – 205, January 2010.

Tony Lindeberg. Feature detection with automatic scale selection. *IJCV*, 30(2): 79–116, 1998.

A. Liu, E. Bullitt, and SM Pizer. 3D/2D registration via skeletal near projective invariance in tubular objects. *Medical Image Computing and Computer-Assisted Intervention - MICCAI'98*, 1496:952–963, 1998. ISSN 03029743.

Xiaofeng Liu, Marius G. Linguraru, Jianhua Yao, and RonaldM. Summers. Organ pose distribution model and an MAP framework for automated abdominal multi-organ localization. In *Medical Imaging and Augmented Reality, Lecture Notes in Computer Science*, volume 6326, pages 393–402, 2010.

LL.Dinguraru and J.A.Pura et al. Multi-organ segmentation from multi-phase abdominal CT via 4D graphs using enhancement, shape and location optimization. In *MICCAI*, volume 13, pages 89–96, 2010.

A.M. Lopez, F. Lumbreras, J. Serrat, and J.J. Villanueva. Evaluation of methods for ridge and valley detection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 21(4):327–335, 1999.

William E. Lorensen and Harvey E. Cline. Marching cubes: A high resolution 3D surface construction algorithm. *SIGGRAPH Comput. Graph.*, 21(4):163–169, August 1987. ISSN 0097-8930. doi: 10.1145/37402.37422. URL `http://doi.acm.org/10.1145/37402.37422`.

G. Malandain and S. Fernández-Vidal. Euclidean skeletons. *Image and Vision Computing*, 16(5):317–327, 1998.

N.Faraj, J.-M. Thiery, and T. Boubekeur. Progressive medial axis filtration. In *SIGGRAPH*, 2013.

Francisco P. M. Oliveira and Joao Manuel R. S. Tavares. Medical image registration: a review. *Computer Methods in Biomechanics and Biomedical Engineering*, 17(2):73–93, 2014. ISSN 10255842, 14768259. doi: 10.1080/10255842.2012.670855.

Kálmán Palágyi and Attila Kuba. A parallel 3d 12-subiteration thinning algorithm. *Graphical Models and Image Processing*, 61(4):199–221, 1999.

H. Park, P.H. Bland, and C.R Meyer. Construction of an abdominal probabilistic atlas and its application in segmentation. *IEEE Trans Med Imaging*, 22(4):483–92, 2003.

R.R. Paulsen, J.A. Baerentzen, and R. Larsen. Markov random field surface reconstruction. *Visualization and Computer Graphics, IEEE Transactions on*, 16(4):636–646, July 2010. ISSN 1077-2626. doi: 10.1109/TVCG.2009.208.

T. Peters and K. Cleary, editors. *Image-Guided Interventions: Technology and Applications*. Springer, 2008.

Stephen M. Pizer, P. Thomas Fletcher, Joshi Sarang, A. G. Gash, J. Stough, A. Thall, Gregg Tracton, and Edward L. Chaney. A method and software for segmentation of anatomic object ensembles by deformable M-Reps. *Medical Physics*, 32(5):1335–1345, 2005.

C. Pudney. Distance-ordered homotopic thinning: A skeletonization algorithm for 3D digital images. *Computer Vision and Image Understanding*, 72(2):404–13, 1998.

T. Rohlfing, CR Maurer, DA Bluemke, and MA Jacobs. Volume-preserving nonrigid registration of MR breast images using free-form deformation with an incompressibility constraint. *IEEE Transactions on Medical Imaging*, 22(6):730–741, 2003. ISSN 02780062, 1558254x. doi: 10.1109/TMI.2003.814791.

Daniel Rueckert, L.I. Sonoda, C. Hayes, D.L.G. Hill, M.O. Leach, and D.J. Hawkes. Nonrigid registration using free-form deformations: Application to breast mr images. *IEEE Transactions on Medical Imaging*, 18(8):712–721, 1999. ISSN 02780062, 1558254x. doi: 10.1109/42.796284.

H. M. Sabry and A. A. Farag. Robust skeletonization using the fast marching method. In *ICIP*, volume 1, pages 437–440, 2005.

Damian J. Sheehy, Cecil G. Armstrong, and Desmond J. Robinson. Shape description by medial surface construction. *IEEE Transactions on Visualization and Computer Graphics*, 2:62–72, 1996. doi: http://doi.ieeecomputersociety.org/10.1109/2945.489387.

Kaleem Siddiqi, Sylvain Bouix, Allen Tannenbaum, and Steven W. Zucker. Hamilton-Jacobi skeletons. *Int. Journal Computer Vision*, 48(3):215–231, 2002.

Marius Staring, Stefan Klein, and Josien P. W. Pluim. A rigidity penalty term for nonrigid registration. *Medical Physics*, 34(11):4098–4108, 2007. ISSN 00942405. doi: 10.1118/1.2776236.

Joshua V. Stough, Robert E. Broadhurst, Stephen M. Pizer, and Edward L. Chaney. Regional appearance in deformable model segmentation. In *Proc. of the 20th int. conf. on Information processing in medical imaging*, volume 4584, pages 532–543, 2007.

M. Styner and G. et al Gerig. Statistical shape analysis of neuroanatomical structures based on medial models. *Med Image Anal.*, 7(3):207–220, 09 2003.

M. Styner and J. A. et al Lieberman. Boundary and medial shape analysis of the hippocampus in schizophrenia. *MIA*, 8(3):197–203, 2004.

H. Sun, A. F. Frangi, H. Wang, and et al. Automatic cardiac mri segmentation using a biventricular deformable medial model. In *MICCAI*, volume 6361, pages 468–475. Springer, 2010.

Hui Sun, Brian B. Avants, Alejandro F. Frangi, Federico Sukno, James C. Gee, and Paul A. Yushkevich. Cardiac medial modeling and time-course heart wall thickness analysis. In *MICCAI Lecture Notes on Computer Science*, volume 5242, pages 766–773, 2008.

H. Sundar, D. Silver, N. Gagvani, and S. Dickinson. Skeleton based shape matching and retrieval. *SMI 2003: Shape Modeling International 2003, Proceedings*, pages 130–139, 2003.

Stina Svensson, Ingela Nyström, and Gabriella Sanniti di Baja. Curve skeletonization of surface-like objects in 3d images guided by voxel classification. *Pattern Recognition Letters*, 23(12):1419–1426, 2002.

Lisa Tang and Ghassan Hamarneh. Smrfi: Shape matching via registration of vector-valued feature images. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8, 2008. ISSN 10636919.

Johan W.H. Tangelder and Remco C. Veltkamp. A survey of content based 3D shape retrieval methods. *Proceedings - Shape Modeling International SMI 2004*, pages 145–156, 2004.

A. Telea and J. van Wijk. An augmented fast marching method for computing skeletons and centerlines. In *Symposium on Data Visualisation*, pages 251–259, 2002.

B.C.S. Tom, S.N. Efstratiadis, and A.K. Katsaggelos. Motion estimation of skeletonized angiographic images using elastic registration. *Medical Imaging, IEEE Transactions on*, 13(3):450–460, Sep 1994. ISSN 0278-0062.

Sergio Vera and Miguel A. et al González. Optimal medial surface generation for anatomical volume representations. In *Lecture Notes on Computer Science*, volume 7601, pages 265–273, 2012.

Sergio Vera, Frederic Perez, Clara Balust, Ramon Trueba, Jordi Rubió, Raul Calvo, Xavier Mazaira, Anandhan Danasingh, Livia Barazzetti, Mauricio Reyes, Mario Ceresa, Jens Fagertun, Hans Martin Kjer, Rasmus Reinhold Paulsen, and Miguel Ángel González Ballester. *Patient Specific Simulation for Planning of Cochlear Implantation Surgery*, pages 101–108. Lecture Notes in Computer Science. Springer, 2014. ISBN 978-3-319-13908-1. doi: 10.1007/978-3-319-13909-8_13.

F. Wilcoxon. Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6):80–83, 1945.

Blake S. Wilson and Michael F. Dorman. Cochlear implants: A remarkable past and a brilliant future. *Hearing Research*, 242(1-2):3–21, 2008. ISSN 03785955, 18785891. doi: 10.1016/j.heares.2008.06.005.

J. Yao and R.M. Summers. Statistical location model for abdominal organ localization. In *Med Image Comput Comput Assist Interv*, volume 12(Pt 2), pages 9–17, 2009.

P.A. Yushkevich, H. Zhang, and J.C. Gee. Continuous medial representation for anatomical structures. *IEEE Trans. Medical Imaging*, 25(12):1547–64, 2006a.

Paul A. Yushkevich, Joseph Piven, Heather C. Hazlett, Rachel G. Smith, Sean Ho, James C. Gee, and Guido Gerig. User-Guided 3D Active Contour Segmentation of Anatomical Structures: Significantly Improved Efficiency and Reliability. *Neuroimage*, 31(3):1116–1128, 2006b.

PaulA. Yushkevich, Hui Zhang, TonyJ. Simon, and James C. Gee. Structure-specific statistical mapping of white matter tracts. *NeuroImage*, 41(2):448–461, 2008.