Atlas Diffeomorphisms via Object Models
(Late Draft)

Rohit R. Saboo

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Computer Science.

Chapel Hill
2011

Approved by:

Stephen M. Pizer, Advisor

Julian G. Rosenman, Reader

Jack Snoeyink, Reader

Ron Alterovitz, Reader

Marc Niethammer, Reader

© 2011
Rohit R. Saboo
ALL RIGHTS RESERVED

Abstract

Rohit R. Saboo: Atlas Diffeomorphisms via Object Models (Late Draft).
(Under the direction of Stephen M. Pizer.)

To tackle the problem of segmenting several closely-spaced objects from 3D medical images, I propose a hybrid of two segmentation approaches: one image-based and one model-based. A major contribution takes the image-based approach by diffeomorphically mapping a fully segmented atlas image to a partially segmented target patient image preserving any 'correspondence' inferred from the partial segmentation of the target. The mapping is produced by solving the steady-state heat flow equation where the temperature is a coordinate vector and corresponding points have the same temperature. Objects carried over from the atlas into the target serve as reasonable initial segmentations and can be further refined by a model-based segmentation method. Good quality segmentations are added to the list of the initial partial segmentations, and the process is repeated.

Another contribution takes the model-based approach in developing shape models of quasi-tubular objects and statistics on those models. Whereas medial models were previously only developed for slab-shaped objects, this contribution provides an approximately medial method to stably represent nearly tubular objects.

I test my method on segmenting objects from 3D Computed Tomography (CT) scans of the head and neck obtained for radiotherapy treatment planning.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 The Segmentation Challenge

Several real world applications, such as surgical planning and radiotherapy treatment planning, require segmentation, or the outlining of objects in three-dimensional (3D) medical images. Automatic segmentation is a challenging, well-studied problem for which several solutions have been proposed. Most solutions are designed with specific applications in mind and are difficult to generalize.

In my application, which is laid out in detail in the next chapter, we face the challenge of segmenting not just one object but a complex of about 40 closely-spaced objects. Some of these objects have poor contrast with respect to adjacent objects, others are extremely variable, and yet others are nearly invisible on a CT scan, so their locations are implied by other objects.

Many techniques have been developed for segmentation. These techniques can be divided into those that use only local image features and those that bring in some form of prior knowledge into the segmentation process. Techniques such as thresholding and edge detection fall into the former category. This thesis is a novel way of incorporating prior knowledge in the segmentation process. Such segmentation methods can be further divided into two broad classes:

- Image-based approaches:

These methods start from an image of an "average"[1] patient, referred to as the *atlas* image. This image is carefully segmented by experts. When a new patient's image[2] is ready for segmentation, the atlas image is warped into the new image using information such as geometric and intensity features. This warp is used to map the segmentations in the atlas image into the new image. Most existing work optimizes the warp by matching intensity features.

The main problem that these methods suffer from is that this optimization happens in a very high-dimensional space. Thus, almost always, one has to use local, greedy methods to do the optimization so that it finishes in a timely manner. This makes the methods susceptible to image artifacts.

- Model-based approaches:

  These methods use a combination of a geometric model and image intensity features to model a shape and then segment the object from the image.

  Such methods suffer from two problems: First, segmenting one organ at a time does not respect interdependencies between neighboring objects. This can lead to undesirable overlap in segmentation results.

  Second, *all* model-based methods optimize an objective function that is typically composed of some geometric and image features. These objective functions are normally not convex, and the number of parameters involved is quite high. As a result, beginning from an initial estimate, they can get stuck in local minima. The better the initial estimate, the better are the chances of getting an accurate segmentation. These initializations typically involve some manual labor and can be very tedious for a complex of about 40 objects.

[1] The word average is used in a loose sense and could simply be the most typical patient or even a choice from among typical patients from various classes.

[2] The patient image is also called the target image.

Both classes of methods can lead to poor correspondences between matching points on the atlas and the target that can in turn lead to poor segmentations. In the next section, I give an overview of my method, which borrows from both of these classes and can deliver segmentations of a complex of objects. It employs correspondence-preserving warps from partially segmented atlases, which is a geometry-based warping method that does not use image intensities and is therefore not afflicted by some of the problems faced by a warping-based method. It then segments the organs using a model-based method, where most of the work is done by the previous step, thereby reducing the chances of getting stuck in a local minimum or neighboring objects overlapping each other after segmentation. This combination of the two brings out the best of both approaches.

1.2 Overview

Previous warping methods try to infer the correspondence and warp the image at the same time. Correspondence is the process of matching points or regions between two or more images; a good correspondence is defined as one that matches regions from one image to similar regions in another image. The definition of a good similarity metric is contentious, and therefore the definition of good correspondence is not well-defined. Several methods have been proposed to infer correspondence. However, there maybe certain places in an image where what constitutes a good correspondence is easy to define – for example the tip of the nose. In certain situations, we are given correspondence, for example, in the form of anatomical features; in other situations, we need to define a similarity metric and compute correspondence ourselves. These are hard problems in their own right. In my warping method, I do not try to solve them, and instead require that the correspondence be given as input (in the form of a mapping between certain points in the images). The warping method in turn strictly respects the given correspondence. By breaking the process into two steps, one can potentially couple the best performing correspondence method (over a small region of the image) with this new

warping method.

In the next few paragraphs, I give a brief overview of all the requirements of the warping method and how each of them is met. In section 1.5 and fig. 1.1, I describe how this correspondence-preserving warping method is embedded in a larger framework to segment a complex of objects. The method is also presented in the form of an algorithm (1).

Algorithm 1 Hierarchical segmentation algorithm

Require: patient image
Require: shape models
Require: segmented atlas image with labeled surface meshes
  return segmented patient image (for organs with shape models)
  Reference objects list $\leftarrow$ {easily segmentable objects}
  Target objects list $\leftarrow$ {objects to be segmented}
  repeat
    for all object $\in$ Reference objects list **do**
      Generate labeled mesh of object in patient image.
      Compute correspondence between object mesh in patient image and atlas image.
    end for
    Produce a correspondence-preserving warp h between patient image and atlas image.
    for all object $\in$ Target objects list **do**
      warped object $\leftarrow$ apply warp h to object.
      Segment object by initializing segmentation method with warped object.
      if object is segmented well **then**
        Reference objects list $\leftarrow$ Reference objects list $\cup$ {object}
        Target objects list $\leftarrow$ Target objects list $-$ {object}
      end if
    end for
  until Segmentations are satisfactory or no improvement is seen.

The larger framework, which segments a complex of objects, starts with a fully segmented atlas. The atlas is augmented with the shape models of all the objects that we wish to segment. The labeled mesh generator, discussed in section 1.3, partially segments the target to produce surface meshes with labeled vertices. These objects, which are also easily segmented in the target because of high contrast, are called *reference objects*, and the remaining not so easily-segmented objects are called *target objects*.

The atlas image is mapped to the target image in two steps. The surface correspondence step takes surface meshes with labeled vertices and infers correspondence between

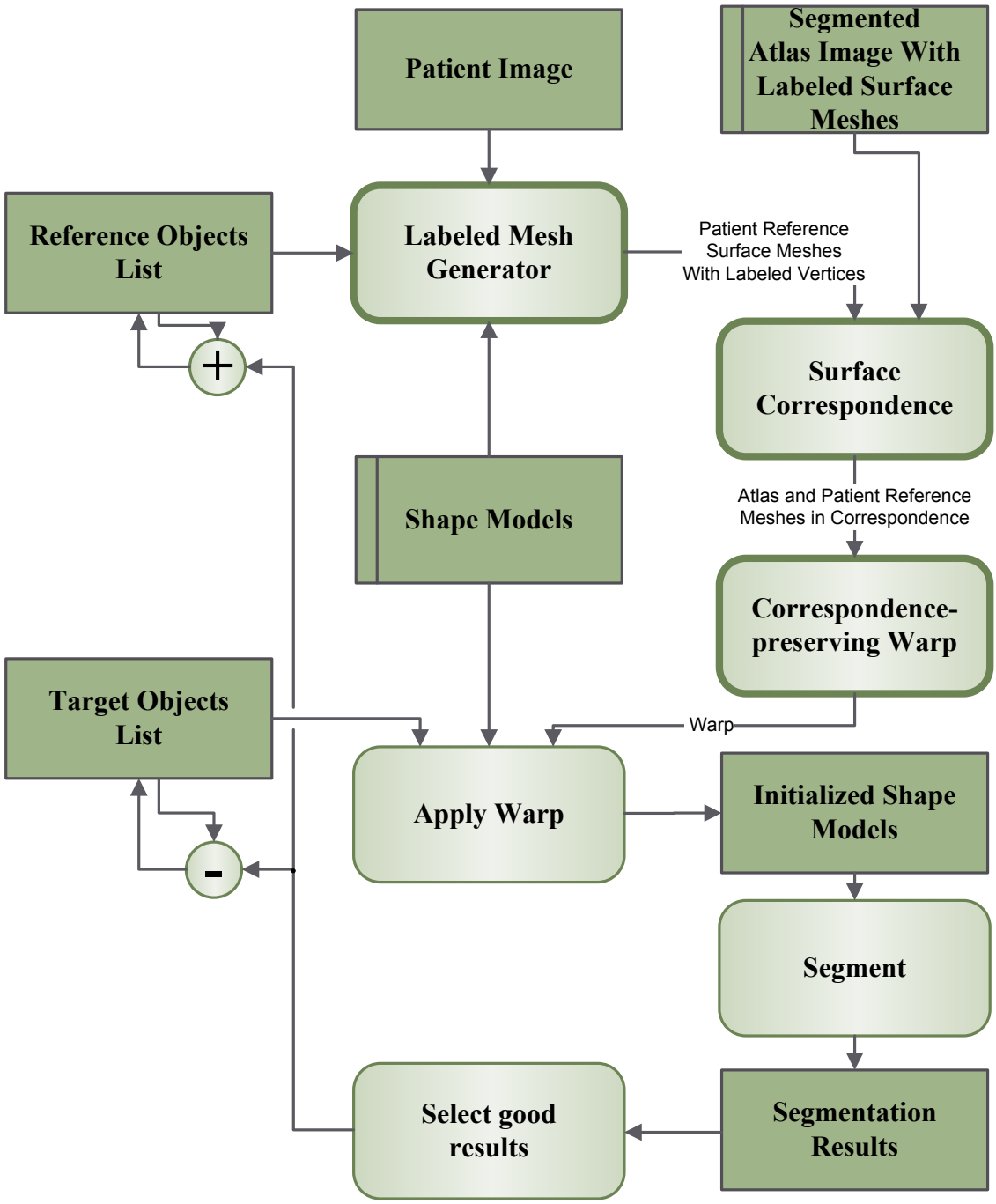


Figure 1.1: An overview of the hierarchical warp method: The labeled mesh generator and the surface correspondence method are shown in fig. 1.2 and 1.3 respectively. The rounded rectangles represent processes, ordinary rectangles are inputs and outputs, and rectangles with a double left edge are precomputed inputs.

repositioned vertices on the reference objects in the atlas and the target. Next, I use these correspondences to drive a warp from the atlas to the target image in my correspondence-preserving warping method; the warp from the correspondence-preserving warping method is based purely on geometric information learned from meshes of the reference objects. These steps are discussed in sections 1.3 and 1.4 respectively.

The resulting warp gives us an initial guess for the segmentation of the target objects. The objects are segmented, and good results are selected. This thesis does not discuss the method of model-based segmentation; Broadhurst [4] and Stough [32] provide details on a particularly efficacious appearance model. The good results, segmentations that appear to match well with the image, are added to the list of the reference objects and removed from the list of the target objects. This loop is called the hierarchical warp method.

1.3 Overview of the Correspondence Method

The inputs and outputs for the correspondence method are as follows:

- Inputs – smooth segmentations of the reference objects in both atlas and target patient images;
- Outputs – correspondence labels on surface points on the same segmentations in both the atlas and the target patient images.

The correspondence is computed object-by-object for each of the objects in the reference set. The better the correspondences of these objects, the better the result from the warp. I compute correspondence by two methods that have worked well for our applications. The user can substitute a method of their liking if they do not prefer these methods for correspondence.

An approximate initial correspondence for any given anatomical object is obtained by modeling the reference objects for it in each of the cases as shape models. This step, shown in fig. 1.2, generates a labeled mesh. The correspondence implied by this labeled mesh is

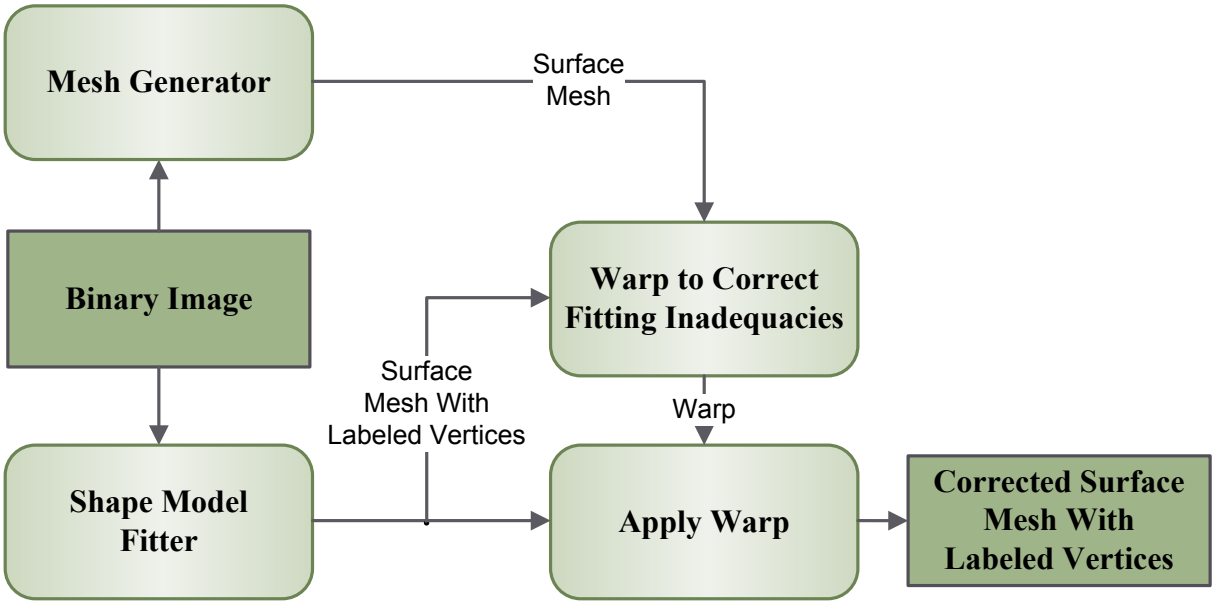


Figure 1.2: An overview of the method that generates labeled meshes from binary images.

not accurate enough for the correspondence-preserving warping method and is improved by an entropy-based correspondence method, shown in fig. 1.3, that moves the individual points on the surface. The entropy-based correspondence method can sometimes get stuck in local minima and initializing it in this manner can yield improved results.

The entropy-based correspondence method requires the computation of curvatures on the surface of objects. To robustly compute surface curvatures, any aliasing artifacts present in the mesh should be removed. A mesh extracted from a binary image is always aliased and needs to be anti-aliased. This necessitates the use of an anti-aliasing procedure on the mesh. I have developed a method that anti-aliases such meshes by a a fourth-order Laplacian of curvature flow.

The reference objects are represented by shape models, m-reps, described by Pizer et al. [28]. Since m-reps attempt to model an object at a larger scale, they may not match the contours of the object precisely. This is corrected by a warping method developed by Yezzi et al. [36]; it warps the surface of the model into the correct surface, carrying over the correspondence implied by the medial models. This correspondence is not accurate; it is improved

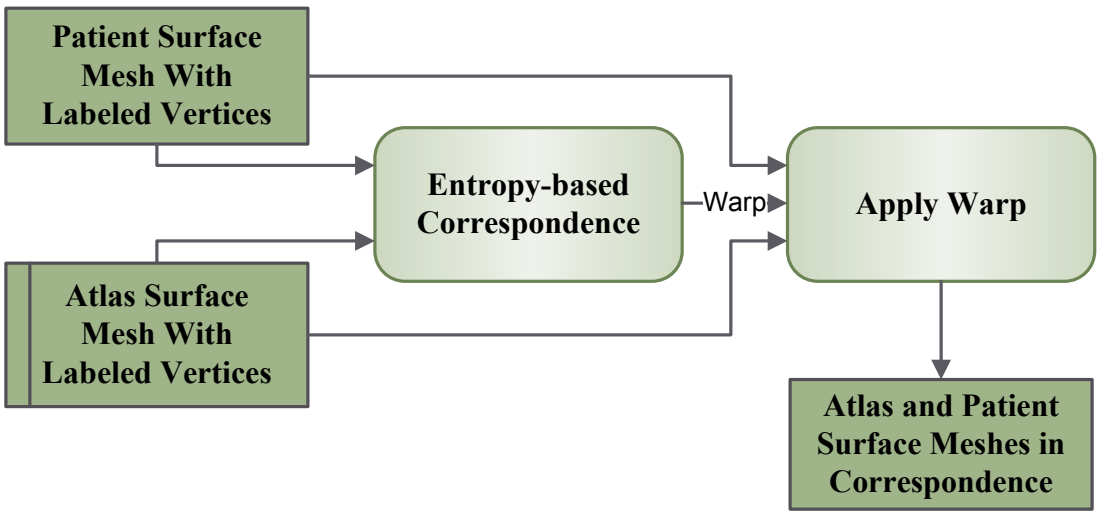


Figure 1.3: An overview of the entropy-based correspondence system that takes in labeled meshes from the atlas and the patient and returns meshes in correspondence.

by the use of an entropy-based correspondence algorithm developed by Oguz et al. [27].

Some objects in the body are tube-like and others are slab-like. Existing work on m-reps can only represent slab-like objects. We need similar abilities of modeling and computing correspondence between tube-like objects that we have for slab-like objects. A method that can model tube-like objects that is consistent with the manner in which slab-like objects can be modeled has been developed, as described in chapter 4.

1.4 Correspondence-preserving Warping Method Overview

The inputs and outputs for the correspondence-preserving warping method are as follows:

- Inputs – segmentations (in the form of boundary meshes) of the reference objects in both the atlas and the target patient's images with correspondence labels;
- Output – a mapping between the atlas and the target patient's image.

At the heart of the warping method, corresponding points between the atlas and the target image are assigned the same tuple of temperatures; a distinct temperature value is assigned

for each dimension, and they are all independent of one another. This is imposed as a boundary condition in the target image; in addition, another set of boundary conditions is imposed at the actual boundaries of the target image. With these boundary conditions, we solve the steady-state heat equation. As a result of solving this equation, a point in the atlas will have a corresponding point in the target image with the same set of temperatures. This correspondence is used to map points between the atlas and the target images.

We also let the heat conduction coefficient vary across the image while solving the heat equation. The precise way in which this is done will be detailed later. Doing so ensures that the mapping between the two images is smooth and is devoid of folds. I will refer to such a mapping as *almost diffeomorphic*.

In the next section, I show how this method is embedded in a larger framework that can segment objects.

1.5 Hierarchical Warp

The inputs and outputs for the hierarchical warping method are as follows:

- Inputs – the target patient's image and a fully segmented atlas image augmented with shape models of the objects;
- Output – a fully segmented target image.

We first identify high-contrast objects that can be easily segmented, such as air cavities and bones. In the image from the target patient, we first segment these easily identifiable objects. These objects form the initial set of the reference objects.

The method described above is then used to infer a warp between the atlas and the target patient. At a result of its application, we have accurate initial guesses for the location of each of the objects.

These guesses are used to drive a model-based optimization process that uses both geometric and image features and tries to segment the individual target objects in this complex.

The segmentations results are automatically evaluated and some of them are now chosen as new reference objects and the whole process is repeated again.

In fig. 1.1, I have presented a graphical overview of this hierarchical segmentation approach.

1.6 Thesis and Claims

Thesis: *Atlas-based segmentation methods that use correspondence produce superior results on a complex of interrelated objects than those that do not. Methods that establish good correspondence, when combined with methods that produce a warp using the steady-state heat equation, efficaciously provide the required correspondence-based warps leading to good segmentations.*

The contributions of this dissertation are the following:

1. The dissertation develops a method, based on heat flow in 4D space, to interpolate between the boundaries of pairs of objects while respecting correspondence. (This is different from the method sketched in section 1.4 and claim 2 below.)

2. Interpolation of objects while warping is not always required; most of the time, we only care about the end result of the warp. My warp method, based on heat flow in 3D space, is guaranteed to produce a smooth, non-folded warp between the two images while respecting positional correspondences on object surfaces; it has been developed out of a desire to provide better guarantees and efficiency than the method described under claim 1.

3. Modeling nearly tubular objects by conventional medial models is not stable. My method models objects that are nearly tubular in shape, giving reasonable correspondence and stable probability distributions on their shape spaces.

4. Shape models such as m-reps do not match the actual boundary precisely because they

try to model the object at a large scale. This can be a problem when they are used as the basis of a warping mechanism. My method based on heat flows improves the correspondence by moving the points to the actual surface when using shape models, specifically m-reps.

5. The correspondence implied by m-reps may not be good enough for correspondence-based warping methods. Even after the points have been moved to the correct surface, they may still need to slide along the surface. My method based on Oguz et al. [27] improves this correspondence.

6. The head and neck area is a challenge for 3D image segmentation. My framework, using the methods developed, can deliver reasonable initializations for segmentation of certain objects (masseter muscle, sternocleidomastoid muscle, and thyroid) in 3D head and neck CT images.

7. Binary images can be efficaciously anti-aliased by a fourth-order Laplacian of curvature flow.

1.7 Overview of Chapters

This chapter provided a brief overview of correspondence-preserving warping methods, their prerequisites, and the manner in which they can be embedded in a larger framework to produce quality segmentations.

In chapter 2, I present the medical problem, head and neck segmentation, that drove the development of this method. I also discuss how existing methods may not live up to challenges presented to us.

Chapter 3 gets the reader up to speed on concepts that include m-reps, thin plate splines, and existing work on registration using partial differentiation equations.

A lot of work had been done on medial models of slab-shaped objects; objects that are

nearly tubular-shaped cannot be modeled effectively by the same methods. In chapter 4, I discuss how to model nearly tubular objects and represent populations of them, and in doing so realize claim 3. These medial models provide an initial approximate correspondence between objects in the atlas image and the patient image.

Chapter 5 discusses the two types of correspondence-enhancing methods: moving points from the m-rep-implied surface to the correct surface, and moving points along the surface. This chapter establishes claims 4 and 5.

In chapter 6 I develop two methods to warp a volume into another volume by respecting correspondence between matching surfaces: The first one solves a heat equation in 4 dimensions; it warps and produces interpolations between the two volumes. This establishes claim 1. The second method solves a heat equation in 3 dimensions; it produces a warp between the two volumes. In doing so, claim 2 is established.

Chapter 7 shows our results on the application problem, and in doing so I establish claim 6.

There are a few other ways in which the framework can be designed. In chapter 8, I present a few of these pipelines and show a few preliminary segmentation results on the same data set.

Finally, chapter 9 concludes with a summary of this dissertation, discusses the contributions, and presents possible future directions.

Chapter 2

Application

In section 2.1, I present the medical problem of head and neck cancer. which provides the main driving problem for this dissertation. Section 2.2 follows with a discussion of why segmentation for treatment planning in head and neck cancer is challenging.

2.1 Treatment of Head and Neck Cancer

There are approximately 40,000 cases of cancers of the head and neck diagnosed each year in the United States. These tumors are usually sensitive to radiation and chemotherapy, and are often treated with both modalities. Several objects in the head and neck are at risk from receiving too much radiation; the parotid (salivary) glands are particularly sensitive. For these patients the major morbidity (treatment complication) is xerostomia, or dry mouth, which can be permanent. Xerostomia can lead to dental problems and poor nutrition, and thus a substantial decrease in the quality of the patient's life.

A major breakthrough in the treatment with radiation therapy occurred about 10 years ago when it was learned how to redirect radiation dose within a patient by physically modulating the radiation beams. This technique, now called intensity modulated radiation therapy, or IMRT, can reduce the radiation dose given to the objects at risk during radiation therapy for head and neck cancer. For the radiation oncologist, the physician overseeing the treatment, the development of an IMRT treatment plan is an important step.

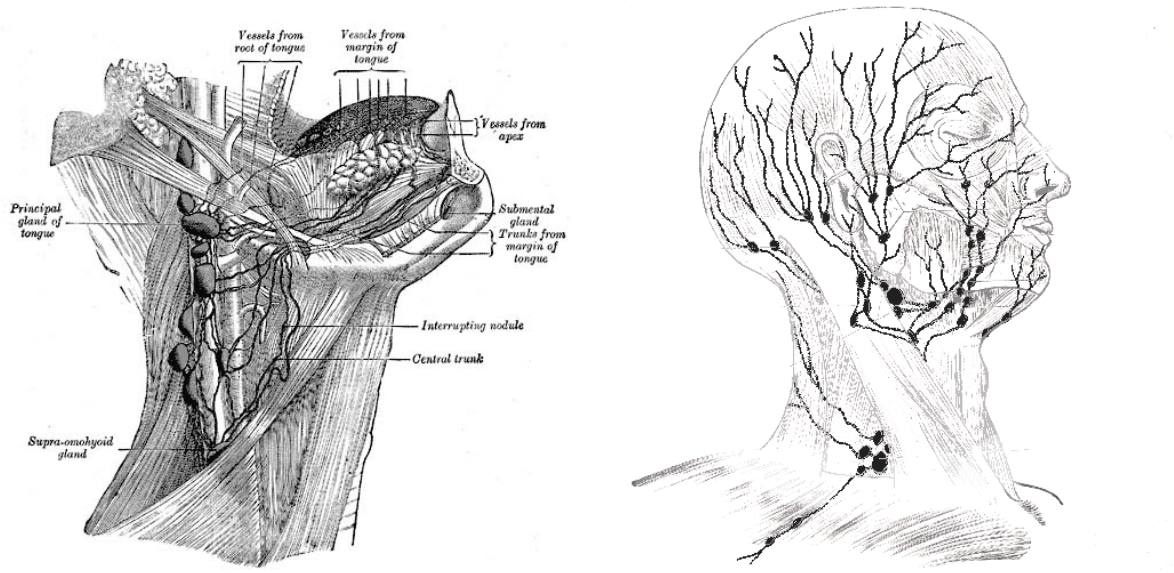


Figure 2.1: Lymphatics of the head and neck (obtained from Gray's anatomy 1918)

The development of the treatment plan is a laborious process. First, dozens of normal objects within the head and neck need to be identified and carefully outlined on each CT slice. This needs to be done because the use of IMRT to reduce the dose to the salivary glands will necessarily increase the dose elsewhere, and the dose tolerance of other organs such as the eyes, spinal cord, larynx, and mandible (jawbone) must be respected. In addition, the salivary glands themselves must be identified. Next, the gross tumor volume (GTV) as seen on the CT scan, or felt during the patient examination, is outlined on each CT slice. But the GTV is not the entire target of the treatment. Head and neck cancer is notorious for spreading invisibly into neighboring objects and down lymphatic pathways in complicated ways. This spread is microscopic and therefore cannot be seen, but if left untreated it will often serve as a nidus of relapse. Determining this larger clinical target volume, or CTV, which includes the GTV plus the microscopic spread is a major task for the radiation oncologist.

Fig. 2.1 shows a schematic of the lymphatic pathways in the head and neck, but as it is only a two-dimensional drawing, the full complexity of the lymphatic pathways (and thus the potential CTV) is not appreciated. To simplify the situation, lymph node chains are collected into groups called levels depending on their surgical accessibility and susceptibility to tumor

Figure 2.2: Three different nodal levels on an axial slice from a head and neck CT scan

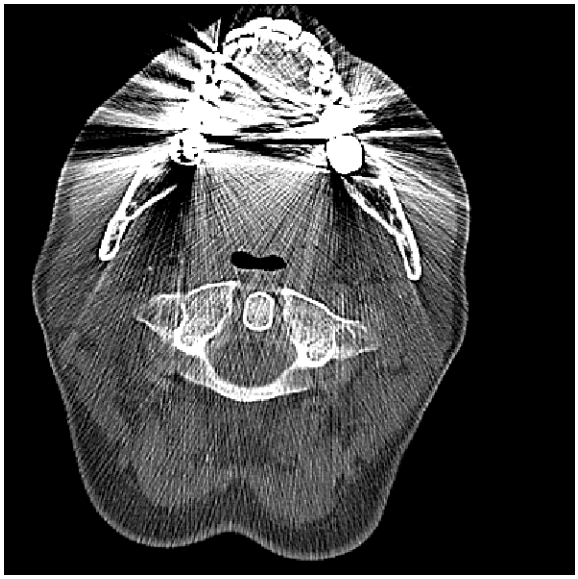


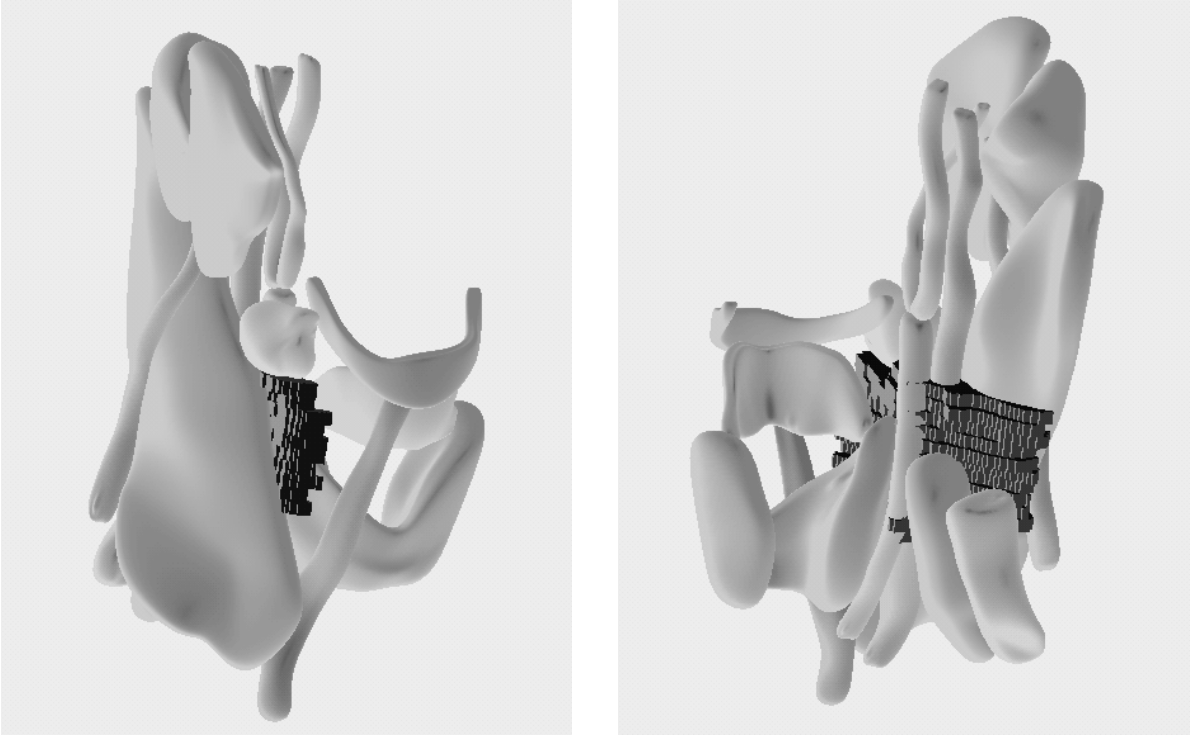
Figure 2.3: Metal streak artifacts in a head and neck CT image

spread from different tumor sites. Figure 2.2 illustrates the distribution of some of the levels on a given CT slice. Drawing these nodal levels on each slice according to the official rules [14] is very difficult, because understanding the lymph regions depends on using its relationship to a variety of neighboring objects. Thus, segmentation of the lymph regions requires the segmentation of at least twenty anatomic objects on each side; typically it is many more than that. Some of these objects are modeled and shown in fig. 2.4. Also shown, is the Level-III nodal region in the form of a block of charcoal-coloured voxels.

In the next section, we discuss the challenges associated with segmenting the head and neck CT images.

2.2 Segmentation of the Head and Neck

In the head and neck region, several closely-spaced objects determine the CTV and objects-at-risk; fig. 2.4 show some of these closely-spaced objects. Several of these objects, such as muscles, are composed entirely of soft tissue. As a result, there is very poor contrast between



(a) Anterior view (b) Posterior view

Figure 2.4: Some of the segmented head and neck organs with the Level-III nodal region

neighboring objects; this can be seen in the axial slice in fig. 2.2 where the bones are easily identifiable, but all the other objects seem to have similar intensities. For any segmentation method utilizing image intensities, this could result in inaccurate segmentations.

Many patients who have head and neck cancer have metal fillings in their teeth. In a CT scan, fillings can create metal streak artifacts, such as those shown in fig. 2.3. This further degrades the image quality and presents another challenge for intensity-based methods. Though there exist several methods to mitigate these artifacts, we do not have much information on whether the existing segmentation methods will succeed on the improved images. I do not have to remove these artifacts from any images as the method presented in this thesis is based only on geometry and not image intensities.

Chapter 1 pointed out that there are two main classes of methods: image-based and model-based. All image-based techniques warp an atlas image. They typically require multiple

atlases since a single atlas may not give accurate segmentations in all cases. The user has to pick an atlas similar anatomy to the patient, a non-trivial task. Even so, their results frequently require the user to edit the final result because these methods perform poorly in the presence of little contrast between neighboring objects, image artifacts, and noise.

Model-based methods can perform more stably when the image quality is not good, but they need a good initial guess for the shape and position of the model.

Therefore, I chose to use a hybrid method that exploits the advantages offered by each class of methods: The hybrid technique starts by warping a fully segmented atlas image into the image of the target patient based on purely geometric cues. This carries over the segmentations into the target image and serves as the initial guess for the model-based segmentation.

Chapter 3

Background

In this chapter, I will present background material on m-reps, thin plate splines, fluid flow registration (warping), and steady-state heat flow that will be helpful in understanding the detailed discussion of the components of my method presented in later chapters. Discrete medial models (m-reps) can model anatomical objects well and provide an object-centric coordinate system, which is necessary for my correspondence-based warping method. Section 3.1 discusses discrete medial models for slab-like objects. Section 3.2 presents some prior work on image registration. In my correspondence-preserving warping method for image registration, I solve the steady-state heat equation; section 3.3 presents some prior work done with heat flow and techniques for speedy convergence of the system.

3.1 M-reps

In this section, I will talk about m-reps for slab-shaped objects. M-reps, which are a specialization of the Blum medial axis representation, model a population of objects and provide correspondence between different members of the population. They were originally designed for slab-shaped objects and are generalized to nearly tubular objects in chapter 4.

Any anatomic object can be approximated by a small collection of non-branching medial manifolds (axes), called *figures*. Each figure is composed of a continuum of 'hubs' symmetrically placed with respect to the object boundary. From each hub there extend two equal length

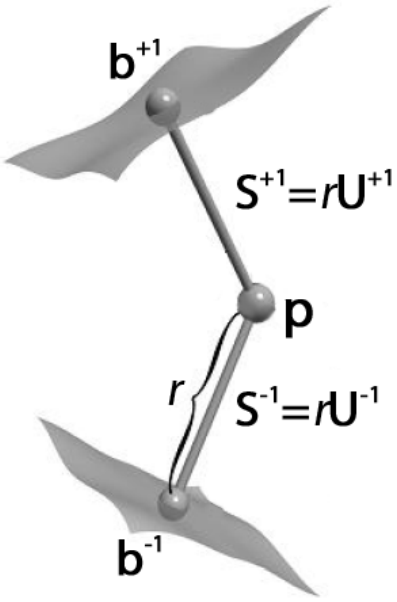


Figure 3.1: A medial atom.

Figure 3.2: An m-rep model of the medial scalene muscle showing the grid of medial atoms. The atoms at the edges of the grid are special – they are coalesced from an ordinary atom and one that is on the edge of the continuous medial sheet.

vectors called *spokes* that end at and are approximately orthogonal to two implied boundary points. The hub and its two associated spokes are known as a *medial atom*. The medial atom, $\mathbf{m}_{ij}$, shown in fig. 3.1, is represented by the four-tuple $(\mathbf{x}, r, \mathbf{U}^{-1}, \mathbf{U}^{+1})$, where $\mathbf{x} \in \mathbb{R}^3$ is the position of the hub, $r \in \mathbb{R}^+$ is the common spoke length, and $\mathbf{U}^{-1}, \mathbf{U}^{+1} \in \mathbb{S}^2$ are the two spoke directions. These two spoke directions give us two opposing points on the surface $\mathbf{b}^{-1}$ and $\mathbf{b}^{+1}$, known as the *implied boundary points*. $\mathbf{U}^{-1}$ and $\mathbf{U}^{+1}$ are normal to the surface at these points.

A discrete m-rep is formed by sampling the medial manifold over a spatially regular grid. In the implementation I used, for ordinary three-dimensional (3D) objects the grid (fig. 3.2) is

a rectangular mesh of medial atoms $\mathbf{m}_{ij}$, where $i \in [1, m]$ and $j \in [1, n]$. The curious reader is referred to Pizer et al. [28] for a much more detailed discussion on m-reps.

Section 3.1.1 presents the means to statistically analyze m-rep populations. This statistical analysis is useful in fitting m-reps to objects and provide correspondence across the population. Section 3.1.2 presents the mechanism in which m-rep models are fitted to anatomical objects. Finally, section 3.1.3 discusses the object-centric m-rep coordinate system, which is useful for my correspondence-based warping method.

3.1.1 M-rep Shape Space

As described in chapter 8 of the medial book [31], m-reps lie in a *Riemannian symmetric shape* space, where every point in this space is an m-rep. Moving along any smooth curve in this space yields a continuously varying m-rep model. Each point in this space is also associated with a tangent space $T_x(M)$ and a *Riemannian metric*, a smoothly varying inner product on this tangent space. Given a tangent plane, the operation that projects an m-rep model on this tangent plane is called a *logarithmic map*, and the operation that projects the model back to the shape space is called an *exponential map.*

Populations of m-rep models are analyzed using Principal Geodesic Analysis (PGA), developed by Fletcher et al. [12]. PGA is a variant of Principal Component Analysis (PCA) that is suited to non-linear spaces. In PGA, m-reps are projected in a metric-preserving way to the tangent space of the Fréchet mean of the population samples. PCA is done in the tangent space, and the principal directions are projected back down on the Riemannian shape space. This analysis yields a mean model and a probability distribution on the shape space for more likely models.

3.1.2 M-rep Training

Training is the process of varying the parameters of the m-rep model so that the implied boundary and the expertly outlined boundary match each other.[1] This process yields models with two desirable characteristics: First, they fit well – they match closely with the expertly outlined boundary; and second, they correspond well with models of other members of the population – the variation in parameters of the models across the population is low. This process, outlined in Merck et al. [25], is briefly presented in this section.

This process starts with the expert manually segmenting a few dozen training images. Sometimes, a set of landmarks is identified on the object; these are typically places on the object that can be easily identified across the entire population. Next, the parameters of the m-rep model are varied inside an optimizer that minimizes a weighted sum of the distances between the m-rep-implied boundary and the actual segmentation, the distances between the landmarks and the corresponding positions on the m-rep surface, and penalties due to boundary or medial surface roughness. The resulting models are called *training models*.

These training models are then iteratively refined. Next, they are statistically analyzed: The modes of variation and a mean are computed. The fitting is repeated for each of the training images by starting with this mean model as the initial model and modifying the model within the shape space to optimize the fit, possibly refining the result. This iterative process yields models that fit well (have a small average distance between the surface of the fitted model and the training image) and correspond well with models of other members of the population (have small standard deviations of the probability distribution of the model parameters).

[1] An accurate description of this method is parameter fitting, but I'll call it training to be consistent with prior literature.

3.1.3 Figural Coordinates and Inter-object Correspondence

All points in the interior of a medially-implied object can be uniquely identified by a set of four coordinates: one running along the major dimension of the medial sheet, another running along the minor dimension, the third indicating the side of the medial sheet and enabling special handling at the edges, and the last one running from the medial sheet towards the surface. This coordinate system determines corresponding positions in the volume of two deformed versions of the model.

The m-rep mesh is considered to represent a continuous sheet of atoms with non-crossing spokes, which imply a closed figural boundary surface. This continuous sheet is obtained by interpolating the grid and is parameterized by $(u, v) \in [(j, j+1) \times (k, k+1)]$ for the part of the mesh with the atom $\mathbf{m}_{jk}$ at its lower left corner.

Methods for atom interpolation are described in Thall [34] and Han [16]; I use the method described by Thall. The result of this interpolation is that every surface point can be assigned a figural coordinate (u, v, ϕ). When ϕ is $-\pi/2$ or $+\pi/2$, it indicates the side. A value of $\phi \in (-\pi/2, +\pi/2)$ indicates a position around the object's crest, whose medial loci form the boundary of the medial grid. The parameters (u, v) are taken from the interpolated atom which corresponds to this surface point. The subdivision is done finely enough that the surface can be represented by a set of tiles with diameters less than the length of all sides of a voxel.

These coordinates can be extended to give unique, object-relative coordinates for the whole region interior to the implied boundary. For this purpose every point (x, y, z) in this region is assigned a figural coordinate (u, v, ϕ, τ), where τ is a measure of distance from the hub along the spoke identified by (u, v, ϕ) as a fraction of the spoke length r. τ is 0 at the hub and 1 at the implied boundary.

Not only do figural coordinates provide a natural way to parametrize the object surface, but also these coordinates yield a natural correspondence between different variants of the same figure. As the model deforms, i.e., its parameters change, the surface points change their physical position in (x, y, z) space. However, they carry over the same figural coordi-

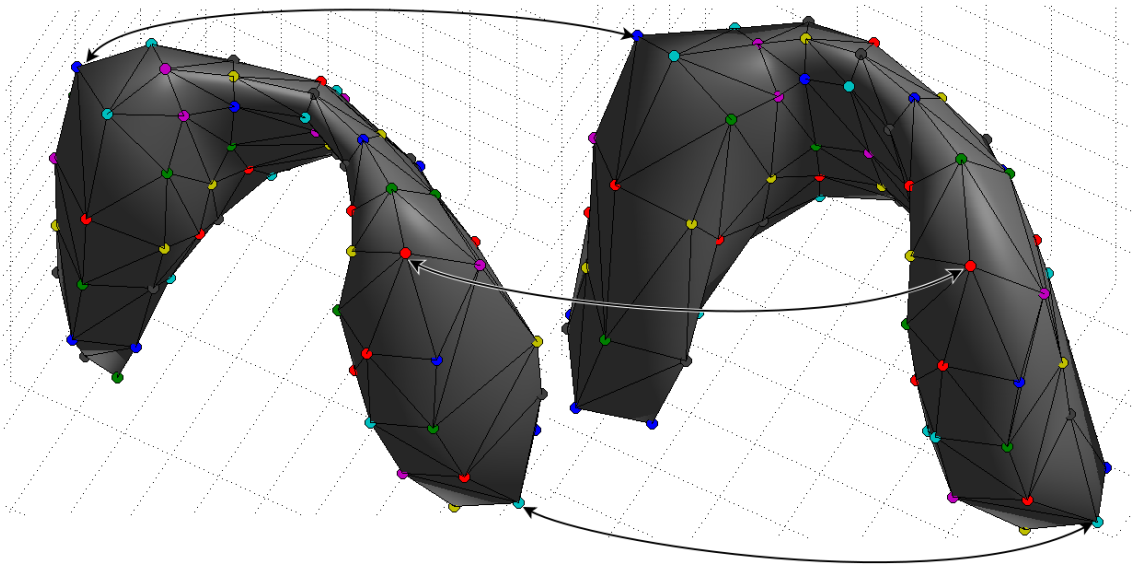


Figure 3.3: Correspondence between points on the surface of two different thyroids.

nates. Consider, for example, thyroids from two different patients. If an m-rep is fitted to one thyroid and then deformed to fit the second thyroid, then the corresponding positions on the two thyroids can be easily defined using this figural coordinate system. Fig 3.3 shows the correspondence for points on the surface of a thyroid. *Thus, the m-reps and their figural coordinates provide a set of corresponding points.*

It is important to refine these correspondences through the entropy-based correspondence step because these correspondences are only approximate. According to one school of thought, good correspondence is characterized by tighter probability distributions of the model parameters. The correspondence inferred by m-reps suffers because they provide only approximate and large-scale fits to the object, and the fitting process is primarily geometric with very few image features.

3.2 Registration Methods

Registration methods can be classified along two axes: first, whether they produce the mappings through displacements (elastic) or integration of velocity fields (fluid); second, whether

they match landmarks, image voxels, or a combination of both through a similarity metric.

By regularizing the displacement or velocity field, these registrations yield a smooth spatial warp. Methods that regularize a displacement field typically do not produce diffeomorphisms, which are one-to-one, smooth, and smoothly invertible mappings; methods that regularize velocity fields, however, most often produce diffeomorphisms.

Image-based registration techniques optimize a function that typically sums data mismatch and mapping irregularity terms. If the data mismatch term includes landmark mismatch, the optimizer will regularize the warp and match the image at the expense of the landmark match. Registration methods that constrain landmarks to match exactly include thin plate splines. Unfortunately thin-plate-spline-based methods do not guarantee diffeomorphisms.

The method of thin plate splines, developed by Bookstein et al. [3], is one of the popular methods that matches landmarks and regularizes displacements; it is discussed in section 3.2.1. It is an important part of another registration method discussed in section 3.2.2, which matches landmark features and regularizes velocity fields; this method comes close to solving the issue of producing correspondence-preserving warps and is therefore of interest here. Of the methods that match image features and regularize velocity fields, the most popular is fluid flow, developed by Joshi et al. [22]; I will not discuss fluid flow as our images have poor image intensity information and therefore fluid flow is not well suited.

3.2.1 Thin Plate Splines

Given a set of landmark points $\mathbf{P}_i$, each with a directional displacement $\mathbf{h}_i$, the thin plate spline $\mathbf{f}(\mathbf{x})$ produces a smooth interpolation of the surface which passes through these landmark points.

The Frobenius norm, which sums the squares of all second-order derivatives of the vector-valued displacement, is a reasonable measure of smoothness of the displacement field. Minimizing this yields a displacement field that is quadratically smooth. Incidentally, this smoothness measure is similar in form to the bending energy associated with a thin steel plate lying

flat on the x, y plane subjected to a small deformation $f(x, y)$ in the z-direction. The elastic energy or the bending energy E stored in such a configuration is given by

$$\iint_{\mathbb{R}^2} \left(\left| \frac{\partial^2 f}{\partial x^2} \right|^2 + 2 \left| \frac{\partial^2 f}{\partial x \partial y} \right|^2 + \left| \frac{\partial^2 f}{\partial y^2} \right|^2 \right) \mathrm{d}x\mathrm{d}y. \tag{3.1}$$

If the steel plate is subjected to a fixed displacement at certain points, i.e., the landmarks $\mathbf{P}_i$, the rest of the plate takes a form such that this bending energy is minimized. Thus, thin plate splines yield the quadratically smoothest possible displacements.

In order to produce a mapping in three dimensions, instead of considering the displacement to be scalar-valued and perpendicular to the '3D plate', consider it to lie within its 'plane'. Thus, we now have a vector-valued displacement function $\mathbf{f}(x, y, z) = (f_x(x, y, z), f_y(x, y, z), f_z(x, y, z))$ at each point (x, y, z) on the plate. The displacement field induced by this function is represented in the following form:

$$(x, y, z) \rightarrow (f_x(x, y, z), f_y(x, y, z), f_z(x, y, z)). \tag{3.2}$$

In thin plate splines, this displacement field minimizes the Frobenius norm of each of f_x, f_y, and f_z.

Applying this idea of displacements to a 3D image, the new warped image is given by

$$I_{warped}(x, y, z) = I(x + f_x(x, y, z), y + f_y(x, y, z), z + f_z(x, y, z)). \tag{3.3}$$

However, if the displacements are large, the result may fold onto itself (i.e., it may be non-diffeomorphic). The following example demonstrates this problem: I matched a set of landmarks between two different patient images and registered them with thin plate splines, which resulted in folded level III lymph levels (shown in green in fig. 3.4). Despite this problem in maintaining diffeomorphicity, the biharmonic thin plate spline operator is an important component in the method described in the next section that guarantees diffeomorphisms.

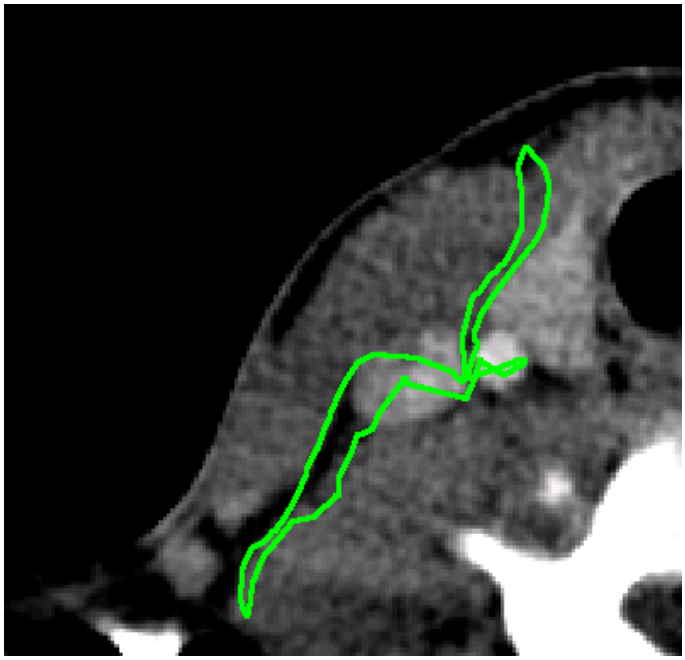


Figure 3.4: A 2D slice from a 3D CT scan in the application of thin plate spline registration between two different patients: The level III lymph levels carried over in the registration are drawn in green. Notice that they are folded.

3.2.2 Large Deformation Diffeomorphisms

As just stated, a major problem of the thin plate spline registration method is that it does not necessarily generate a diffeomorphism. (Other spline methods also suffer from the same problem.) When the deformations between the atlas and the image under analysis are large and curved, the transformations introduce folding and do not preserve the topology of the atlas. The diffeomorphic landmark mapping framework developed by Joshi and Miller [22] can overcome this.

This framework defines a time-indexed transformation $\mathbf{h}(\mathbf{x}, t)$ mapping the atlas to the target by integrating a velocity vector field:

$$\mathbf{h}(\mathbf{x}, t) = \mathbf{x} + \int_0^t \mathbf{v}(\mathbf{h}(\mathbf{x}, t), t)\, \mathrm{d}t. \tag{3.4}$$

Given a set of landmark points $\mathbf{x}_i$ with paths $\mathbf{h}(\mathbf{x}_i, t)$ describing their motion from the atlas to the target and a linear differential operator L that provides a measure of smoothness

of a velocity field, the diffeomorphic transformation is given by the velocity vector field from the minimization:

$$\hat{\mathbf{v}}(\mathbf{x},t) = \arg\min_{\mathbf{v}} \int_0^1 ||L\mathbf{v}(\mathbf{x},t)||^2 \,\mathrm{d}t$$
$$\text{subject to: } \mathbf{v}(\mathbf{h}(\mathbf{x}_i,t)) = \frac{\mathrm{d}\mathbf{h}(\mathbf{x}_i,t)}{\mathrm{d}t}. \tag{3.5}$$

Many linear differential operators have been used in the literature; the biharmonic thin plate spline (TPS) operator is the most common because it has a closed-form solution. Given the complete space-time paths of the landmark points, the closed-form solution for the velocity fields is given as a combination of an affine motion $A\mathbf{x} + T$ and a weighted superposition of Green's functions of the differential operator $LL^\dagger$, where $L^\dagger$ is the adjoint of the linear differential operator L, with weights $\beta_i(t)$:

$$\mathbf{v}(\mathbf{x},t) = \sum_{i=1}^{N} \beta_i(t) K(\mathbf{h}(\mathbf{x}_i,t),\mathbf{x}) + A\mathbf{x} + T. \tag{3.6}$$

The weights $\beta_i(t)$ and the affine motion $A\mathbf{x} + T$ are chosen so that the velocity field satisfies the set of constraints $\mathbf{v}(\mathbf{h}(\mathbf{x}_i,t)) = \mathrm{d}\mathbf{h}(\mathbf{x}_i,t)/\mathrm{d}t$ and any additional boundary conditions. In three dimensions, using the TPS operator with zero boundary condition at infinity results in a Green's function $K(\mathbf{x},\mathbf{y})$ given by $1/|\mathbf{x}-\mathbf{y}|$ and the conditions

$$\begin{aligned} \sum_{i=1}^{N} \beta_i(t) &= 0 \text{ and} \\ \sum_{i=1}^{N} \beta_i(t)\mathbf{h}(\mathbf{x}_i) &= 0. \end{aligned}$$

For each time t, the weights $\beta_i(t)$ can be estimated by a solution of a linear system of equations. For details see [22].

Discrete Integration of Velocity Fields.

For a practical computer implementation the continuous integral of the velocity vector fields is discretized in time as follows: Let $t_j, j = 0, \cdots, M$ be a discretization of the interval $[0, 1]$; then, given a velocity field $\mathbf{v}(\mathbf{x}, t)$, the integral in equation 3.4 can be written recursively as

$$\mathbf{h}(\mathbf{x}, t_{j+1}) = \mathbf{h}(\mathbf{x}, t_j) + \int_{t_j}^{t_{j+1}} \mathbf{v}(\mathbf{h}(\mathbf{x}, t), t)\, \mathrm{d}t$$
$$\text{with } t_j = \frac{j}{M}, \quad j = 0, \cdots, M.$$

The integral above is approximated, resulting in

$$\mathbf{h}(\mathbf{x}, t_{j+1}) = \mathbf{h}(\mathbf{x}, t_j) + \frac{1}{M}\mathbf{v}(\mathbf{h}(\mathbf{x}, t_j), t_j), \tag{3.7}$$

with the initial condition $\mathbf{h}(\mathbf{x}, 0) = \mathbf{x}$. Using this discretization, the generation of a large deformation diffeomorphic transformation can be seen as a repeated application of a TPS interpolation. Substituting equation 3.6 into equation 3.7 yields a solution of a sequence of thin plate spline problems:

$$\frac{\mathbf{h}(\mathbf{x}, t_{j+1}) - \mathbf{h}(\mathbf{x}, t_j)}{1/M} = \sum_{i=1}^{N} \left(\beta_i(t_t) K(\mathbf{h}(\mathbf{x}_i, t_j), \mathbf{x}) + A_j \mathbf{x} + T_j\right). \tag{3.8}$$

This can be interpreted as a repeated application of a spline interpolation for incremental small motions of the landmark along the given paths.

The possibly curved landmark paths need to be provided through some mechanism. The method detailed above would yield a diffeomorphic mapping if the provided landmark paths do not cross each other. Generating landmark paths that are guaranteed to not cross is a non-trivial task; it is the Achilles' heel of this method.

3.3 Steady-state Heat Flow

Is a diffeomorphism really what we want? It may be too restrictive and a smooth, non-folding warp would suffice. My method constructs a steady-state heat flow problem whose solution yields the necessary landmark-matching smooth, non-folding warp. Heat flow was first used by Jones et al. [21] to compute cortical thickness. Yezzi et al. [36] improved the method and used it to compute tissue thickness. Dinh et al. [11] interpolated 3D objects by solving a heat flow equation. The steady-state heat flow partial differential equation takes the form of Laplace's equation. Section 3.3.1 presents a little background on the Laplace's equation, and section 3.3.2 talks about techniques to solve this equation.

3.3.1 Laplace's Equation

Laplace's equation is widely used in physics to solve gravitational, electrodynamic, thermodynamic, fluid flow, and other systems. It is a second-order partial differential equation of a scalar field ϕ, written using the Laplacian operator Δ:

$$\Delta\phi = \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2} = 0. \tag{3.9}$$

This equation is typically accompanied by a fixed value for the scalar field on some boundaries. Solutions of this equation are called *harmonic functions* in mathematics and *potential functions* in physics.

If the boundary conditions are represented as two surfaces with two different potentials, the equi-potential surfaces can be seen as interpolated versions of these two surfaces. Further, flowing along the flux lines, also called *streamlines*, correspondence can be inferred between these two surfaces. Fig. 3.5 shows these equi-potential surfaces and streamlines for a sample problem. This property was first exploited by Jones et al. [21] to compute cortical thickness. However, the correspondence inferred in this way has no anatomical bearing, and in this form there is no way to enforce correspondence between certain points.

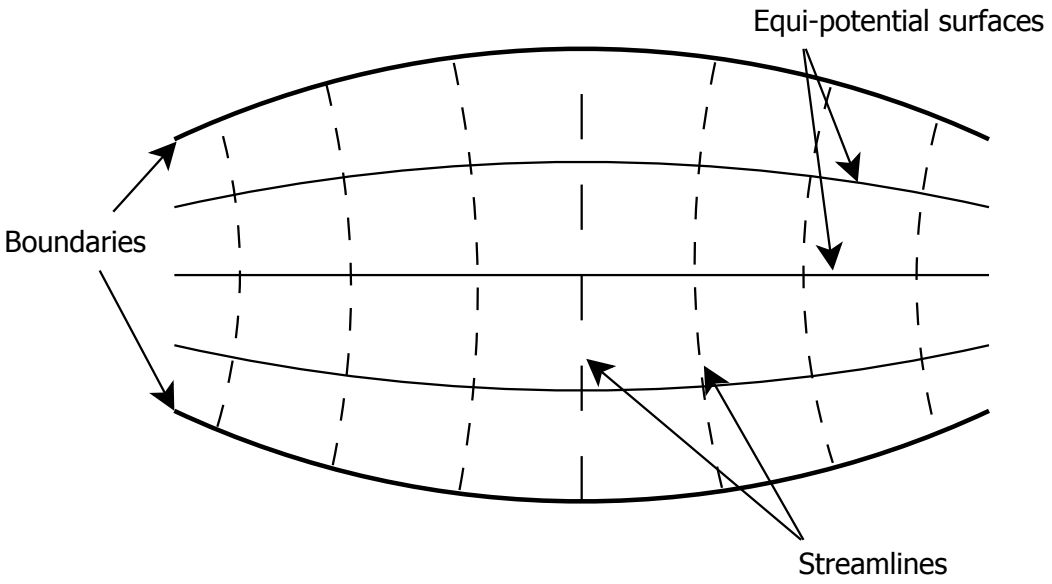


Figure 3.5: Solution of Laplace's equation: The top and bottom curves are fixed value boundary surfaces. Solid lines indicate equi-potential lines, and dotted lines show the streamlines.

Laplace's equation is an elliptic partial differential equation. A useful property of an elliptic PDE is that the solution is necessarily smooth – in the case of Laplace's equation, two orders smoother than the boundary conditions.

3.3.2 Solving Laplace's Equation

For simplicity, consider the solution of Laplace's equation in two dimensions. The numerical methods in two dimensions are the same as in three dimensions.

The 2D Laplace's equation, where ϕ is the temperature at a point in space, is given by

$$\Delta\phi = \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} = 0. \tag{3.10}$$

Its discretized form, where $\phi_{i,j}$ is the value of the temperature at grid location (i, j) and h is

the spacing of the grid, is given by

$$\frac{\phi_{i+1,j} + \phi_{i-1,j} + \phi_{i,j+1} + \phi_{i,j-1} - 4\phi_{i,j}}{h^2} = 0. \tag{3.11}$$

Such an equation is typically solved by a relaxation method such as Jacobi iterations:

$$\phi_{i,j}^{t+1} = \frac{\phi_{i+1,j}^{t} + \phi_{i-1,j}^{t} + \phi_{i,j+1}^{t} + \phi_{i,j-1}^{t}}{4}, \tag{3.12}$$

where $\phi_{i,j}^{t}$ is the value of the temperature at location (i, j) in iteration t.

Jacobi iterations can run in parallel but take a lot of time to converge. This is improved by two methods called red-black ordering and Successive Overrelaxation (SOR).

The method of red-black ordering exploits the fact that the value of a new iterate is only dependent on its four non-diagonal neighbors (in two dimensions). Imagine the entire location grid as a chessboard (with red and black colors); the value at the red squares is dependent only on its black neighbors and vice-versa. Thus, one could update all red positions in parallel and then all black positions in parallel.

$$\phi_{i,j}^{t+1} = \frac{\phi_{i+1,j}^{t} + \phi_{i-1,j}^{t} + \phi_{i,j+1}^{t} + \phi_{i,j-1}^{t}}{4} \quad \forall \text{ red points } (i, j), \tag{3.13}$$

$$\phi_{i,j}^{t+1} = \frac{\phi_{i+1,j}^{t+1} + \phi_{i-1,j}^{t+1} + \phi_{i,j+1}^{t+1} + \phi_{i,j-1}^{t+1}}{4} \quad \forall \text{ black points } (i, j). \tag{3.14}$$

By itself, red-black ordering increases the convergence speed only by a factor of two. The second improvement, which requires red-black ordering, is called *Successive Overrelaxation* (SOR). In Jacobi iterations, the value is relaxed by a factor of 1. In SOR, the iterates are 'over-relaxed' by a factor of ω, the Chebyshev acceleration. For an optimum value of ω, the iterations converge extremely fast. The new update equations are given by

$$\phi_{i,j}^{t+1} = \phi_{i,j}^{t} + \omega \times \left(\frac{\phi_{i+1,j}^{t} + \phi_{i-1,j}^{t} + \phi_{i,j+1}^{t} + \phi_{i,j-1}^{t} - 4\phi_{i,j}^{t}}{4} \right) \quad \forall \text{ red points } (i, j), \tag{3.15}$$

$$\phi_{i,j}^{t+1} = \phi_{i,j}^{t} + \omega \times \left(\frac{\phi_{i+1,j}^{t+1} + \phi_{i-1,j}^{t+1} + \phi_{i,j+1}^{t+1} + \phi_{i,j-1}^{t+1} - 4\phi_{i,j}^{t}}{4} \right)$$
$$\forall \text{ black points } (i,j). \tag{3.16}$$

Obtaining a fast solution is important for any application deployed in a clinical setting. Therefore it is important that my method be built on fast algorithms. The methods discussed above have been extensively used and developed in literature for solving Laplace's equation. They can be easily parallelized if speed were to become a critical issue. To increase the convergence rates further, I solve Laplace's equation using Jacobi iterations along with SOR and red-black ordering in a multi-scale fashion detailed in chapter 6.

Chapter 4

Quasi-tubular Medial Models

4.1 Introduction

In the human body, the blood vessels, the bronchi, and the colon are examples of nearly tubular objects. Segmenting these objects is an important task in medical imaging, and learning probability distributions on their populations is useful to segmentation algorithms [6]. Representing these objects with slab-shaped m-reps is unstable. However, most of them can be represented as a tube at the large scale with smaller scale changes understood as deviations from the tube. Some of these are shown in fig. 4.1.

A good statistical model for a population of objects consists of good geometric and image intensity representations. A good geometric model should stably and uniquely represent a population of objects and not just a single member, and it should provide correspondence between different members of the population. A good image intensity model should be able to represent image intensity distributions around the object for a good-quality segmentation. All these criteria are satisfied by the quasi-tubular medial models developed in this chapter along with the image intensity distribution model developed by Broadhurst [4] and Stough [32].

In this chapter, I first present a broad classification of tubular models and an overview of the image intensity model I use. This is followed by the several definitions associated with a nearly tubular objects from previous literature, and how they fail to satisfy all the conditions

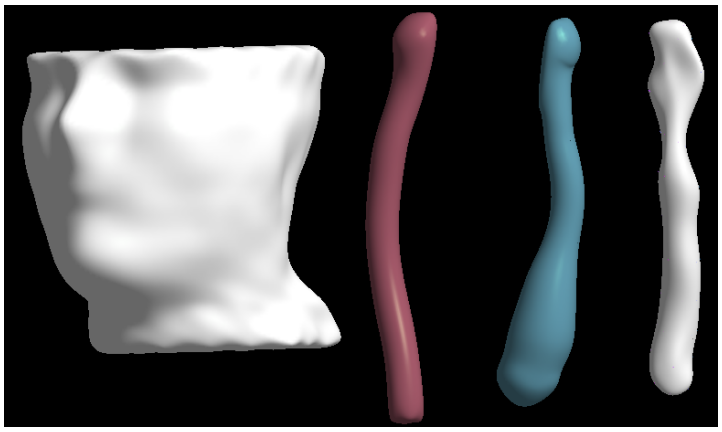


Figure 4.1: Renderings of quasi-tube models fitted to different objects. From the left to right the objects are sections of the head and neck's skin surface extracted from a 3D CT scan, the carotid artery, the internal jugular vein, and a section of the upper airway.

of a good statistical model. I then discuss my representation of a tube and representation of a population of tubular objects, and I see how these representations perform in practice.

The first part of a good statistical model for tubes is its geometric representation. There are several tubular models discussed in the literature. They fall into three broad categories – medial, skeletal, and boundary representations. The medial tubular model, defined by Koenderink [24], is the envelope of a set of spheres centered on a space curve. Skeletal tubular models include swept surfaces and generalized cylinders. Boundary representations will not be discussed here because they are not specific to tubular objects.

Generalized cylinders, a type of skeletal tubular models, incorporate a notion of a skeletal curve – a one-dimensional (1D) curve along which a cross-section is swept to generate a volume. The several definitions of a generalized cylinder differ in restrictions on the skeletal curve and on the cross-section function. A tubular generalized cylinder has a circular cross-section that may vary in size and have a bent skeletal curve. A point on a medial curve has an entire curve on the boundary equidistant from it, whereas a point on a skeletal curve can be much closer to one side than the other.

The second part of a good statistical model is the representation of image intensity patterns around objects of a population. 3D CT images of the human body and the head and neck in

particular often have poor contrast between soft tissue objects. This increases the signal-to-noise ratio and adversely affects the stability of a segmentation method that uses image intensities from a small region. For example, the physically motivated model by Terzopoulos et al. [33] uses image intensity forces that may not behave stably in the presence of poor contrast. A more stable approach would be to acquire image intensities over several large regions spread around the object of interest.

Segmentations can be stabilized in the presence of poor intensities by limiting the deformation to credible shapes. Doing so requires the knowledge of the different shapes the object can possibly take. Statistics is a suitable way to characterize these variations in shape and intensities. Our group has employed a statistical method of modeling intensities, developed by Broadhurst [4] and Stough [32], that leads to a more stable segmentation process in the presence of poor contrast and noise.

In this chapter, I discuss geometry and statistics for the medial definition of tubes, and I then extend it to support deviations from the tubular object. Section 4.1.1 details prior work done on modeling tubular objects, and section 4.1.2 presents the segmentation objective that serves as the driving problem.

4.1.1 Prior Work on Modeling Tubular Objects

There are several definitions associated with center-line-based tubular models. The definitions and the methods that implement them differ by whether the center-line and cross-sections are restricted, whether the center-line is unique, and whether the method is robust with respect to the number of training samples. Several of these definitions and their shortcomings are presented in the next few paragraphs.

Generalized cylinders, also known as generalized cones, were fist proposed by Binford [2] with special instances studied extensively in computer vision. A straight homogenous generalized cone [35] is the surface obtained by sweeping a fixed cross-section along a straight line while possibly scaling it, whereas a straight homogenous generalized cylinder may have

a cross-section that can change shape. Huang et al. [19, 18] discuss generalized tubes, which are constructed by sweeping a fixed cross-section along a curve with certain constraints.

Several center-line-based methods do not generate unique center-lines. To model populations, however, there must be a unique center-line found in a principled manner, so that variation introduced by the modeling process is not reflected in the statistics. O'Donnell et al. [26] discuss a novel method of generalized cylinders that works around part of this uniqueness problem by starting with a base cross-section that may be anisotropically scaled. Although they use only two scaling parameters, their method can be extended to produce arbitrary scaling. Further, they allow for local deformations of the cross-section by a spline function on the surface. Terzopoulos et al.'s [33] deformable model, which uses image-based and regularity forces to deform the model, can represent a wide range of tubular objects. However, none of them have discussed any means to compute statistics of their objects. It is also not trivial to extend any of them to include a statistical representation of the object they are modeling.

The subtle differences between a skeletal and a medial curve for a generalized cylinder are sometimes overlooked; the two terms are used interchangeably in some of the cited work. A skeletal curve is the locus of the centers of an arbitrary cross-section, whereas a medial curve is the locus of the centers of swept circles. (More generally, a medial curve is the locus of the centers of swept spheres.) Sweeping a non-circular cross-section typically results in a 2D medial surface. When the medial surface degenerates to a curve, we call it a *medial axis*. Having a true medial axis representation overcomes the issue of finding a unique center-line. The class of generalized cylinders whose medial axis is a curve is restricted to those with a circular cross-section.

A generalized cylinder and an object with a well-defined medial axis are closely related. When we sweep a constant circular cross-section along an appropriately smooth curve, the curve forms the medial axis of the generated surface. However, if we sweep a non-constant circular cross-section along a curve, the generated object may not have a curve as its medial locus though some such objects (generated by sweeping spheres of varying sizes) will have

a curve as the medial locus. Even when a generalized cylinder does not have a curve as its medial axis, it is useful to find an approximate medial axis for that object because the orientation of the medial surface across a population can be unstable.

Not all combinations of cross-sections and curves form a legal generalized cylinder. A sweep of a cross-section along a curve may result in two adjacent cross-sections crossing each other near a sharp bend on the curve. Such instances of the generalized cylinder are illegal, restricting the range of permissible cross-sections and medial curves. Damon [8] has described a method in the swept surface paradigm using a shape operator that can be used to detect these illegal generalized cylinders.

Finally, a quasi-tubular object can be thought of as an object that is modeled as deviations from a tubular object. In the general case, it is an object with a star-shaped[1] cross-section that may not be close to circular but varies slowly along the medial axis.

No specialized means of performing statistics has been developed for the generalized cylinders and swept surface models discussed above; they are best suited for modeling *individual* quasi-tubular objects and not *populations* of them. Because statistical descriptions can stabilize segmentations, there has been some work on statistically modeling tubes. One of the methods, called generalized stochastic tubes, developed by Huang et al. [20], aids in the segmentation of blood vessels but is specialized for this application. Another method, developed by de Bruijne et al. [10], retrofits Active Shape Models with center-lines but lacks a robust image appearance model.

With statistical shape models a special concern is their robustness against the number of training samples, since in medicine acquiring these training samples can be very expensive. As mentioned by Joshi et al. [23], the orientation of the narrow medial sheet of objects with a nearly circular cross-section is sensitive to small changes in the boundary and will result in a population with broad variation. I avoid this variability by approximating the narrow medial

[1] A star-shape is one where there exists a point in the interior of the shape such that every line connecting a point on the boundary with this point does not intersect the boundary at any other point.

sheet by a medial curve. Thus, the method described here uses statistical shape models for which the probability estimation is particularly robust against the number of training samples.

Several alternative methods exist that focus on the extraction of the center-line from image data. Examples include the cores methods (height ridges of medial strength) developed by Aylward et al. [1] and by Fridman et al. [13] to segment the vascular tree. Most existing work on segmenting the vascular tree uses a tube-following paradigm, where the tube is segmented a piece at a time. The two papers just cited use a medial model for a tube. However, their model is restricted to tubes with a cross-section that does not vary quickly as one moves along the axis. They use a model called 'the core', which is an m-rep extracted as a height ridge of a medial strength function on medial atoms. The medial atoms are characterized by a point and a set of equi-length spokes emanating from the point in the shape of a cone. These two papers compute a medialness measure to follow the medial axis of the tube after the core has been initialized. These methods, however, are tailored to suit the vascular and the bronchial trees, which have higher contrast and are narrower when compared with rectums, upper airways, and the other objects that I am interested in here. Also, the latter set of objects do not branch and can have cross-sections that can be far from circular. Rectums in particular can vary greatly in shape due to the presence of gas and other non-tissue material. Another difference between the objects is related to the surface curvatures: The curvature of the surface in blood vessels changes slowly along the circumferential direction and stays almost the same as that of the medial curve in the axial direction. In rectums and other objects, the surface curvature is not related closely to the medial curvature.

4.1.2 The Driving Problem: Segmentation of Quasi-tubes

The method of segmentation via posterior optimization of m-reps, developed by Pizer et al. [28], has been successful in dealing with slab-shaped objects with a lot of variability and poor contrast. We develop a new method that draws on the strengths and ideas from these methods but represents a tube-like object with a discretely sampled medial space curve and

then models quasi-tubes as deviations from these tubes.

The segmentation method can be divided into two parts – training and the actual segmentation itself. During training, a rough m-rep model of the object is allowed to vary inside an optimizer that favors smooth models with a regularly spaced discrete medial mesh and that match well with the image data. The resulting models are known as training or fitted models.

Next, these training models are statistically analyzed. The variation in the shape space of the models is studied using Principal Geodesic Analysis (PGA), developed by Fletcher et al. [12], which is a form of Principal Component Analysis (PCA) suited for non-linear spaces. The result is a mean shape $\overline{\mathbf{m}}$ and a prior $p(\mathbf{m})$ for segmentation. At the same time, the region around the object is divided into small parts, and the distribution of intensities in each region is studied with the help of local region intensity quantile functions, developed by Broadhurst [4] and Stough [32]. We then apply PCA on these quantile functions to produce a likelihood function $p(\mathbf{I}|\mathbf{m})$ for segmentation.

When an image is to be segmented, the mean model is placed close to the real organ with the help of landmarks or manually. The model is allowed to deform along its principal modes of variation in an optimizer that favors likely shapes and intensity distributions around and within the object. The objective function maximized is the weighted sum of $\log p(\mathbf{m})$ and $\log p(\mathbf{I}|\mathbf{m})$ with the weights chosen to make the two terms have equal variance. This is a variant of the method of posterior optimization.

Similarly, for quasi-tubes we start with a tubular medial model that during training is deformed to a quasi-tubular model. Shape and intensity statistics are obtained, and then these probability distributions are used to drive the segmentation. As with slabs, the objective function is $w_1 \log p(\mathbf{m}) + w_2 \log p(\mathbf{I}|\mathbf{m})$.

The remainder of the chapter is organized as follows: Section 4.2 describes the representation and geometry of tubular medial models. Section 4.3 describes the way in which we estimate probability distributions on these models. Section 4.4 presents a study on the performance of the tube model in simulated rectums. Section 4.5 details the modeling of the

deviations from a tubular to a quasi-tubular model. Section 4.6 gives more details of the training and segmentation approaches. Finally, section 4.7 presents quantitative and qualitative results obtained by applying my method on real data obtained from CTs of rectums.

4.2 Medial Models for Tubes

This section describes the representation and geometry of tubular medial models. The structure of a tubular m-rep is described in the following manner: A first order tube m-rep is a continuous space curve with a cone placed at every point along the curve. The axis of the cone is tangential to the space curve at the tip of the cone. Sweeping the edges of the cone bases gives the boundary of the modeled object, which is orthogonal to the rays from the cone tip to the cone base. The cones may have a half cone angle greater than $\pi/2$ but less than π. They are not allowed to intersect each other. Damon [7] has provided us with tools that can be used to measure local self-intersection (folding) of the object implied by the medial surface of a slab m-rep. In section 4.2.1 we adapt these tools to do the same for tubular m-reps.

In practice, we represent the medial model of a tube by discretely sampling the space curve of cone tips. Each sample, shown in fig. 4.2, is called an atom. Associated with each sample is its position in space, $\mathbf{p} = (x, y, z)$, and a cone with its tip positioned on the sample. The cone in turn is represented by its bisector, $\widehat{\mathbf{U}}_0$ given by two angles $(U_{0,1}, U_{0,2})$, the half cone angle, θ, and the length of its inclined surface rays, r. The bisector of the cone always points along increasing arc length; thus, the bisector points in the same direction when θ changes across $\pi/2$. To keep the discrete samples regularly spaced, while developing the models, we impose a penalty, called *irregularity penalty*, on the model that penalizes atoms moving away from the average of its neighbors.

The atoms at the two ends of the chain have an additional parameter describing the curvature of the cap at that end. However, when we are modeling open tubes, the end-atoms don't have any special properties and behave just like any other atom.

A continuous medial curve $\gamma(u)$ is interpolated from these atom positions and cone bisector vectors using piecewise cubic Hermite splines. The cone bisector vectors, scaled by the mean of the distance between the position of the atom and its two neighbors, serve as the tangents in the Hermite interpolation.

Cones are rotationally symmetric, which prevents uniquely identifying points on them. This symmetry is broken by a parameter ϕ that rotationally orients the entire tube along its length. One of the atoms in the tube is designated as a base atom. Usually this atom is close to a feature in the object that can help fix the rotational orientation. This feature can be an anatomic entity such as a part of a bone, a certain neighboring organ, or a tissue that can be easily identified in the entire population. Whenever the cone for this atom is rotated around its bisector, all the other cones are sympathetically rotated (see section 4.2.3). This is needed for correspondences that depend upon the position along the circumference of the tube such as those required for quasi-tube statistics.

The surface implied by the tubular medial model is called the *medially implied surface* and generated as follows: The circles at the base of the cones are discretely sampled. These surface points are interpolated to generate the fine and smooth mesh called the medially implied surface. The medially implied surface for a tube model of a rectum is shown in fig. 4.3.

Each position on the surface of the tube and its interior is uniquely represented by a set of coordinates, and this coordinate system provides correspondence between deformed versions of the tube. There is a parameter u that ranges from 0 to one less than the number of primitives and increases along the length of the tube. The second parameter ϕ ranges from 0 to 2π along the circumference of the tube. At the end-points, ϕ could take any value; for ease of distinction, it is set to -1 (a value that is not between 0 and 2π). A third parameter t that is 1 throughout the length of the tube changes gradually to 0 at the start of the end caps to the tips.

Section 4.2.1 highlights a geometric operator that is useful for detecting models with self-intersecting surfaces and for interpolating between atoms in the medial axis. Section 4.2.2 describes a geometric measure, which I call curviness, that is minimized to avoid wavy medial

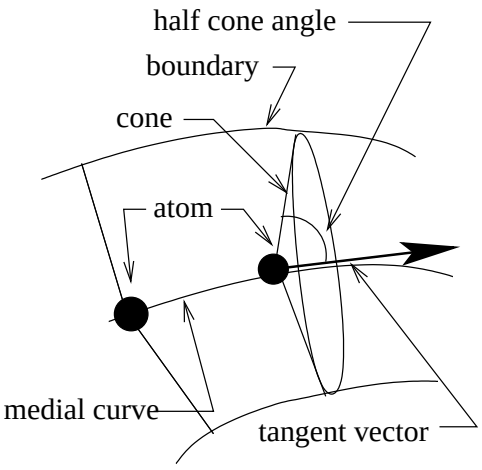


Figure 4.2: Representation of a tube atom

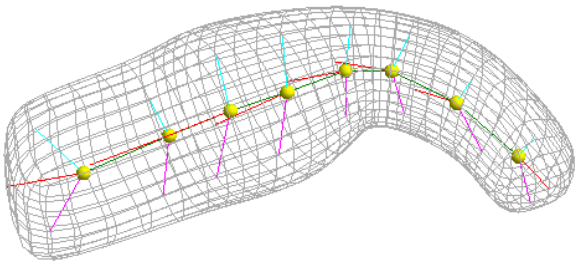


Figure 4.3: A mean model of a rectum from one of our studies showing the medially implied surface as a wireframe.

axes. Finally, section 4.2.3 describes the process of untwisting the tube model.

4.2.1 Geometry on tubular models

Damon defined a shape operator for swept surfaces [7] that is useful to ensure that the swept surface model does not self-intersect. In this section, I present a shape operator for the tubular medial model.

Consider the medial atom cone to be composed of a set of spokes $\mathbf{S}$ extending from the tip to the base. These spokes are parameterized by the arc angle $\phi \in [0, 2\pi)$ on the circumference of the base made with respect to a zero reference point on the circumference. Define the 1D radial shape operator for a tube as

$$S_{rad}(\phi) \quad = \quad -proj_{\mathbf{S}} \left(\frac{\partial \mathbf{U}(\phi)}{\partial s} \right), \tag{4.1}$$

where $proj_{\mathbf{S}}$ is projection along the spoke vector $\mathbf{S}$, $\mathbf{U}$ is the corresponding unit spoke vector, and s is an arc-length parametrization of the medial curve. S_{rad} and r are also functions of u.

The derivative may be written in the form

$$\frac{\partial \mathbf{U}(\phi)}{\partial s} \quad = \quad a \cdot \mathbf{U} - \kappa_{rad} \cdot \gamma'(s), \tag{4.2}$$

where κ_{rad} is the principal radial curvature. Equations 4.1 and 4.2 imply $S_{rad}(\phi) = \kappa_{rad}(\phi)$.

An important use of this shape operator is in detecting illegal models – an illegal model is one where some of the spokes cross each other and the surface has folded onto itself. Damon [7] shows us that the spokes will cross each other if and only if $r \times \kappa_{rad} > 1$. Unlike Damon's shape operator, this shape operator is a function of the angle ϕ. However, it suffices to evaluate this for the angle corresponding to the direction of the curve normal, because if the spokes do cross each other they will definitely do so in the direction of maximum bending.

The model is illegal if $\exists\, u$ s.t. $r \times S_{rad} \geq 1$. This condition can be relaxed as a penalty to make it suitable for an optimizer that expects a continuous objective function. The penalty is the p-norm of the individual measurements $\max(0, r \times S_{rad} - \beta)$ along the tube for a certain threshold $\beta \in [0, 1)$. Experimentally, $p = 6 - 8$ and $\beta = 0.8 - 0.9$ have been observed to produce good results. Larger values of p are sometimes useful when we wish to make the aggregate measure more sensitive to local problems.

The S_{rad} operator can also be used to interpolate between two consecutive atoms on the same medial manifold. Han et al. [17] have used this operator to interpolate between atoms on a 2D medial manifold. We have adapted the interpolation method to generate interpolated atoms for tubes. An important use is in interpolating atoms to improve correspondence between models.

4.2.2 Geometric Penalty - Curviness

A wavy medial curve results in crooked-looking models. It also necessitates the use of significantly more samples in the computation of the illegality penalty. Crooked-looking models are typically not anatomically probable. In this section, I develop the curviness measure that

alleviates this problem.

This penalty, *curviness*, C, is defined by the p-norm of the total curvature over the entire length of the medial curve:

$$T = \sqrt{\kappa^2 + \tau^2}, \tag{4.3}$$

$$\text{where } \kappa = \frac{|\gamma'(u) \times \gamma''(u)|}{|\gamma'(u)|^3} \tag{4.4}$$

$$\text{and } \tau = \frac{[\gamma'(u)\ \ \gamma''(u)\ \ \gamma'''(u)]}{|\gamma'(u) \times \gamma''(u)|^2}. \tag{4.5}$$

$$C = \left(\frac{1}{n-1} \int_0^{n-1} T^p \, \mathrm{d}u \right)^{1/p}, \tag{4.6}$$

where T is the total curvature, κ is the curvature, and τ is the torsion of the interpolated medial axis $\gamma(u)$. Different values of p between 2 and 10 are appropriate depending upon how much we need to penalize individual sharp bends. However, larger values of p create steeper objective function gradients, which may not be good in an optimizer.

The curviness measure is rotation- and translation-invariant. Dividing by the arc length or the average inter-atom distance makes it scale-invariant too.

4.2.3 Untwisting the Tube Model

The representation does not prevent a tube model from twisting around its axis. This freedom is good if the objects twist across the population. However, in quasi-tube statistical analysis, for objects that do not twist this is an unwanted source of variation. Fig. 4.4 shows the surface of a twisted bowling pin model, and fig. 4.5 shows the surface for an untwisted version of the same model. Bowling pins ordinarily do not have twisted surfaces. Therefore, this variation is unwanted and should be removed.

To remove this source of variation, one of the atoms in the tube model is chosen as the base atom. The rotational orientation of the cone around the axis is fixed with the help of a landmark – a consistent anatomical or geometric feature across the population. The cone

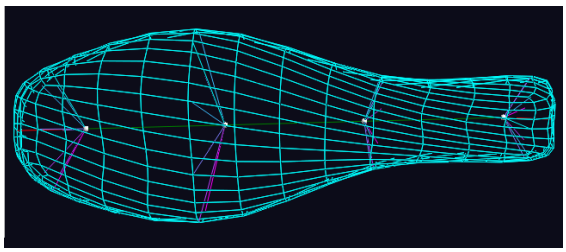


Figure 4.4: Surface of a twisted bowling pin

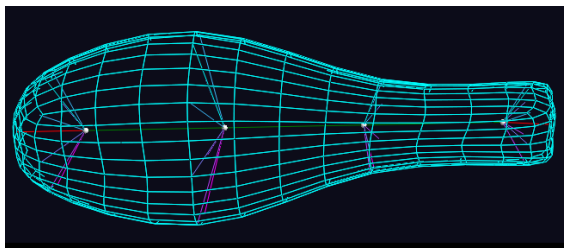


Figure 4.5: Surface of the same bowling pin, now untwisted

of an adjacent atom is rotated so that the distance between the corresponding ϕ positions on the edge of the base of the cones is minimized. This process is continued till the change is propagated to all atoms.

4.3 Shape Representation and Statistics of Tubes

In order to do statistics on tubular models, we need to understand spaces containing credible tubes and probability distributions on these spaces. These matters for *quasi*-tubular spaces are covered in section 4.5.

A tube atom $\mathbf{m}$, shown in fig. 4.6, is represented by the tuple $\mathbf{M} = \langle \mathbf{P}, \hat{\mathbf{U}}_0, \theta, r \rangle$. A tube consisting of n atoms is represented by n such tuples concatenated together. Here the cone vertex $\mathbf{P}$ belongs to the group $\mathbb{R}^3$, the cone axis $\hat{\mathbf{U}}_0$ is a point on the two-dimensional (2D) sphere $\mathbb{S}^2$, the half cone angle θ ranges from 0 to π and is related to the group $\mathbb{RP}^1$, and the cone length r belongs to $\mathbb{R}^+$. All of the groups are Lie groups, and except for $\mathbb{R}^3$ they are not Euclidean manifolds. A Lie group has a differentiable group operator, an inverse element, and the identity element. Using the group operators and by the action of the inverse on an element of the group on the element itself, we obtain the identity element. These properties assist in developing a distance metric in this space and the means of projecting back and forth between this shape space and a tangent plane. Several of the results presented in this section are similar to those worked out by Fletcher et al. [12] for slabular m-reps.

The path with the shortest distance between two points in a manifold is known as the

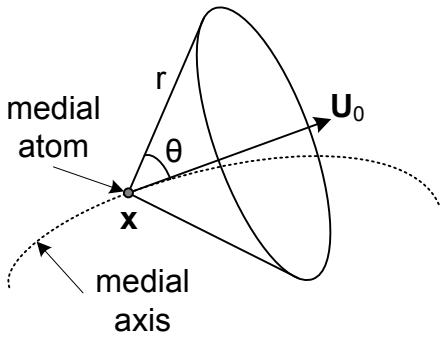


Figure 4.6: The structure of a medial atom.

geodesic between them. The length of this path is called the *geodesic distance*. If we have a suitable mapping between the manifold and a tangent space, this mapping can be used to measure the geodesic distance on the tangent plane. The map that takes us from the manifold to the tangent space is called the Logarithmic chart, and the reverse map is known as the Exponential chart. The maps of the individual components are given by the equations presented below. The map for the atom is simply the direct product of these maps. The maps for the position are identity functions and that for the radius are the ordinary logarithm and exponential functions. In the following equations, $\theta \in (0, \pi)$, $\mathbf{U}_0 = \langle u_1, u_2, u_3 \rangle \in \mathbb{S}^2$, and $\mathbf{U'}_0 = \langle u'_1, u'_2 \rangle \in \mathbf{T}_{(0,0,1)}\mathbb{S}^2$.

$$\mathrm{Log}(\mathbf{x}) = \mathbf{x}, \tag{4.7}$$

$$\mathrm{Exp}(\mathbf{x}') = \mathbf{x}, \tag{4.8}$$

$$\mathrm{Log}(r) = \log(r), \tag{4.9}$$

$$\mathrm{Exp}(r') = \exp(r), \tag{4.10}$$

$$\mathrm{Log}(\theta) = \tan\left(\theta - \frac{\pi}{2}\right), \tag{4.11}$$

$$\mathrm{Exp}(\theta') = \tan^{-1}(\theta) + \frac{\pi}{2}, \tag{4.12}$$

$$\mathrm{Log}(\mathbf{U}_0) = \left(u_1 \frac{\alpha}{\sin\alpha}, u_2 \frac{\alpha}{\sin\alpha}\right), \text{and} \tag{4.13}$$

$$\mathrm{Exp}(\mathbf{U}'_0) = \left(u'_1 \frac{\sin|u'|}{|u'|}, u'_2 \frac{\sin|u'|}{|u'|}, \cos|u'|\right), \tag{4.14}$$

where $\alpha = \cos^{-1}(u_3)$, and $|u'| = \sqrt{u_1'^2 + u_2'^2}$. All the above maps are taken centered at the identity element. The identity element for the group $\mathbb{R}^3$ is 0, for $\mathbb{R}^+$ it is 1, and for $\mathbb{S}^2$ it is the point $(0, 0, 1)$. For the group of θ, the identity element is $\pi/2$. To obtain the chart for a tangent plane centered at a point $\mathbf{m}$ different from the identity element, we need to apply the inverse of that point to the element in order to move the tangent plane to the identity element. We use $\mathrm{Log}_{\mathbf{m}}$ and $\mathrm{Exp}_{\mathbf{m}}$ as the notation in this case. To make the units of all the components commensurate in the Log map, we multiply the unitless quantities with the mean radius taken over all the corresponding atoms in the population. The geodesic distance is then defined as the L2-norm of the difference of these normalized atoms projected into the tangent space.

The Fréchet mean μ of a set of atoms is defined as the atom that minimizes the sum of squared geodesic distances from all the corresponding atoms from the population:

$$\mu = \underset{\mathbf{m}\in\mathbf{M}}{\operatorname{argmin}} \sum_i |\mathrm{Log}_{\mathbf{m}}(\mathbf{m}_i)|^2. \tag{4.15}$$

To compute the modes of the variation of the model, the Log map of all these atoms is computed, and they are projected on the tangent plane centered on the mean. PCA is performed on these projected atoms, and the first few modes that represent more than 90% of the total variation are kept. Fig. 4.7 shows the shape variation along the first mode of variation of the rectum from our study.

In a tube model, an atom can move around the medial curve with only a little change in the surface. This sliding of the atoms around the medial curve is somewhat arbitrary and can lead to unwanted noise and modes in the statistics. Before computing statistics, correspondence between the several tube models in the population is improved by sliding the atoms around the medial curve. There are several schemes to choose from: One scheme aims to minimize the variability of the distribution of the atoms; another scheme re-arranges the atoms so that they are all equally spaced. Correspondence may also be imposed by user-specified landmarks. Further, a hybrid of these strategies may be employed. The choice of a scheme is dependent

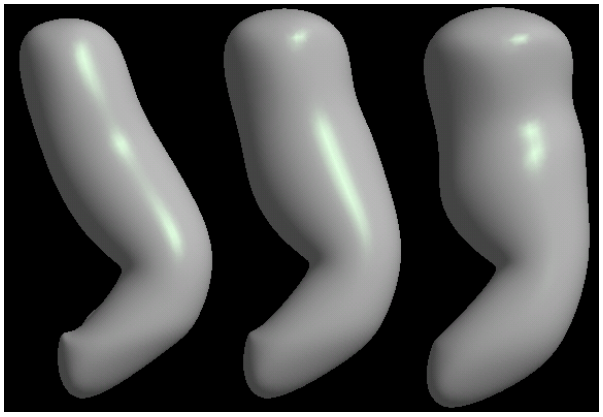


Figure 4.7: Mean model of a rectum (center) deformed by ± 1.5 standard deviations along the first mode of variation (left and right), which resembles the anatomical shape change due to bloating by gas.

on the application. In segmenting rectums, the scheme of interpolating atoms so that they are equally spaced was used. I have not evaluated the impact of the different correspondence schemes on the quality of segmentations in this work.

In several tubular objects, the end may be arbitrarily cutoff. This can lead to undesirable variation in statistics that is related to the position of the ends. Removing this unwanted variation is discussed in the following section.

4.3.1 End Atom Normalization

Organs in the shape of tubes in the human body are usually quite long. However, in applications such as radiotherapy, surgery, and treatment planning, one is interested in a small section. During the step of manual segmentation for generating training images for the model, the cutoff may be somewhat arbitrary. If we compute shape statistics of models trained on these images, the variable position of the ends may become a dominant and misleading mode of variation after PGA.

To remove this variation, the mean of the length of the medial curve between the first and the second atoms across all the training models is computed. Next, the first atom of each model is interpolated or extrapolated so that it is at the mean distance from the second atom

along the medial curve. This is repeated for the other end. If $\mathbf{m}_{0,i}$ and $\mathbf{m}_{1,i}$ are the first and second atoms from the i^{th} model, then the new first atom, $\mathbf{m'}_{0,i}$ is given by the following relation:

$$\begin{aligned} l_i &= \int_0^1 \gamma_i(u)\,\mathrm{d}t, \\ \mathbf{m'}_{0,i} &= \operatorname*{Exp}_{\mathbf{m}_{1,i}}\left(\frac{\bar{l}}{l_i}\operatorname*{Log}_{\mathbf{m}_{1,i}}(\mathbf{m}_{0,i})\right), \end{aligned} \tag{4.16}$$

where $\bar{l}$ is the mean of l_i over all the training samples. The same equation can be used for both interpolation and extrapolation.

4.4 Synthetic Rectums Study

In order to test the method of tubular medial models, Joshua Levy trained some models on real rectum data and obtained a mean model with shape statistics from these trained models. These statistics were then sampled to produce 99 synthetic rectum models that were truly tubular. Grayscale images for these models with additive Gaussian noise were generated to train the intensity statistics on. We first trained tubular medial models on each of the 99 cases and computed shape statistics on these in a leave-one-out manner. Corresponding image statistics were also computed in a leave-one-out fashion. Landmarks corresponding to the two ends of the rectum and the point closest to the prostate were also automatically generated and used for the initialization during the segmentation.

Some of these results are shown in fig. 4.8. In fig. 4.9, the performance of this method over all the samples is graphed. The green line shows how far away the trained models are from the ground truth using the average distance measure. The red line shows the result of the segmentation. The length of a side of a voxel in the simulated data is 2 mm. Thus on average, all the results are less than half a voxel away from the ground truth.

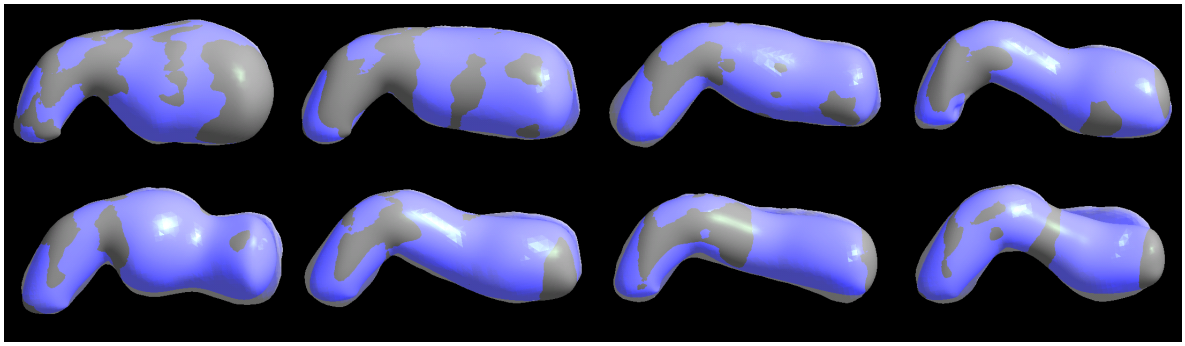


Figure 4.8: Set of segmentation results in decreasing order of performance on synthetic rectum images: The segmented rectum is shown in a gray color, and the ground truth is shown by a transluscent blue color. The average distances between the two from left to right in units of mm are: 0.9, 1.2, 1.4, and 4.0.

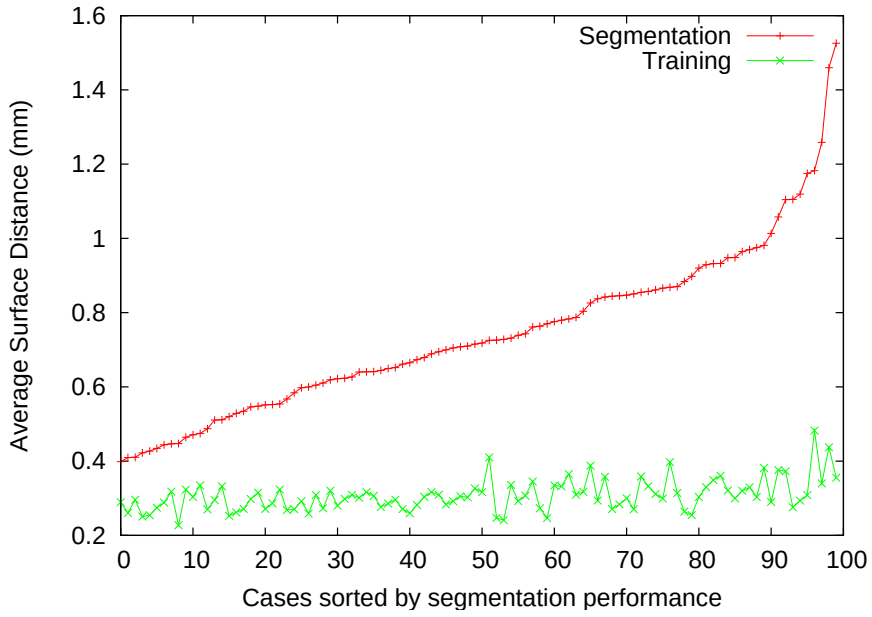


Figure 4.9: The red line shows the average distance our segmentation results were from the ground truth on synthetic rectum data. The green line shows the same for the trained models. The length of a side of a voxel is 2 mm.

4.5 Shape Representation and Statistics of Quasi-tubes

In this section, the deviations from a tube to a quasi-tube are modeled.

Several objects in the real world can be modeled as deviations from a tubular object. Take the head and neck for example: Start with a cylinder, make the cylinder bulge out in certain

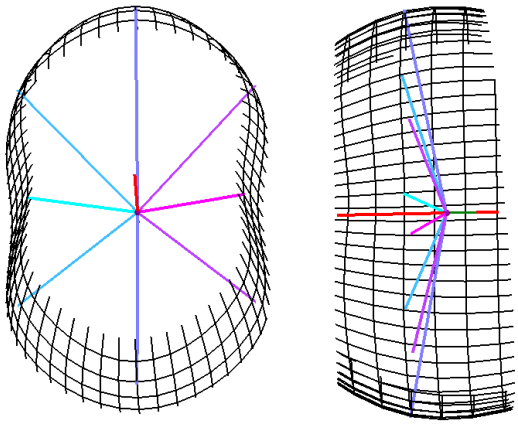


Figure 4.10: A quasi-tube atom with spokes of varying length and a cut-away section of the medially implied surface shown in two different orientations.

regions to produce features such as the nose and the lips, and make it cave in to produce the eye sockets and other cavities. Sections of the colon, blood vessels, the bronchial tree, and many other organs in the body can be thought of as quasi-tubes.

The deviation from a tube to a quasi-tube is accomplished by altering the cross-section of the tube. A tube atom is represented by a cone of spokes. All of the spoke ends lie in a single plane, α, forming a cross-section of the tube. The cross-section can be altered by changing the length of each spoke and inclining it in the plane formed by itself and the axis in such a way that the tips of all the spokes continue to lie in the plane α. Fig. 4.10 shows a quasi-tube atom.

Changing the cross-sectional shape in this way makes the computation of the shape operator in the circumferential direction straightforward, and the component along the axis is still given by the 1D shape operator defined by equation 4.1 as $\mathbf{U}$ is a function of the spoke angle ϕ. Further, because the spokes vary in length, the object is no longer strictly Blum medial but skeletal. However, Damon's proofs concerning the use of the shape operator S_{rad} are valid for skeletal objects, too.

To compute the statistical variation in the spokes across the population, the change in the length of the spokes is represented as a multiplicative parameter that belongs to the group $\mathbb{R}^+$.

The Log and Exp maps for this parameter have the same form as that for the radius r.

The segmentation step is divided into two scales: At the large scale, a mean of each quasi-tube atom across all population samples is used. This gives a large scale model with each cross-section having a different shape. At the small scale, the individual quasi-tube atoms are allowed to vary to allow for small changes in each cross-sectional shape.

4.6 Training and Segmentation

I have discussed how we geometrically and statistically model tubes and quasi-tubes. I now discuss the training of these statistical geometric models and the appearance model, so that we may apply the techniques to learn these models and eventually segment real images.

To fit quasi-tube models to training cases and derive a probability distribution on that object, we followed the general approach described in Merck et al. [25]. Expert outlines designating the target object in several 3D images were converted into binary images. An initial model that somewhat resembled the object was created. Sometimes landmarks were added to initialize our models. The parameters for these models were then varied inside an optimizer. The objective function was set to a sum of image match and geometric penalty terms. The image match used was the average of the sum of squared distances between a point on the object's surface and the closest point on the expertly outlined object. In certain places [2] such as the crest, the reverse distance was used, as the original distance is artificially low. When landmarks were used, the distances between them and the corresponding positions on the model were taken into account. The geometric penalty terms consisted of a combination of the irregularity, S_{rad} and the curviness penalties mentioned in sections 4.2, 4.2.1, and 4.2.2 respectively. The weights for the geometric penalties were relaxed proportionally to the quality of the fit of the model.

Next, the statistical shape model from all the models in the population was computed. To

[2] See Merck et al. [25] for a full discussion of the calculation of this distance function.

achieve this, first, the Log map of the atoms is computed. The spoke deviations are represented as a multiplicative parameter that belongs to the group $\mathbb{R}^+$. The Log map of the deviations is the Log map for the group $\mathbb{R}^+$ as described in section 4.3. Next, the mean across the entire population is computed. The atoms are projected at the tangent plane centered on the mean. The statistical shape model is computed as described in section 4.3 except that the spoke deviations are only used in the computation of the mean and not in the computation of the modes of variation of the shape. These comprise the large scale shape variation. The spokes are then allowed to vary at the smaller scale. This separation of variation at two different scales yields more stable statistics.

The appearance model used for segmentation had been developed by Broadhurst [4] and Stough [32]. In that model, the object's surface is divided into several regions. An intensity histogram for each region is computed and converted into quantile distributions. The distribution of these quantile functions is then analyzed with the help of Principal Component Analysis (PCA). As tubes can be arbitrarily cut off, intensities at the ends of tubes could be optionally ignored.

The segmentation used the method of posterior optimization where the statistical shape model gave the prior and the appearance model gave the likelihood.

4.7 Application and Results

Segmenting rectums from 3D CT scans is important for adaptive radiotherapy treatment for prostate cancer. It is important for the patient's health and quality of life that the rectum does not receive too much radiation. The rectum changes shape significantly from day to day due to the presence of gas and faeces pushing around the abutting prostate. The presence of gas also creates a large variation in the intensity distributions and necessitates special handling of the interior intensity distributions in the rectum.

In the experiment conducted, the data came from different days of several patients. For

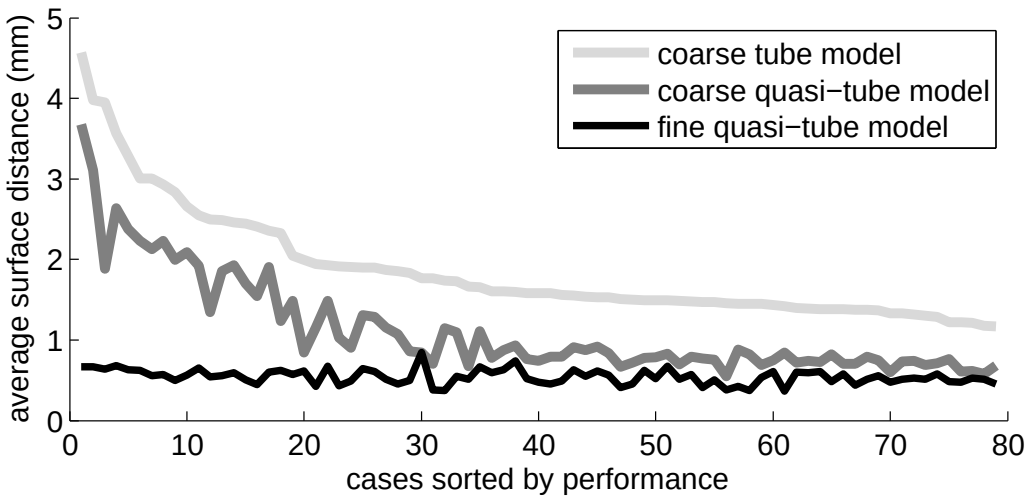


Figure 4.11: Average surface distance between the training models and the expert segmentation for the several cases (sorted by increasing performance of the coarse tube model) at different stages of fitting – coarse tube fitting, coarse quasi-tube fitting, and finer scale quasi-tube fitting.

each patient, the statistical shape model was built with the help of data from just that patient. The study was done in a leave-one-out manner, i.e., the statistical shape model for each day of a patient was built by using data from the other days. Even though it is clinically not possible to train using data coming from future day scans, it suffices for the purposes of this study. If more training data were available, cross-patient statistics could have been incorporated to get over this limitation.

3D CT images with a resolution of $0.98 \times 0.98 \times 3\ \mathrm{mm}^3$ of 5 patients with 13–18 images for each patient for a total of 79 images were obtained. Next, I trained quasi-tubular medial models on manual segmentations of rectums from this data. For the training, I used a hierarchical training process where first, a coarsely sampled tubular medial model was fitted to the data. Next, the individual spokes were allowed to vary, generating quasi-tubular models. After this the model was subdivided by generating interpolated atoms halfway between previously existing atoms. This subdivided model was refined to better fit the binary image. Further subdivision yielded only marginal improvement and was therefore not considered.

The quantitative results of the training process are shown in fig. 4.11. The median average

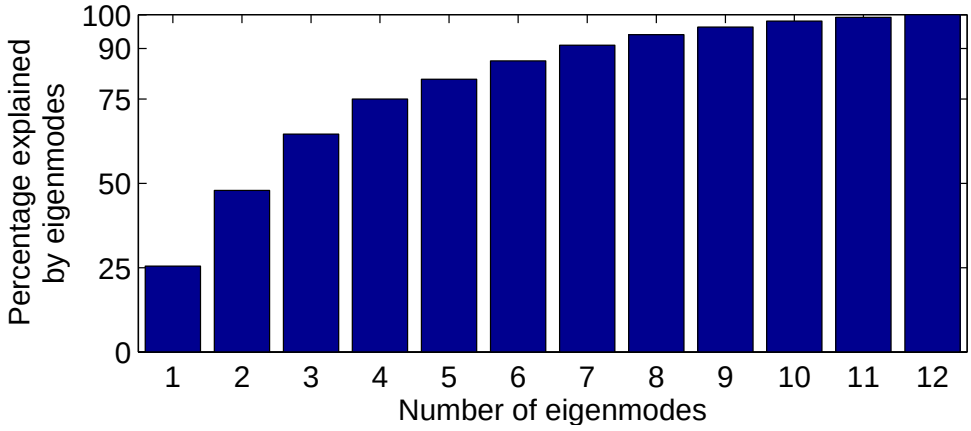


Figure 4.12: Graph showing cumulative shape variation (as defined by the eigenvalues) captured by the first few eigenmodes: the first seven modes capture about 90% of the shape variation.

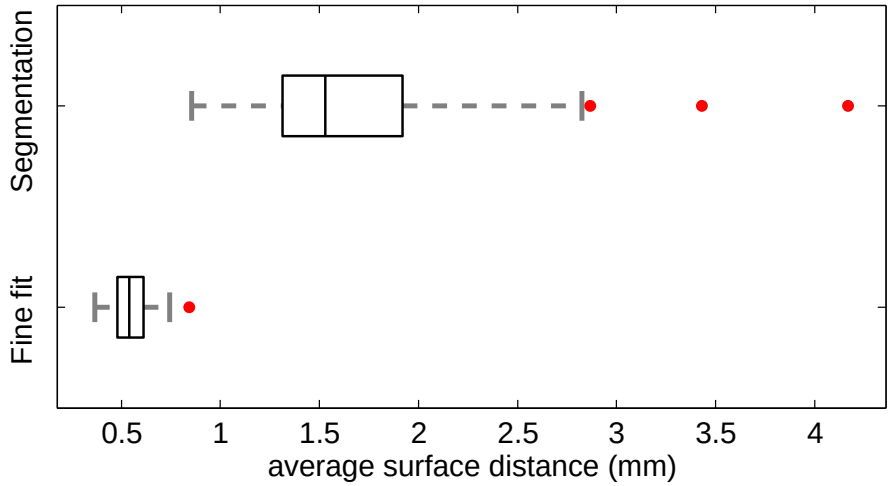


Figure 4.13: Boxplot of distribution of segmentation results versus training results. The box goes from 25% to 75% quartile with a line for the median. The two lines at the end go 1.5 times the intra-quartile range (except when there is no more data). The red dots are outliers.

surface distances for the coarse tube, coarse quasi-tube, and the fine quasi-tube models were 1.58, 0.84, and 0.54 mm with standard deviations of 0.70, 0.62 and 0.09 mm. This shows that rectums can be modeled fairly well and in cases where the tube model is lacking, the quasi-tubular model does a much better job.

In fig. 4.12, I show the cumulative shape variation captured by the first few eigenmodes

Figure 4.14: Each row shows the outline (white or black) of our segmentation on two different axial slices of the same image. Note the poor contrast in the slices in the right column. The slices on the right are inferior with respect to those on the left. The first three rows are typical results, and the last row is one of the better segmentations.

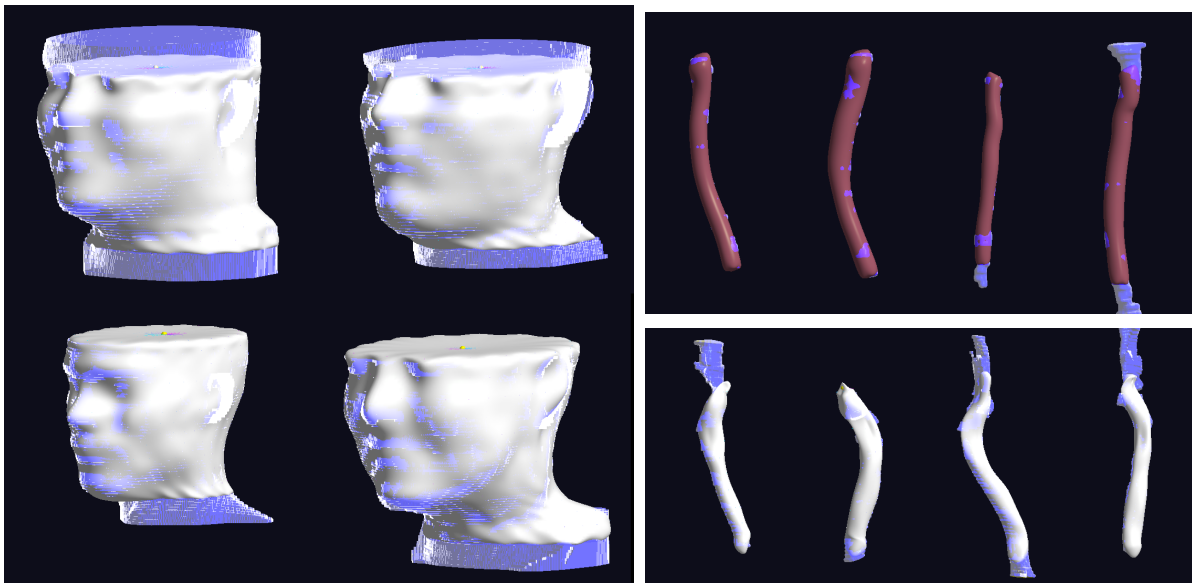


Figure 4.15: Quasi-tubular medial models (in white) fit to sections of the head and neck skin surface (left), common carotid arteries (top right), and upper airway sections (bottom right) vs. the manually segmented objects (in translucent blue)

for a quasi-tubular model trained over the data of one of the patients. In about seven modes, most of the shape variation is effectively captured. In fig. 4.7, I showed the variation of the rectum along the first principal direction, which resembles anatomical shape changes due to filling of gas.

During segmentation, I used a similar hierarchical approach. The model was first initialized semi-automatically. Next, the model was segmented in the shape space of the coarse models. The result was then used to drive the segmentation in the shape space of the finely sampled trained medial models. In fig. 4.13, I show the segmentation results versus the quality of fine training fits for all the patients. The median of the segmentation results was 1.53 mm. The segmentation results for a few of these is shown in fig. 4.14. Notice the complete lack of contrast in the inferior slices shown.

A method based on these ideas along with certain proprietary changes was developed at Morphormics[3] with significant improvements in the quality of rectum segmentations.

I have also fit quasi-tubular models to various objects of the head and neck. In fig. 4.15, I

[3]See the company web-site http://www.morphormics.com/ for more details.

show quasi-tubular models fit to sections of the skin surface and upper airway. The expertly contoured outline is shown in translucent blue. The average surface distance for the fitted models of the skin surface, common carotid artery, internal jugular vein, and the upper airway section were 1.22, 0.8, 1.3, and 1.13 voxels.

4.8 Discussion

I have talked about how ordinary medial models designed for slab-shaped objects can be unstable for nearly tubular objects and how quasi-tubular medial models are more stable for such objects. However, for what organs should we use quasi-tubular medial models? The instability increases as the medial sheet narrows. Therefore, when the ratio of breadth to length of the medial sheet falls below a certain threshold, we should switch from using an ordinary medial model to a quasi-tubular medial model.

I have developed a new method for modeling populations of nearly tubular objects as a tubular medial model with deviations from perfect tubularity described by local changes. Application to real world rectum data demonstrated the effectiveness of this method. We have seen that this representation is also effective for elongated objects with distinctly non-circular but slowly varying cross-sections.

Further, I showed how we can study population variations by doing statistics in the non-linear space in which these quasi-tubular medial models lie. I also discussed the radial shape operator that is needed for studying skeletal geometry on these models.

The rectum is a challenging organ to segment due to immensely varying shape and poor image contrast. Several objects in the head and neck – skin surface, upper airway, jugular vein, and carotid artery – provide difficult modeling challenges. I have shown that these quasi-tubular models can be trained to within sub-voxel accuracy and give reasonable segmentations.

Chapter 5

Correspondence

As stated in chapters 1 and 3, the correspondence between the points on the surfaces implied by m-rep objects is neither precise nor accurate. Precision is related to the level of detail in the modeled objects, and accuracy is related to correctness in the reproduction of the modeled object. The limitation in precision is due to m-reps' modeling anatomical structures at a large scale – by design they do not precisely follow the expertly outlined contours. The inaccuracy is due to the nature of the fitting process – a trade-off between geometric typicality and data match, which is itself a compromise between local data match at several locations. As the warping method detailed in chapter 6 preserves the given correspondence, its result can be only as good as the given correspondence. A method to increase the precision is detailed in section 5.1, and a method to increase the accuracy is detailed in section 5.2.

5.1 Enhancing the M-rep Fit

The m-rep fit enhancing method, depicted in fig. 1.2, takes as input a smooth segmentation of an object (in the form of a mesh) with surface labels and a manual segmentation (again in the form of a mesh); it produces as output a warped version of the segmentation with the corresponding labels that matches the manual segmentation closely.

The inputs and outputs for this method are as follows:

- Inputs – a labeled mesh from an automatic segmentation and the binary image from the manual segmentation;
- Outputs – the input mesh with the labels carried over and vertices moved to match the manually segmented binary image.

This method moves mesh elements from an initial segmentation to a user-approved binary boundary. The initial mesh can be generated by an m-rep, and the user-approved boundary can be a manual segmentation, which may be anti-aliased by a method described in section 5.2. Surface labels are provided with the initial mesh and are carried with the mesh elements when they are moved.

First, I will detail this method, and then, in section 5.1.1, I will show results for applying the same to objects from the head and neck region.

The method propagates vertices of a mesh representing one surface along with neighbor and label information while warping one surface into another. To successfully propagate this information, it is important that the paths of propagation do not cross each other. Such non-crossing paths can be obtained from the streamlines of the solution of a diffusion equation. One of the types of a diffusion equation is Laplace's equation, which was discussed in section 3.3.1. Thus, streamlines of Laplace's equation can propagate labels from one surface to another. Setting one surface to a fixed value of 0 and another surface to a value of 1, and solving the Laplace's equation will produce equi-potential surfaces between the two surfaces with values between 0 and 1. Streamlines, which are perpendicular to these equi-potential surfaces, will be formed connecting the two surfaces. In order to produce the warped mesh, the vertices of one of the surfaces are propagated along these streamlines along with neighbor and label information.

A steady-state heat flow equation, which takes the form of Laplace's equation, was used by Dinh et al. [11] to interpolate 3D structures. It was also used by Jones et al. [21] and Yezzi et al. [36] to measure cortical and tissue thickness respectively. In my case, I want to propagate labels from an automatically segmented mesh M_0 to a manually segmented mesh

M_1. The shape model used produces good quality meshes. Therefore, it can be assumed that the two meshes are not significantly different from each other. Therefore, I use a simplified version of Dinh's method to solve the steady-state heat flow problem on a four-dimensional (4D) rectangular grid instead of a tetrahedral mesh. Solving Laplace's equation on a two-dimensional rectangular grid was detailed in section 3.3.2. The solution for four dimensions is a simple extension. To propagate labels, I solve a transport equation that propagates the labels along paths orthogonal to the equi-potential surfaces. The solution is similar to the one described by Dinh.

First, a 4D region with coordinates (x, y, z, t) is set up in which a steady-state heat flow problem is formulated and solved. The mesh M_0 is placed in the 3D facet at $(x, y, z, 0)$ with its temperature set to 0; this serves as a boundary condition. The mesh M_1 is placed in the 3D facet at $(x, y, z, 1)$ with its temperature set to 1; this serves as another boundary condition. The space is discretized to form a 4D rectangular grid with a resolution that balances computation time with precision. Von Neumann boundary conditions are used at the extremities of the grid. A steady-state heat flow problem is solved with these conditions to obtain the temperature T at every position in this 4D space.

Next, this temperature distribution is used to propagate the labels from the mesh M_0 to the mesh M_1. Let the labels be represented by a continuous function $\mathbf{L}$. $\mathbf{L}$ maps into a two-dimensional (2D) set such as $\mathbb{R}^2$ or $\mathbb{S}^2$ in order to represent a 2D surface. Let L_0 and L_1 be the individual components of this function. ∇L_i will always be non-zero along a surface and zero orthogonal to the surface. This leads to the following transport equation with the initial condition being the labels given for mesh M_0:

$$\begin{aligned} \nabla L_0 \cdot \nabla T &= 0, \\ \nabla L_1 \cdot \nabla T &= 0. \end{aligned} \tag{5.1}$$

This equation can be solved by the method of characteristics. The family of characteristic

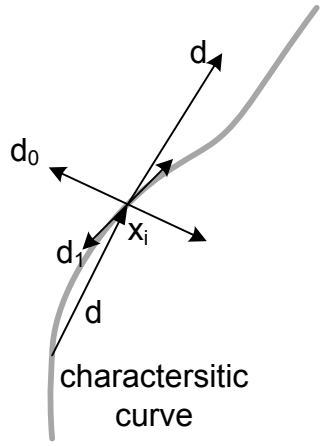


Figure 5.1: Three of the four differencing directions at point $\mathbf{x}_i$: $\mathbf{d}_0$ and $\mathbf{d}_1$ are orthogonal to $\mathbf{d}$ along which $\nabla T|_{\mathbf{x}_i}$ is computed.

curves, also called streamlines, for the transport equation above is given by the following equation:

$$\frac{\mathrm{d}x}{\partial T/\partial x} = \frac{\mathrm{d}y}{\partial T/\partial y} = \frac{\mathrm{d}z}{\partial T/\partial z} = \frac{\mathrm{d}t}{\partial T/\partial t}. \tag{5.2}$$

Along each streamline, the labels stay the same. This is used to transport the labels from mesh M_0 to mesh M_1. The streamline can be obtained by the following Euler integration with step size h:

$$\mathbf{x}_{i+i} \;=\; \mathbf{x}_i + h\nabla T|_{\mathbf{x}_i}. \tag{5.3}$$

For good accuracy, the differencing for ∇T has to be chosen carefully. If we knew the correct value of ∇T, forward differencing in the direction of ∇T will yield the correct value. However, we can only estimate the direction of ∇T based on prior temperature gradient directions. I forward difference in the prior gradient direction, $\nabla T|_{\mathbf{x}_{i-1}}$ and central difference in directions orthogonal to it. Let $\mathbf{d} = (x, y, z, w)$ be the normalized gradient direction. The directions $\mathbf{d}_0 = (-y, x, -w, z)$, $\mathbf{d}_1 = (w, z, -y, -x)$, and $\mathbf{d}_2 = (-z, w, x, -y)$ form a mutu-

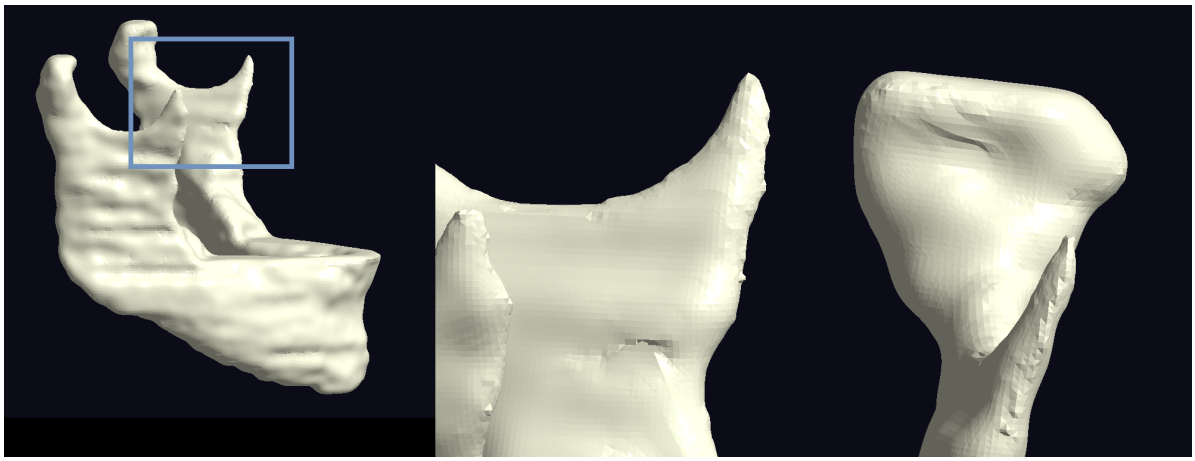


Figure 5.2: A lateral view of a mandible with the condyles enclosed in a box (left), zoomed in versions of the condyles – lateral view (center) and anterior view (right). The condyles are typically a voxel thick and pose numerical issues for the heat flow method.

ally orthogonal set of vectors. This results in the following differencing equation:

$$\begin{aligned}
\nabla T|_{\mathbf{x}_i} \quad = \quad & (T(\mathbf{x}_i + h \cdot \mathbf{d}) - T(\mathbf{x}_i)) \cdot \mathbf{d} \\
& (T(\mathbf{x}_i + \frac{h}{2} \cdot \mathbf{d}_0) - T(\mathbf{x}_i - \frac{h}{2} \cdot \mathbf{d}_0)) \cdot \mathbf{d}_0 \\
& (T(\mathbf{x}_i + \frac{h}{2} \cdot \mathbf{d}_1) - T(\mathbf{x}_i - \frac{h}{2} \cdot \mathbf{d}_1)) \cdot \mathbf{d}_1 \\
& (T(\mathbf{x}_i + \frac{h}{2} \cdot \mathbf{d}_2) - T(\mathbf{x}_i - \frac{h}{2} \cdot \mathbf{d}_2)) \cdot \mathbf{d}_2. \qquad (5.4)
\end{aligned}$$

Fig. 5.1 depicts three of these four differencing directions visually.

5.1.1 Results

I applied this method to several organs in the head and neck and measured the average increase in the quality of the fit for each of these organs. The quality of fit can be measured by the median of the distance between the surfaces of the automatic segmentation and the manual segmentation. Table 5.1 shows the average initial, average final, and average improvement in this measure for several organs. The decrease in this measure is indicative of an increase in the quality of the fit of these organs. Ideally, the distance should fall to zero. However, extremely

Organ	Average median surface distance (mm)		
	Initial	Final	Decrease
skin surface	9.38	0.85	8.53
mandible	1.09	0.34	0.74
trachea	1.57	0.14	1.42
left clavicle	1.18	0.31	0.87
right clavicle	1.17	0.34	0.83

Table 5.1: Average initial, final, and decrease in the median of the surface distances for several organs by enhancing the m-rep fit using a heat flow based method.

thin regions such as the condyles of the mandibles shown in fig. 5.2, and the resolution of the grid at which the method can be run staying within memory limits and with reasonable computation times (on the order of tens of minutes) prevent this measure from falling to zero.

5.2 Entropy-based Correspondence

The entropy-based correspondence method, depicted in fig. 1.3, takes as input labeled segmentations of the organ from both the atlas and the patient and produces a labeled segmentation of the patient's organ with improved correspondence to the atlas segmentation.

The inputs and outputs for this method are as follows:

- Inputs – smooth segmentations of the object in the atlas and the target patient images in the form of labeled meshes (output of previous method);
- Outputs – correspondence labels on the boundary vertices on the same segmentations in both the atlas and the target patient images.

The improved matching of the mesh to the manual segmentation from the previous method still may not accurately reflect the anatomical correspondence. If there were a solution to detecting and matching hundreds of anatomical landmarks automatically, most of this thesis would be redundant. We can, however, identify geometry-based correspondence in the objects, which may closely resemble anatomical correspondence. Geometric attributes such as

position and curvatures can be used to improve the correspondence. These geometric cues are taken into account by an entropy-based correspondence method, developed by Oguz et al. [27].

First, I will overview Oguz's entropy-based correspondence method, and then, in section 5.2.1, show the result of applying it to objects from the head and neck.

Good correspondence between surface points across a population (ensemble) may be identified by having features that are close to each other. Further, it is desirable to have these correspondences between points that cover the surface as uniformly as possible, so that no place is left out. The points are represented by a set of particles, P^k on the surface of the k^{th} object. The correspondence is measured by the entropy of the object features at the particle locations across the entire population; the lower the entropy, the closer is the match between object features at particle locations between different objects of the population. The uniformity in the spread is measured by the entropy of the particle positions; the higher the entropy, the more uniformly the particles are spread out. Thus, if H represents the entropy function and Z is the collection of object features at the particle locations, we need to minimize the entropy of the collection of object features at the particle locations $H(Z)$ and maximize the entropy of of the particle positions $\sum_k H(P^k)$. Thus, we need to minimize Q in

$$Q = H(Z) - \sum_k H(P^k). \tag{5.5}$$

Both the terms are in units of entropy (bits). Experimentally, it has been shown that weighting the two terms equally produces good results.

Spatial location and surface curvature are good examples of geometric features. In [5], Cates has used as Z the positions of the particles in 3D space after all the objects have been aligned through a similarity transform. In [27], Oguz has used curvature-based features in addition to these location-based features and has shown that the correspondence produced is more meaningful. In applying this method to my problem, I use two curvature-based features

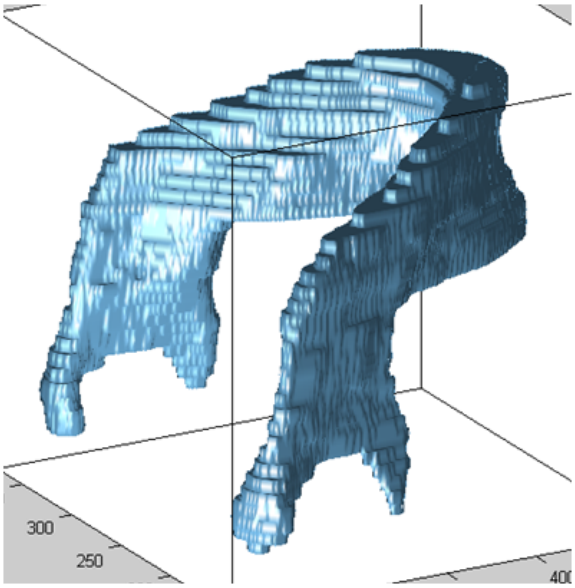


Figure 5.3: A manually segmented mandible. Note the severe aliasing.

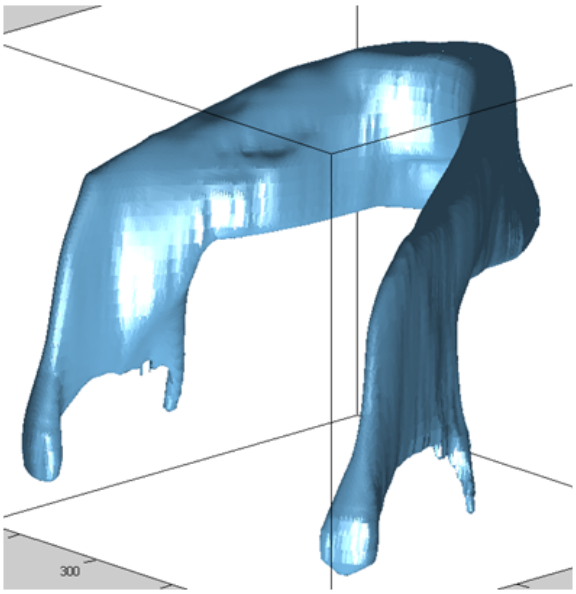


Figure 5.4: An anti-aliased version of the mandible shown on the left.

– mean curvature H and Gaussian curvature K – in addition to the location-based features.

To measure these features, the particles need to lie on the actual object boundary. The errors left by the previous method are corrected by projecting the point to the closest location on the real boundary before the optimization process begins. This reduces the errors reported in table 5.1 to zero.

This method is dependent on position and curvatures. Curvatures are hard to extract from unprocessed aliased manual segmentations. An aliased mandible is shown in fig. 5.3. To extract meaningful curvatures, this aliased segmentation needs to be smoothed. Several smoothing techniques exist in the literature, but most shrink the object or change the topology of thin regions. A full discussion of such preprocessing techniques is not relevant to this thesis. I have developed a method that anti-aliases the object by smoothing the curvature and constraining the movement to lie within the voxel limits. It evolves a signed distance function of the object, while restricting voxels that could through evolution change the topology of the object. Thus, this method maintains geometry and topology while staying within the constraints. An anti-aliased mandible is shown in fig. 5.4. A full paper [30] on this method is in preparation

by me as of the time of writing this thesis.

5.2.1 Results

I applied the above series of methods in order to produce surfaces with correspondence between them for several structures in the head and neck for six patients – the skin surface, mandible, trachea, and left and right clavicles. Table 5.2 shows the entropy of the distribution of corresponding surface points and the absolute distance they moved as a result of the entropy-based correspondence method and the projection operation from their positions after applying the m-rep fit enhancing method. A smaller value for the entropy indicates better performance by the entropy-based correspondence method. A larger movement of the surface points suggests the usefulness of the entropy-based correspondence method towards improving the correspondence between the points on the different surfaces.

Fig. 5.5, 5.6, and 5.7 show the corresponding points (highlighted by the same color) on tracheas, left clavicles, and right clavicles respectively of the six patients. The points in correspondence across the different patients visually correspond well.

The head and neck skin surface has a large entropy compared to the other objects, because of the somewhat arbitrary cutoff at the top and the bottom and problems with finding correspondence near the ears, which are very difficult to represent.

The entropy for the mandibles is also somewhat high. This is explained by the significantly different shape of the mandible of the patient in the first column and the second row in fig. 5.8. Further, the entropy-based correspondence method has difficulties in moving particles on the surface of a very thin object – at times the particles jump from one side to the other. This particular patient's mandible is also thin – at times as thin as a voxel. Some of the particles visible at the front of this mandible and not found at the front of other mandibles actually belong to the surface on the side facing away. Barring this exception, the correspondence visually looks good on all the other mandibles.

Organ	Entropy	Average movement (mm)
skin surface	12.10	1.72
mandible	10.57	0.75
trachea	8.99	0.38
left clavicle	9.91	0.65
right clavicle	9.53	0.82

Table 5.2: Average ensemble entropy and movement for the points. The length of a side of a voxel is approximately 0.4 mm.

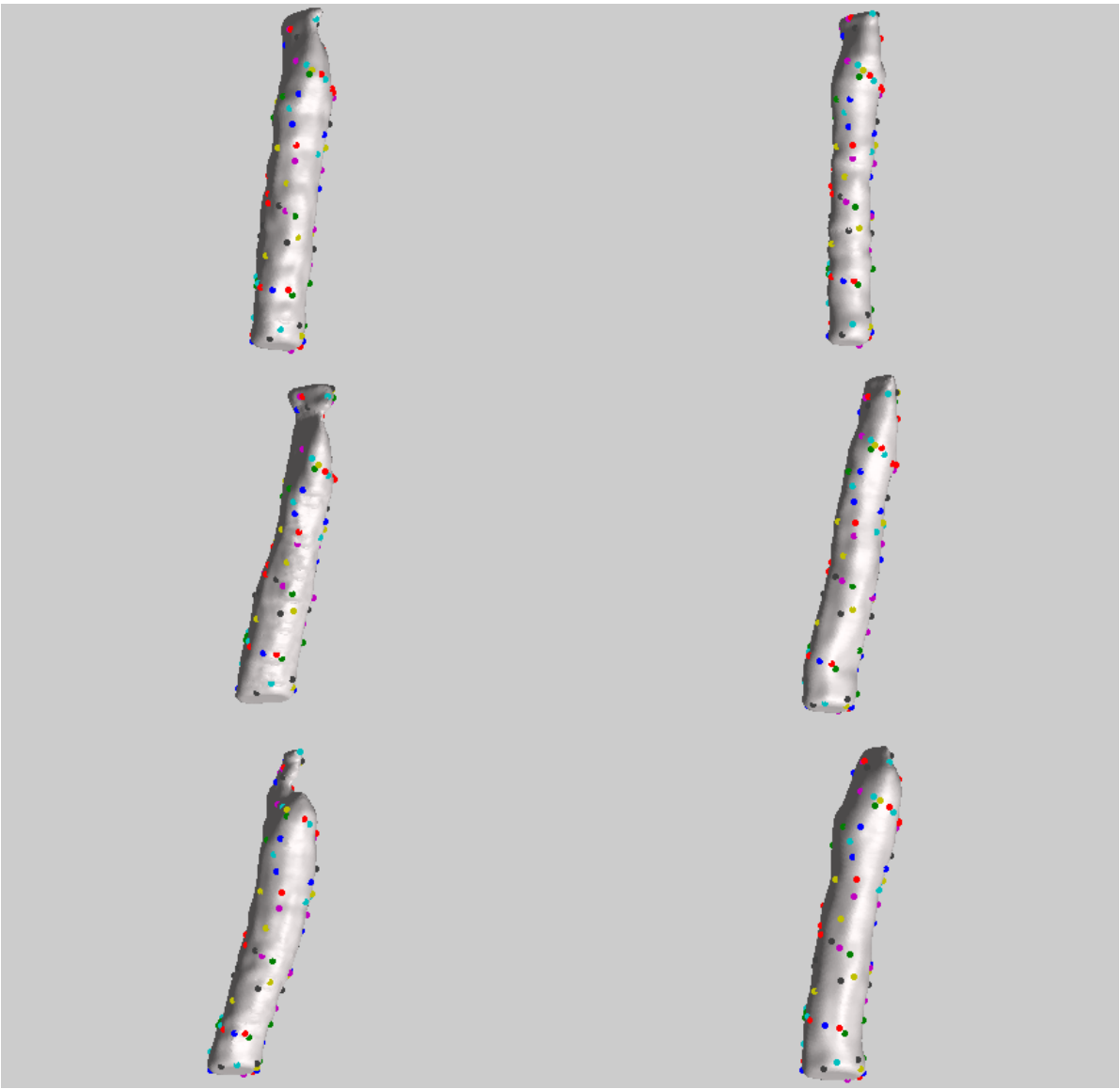


Figure 5.5: Corresponding points after entropy-based correspondence between tracheas of six different patients.

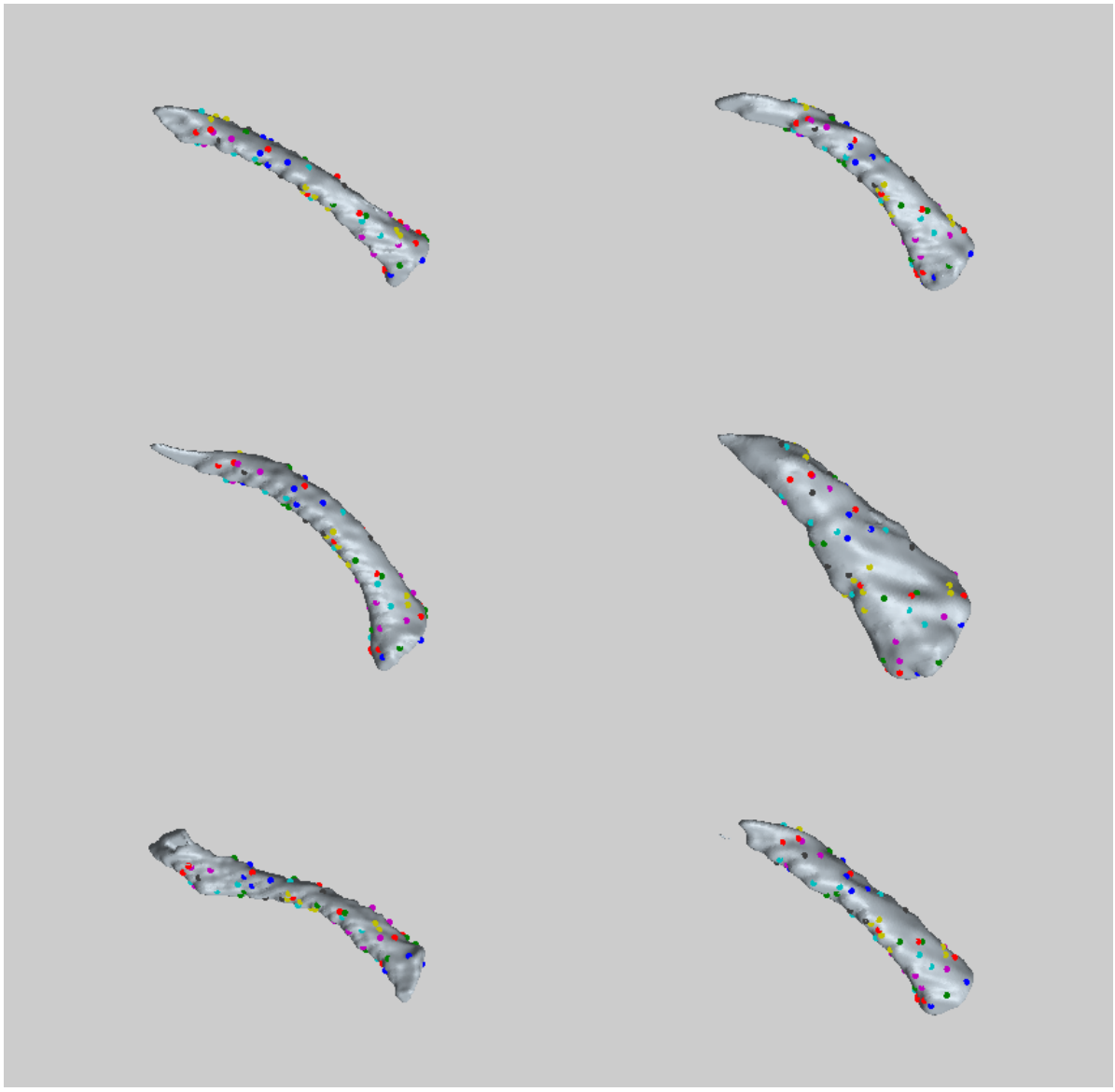


Figure 5.6: Corresponding points after entropy-based correspondence between left clavicle bones of six different patients.

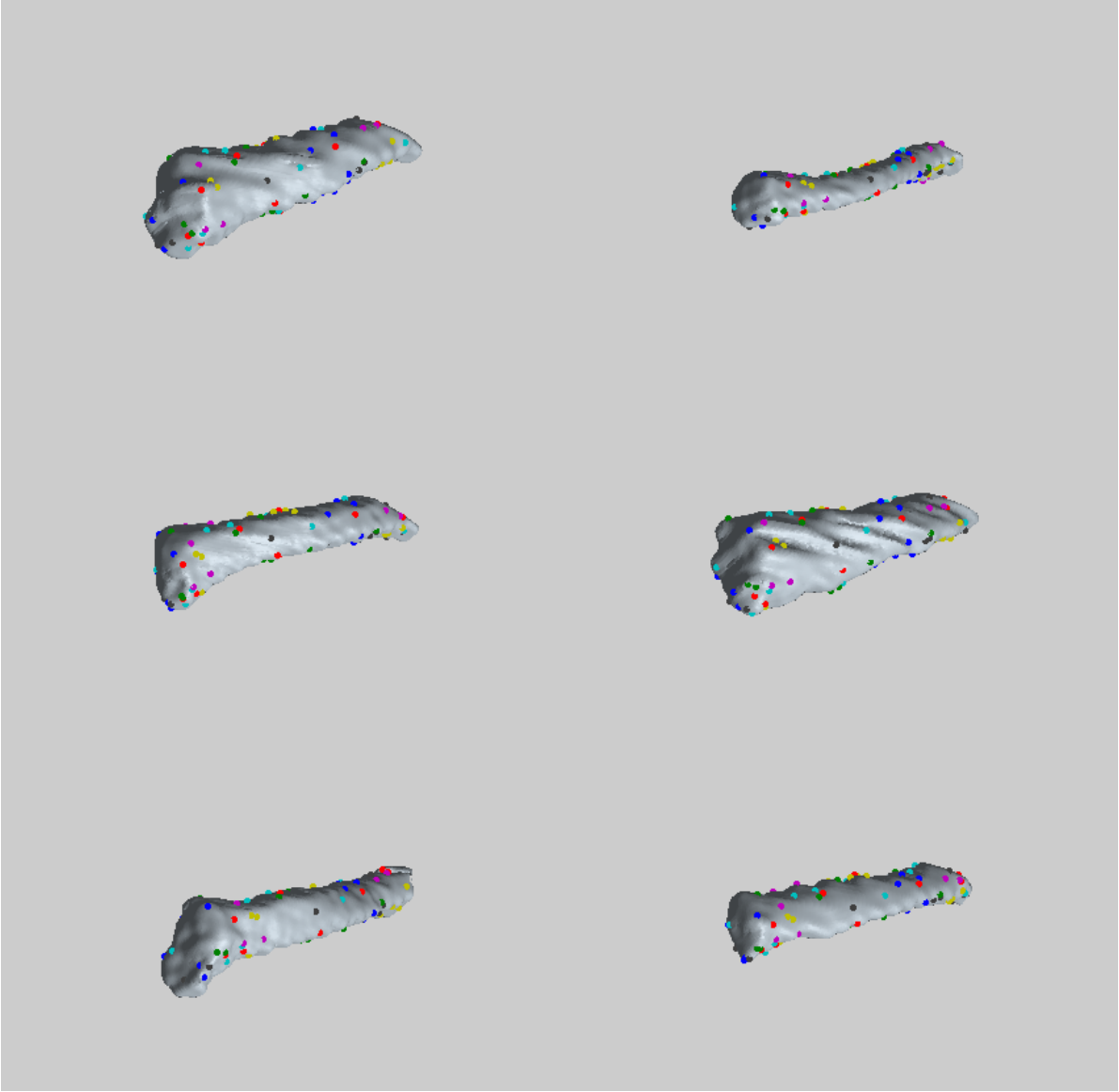


Figure 5.7: Corresponding points after entropy-based correspondence between right clavicle bones of six different patients.

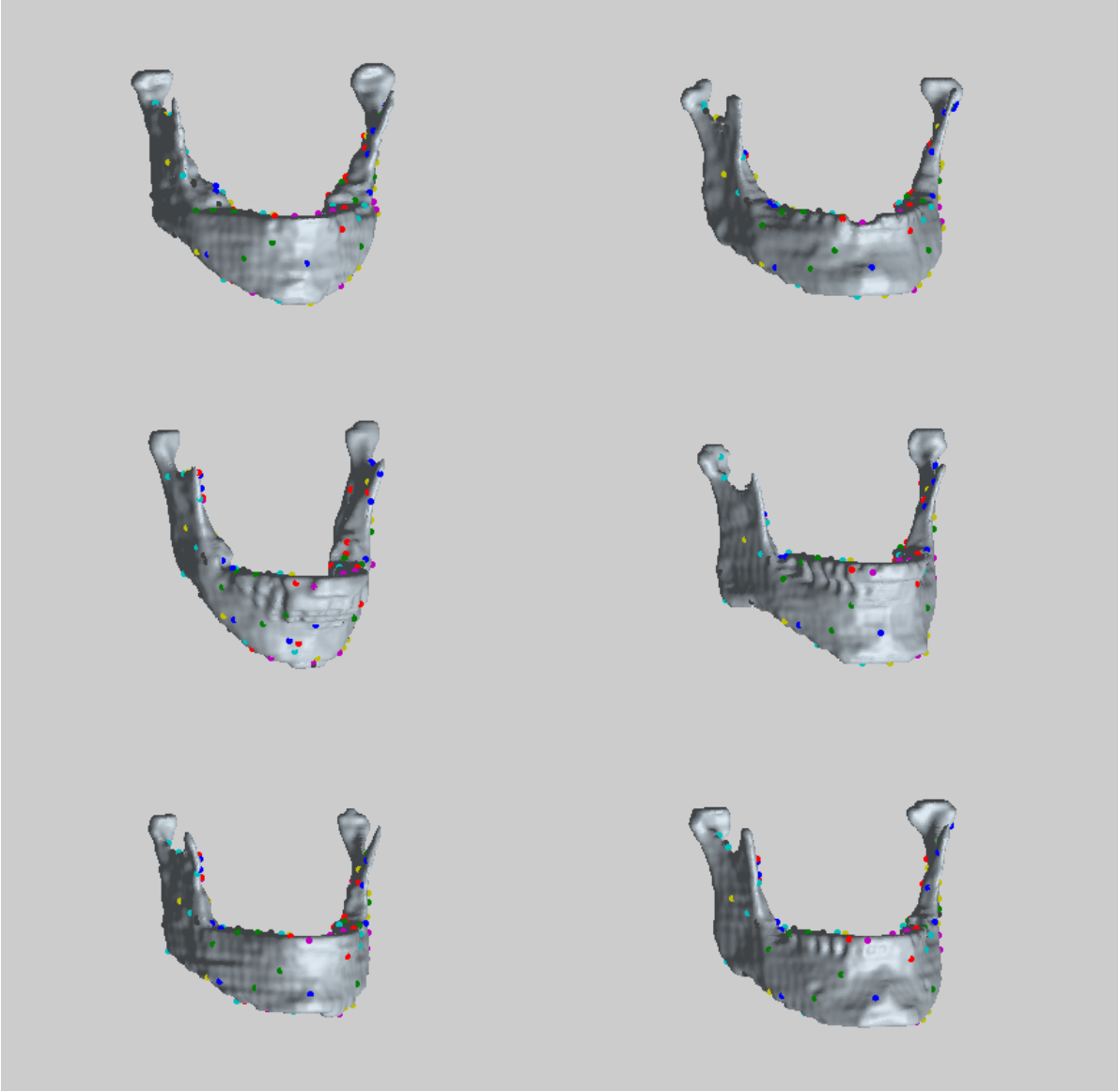


Figure 5.8: Corresponding points after entropy-based correspondence between mandibles of six different patients.

Chapter 6

Non-folding, Smooth Volume Warp

In the previous chapters, I have detailed a method that produces correspondence between two surfaces. In this chapter I develop a method, based on steady-state heat flow, that produces an almost diffeomorphic warp, specifically a warp that does not have folds and is smooth, between two images while respecting this correspondence between them.

The warping method assigns the same unique temperature tuple for corresponding points between two images, and through the solution of a steady-state heat flow problem on one of them, it extends this correspondence to the rest of the image. Corresponding points between the atlas and the target image are assigned the same tuple of temperatures; a distinct temperature value is assigned for each dimension, and they are all independent of one another. Section 6.1 details the method of creating a temperature distribution, first on the atlas image and then through implication of the correspondences on object boundaries in the target image. These object boundary temperatures in the target image are used as boundary conditions in a steady-state heat flow problem. Section 6.2 details the method of solving this problem. As a result of solving this, a point in the atlas will have a corresponding point in the target image with the same tuple of temperatures. This correspondence is used to map points in the target image back to the atlas image.

The mapping obtained in this way may be folded. Section 6.3 develops a measure of foldedness. In order to lower this measure, I let the heat conduction coefficient vary across

the image while solving the steady-state heat flow problem. The precise way in which this is done is detailed in section 6.4. In section 6.5 I present results generated by this method on a synthetic data set. Finally, in section 6.6 I present some discussion and directions for future work on this method.

6.1 Temperature Distribution

This section details the method of creating a temperature distribution on the source and target images. The coordinate system provides one such way to uniquely assign temperatures to every point in n-dimensional space, such that the temperature changes smoothly over space; the coordinate system is a smooth function that assigns a unique n-tuple of real numbers to every point in n-dimensional space. Specifically, I pick the tuple of temperatures from the Cartesian coordinate system.

In 3D, a point at location (x, y, z) is assigned the tuple of temperatures (x, y, z) in the source image $\mathcal{A}$. The same temperature tuple (x, y, z) is assigned to the corresponding point at location (x', y', z') in the target image $\mathcal{B}$. This is done for each point on each surface for which we have a corresponding point in the target image.

However, the steady-state heat flow solution is computed on a finite resolution grid, where the points on the target image will rarely co-incide with the grid points. Here, the temperatures at the grid points need to be estimated from the points in the target image that have a temperature assigned; the simple method above that works over continuous domains does not work anymore. We need a way to interpolate temperatures. All close points in the target image contribute in determining its temperature. In order to estimate the temperature at a distance from a point, we need to estimate a local temperature distribution around each point. For each point in the target image, $\mathbf{p}$, a temperature tuple distribution $\mathbf{T_p}(\mathbf{x}) = (T^0_{\mathbf{p}}(\mathbf{x}), T^1_{\mathbf{p}}(\mathbf{x}), T^2_{\mathbf{p}}(\mathbf{x}))$ is estimated from a small, corresponding region, where $\mathbf{x}$ is a point in the neighborhood of $\mathbf{p}$ in the space of the target image. The contribution from this distribution is weighted by a

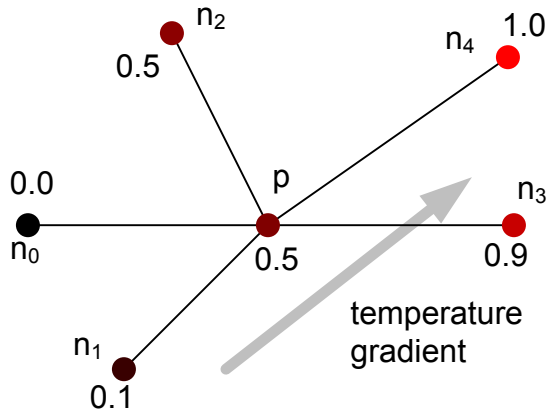


Figure 6.1: A point $\mathbf{p}$ is shown with its neighboring points $\mathbf{n}_0 \ldots \mathbf{n}_4$ in the space of the target image. The colors range from black for a low temperature to red for a high temperature. The gray arrow shows the direction of the local temperature gradient.

weighting function $w(d)$ of the distance d to the center of the grid point where the resulting temperature tuple is being evaluated. The target points $\mathcal{N}_\mathbf{x}$ that lie closer to the grid point $\mathbf{x}$ than to any other grid point contribute to the temperature distribution $\mathbf{T}(\mathbf{x})$ at the grid point $\mathbf{x}$:

$$\mathbf{T}(\mathbf{x}) = \frac{\sum_{\mathbf{p}\in\mathcal{N}_\mathbf{x}} \mathbf{T}_\mathbf{p}(\mathbf{x}) w(|\mathbf{p}-\mathbf{x}|)}{\sum_{\mathbf{p}\in\mathcal{N}_\mathbf{x}} w(|\mathbf{p}-\mathbf{x}|)}. \tag{6.1}$$

The weighting function $w(d)$ is designed such that it falls off to 0 at a distance $\sqrt{3}$ image voxel units away; $\sqrt{3}$ is chosen because it is the maximum distance between two points on a unit cube. This weighting function is built from a Gaussian function $G(d)$ with a mean of 0 and a standard deviation σ of 0.5:

$$G(d) = \frac{1}{\sqrt{2\pi}\,\sigma} \mathrm{e}^{-\frac{d^2}{2\sigma^2}}, \tag{6.2}$$

$$w(d) = \max\left\{0, \frac{G(d) - G(\sqrt{3})}{1 - G(\sqrt{3})}\right\}. \tag{6.3}$$

The local temperature distribution $\mathbf{T}_\mathbf{p}(\mathbf{x})$ in the space of the target image for a point $\mathbf{p}$ is estimated from the temperature differences with neighboring points in a manner that is not sensitive to any particular temperature. Fig. 6.1 shows the point $\mathbf{p}$ along with five neighboring

points $\mathbf{n}_0 \ldots \mathbf{n}_4$. The underlying temperature gradient $\nabla \mathbf{T}_\mathbf{p}$ is shown by the gray arrow. The temperature at a neighboring point $\mathbf{n}_i$ is related to the temperature at point $\mathbf{p}$ by the relation

$$\mathbf{T}_{\mathbf{n}_i} = \mathbf{T}_\mathbf{p} + (\mathbf{n}_i - \mathbf{p}) \cdot \nabla \mathbf{T}_\mathbf{p}. \tag{6.4}$$

For m neighboring points, where $m > 3$, this results in an over-specified system of equations, with which we need to solve for the unknown $\nabla \mathbf{T}_\mathbf{p}$:

$$\underbrace{\begin{bmatrix} \mathbf{n}_0 - \mathbf{p} \\ \vdots \\ \mathbf{n}_i - \mathbf{p} \\ \vdots \\ \mathbf{n}_m - \mathbf{p} \end{bmatrix}}_{\mathbf{D}} \underbrace{\begin{bmatrix} \frac{\partial T^0_\mathbf{p}}{\partial x} & \frac{\partial T^1_\mathbf{p}}{\partial x} & \frac{\partial T^2_\mathbf{p}}{\partial x} \\ \frac{\partial T^0_\mathbf{p}}{\partial y} & \frac{\partial T^1_\mathbf{p}}{\partial y} & \frac{\partial T^2_\mathbf{p}}{\partial y} \\ \frac{\partial T^0_\mathbf{p}}{\partial z} & \frac{\partial T^1_\mathbf{p}}{\partial z} & \frac{\partial T^2_\mathbf{p}}{\partial z} \end{bmatrix}}_{\text{unknown } \nabla \mathbf{T}_\mathbf{p}} = \underbrace{\begin{bmatrix} T^0_{\mathbf{n}_0} - T^0_\mathbf{p} & T^1_{\mathbf{n}_0} - T^1_\mathbf{p} & T^2_{\mathbf{n}_0} - T^2_\mathbf{p} \\ & \vdots & \\ T^0_{\mathbf{n}_i} - T^0_\mathbf{p} & T^1_{\mathbf{n}_i} - T^1_\mathbf{p} & T^2_{\mathbf{n}_i} - T^2_\mathbf{p} \\ & \vdots & \\ T^0_{\mathbf{n}_m} - T^0_\mathbf{p} & T^1_{\mathbf{n}_m} - T^1_\mathbf{p} & T^2_{\mathbf{n}_m} - T^2_\mathbf{p} \end{bmatrix}}_{\mathbf{T}_{\mathbf{n}_i} - \mathbf{T}_\mathbf{p}}. \tag{6.5}$$

The system above is solved for the unknown $\nabla \mathbf{T}_\mathbf{p}$ by a weighted least squares method, where the weights $\mathbf{W}$ are given by the weighting function $w(d)$ of the distance of the neighboring point $\mathbf{n}_i$ to the point $\mathbf{p}$:

$$\mathbf{w} = \begin{bmatrix} w(|\mathbf{n}_0 - \mathbf{p}|) \\ \vdots \\ w(|\mathbf{n}_i - \mathbf{p}|) \\ \vdots \\ w(|\mathbf{n}_m - \mathbf{p}|) \end{bmatrix}. \tag{6.6}$$

Therefore, we have

$$\nabla \mathbf{T}_\mathbf{p} = \underset{\nabla \mathbf{T}_\mathbf{p}}{\operatorname{argmin}}((\mathbf{T}_{\mathbf{n}_i} - \mathbf{T}_\mathbf{p}) - \mathbf{D} \times \nabla \mathbf{T}_\mathbf{p})' \times \operatorname{diag}(\mathbf{w}) \times ((\mathbf{T}_{\mathbf{n}_i} - \mathbf{T}_\mathbf{p}) - \mathbf{D} \times \nabla \mathbf{T}_\mathbf{p}), \tag{6.7}$$

where $\operatorname{diag}(\mathbf{w})$ returns a diagonal matrix with elements from the vector $\mathbf{w}$. Only the points for which the weight is greater than zero are considered as neighbors in the system of equations

above.

A steady-state heat flow solution is found by the method described in the next section with the temperature distribution described in this section as a set of fixed boundary conditions.

6.2 Solving the Steady-state Heat Flow Equation

The temperatures at the corresponding points are used as one set of boundary conditions. Another set of Dirichlet boundary conditions with the temperature picked from the same co-ordinates in the source image is added for the actual boundaries of the image. This system is then solved by the method of red-black successive over-relaxations, discussed in section 3.3.2, in a multi-scale fashion.

The grid over which the system is to be solved is first down-sampled to a smaller grid. Red-black SOR iterations are performed on this grid, and a solution is obtained. (See Numerical Recipes [29] for a discussion of this method.) Next, the solution is interpolated to a higher resolution, and the iterations are continued at this higher resolution. The process is continued in this manner until the solution is obtained at the desired grid resolution.

This process is repeated for each of the members of the tuple of temperatures.

6.3 How Folded Is It?

In this section, I develop a quantitative measure of the amount of foldedness in a warp. The warp obtained as a result of the method above may be folded. To remove any possible folds, we need to locally distort the space. This is done by changing the heat conduction coefficient at every point in space inside an optimizer. Doing this requires an objective function that measures the foldedness of the warp. I first present this measure for a continuous domain and then, in the next section, present it for a discrete domain.

The volume of the folded region serves well as a measure of foldedness. This volume can be obtained by the Jacobian of the warp. The determinant of the Jacobian of a warp is

the factor by which the volume has locally grown or shrunk. If it is negative, the volume has folded onto itself, and the warp is no longer a one-to-one mapping. For a warp $\mathbf{h}(\mathbf{x}) = (h_x(\mathbf{x}), h_y(\mathbf{x}), h_z(\mathbf{x}))$ mapping points in image $\mathcal{B}$ into image $\mathcal{A}$, the Jacobian is given by

$$\mathbf{J} = \begin{bmatrix} \frac{\partial h_x}{\partial x} & \frac{\partial h_y}{\partial x} & \frac{\partial h_z}{\partial x} \\ \frac{\partial h_x}{\partial y} & \frac{\partial h_y}{\partial y} & \frac{\partial h_z}{\partial y} \\ \frac{\partial h_x}{\partial z} & \frac{\partial h_y}{\partial z} & \frac{\partial h_z}{\partial z} \end{bmatrix}. \tag{6.8}$$

Thus, the measure of foldedness $N(\mathbf{h})$ can be obtained by integrating, over the entire volume, the negative values of the determinant of the Jacobian of the warp:

$$|x|^{-} = \begin{cases} 0 & \forall x \geq 0 \\ x & \forall x < 0 \end{cases} \tag{6.9}$$

$$N(\mathbf{h}) = -\iiint |\,|\mathbf{J}(\mathbf{h})|\,|^{-}\,\mathrm{d}V. \tag{6.10}$$

Computing this measure accurately over the discrete domain is somewhat tricky. The next section describes how this measure of foldedness is computed for the discrete domain.

6.3.1 Measure of Foldedness in the Discrete Case

In the discrete domain, the space is divided into cubes, where each corner of a cube is a point of the grid on which the warp is computed. Each cube is further subdivided into five tetrahedra. This subdivision is shown in fig. 6.2. This measure is similar to computing the Jacobian on a discrete domain, which was first presented in Haber and Modersitzki [15].

In order to find the volume of the folded region, we now need to find the volume of the

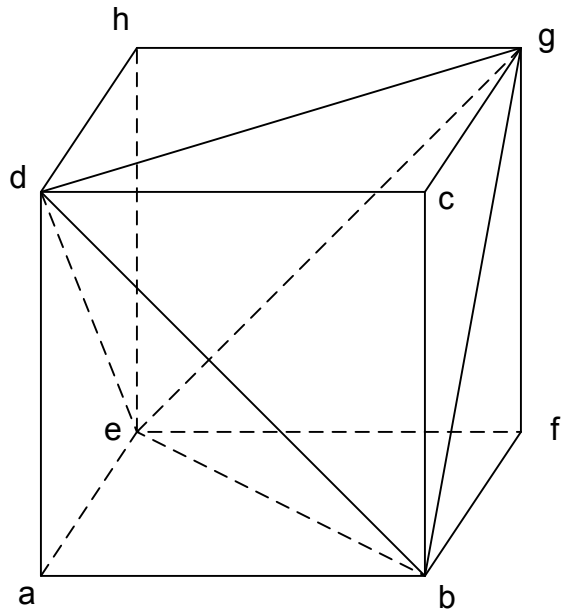


Figure 6.2: Subdivision of a cube into five tetrahedra: abed, hedg, befg, bcgd, and gbde. a,b,c,d,e,f,g, and h are points on the grid where the warp is computed.

folded tetrahedra $\mathbf{abcd}$ when mapped back into image $\mathcal{A}$, which is given by

$$\begin{aligned} V &= \frac{1}{6}\left[\, \mathbf{h}(\mathbf{b}) - \mathbf{h}(\mathbf{a}),\ \mathbf{h}(\mathbf{c}) - \mathbf{h}(\mathbf{a}),\ \mathbf{h}(\mathbf{d}) - \mathbf{h}(\mathbf{a}) \,\right] \\ &= \frac{1}{6}\left(\mathbf{h}(\mathbf{b}) - \mathbf{h}(\mathbf{a})\right) \cdot \left(\left(\mathbf{h}(\mathbf{c}) - \mathbf{h}(\mathbf{a})\right) \times \left(\mathbf{h}(\mathbf{d}) - \mathbf{h}(\mathbf{a})\right)\right). \end{aligned} \tag{6.11}$$

If the vectors $\mathbf{ab}$, $\mathbf{ac}$, and $\mathbf{ad}$ form a right-handed coordinate system, this volume will be negative if the mapped tetrahedra has folded, i.e., the mapped vectors have changed their handedness; if the vectors form a left-handed coordinate system, the volume will be positive if the mapped tetrahedra has folded. I start by choosing an ordering of the vertices a,b,c, and d of the tetrahedra such that the vectors form a right-hand coordinate system. Thus, the measure of foldedness in the discrete domain is obtained by summing the values of all tetrahedra i with negative volumes:

$$N(\mathbf{h}) = -\sum_i \left| \frac{1}{6}\left[\, \mathbf{h}(\mathbf{b}_i) - \mathbf{h}(\mathbf{a}_i),\ \mathbf{h}(\mathbf{c}_i) - \mathbf{h}(\mathbf{a}_i),\ \mathbf{h}(\mathbf{d}_i) - \mathbf{h}(\mathbf{a}_i) \,\right] \right|^{-}. \tag{6.12}$$

6.4 Removing Folds from a Warp

This section details the way in which the warp obtained by solving a steady-state heat flow equation is devoided of its folds. This is achieved by solving the steady-state heat flow problem on an inhomogenous medium. The conductivity c of the medium is changed locally such that the foldedness of the warp is reduced to 0.

The inhomogenous steady-state heat flow problem for temperature ϕ is given by the following equation:

$$\begin{aligned} \nabla(c\nabla\phi) &= 0, \\ \implies \nabla c \cdot \nabla\phi + c\Delta\phi &= 0. \end{aligned} \tag{6.13}$$

The discrete form of this equation, with spacing h, in 2D is given by

$$\begin{aligned} 0 = &\left(\frac{c_{i+1,j} - c_{i-1,j}}{2h}\right)\left(\frac{\phi_{i+1,j} - \phi_{i-1,j}}{2h}\right) + \\ &\left(\frac{c_{i,j+1} - c_{i,j-1}}{2h}\right)\left(\frac{\phi_{i,j+1} - \phi_{i,j-1}}{2h}\right) + \\ &c_{i,j}\frac{\phi_{i+1,j} + \phi_{i-1,j} + \phi_{i,j+1} + \phi_{i,j-1} - 4\phi_{i,j}}{h^2}. \end{aligned} \tag{6.14}$$

It is solved in the same manner as the steady-state heat flow equation for a homogenous medium.

I minimize the foldedness of the warp $\mathbf{h}$ by minimizing the function $N(\mathbf{h})$, which measures the total folded volume in the warp, over a discrete grid of conductivity coefficients c_i. Several different combinations of c_i may be able to achieve the minimum of 0 for $N(\mathbf{j})$. A combination is chosen such that the change in c_i from its original values c_i^0 is kept to a minimum; specifically, I minimize the p-norm of the change in $c - i$:

$$\{c_i\} = \underset{\{c_i\}}{\operatorname{argmin}} \sum |c_i - c_i^0|^p \text{ s.t. } N(\mathbf{h}) = 0. \tag{6.15}$$

Ordinarily an optimization might use a gradient descent scheme. However, computing the $N(\mathbf{h})$ is an expensive operation – it requires solving the steady-state heat flow problem albeit using the last available solution. Thus the optimization above can be prohibitively expensive. Therefore, a greedy version of the gradient descent method is used to update the value of the coefficients c_i.

In each round of the optimization process, the derivative $\partial N(\mathbf{h})/\partial c_i$ is computed numerically for a conductivity coefficient c_i at a time. If there is an improvement observed in the foldedness, c_i is updated by a small multiplicative (or divisive) factor in the direction of improvement. If no improvement is observed, the coefficient c_i is marked to not be visited again for the next r rounds. Empirically, it has been observed that a coefficient that does not produce an improvement in one round rarely produces an improvement in later rounds of the optimization process. After processing for the coefficient c_i is complete, the optimization progresses to another coefficient c_j. The optimization ceases when no more improvement is possible.

The order in which the coefficients are chosen in the optimization process can significantly affect the convergence speed. I order the coefficients based on their proximity to locations where the warp is folded; the closest coefficients are chosen first in the optimization order. This not only tries to keep the change in c_i to a minimum, but also ensures that the time spent by the optimization process is on coefficients closest to the folded region, i.e., on the coefficients that have the highest chance of reducing the foldedness.

At the end of this optimization process, typically $N(\mathbf{h})$ converges to 0 guaranteeing that the warp is smooth and devoid of folds. However, if it were to not converge to 0, the warp will have folds.

The next section details the result of applying this method on synthetic concentric sphere data.

6.5 Experiments

I tested this method by generating warps between pairs of concentric spheres of different radii and evaluating the warp of the space in between these concentric spheres.

I generated two concentric spheres centered at the origin with radii r_1 and r_2, such that $r_2 > r_1$. I scaled the outer sphere by a factor α so that its new radius is αr_2. The space in between the two given spheres can be understood as a continuum of spheres centered on the origin with radii between r_1 and r_2. A spatial grid of size 64^3 was used to compute the correspondence-preserving smooth, non-folding warp. In this grid, the diameter of the outer sphere spans a distance of 52.5 grid points. After computing and applying the warp, I measured the deviation of the warped spheres from a spherical shape and the distance traveled by the points along the surface of each sphere from where they were expected to be. The first measure reports the movement of points in a direction orthogonal to the surface of the sphere, and the second measure reports movement of points in a direction along the surface of the sphere.

The deviation from a spherical shape is computed in the following manner: Ten spheres are picked in between the two given spheres. Their radii are computed as the mean distance of the points on the warped spheres from the origin. The standard deviation of the difference and the maximum absolute difference between the distance of the points from the origin and the radius of the sphere they belong to is recorded. The result in shown in fig. 6.3 as a function of the scale α. The deviation of the warped sphere from a spherical shape is quite small, which shows that the warp is uniform in all radial directions.

The distance traveled by the points along the surface of each sphere from where they were expected to be is computed by measuring the angle subtended at the sphere center $\mathbf{c}$ by which the surface points moved away from their true position (say from $\mathbf{p}$ to $\mathbf{p}'$) and multiplying it with the warped radius r of the sphere to which the surface point belonged, i.e., the distance

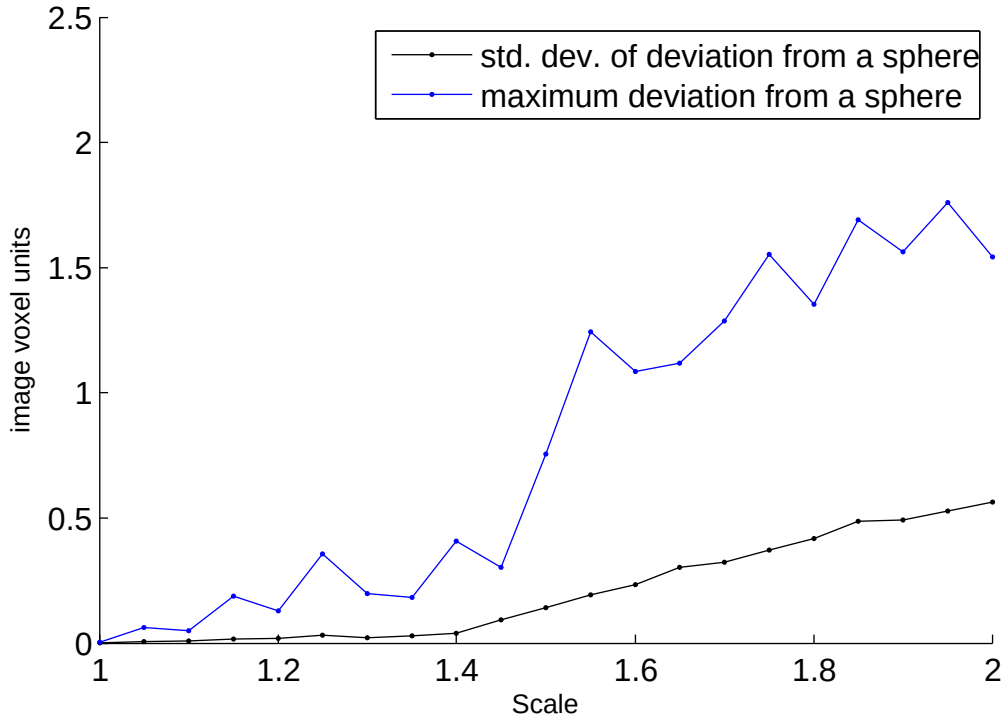


Figure 6.3: The standard deviation and the maximum deviation of the concentric spheres from a spherical shape in a warp field between two spheres as the outer sphere was scaled by a factor α. The data is presented in units of image voxels.

travelled is

$$r \times \cos^{-1}\left(\frac{(\mathbf{p}-\mathbf{c})\cdot(\mathbf{p}'-\mathbf{c})}{|\mathbf{p}-\mathbf{c}||\mathbf{p}'-\mathbf{c}|}\right). \tag{6.16}$$

The mean and standard deviation of this measure were recorded and are shown in fig. 6.4. From this, we can see that the distance traveled by the points along the surface of each sphere is quite small; thus, the warp does not produce unwanted expansions or contractions in space in any single radial direction.

Finally, in fig. 6.5, I show the radii of the warped spheres as a function of their original radius for the scale factor of 2. This relationship is almost linear and shows that the warp in the space between the spheres is almost uniform.

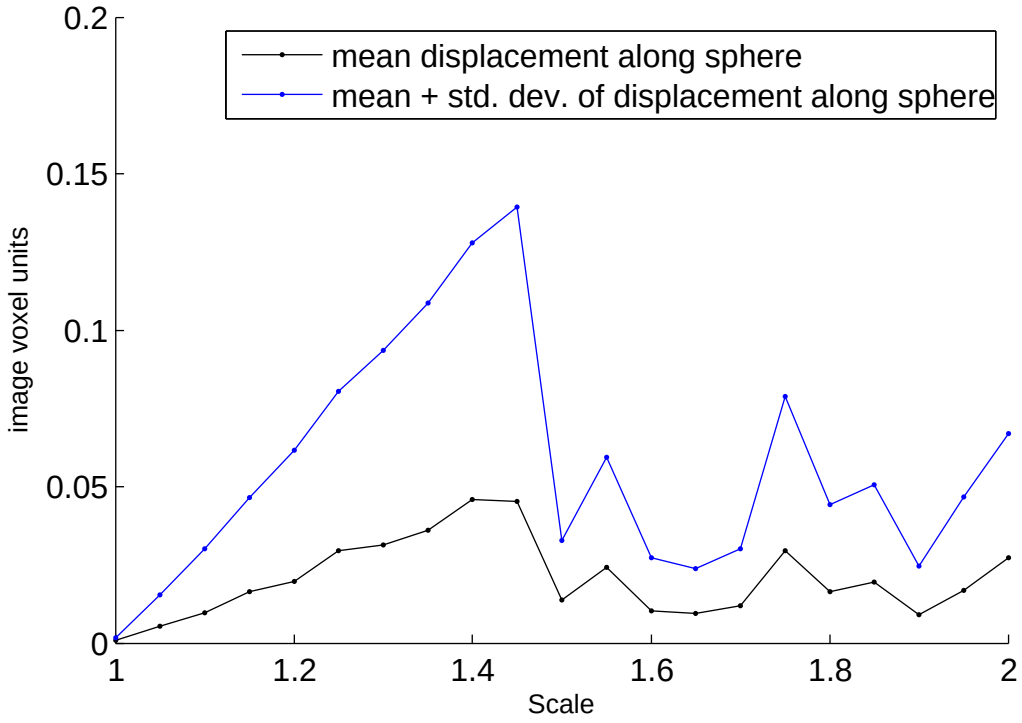


Figure 6.4: The mean and the standard deviation of the distance moved by points along the surface of concentric spheres from their true position in a warp field between two spheres as the outer sphere was scaled by a factor α. The data is presented in units of image voxels.

6.6 Discussion

The entire method is summarized in algorithm 2.

From, the experiments shown in the previous section on synthetic data, we can see that the method produces reasonable smooth, non-folded warps. Nevertheless, there are three ways in which this method can be improved.

The first way in which this method can be improved is by changing the energy function to include other terms such as an image match term so as to be able to produce smooth, non-folded warps that not only respect correspondence but also match image intensities. An example of such an energy function that introduces a low-weighted image match term ($w_b << w_a$) is

$$E(\mathbf{h}) = w_a N(\mathbf{h}) + w_b \iiint |I_{\mathcal{A}}(\mathbf{h}(\mathbf{x})) - I_{\mathcal{B}}(\mathbf{x})|^2 \ \mathrm{d}V \ . \tag{6.17}$$

Algorithm 2 Algorithm for the steady-state-heat-flow-based smooth, non-folding warping method.

Require: partial correspondence between atlas image $\mathcal{A}$ and patient (target) image $\mathcal{B}$
Require: region of interest (ROI) in target image over which the warp is requested
return a warp field $\mathbf{h}$ mapping points in the target image to points in the atlas image.
for all point $\mathbf{p} \in$ partial correspondence list **do**
Compute local temperature gradient at $\mathbf{p}$ using equation 6.5 in the target image.
end for
for all grid points $\mathbf{x} \in \mathcal{B}$ **do**
Compute temperature distribution at $\mathbf{x}$ using equation 6.1.
end for
Fix boundaries of ROI in image $\mathcal{B}$ to the temperature distribution in image $\mathcal{A}$.
$\forall\, i \quad c_i \leftarrow 1$.
$\mathbf{I} \leftarrow$ down-sampled version of initial temperature distribution.
repeat
Compute warp $\mathbf{h}$ using image $\mathbf{I}$ as the initial distribution.
Compute foldedness $N(\mathbf{h})$ of the warp $\mathbf{h}$.
if $N(\mathbf{h}) > 0$ **then**
Find locations of folding.
Order coefficients c_i by proximity to locations of folding.
while $N(\mathbf{h}) > 0$ **do**
Find next (first) available coefficient c_i in the ordering. If all coefficients have been checked, move back to the first coefficient, thus, starting the next round of optimization. If no coefficient is available to be optimized in current round, break out of this loop.
if $N(\mathbf{h}; c_i \times \alpha) < N(\mathbf{h}; c_i)$ **then**
$c_i \leftarrow c_i \times \alpha$.
Move to start of loop.
end if
if $N(\mathbf{h}; c_i/\alpha) < N(\mathbf{h}; c_i)$ **then**
$c_i \leftarrow c_i/\alpha$.
Move to start of loop.
end if
Mark coefficient c_i as unusable for next r rounds.
end while
end if
Interpolate warp $\mathbf{h}$ to next higher resolution to produce image $\mathbf{I}$.
Interpolate coefficient field c_i to next higher resolution.
until highest desirable resolution has been reached.
return warp $\mathbf{h}$

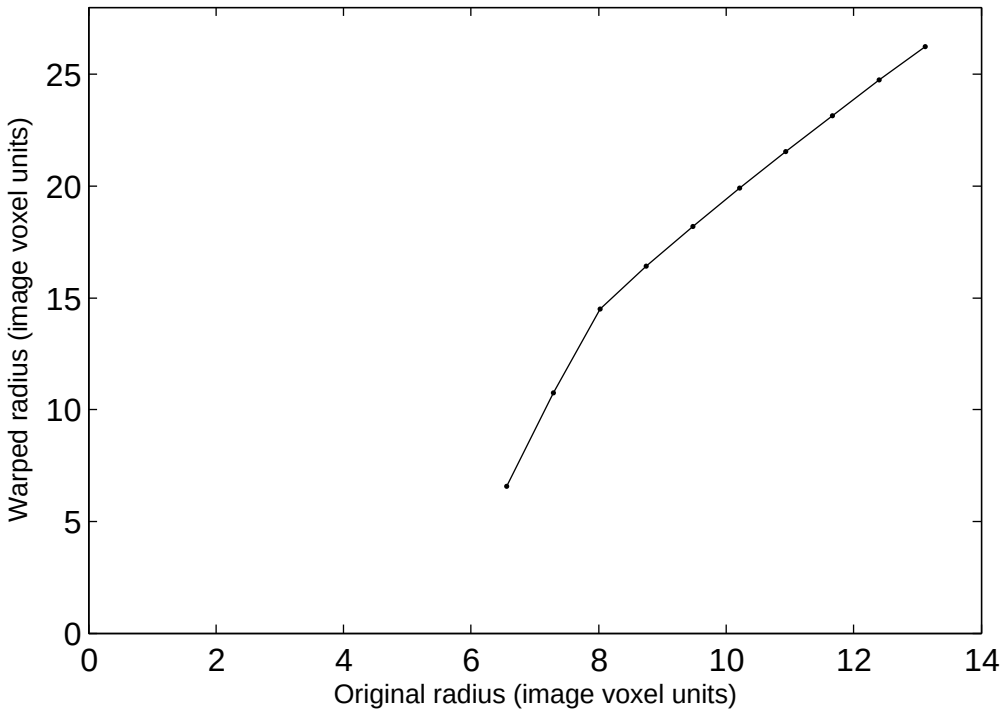


Figure 6.5: The radii of the warped concentric spheres as a function of their original radii in a warp field between two spheres as the outer sphere was scaled by a factor $\alpha = 2.0$. The data is presented in units of image voxels. The relationship is almost linear.

One could conceivably include other energy terms such as a modified version of the foldedness energy function that penalizes significant volume changes. Haber and Modersitziki [15] present one such example.

The second improvement arises from a better representation of the conduction coefficients and other optimization algorithms for changing the c_i field. The field can be represented parametrically, for example using multiscale wavelets, which has the potential for making the optimization significantly faster and producing warps that may be smoother.

Finally, this method can be improved by reimplementing it for a GPU. The current implementation is written in MATLAB and can take anywhere between a few minutes to a few hours. It uses a Laplacian solving method – red-black iterations – that can be parallelized, but the implementation does not exploit this. Reimplementing this method for a GPU will help in

exploiting the underlying parallelism in the solver and greatly increase its speed.

Chapter 7

Results

In this chapter, I apply the hierarchical segmentation method developed in the previous chapters to the problem of head and neck segmentation. First, in section 7.1, I present a brief overview of the method. Next, in section 7.2, I point out some of the objects in the head and neck region that are segmented and landmark points on the same. Finally, in section 7.3, I show the result of applying the correspondence-preserving warping method along with the hierarchical segmentation method on 3D CT images of six patients.

7.1 Overview

This section presents a review of the hierarchical segmentation method shown in fig. 1.1.

The overall framework, which segments a complex of objects, starts with a fully segmented atlas. The atlas is augmented with the shape models of all the objects that we wish to segment. The labeled mesh generator, introduced in section 1.3 and developed in chapter 4 for nearly tubular objects, partially segments the target to produce surface meshes with labeled vertices. These objects, which are also easily segmented in the target because of high contrast, are called *reference objects*, and the remaining not so easily-segmented objects are called *target objects*.

The atlas image is mapped to the target image in two steps. The surface correspondence step takes surface meshes with labeled vertices and infers correspondence between reposi-

tioned vertices on the reference objects in the atlas and the target. Next, these correspondences are used to drive a warp from the atlas to the target image in my correspondence-preserving warping method; the warp from the correspondence-preserving warping method is based purely on geometric information learned from meshes of the reference objects. These steps were introduced in sections 1.3 and 1.4, and developed in chapters 5 and 6 respectively.

In the hierarchical warp method, the resulting warp yields an initial guess for the segmentation of the target objects. The objects are then segmented, and good results are selected. The good results are added to the list of the reference objects and removed from the list of the target objects. The entire process is then repeated.

In this thesis, after initialization, I do not perform a real segmentation, which would only test the efficacy of the segmentation method that is used and is not the focus of the thesis; instead, I replace one of the warped target objects by its true shape model and use it to complete the loop.

The objects that form the reference objects and the target objects list are presented in the next section.

7.2 Head and Neck Anatomy

The head and neck region has a complex anatomy. There are about forty objects of interest in the head and neck region. This section points out the objects that are either used as reference objects or target objects. It also points out special points on these objects that are easily identifiable across the population and are used as landmarks for training shape models of the objects.

The easily segmentable objects form the reference objects. These objects typically are either bones or air cavities, which have high contrast in CT images. Some of the easily segmentable objects (fig. 7.1) include the skin surface, mandible (jawbone), trachea, and clavicles. These serve as reference objects. Certain other objects in the head and neck such as

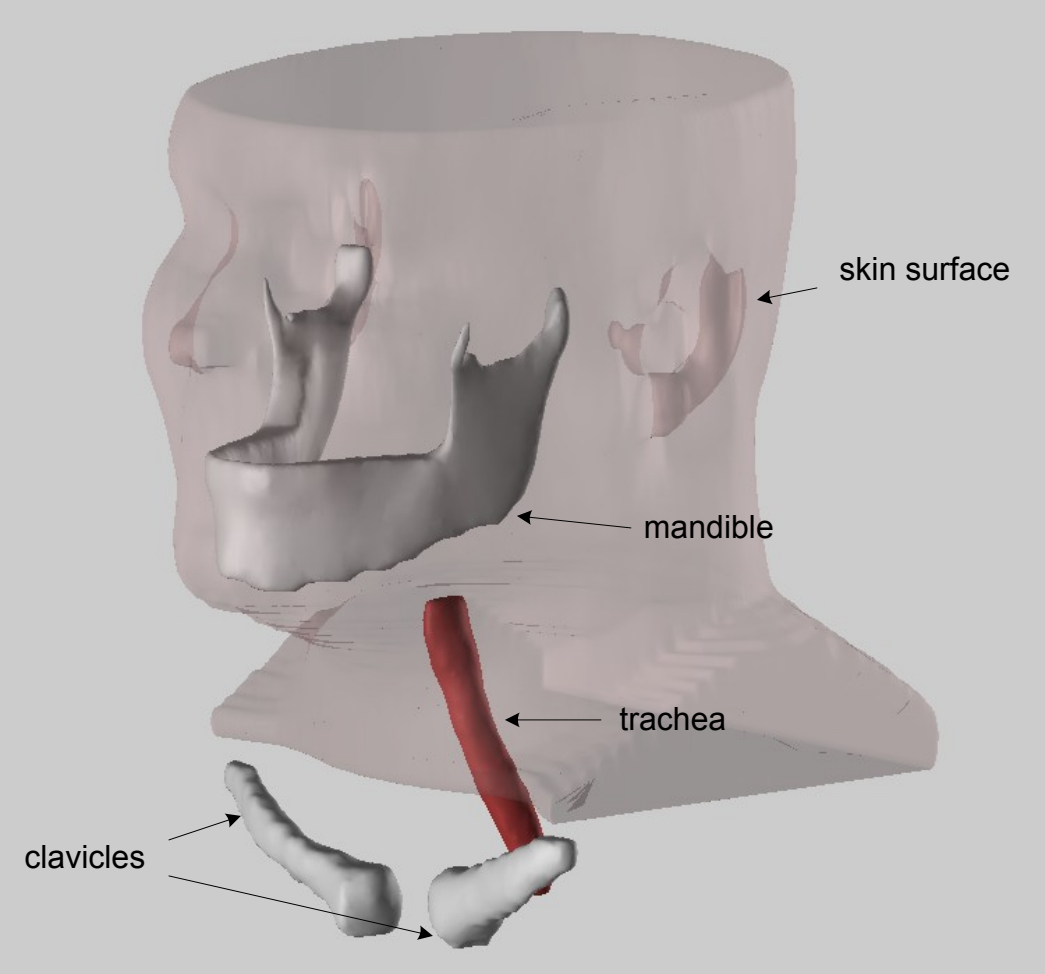


Figure 7.1: Reference objects

the vertebrae are also easily segmentable, but developing shape models for them is tricky; therefore, they are not considered as reference objects.

The skin surface as used here extends from the bridge of the nose to the top of the neck. It has four landmarks that are quite consistent across the population – bridge of the nose, tip of the nose, tip of the chin, and 'top of the neck'. The top of the neck is the point where the neck meets the face in the central sagittal plane. The top of the nose and top of the neck are hyperbolic points (one of the principal curvatures is negative and the other is positive). The tip of the nose and the tip of the chin are convex points (both principal curvatures are positive). These landmarks are shown for a sagittal slice of a CT image of a patient in fig. 7.2.

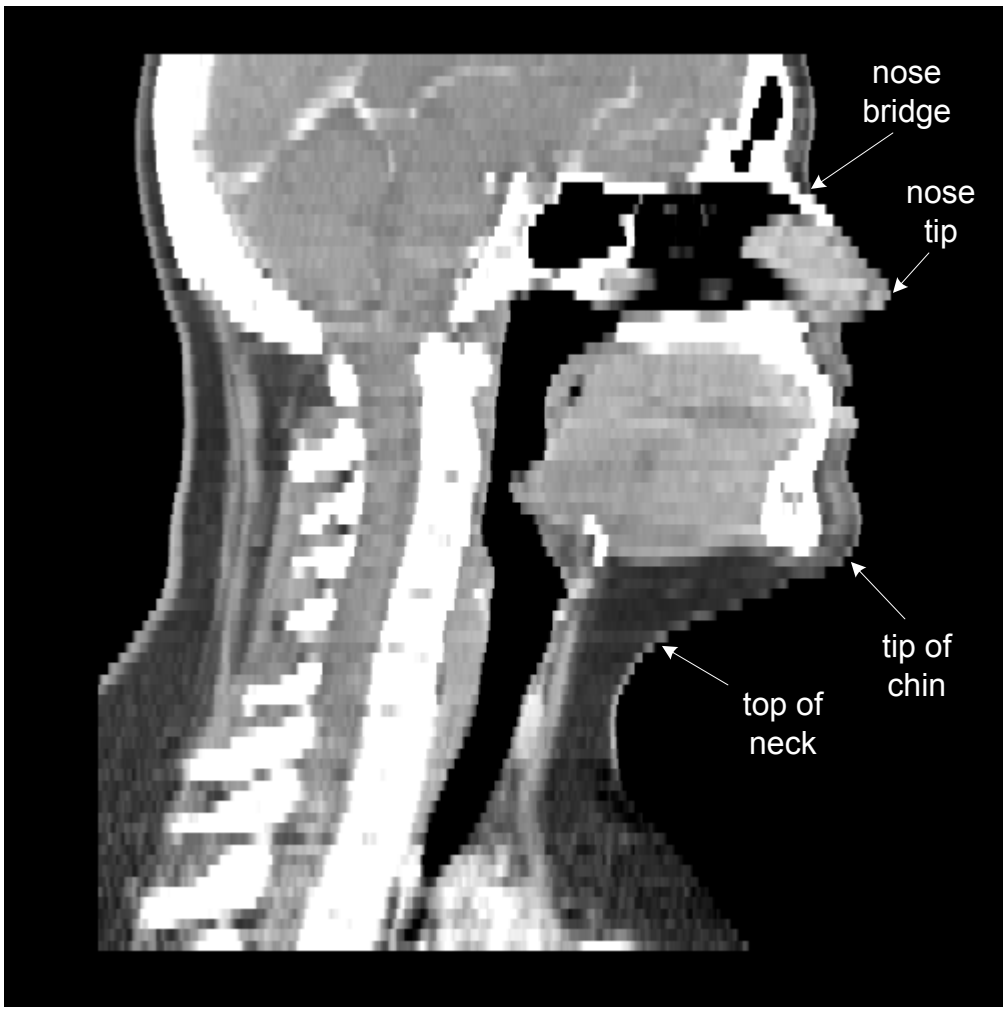


Figure 7.2: Landmarks on the skin surface shown on a sagittal slice of a CT image of a patient.

The trachea is the part of the wind pipe that extends below the larynx. Only the region between the vertebrae C1 and T2 is modeled. These vertebrae also serve as landmarks for the shape model. For an infant, the larynx is around the second or third vertabra, and it descends as a s/he grows. This is important to remember if objects from a child's image need to be modeled. The trachea lies in the neck region and is therefore useful for providing correspondence for organs in the neck region where the skin surface is not modeled.

The target objects over which measurements are taken include the sternocleidomastoid muscle, thyroid, internal jugular vein, internal carotid artery, and the masseter muscle. These objects are shown in fig. 7.3. Some of these objects are present in both sides of the head and

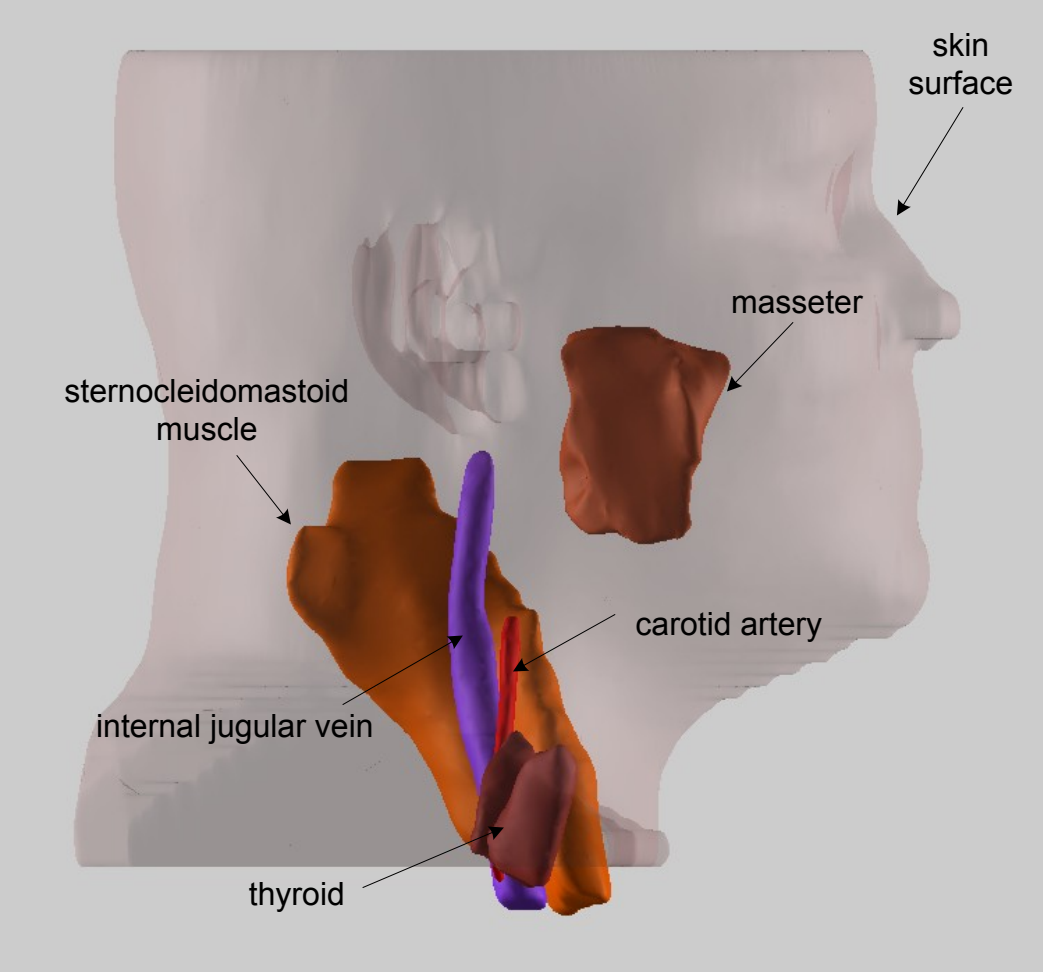


Figure 7.3: Target objects

neck, but I measure results only on the objects present on the right side because of the lack of a manual segmentation to compare against for the ones on the left side in the data set.

For the next round of improving the initializations, I use the sternocleidomastoid muscle on the right side of the head and neck. This muscle connects the sternum, clavicle, and the mastoid. It stretches across a large region of the head and neck. Also, it has reasonable contrast at its boundary. Thus, it may be relatively easy to segment and also be a useful structure to include in the reference objects list for the next round of initialization.

Object	Volume overlap (%)		Root mean square distance (mm)	
	mean	std. dev.	mean	std. dev.
skin surface	96.91	2.82	0.33	0.036
mandible	92.70	1.39	0.48	0.040
trachea	95.58	1.05	0.34	0.077
left clavicle	93.43	1.27	0.46	0.050
right clavicle	93.11	1.48	0.47	0.052

Table 7.1: Match measures between atlas reference objects and warped reference objects from the patient image.

7.3 Application and Results

For this experiment, I used CT images of the head and neck region from six patients. These images have an inter-voxel distance in the axial plane between 0.3 and 0.5 mm, and a separation of approximately 3 mm between two axial slices. All 15 possible pairs between these patients were formed. One of the patients served as an atlas, and the other served as the target patient for each of these pairs. The experiment was repeated with the roles of the patients reversed in each pair. First, match measures of the reference objects are presented. Next, match measures of the target objects are presented.

Fig. 7.4, 7.5, 7.6, 7.7, and 7.8 show the skin, mandible, trachea, left clavicle, and right clavicle respectively of the atlas (in blue) and a warped patient image (in red). Mostly, the two surfaces are very close to each other. However, the two surfaces do not match well in certain locations. The nose of the atlas patient has been arbitrarily cut off, and the match near the nose is poor. The ears are very thin and variable across the patients, and the match near the ears is also poor. The condyles of the mandible do not match well because of the difficulty in representing these thin structures.

Next, on these reference objects, I computed two different match measures – root mean square surface distance and volume overlap. The mean square surface distance measures the mean of the square of the distances between the points on one surface $\mathcal{A}$ and the closest point on the other surface $\mathcal{B}$. Because surface distance measures are asymmetric, the final measure

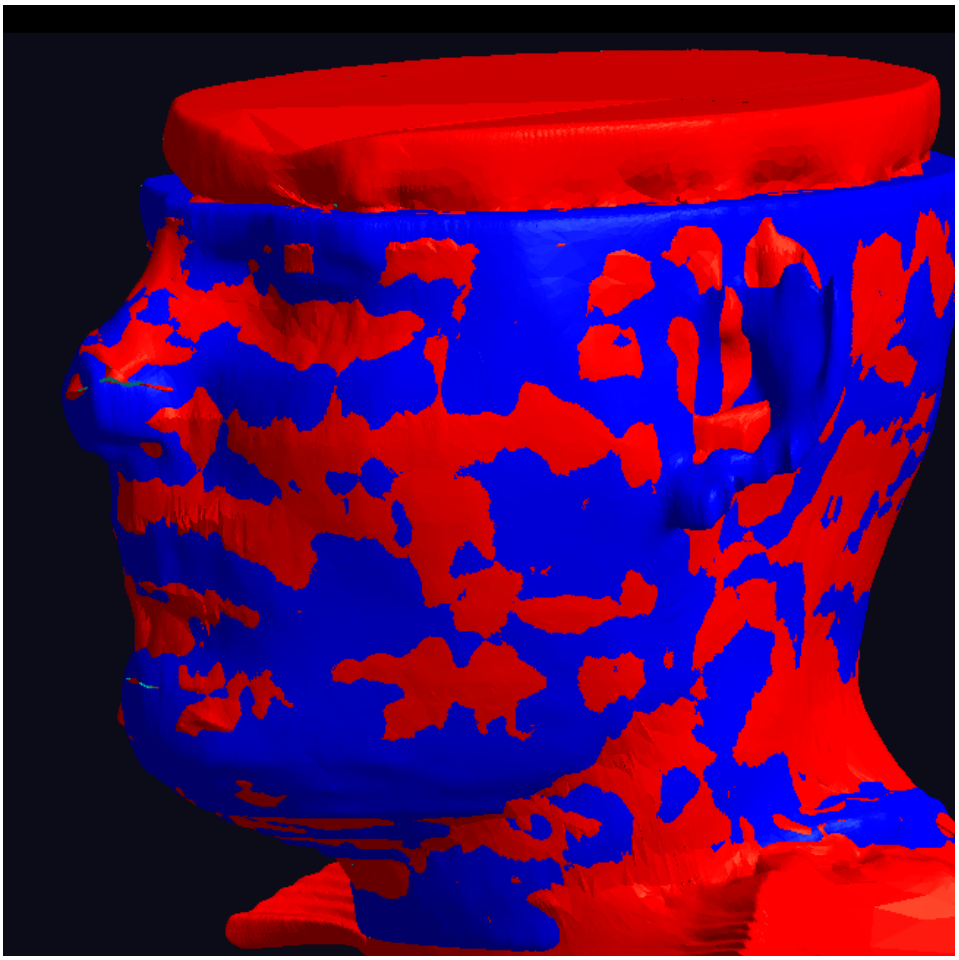


Figure 7.4: Atlas' skin surface in blue with the warped patient's skin surface overlaid in red. Most of the two surfaces match well. However, the nose of the atlas patient has been arbitrarily cut off, and the match near the nose is poor. Also, the ears are very thin and variable across patients, and the match near the ears is also poor.

reports the root of the mean of the distance from surface $\mathcal{A}$ to surface $\mathcal{B}$ and the distance from surface $\mathcal{B}$ to surface $\mathcal{A}$. Surface distance measures can be accurately computed only when the two objects overlap. If a section of one of the objects lies completely outside the other object, the surface distance measure reported will be inaccurate. Therefore, it is important to measure the overlap between the two objects, as the greater the overlap, the more accurate the surface distance reported between them will be. The volume overlap measures the overlap between the two objects as the ratio of the intersection of the segmented volume with the truth volume to the average of the two volumes. The volume overlap measure, by itself, is not a very good

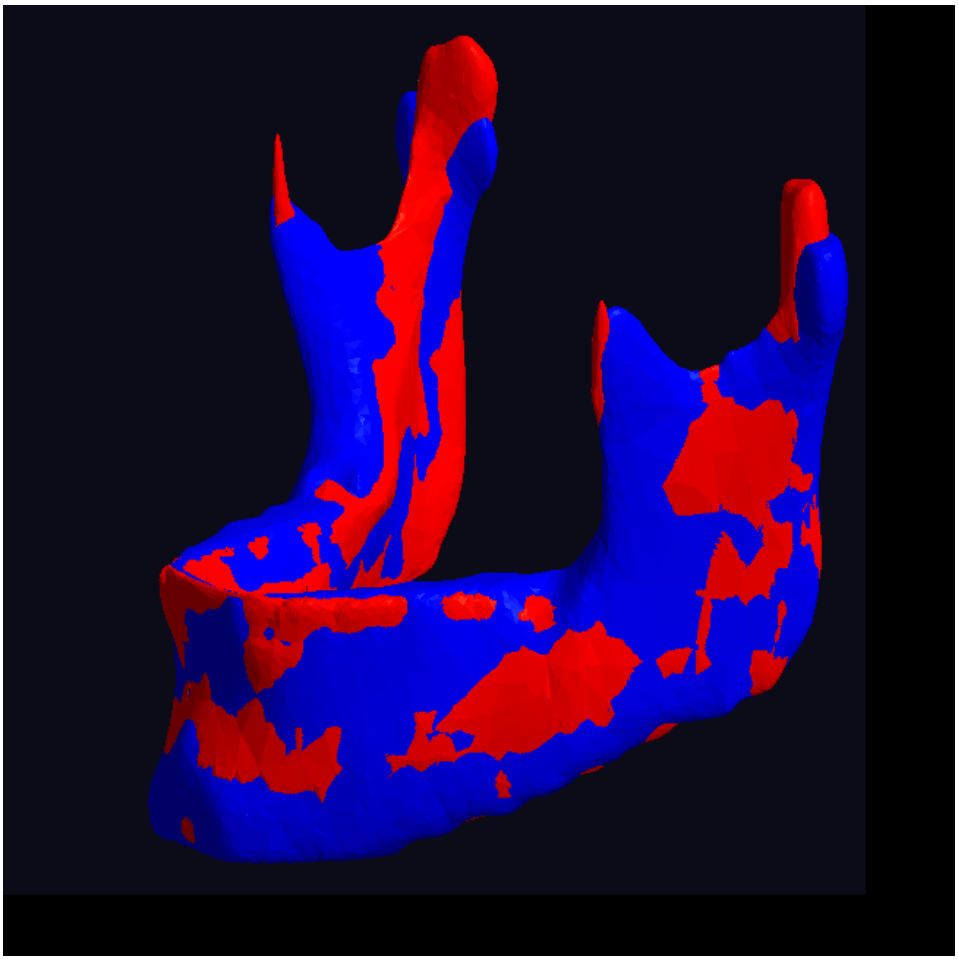


Figure 7.5: Atlas' mandible in blue with the warped patient's mandible overlaid in red. Most of the warped mandible matches with the atlas mandible. However, the condyles do not match well because of the difficulty in representing these thin structures.

measure of match, because for thin, elongated objects, the measure may be quite low though visually the overlap may look quite good; however, for large blob-shaped objects with the same measure value, visually the overlap may look quite bad.

Table 7.1 shows these match measures for the reference objects. Ideally, they should have a 100% volume overlap and a 0 root mean square surface distance between them, but imprecision introduced by the (finite) resolution at which the different components of the method operate introduce small errors.

Fig. 7.9, 7.10, 7.11, 7.12, and 7.13 show the internal jugular vein, masseter muscle, thy-

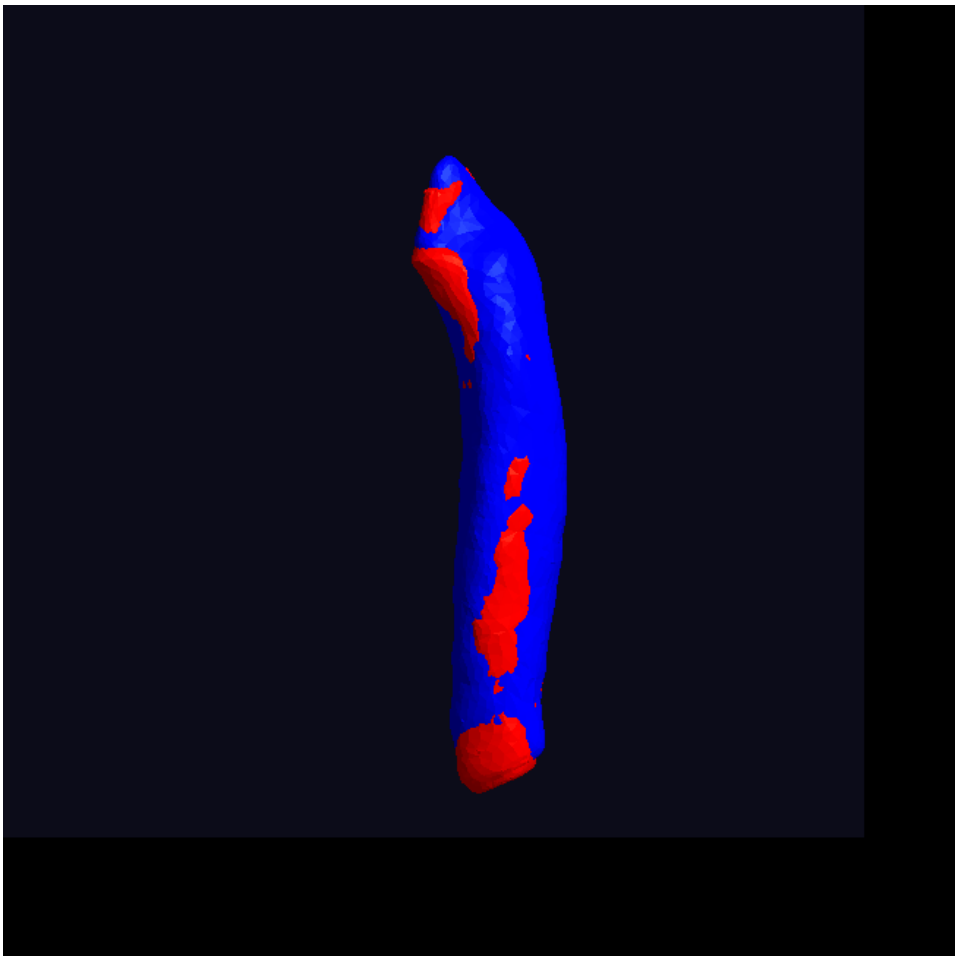


Figure 7.6: Atlas' trachea in blue with the warped patient's trachea overlaid in red. The two surfaces match well almost everywhere.

roid, left sternocleidomastoid muscle, and right sternocleidomastoid muscle respectively of the atlas image (in blue) and the warped patient image (in red). The surfaces shown in cyan are the results of applying the warp by including the right sternocleidomastoid muscle in the list of reference objects.

Table 7.2 lists the same numerically using two different measurements – volume overlap and average surface distance; table 7.3 lists the same results after the right sternocleidomastoid muscle was used as a reference object. Objects that were not close to any reference object but were close to the right sternocleidomastoid muscle – right internal jugular vein and right internal carotid artery – show an improved match with the corresponding atlas object.

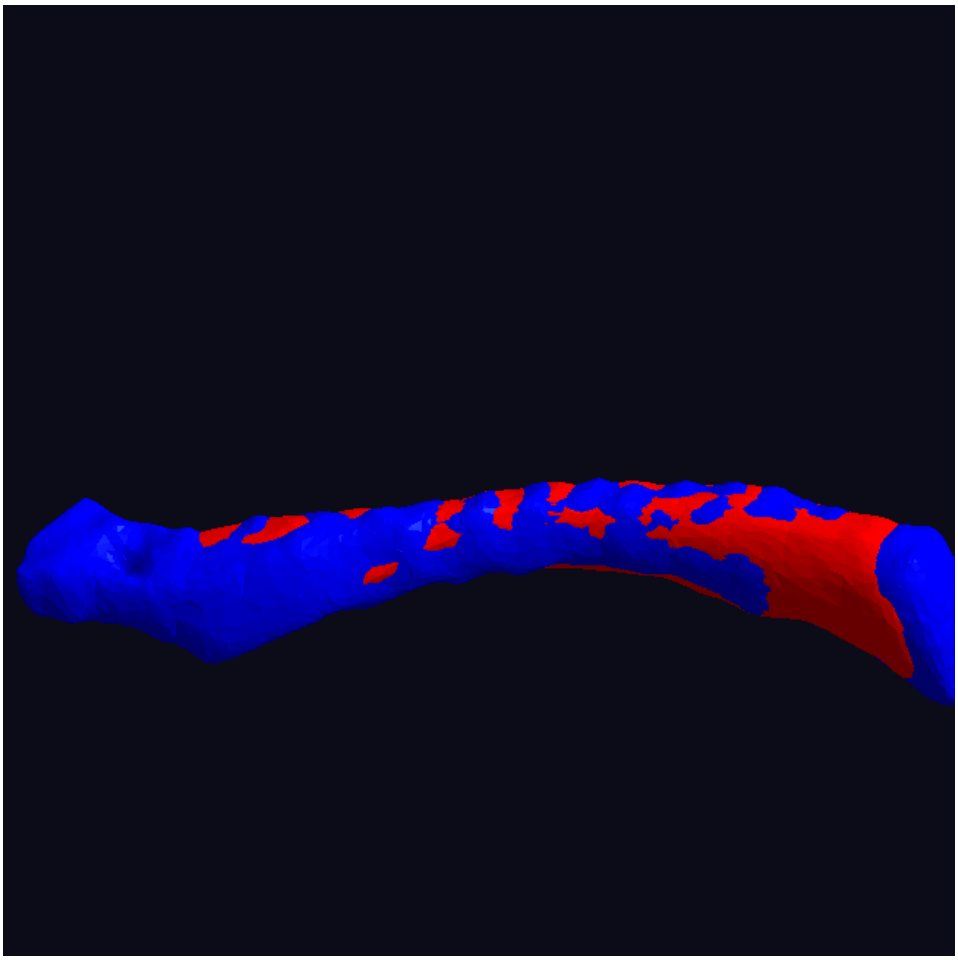


Figure 7.7: Atlas' left clavicle in blue with the warped patient's left clavicle overlaid in red. The two surfaces match well almost everywhere.

The location of veins and arteries is variable across patients. The warped internal jugular vein and carotid artery do not match well with the ones from the atlas. They are both close to the sternocleidomastoid muscle. After using the sternocleidomastoid muscle as a reference object, the match improves considerably validating the hierarchical segmentation method. Other organs such as the warped masseter muscle and the thyroid, which lie adjacent to the reference objects mandible and trachea respectively, match well with the ones from the atlas, and the results do not differ much if the sternocleidomastoid muscle is used or not.

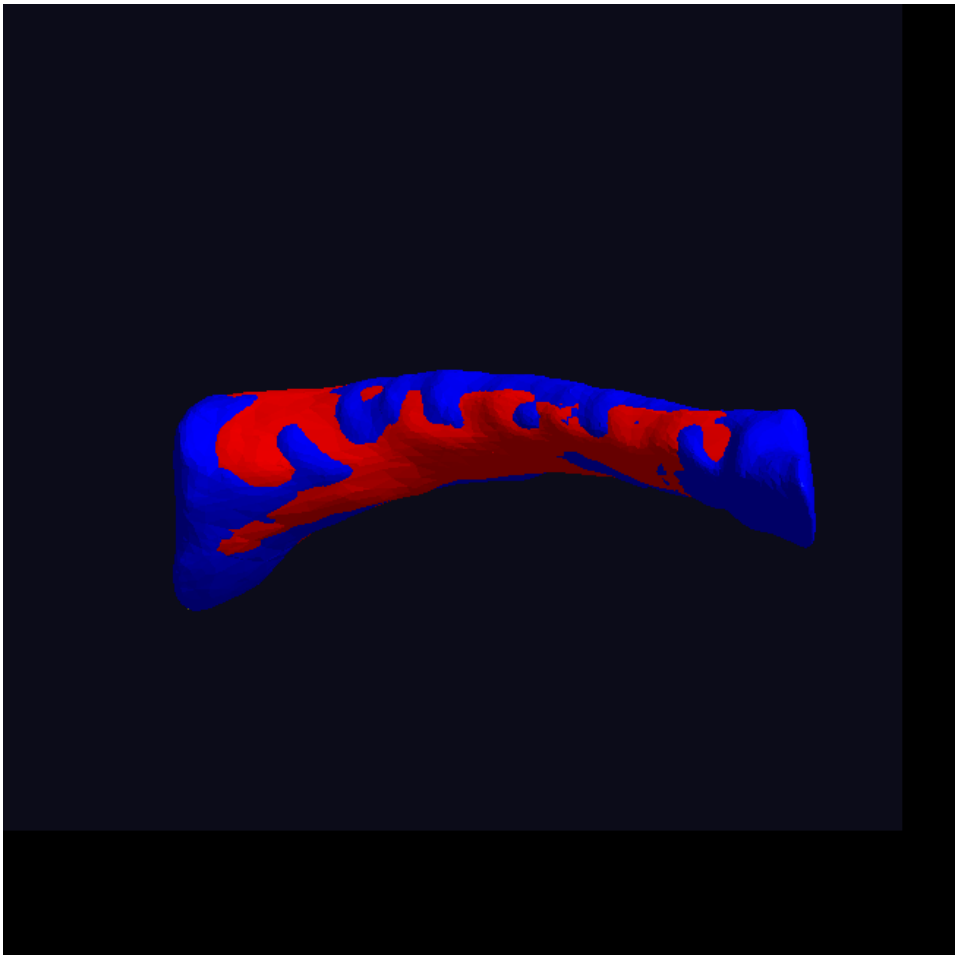


Figure 7.8: Atlas' right clavicle in blue with the warped patient's right clavicle overlaid in red. The two surfaces match well almost everywhere.

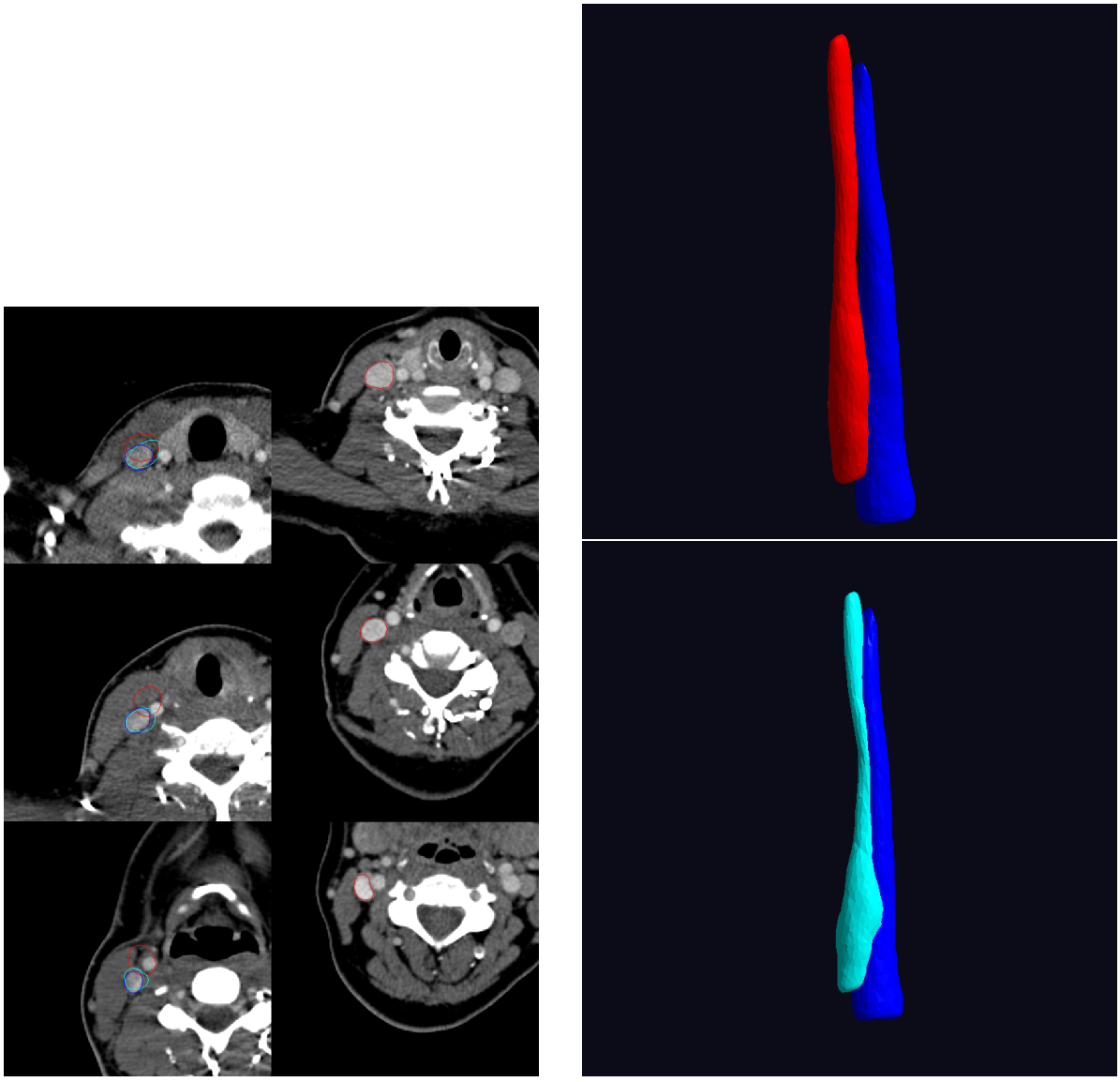


Figure 7.9: Internal jugular veins in the atlas (blue), warped from patient image (red), and warped from patient image after using the sternocleidomastoid muscle as a reference object (cyan). The left image shows the results on different axial slices. The right images show the same in 3D. There is a noticeable improvement in the match between the warped internal jugular vein and the atlas one after the sternocleidomastoid muscle was used as a reference object.

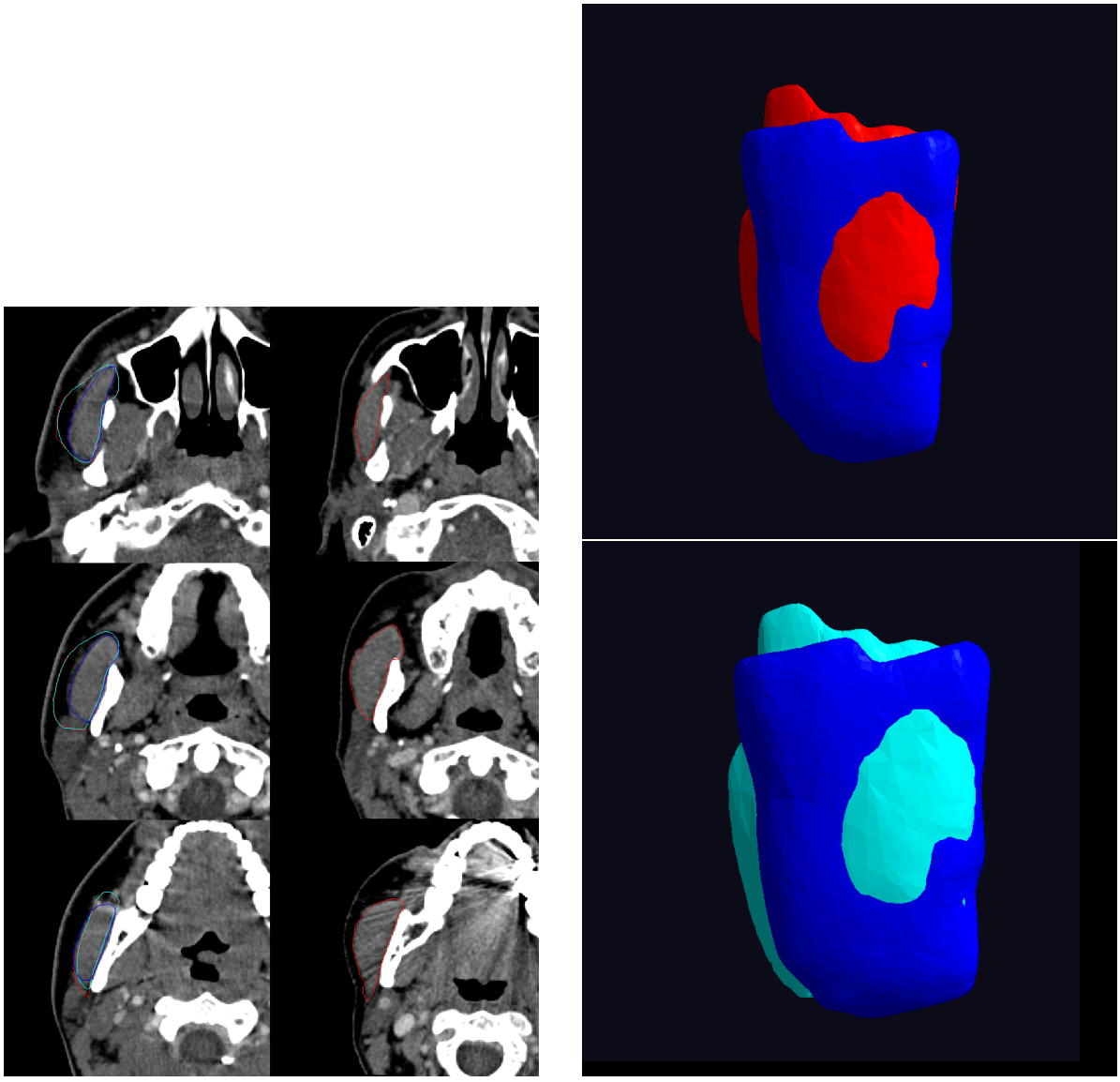


Figure 7.10: Masseter muscle in the atlas (blue), warped from patient image (red), and warped from patient image after using the sternocleidomastoid muscle as a reference object (cyan). The left image shows the results on different axial slices. The right images show the same in 3D. The masseter muscle is adjacent to the mandible, which is a reference object. The warped masseter appears to match well with the atlas masseter. Further, there is not much improvement after the sternocleidomastoid muscle was used as a reference object.

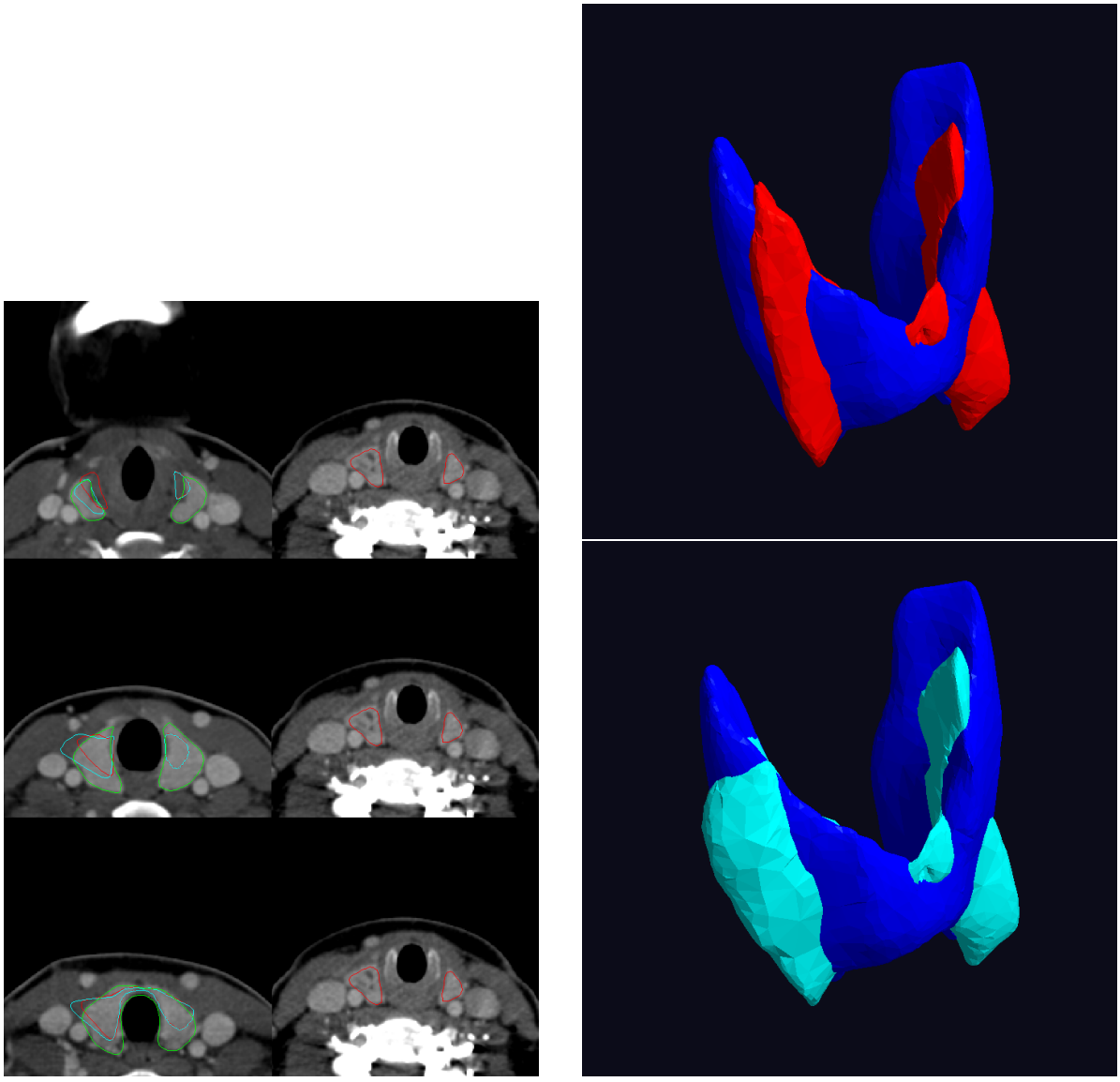


Figure 7.11: Thyroid in the atlas (blue), warped from patient image (red), and warped from patient image after using the sternocleidomastoid muscle as a reference object (cyan). The left image shows the results on different axial slices. The right images show the same in 3D. The thyroid lies next to the trachea. The warped thyroid appears to match well with the atlas thyroid. There is not much improvement after the sternocliedomastoid muscle was used as a reference object.

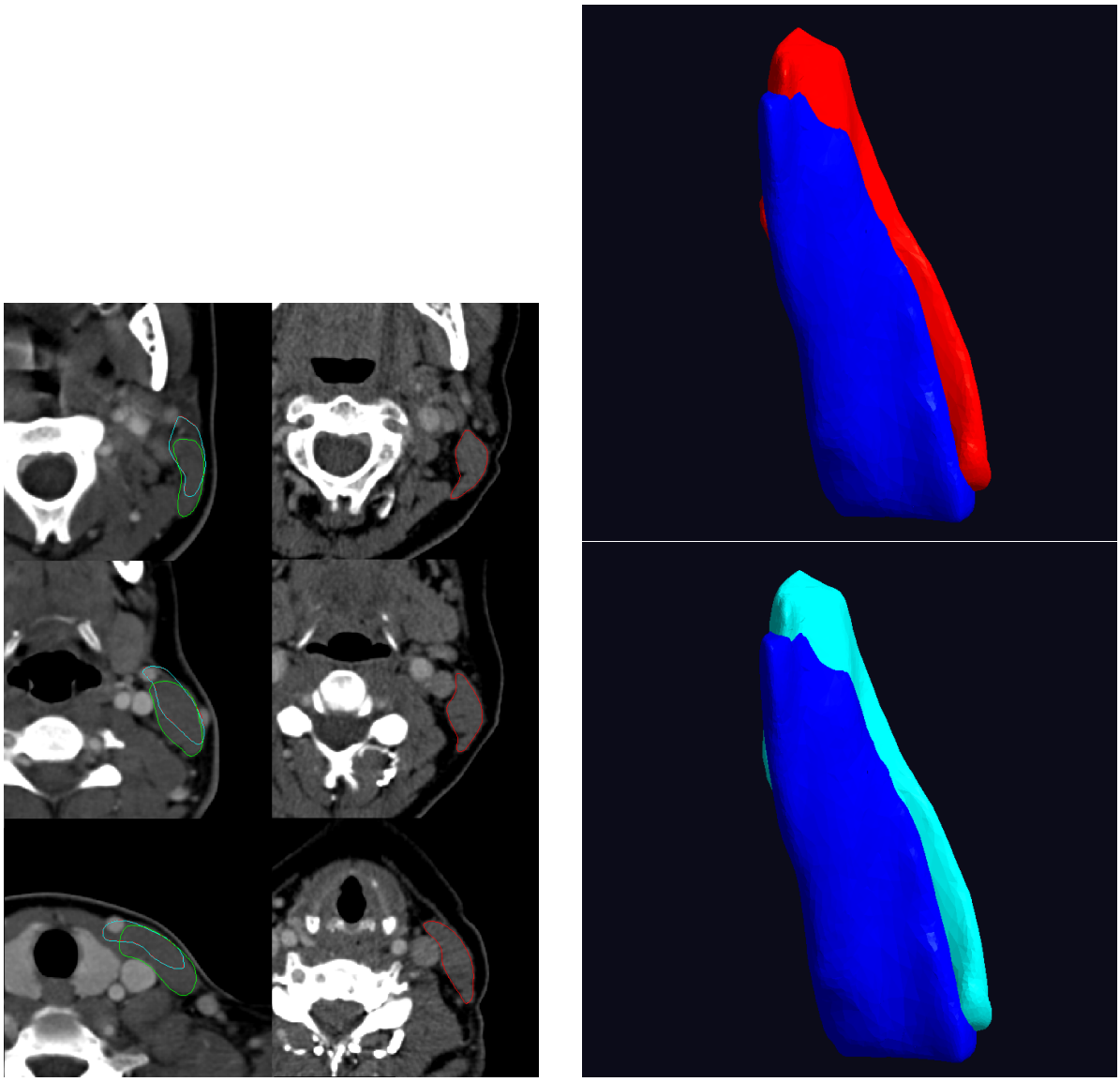


Figure 7.12: Left sternocleidomastoid muscle in the atlas (blue), warped from patient image (red), and warped from patient image after using the sternocleidomastoid muscle as a reference object (cyan). The left image shows the results on different axial slices. The right images show the same in 3D. The warped left sternocleidomastoid muscle matches well with the one from the atlas. The left sternocleidomastoid muscle lies far from the right sternocleidomastoid muscle. Therefore, there is not much difference in the results after using the right sternocleidomastoid muscle as a reference object.

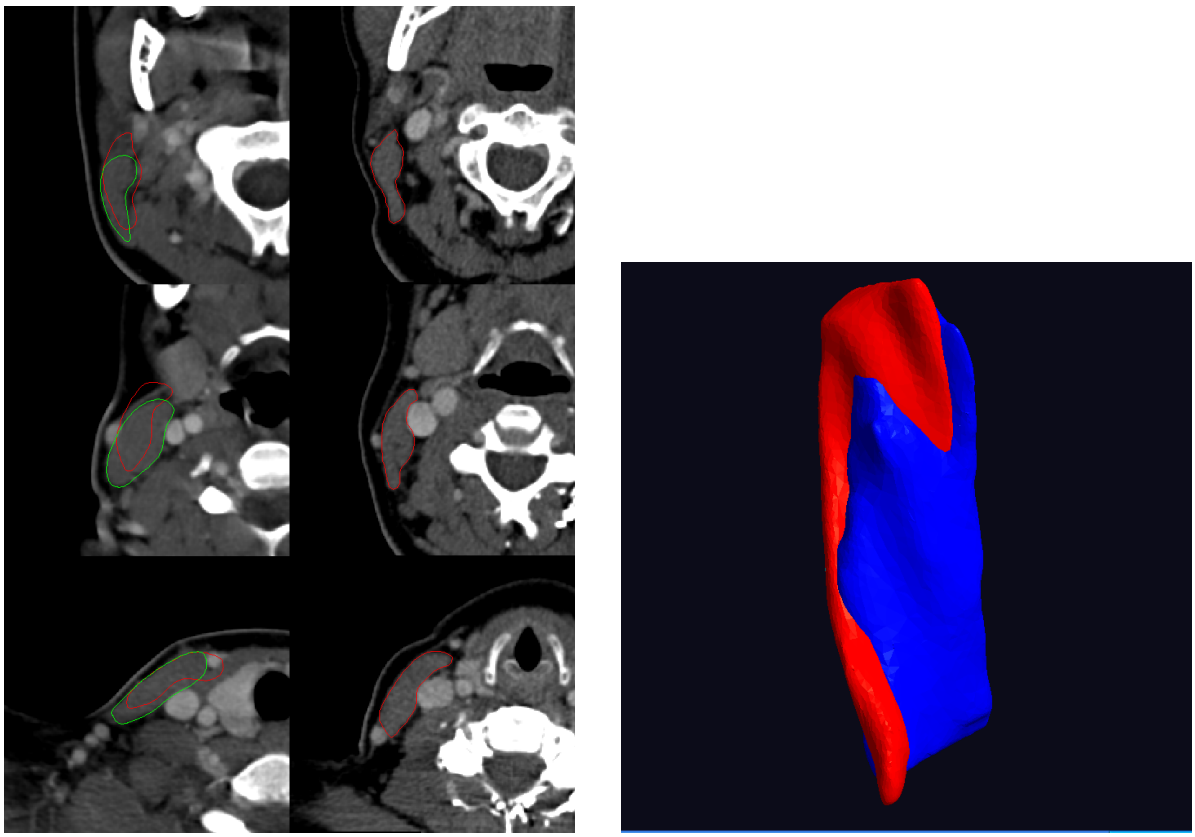


Figure 7.13: Right sternocleidomastoid muscle in the atlas (blue) and warped from patient image (red). The left image shows the results on different axial slices. The right image shows the same in 3D. The two surfaces match well almost everywhere.

Object	Volume overlap (%)		Root mean square distance (mm)	
	mean	std. dev.	mean	std. dev.
sternocleidomastoid muscle	28.50	20.72	0.28	0.061
thyroid	43.43	9.56	0.38	0.037
internal jugular vein	15.10	11.59	-	-
internal carotid artery	0.86	2.22	-	-
masseter muscle	63.57	10.01	0.38	0.049

Table 7.2: Match measures between atlas target objects and warped target objects from the patient image. Because of inadequate correspondence of nearest neighbor points in the available method for computing root mean square distances computed by this program for poorly overlapping objects such as the internal jugular veins and the carotid arteries produce uninformative results and are therefore not reported.

Object	Volume overlap (%)		Root mean square distance (mm)	
	mean	std. dev.	mean	std. dev.
thyroid	47.53	7.92	0.39	0.035
internal jugular vein	48.28	11.66	0.43	0.062
internal carotid artery	3.6	7.18	0.29	0.083
masseter muscle	63.42	10.02	0.38	0.049

Table 7.3: Match measures between atlas target objects and warped target objects from the patient image after the sternocleidomastoid muscle was added to the list of reference objects.

Chapter 8

Interpolating Methods for Correspondence-preserving Warps

A warp can be generated in two ways – one that involves both temporal interpolation and spatial computations, and another that only involves spatial computations. The former can provide interpolations between the two images, while the latter cannot. The main correspondence-preserving warping method presented in this thesis produces an almost diffeomorphic warp, one that is smooth and is not folded, by solving an elliptic PDE in the space of the target patient image. In this chapter, I briefly present two other methods, developed with the goal of producing correspondence-preserving diffeomorphisms (one-to-one, smooth, and smoothly invertible mappings) that take the approach of interpolation. Though the warps were intended to be diffeomorphic, the first of the two methods falters on providing strong guarantees for the same.

In sections 3.2.1 and 3.2.2, I discussed a framework that, given a set of non-intersecting (at the same interpolation time) temporally interpolating paths between a source and a target image, could extrapolate this interpolation to the entire image. If the given paths do not intersect each other, then the entire warp is guaranteed to be diffeomorphic.

The first method, detailed in section 8.1, attempts to generate these non-crossing paths by interpolating between shape models contained in their shape space; in particular it uses m-reps. The second method, detailed in section 8.2, formulates a different form of the steady-

state heat flow problem that only delivers the diffeomorphism on the surface of the object and operates in four dimensions (not the three dimensions of the method in chapter 6) to generate the interpolations between the two images.

8.1 Correspondence-preserving Warps by Geodesic Interpolation between M-reps

This section details a method that uses paths in shape space between two m-reps to produce interpolations between them and to use these interpolations with the large deformation framework (LDDM, see section 3.2.2) to produce a warp between the two images in which the objects modeled by the two m-reps lie.

Equipping the shape space with a shape metric enables us to determine the length of the path in shape space between two m-reps. A geodesic path, introduced for quasi-tubular medial models in section 4.3, is the shortest path in the medial shape space between two points representing an m-rep each. Interpolating along geodesics between two m-reps yields the geodesic interpolation between them.

This geodesic interpolation between m-reps yields temporal paths between the two m-reps for every point on their implied surface. Thus for every surface point $\mathbf{x}_i$, we have a curve $\mathbf{h}(\mathbf{x}_i, t)$. These curves when used with the LDDM framework yield a similar curve $\mathbf{h}(\mathbf{x}_i, t)$ for every point $\mathbf{x}_i$ in three-dimensional (3D) space through interpolating the given paths by thin plate splines. This new set of curves represents a warp for the entire image.

In the rest of this section, I will detail the method of geodesic interpolation, review the LDDM framework, present a set of results obtained by the application of this method in conjunction with LDDM, and compare this correspondence-preserving warping method with the main corresponding-preserving warping method presented in the thesis.

8.1.1 Geodesic Interpolation between M-reps

M-reps lie in a *Riemannian symmetric shape* space, where every point in this space is an m-rep. Thus, stepping along any smooth curve in this space yields a continuously varying m-rep model. Each point in this space is also associated with a tangent space $T_x(M)$ and a *Riemannian metric*, a smoothly varying inner product on this tangent space. It suffices to say here that a *geodesic curve* is a special curve, which minimizes the *distance* between the two m-reps, which forms its end-points. This distance is also known as the *geodesic distance* or the *Riemannian distance*.

For any tangent vector $v \in T_xM$ there exists a unique geodesic $\gamma_v(t)$, v being its initial velocity. The *Riemannian exponential map* maps v to the point at $t = 1$ along the geodesic γ_v. The geodesic has a varying direction but a constant speed given by $|| \left[\frac{d\gamma_v}{dt}\right]_t || = ||v||$. Therefore, the exponential map, $\mathrm{Exp}_x(v)$, preserves distances from the initial point, i.e., $d(x, \mathrm{Exp}_x(v)) = ||v||$. In the neighborhood of zero, the exponential map is a diffeomorphism. The inverse in this neighborhood is known as the *Riemannian log map*, which is denoted as Log_x. Thus, for a point y in the neighborhood of x the geodesic distance between them is given by $d(x, y) = ||\mathrm{Log}_x(y)||$.

We can march along the geodesic between two m-reps by marching along the geodesics between the corresponding pairs of medial atoms $\mathbf{m}_1$ and $\mathbf{m}_2$. A medial atom $\mathbf{m}(t)$ that is fractional distance t along the geodesic between $\mathbf{m}_1$ and $\mathbf{m}_2$ is given by

$$\mathbf{m}(t) = \underset{\mathbf{m}_1}{\mathrm{Exp}}\left(t \underset{\mathbf{m}_1}{\mathrm{Log}}(\mathbf{m}_2)\right). \tag{8.1}$$

Interpolated m-reps obtained from a pair of m-reps, one from the atlas image and the other from the patient image, yield paths for points on their surfaces. Thus, we can get a set of paths that can serve as the paths in the LDDM framework (a review of which is presented in the next section) to extrapolate this interpolation to the entire space. These paths are not guaranteed to never cross each other, but the likelihood of a crossing is extremely low for

ordinary anatomical objects. In the tests that I had performed, I had not encountered a situation where these paths crossed each other.

8.1.2 Large Deformation Diffeomorphism Framework

In this framework a time-indexed transformation $h(\mathbf{x}, t)$ mapping the atlas to the target is defined via an integration of a velocity vector field, given by

$$\mathbf{h}(\mathbf{x}, t) = \mathbf{x} + \int_0^t \mathbf{v}(\mathbf{h}(\mathbf{x}, t), t)\, \mathrm{d}t \ . \tag{8.2}$$

Given a set of landmark points $\mathbf{x}_i$ with respective non-crossing paths $\mathbf{h}(\mathbf{x}_i, t)$ describing the motion of the landmark points from the atlas to the target, the diffeomorphic transformation $\mathbf{h}$ is obtained by finding a velocity vector field that minimizes an energy function that is related to the smoothness (as defined by a linear differential operator L) of the velocity vector field , while maintaining the given landmark paths. Mathematically, it can be written as

$$\begin{aligned} \hat{\mathbf{v}}(\mathbf{x}, t) &= \arg\min_{\mathbf{v}} \int_0^1 ||L\mathbf{v}(\mathbf{x}, t)||^2 \, \mathrm{d}t \\ subject\ to: \ \mathbf{v}(\mathbf{h}(\mathbf{x}_i, t)) &= \frac{\mathrm{d}\mathbf{h}(\mathbf{x}_i, t)}{\mathrm{d}t}. \end{aligned} \tag{8.3}$$

Many linear differential operators have been used in the literature; the biharmonic thin plate spline (TPS) operator is the most common because it has a closed-form solution. In addition, plugging in the biharmonic thin plate spline operator generates quadratically smooth interpolations.

As discussed in the previous section, these landmark points are obtained by geodesic interpolation of m-reps. The next section presents the result of using this method on the head and neck segmentation problem presented in this thesis.

8.1.3 Application and Results

3D CT head and neck images of five different patients were expertly outlined. For each of these patients, a few of the organs were segmented. Eleven random pairs of patients were chosen from these five patients. In each pair, one of the patients served as the target patient and the other as the source or the atlas. The easily identifiable objects – skin surface, mandible, clavicles, and the trachea served as the reference objects. A warp was computed from the atlas to the target based on these objects. The computed warp was then used to transfer all the other objects to the target patient.

In table 8.1, I have listed the mean and the standard deviation of the volume overlaps obtained for some of the reference objects. The volume overlap measures the ratio of the intersection of the segmented volume with the truth volume and their average. For thin, elongated objects, the measure may be quite low though visually the overlap may look quite good. However, the same number for a large blob-shaped object may turn out to be visually quite bad. Despite this limitation, volume overlap measures were the only ones recorded when this experiment was run. A higher standard deviation in the volume overlap indicates that the result may not be quite accurate or representative of the population. The volume overlap measure can be used as a cutoff to predict how well the object may be segmented if this result were used as an initialization. Thus, a number below a certain threshold may indicate a bad result and one above will indicate a good result. This number can only be determined empirically and will most likely differ for each object. Studies to find an appropriate number were not performed as part of this experiment. The objects listed in this table should have near complete overlap; however, due to inadequate m-rep fits, the volume overlaps are less than expected.

In table 8.2, I have listed the mean and the standard deviation of the volume overlaps for some of the warped objects. These measures cannot be compared with the results in chapter 7 because the data sets used are not the same. We can see that certain objects such as the masseter and the thyroid can be very well predicted for the purposes of initialization. Other

object	volume overlap (%)	
	mean	std. dev.
left clavicle	81.1	2.35
right clavicle	84.2	2.61
mandible	78.7	3.70
trachea	67.6	6.75

Table 8.1: Mean and standard deviation of volume overlaps of each of the different objects used to determine the warp.

objects such as the sternocleidomastoid muscle have a reasonable prediction for initializing a segmentation method. Veins and arteries express a lot of variability across patients, and they are not predicted well by this method.

8.1.4 Discussion

We can see that the method of geodesic interpolation generates good initializations for some objects and poor initializations for others. Besides that, the method does have its disadvantages.

The first disadvantage is that the guarantee of a diffeomorphic warp is lost. Geodesic interpolation between m-reps does not provide any guarantees that the interpolated m-reps will be legal, specifically that their surfaces do not self-intersect. Self-intersecting surfaces will result in crossing paths being provided to the LDDM framework. However, these crossing paths are easy to detect, they do not occur frequently, and can be easily remedied. When two crossing paths are detected, both these paths and the paths between corresponding points in a small neighborhood around them are discarded from the set of paths contributing to the landmark paths in the LDDM framework.

The second disadvantage stems from the imprecise and inaccurate correspondence provided by m-reps. As stated in section 3.1.3, m-reps provide only approximate and large-scale fits to the object, and the fitting process is primarily geometric with very few image features. Interpolating between m-reps that have poor correspondence generates m-reps that also have

object	volume overlap (%)	
	mean	std. dev.
C3 vertebra	38.8	-
common carotid	8.35	5.26
cricoid cartilage	30.5	-
external jugular vein	2.52	-
hyoid bone	12.2	-
internal carotid	2.26	-
internal jugular	20.1	14.7
lateral pterygoid gland	28.0	-
masseter	64.8	3.0
medial scalene muscle	21.3	-
parotid glands	46.1	9.6
left sternocleidomastoid muscle	41.7	-
right sternocleidomastoid muscle	35.6	15.8
thyroid	56.4	5.23

Table 8.2: Mean and standard deviation of volume overlaps of each of the different objects on which the warp was applied. For structures for which I had only one sample, only the mean is listed and the standard deviation is not listed.

poor correspondence between points on their surfaces. The method detailed in chapter 5, with certain modifications, may be used to remedy this poor correspondence.

This method does offer the advantage of providing an interpolation between the source and target images.

8.2 Correspondence-preserving Heat-flow-based Interpolation between Images

The method presented in the previous section used paths obtained from interpolating m-reps in the LDDM framework. The method presented in this section obtains these LDDM-initializing paths from a different formulation of the steady-state heat flow problem that only delivers the diffeomorphism on the surface of the object and operates in four dimensions (as compared to the three dimensions of the method in chapter 6).

In this method, a special temperature distribution is laid down on the surface of the source and target objects such that corresponding points have the same temperatures. These are in turn embedded in a four-dimensional (4D) space made from the three usual spatial dimensions and a dimension representing the interpolative parameter. The source and target objects are respectively places at two values of the interpolative parameter. A steady-state heat flow solution is computed. This yields two-dimensional (2D) isotherms connecting the source and target points. Finally, an A* (A star) shortest path finding algorithm yields the necessary paths between the source and target object surfaces' points, which can be used with the LDDM framework.

A special temperature distribution, discussed in section 8.2.1, is laid down on the surface of the source object. The same temperatures are laid down on the corresponding surface points of the target object. The source object is embedded in one 3D facet of a 4D space, and the target object is embedded in the opposite 3D facet of this space. The temperatures of their surfaces are fixed; these serve as a set of Dirichlet boundary conditions. A set of von Neumann boundary conditions is imposed at the boundaries of this 4D space. A steady-state heat flow solution is obtained via the numerical methods described in section 3.3.2 and 6.2. This results in '2D isotherms' connecting the surface points on the source and target images. By their very definition – *surfaces of constant temperature* – no two isotherms can cross each other. Thus, paths between pairs of points of different temperatures will not cross each other.

To obtain one-dimensional (1D) curves connecting the source and target points, the next step of the computation needs to find a path through these 2D isotherms from each source point to the corresponding target point. The path I chose is the shortest path because it is unique and cannot cross another shortest path between another pair of corresponding points that lie on the same isotherm. The shortest paths are found by an A* search algorithm. Given a source point $\mathbf{a}$ and a target point $\mathbf{b}$, the A* algorithm uses a best-first search to find the least-cost path between them. In this case, cost is equivalent to distance, denoted by $s(\mathbf{a}, \mathbf{b})$. Given a heuristic $h(\mathbf{c}, \mathbf{b})$ that estimates the distance between points $\mathbf{c}$ and $\mathbf{b}$, for any other

point $\mathbf{c}$ that lies on the same two-dimensional isotherm containing $\mathbf{a}$ and $\mathbf{b}$, the A* algorithm estimates the distance $s(\mathbf{a}, \mathbf{b})$ through the equation

$$s(\mathbf{a}, \mathbf{b}) = s(\mathbf{a}, \mathbf{c}) + h(\mathbf{c}, \mathbf{b}). \tag{8.4}$$

If this heuristic always underestimates the distance, A* search is guaranteed to find the shortest path. The ordinary Euclidean distance between the points serves as a very good heuristic because the real distance between two points on a possibly curved manifold will always be greater than or equal to the Euclidean distance between them.

Sometimes, the steady-state heat flow solution may not contain an isotherm connecting the corresponding points. This will present in the form of an A* search that runs too long. In this case, search for a path connecting the pair of points is dropped, and the shortest paths found for a small neighborhood around these points are also discarded from the set of paths that will contribute to the landmark paths in the LDDM framework. It would be an extremely rare occurrence that no isotherm would be found connecting any pair of corresponding points between the source and target images. Thus, a diffeomorphism can almost always be found between the source and target images.

8.2.1 The Temperature Distribution

First, consider a 1D curve embedded in a 2D space for simplicity. I will then extend this to ordinary 2D surfaces embedded in 3D spaces.

Every surface point is assigned a temperature as a function of the parameter τ of the curve, $T = f(\tau)$. For an open curve, f should be monotonic. A monotonic function ensures a unique temperature for each point on the curve. The identity function is an example of a monotonic function that is typically used because of its simplicity. For a closed curve, a monotonic function will result in discontinuity at the point of the parameter discontinuity. Therefore, for

a closed curve, if the parameter τ ranges from 0 to 1, a modified function g is used:

$$g(\tau) = \begin{cases} f(\tau) & \forall \quad 0 \leq \tau < 0.5 \\ f(1-\tau) & \forall \ 0.5 \leq \tau < 1 \end{cases} \tag{8.5}$$

This function g is continuous everywhere as long as f is continuous.

The points at the 'poles' – the ones with extreme temperatures – need to be explicitly connected by an isotherm, which is provided as a fixed boundary condition.

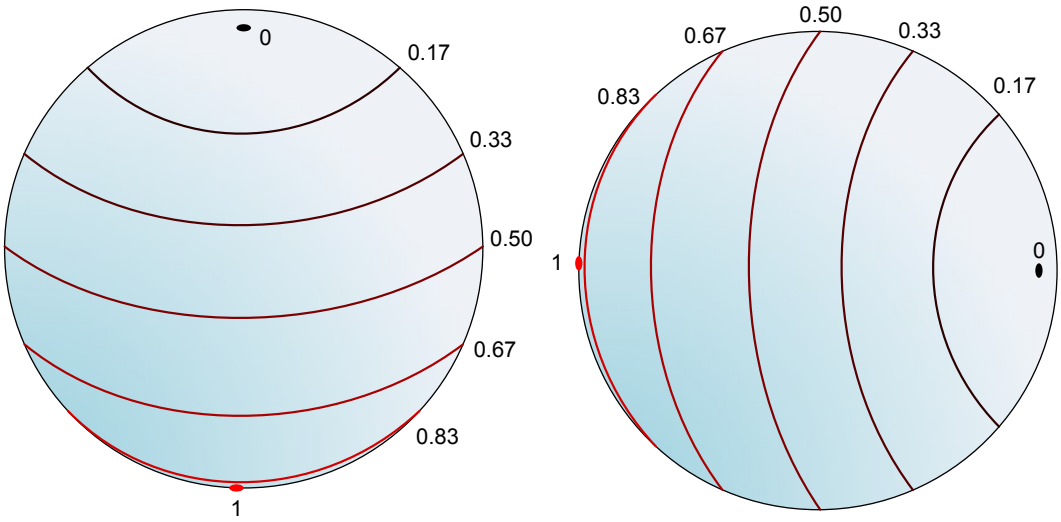


Figure 8.1: Initial temperature distribution for the two temperature components on a spherical surface. All points on a latitude are assigned the same temperature value. The poles are separated as far apart as possible for the two components.

For a 2D surface embedded in 3D space, a 2D temperature distribution is used, where each of the components satisfies the properties listed above. Such a temperature distribution can be seen in fig. 8.1. Such a temperature distribution has four poles – two for each of the components.

8.2.2 Discussion

Only preliminary trials were run using this heat-flow-based method. It was dropped in favor of the method presented in the previous chapters in this thesis because of a few disadvantages.

The whole process is computationally more expensive for three reasons: First, the A* search was computationally very expensive in the manner implemented. This could be remedied by reimplementing the search in a parallel manner since the shortest path finding is independent for each pair of corresponding points. Second, computing the steady-state heat flow solution for a 4D space takes significantly much more time and space than the 3D steady-state heat flow solution required by the method presented in chapter 6. Further, compared with geodesic interpolation, which runs on the order of seconds, this method can take up to an hour. Third, the process of producing the actual warp by using the paths generated in this method in an LDDM framework takes an additional amount of time on the order of tens of minutes. These components may be sped up by implementing on a Graphics Processing Unit (GPU), but the same holds true for the method of chapter 6.

The process of connecting and choosing the poles is somewhat arbitrary and is a disadvantage for the main application of this thesis, where one might have to connect poles of several closely-spaced objects. It may be possible to get rid of the poles entirely for closed objects by using temperature values that come from a group that matches the topology of the object. For example, for an object with spherical topology, the temperature distribution could be sampled from the group of points on a unit sphere $\mathbb{S}^2$. However, this temperature distribution cannot be extended to an ensemble of closely-spaced, possibly nested objects.

The main advantage this method has over all the other methods is the control provided to the user over the interpolation process. This can, for example, be important for an artist who wants to interpolate two- and three-dimensional objects where s/he desires control over the interpolation of certain points. Additional fixed boundary conditions along these paths can be easily introduced to take care of the artist's desires.

Chapter 9

Discussion and Conclusion

Several methods for segmentation using registration techniques have been proposed in the previous literature. However, sometimes they failed either on providing good correspondence between the images being registered or on producing a smooth, non-folded warp. The major contribution of this dissertation is in producing a correspondence-preserving smooth, non-folding warping method.

This dissertation has also worked towards automatic initialization of a complex of interrelated objects for segmentation. This is achieved by providing the correspondence-preserving smooth, non-folding warping method with correspondence between surfaces of some objects in the atlas and the target image that is initially obtained from medial models of these objects and is improved by using Oguz's [27] entropy-based correspondence method. To use medial models in the entropy-based correspondence method, the somewhat inadequate medial fits to the corresponding 3D objects have to be corrected. This dissertation introduced a steady-state-heat-flow-based method based on Dinh et al. [11] to do the same. This dissertation also detailed a method to model nearly tubular objects as medial models.

All of these contributions/methods were introduced in chapter 1. They are revisited and summarized along with their shortcomings and directions for future work in section 9.1 of this chapter. Some of their shortcomings and directions for future work were also explored in the individual chapters.

In my experiments, the warps have been using an individual patient's image as the atlas image, and the statistical shape models and the correspondence on them has been constructed from a limited number of training cases. This resulted in a less-than-adequate performance by some of the methods introduced in this dissertation. From the experimental standpoint, a mean atlas image and better trained statistical shape models would be a boon. This is discussed in section 9.2.

Throughout, I have emphasized the importance of developing a method that produces a smooth, non-folded warp, which is possibly even diffeomorphic. A diffeomorphism is one-to-one and smooth. However, from an application standpoint, is it necessary to have a mapping that is smooth everywhere? This question is addressed in section 9.3.

9.1 Summary of Contributions

In this section, I restate the claims made in chapter 1 and summarize the contributions of this dissertation towards fulfilling those claims.

9.1.1 Interpolating Objects using 4D Steady-state Heat Flow

Claim 1: *My method based on heat flow in 4D space can be used to interpolate objects while respecting correspondence.*

In this method, discussed in section 8.2, a special temperature tuple distribution is laid down on the surface of the source and target objects such that corresponding points have the same temperature tuples. These are in turn embedded in opposite 3D facets of a 4D space. A steady-state heat flow solution is computed. This yields two-dimensional isotherms connecting the source and target points. Paths in these two-dimensional isotherms are used to interpolate between the source and the target objects. Amidst the several paths, the shortest paths are chosen because they are unique and cannot intersect other shortest paths on the same manifold. An A* shortest path finding algorithm yields these shortest paths.

The main feature of this method is the control provided to the user over the interpolation process. This can, for example, be important for an artist who wants to interpolate two- and three-dimensional objects where s/he desires control over the interpolation of certain points. Additional fixed boundary conditions along these paths can be easily introduced to take care of the artist's desires.

The challenges of this method are in establishing the desirable temperature distribution in a multi-object case. The method was implemented and applied to a few simple object configurations; it was seen to produce credible results in these cases. This establishes claim 1.

9.1.2 Smooth, Non-folding Volumetric Warps using 3D Steady-state Heat Flow

Claim 2: *Interpolation of objects while warping is not always required; most of the time, we only care about the end result of the warp. My warp method, based on heat flow in 3D space, is guaranteed to produce a smooth, non-folded warp between the two images while respecting positional correspondences on object surfaces; it has been developed out of a desire to provide better guarantees and efficiency than the method described under claim* 1.

My method, described in chapter 6, assigns the same unique temperature tuple for corresponding points between two images, typically an atlas and a target image, and via a steady-state heat flow solution on the target image it extends this correspondence to the rest of the image. Corresponding points between the atlas and the target image are assigned the same tuple of temperatures; the dimensionality of the tuple is the same as the number of spatial dimensions, which is three. The different temperature components of the tuple are independent of each other. The temperature distribution laid on the source and target images is detailed in section 6.1. The temperature distribution in the target image is used as a boundary condition for a steady-state heat flow problem. Solving this steady-state heat flow problem, as detailed in section 6.2, and varying the underlying conductivity yields for each point in the atlas, exactly one point in the target image with the same set of temperatures. This smooth one-to-one

mapping is used to map points between the source and the target images.

The warp from the target image to the source image, obtained from the solution of a steady-state heat flow problem on a *homogenous* medium, may not be injective, i.e., have folds. To devoid this warp of folds, the heat conduction coefficient at every point in space is varied inside an optimizer with an objective function that measures the foldedness of the warp. A measure of foldedness was developed in section 6.3.

This method uses a steady-state heat flow solver described in sections 3.3.2 and 6.2. The speed of the method is determined by the speed of the solver. Although, the solver can be parallelized, the actual implementation executes in a serial manner. This can be easily rectified by rewriting the solver for a GPU, which can result in a significant reduction in the processing time.

The method was implemented and applied to several objects in the head and neck. The results, which are quite encouraging, are shown in chapter 7. This establishes claim 2.

9.1.3 Quasi-tubular Medial Models

Claim 3: *Modeling nearly tubular objects by conventional medial models is not stable. My method models objects that are nearly tubular in shape, giving reasonable correspondence and probability distributions on their shape spaces.*

Nearly tubular objects may be thought of an object produced by sweeping a nearly circular cross-section, which may change in shape, along a possibly curved line. Bronchi, blood vessels, and the colon are examples of tubular objects in the human body. Even the head and neck skin surface 4.1 can be thought of as a tubular object.

A medial model may be thought of as the locus (medial manifold) of the centers of the spheres that are tangential to the boundary being modeled. An m-rep discretizes the medial manifold into atoms (fig. 3.1) spread over a grid (fig. 3.2). Spokes radiating outwards from the atoms generate the modeled boundary.

The medial manifold of nearly tubular objects can be unstable (See Joshi et al. [23]). This

necessitated the development of specialized medial models that can stably model a population of nearly tubular objects.

In chapter 4, I have developed a quasi-medial model for nearly tubular objects. The modeling is done in two steps: First, a tubular model is constructed, and then deviations from the tubular model to match the actual boundary are incorporated. The tubular model consists of several atoms strung together from which now radiate a cone of spokes instead of two spokes as in medial models for slab-like objects. The deviation from a tubular to a nearly tubular model is achieved by letting all these spokes vary such that their ends continue to lie in the same plane. Means of training these models and statistically analyzing them were detailed. Finally, models of the head and neck, carotid arteries, jugular veins, and the rectum were generated from manually outlined patient images with this model. It was also used to segment rectums from 3D CT scans of patients. These results establish claim 3. Nevertheless, there are several avenues for improvement in this model.

Recall that a tube atom consists of a cone of a spokes, and these spokes can change their inclination with respect to the cone axis. This freedom in changing the inclination of the spokes may be unnecessary and could lead to a mode of variation in the statistics composed of noise from the training process. One could possibly represent the same tube model by having cones of differing inclinations. However, the structure of the training process leads to the trained model having cones with inclinations very similar to the initial model. This ensures that this variation can be kept at the minimum by judiciously choosing the initial model. Nevertheless, this variation was eliminated by restricting the angle of the spokes to 90° with respect to the cone axis. A model that included this restriction along with certain proprietary changes was developed at Morphormics[1] with significant improvements in the quality of rectum segmentations.

Several structures such as blood vessels and bronchi form branching tubes. The quasi-tubular medial models developed here require an extension to model a branching tube. Han et

[1]See the company web-site http://www.morphormics.com/ for more details.

al. [16] has developed a method that can model branching structures for slab-shaped medial models. Concepts from his method may be borrowed to develop quasi-tubular medial models for branching tubes.

9.1.4 Enhancing the M-rep Fit

Claim 4: *Shape models such as m-reps do not match the actual boundary precisely because they try to model the object at a large scale. This can be a problem when they are used as the basis of a warping mechanism. My method based on heat flows improves the correspondence by moving the points to the actual surface, when using shape models, specifically m-reps.*

The correspondence between the points on the surfaces implied by m-rep objects is neither precise nor accurate. Precision is related to the level of detail in the modeled objects, and accuracy is related to correctness in the reproduction of the modeled object. The limitation in precision is due to m-reps modeling anatomical structures at a large scale – by design they do not precisely follow the expertly outlined contours. The inaccuracy is due to the nature of the fitting process – a trade-off between geometric typicality and data match, which is itself a compromise between local data match at several locations. M-reps are used to model objects that are in turn used in a warping method that produces a warp strictly matching the input boundaries. If the input boundaries are not precise, this warp will reflect that imprecision. This necessitates the correction of the m-rep boundary so that it matches the input boundary precisely.

The labels on the m-rep-implied boundary surface are used to infer correspondence and are vital to the warping mechanism used. Thus, while correcting the m-rep boundary, the labels on the boundary need to be carried over correctly. Steady-state heat flows have been used before to carry over textures between different objects by Dinh et al. [11]. In section 5.1, I used a similar technique to correct the m-rep-implied boundary while carrying over the labels.

The m-rep boundary correction method moves the mesh elements comprising the segmented boundary, which in this case comes from m-reps, to a user-approved boundary. Sur-

face labels provided with the initial mesh are carried with the mesh elements when they are moved.

Non-crossing paths from a solution of a diffusion equation can be used to propagate labels from one surface to another. I used the Laplace's equation, which is a form of a diffusion equation. Streamlines of Laplace's equation never cross each other and provide a mechanism to propagate labels from one surface to another. Setting one surface to a fixed value of 0 and another surface to a value of 1 will produce equi-potential surfaces between the two surfaces of values between 0 and 1. Streamlines, which are perpendicular to these equi-potential surfaces, will be formed connecting the two surfaces. In order to produce the warped mesh, the vertices of one of the surfaces are propagated along these streamlines along with the neighbor and label information.

The correction to the m-rep fit, however, is limited to the resolution of the grid on which the steady-state heat flow problem is solved. The higher the resolution of the grid used for the solution, the longer will be the convergence time and the memory requirements for the solver for the steady-state heat flow solver. However, much of the space in this grid is too far away from the object. A small grid could be constructed that extends to a small distance from the object boundaries both towards the interior and the exterior. The size of this grid would be related to the surface area of the objects as opposed to the volume of the objects for the former grid. This could substantially reduce the time and space requirements of the solver as well as allow for the use of a grid with a higher resolution.

The results obtained from applying this method to objects from the head and neck, detailed in section 5.1.1, establish claim 4.

9.1.5 Entropy-based Correspondence to Improve M-rep-implied Correspondence

Claim 5: *The correspondence implied by m-reps may not be good enough for correspondence-based warping methods. Even after the points have been moved to the correct surface, they*

may still need to slide along the surface. My method based on Oguz et al. [27] improves this correspondence.

M-rep-implied boundaries do not fit the boundary of the modeled object accurately. The error in accuracy is due to the nature of the fitting process – a trade-off between geometric typicality and data match, which is itself a compromise between data match at several locations. As the warping method detailed in chapter 6 preserves input correspondence, the result can only be as good as the input correspondence. This necessitates the use of a method that can improve the correspondence to better reflect the underlying anatomical correspondence.

Anatomical correspondence may be closely related to geometry-based correspondence in the objects. For example both the principal curvatures at the nose tip attain a local maximum. Features such as curvatures and location can be used as cues for a correspondence method. Oguz et al. [27] have developed a method that takes just such cues into account along with the labeled segmentations of the organ from the atlas (the atlas here contains several segmentations) and the patient as inputs. It produces a labeled segmentation of patient's organ with improved correspondence to the atlas segmentation. This method has been discussed in section 5.2.

This method has two main shortcomings: One of the shortcomings of this method is its inability to properly handle images with varying spacing and sizes. Real-world CT images typically never have the same spacing or size. Any resampling that is performed deteriorates the quality of the input image. However, because of the limitation of this method, images are forced to be resampled. Therefore, the implementation of the entropy-based correspondence method should be modified so that it can accept images of varying spacing. The other shortcoming is an inability of this method to operate correctly on thin objects.

The results from applying this method to objects in the head and neck, detailed in section 5.2.1, establish claim 5.

9.1.6 Hierarchical Segmentation Results

Claim 6: *The head and neck area is a challenge for 3D image segmentation. My framework, using the methods developed, can deliver reasonable initializations for segmentation of certain objects in 3D head and neck CT images.*

An experiment, described in chapter 7, was run with the entire hierarchical method on a set of six patient images. All 15 possible pairs between these patients were formed. One of the patients served as an atlas, and the other served as the target patient for each of these pairs. The experiment was repeated with the roles of the patients reversed in each pair. Measurements of the overlap of the reference objects were presented in table 7.1. Measurements of the overlap of the target objects were presented in table 7.2. The same overlap measures were presented in table 7.3 after using an additional reference object.

The results presented in chapter 7, and the improvement in the results for organs such as the jugular veins and carotid arteries, which are close to the sternocleidomastoid muscle, after using the sternocleidomastoid muscle as a reference object provides evidence for the efficacy of the hierarchical segmentation approach, and they establish claim 6.

9.1.7 Anti-aliasing Binary Volumes

Claim 7: *Binary images can be efficaciously anti-aliased by a fourth-order Laplacian of curvature flow.*

The entropy-based correspondence method requires the computation of curvatures on the surface of objects. To robustly compute surface curvatures, any aliasing artifacts present in the mesh should be removed. A mesh extracted from a binary image is always aliased and needs to be anti-aliased. This necessitates the use of an anti-aliasing procedure on the mesh. I have developed a method that anti-aliases such meshes by a a fourth-order Laplacian of curvature flow. However, this method is not crucial to the functioning of the several components presented in this dissertation and is not presented here. A paper [30] on this method, which establishes claim 7, is in preparation at the time of writing this dissertation.

9.1.8 Thesis Statement

Thesis: *Atlas-based segmentation methods that use correspondence produce superior results on a complex of interrelated objects compared to those that do not. Methods that establish good correspondence, when combined with methods that produce a warp using the steady-state heat equation, efficaciously provide the required correspondence-based warps leading to good segmentations.*

Claim 1 demonstrates a method that can produce smooth, non-folded volumetric warps while strictly respecting input correspondence using steady-state heat flows. Claim 3 shows that medial models for nearly tubular objects can be developed, which along with previously developed medial models for slab-shaped objects can provide the necessary correspondence between two sets of images. These correspondences may not be accurate and precise. Claims 4 and 5 show that these correspondences can be improved by the combined use of another kind of steady-state heat flow problem that corrects the m-rep-implied boundary and an entropy-based correspondence method that improves the correspondence implied by m-reps. Finally claim 6 shows that the above combination of methods can be used towards segmentations of a complex of interrelated objects in the head and neck.

9.2 Number of Training Cases

Statistical methods such as statistical shape models, specifically m-reps, and entropy-based correspondence, typically perform better with a larger number of training cases. At present we are limited to six training images. This greatly limits the capabilities of the two statistical methods listed above.

First, statistical shape models, specifically medial models, typically produce more stable models with an increase in the amount of the input data. The statistical methods developed for quasi-tubular medial models could also greatly benefit from more training cases.

Second, the entropy-based correspondence method computes the entropy of the distribu-

tion of the corresponding points across several images. The entropy of the distribution will become more stable as the number of training cases increase. This in turn can lead to better correspondence between the different instances of the same object.

Eventually, it would be good to build an atlas image that is representative of the population, instead of using another patient as an atlas. Such an atlas image can be built using Davis's method [9], which produces a mean of images by imposing a metric on the space of images.

9.3 Diffeomorphism – an Appropriate Mapping?

Warps between an atlas and a patient image or between two images of the same patient at different times have typically been modeled as diffeomorphisms. This is mathematically convenient. However, certain anatomical changes do not follow a diffeomorphism. A diffeomorphism implies a smooth, one-to-one correspondence between two images. One of the ways in which this continuity and smoothness of the mapping is broken by organs sliding over one another between days. The sliding transform is defined as a one-to-one mapping that is smooth everywhere but can possibly be discontinuous along a finite set of co-dimension 1 surfaces. This has been a challenging problem in medical image analysis, but it may be solved using the correspondence-preserving steady-state heat flow based warping method that I have developed.

Let the surface along which sliding takes place be modeled by some other mechanism. The generated warp should be allowed to be discontinuous along this surface. In the steady-state heat flow formulation, this can be modeled with the presence of an infinitely thin insulator. This insulator in turn is represented by disallowing heat flow orthogonal to the surface. In the grid-based solution of the solver, this is modeled by disabling communication between neighboring nodes on opposite sides of the surface.

Bibliography

[1] Stephen Aylward, Stephen Pizer, David Eberly, and Elizabeth Bullitt. Intensity ridge and widths for tubular object segmentation and description. *MMBIA*, page 131, 1996. 38

[2] T. O. Binford. Visual perception by computer. In *Proceedings of the IEEE Conference on Systems and Control*, 1971. 35

[3] Fred L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–585, 1989. 24

[4] Robert Broadhurst, Joshua Stough, Stephen Pizer, and Edward Chaney. A statistical appearance model based on intensity quantiles histograms. *ISBI*, 2006. 6, 33, 35, 39, 53

[5] Joshua E. Cates, P. Thomas Fletcher, Martin Andreas Styner, Martha Elizabeth Shenton, and Ross T. Whitaker. Shape modeling and analysis with entropy-based particle systems. In *Proceedings of Information Processing in Medical Imaging*. Springer, 07 2007. 65

[6] Timothy F. Cootes, A. Hill, Christopher J. Taylor, and J. Haslam. The use of active shape models for locating structures in medical images. In *Proceedings of Information Processing in Medical Imaging*, pages 33–47. Springer, 1993. 33

[7] James N. Damon. Determining the geometry of boundaries of objects from medial data. *International Journal of Computer Vision*, 63:45–64, June 2005. 40, 42, 43

[8] James N. Damon. Swept regions and surfaces: Modeling and volumetric properties. *Theoretical Computer Science*, 392(1-3):66 – 91, 2008. Computational Algebraic Geometry and Applications. 37

[9] Brad C. Davis, P. Thomas Fletcher, Elizabeth Bullitt, and Sarang C. Joshi. Population shape regression from random design data. *International Journal of Computer Vision*, 90(2):255–266, 2010. 125

[10] M. de Bruijne, B. van Ginneken, M. Viergever, and W. Niessen. Adapting active shape models for 3D segmentation of tubular structures in medical images. *IPMI*, 2732:136–147, 2003. 37

[11] Huong Quynh Dinh, Anthony Yezzi, and Greg Turk. Texture transfer during shape transformation. *ACM Transactions on Graphics (TOG), Volume 24 Issue*, 2:289–310, 2005. 29, 60, 115, 120

[12] Thomas P. Fletcher, Cong Lin Lu, Stephen M. Pizer, and Sarang C. Joshi. Principal geodesic analysis for the study of nonlinear statistics of shape. *IEEE Transactions on Medical Imaging*, 23:995–1005, 2004. 20, 39, 45

[13] Yonatan Fridman, Stephen M. Pizer, Stephen Aylward, and Elizabeth Bullitt. Segmenting 3D branching tubular structures using cores. In *MICCAI*, pages 570–577, November 2003. 38

[14] Vincent Gregoire and Peter Levendag , et al. CT-based delineation of lymph node levels and related CTVs in the node-negative neck: DAHANCA, EORTC, GORTEC, NCIC, RTOG consensus guidelines. *Radiother Oncol.*, 69(3):227–236, Dec 2003. 15

[15] Eldad Haber and Jan Modersitzki. Image registration with guaranteed displacement regularity. *International Journal on Computer Vision*, 2006. 77, 85

[16] Qiong Han. *Proper Shape Representation of Single Figure and Multi-Figure Anatomical Objects*. PhD thesis, University of North Carolina at Chapel Hill, Chapel Hill, NC, 2008. 22, 120

[17] Qiong Han, Stephen M. Pizer, and James N. Damon. Interpolation in discrete single figure medial objects. In *Proceedings of Computer Vision and Pattern Recognition*, page 85, 2006. 43

[18] Qian Huang and G. C. Stockman. Generalized tube model: recognizing 3D elongated objects from 2D intensity images. In *CVPR*, pages 104–109, Jun 1993. 36

[19] Qian Huang and G.C. Stockman. Model-based automatic recognition of blood vessels from mr images and its 3D visualization. In *Proceedings of the IEEE International Conference on Image Processing*, volume 3, pages 691–695, Nov 1994. 36

[20] Qian Huang and G.C. Stockman. Generalized stochastic tube model: tracking 3D blood vessels in MR images. In *ICPR*, volume 2, pages 156–160, 9-13 Oct 1994. 37

[21] Stephen E. Jones, Bradley R. Buchbinder, and Itzhak Aharon. Three-dimensional mapping of cortical thickness using laplace's equation. *Human Brain Mapping*, 11:12–32, 2000. 29, 60

[22] Sarang C. Joshi and Michael I. Miller. Landmark matching via large deformation diffeomorphisms. *IEEE Transactions on Image Processing*, 9(8):1357–1370, August 2000. 24, 26, 27

[23] Sarang C. Joshi, Stephen M. Pizer, Thomas P. Fletcher, P. Yushkevich, A. Thall, and J.S. Marron. Multiscale deformable model segmentation and statistical shape analysis using medial descriptions. *IEEE Transactions on Medical Imaging*, 21(5):538–550, May 2002. 37, 118

[24] Jan J. Koenderink. *Solid Shape*. MIT Press, 1990. 34

[25] Derek Merck, Gregg Tracton, Rohit Saboo, Joshua Levy, Edward Chaney, Stephen Pizer, and Sarang Joshi. Training models of anatomic shape variability. *Medical Physics*, 35(8):3584–3596, 2008. 21, 52

[26] T. O'Donnell, T.E. Boult, Xi-Sheng Fang, and A. Gupta. The extruded generalized cylinder: a deformable model for object recovery. In *CVPR*, pages 174–181, 21-23 Jun 1994. 36

[27] Ipek Oguz, Joshua Cates, Thomas Fletcher, Ross Whitaker, Derek Cool, Stephen Aylward, and Martin Styner. Cortical correspondence using entropy-based particle systems and local features. *IEEE International Symposium on Biomedical Imaging*, 2008. 8, 11, 65, 115, 122

[28] Stephen M. Pizer, Thomas P. Fletcher, Yonatan Fridman, Daniel S. Fritsch, Graham A. Gash, John M. Glotzer, Sarang C. Joshi, Andrew Thall, Gregg Tracton, Paul Yushkevich, and Edward L. Chaney. Deformable m-reps for 3D medical image segmentation. *IJCV - Special UNC-MIDAG issue*, 55(2):85–106, Nov-Dec 2003. 7, 20, 38

[29] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press, 2007. 76

[30] Rohit R. Saboo, Marc Niethammer, and Stephen M. Pizer. Anti-aliasing binary volumes using fourth-order Laplacian of curvature flow. *in preparation*. 66, 123

[31] Kaleem Siddiqi and Stephen Pizer. *Medial Representations*, chapter 8. Springer, 2008. 20

[32] Joshua Stough, Robert E. Broadhurst, Stephen M. Pizer, and Edward L. Chaney. Regional appearance in deformable model segmentation. In *IPMI*, 2007. 6, 33, 35, 39, 53

[33] Demetri Terzopoulos, Andrew Witkin, and Michael Kass. Constraints on deformable models: recovering 3D shape and nongrid motion. *Artif. Intell.*, 36(1):91–123, 1988. 35, 36

[34] Andrew Thall. *Deformable Solid Modeling via Medial Sampling and Displacement Subdivision*. PhD thesis, University of North Carolina at Chapel Hill, 2004. 22

[35] F. Ulupinar and R. Nevatia. Shape from contour: straight homogeneous generalized cones. In *Proceedings of Computer Vision*, pages 582–586, Dec 1990. 35

[36] Jr. Yezzi, Anthony J. and Jerry L. Prince. An Eulerian PDE approach for computing tissue thickness. *IEEE Transactions on Medical Imaging*, 22(10):1332–1339, October 2003. 7, 29, 60