

Spine Segmentation in Computed Tomography Images using Geometric Flows and Shape Priors

Pulkit Khandelwal

McGill University
School of Computer Science
Faculty of Science

Submitted in partial fulfillment of the requirements for the degree
of Master of Science in Computer Science

August 2019

©Pulkit Khandelwal, 2019

Abstract

Surgical treatment of spine-related injuries requires the placement of pedicle screws. Precise localization of the individual vertebrae and surrounding tissues is thus essential to avoid damage to nearby regions. Image-guided surgery can help in surgical planning and thus improve prognosis. During surgery, preoperative patient scans are registered to intraoperative scans which allows surgeons to track the location of the surgical tools and better visualize their position with respect to the actual anatomy. In this thesis, we present a semi-automated pipeline to segment the human spinal column in computed tomography scans. These segmented anatomical structures would thus act as a model to which the intraoperative scans are later registered. We incorporate a shape prior into geometric active contours to augment the segmentation produced using region and boundary based terms. We have also applied ideas based on anisotropic diffusion and flux computation to preprocess the volumes to address the challenges faced when working with CT scans, such as region inhomogeneities within and outside the spine and a lack of signal at the vertebral boundaries due to partial volume effects. We validate our approach on three datasets and achieve results comparable to the state of the art. Our method also provides good localization and segmentation of the spinal canal and intervertebral discs as offshoots.

Resumé

Les opérations chirurgicales des lésions de la colonne vertébrale nécessitent la mise en place de vis pédiculaires. Localiser précisément les vertèbres et les tissus environnants est donc indispensable pour éviter d'endommager les régions voisines. Guider le chirurgien par l'image aide à la planification chirurgicale et améliore donc le pronostic. Pendant l'opération, des scans préopératoires des patients sont recalés sur des images captées en direct, ce qui permet aux chirurgiens de localiser leurs outils et mieux visualiser leur position par rapport à l'anatomie du patient. Dans cette thèse, nous présentons un processus semi-automatisé pour segmenter des scanners de tomodensitométrie et en extraire la colonne vertébrale. Ainsi segmentées, ces structures peuvent alors servir de modèle pour un recalage futur au moment de l'opération. Nous utilisons des contours géométriques actifs, s'appuyant sur la détection des contours, des régions, et utilisant une connaissance a priori de ces formes afin d'affiner la segmentation. Nous avons également utilisé des principes de la diffusion anisotrope et de calculs de flux afin de prétraiter les volumes. En effet, il s'agit de répondre aux défis usuels de la tomodensitométrie, comme par exemple l'hétérogénéité à l'intérieur et à l'extérieur de la colonne vertébrale ou les perturbations liées à la limite de résolution. Nous validons notre approche sur trois jeux de données et nous obtenons des résultats comparables à l'état de l'art. De notre méthode découle aussi une bonne localisation et une bonne segmentation du canal rachidien et des disques intervertébraux.

Acknowledgements

I would like to thank my supervisors Prof. Kaleem Siddiqi and Prof. Louis Collins for their constant support over the past three years. They have shaped my thought process and have stirred confidence in me to pursue a possible career in academia. Kaleem has helped me understand how important it is to have a solid foundation with strong fundamentals. Louis has taught me how to formulate a problem statement and ask good questions worthy of spending time on. I would also like to thank Prof. Peter Savadjiev, my thesis examiner, for his invaluable comments, which helped shaped the structure of this thesis. I spent Fall 2017 interning at Planet Labs in San Francisco, where I was mentored by Katherine Scott. I would like to thank her for introducing me to the world of software engineering and practices. Katherine has helped me become a better programmer, which in turn helped me in writing cleaner code for this thesis.

This thesis is theoretically motivated, experimentally validated and clinically driven. I hope that I keep working in this fashion. I would like to thank the open source community which has helped me fill in gaps in my understanding of concepts.

I also owe a lot to the city of Montréal. The great café culture, the cute bookstores, craft beer pubs, and the scenic bike paths have helped me keep myself a bit sane. To all the awesome baristas in Montréal: take a bow!

I would like to thank sources of funding from Kaleem and Louis.

Last but not the least, I owe a lot to the constant love and support of my parents and my sister.

Dedication

This thesis is dedicated to the people who are less privileged than I am and to whom education is a far away possibility. I hope one day each and every member of our society gets an opportunity to read, write and learn. This thesis is also dedicated to all the unseen forces of the universe which have helped me get to where I am today.

Abbreviations

IGS Image-guided surgery

DRA Dynamic Reference Array

CT Computed Tomography

MRI Magnetic Resonance Imaging

US Ultrasound

CSI Computational Methods and Clinical Applications for Spine Imaging

HU Hounsfield Unit

GAC Geodesic Active Contour

ASM Active Shape Model

PCA Principal Component Analysis

MRF Markov random Field

ROI Region of Interest

MAP Maximum a Posteriori

SDF Signed Distance Function

PDE Partial Differential Equation

KDE Kernel Density Estimation

SIM Statistical Interspace Model

CPU Central Processing Unit

GPU Graphics Processing Unit

ASSM Average Symmetric Surface Distance

HD Hausdorff Distance

HD95 Hausdorff Distance 95th percentile

AHD Average Hausdorff Distance

Symbols

\mathbf{E} represents an energy functional

\mathbf{F} represents a speed function

\mathcal{N} normal vector

\mathcal{T} tangent vector

\mathbf{V} a vector field

\mathbf{R}^n real coordinate space

Ω image domain

κ curvature

g stopping potential

$\mathbf{C}(p, t)$ family of smooth curves

H Heaviside function

δ Dirac-delta function

p arc-length parameterization

div divergence

B Bhattacharya coefficient

ψ signed distance function of a shape to build a shape prior

$\hat{\phi}$ average shape computed from ψ_i

ϕ_i eigenmodes representing shape variation

Contents

Abstract	i
Resumé	ii
Acknowledgements	ii
Abbreviations	v
Symbols	vii
1 Introduction	1
1.1 Motivation and Overview	1
1.2 Anatomy of the Spine	3
1.3 Imaging Modality: Computed Tomography	4
1.4 Contributions	5
2 Background	7
2.1 Literature Review: Vertebrae Segmentation	7
2.1.1 Image processing based segmentation	8

2.1.2	Active Surfaces	8
2.1.3	Appearance and Shape Modelling	11
2.1.4	Optimization: Graph-cuts, Markov Models and Variational Methods	13
2.1.5	Machine and Deep Learning	14
2.2	Geometric Flows and Shape based Image Segmentation	16
2.2.1	Boundary-based flows	21
2.2.2	Region and texture-based flows	22
3	Geometric Flows and Shape Priors for Spine Segmentation	28
3.1	Pre-processing of CT volumes	28
3.1.1	Challenges posed in spine segmentation	28
3.1.2	Non-linear Anisotropic Diffusion Filtering	30
3.1.3	Flux-maximizing Feature Map (flux-map)	33
3.2	Region-based segmentation of the spine	34
3.3	Shape-prior based flows	37
3.4	Segmentation procedure	45
4	Experimental Results and Discussion	47
4.1	Spine segmentation using region-based flow with shrink-wrapping: discussion	48
4.2	Shape-prior based vertebrae segmentation: discussion	63
5	Conclusion	68
5.1	Concluding remarks	68

5.2	Future Work	71
Appendix A Numerical Methods for Level-sets		72
A.1	Solution to the Hamilton-Jacobi Equation and Upwind scheme	72
A.2	Narrow band implementation and CFL condition	74
A.3	Typical front evolution update equation	74
Appendix B Fast Marching Methods		76
Appendix C Additive Operator Splitting		78
Appendix D Sparse Field Methods		81
Appendix E Segmentation Evaluation Metrics		83
E.1	Dice Coefficient	83
E.2	Average Symmetric Surface Distance (ASSD)	84
E.3	Hausdorff Distance (HD)	84
E.4	Average Hausdorff Distance (AHD)	84
E.5	Hausdorff Distance 95th percentile (HD95)	85
Appendix F Software used in this work		86
Bibliography		86

List of Tables

4.1	Spine segmentation results reported for Database 1	56
4.2	Spine segmentation results reported for Database 2	57
4.3	Spine segmentation results reported for Database 2: [Thoracic vs Lumbar Vertebrae comparison]	58
4.4	Spine segmentation results reported for Database 3	66

List of Figures

- 1.1 Image-guided surgery using iUS (intraoperative ultrasound): a) - the tracked iUS probe; b) acquiring tracked iUS in surgery within the surgical cavity (can also be acquired on the skin); c) the iUS image of a porcine phantom, overlaid on the preoperative CT before registration (used to show proposed technique on a human); d) iUS after registration; e) IGS navigation using the Medtronic system; e) the Medtronic IGS navigation display. Source of the image and description: [Goulet, 2010] 3

- 1.2 Anatomy of a Vertebra. Source: [Wikipedia, 2015] 4

- 2.1 Region-based segmentation of the spleen using [Rousson et al., 2003] on a sample CT image. Top left image shows the initialization with some manually placed seeds. The top right image shows the final obtained segmentation. The bottom left image is the groundtruth. The bottom right image shows the segmentation overlaid on top of the groundtruth. 26

- 3.1 Challenges in spine segmentation: the image on the extreme left shows anterior-posterior subluxation for vertebrae L4 and L5. The image in the middle show gaps on the vertebral boundary. The image on the right has missing boundaries as well. Also, notice that the intensity distribution is remarkably different in the rightmost image when compared to the first two images, as these images were acquired using different scanners and protocols. 29

- 3.2 Left column: sagittal and coronal slices of the original image, Middle column: edge-enhanced image (3.1.2) and Right-most column: flux-map (3.1.3). See text for a discussion. 35
- 3.3 Sample gold standard volumes comprising the training set of groundtruth segmentations of L4 vertebra. Notice the differences between the vertebral shapes. This is a representative database, with variations in the vertebral processes and in the thickness of the vertebral bodies. 40
- 3.4 Contribution of each of the eigenvectors to the total shape variability being captured by the vertebrae database. The red line indicates the 98th percentile, i.e., the first 15 eigenvectors account for 98% of shape variability in the database. 41
- 3.5 Shape Variations. The first four **eigenshapes** are depicted with first and second negative and positive shape variations obtained by $\bar{\phi} \pm \sigma i \lambda_i \phi_i$. Each row depicts one eigenshape. Left to right: in each row, one can see how the eigenshape changes. There is a visible change in the structure of the vertebral processes. Top to bottom: there are clear visible differences amongst the different eigenshapes, especially in the vertebral processes. 41
- 3.6 A flowchart summarizing the two segmentation pipelines. 46
- 4.1 Spine segmentation evaluation metrics using the region-based flow with shrink-wrapping for Database 1. The minimum, median, inter-quartile range, and the maximum values can be clearly seen. The values beyond those can be interpreted as outliers. 49

- 4.2 [Top row: left to right] The first image depicts the initialization by placing voxels. The second image shows the extracted surface using the region-based flow, which has cavities inside since the segmentation is concentrated on boundaries of the vertebrae. The third image shows a bounding-box placed around the extracted surface, which initializes the process of shrink-wrapping around the middle segmentation. [Bottom row: left to right] The leftmost image is the shrink-wrapped spine segmentation, which now contains a vertebral volume. The middle image shows a groundtruth image obtained by manual segmentation. The rightmost image shows an overlay of the obtained segmentation on the groundtruth. 51
- 4.3 [Left to right and top to bottom] Surface extraction process for Database 1, using the region-based flow on the flux-map, as explained in Section 3.2. The entire evolution procedure can be seen from initialization to the final extracted surface. 52
- 4.4 Sagittal slices: The top-left image shows the initial seeds placed on the edge-enhanced flux-map for an example from Database 1. The top-right image shows the extracted surface (vertebral boundary). The bottom left image shows the shrink-wrapped segmented spine. Notice how the cavities visible in the top-right image are now filled, and the surface holes are closed. The bottom-right image depicts the groundtruth. 53
- 4.5 Sample sagittal, coronal and axial segmented slices for Database 1. 54
- 4.6 Sample sagittal, coronal and axial segmented slices for the region not only limited to the available groundtruth for Database 1. Here, sacrum and thoracic (T12) is shown. 54
- 4.7 Three examples of the segmented spine (red) overlayed on the groundtruth (blue) for Database 1. 55

- 4.8 [Left to right and top to bottom] Surface extraction process for Database 2, using the region-based flow on the edge-enhanced image, as explained in Section 3.2. The entire evolution procedure can be seen from initialization to the final segmented upper thoracic vertebrae. 58
- 4.9 Left: Shown in light grey is the vertebrae region that is common to both, the groundtruth and the segmentation obtained by our method. The extra regions segmented such as the ribs, the coccyx and the sacrum, and other voxels in the spinal column are shown in red. Middle: A zoomed-in image of the upper thoracic region. Right: The segmented spine (shown in red) overlaid on top of the groundtruth (shown in light grey) after cropping-out the regions manually using ITK-SNAP. 59
- 4.10 The boundary of the segmentation (yellow) and groundtruth (green) overlaid on top of the image for a subject from Database 2 (CSI). *From left to right*: parts of upper thoracic, middle thoracic and lumbar region. [From top to bottom] axial, sagittal and coronal views. 60
- 4.11 Spine segmentation evaluation metrics using the region-based flow on the **lumbar region** with shrink-wrapping for Database 2. The minimum, median, inter-quartile range, and the maximum values can be clearly seen. The values beyond those can be interpreted as outliers. 61
- 4.12 Spine segmentation evaluation metrics using the region-based flow on the **thoracic region** with shrink-wrapping for Database 2. The minimum, median, inter-quartile range, and the maximum values can be clearly seen. The values beyond those can be interpreted as outliers. 62
- 4.13 Sample slices shown for the region-based surface extraction on the flux-map for Database 3. 63
- 4.14 Sample slices shown for the shrink-wrapping process for Database 3. 64

4.15	Shape-based segmentation evaluation metrics for Database 3. The minimum, median, inter-quartile range, and the maximum values can be clearly seen. The values beyond those can be interpreted as outliers.	65
4.16	Sample slices shown for the shape-prior based segmentation procedure for a subject from Database 3.	66
4.17	Three examples of segmentation (red) overlaid on the groundtruth (blue), for the shape prior based segmentation of L4 vertebra in Database 3. . . .	66
4.18	[<i>Left to right and top to bottom</i>] Shape-prior based segmentation process for a subject from Database 3. The entire evolution procedure is shown from the mean-shape initialization, to the final segmented volume.	67
5.1	Leftmost image: sagittal slice of femur in CT image. The flux-map is shown in the middle, and the right most image shows the segmented region using the region-based surface extraction and shrink-wrapping process.	69

Chapter 1

Introduction

1.1 Motivation and Overview

Each year hundreds of thousands of people around the world are surgically treated for spine-related injuries. Some of the most common injuries are spinal tumors, herniated discs, scoliosis, spinal stenosis, injuries to the craniocervical junction and osteoporosis. These surgical procedures often require placement of pedicle screws to provide better mechanical stability and to fuse the different vertebrae where required. It is thus very important for surgeons to visualize and keep track of where the surgical tools are with respect to the patient's anatomy so that there are no unwanted ramifications of the surgery such as damage to nearby organs or tissue due to screw malpositioning. To reduce these errors, image-guided surgery (IGS) might prove to be extremely useful.

In image-guided surgery [Drouin et al., 2017], a patient's preoperative images (Magnetic Resonance Imaging or Computed Tomography) are correlated with the patient's actual anatomy on a computer workstation. A typical IGS setup is composed of a computer workstation, camera dynamic reference array (DRA) and a number of specialized surgical tools [Goulet, 2010]. Prior to the surgery the patient is scanned to obtain 3D models of the spine, which are visualized on the workstation. During the surgery, the DRA is fixed above the region where the surgery is to be performed in the vertebra. The

camera then determines the relative 3D positions between the DRA and the surgical tools. Now, a registration step is required which defines a spatial transformation between the preoperative 3D models with the manually identified landmarks on the vertebrae of the patient.

Thus, the surgeon can visualize the position of the surgical tools with respect to the patient's anatomy in real-time. Also, instead of manually identifying the landmarks on the preoperative images and the patient's body, one can scan the patient intraoperatively using either CT or ultrasound. Thus, IGS helps in planning surgical navigation. In this thesis, we focus on the first part of IGS which involves segmentation of the preoperative CT scans to obtain 3D models of the spine. In particular, we incorporate knowledge of the geometry of the vertebrae with information gained from the distributions of the image intensities in the surrounding neighbourhood of the spine. We accomplish this via a formulation of geometric flows based on a level-set framework, to be discussed in Section 2.2. These methods are not only much more intuitive and practical than previous patch-based and complex shape-based methods but also prove to be accurate and precise. Though, in the past there has been some use of geometric level surfaces for vertebral segmentation in CT, none of the approaches thus far follow a systematic approach to design a segmentation pipeline which overcomes the challenges posed by CT images in image segmentation (as discussed in Chapter 3) or validates the practicality of the method on clinical data.

The rest of the present chapter gives a brief overview of the anatomy of the spine and how images are reconstructed by computed tomography (CT). Chapter 2 provides a thorough literature review of the previous methods in vertebrae segmentation in CT images. The second part of Chapter 2 presents an introduction to level-set methods. We review the mathematical formulation of how partial differential equations are used to segment regions of interest in images and higher dimensional volumes. We explore several different sources of information which should be included in the energy functional to obtain reasonable segmentations of the spine. Chapter 3 describes the proposed pipeline. First, we argue for the importance of anisotropic diffusion and flux-based methods to preprocess the

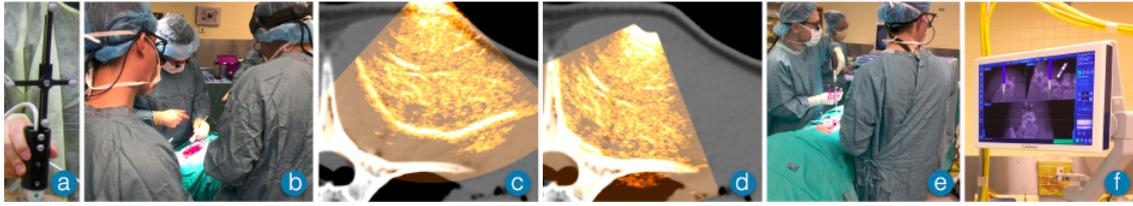


Figure 1.1: Image-guided surgery using iUS (intraoperative ultrasound): a) - the tracked iUS probe; b) acquiring tracked iUS in surgery within the surgical cavity (can also be acquired on the skin); c) the iUS image of a porcine phantom, overlaid on the preoperative CT before registration (used to show proposed technique on a human); d) iUS after registration; e) IGS navigation using the Medtronic system; f) the Medtronic IGS navigation display. Source of the image and description: [Goulet, 2010]

CT volumes. We then show how one can obtain entire spine and individual vertebrae segmentations using various energy functionals incorporating shape, region, and boundary terms. In Chapter 4 we discuss the results of the validation of the proposed pipeline applied to three datasets. We highlight the advantages of the approach proposed in this thesis over existing methods, and also address where this pipeline might fail. Chapter 5 concludes this thesis with a discussion of possible extensions to this work. In the appendices, the reader can find the details of the numerical methods used to implement these geometric flows using level-set ideas, the metrics used to evaluate the results, and the open source software packages used for visualization.

1.2 Anatomy of the Spine

We briefly review the anatomy of the human spine with the goal of understanding the challenges faced when segmenting the vertebral column in CT volumes. The human vertebral column consisting of thirty-three vertebrae is divided into five regions: cervical, thoracic, lumbar, sacrum and coccyx. There are seven cervical vertebrae, twelve thoracic vertebrae, and five lumbar vertebrae. These twenty-four vertebrae are separated from each other by intervertebral discs. The other nine vertebrae in the sacrum and coccyx regions are fused. Each individual vertebrae consists of two major parts: the vertebral body and the vertebral arch. The vertebral arch constitutes the posterior surface because it faces the back of the person. A cavity is present between these two parts, known as the

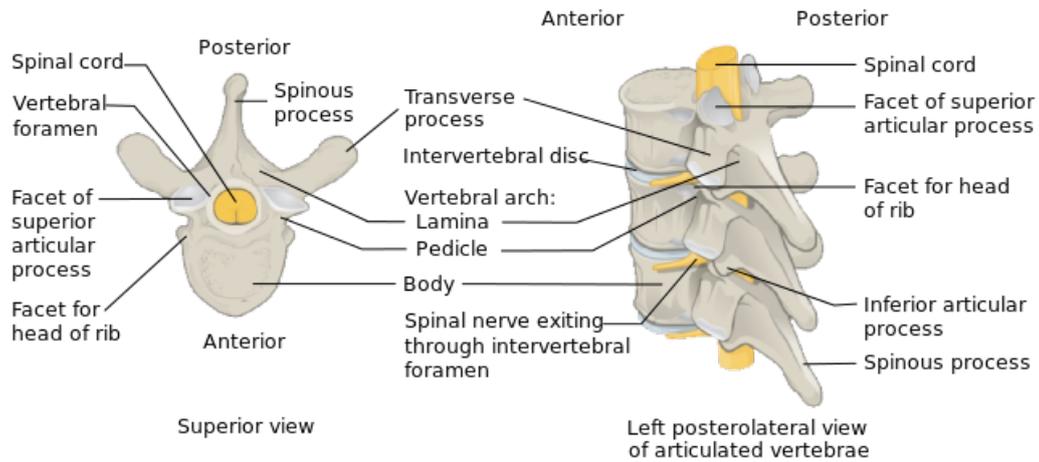


Figure 1.2: Anatomy of a Vertebra. Source: [Wikipedia, 2015]

vertebral foramen, where the spinal cord is found.

The vertebral arch is formed by a pair of pedicles and a pair of laminae and supports seven processes, four articular, two transverse, and one spinous. Two transverse processes and one spinous process are posterior to the vertebral body. Figure 1.2 depicts the anatomy of the spine and labels the different parts of a typical vertebra.

1.3 Imaging Modality: Computed Tomography

Computed Tomography, an imaging modality developed by Geoffrey N. Hounsfield is based upon the discovery of X-rays by Wilhelm Rontgen. Conventional X-ray imaging systems can only be used to obtain 2D images. To obtain information regarding depth, a CT scan produces an n-dimensional (where n is the depth) volume of the underlying structure of the region of interest. Although radiation doses in CT are relatively high, it has had a remarkable impact in the field of radiology. It has helped surgeons to help in the early diagnosis of tumors and track the progress of diseases.

CT helps in the reconstruction of the cross-sectional planar images of the body. In the simplest form, a CT scanner consists of a collimator (which emits X-rays), and a detector. This setup is rotated around the patient to obtain images at various angles. The X-rays

penetrates through the body and the detector measures the transmitted radiation. The degree to which the different organs and tissues attenuate the signal affects the final image reconstruction. For example, a dense structure, such as a bone, would absorb most of the X-rays, i.e., it attenuates the most. On the other hand, soft tissues tend to attenuate the least, i.e., they let the signal pass through the most. The X-rays collected on the detector are then converted to an electrical signal, which is then converted to digital images. Hence, the image acquisition process is known as computed tomography.

The higher the attenuation the larger the μ value (attenuation coefficient). For every slice scanned, the summed attenuation values are recorded in each direction. The image is then reconstructed using these μ values and called (filtered) Back Projection. Using Back Projection, the final grayscale value for each voxel is obtained. The attenuation coefficients are expressed in Hounsfield units (HU) and are scaled relative to the μ value of water. HU typically falls in the range -1000 to +1000. The window size (W) and level (L) are two visualization parameters set by the radiographer. If the window size is reduced then the contrast between the organs is enhanced. The centre of the window is adjusted using the Level (L) and this controls the contrast within an organ.

There are various kinds of CT scanners such as fan-beam, spiral or helical, and multi-slice. Iodine and barium are the most commonly used contrast agents. Contrast agents help in enhancing structures of interest and making the boundaries of structures more pronounced.

1.4 Contributions

The key contributions of this thesis are:

- the design of a novel CT image pre-processing pipeline based on anisotropic diffusion and flux-based geometric flows.
- the development of a level-set formulation to design a new clinically oriented pipeline

to segment the entire spine and also incorporate a shape-prior to segment individual vertebrae in computed tomography scans.

- the evaluation of the proposed methods on one in-house dataset and two publicly available datasets. We demonstrate that the method is fast and flexible, and requires very little effort on the part of the user to set up the segmentation procedure, which could potentially be useful to translate into a clinical setting.

Chapter 2

Background

We will begin our discussion by reviewing the previous approaches in the literature for vertebrae segmentation in CT images. We shall review the methods and results of the popular approaches, and discuss where they fail or improve upon other techniques. In the second section, we shall review the use of partial differential equations within a level-set framework for the task of image segmentation.

2.1 Literature Review: Vertebrae Segmentation

Here, we shall briefly review the previous methods for vertebrae segmentation. The methods can be broadly classified in the following categories: thresholding based techniques, active contours and level-sets, optimization techniques such as graph-cuts, deformable shape-based models, classical machine learning methods such as k-means, random forests and support vector machines used with handcrafted features, and data-driven deep learning methods.

The brief introduction to spinal anatomy in the previous chapter suggests that vertebrae segmentation is a non-trivial task. Different structures can be of interest in spine segmentation including intervertebral discs, the spinal cord, each vertebra or specific parts of each

vertebrae. The segmentation of one structure could potentially be helpful in identifying other structures.

2.1.1 Image processing based segmentation

Thresholding based methods are widely used to separate regions based on their raw intensity values. A fully automatic 3D segmentation method using adaptive thresholding was proposed in [Zhang et al., 2010]. An initial segmentation is first performed using a sensible threshold value to partition the image into classes containing bone vs. non-bone classes. This is then followed by automatic flood-filling and finally, the segmentation is refined by an iterative and adaptive thresholding step based on local connectivity and statistics. The method is fast and has been applied to several CT vertebrae datasets.

An interactive segmentation tool was developed in [Kaminsky et al., 2004] using different techniques such as logical and morphological operators, filtering, region growing, affine and rigid transformations. A very widely used open source software, *Slicer3D* [Kikinis et al., 2014], has also been used to segment the spine using simple thresholding.

2.1.2 Active Surfaces

Active contours (surfaces) evolve a deformable model to extract the region of interest in an image (volume). The evolution of the contour is governed by information extracted from image intensities, local and global regional characteristics, geometrical properties or shape-priors. We will discuss active surfaces using the level-set based formulation in Section 2.2.

Typical difficulties faced in vertebrae segmentation include the presence of double edges at the boundary. The outer edge tends to have a much higher density than the inner edge. [Tan et al., 2008] use a cascade of level-sets to segment vertebrae by exploiting image information at multiple scales. Each level-set follows the geodesic active contour

(GAC) [Caselles et al., 1995] formulation, to be discussed in Chapter 3, differing only in the criteria used for the gradient term. They have designed a pipeline which involves pre-processing using Gaussian smoothing. The series of level-set curve evolution first roughly segments the interior of the vertebra using a sigmoid for gradient information. The second level-set then uses a Laplacian as the speed function. The cortical bone is then segmented using a third level-set curve evolution. This cascade is finally post-processed via simple morphological operations such as erosion and dilation. This entire pipeline is applied again to a sub-sampled version of the original CT image. This resulting sub-sampled segmentation is then super-sampled back to the original scale and refined with yet another level-set evolution to give the final segmentation. This pipeline was originally developed to evaluate Ankylosing Spondylitis (a form of spinal disease). This method has a very detailed discussion on hyperparameter tuning for choosing the relative weights for the three terms in the GAC formulation. However, it lacks a formal motivation for how these parameters overcome the difficulties or affect the segmentation.

An edge and region-based level-set (ERBLS) with an Otsu adaptive threshold automatic initialization method was proposed by [Huang, 2013] which reconstructs 3D vertebral models from 2D axial segmented slices. The Otsu threshold [Otsu, 1979] gives a reasonable segmentation of the vertebral bone, since the bone image intensity is brighter than the rest of the structure. The level-set is then initialized with this rough segmentation. The initial level set inside the vertebral body evolves by splitting and merging. However, one might wonder why is the level-set curve evolution is required in this formulation if the thresholding itself nicely delineates the vertebral boundary. Secondly, the proposed method was not directly applied to 3D volumes but rather was used to segment each individual 2D slice, and hence does not exploit all the information available in a full 3D volume.

[Lim et al., 2013] introduced a *Willmore Force* term with a kernel density estimation based shape-prior within a level-set framework to produce good segmentation of the lumbar region in 3D. The shape energy is formulated as a sum of Gaussians of shape dissimilarity measures between the signed distance function (SDF) of a mean-shape and the SDF

of each individual vertebrae in the training set. The Willmore flow is based on a boundary energy functional which is a function of mean curvature, capturing the deviation of the evolving surface from local sphericity. This work has produced impressive results with an average Dice score of 89.00%. We believe that even better results could be obtained by incorporating region based information in the energy functional. Furthermore, a sensible pre-processing step can lead to the incorporation of smarter energy terms.

A related method was proposed in [Kim and Kim, 2009] using 3D deformable fences (3DDF). This approach constructs 3D fences that separate adjacent vertebrae from valley-emphasized Gaussian images. Initial curves for the 3D fences are extracted from intervertebral discs, detected with anatomical characteristics, then optimized using a deformable model. A minimum cost path finding method attempts to correct any erroneous curves trapped into a local minimum. The final volume is labeled with the help of the 3D fences by a fence-limited region growing method. This approach is quite complex to follow or reproduce as there are a number of heuristics to take care of, such as selecting the correct range of threshold values to produce a valley-emphasized Gaussian image. Moreover, to segment the vertebrae, the other structures such as the discs and the spinal cord are to be separately identified and segmented, which requires alternating between 2D and 3D segmentation. A number of steps then follow to correct erroneous segmentation in each part of the spinal column.

In an evaluation of the Medical Image Computing and Computer Assisted Intervention (MICCAI) Computational Methods and Clinical Applications for Spine Imaging (CSI) 2014 challenge¹, the authors in [Hammernik et al., 2015a] have proposed a better and simpler variational level set formulation which reports an average Dice score of $93.00 \pm 0.04\%$. A bone prior is obtained from the histogram profiles of the bone and other soft tissues. Separately, a mean-shape prior is obtained by aligning all the groundtruth segmentations of the training set of vertebrae. The thresholded bone prior is then registered to the mean-shape. A total variation energy functional is formulated which incorporates an edge detection function using a structure tensor, the mean-shape and a bone prior.

¹<http://csi-workshop.weebly.com/>

2.1.3 Appearance and Shape Modelling

Variation in shape complexity and pixel information can be captured from a training dataset of groundtruth segmented volumes to create a representative model of a structure of interest. Here, we shall review the most relevant literature for shape and appearance-based segmentation of the vertebrae. Choosing appropriate landmarks and having an efficient registration scheme is vital in model-based segmentation procedures. Furthermore, some methods segment each vertebra individually while others aim at segmenting a section of the spine (for example, the lumbar region) in one step.

One of the earlier approaches is described in [Klinder et al., 2009] where a pipeline for automatically detecting, identifying, and segmenting vertebrae in CT images has been proposed. Prior knowledge is incorporated based on models obtained via a generalized Hough transform and appearance-based models. A deformable approach is then used to perform segmentation by minimizing an energy functional based on image gradients.

In [Ibragimov et al., 2015] and then in a follow-up work in [Korez et al., 2015], the authors have combined detection with a shape-constrained model to segment the vertebrae. The detection is based on interpolation theory, comprising of an alignment step between a 3D mean shape mesh and each vertebra using an objective function, which computes the dot product of the normal vector at a vertex of the mean shape mesh and a Haar-like gradient at that vertex location in the CT image. A mesh deformation step is then employed to improve upon the alignment results and get a more accurate segmentation. This method achieves an average Dice score of $93.76 \pm 1.61\%$.

A part based ASM decomposition and conditional shape model based segmentation procedure was proposed in [Pereñez et al., 2015]. A Voronoi approach is used to decompose the vertebrae into different parts. This helps include fine details and account for shape variability at all the vertebral regions. Conditional models are used to model the statistical inter-relationships to establish statistical coherence of the different parts. Segmentation is then obtained using an ASM fitting procedure.

In a promising work by [Ibragimov et al., 2014], a landmark based shape representation using concepts from theory of transport and a landmark detection algorithm based on game theory is combined to segment vertebrae in 3D, which results in a Dice score of 93.50%.

The authors in [Shalaby et al., 2014] introduce a 2D vertebral body segmentation approach based on a shape prior constructed using principal component analysis (PCA) on the signed distance function of the training groundtruth segmentation images. An initial segmentation is then obtained using a tensor level set approach exploiting image intensity, Gaussian smoothed values and Gabor features for gradient information. This initial segmentation is registered with the shape prior to obtain the final segmentation.

In [Forsberg, 2015] the authors propose an atlas-based registration method which achieves an average Dice score of $94.00 \pm 0.03\%$. The method makes use of a pre-processing step which consists of spinal canal tracking and disc detection for estimating the local pose of each vertebra. This step is required for the initial alignment of the atlases before the registration of the atlases to the target data set. The final step is to merge labels of the deformed atlases into a single volume.

Superquadrics provide an elegant way to model plausible 3D shapes with a certain number of parameters. In [Stern et al., 2011] the vertebral body shape is modeled using a superquadric in the form of an elliptical cylinder. Using a set of 31 parameters the vertebral body anatomy is modeled for shape and pose. The initial model is deformed and aligned to segment the 3D image based on a similarity measure. The authors have carefully designed a criterion based on minimizing the overlap between the probability distributions of intensities inside and outside the vertebra; thus combining intensity and gradient information. This method reports a Dice score of 94.50%.

A statistical multi-vertebrae shape and pose model was developed in [Rasoulia et al., 2013]. A novel registration-based technique is used to segment the CT images of the spine rather than each individual vertebrae. The poses are represented by similarity transformations forming a Lie group. A group-wise Gaussian Mixture Model based registration technique,

combined with, partial and full Generalized Procrustes analysis with PCA was used to obtain the statistical multi-object shape+pose model. An initial segmentation using Canny edge detection is used to initialize the segmentation procedure. The model is then registered to this rough segmentation using an expectation-maximization algorithm.

In [Castro-Mateos et al., 2015] a statistical inter-space model was introduced, which directly addresses the neighbouring relationship between regions in different objects or regions of a multi-object structure by learning the statistical distribution between them to segment the vertebral column, and reports an average Dice score of 90% and a low Hausdorff error on the 2014 MICCAI challenge. They try to segment complex geometries involving overlaps, excessive separation of neighbouring structures and unrealistic orientations between neighbouring objects.

In summary, there has been a good amount of work based on shape modeling drawing on concepts from computer graphics, Lie groups, interpolation theory and game theory. We have seen new registration techniques being developed with novel similarity metrics. But, in this thesis we argue that shape and pose variations to segment each individual vertebra, and even the entire spine, can still be exploited in a much simpler setting, in particular, within a variational level-set framework, as described in Chapter 3.

2.1.4 Optimization: Graph-cuts, Markov Models and Variational Methods

Optimization-based methods to create 3D models have been quite successful. In [Kadoury and Paragios, 2010] a training set of prior mesh models is used to develop a low dimensional manifold embedding which establishes patterns of global shape variations, followed by the capture of appearance. At inference time, a higher-order Markov Random Field (MRF) is employed using unary and pairwise potentials to measure the similarity between data and shape.

A probabilistic model which approximates the marginal densities of the vertebral body and

the background using shape and appearance variations was developed in [Aslan et al., 2010]. A matched filter was used to detect the vertebral body. A graph-cuts formulation which integrates a linear combination of Gaussians (LCG), a Markov Gibbs Random Field (MGRF), and a distance probabilistic model obtained from 3D shape prior is used.

In the work by [Hammernik, 2015], the author proposed to solve a convex variational energy functional within an optimization framework. The variational energy functional is inspired by the PDE based image denoising methods and level-set based segmentation framework. The authors first smooth the image using the total variation norm image denoising model, referred to as the Rudin-Osher-Fatemi (ROF) model [Rudin et al., 1992] followed by the minimization of an energy functional incorporating a mean-shape model, bone priors and a structure tensor based anisotropic edge term. The optimization method used is based on Dijkstra's primal-dual formulation. This formulation involves a convex energy functional which guarantees a non-unique global convergence.

2.1.5 Machine and Deep Learning

Classical machine learning approaches often rely on the design of handcrafted features based on literature from the computer vision community, and then use classification techniques such as Support Vector Machines (SVMs), random forests, or logistic regression.

A learning-based method evaluated on the MICCAI CSI 2014 challenge is proposed in [Chu et al., 2015b] to automatically localize and segment 3D vertebral bodies. For the localization of 3D vertebral bodies, a random forest regressor is used to aggregate votes from a set of randomly sampled image patches to get a probability map of the center, which is then regularized by a Hidden Markov Model (HMM) to remove any ambiguity. This provides a region of interest (ROI) for the segmentation step, where the random forest classification is used to estimate the likelihood of being in the foreground or a background. As an end result of a detection pipeline, [Kelm et al., 2013] developed a segmentation approach based on iterative marginal space learning incorporating pose prior information.

The authors in [Wang et al., 2015] proposed a unified framework to segment multiple structures, i.e., axial and sagittal vertebral slices and discs in multiple modalities using a single learned model. This method leverages features extracted based on texture (WISIFT [Badino et al., 2012]), intensity patterns (WISURF [Badino et al., 2012]), semantic context (GIST [Oliva and Torralba, 2001]) and shape (HOG [Dalal and Triggs, 2005]) properties to train a multi-kernel multi-dimensional support vector regressor to predict the boundary of the structure of interest.

In the past few years, medical image analysis has been heavily influenced by data-driven deep learning methods. A major disadvantage of this paradigm is that these networks are not directly interpretable and as such the learnt features might not make sense to a clinician or a radiologist. The most recent work [Lessmann et al., 2019] has proposed an iterative instance segmentation approach that uses a fully convolutional neural network (ConvNet) achieving a Dice score of $96.3 \pm 1.3\%$ for the MICCAI CSI 2014 challenge dataset. The vertebrae are segmented one after the other by combining the network with a memory component which acts as a prior and is used to iteratively search for the next vertebra to be segmented. The network performs multiple tasks: segmentation of a vertebra, regression of its anatomical label and prediction of whether the vertebra is completely visible in the image. [Korez et al., 2016] designed a 3D ConvNet architecture which is used to learn the appearance of the vertebral bodies of MR images to generate 3D spatial probability maps, which guide the deformable model towards the boundaries of the vertebral bodies. In [Janssens et al., 2018] a cascade of fully convolutional networks (FCNs) is used to first localize the bounding box around the lumbar region together with a segmentation network to label the pixels as foreground or background.

Later, we shall show in Chapter 3 how our region-based level-set method can be used to obtain full spine segmentation which might then act as groundtruth labels for training deep networks, or would ease the workload of neurologists or radiologists. The segmentations produced are based on the local geometry of the vertebrae and the surrounding tissues.

2.2 Geometric Flows and Shape based Image Segmentation

In this section, we review the level-set based geometric flow formulation. Active contours extract the boundary of the region of interest by deforming a contour (in 2D) or a surface (in 3D) governed by certain forces, influenced by image and curve information. They can be either implicit or explicit depending upon their formulation. Level-sets and Fast Marching Methods are implicit. Level-sets are embedded in a higher dimensional manifold and solve an initial value problem, while fast marching methods are based on a boundary value partial differential equation. Snakes, which will be discussed below, are typically explicitly parameterized.

Explicit Models

Parametric snake [Kass et al., 1988] models can be described as a contour $\mathbf{v} = (x(s, t), y(s, t))$ that is time-varying and parameterized by $s \in [0, 1]$. In other words, it is an energy minimizing spline guided by external and internal guiding forces. The energy functional can then be written as

$$E_{snake}^* = \int_0^1 [E_{int}(\mathbf{v}(s)) + E_{image}(\mathbf{v}(s)) + E_{constraint}(\mathbf{v}(s))] ds. \quad (2.1)$$

Internal energy (E_{int}) makes the snake act like a membrane which controls the elasticity and the stiffness of the snake. The image-based features (E_{image}) exploit line, edge and curvature information such as stopping potentials. The snake can be constrained by external forces ($E_{constraint}$), including user interaction. The energy minimizing functional gives rise to a set of Euler equations for the evolving contour as given in [Kass et al., 1988]:

$$\alpha x_{ss} + \beta x_{ssss} + \frac{\partial E_{ext}}{\partial x} = 0 \quad (2.2)$$

$$\alpha y_{ss} + \beta y_{ssss} + \frac{\partial E_{ext}}{\partial y} = 0 \quad (2.3)$$

where, E_{ext} is an external energy functional and, α and β are constants. The introduction

of snakes have laid the foundation for many image segmentation methods. Researchers have built upon the initial model to develop smarter energy functionals and priors.

An important snake-based model, based on a Gradient Vector Flow (GVF), was proposed by [Xu and Prince, 1998], which solves the problems associated with initialization and boundary convergence. The energy functional to be minimized for the GVF field vector $\mathbf{v}(x, y) = [u(x, y), v(x, y)]$ is

$$\iint \mu(u_x^2 + u_y^2 + v_x^2 + v_y^2) + |\nabla f|^2 |\mathbf{v} - \nabla f|^2 dx dy \quad (2.4)$$

which is then solved by using the associated Euler-Lagrange formulation equations,

$$\mu \nabla u^2 - (u - f_x)(f_x^2 + f_y^2) = 0, \quad (2.5)$$

$$\mu \nabla v^2 - (v - f_y)(f_x^2 + f_y^2) = 0. \quad (2.6)$$

An edge-map $f(x, y)$ is derived by computing the magnitude of the gradient of the Gaussian smoothed image. The gradients of the edge-map ∇f point toward the edges and, generally have large magnitudes only in the immediate vicinity of the edges and have a value close to zero in homogeneous regions. The above energy functional has the following property. When ∇f is small, the energy is dominated by the sum of squares of the partial derivatives of the vector field, which yields a slowly varying field. On the other hand, when ∇f is large, the second term dominates and it keeps the vector \mathbf{v} equal to the gradient of the edge-map. Various snake-based formulations have been successfully applied to medical imagery with complex anatomy.

Implicit Models

[Osher and Sethian, 1988] introduced level-sets for capturing moving fronts to the computational physics community. We shall review this formulation, as applied to image analysis. We will further see how active surfaces are divided into different categories depending upon the image information they are designed to capture: boundary, region,

shape or topology of the structures in the image. Examples will be given for each of the three categories from the literature.

A time-dependent initial value partial differential equation problem is used to model an evolving front (contour) moving with a speed F . The speed function may depend upon factors such as the local geometrical properties of the front such as curvature, global properties that depend upon shape and position of the front, or any other independent property. The front is neither strictly positive nor negative and can move both forward and backward, with the possibility of passing through a point (x, y) several times. The initial position of the front is embedded as the zero level set of a higher dimensional function ϕ . Thus, we can think of the front to be given by the zero level-set of the time-dependent level-set function ϕ . To come up with a mathematical formulation, we require that the zero level-set of the evolving front, i.e., $\phi(x(t), t) = 0$ should be the evolving front. Differentiating this and applying the chain rule gives us:

$$\phi_t + \nabla\phi(x(t), t) \cdot \mathbf{x}'(t) = 0, \quad (2.7)$$

where, F is the speed in the outward normal direction, which gives, $\mathbf{x}'(t) \cdot n = F$ where $n = \frac{\nabla\phi}{|\nabla\phi|}$. Thus, we end up with an evolution equation for ϕ ,

$$\phi_t + F |\nabla\phi| = 0 \quad (2.8)$$

given $\phi(x, t = 0)$. Thus, Equation 2.8 describes the time evolution of the level-set function ϕ such that the zero level set of this evolving function is always identified with the propagating interface.

It should also be noted that a boundary value problem can be solved which leads to another class of computational techniques to model the evolving front, namely the Fast Marching Method [Sethian, 1996]. Assume that the front always moves outward; we can then characterize the position of the expanding front by computing its arrival time $T(x, y)$

as it crosses each point (x, y) . The equation can be written as

$$|\nabla T| F = 1. \quad (2.9)$$

$T = 0$ on τ , where τ is the initial location of the interface. This remains out of the scope of this thesis, but a short description has been given in Appendix B.

Let us now discuss the level-set formulation. Let γ be a simple, smooth, closed initial curve in R^2 and let $\gamma(t)$ be the one-parameter family of curves generated by moving γ along its normal vector field with speed F . Here, F is the given scalar function. We can thus write, $\mathbf{n} \cdot \mathbf{x}_t = F$, where \mathbf{x}_t is the position vector of the curve, t is time and \mathbf{n} is the unit normal to the curve. Let us now parameterize this equation. Let the speed function F be dependent on only its curvature κ , i.e. $F = F(\kappa)$. Let the parameterized position vector be $\mathbf{x}(s, t)$, where $0 \leq s \leq S$ assuming periodic boundary conditions $\mathbf{x}(0, t) = \mathbf{x}(S, t)$. Also, let $\mathbf{n}(s, t)$ and $\kappa(s, t)$ be the parameterized outward normal and curvature respectively at every point on the evolving curve. The individual components $\mathbf{x} = (x, y)$ are given by a Lagrangian representation describing the moving front:

$$x_t = F \left[\frac{x_s y_{ss} - y_s x_{ss}}{(x_s^2 + y_s^2)^{3/2}} \right] \left(\frac{y_s}{(x_s^2 + y_s^2)^{1/2}} \right) \quad (2.10)$$

$$y_t = -F \left[\frac{x_s y_{ss} - y_s x_{ss}}{(x_s^2 + y_s^2)^{3/2}} \right] \left(\frac{x_s}{(x_s^2 + y_s^2)^{1/2}} \right). \quad (2.11)$$

An early introduction of level-sets to image segmentation can be found in the work by [Malladi et al., 1995], which expands upon the above Lagrangian formulation. The key idea in the level-set approach is to represent the moving front $\gamma(t)$ as the level set $\phi = 0$ of a function ϕ . Thus, for a moving closed hypersurface $\gamma(t)$, an Eulerian formulation can be proposed for the motion of the hypersurface propagating along its normal direction with a speed of F . The propagating surface can be embedded as the zero level-set of a higher dimensional function ϕ . Let $\phi(x, t = 0)$, where $x \in \mathbb{R}^N$ is defined by,

$$\phi(x, t = 0) = \pm d, \quad (2.12)$$

where d is the distance from x to $\gamma(t = 0)$ and the plus or minus sign is chosen if the point x is outside or inside the initial hypersurface $\gamma(t = 0)$, respectively. Thus, we have a function $\phi(x, t = 0) : \mathbb{R}^N \rightarrow \mathbb{R}$ which gives us,

$$\gamma(t = 0) = \{x \mid \phi(x, t = 0) = 0\}. \quad (2.13)$$

A typical level-set embedding function is chosen as the signed euclidean distance function(SDF):

$$\phi(x) = \begin{cases} -d, & x \in \Omega^+ \\ +d, & x \in \Omega^- \\ 0, & x \in \Omega^0 \end{cases} \quad (2.14)$$

where Ω^+ is the region inside and Ω^- outside the front. As an example, imagine that the circle is the zero level set $\phi = 0$ of an initial surface $z = \phi(x, y, t = 0)$ in \mathbb{R}^3 . Thus we obtain Equation 2.8 as the evolution equation and refer to it as a Hamilton-Jacobi type of equation. This formulation of the propagating front as the zero level set of a higher dimensional surface has several advantages. First, the evolving front $\phi(x, t)$ always remains a function as long as the speed function F remains smooth. Second, since $\phi(x, t)$ remains a function as it evolves, we may use a discrete grid in the domain of x and substitute finite difference approximations for the spatial and temporal derivatives as explained in Appendix A. Third, intrinsic geometrical properties of the front may be easily determined from the level-set function ϕ such as the curvature. Moreover, sound numerical methods can be easily incorporated to account for the change of topology, breaks, merges, or singularities formed by the evolving curve. Finally, the level-set approach can be easily extended to higher dimensions, and as such level-sets have found its application in the segmentation of large volumetric data.

We use a level-set based framework for multi-dimensional medical imagery in this thesis. In Appendix A, we briefly discuss the numerical implementation of level-set methods. We shall now look at the design of various speed functions depending upon boundary and region terms, and the corresponding energy functional, along with qualitative examples.

In general, PDE based segmentation is dependent on the design of an energy functional given in a variational setting for the continuous case, which is then discretized and solved using numerical methods such as level-sets. We leave the discussion of using shape-priors as a cue until Chapter 3.

2.2.1 Boundary-based flows

A geometric flow could be designed in such a way such that the propagating interface stops at the boundaries of the desired object while moving freely in homogeneous regions. Thus to detect edges, a typical term is $g = \frac{1}{1+|\nabla I_\sigma|^n}$, first used for diffusion filtering in [Perona and Malik, 1990], which represents the inverse of the gradient magnitude of a Gaussian blurred image (I_σ). Intuitively, in a region of constant intensities, this term would be quite large as the gradient magnitude is negligible, and as such the front would propagate according to the other terms which influence the flow. On the other hand, the gradient would be high at the edges and this high gradient would lead to the term being close to zero which would help the flow slow down and eventually stop at the desired boundary.

[Malladi et al., 1995] use this term g in their level-set update equation as follows:

$$\frac{\partial \phi}{\partial t} = g (F_A + F_G) |\nabla \phi|, \quad (2.15)$$

where, F_A controls the expansion and contraction of the flow, and F_G depends on the geometry of the curve such as its curvature κ .

[Caselles et al., 1995] use a similar term in a level-set update equation:

$$\frac{\partial \phi}{\partial t} = g (c + \kappa) |\nabla \phi| + \nabla u \cdot \nabla g, \quad (2.16)$$

where the mean curvature κ is as given below and can be used as the speed function F

in 2.8,

$$\begin{aligned} \kappa &= \nabla \cdot \frac{\nabla \phi}{|\nabla \phi|} \\ &= \frac{(\phi_{yy} + \phi_{zz})\phi_x^2 + (\phi_{xx} + \phi_{zz})\phi_y^2 + (\phi_{xx} + \phi_{yy})\phi_z^2 - 2\phi_x\phi_y\phi_{xy} - 2\phi_x\phi_z\phi_{xz} - 2\phi_y\phi_z\phi_{yz}}{(\phi_x^2 + \phi_y^2 + \phi_z^2)^{3/2}}. \end{aligned} \quad (2.17)$$

Equation 2.16 can equivalently be written as,

$$C_t = g(c + \kappa)\mathcal{N} - (\nabla g \cdot \mathcal{N})\mathcal{N}, \quad (2.18)$$

where C_t is the evolving level-set. The inverse gradient function only helps in ideal cases. Practically, the edges don't have constant value throughout and are also marred by small gaps. The term $\nabla g \cdot \mathcal{N}$ in the above equation helps to detect boundaries with high variation in gradient magnitude values and the presence of small gaps.

2.2.2 Region and texture-based flows

Boundary-based flows leak through regions when the edge potential is low or the edges are not sharp. The image domain can also be partitioned based on the region statistics between the different subsets of the domain. Region-based flows exploit the geometrical and statistical properties of the local regions and are often modeled as an optimization problem. [Chan and Vese, 2001] considered a *minimum partition problem* to minimize an energy functional whose stopping term extracts information not only limited to the edges,

$$\begin{aligned} F(c_1, c_2, \mathbf{C}) &= \mu \text{Length}(\mathbf{C}) + v \text{Area}(\text{inside}(\mathbf{C})) \\ &= \lambda_1 \int_{\text{inside}(\mathbf{C})} |u_0(x, y) - c_1|^2 dx dy + \lambda_2 \int_{\text{outside}(\mathbf{C})} |u_0(x, y) - c_2|^2 dx dy \end{aligned} \quad (2.19)$$

The energy can be re-written as:

$$\begin{aligned}
F(c_1, c_2, \mathbf{C}) &= \mu \int_{\Omega} \delta(\phi(x, y)) |\nabla \phi(x, y)| dx dy + v \int_{\Omega} H(\phi(x, y)) dx dy \\
&+ \lambda_1 \int_{\Omega} |u_0(x, y) - c_1|^2 H(\phi(x, y)) dx dy \\
&+ \lambda_2 \int_{\Omega} |u_0(x, y) - c_2|^2 (1 - H(\phi(x, y))) dx dy,
\end{aligned} \tag{2.20}$$

where, δ is the Dirac-delta function and H is the Heaviside function. An evolving curve C in an image domain Ω can have a bi-partition defined as *inside*(C) and *outside*(C), respectively, with each having approximately piece-wise constant intensities. Let an object of interest be represented by intensity value u_0^i and the background with an intensity value u_0^o . Thus, we have $u_0 = u_0^i$ inside the object and $u_0 = u_0^o$ outside the object. In Equation 2.19, c_1 and c_2 are the mean intensity values inside and outside the propagating interface. Thus, the objective is to minimize an energy functional $F(c_1, c_2, \mathbf{C})$, i.e. $\text{Inf}_{c_1, c_2, \mathbf{C}} F(c_1, c_2, \mathbf{C})$.

The corresponding gradient descent update equation is as follows:

$$\frac{\partial \phi}{\partial t} = \delta(\phi) \left[\mu \nabla \cdot \frac{\nabla \phi}{|\nabla \phi|} - v - \lambda_1 (u_0 - c_1)^2 + \lambda_2 (u_0 - c_2)^2 \right], \tag{2.21}$$

where, λ_1 and λ_2 are the weights for the inside and outside terms. This region-based formulation perform good on images with regions of different intensities with possibly blurred boundaries. The major drawback of this model is that it assumes piece-wise constant intensities in the two sub-domains of the image and thus it might fail when applied to images with a considerable amount of noise. This calls for a regularization term which handles noise.

Another class of region-based level set algorithms makes use of probability densities to reason about the different regions, as given by [Rousson, 2004]. Using a *Maximum a posteriori* (MAP) formulation, we can write,

$$p(P(\Omega) | I) = \frac{p(I | P(\Omega))}{p(I)} p(P(\Omega)) \tag{2.22}$$

where $p(P(\Omega))$ is the probability of partition $P(\Omega)$ of the image domain and $p(I)$ is the probability of an image I . The term $p(I|P(\Omega))$ represents the *a posteriori* segmentation probability of an image I given the partition $P(\Omega)$. The image domain Ω can be composed of N regions:

$$p(I|P(\Omega)) = p(I|\Omega_1, \dots, \Omega_N) = p(I | \Omega_1)p(I|\Omega_2) \dots \dots p(I | \Omega_N), \quad (2.23)$$

where $p(I|\Omega_x)$ denotes the probability of having the image I when Ω_x corresponds to a region of interest. An energy functional can be formulated by taking the negative logarithm of the posterior,

$$E([\Omega_1, \Omega_2, \dots, \Omega_N, \theta_1, \theta_2, \dots, \theta_N]) = - \sum_i \int_{\Omega_i} \log(p(I(x) | \theta_i)) dx + \mu |\mathbf{C}|, \quad (2.24)$$

where $p(I|\theta_i)$ represents the *a posteriori* intensity distribution of a region Ω_i parameterized by θ_i . The term $\mu |\mathbf{C}|$ is obtained by taking the logarithm of a regularization constraint, $p(P(\Omega)) \propto e^{-\mu |\mathbf{C}|}$, which favors a short length $|\mathbf{C}|$ of the partition boundary. Consider the case when $N = 2$ with the region partitions being Ω_1 and Ω_2 , and ϕ being the signed distance function, the energy functional becomes,

$$\begin{aligned} E(\phi, [\theta_1, \theta_2]) &= - \int_{\Omega} (H(\phi) \log(p(I(x) | \theta_1)) + (1 - H(\phi)) \log(p(I(x) | \theta_2))) dx \\ &+ \mu \int_{\Omega} |\nabla H(\phi(x))| dx. \end{aligned} \quad (2.25)$$

The evolving equation is thus,

$$\frac{\partial \phi}{\partial t}(x) = \delta(\phi) \left[\mu \nabla \cdot \frac{\nabla \phi}{|\nabla \phi|} - \log \frac{p(I(x)|\theta_1)}{p(I(x) | \theta_2)} \right], \quad (2.26)$$

where δ is the Dirac-delta function and H is the Heaviside function. The parameters θ_i are updated after every iteration.

In the above image partitions, the individual regions were modelled by a single Gaussian as $p_i(I | \mu_i, \sigma_i) = \frac{1}{\sqrt{(2\pi\sigma_i^2)}} e^{-\frac{(I-\mu_i)^2}{2\sigma_i^2}}$. The variance term σ allows to capture regions with

different variances and noisy images are easily handled, which was a major drawback of the piece-wise constant assumption in the Chan-Vese formulation.

The Gaussian distribution assumption on image intensities gives a rather limited outlook to model the region distributions. This is because the Gaussian distribution assumes that the image regions are piece-wise continuous, but is often not the case in textured images. Kernel-based estimation of the distributions [Kim et al., 2005] within a variational framework gives,

$$\frac{\partial \phi}{\partial t}(x) = \delta(\phi) \left[\mu \frac{\nabla \phi}{|\nabla \phi|} - \log \frac{p_{c_{in}} I(x)}{p_{c_{out}} I(x)} + \sum_{C_x = C_{in}, C_{out}} \frac{1}{|C_x|} \int_{C_x} \frac{g_\sigma(I(x) - I(\hat{x}))}{p_{C_x}(I(\hat{x}))} d\hat{x} \right], \quad (2.27)$$

where the g_σ is a Gaussian kernel and the Parzen window is given by,

$$p_i(I) = \frac{1}{|\Omega_i|} \int_{\Omega_i} g_\sigma(I - \hat{I}(x)). \quad (2.28)$$

Several other region-based curve evolution techniques have been formulated.

[Michailovich et al., 2007] explicitly model the intensity distribution inside and outside the evolving front by minimizing the distance between two probability distributions using the Bhattacharya distance ($-\log(B)$) given by the Bhattacharya coefficient (B),

$$B = \int_{\mathcal{R}^N} \sqrt{p_1(\mathbf{z})p_2(\mathbf{z})} dz, \quad (2.29)$$

where $p_1(z)$ and $p_2(z)$ are two probability distributions. Letting “-” and “+” denote the foreground and the background regions being separated, the following update is applied for an evolving front given by the SDF ϕ as:

$$\begin{aligned} \frac{\partial \phi}{\partial t} &= \frac{1}{2} \delta(\phi(x)) \tilde{B}(\phi(x)) \left(\frac{1}{A_-} - \frac{1}{A_+} \right) \\ &+ \frac{1}{2} \int_{\mathbf{z} \in \mathcal{R}^N} K_+(\mathbf{z} - I(x)) \frac{1}{2} \sqrt{\frac{P_-(\mathbf{z} | \phi(x))}{P_+(\mathbf{z} | \phi(x))}} d\mathbf{z} \\ &- \frac{1}{2} \int_{\mathbf{z} \in \mathcal{R}^N} K_-(\mathbf{z} - I(x)) \frac{1}{2} \sqrt{\frac{P_+(\mathbf{z} | \phi(x))}{P_-(\mathbf{z} | \phi(x))}} d\mathbf{z}, \end{aligned} \quad (2.30)$$

where $\tilde{B}(\phi(x)) = \int_{\mathbf{z} \in \mathcal{R}^N} \sqrt{P_-(\mathbf{z} | \phi(x)) P_+(\mathbf{z} | \phi(x))}$, A_- and A_+ are the areas of the two competing regions and $I(x)$ is the image under consideration.

Figure 2.1 provides an example of spleen segmentation² in a CT image using a region based flow approach described in [Rousson et al., 2003]. We shall look into these methods in some detail when we study the region-based spine segmentation method in Chapter 3. [Mukherjee and Acton, 2015] use Legendre polynomials to generalize the Chan-Vese formulation.

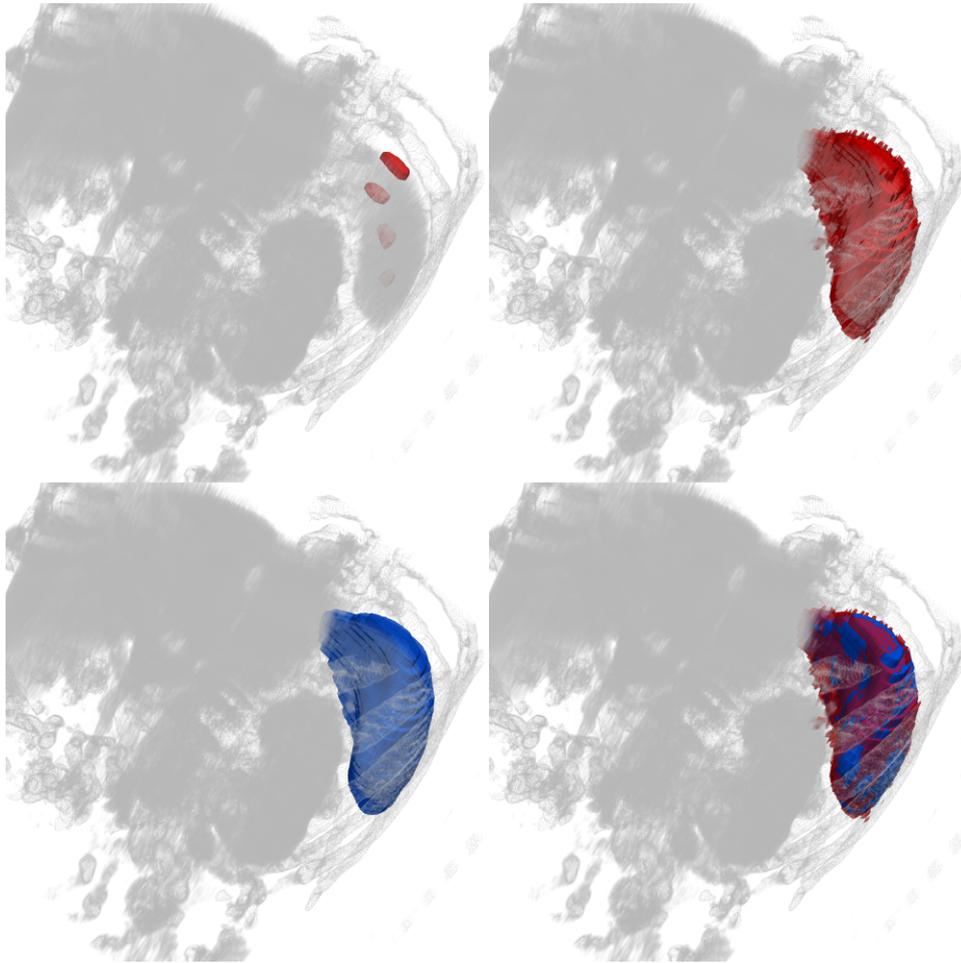


Figure 2.1: Region-based segmentation of the spleen using [Rousson et al., 2003] on a sample CT image. Top left image shows the initialization with some manually placed seeds. The top right image shows the final obtained segmentation. The bottom left image is the groundtruth. The bottom right image shows the segmentation overlaid on top of the groundtruth.

There are also other cues which could possibly be incorporated into an energy minimizing formulation, as additional feature vectors. Filters such as a Gabor [Turner, 1986], wavelets

²This image was obtained from <http://medicaldecathlon.com/>

and (non) linear structure and diffusion tensors have been used to handle non-smooth images. These cues can be easily coupled with boundary and region-based information within a variational formulation. A good reference can be found in [Cremers et al., 2007]. We shall further discuss non-linear structure tensors in Chapter 3, but for the particular case of non-linear diffusion filtering for image pre-processing.

Chapter 3

Geometric Flows and Shape Priors for Spine Segmentation

3.1 Pre-processing of CT volumes

3.1.1 Challenges posed in spine segmentation

Spine segmentation is a difficult task as several factors pose challenges. The spine is a complex structure, with each vertebra varying in shape. This often initiates a need to develop an explicit shape-prior customized for each vertebrae. In addition, the scans of patients with severe trauma might have different vertebrae fused together or have parts of the spine severely fractured. This makes it even harder to segment individual vertebrae separately. Furthermore, there might be gaps in the images of the vertebral bodies due to a loss of signal during the acquisition of the volumes or due to insufficient bone density leading to low edge contrast and high image noise. The outer cortical bone is much denser than the inner bone marrow. This in turn can cause the appearance of a double boundary and non-uniform thickness. The cortical bone appears as a distinctively bright intensity region, while the trabecular bone is similar to soft tissues. The intensity distribution inside and outside the vertebrae is very inhomogeneous and as such there

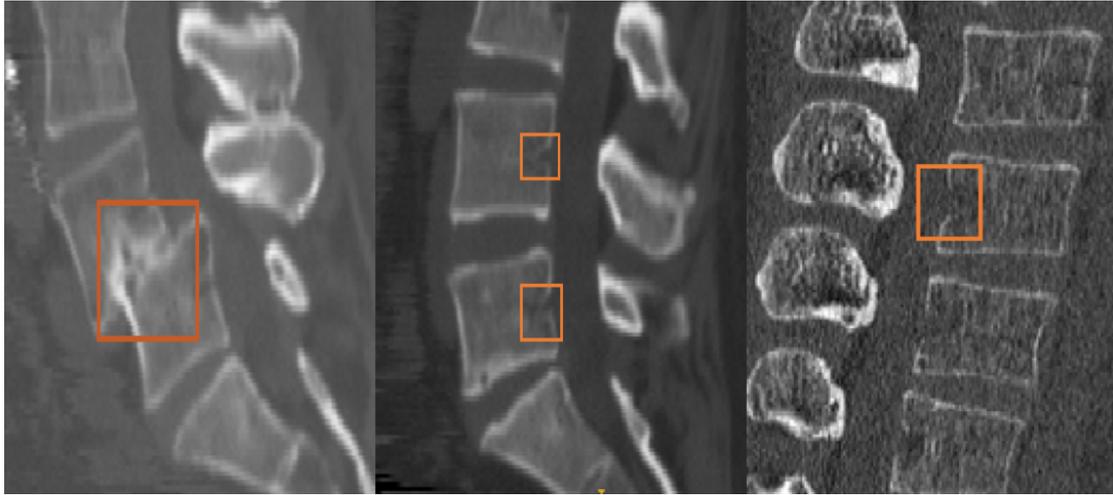


Figure 3.1: Challenges in spine segmentation: the image on the extreme left shows anterior-posterior subluxation for vertebrae L4 and L5. The image in the middle show gaps on the vertebral boundary. The image on the right has missing boundaries as well. Also, notice that the intensity distribution is remarkably different in the rightmost image when compared to the first two images, as these images were acquired using different scanners and protocols.

is no clear separation between these intensity distributions. Another key consideration while dealing with medical imagery is that the data itself comes from different centres acquired using variations in protocols with different scanner specifications. Thus, we employ smarter pre-processing methods to overcome these challenges, as discussed in the upcoming sections. See Figure 3.1 for some qualitative examples highlighting these challenges. Throughout this chapter, we will discuss ways in which we overcome these challenges in our segmentation pipeline.

A preliminary pre-processing step is the intensity normalization. We normalize the intensity values to a range between 0 and 1. First, we cap off the values above the 95th percentile as 1 and values below the 5th percentile as 0. The rest of the Hounsfield Units are now scaled between 0 and 1. In some images, we choose an appropriate percentile range to obtain some visible contrast between the vertebral column and rest of the image, as discussed in Chapter 4.

3.1.2 Non-linear Anisotropic Diffusion Filtering

Typically a Gaussian filter is used to smooth an image, to have smoothly varying gradients and avoid singularities. This often leads to an unnecessary loss in detail, for example, a Gaussian filter might substantially blur the edges of the vertebrae. Some of the earliest filtering work goes back to the method proposed in [Perona and Malik, 1990] where they designed a non-linear anisotropic edge preserving diffusion method. We also propose to use anisotropic diffusion-based filters to smooth the CT images. We make use of coherence and edge-enhanced diffusion filters [Weickert and Scharr, 2002] based on diffusion and structure tensors. These filters have the advantage that they make the vertebrae edges sharper and produce smooth homogeneous regions inside and outside the vertebrae, as they adapt to the underlying image structure. Intuitively, the edges are smoothed by elongated kernels and Gaussian-like kernels are used to smooth uniform regions.

In particular, we use the method proposed by [Kroon and Slump, 2009] which we describe below. Let I denote the 3D image, where each voxel is indexed by the 3-tuple (x, y, z) . For every voxel, we compute the following in an iterative step-wise fashion for a user-specified number of iterations,

- **Step 1: Structure Tensor**

For every voxel, we compute a 3 x 3 positive semi-definite structure tensor given by,

$$J(\nabla I) = K_j * (\nabla I \cdot \nabla I^T), \quad (3.1)$$

where ∇I is the image gradient and K_j is a Gaussian weighted function. The eigenvectors and eigenvalues are computed by eigen-analysis of the structure tensor,

$$J(\nabla I) = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{pmatrix} \cdot \begin{pmatrix} \mu_1 & 0 & 0 \\ 0 & \mu_2 & 0 \\ 0 & 0 & \mu_3 \end{pmatrix} \cdot \begin{pmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \mathbf{v}_3^T \end{pmatrix}. \quad (3.2)$$

The three eigenvectors $\mathbf{v}_1, \mathbf{v}_2$, and \mathbf{v}_3 give the local image orientation with the

three eigenvalues μ_1, μ_2 , and μ_3 describing the average contrast in those directions. Together with the corresponding eigenvalues, the eigenvectors characterize the local structural features of the image within a small neighborhood. Each eigenvalue reflects the gray level variance in the direction of the corresponding eigenvector. The eigenvector \mathbf{v}_1 , corresponding to the largest eigenvalue μ_1 defines a local average over the directions of maximum variance.

- **Step 2: Diffusion Tensor**

A diffusion tensor is constructed using the eigenvectors and eigenvalues obtained in Step 1. Let \mathbf{D} be the diffusion tensor computed as,

$$\mathbf{D} = \begin{pmatrix} D_{11} & D_{12} & D_{13} \\ D_{12} & D_{22} & D_{23} \\ D_{13} & D_{23} & D_{33} \end{pmatrix}. \quad (3.3)$$

There are several methods to compute the values of the diffusion tensor but here we are interested in the edge-enhancing method, which leads to the following set of equations,

$$\begin{aligned} D_{11} &= \lambda_1 v_{11}^2 + \lambda_2 v_{21}^2 + \lambda_3 v_{31}^2 \\ D_{22} &= \lambda_1 v_{12}^2 + \lambda_2 v_{22}^2 + \lambda_3 v_{32}^2 \\ D_{33} &= \lambda_1 v_{13}^2 + \lambda_2 v_{23}^2 + \lambda_3 v_{33}^2 \\ D_{12} &= \lambda_1 v_{11}v_{12} + \lambda_2 v_{21}v_{22} + \lambda_3 v_{31}v_{32} \\ D_{13} &= \lambda_1 v_{11}v_{13} + \lambda_2 v_{21}v_{23} + \lambda_3 v_{31}v_{33} \\ D_{23} &= \lambda_1 v_{12}v_{13} + \lambda_2 v_{22}v_{23} + \lambda_3 v_{32}v_{33} \end{aligned} \quad (3.4)$$

where the values λ_1, λ_2 and λ_3 are given in [Weickert and Schar, 2002] and are computed at every iteration as follows:

$$\begin{aligned} \lambda_1 &:= c_1 \\ \lambda_2 &:= c_1 + (1 - c_1)e^{\frac{-c_2}{(\mu_2 - \mu_3)^2}} \\ \lambda_3 &:= c_1 + (1 - c_1)e^{\frac{-c_2}{(\mu_1 - \mu_3)^2}} \end{aligned} \quad (3.5)$$

where $c_1 \in (0, 1)$ and $c_2 > 0$ are the smoothing and the edge-enhancement constants. The selection of the parameters λ_i can be understood as follows. The magnitude of diffusion in the orientation perpendicular to the gradient is determined by the difference $\mu_1 - \mu_3$ and $\mu_2 - \mu_3$. This difference is large for flow-like structures due to a high variance in one direction and consequently, such structures are strongly enhanced. Areas containing only noise are characterized by μ_i and therefore remain unmodified [Frangakis and Hegerl, 2001].

- **Step 3: Diffusion scheme**

The image is diffused for a certain number of iterations, specified by the user according to the equation below,

$$\frac{\partial I}{\partial t} = \nabla \cdot (\mathbf{D} \nabla I) \quad (3.6)$$

where $\nabla \cdot (\mathbf{D} \nabla I)$ can be solved using various formulations as described in [Kroon et al., 2010] and [Weickert and Scharr, 2002]. We use the Rotation Invariant Scheme computed as,

$$\nabla \cdot (\mathbf{D} \nabla I) = \partial_x j_1 + \partial_y j_2 + \partial_z j_3 \quad (3.7)$$

where the components j_1 , j_2 and j_3 are given by,

$$\begin{aligned} j_1 &:= D_{11}(\partial_x u) + D_{12}(\partial_y u) + D_{13}(\partial_z u) \\ j_2 &:= D_{12}(\partial_x u) + D_{22}(\partial_y u) + D_{23}(\partial_z u) \\ j_3 &:= D_{13}(\partial_x u) + D_{23}(\partial_y u) + D_{33}(\partial_z u). \end{aligned} \quad (3.8)$$

We refer the interested reader to the comprehensive numerical update scheme given in a longer version of the work [Kroon et al., 2010] which also describes the implementation.

3.1.3 Flux-maximizing Feature Map (flux-map)

Once, we have a diffused volume, a *flux-map* is produced which enhances the image contrast near the edges. Flux maximizing geometric flows were first proposed by [Vasilevskiy and Siddiqi, 2002] and later [Law and Chung, 2009] gave a $O(N \log N)$ running time algorithm for both spherical and anisotropic oriented flux computation compared to the traditional $O(s^2 N)$, where s is the radius of the disc, centered at a voxel, along whose boundary the flux is computed. In other words, the flux is computed as follows, at every voxel, consider a sphere of some radii which covers expected width of the vertebral boundary. We compute the outward flux using Eq. 3.9 and divide by the number of entries in the discrete sum. Now at each location, we select the flux value with the largest magnitude, over the range of radii considered. We limit our discussion to the key equations and the relevance of the method within the proposed segmentation pipeline, and refer the reader to both the papers for an in-depth mathematical formulation.

Let $\mathbf{C} = \mathbf{C}(p, t)$ be a smooth family of curves evolving in a plane parameterized by p and t where $0 \leq p \leq 1$ where the tangent \mathcal{T} and the inward normal \mathcal{N} are given by,

$$\mathcal{T} = \frac{(x_p, y_p)}{\|C_p\|} = (x_s, y_s), \mathcal{N} = \frac{(-y_p, x_p)}{\|C_p\|} = (-y_s, x_s)$$

and s is the arc-length parametrization of the curve. A vector-field $\mathbf{V} = (\mathbf{V}_1(x, y), \mathbf{V}_2(x, y))$ is defined for each point (x, y) in \mathcal{R}^2 . The total inward flux is then given by,

$$Flux(t) = \int_0^1 \langle \mathbf{V}, \mathcal{N} \rangle \|C_p\| dp = \int_0^{L(t)} \langle \mathbf{V}, \mathcal{N} \rangle ds, \quad (3.9)$$

where $L(t)$ is the euclidean length of the curve. The flux computation is easily extended to three dimensions as: let S be a surface embedded in \mathcal{R}^3 with coordinates indexed by (u, v) , $S_u = \frac{\partial S}{\partial u}$ and $S_v = \frac{\partial S}{\partial v}$ and the infinitesimal area is given by,

$$dS = (\|S_u\|^2 \|S_v\|^2 - \langle S_u, S_v \rangle^2)^{1/2} du dv.$$

Given a vector field $\mathbf{V} = (\mathbf{V}_1(x, y, z), \mathbf{V}_2(x, y, z), \mathbf{V}_3(x, y, z))$ defined at each point (x, y, z) in \mathcal{R}^3 , one can easily define the flux in 3D by the following surface integral:

$$Flux(t) = \int_0^1 \int_0^1 \langle \mathbf{V}, \mathcal{N} \rangle ds. \quad (3.10)$$

The main result proposed by [Vasilevskiy and Siddiqi, 2002] is: “The direction in which the inward flux of the vector field \mathbf{V} through the curve C is increasing most rapidly is given by $\frac{\partial \mathbf{C}}{\partial t} = div(\mathbf{V}) \mathcal{N}$. In other words, the gradient flow which maximizes the rate of increase of the total inward flux is obtained by moving each point of the curve in the direction of the inward normal by an amount proportional to the divergence of the vector field.”

We exploit this result to enhance the contrast of the vertebral edges. We do not evolve the flow equations mentioned in these two articles. Rather, we obtain the flux at every voxel of the edge-enhanced diffused image and generate a flux-based feature image (henceforth called **flux-map**) based on the fast implementation by [Law and Chung, 2009].

Figure 3.2 shows some qualitative examples which highlight the differences between the images before and after running the non-linear anisotropic diffusion scheme on the normalized vertebrae scans and the obtained flux-map using 3.9 based on the implementation of [Law and Chung, 2009]. The most notable advantages of this scheme include the closing of many of the interrupted edges, and the denoising of inhomogeneous regions. The edges also appear to be sharper than before. Qualitatively, the edge-enhanced flux-map looks piece-wise linear.

3.2 Region-based segmentation of the spine

One could argue that the edge-enhanced non-linear diffusion filtering and the flux-map might enhance spurious and unwanted edges inside the vertebral body. While this might be true, it does not affect our region or shape based segmentation pipeline, as will be

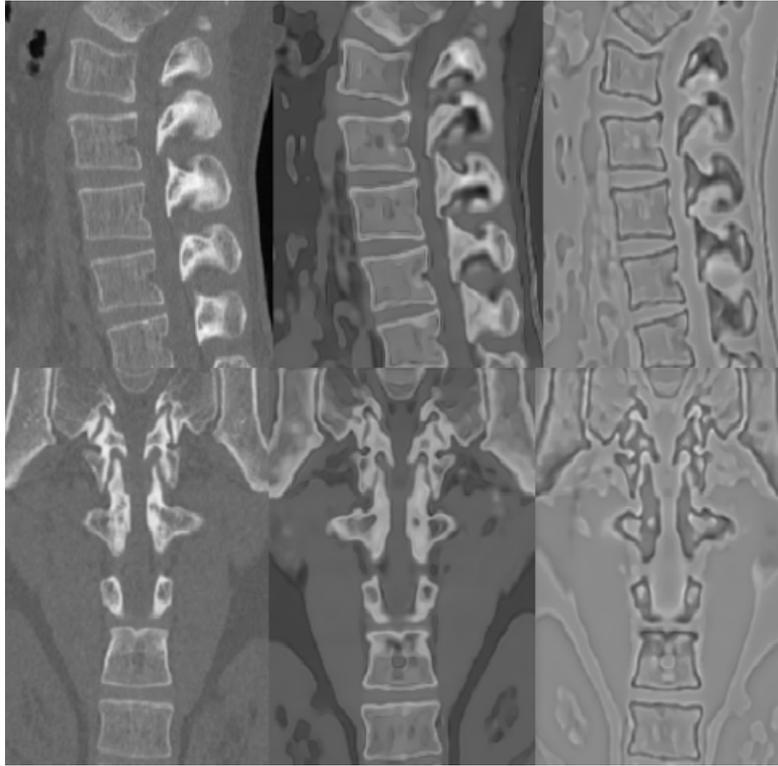


Figure 3.2: Left column: sagittal and coronal slices of the original image, Middle column: edge-enhanced image (3.1.2) and Right-most column: flux-map (3.1.3). See text for a discussion.

apparent from the formulation described in the next few sections. We leave the qualitative and quantitative discussion of the experimental results to Chapter 4.

Once a flux-map is obtained as described in the previous section, the user is required to place initial seeds at certain locations on the vertebral boundary or at the intersection of the vertebral body and the processes. Although placing a single voxel should suffice, it is recommended to have multiple seeds initialized to speed up the surface evolution process.

Now, we run a region-based flow with a sparse field method [Appendix D] for stable numerical updates. The main aim is to extract the surface of the spine, which has a distinct intensity distribution from that of the vertebral body and the region just outside the spine in the flux-map. We have a variety of region-based energy functionals implemented but we found that they were similar in performance in our applications as the filtered flux-map image is piece-wise constant. The results reported in this work are based on the level-set surface evolution proposed in [Chan and Vese, 2001], as discussed in Chapter 2. Some of the region-based flows which work well in our formulation are detailed in

[Michailovich et al., 2007], [Sandhu et al., 2008], [Dambreville et al., 2007], [Rathi et al., 2006] and [Rousson et al., 2003].

So far, we have placed seeds at the intersection of the vertebral boundaries and let them evolve using a region-based flow. This results in surface extraction of the spine, i.e., a thin sheet of the vertebral edges is reconstructed along the spinal boundary. Note that there might be some small gaps present on this extracted surface due to loss of signal in the CT images.

The next step is to fill the holes on the extracted surface and also the interior of the vertebra by the process of *shrink-wrap*. For this, a bounding box is used to initialize the geodesic active contour (GAC) [Caselles et al., 1995] based surface evolution. The inward flow wraps around the extracted surface and its zero level-set gives the final segmentation of the spine. This fills the interior of each of the vertebrae present in the image with the added advantage of closing small gaps present on the surface of the spine.

The *shrink-wrapping* problem could also be formulated as a surface reconstruction problem, where one can use the level-set update as given in [Zhao et al., 2001] which uses the SDF as the stopping criteria instead of the inverse of the gradient used in GAC. The formulation is very similar to GAC, as given below:

$$\frac{\partial \phi}{\partial t} = |\nabla \phi| \nabla \cdot \left[\frac{\nabla \phi}{|\nabla \phi|} \right] = |\nabla \phi| \left[\nabla d \cdot \frac{\nabla \phi}{|\nabla \phi|} + d \nabla \cdot \frac{\nabla \phi}{|\nabla \phi|} \right], \quad (3.11)$$

where d is the stopping criteria, computed by the SDF of the extracted spinal surface.

Note that we have based the implementation of the GAC using additive operator splitting [Appendix C] updates. We have also employed the morphological active contour formulation by [Marquez-Neila et al., 2013] to obtain much smoother segmented volumes. Note that in cases where there are large holes present on the extracted boundary, we did not shrink-wrap such extracted surfaces because the GAC leaks through large holes. Instead, we use the region-based flow directly on the edge-enhanced (not the flux-map) image followed by the shrink-wrapping step to get complete spine segmentations.

3.3 Shape-prior based flows

Using only a boundary or a region-based term might not be sufficient to extract the desired region of interest in an image. Additional cues such as texture-based features modelled by Gabor filters or local binary patterns could help. Another important ingredient is prior knowledge about the desired object shape and its characteristics in the evolving front. Complex anatomical structures, such as the spine provide a suitable test case, because they are largely rigid objects with a fixed part structure. Images with occlusions or missing parts of the desired objects could also benefit from priors. In this thesis, we use shape-priors to segment L4 lumbar vertebrae in a challenging dataset of patients with trauma as discussed in Chapter 4. In this dataset the trauma typically resulted in shifted vertebrae and dislocated spines. The pose variations modeled from the shape priors help in demarcating anatomy, when it does not adhere to an expected norm.

The shape-prior framework consists of two stages, the first being shape modelling stage, where the best possible shape characteristics are learnt from a training set, and the second being a constrained segmentation step using an energy functional minimization algorithm which optimizes the required model parameters.

One of the earliest approaches to shape-based segmentation was proposed in [Cootes et al., 1995] based on landmark points. Active shape models are statistical models which deform a shape to iteratively fit the desired region of interest. A set of landmark contour points are selected in a cohesive fashion across a set of training contours (or surfaces in 3D) obtaining one-to-one correspondences between each contour in the training set. A principal component analysis (PCA) is then applied on the landmark representations of the training set to reduce the dimensions and obtain the principal modes of variations with respect to a mean shape. This method works reasonably well, but it handles varying topology poorly and also requires careful placement of control landmark points which is a cumbersome task.

Most of the shape-prior based modelling in the literature is based on this PCA based

modelling. In this section, we review the key ideas of how a shape-prior is constructed from a training set. We then look at how this shape model is used within a level-set framework for segmentation. We will limit our discussion to the work by [Leventon et al., 2002] and [Tsai et al., 2003].

Principal Component Analysis

We review the PCA based modelling as proposed in [Tsai et al., 2003]. Consider a training set of n manually labelled binary masks $\{I_1, I_2, I_3, \dots, I_n\}$. Here, we shall consider a set of groundtruth segmentations of an individual vertebra (say, L4). These different masks are first *jointly aligned* using the following transformations to remove any pose variations while calculating the pose parameters $\mathbf{p} = [a \ b \ c \ h \ \theta]^T$, with a, b, c corresponding to components of 3D translation, h denoting a scaling constant and θ corresponding to 3D rotation respectively. In three dimensions, $\tilde{I}(\tilde{x}, \tilde{y}, \tilde{z})$ is the transformed image based on the pose parameters \mathbf{p} , where the transformed coordinates are given by,

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & a \\ 0 & 1 & 0 & b \\ 0 & 0 & 1 & c \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} h & 0 & 0 & 0 \\ 0 & h & 0 & 0 \\ 0 & 0 & h & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \theta & -\sin \theta & 0 & 0 \\ \sin \theta & \cos \theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}. \quad (3.12)$$

The series of transformations can be represented as a matrix $T[\mathbf{p}]$ to be the product of three matrices: a translation matrix $M(a, b, c)$, a scaling matrix $H(h)$ and a rotation matrix $R(\theta)$, thus mapping the coordinates $(x, y, z) \in \mathcal{R}^3$ to $(\tilde{x}, \tilde{y}, \tilde{z}) \in \mathcal{R}^3$. Note the $R(\theta)$ in Eq. 3.12 is given for the rotation about the z-axis. The standard transformations are also defined for rotations about x-axis and y-axis. In practice, a given order of rotations is fixed for the purpose of transforming the coordinates.

The following energy functional can be minimized using gradient descent:

$$E_{align} = \sum_{i=1}^n \sum_{j=1, j \neq i}^n \left[\frac{\int \int_{\Omega} (\tilde{I}_i - \tilde{I}_j)^2 dA}{\int \int_{\Omega} (\tilde{I}_i + \tilde{I}_j)^2 dA} \right] \quad (3.13)$$

where Ω is the image domain and thus align the binary volumes.

Using these aligned binary volumes, the signed distance function is computed for every shape in the training set $\{\Psi_1, \Psi_2, \Psi_3, \dots, \Psi_n\}$ with negative distances inside the vertebrae and positive outside representing the background. The average of these SDFs is $\bar{\phi} = (1/n) \sum_i \Psi_i$. To extract the shape variability, $\bar{\phi}$ is subtracted from each of the n SDFs to create n mean-offset functions $\{\tilde{\Psi}_1, \tilde{\Psi}_2, \tilde{\Psi}_3, \dots, \tilde{\Psi}_n\}$. These are the functions with which the variabilities in the shape population is captured via PCA.

We form n column vectors, $\tilde{\Phi}_i$ consisting of the n shapes $\tilde{\Psi}_i$. The image is converted to a column vector by arranging the voxels lexicographically. The total number of voxels or pixels in each image is N . A shape variability matrix is then defined, $S = [\tilde{\Phi}_1 \tilde{\Phi}_2 \dots \tilde{\Phi}_n]$ where S is a $N \times n$ matrix. An eigenvalue decomposition follows,

$$\frac{1}{n} \mathcal{S} \mathcal{S}^T = \mathcal{U} \Sigma \mathcal{U}^T, \quad (3.14)$$

where \mathcal{U} is a $N \times n$ matrix whose columns represent the n orthogonal modes of variations and Σ is an $n \times n$ diagonal matrix whose diagonal elements represent the corresponding eigenvalues. Calculating the eigenvalues and eigenvectors using the matrix $\mathcal{S} \mathcal{S}^T$ is costly. We can use the matrix $\mathcal{W} = \mathcal{S}^T \mathcal{S}$ to reduce the computational burden. If \mathbf{v} is an eigenvector then it is straightforward to see that $\mathcal{S} \mathbf{d}$ is the corresponding eigenvector of $\mathcal{S} \mathcal{S}^T$ and the corresponding eigenvalues of both the matrices remain the same.

We get n different eigenmodes $\{\Phi_1, \Phi_2, \Phi_3, \dots, \Phi_n\}$ which can be permuted to get the different *eiegnshapes* which capture variations in shape and pose. Normally, one limit to some $k \leq n$ which is sufficient to capture a variety of shape variations.

$$\Phi[w] = \bar{\Phi} + \sum_{i=1}^k w_i \phi_i, \quad (3.15)$$

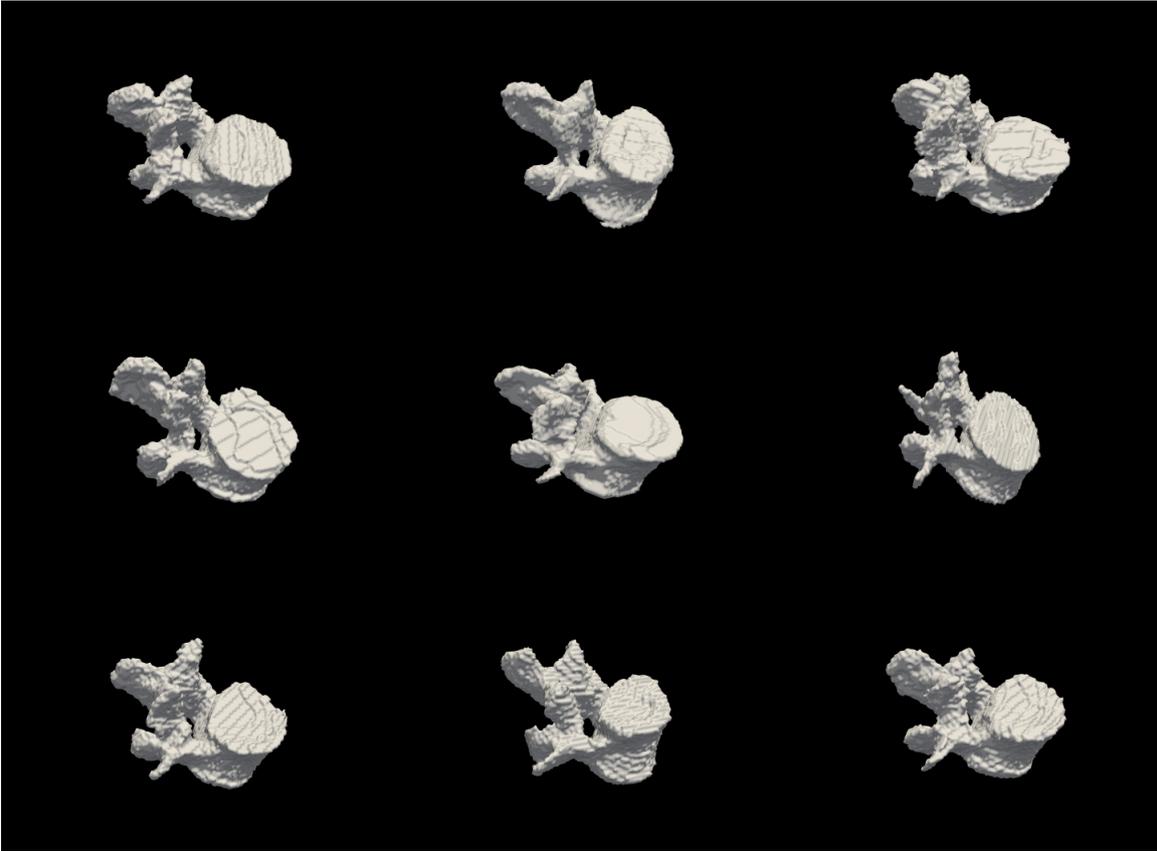


Figure 3.3: Sample gold standard volumes comprising the training set of groundtruth segmentations of L4 vertebra. Notice the differences between the vertebral shapes. This is a representative database, with variations in the vertebral processes and in the thickness of the vertebral bodies.

where $\mathbf{w} = \{w_1, w_2, w_3, \dots, w_n\}$ are the weights for the k eigenshapes with the variances $\{\sigma_1^2, \sigma_2^2, \sigma_3^2, \dots, \sigma_n^2\}$.

See Figures 3.3 through 3.5 for a sample shape population used to build the shape model using L4 vertebrae.

Integrating the PCA shape model into the Level-set framework

The PCA-based shape model described above is built offline. This model has to be integrated into a level-set framework to obtain a constrained segmentation. Usually edge and region-based terms are used in conjunction with the shape model to optimize the parameters to obtain a final segmentation.

[Leventon et al., 2002] were amongst the first to combine the edge and curvature infor-

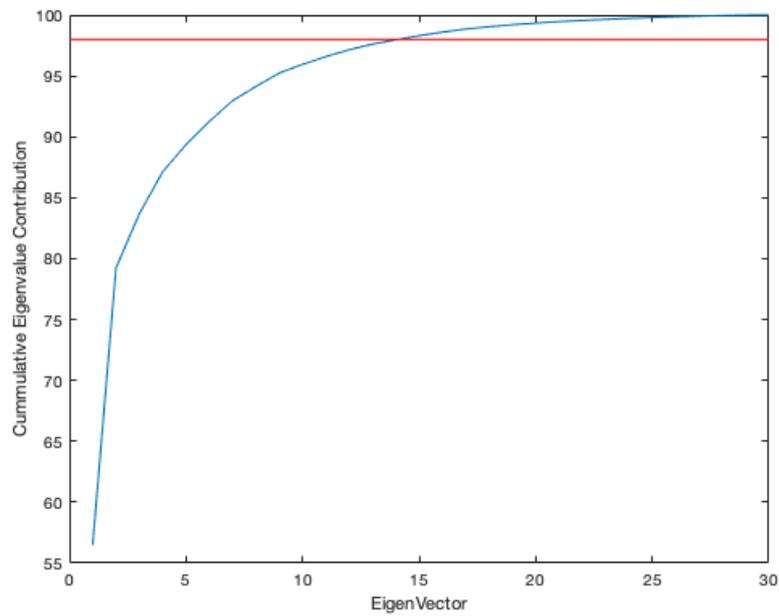


Figure 3.4: Contribution of each of the eigenvectors to the total shape variability being captured by the vertebrae database. The red line indicates the 98th percentile, i.e., the first 15 eigenvectors account for 98% of shape variability in the database.

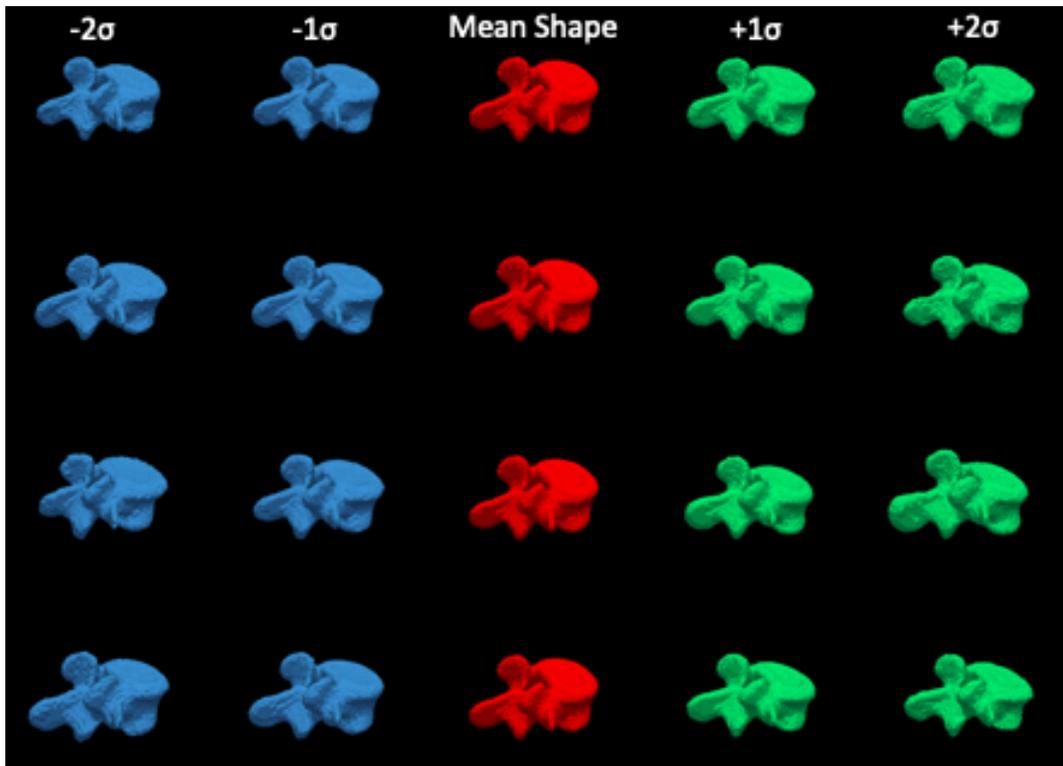


Figure 3.5: Shape Variations. The first four **eigenshapes** are depicted with first and second negative and positive shape variations obtained by $\bar{\phi} \pm \sigma i \lambda_i \phi_i$. Each row depicts one eigenshape. Left to right: in each row, one can see how the eigenshape changes. There is a visible change in the structure of the vertebral processes. Top to bottom: there are clear visible differences amongst the different eigenshapes, especially in the vertebral processes.

mation in a GAC evolution with the shape model using a *Maximum a Posteriori* (MAP) approach,

$$\alpha_{MAP}, p_{MAP} = \underset{\alpha, p}{\operatorname{argmax}} P(\alpha, p \mid u, \nabla I) \quad (3.16)$$

where, u is the evolving surface, ∇I is the gradient of the image, α and p are the shape and pose parameters, respectively, which completely determine the final segmentation (the final curve denoted by u^*). This MAP formulation is solved using Bayes' rule and α_{MAP}, p_{MAP} are the estimated shape and pose parameters.

Another approach is to iteratively minimize an energy functional using gradient descent to optimize the pose \mathbf{p} and shape \mathbf{w} parameters. [Tsai et al., 2003] computes the gradient of the region-based energy functional defined by [Chan and Vese, 2001] with respect to the parameters \mathbf{p} and \mathbf{w} , which is given by the following two equations:

$$\nabla_{\mathbf{w}} E_{CV} = -2(\mu \nabla \mathbf{w} S_u + v \nabla \mathbf{w} S_v) + (\mu^2 \nabla \mathbf{w} A_u) + v^2 \nabla \mathbf{w} A_v), \quad (3.17)$$

$$\nabla_{\mathbf{p}} E_{CV} = -2(\mu \nabla \mathbf{p} S_u + v \nabla \mathbf{p} S_v) + (\mu^2 \nabla \mathbf{p} A_u) + v^2 \nabla \mathbf{p} A_v), \quad (3.18)$$

where R^u and R^v are the regions inside and outside the evolving front respectively. μ and v represents the mean intensities in the two regions. S and A represents the sum of intensities and the area for a given region, respectively. The parameters are updated using the following equations at every iteration:

$$\mathbf{w}^{t+1} = \mathbf{w} - \alpha_{\mathbf{w}} \nabla_{\mathbf{w}} E, \quad (3.19)$$

$$\mathbf{p}^{t+1} = \mathbf{p} - \alpha_{\mathbf{p}} \nabla_{\mathbf{p}} E, \quad (3.20)$$

where $\alpha_{\mathbf{w}}$ and $\alpha_{\mathbf{p}}$ are the learning rates. At the end of an iteration, a new shape is computed using,

$$\phi[\mathbf{w}, \mathbf{p}](x, y) = \bar{\Phi}(\tilde{x}, \tilde{y}) + \sum_{i=1}^k w_i \phi_i(\tilde{x}, \tilde{y}) \quad (3.21)$$

where,

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \\ 1 \end{pmatrix} = T[\mathbf{p}] \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}.$$

The zero level-set of ϕ represents the evolving surface and upon convergence the final segmentation. *It is this formulation which we use in our shape-prior based segmentation for the L4 vertebra.*

We now review some of the other relevant shape-prior based methods in the literature. The method proposed in [Chan and Zhu, 2005] drives the evolving contour towards the mean shape and also has a separate Chan-Vese region term in the formulation but does not learn pose and shape variations from a general population. It is thus inefficient in handling complex anatomies. The major disadvantage of the above mentioned three methods, [Leventon et al., 2002], [Tsai et al., 2003] and [Chan and Zhu, 2005] is in their incapacity to handle local edge variations. This is because the energy models region information only and lacks a term to attract the evolving surface to the edges of the object under consideration. As discussed in Chapter 2, [Lim et al., 2013] use a PCA and KDE based shape prior with a boundary based term which constrains the flow to lie within the object of interest and not leak across the edges. The contribution of all the three terms, based on boundary, region, and shape, was proposed in [Bresson et al., 2006]. We review the self-explanatory energy functionals below. [Chan and Zhu, 2005] proposed an energy functional which combines the Chan-Vese functional with a shape term, driving the evolving curve to the mean-shape of the groundtruth training samples as,

$$\frac{\partial \phi}{\partial t} = -[(u - c_1)^2 - (u - c_2)^2 + 2\lambda(H(\phi) - H(\psi))] \delta(\phi), \quad (3.22)$$

where $H(\psi)$ is the Heaviside function of the mean-shape (here, computed by taking the average of all the aligned binary images). [Lim et al., 2013] combines a kernelized shape-

prior term with an edge-mounted Willmore force term as follows:

$$\begin{aligned}
\frac{\partial \phi}{\partial t} &= \frac{\partial E(\phi)}{\partial \phi} = \frac{\partial(\lambda E_s + E_w)}{\partial \phi}, \\
\frac{\partial E_s}{\partial \phi} &= \frac{\sum_{i=1}^N \alpha_i \frac{\partial d^2(H(\phi), H(\phi_i))}{\partial \phi}}{2\sigma^2 \sum_{i=1}^N \alpha_i}, \\
\frac{\partial E_w}{\partial \phi} &= -g(I) \|\nabla \phi\| (\Delta_M h + h(t) (\|S(t)\|_2^2 - \frac{1}{2}h(t)^2)),
\end{aligned} \tag{3.23}$$

where h is the mean curvature on M (a d -dimensional surface embedded in R^{d+1}). $\Delta_M h$ is the Laplace Beltrami operator given by: $\Delta h - h \frac{\partial h}{\partial n} - \frac{\partial^2 h}{\partial n^2}$, with $n = \frac{\nabla \phi}{\|\nabla \phi\|}$.

$S = (I - n \otimes n)(\nabla \times \nabla)\phi$ is the shape operator on ϕ . $\|S\|_2$ is the Frobenius norm of the shape operator S .

The set of flow equations which simultaneously evolve, combining boundary, region and shape terms, as proposed in [Bresson et al., 2006], are as follows:

$$\begin{aligned}
F &= \beta_s F_{shape}(C, \mathbf{x}_{pca}, \mathbf{x}_T) + \beta_b F_{boundary}(C) + \beta_r F_{region}(\mathbf{x}_{pca}, \mathbf{x}_T, u_{in}, u_{out}), \\
F_{shape} &= \int_0^1 \phi^2(\mathbf{x}_{pca}, h_{\mathbf{x}_T} C(q)) |C'(q)| dq, \\
F_{boundary} &= \int_0^1 g(|\nabla I(C(q))|) |C'(q)| dq, \\
F_{region} &= \int_{\Omega_{in}(\mathbf{x}_{pca}, \mathbf{x}_T)} (|I - u_{in}|^2 + \\
&\quad \mu |\nabla u_{in}|^2) d\Omega + \int_{\Omega_{out}(\mathbf{x}_{pca}, \mathbf{x}_T)} (|I - u_{out}|^2 + \mu |\nabla u_{out}|^2) d\Omega,
\end{aligned}$$

where q is the curve parameterization and the rest of the symbols have the usual standard meaning. An in-depth discussion will follow in Chapter 4, where we experimentally validate the methods described in the present chapter on several spine datasets. We also provide extensive visualizations to see the segmentation process and the results.

3.4 Segmentation procedure

There are several components in our geometric flow and shape-prior based segmentation pipeline, as summarized in the flowchart in Figure 3.6. We briefly review the two stage-pipeline below.

The normalized input image is pre-processed. First, the edge-enhanced image is obtained using the edge-enhancing flow (Section 3.1.2). The flux-map is then computed on the edge-enhanced image using an anisotropic flux method (Section 3.1.3). Now, the user can choose one of the two pipelines. Pipeline 1 is used for the segmentation of a desired region of the spine. Pipeline 2 is used for the segmentation of an individual vertebrae.

Pipeline 1: the user places seeds on the the flux-map image to initialize the region-based Chan-Vese segmentation flow (Section 3.2). This results in the surface extraction of the particular region of interest in the spinal column (for example, the lumbar or thoracic region). The user then initializes a bounding-box around the extracted surface to perform a type of shrink-wrapping, an inward flow wraps around the extracted surface and its zero level-set gives the final segmentation of the spine, which fills the interior of each of the vertebrae present in the image with the added advantage of closing small gaps present on the surface of the spine. For this process, we use the geodesic active contour flow (Equation 2.18) and obtain complete segmentations. Note that if there are large holes present on the extracted surface, the user can place seeds directly on the edge-enhanced image and run the Chan-Vese region-based segmentation to get segmentations of the spine (with some holes inside the vertebral bodies), followed by the shrink-wrapping step.

Pipeline 2: the PCA shape-prior is built offline. The user initializes the shape-prior based segmentation method by placing the mean-shape (of the groundtruth vertebrae population) near the vertebrae to be segmented, on the flux-map. The individual vertebra is then segmented using the shape-prior based segmentation method, as described in Section 3.3.

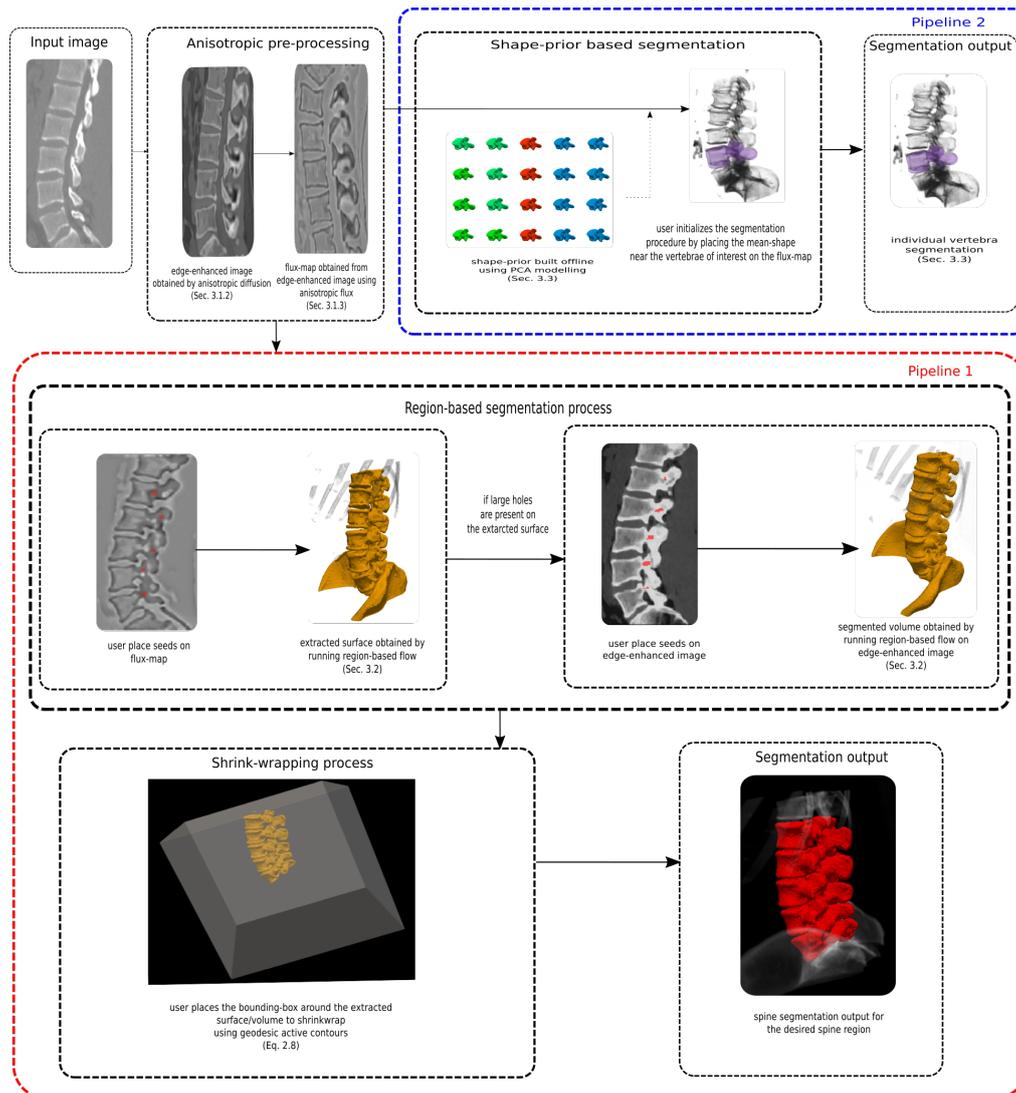


Figure 3.6: A flowchart summarizing the two segmentation pipelines.

Chapter 4

Experimental Results and Discussion

In this chapter we discuss our results in detail and compare the advantages and limitations with respect to other methods. We benchmark the performance of our proposed two-part segmentation pipeline on three databases.

Database 1: Healthy cases [Lumbar Vertebrae]

A publicly available database¹ [Ibragimov et al., 2014] of 50 vertebrae was extracted from 10 axially reconstructed CT images of lumbar spine, with in-plane voxel size between 0.282 and 0.791 mm, and slice thickness between 0.725 and 1.530 mm. The lumbar vertebrae (L1-L5) were manually segmented to obtain a binary mask for each vertebra.

Database 2: CSI MICCAI challenge [Thoracal-lumbar vertebrae]

The Department of Radiological Sciences, University of California, Irvine, School of Medicine² acquired data on the Philips or the Siemens multidetector CT scanners. The data sets are spine CT images acquired during daily clinical routine work in a trauma center from 10 young adults (16-35 years old). The datasets were acquired without intravenous contrast. The in-plane resolution is between 0.3125 mm and 3613 mm and the slice thickness is 1mm. In each scan, all 12 thoracic and 50 lumbar vertebrae, totalling, 120 thoracic and 50 lumbar vertebrae across 10 subjects have been segmented and provided

¹Database 1 can be downloaded at <http://lit.fe.uni-lj.si/tools.php?lang=eng>

²Dataset 2 at <http://spineweb.digitalimaginggroup.ca/spineweb/index.php?n=Main.Datasets>

as groundtruth references.

Database 3: Trauma cases

The Montréal Neurological Institute (MNI) provided CT images of the lumbar region of 30 patients. These were cases which had suffered a spinal injury, typically leading to a dislocation of the spine with a corresponding shift to the positions of one or more vertebrae. The images were acquired with 130 kVp tube potential and 175 A tube current. One image was acquired with 225 A tube current. Manual voxel-based segmentations of the L4 vertebra was provided for each image. Because the manual segmentations is voxel-based the precision is limited by the voxel size. The in-plane resolution is 512 x 512 voxels with a voxel size of 0.352 x 0.352 mm. The number of slices ranges from 55 to 200 with a slice thickness of 0.998 to 2.002 mm.

The metrics used to compare the obtained segmentation and the available groundtruth are summarized in Appendix E. We use the toolbox of [Taha and Hanbury, 2015] and [Maier et al., 2019] to produce the comparison metrics for the purpose of the evaluation of our segmentation results.

4.1 Spine segmentation using region-based flow with shrink-wrapping: discussion

We evaluate the proposed region-based spine segmentation pipeline on the obtained edge-enhanced flux-maps, as discussed previously in Section 3.2 on the healthy population in Databases 1 and 2.

We first discuss the results on Database 1. We report an average Dice score of $92.36 \pm 0.0131\%$ with the average symmetric surface distance at 0.7884 ± 0.1385 mm. The average Hausdorff and its 95th percentile are at 0.1385 ± 0.0845 mm and 2.1638 ± 0.6464 mm respectively. The average surface distance is 0.6626 ± 0.1371 mm. Figure 4.1 summarizes the obtained metrics used to evaluate the segmentation results as a box-plot.

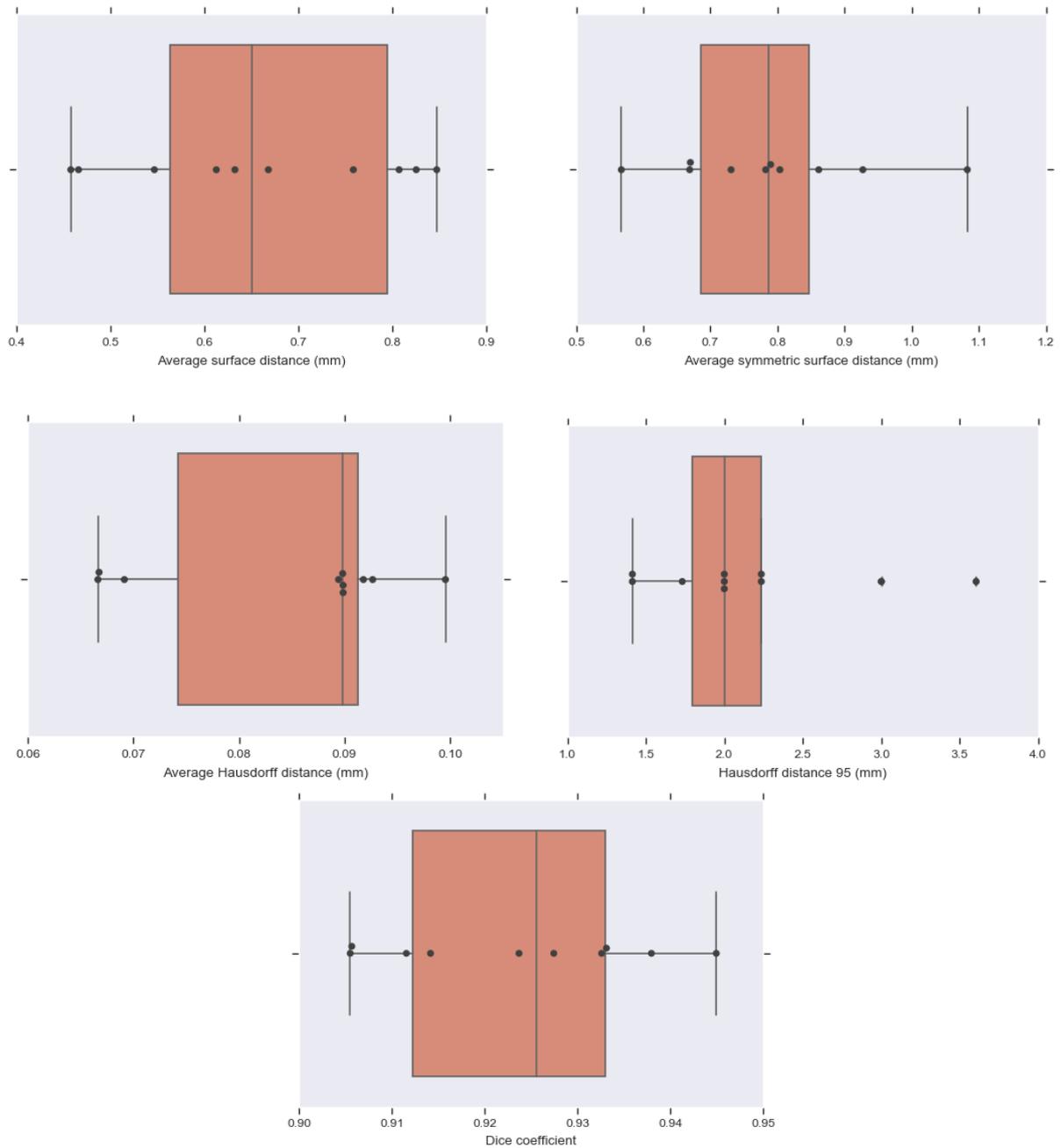


Figure 4.1: Spine segmentation evaluation metrics using the region-based flow with shrink-wrapping for Database 1. The minimum, median, inter-quartile range, and the maximum values can be clearly seen. The values beyond those can be interpreted as outliers.

See Figure 4.2 for an example of the region-based segmentation procedure. The initial seeds are manually placed by the user at the intersection of the spinous processes and the vertebral body. Although a single seed should suffice, in order to speed up the segmentation process, we place multiple seeds across the spine. Next, the surface of the spine is extracted using the region-based flow on the flux-map. Figure 4.3 shows the surface extraction process in 3D, where the initial seeds evolve over time to traverse the

boundary of the spine in the flux-map. The extracted surface contains some holes due to the lack of signal at the boundary, as can be also be seen in a 2D sagittal view in Figure 4.4. To fill up the interior volume of the spine and thus obtain a complete segmentation of every vertebrae, we use the shrink-wrapping method by placing a bounding-box around the extracted surface. This step also helps in closing of the holes present on the boundary. The complete segmentation for the lumbar region is shown in Figures 4.2 and 4.3 in 3D, while Figures 4.4 through 4.6 show 2D slices. Further, an overlay of the segmented volume on the groundtruth is shown in Figure 4.7. In case where there are large holes present on the extracted surface, the GAC might leak into the vertebral bodies upon shrink-wrapping. Therefore, the user can directly place seeds on the edge-enhanced image to get filled volumes of the spine. The shrink-wrapping step then follows to fill any holes inside the vertebral body (which are present due to inhomogeneous intensity distribution).

The major advantage is that the flow traverses on the vertebral boundary providing a very accurate localization of the surface. The method also helps in delineating the spinal cord and the intervertebral discs. Further, the method does not require any training data and is simple to implement with a clear mathematical underlying. The method is easily reproducible as these geometric flows are implemented without heuristics. A positive offshoot of this method is that it can successfully segment any vertebral region of the spine or alternatively any visible spinal region in a given volume. So, if one specifically wants to segment a small localized region for spinal surgical planning, this can be achieved by cropping the desired region of interest and then running the region-based flow on the edge-enhanced flux-map. This is of significant clinical relevance, as this method can be used to obtain a 3D model of the spine. These 3D models can then be used for downstream tasks.

It is surprising to discover that there are very few papers which have employed level-set based segmentation for CT spine segmentation. In particular, [Lim et al., 2013] use an edge-mounted Willmore flow with kernelized shape-prior based level-set method for lumbar vertebrae segmentation and [Tan et al., 2008] use a series of GAC based level set method to segment the vertebral body. [Lim et al., 2013] show for 2D slices that GAC

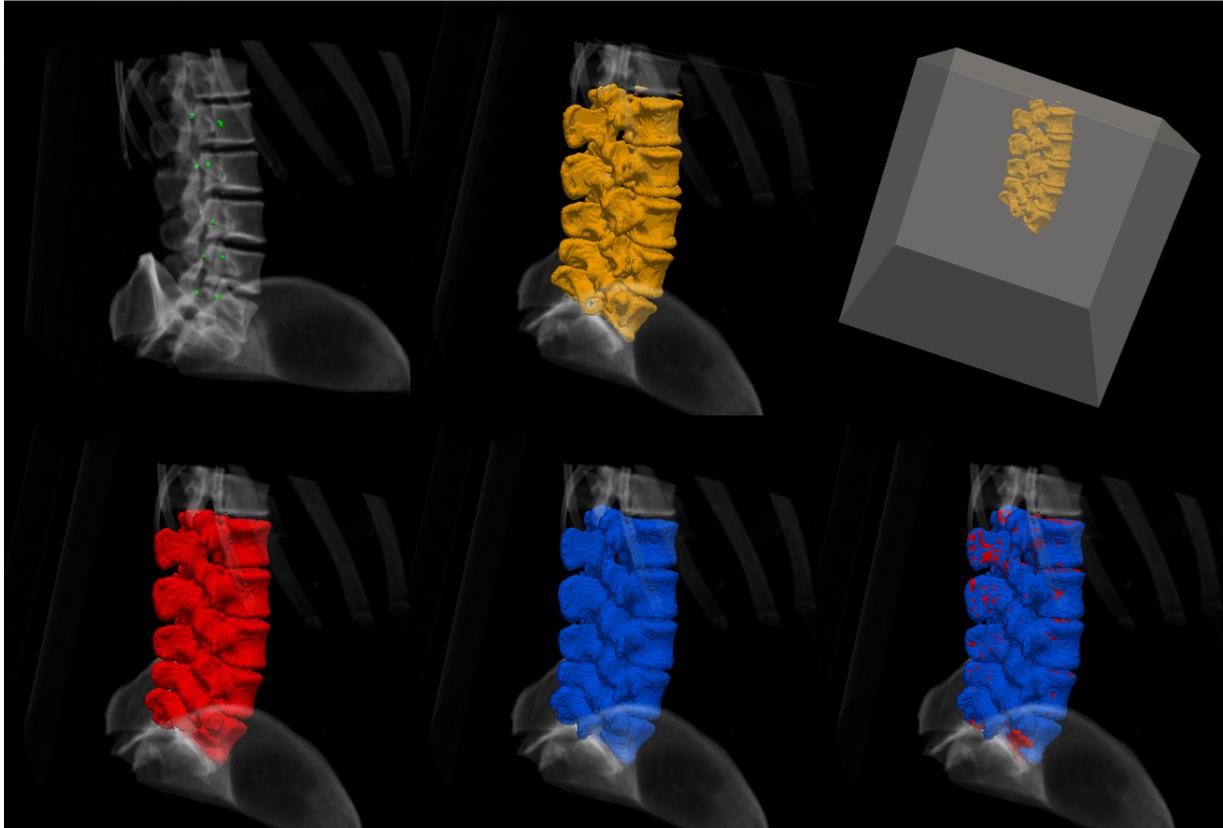


Figure 4.2: [Top row: left to right] The first image depicts the initialization by placing voxels. The second image shows the extracted surface using the region-based flow, which has cavities inside since the segmentation is concentrated on boundaries of the vertebrae. The third image shows a bounding-box placed around the extracted surface, which initializes the process of shrink-wrapping around the middle segmentation. [Bottom row: left to right] The leftmost image is the shrink-wrapped spine segmentation, which now contains a vertebral volume. The middle image shows a groundtruth image obtained by manual segmentation. The rightmost image shows an overlay of the obtained segmentation on the groundtruth.

and Chan-Vese based methods fail to segment the CT vertebrae. This is in direct contrast to our proposed method. The key reason why the GAC and Chan-Vese based formulation does not work for them is because they do not use appropriate filtering methods as used in our study. Furthermore, they use a doughnut shaped initialization, which takes a substantial amount of time to manually initialize, whereas we use a much simpler one-click (one voxel suffices) initialization. Our method outperforms these approaches, which report a Dice score of $89.32 \pm 1.70\%$ with a Hausdorff distance of 14.03 ± 1.40 mm for the lumbar region³. Furthermore, these methods are not guaranteed to perform well on cases with trauma because the shape prior used does not model shape variations of the

³The dataset used in [Lim et al., 2013] is not publicly available but is quite similar to the Database 1 used in our work.



Figure 4.3: [Left to right and top to bottom] Surface extraction process for Database 1, using the region-based flow on the flux-map, as explained in Section 3.2. The entire evolution procedure can be seen from initialization to the final extracted surface.

population but only accounts for each vertebrae in the training set individually.

[Lessmann et al., 2019] proposed a deep learning based method which iteratively segments the different vertebrae using a *memory network*, one after the other, keeping the infor-

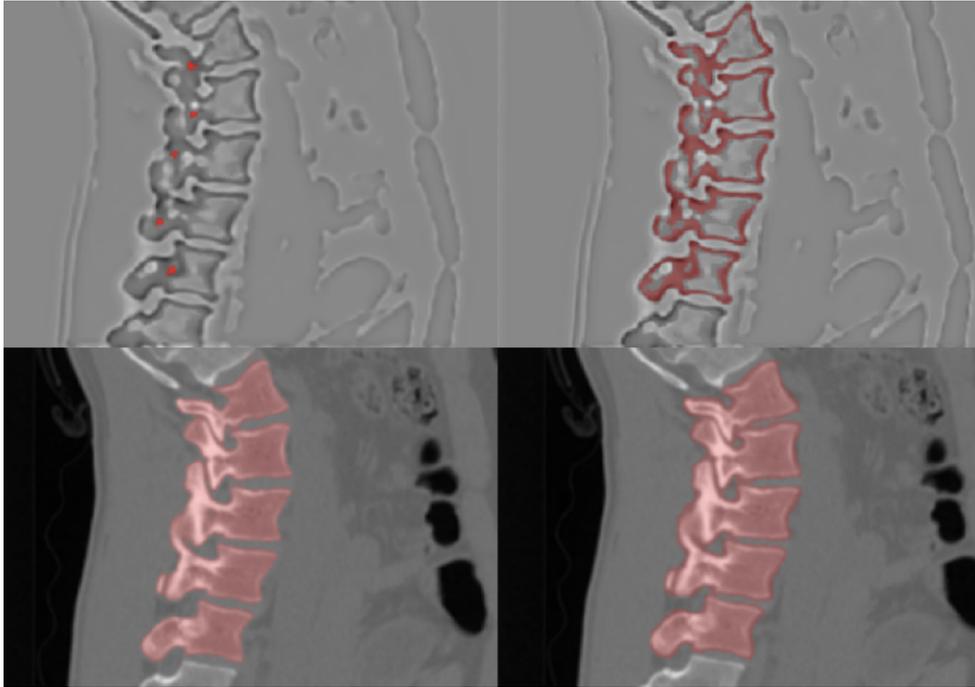


Figure 4.4: Sagittal slices: The top-left image shows the initial seeds placed on the edge-enhanced flux-map for an example from Database 1. The top-right image shows the extracted surface (vertebral boundary). The bottom left image shows the shrink-wrapped segmented spine. Notice how the cavities visible in the top-right image are now filled, and the surface holes are closed. The bottom-right image depicts the groundtruth.

mation from the already segmented vertebrae. This involves keeping a track of nearby segmented vertebrae, before segmenting the next vertebra. In contrast, we do not require any such information and segment the full spine or a desired individual vertebra. Thus, their method does not process the entire volume at once, unlike ours. Our method’s performance is comparable to theirs but it does not outperform them in terms of Dice coefficient. However, our ASSD of 0.7884 mm is well within the acceptable surface error of 2mm for facet joint injection procedures [Gill et al., 2012]. In their formulation, the images were cropped to the region which contains the vertebrae to be segmented to restrict the field of view. In contrast, we do not require any such intervention.

Comparing our method to shape and landmark based methods, our approach does not require any complex shape-modelling and the use of constrained optimization methods as proposed in [Rasoulia et al., 2013], [Ibragimov et al., 2014], [Korez et al., 2015], [Castro-Mateos et al., 2015] and [Pereañez et al., 2015]. These methods require lots of manually selected landmark points which, are cumbersome to obtain. Secondly, the op-

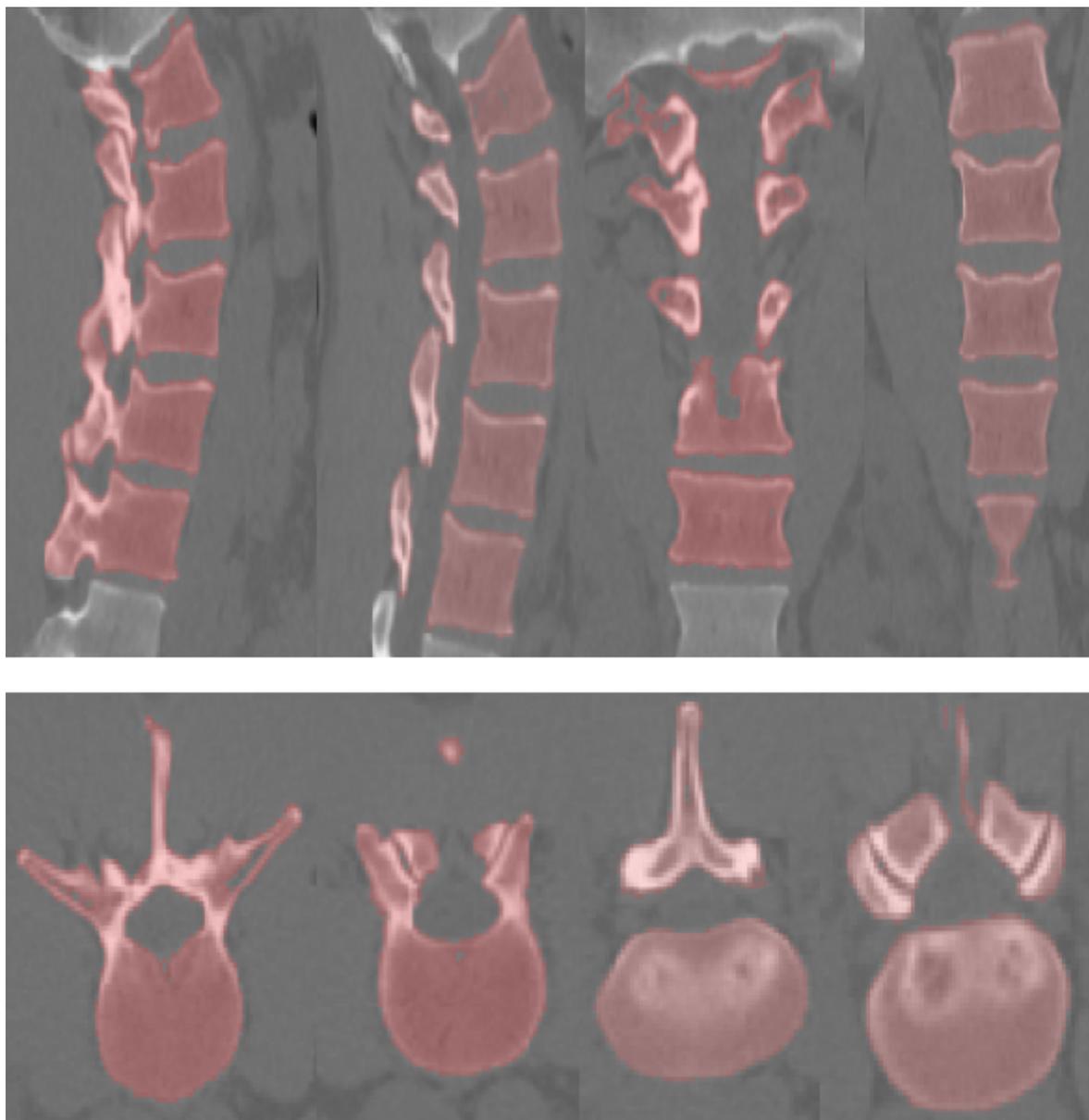


Figure 4.5: Sample sagittal, coronal and axial segmented slices for Database 1.

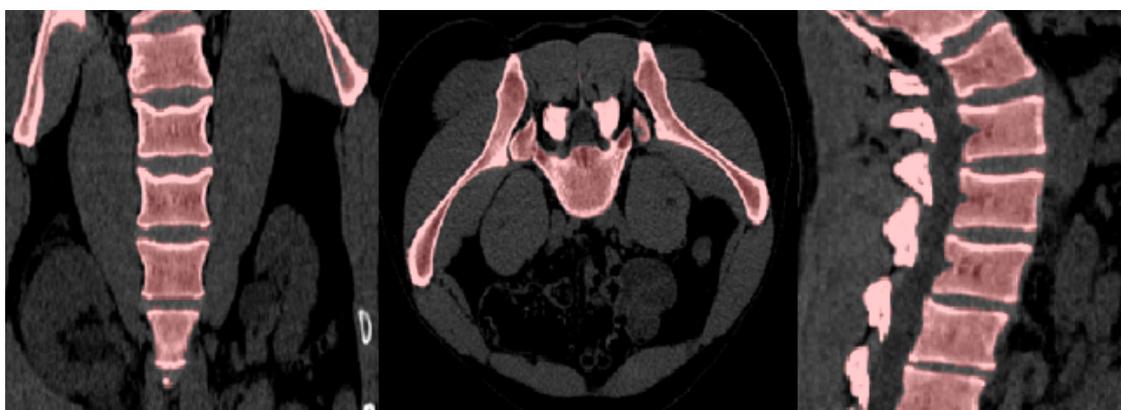


Figure 4.6: Sample sagittal, coronal and axial segmented slices for the region not only limited to the available groundtruth for Database 1. Here, sacrum and thoracic (T12) is shown.

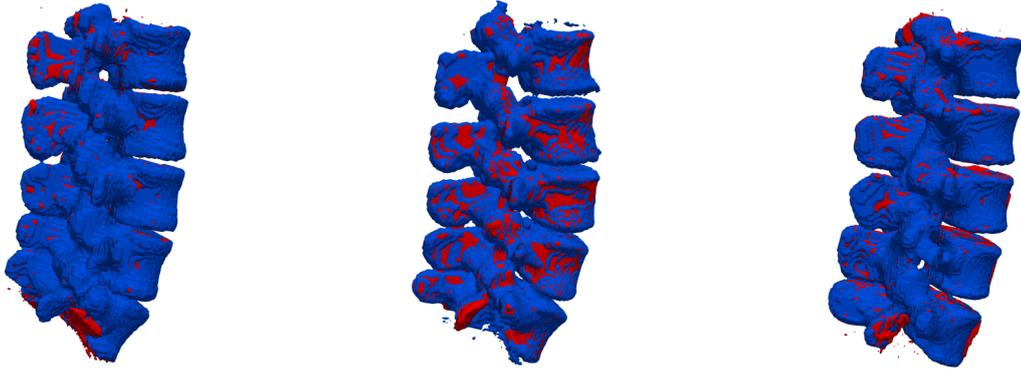


Figure 4.7: Three examples of the segmented spine (red) overlaid on the groundtruth (blue) for Database 1.

timization methods are local and do not guarantee convergence. [Rasoulia et al., 2013] makes a restrictive assumption that there is a strong correlation between the shapes and poses of different vertebrae, within the same patient. We do not make any such assumptions in our region-based surface extraction approach and we outperform⁴ them on the comparable metric of HD95 with 2.1638 ± 0.6464 mm vs 3.80 ± 1.61 mm on Database 1.

The Statistical Interspace Models (SIMs) method proposed by [Castro-Mateos et al., 2015] models the interspace between different vertebrae, to avoid overlap between them. Our region-based method does not require any explicit account for this as it segments each vertebrae exactly on its surface, via a much simpler procedure. [Korez et al., 2015] proposed a landmark-based segmentation method based on transportation theory and [Ibragimov et al., 2014] uses interpolation theory. We report very comparable segmentation metrics in terms of Dice coefficients and Hausdorff distances compared to these three approaches. [Pereañez et al., 2015] proposed to segment the three parts of the vertebrae, that is, the vertebral body and the two processes individually. Again, we do not require any such conditional decomposition.

Table 4.1 provides a comparison of different metrics between our method and other approaches. The proposed segmentation procedure is very fast even on a CPU. The edge-enhancing pre-processing step takes around 4 minutes and another 15 seconds to compute the flux-map. These pre-processing steps are done offline. The surface extraction

⁴The dataset used in [Rasoulia et al., 2013] is not publicly available but is quite similar to the Database 1 used in our work.

and shrink-wrapping procedure takes around 3 minutes and a minute respectively on an Intel(R) Xeon(R) CPU at 3.50GHz with 12 cores for the entire lumbar region in 3D.

Table 4.1: Spine segmentation results reported for Database 1

Paper	Dice Coefficient	Avg. HD(mm)	HD95(mm)	ASSD(mm)	ASD(mm)
[Lessmann et al., 2019]	$96.50 \pm 0.8\%$	-	-	0.2 ± 0.0	-
[Rasoulian et al., 2013]	-	-	$3.8 \pm 1.6\%$	-	-
[Ibragimov et al., 2014]	$93.70 \pm 1.12\%$	-	-	-	0.75 ± 0.13
[Korez et al., 2015]	$95.30 \pm 1.4\%$	-	-	-	-
[Chu et al., 2015a]	$91.00 \pm 7.0\%$	-	-	-	0.9 ± 0.3
Ours	$92.36 \pm 0.0131\%$	0.1385 ± 0.0845	2.1638 ± 0.6464	0.7884 ± 0.1385	0.6626 ± 0.1371

We now discuss the results on the CSI challenge dataset (Database 2). The thoracic and lumbar vertebrae are segmented separately. We divide the thoracic region into three groups, i.e., upper, middle and lower thoracic regions. Instead of normalizing the images using the 5th and 95th percentile threshold as described earlier, the thoracic region in the CSI database have been normalized using an appropriate selection for the lower and upper threshold values. The intensities are then scaled between 0 and 1. This also has the advantage that we obtain a better contrast between the vertebral column and the rest of the image volume.

We did not obtain good segmentations for the thoracic vertebrae if the region-based flow is used on the flux-map. We observed the presence of several holes on the vertebral boundary if the region-based flow is used on the flux-map. This is due the lack of signal at the boundaries, which leads to gaps in the final extracted surface. As a consequence, the shrink-wrapping process which uses GAC leaks through the larger holes. Additionally, the GAC based shrink-wrapping method gets stuck inside the intervertebral discs (IVD) due to the fact that the IVDs in the thoracic region are thinner than those in the lumbar region. Therefore, for the thoracic region, we applied our region-based flow directly on the edge-enhanced image. This leads to accurate segmentation of the vertebrae and their processes with a few holes present inside the vertebral bodies due to the fact that there are certain piece-wise inhomogeneities present, which have earlier been substantially reduced because of the pre-processing step. After this, we run the shrink-wrapping step to get complete vertebral segmentations, filling small holes in the vertebral bodies, if any.

Figure 4.8 shows the region-based segmentation evolution process on the edge-enhanced image in the upper thoracic region for a random subject. Figure 4.9 depicts the entire spine overlaid on the groundtruth volume. Notice that in the thoracic region, the ribs get segmented too as they are attached to the vertebrae. The advantage of our method is that one can also segment the coccyx and the sacrum bones. The middle figure shows the zoomed in thoracic region. Using ITK-SNAP [Yushkevich et al., 2006], we manually crop out the ribs, the coccyx and the sacrum bones for the purpose of computing the evaluation metrics. Finally, the rightmost figure shows the thoracic and the lumbar regions after cropping out the extraneous regions, overlaid on the groundtruth. Figure 4.10 shows the boundary of the groundtruth and the segmented volume overlaid on top of a 2D slice.

Table 4.2 reports the metrics averaged over the entire thoracic and lumbar regions for the CSI database. We report an average Dice score of $95.97 \pm 0.0159\%$ which is very close to the current state of the art Dice score of $96.30 \pm 1.3\%$. Our method performs better than the non-deep learning methods both in terms of volume overlap and surface distances.

Table 4.2: Spine segmentation results reported for Database 2

Paper	Dice Coefficient	Avg. HD(mm)	HD95(mm)	ASSD(mm)	ASD(mm)
[Lessmann et al., 2019]	$96.30 \pm 1.3\%$	-	-	0.1 ± 0.1	-
[Hammernik et al., 2015a]	$93.00 \pm 0.04\%$	-	-	-	0.47 ± 0.54
[Korez et al., 2015]	$94.40 \pm 2.1\%$	-	-	0.3 ± 0.1	-
Ours	$95.97 \pm 0.0159\%$	0.04522 ± 0.0599	1.7211 ± 1.0528	$0.3454 \pm 0.1506\%$	0.4176 ± 0.2383

Figures 4.11 and 4.12 and Table 4.3 report results separately for the three thoracic regions and the lumbar region. The segmentation results are quite similar for these two regions. Our method performs better than [Hammernik et al., 2015b] for the lumbar and the thoracic regions, and thus the entire spine. We acknowledge that our work is similar in nature to the work proposed by the thesis of [Hammernik, 2015] and then the follow-up CSI workshop submission [Hammernik et al., 2015b]. The authors have used the ROF model to de-noise the image anisotropically. They solve a variational energy functional in a convex optimization framework. The energy functional consists of a structure tensor based GAC term, a mean-shape model and a bone prior map.

Our method though similar in nature, uses level-set framework. [Hammernik et al., 2015b]

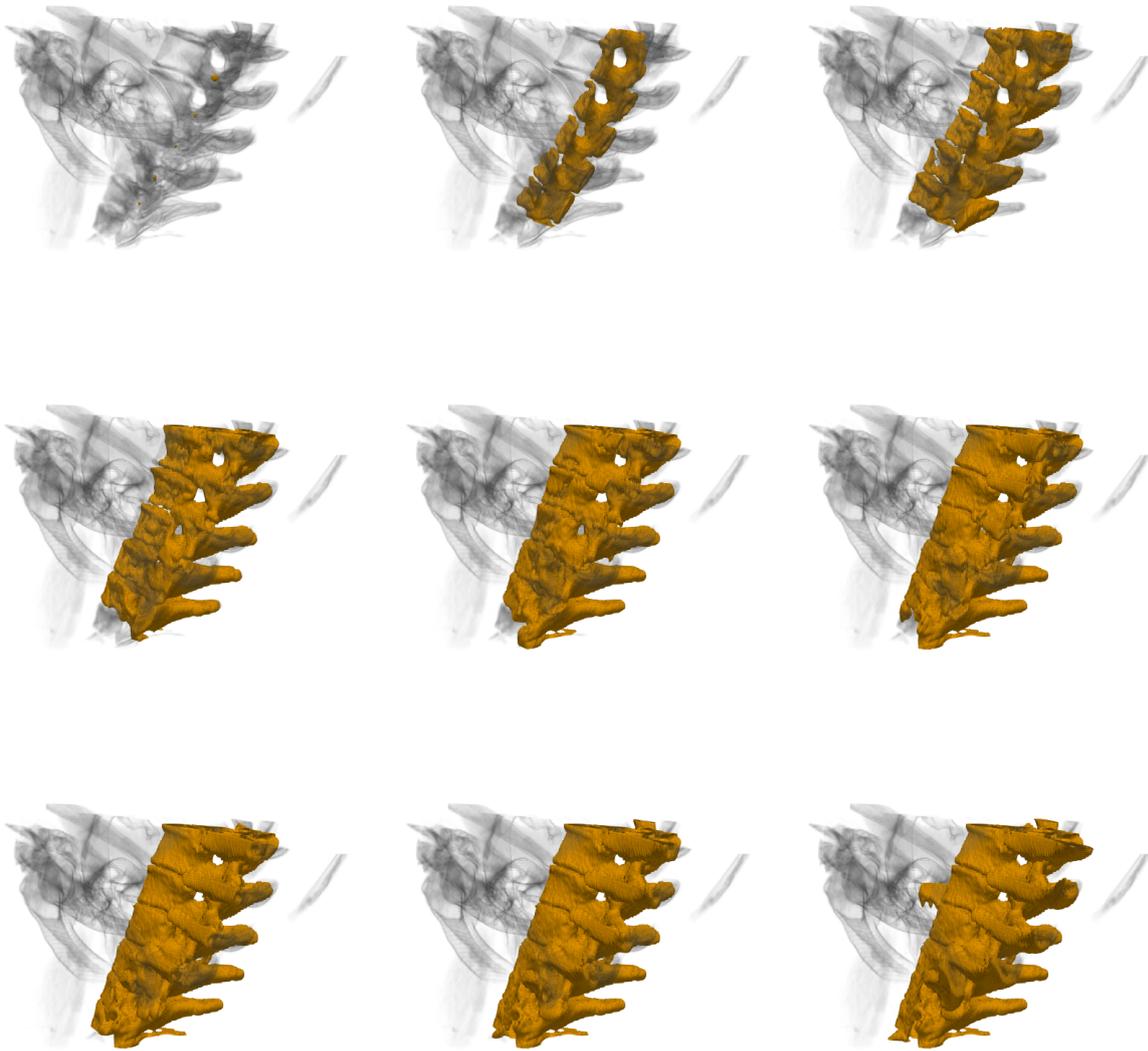


Figure 4.8: [Left to right and top to bottom] Surface extraction process for Database 2, using the region-based flow on the edge-enhanced image, as explained in Section 3.2. The entire evolution procedure can be seen from initialization to the final segmented upper thoracic vertebrae.

Table 4.3: Spine segmentation results reported for Database 2: [Thoracic vs Lumbar Vertebrae comparison]

Vertebral Region	Dice Coefficient	Avg. HD(mm)	HD95(mm)	ASSD(mm)	ASD(mm)
Thoracic	$95.62 \pm 0.0179\%$	0.0525 ± 0.0721	1.7071 ± 1.1202	0.3285 ± 0.1534	0.4382 ± 0.2691
Lumbar	$96.56 \pm 0.0094\%$	0.0328 ± 0.0247	1.7448 ± 0.9276	0.3739 ± 0.1413	0.3827 ± 0.1687

use GPUs to implement the computationally expensive convex optimization step. In contrast, our method is not limited by hardware constraints. Additionally, the ribs might be segmented in their method as well, which the authors propose to remove by providing foreground and background user scribbles. Their method require a groundtruth to build

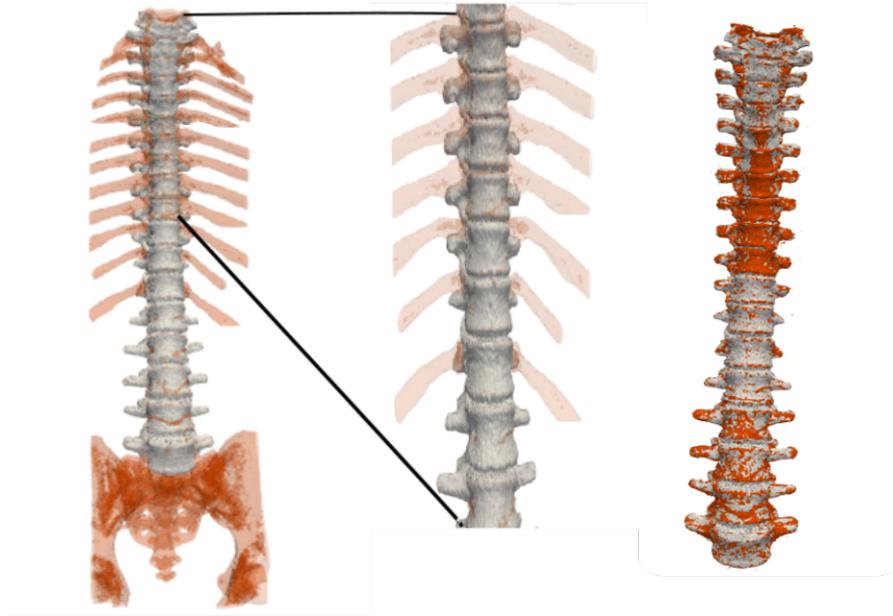


Figure 4.9: Left: Shown in light grey is the vertebrae region that is common to both, the groundtruth and the segmentation obtained by our method. The extra regions segmented such as the ribs, the coccyx and the sacrum, and other voxels in the spinal column are shown in red. Middle: A zoomed-in image of the upper thoracic region. Right: The segmented spine (shown in red) overlaid on top of the groundtruth (shown in light grey) after cropping-out the regions manually using ITK-SNAP.

the shape model. They construct the mean-shape model using a leave one out cross validation scheme, i.e., they use nine vertebrae to construct each a shape model and segment one of the held-out subject. In contrast, we do not require any training data. They report that the IVDs are incorrectly segmented in the upper thoracic regions due to thinner spacing between the adjacent vertebrae. We do not see any such mis-segmentation in our experiments. In our method, we do require some manual cropping of the extraneous segmentations such as the ribs. We argue that some manual intervention is acceptable since the accurate segmentations outweigh the little time required to crop out the ribs. Whereas it is an arduous task to segment a vertebrae manually, removing the ribs, the coccyx or the sacrum requires very little effort. Alternatively, these regions could be prevented from being completely segmented by limiting the region of interest to the vertebral column prior to the initialization of the segmentation procedure.

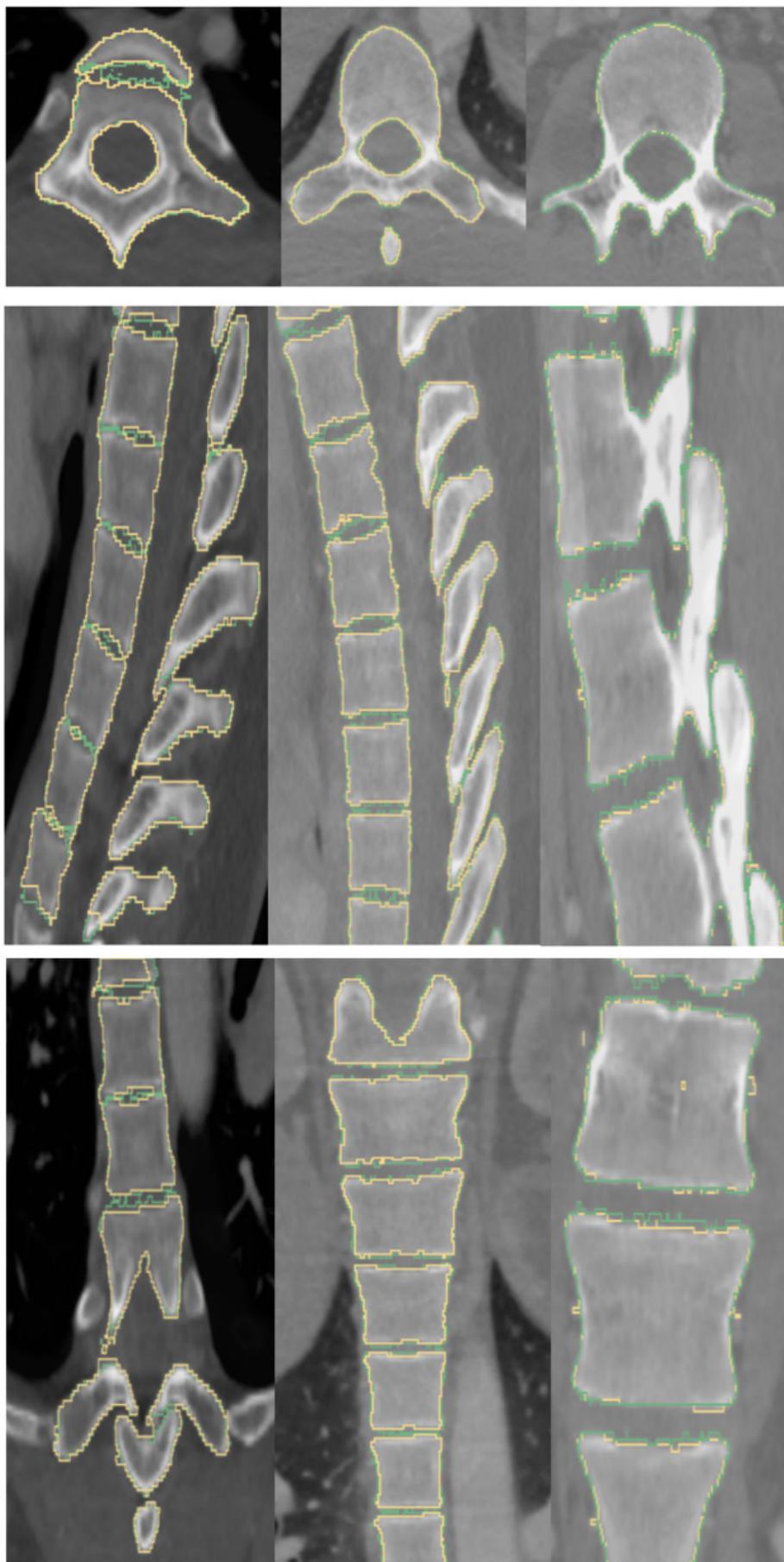


Figure 4.10: The boundary of the segmentation (yellow) and groundtruth (green) overlaid on top of the image for a subject from Database 2 (CSI). *From left to right*: parts of upper thoracic, middle thoracic and lumbar region. [*From top to bottom*] axial, sagittal and coronal views.

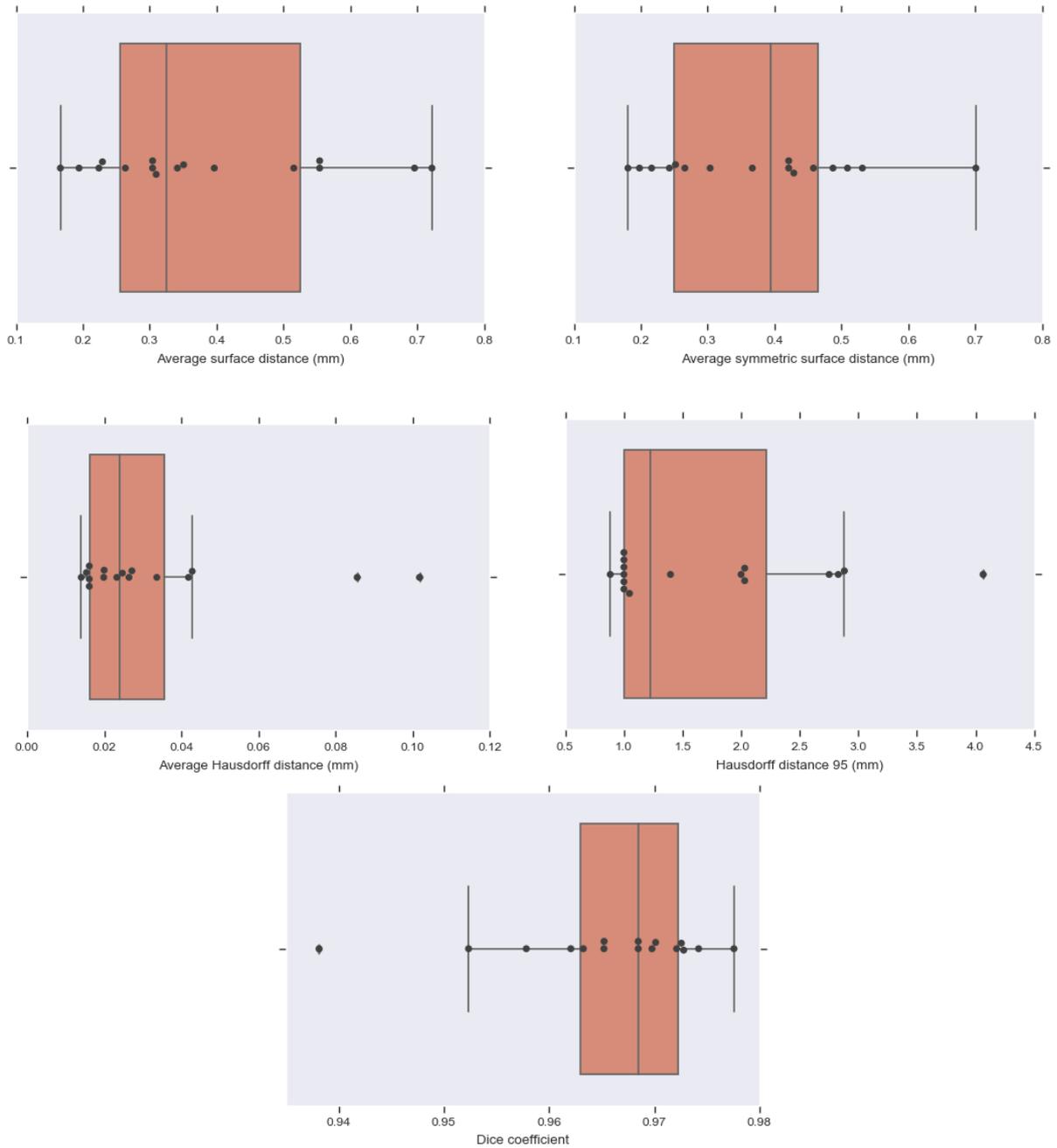


Figure 4.11: Spine segmentation evaluation metrics using the region-based flow on the **lumbar region** with shrink-wrapping for Database 2. The minimum, median, inter-quartile range, and the maximum values can be clearly seen. The values beyond those can be interpreted as outliers.

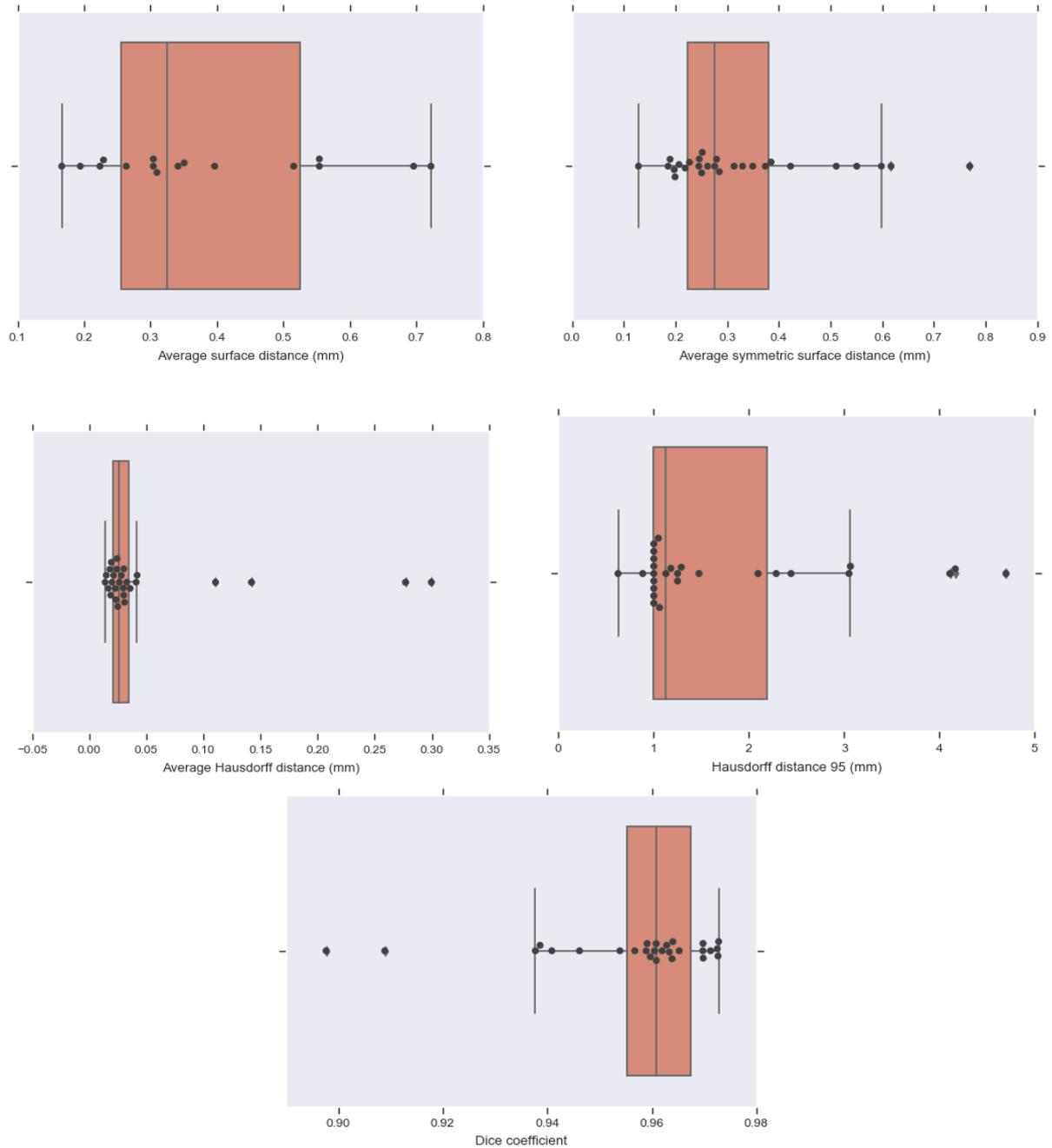


Figure 4.12: Spine segmentation evaluation metrics using the region-based flow on the **thoracic region** with shrink-wrapping for Database 2. The minimum, median, inter-quartile range, and the maximum values can be clearly seen. The values beyond those can be interpreted as outliers.

4.2 Shape-prior based vertebrae segmentation: discussion

We evaluate the shape-prior based segmentation method as discussed in Section 3.3 on Database 3. To motivate the need to use a shape-prior, we first present some results of the region-based segmentation pipeline on this challenging database of patients with trauma. A 2D slice obtained by the region-based surface extraction method is shown in Figure 4.13 and the filled volume is shown in Figure 4.14. The segmentations are very accurate, as the flow evolves along the vertebral boundary, but it is hard to demarcate each individual vertebra separately because several vertebrae are either fused, fractured or dislocated. A shape-prior based segmentation method applied to each individual vertebra can help eliminate this problem, though the shape-prior based formulation has its own disadvantages, which we shall discuss in Chapter 5.

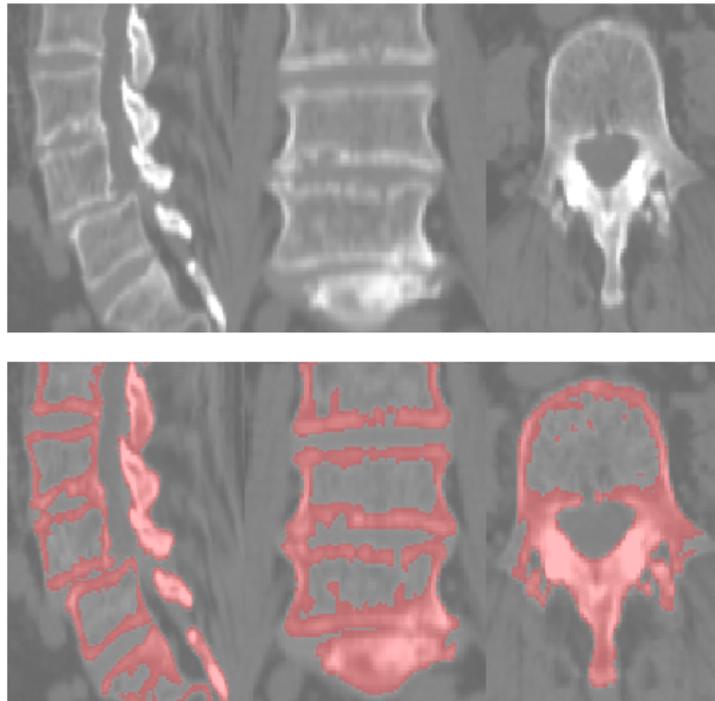


Figure 4.13: Sample slices shown for the region-based surface extraction on the flux-map for Database 3.

We use a six-fold cross-validation approach. We build the shape prior with 25 vertebrae and evaluate on the remaining five. This procedure is then repeated for the different

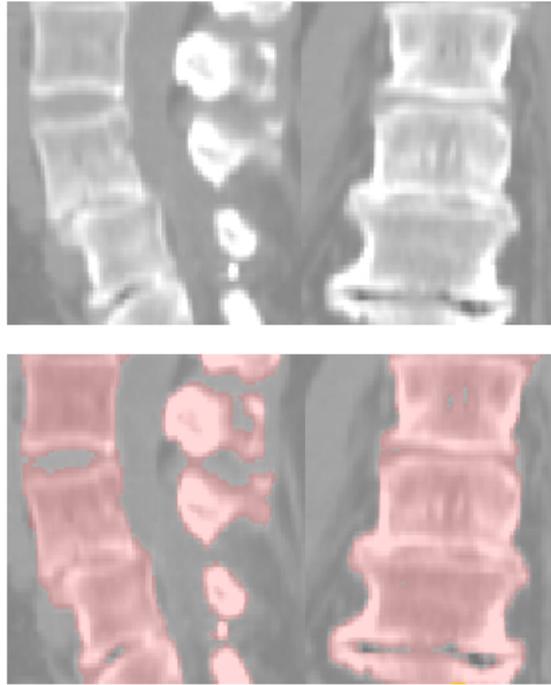


Figure 4.14: Sample slices shown for the shrink-wrapping process for Database 3.

sets. A total of seven parameters are optimized by the level-set formulation: three for transformation, three for rotation, and one for scale. We report an average Dice score of $84.00 \pm 0.0310\%$. We report an average symmetric surface distance at 1.3149 ± 0.2421 mm. The average Hausdorff and its 95th percentile are at 0.2865 ± 0.09574 mm and 3.7244 ± 1.4368 mm respectively. The average surface distance is 1.4605 ± 0.2662 mm. Figure 4.15 gives a box-plot for every metric used to evaluate the segmentation results.

Figure 4.16 shows some slices in 2D for the shape-prior based segmentation obtained. Observe that the boundary demarcated by this shape-prior formulation is not precise but still the ASSD is well within the clinically accepted surface distance of 2 mm. Figure 4.17 depicts the segmentation overlaid over the groundtruth. Figure 4.18 depicts the shape prior segmentation procedure. The algorithm is initialized with the mean shape of the training population placed near the region of interest.

This PCA based shape-prior level set method is not only fast, but also requires, very little effort to set up. The PCA modelling step takes around a minute to complete for a set of 25 vertebrae. This step is done offline, before the actual segmentation step. It takes a couple of minutes to segment an individual vertebra, which is the online step. A

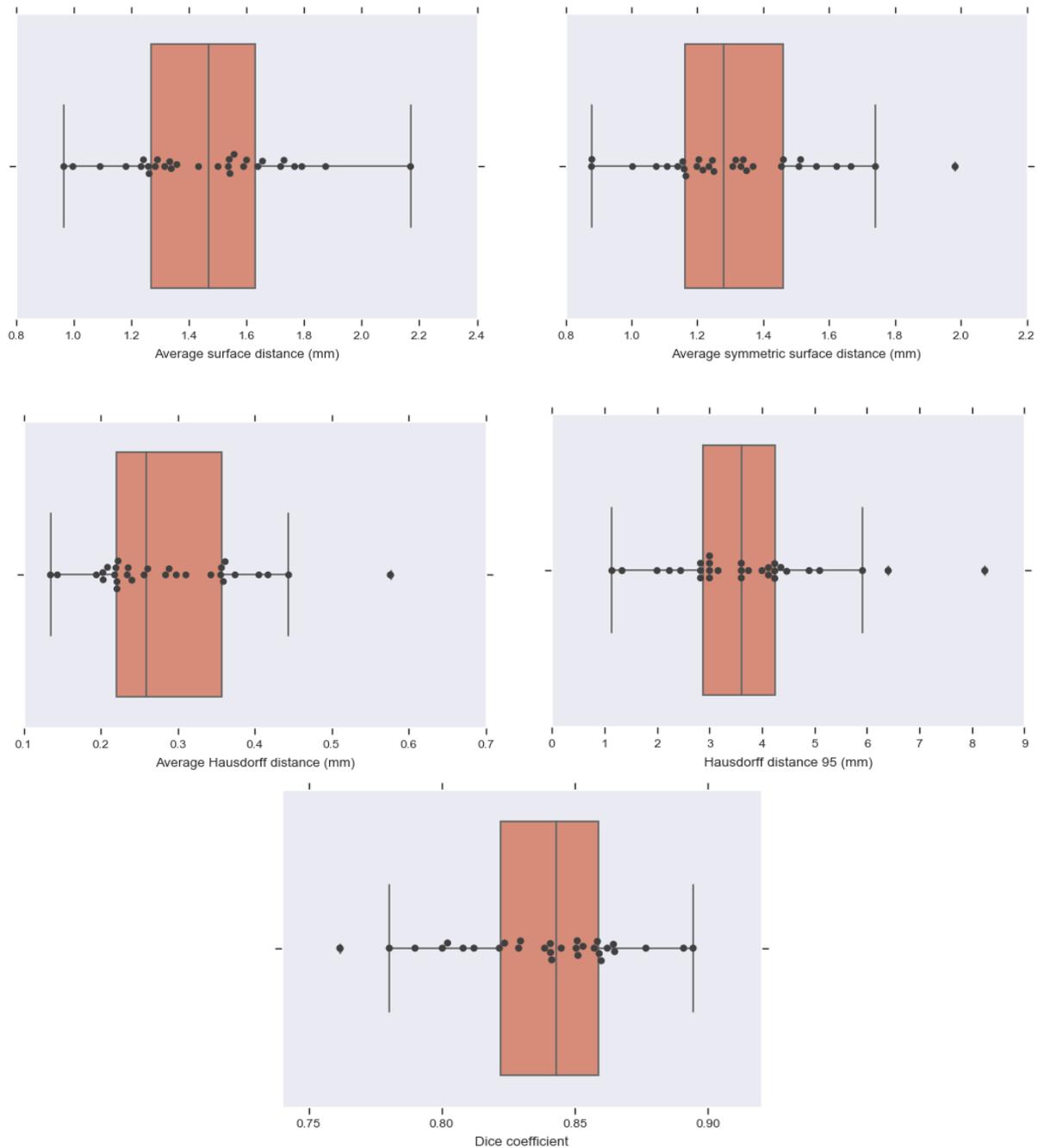


Figure 4.15: Shape-based segmentation evaluation metrics for Database 3. The minimum, median, inter-quartile range, and the maximum values can be clearly seen. The values beyond those can be interpreted as outliers.

shape-prior can also be built for all the vertebrae in the lumbar region together.

The active shape and appearance model based method used in [Stephansen, 2012] requires a cumbersome procedure of obtaining manual landmarks. We outperform this method using the proposed shape-prior method both in terms of Dice score and surface distance errors. See Table 4.4 for a comparison between our method and [Stephansen, 2012].

Table 4.4: Spine segmentation results reported for Database 3

Paper	Dice Coefficient	Avg. HD(mm)	HD95(mm)	ASSD(mm)	ASD(mm)
[Stephansen, 2012]	80.86% (median)	-	-	-	-
Ours	$84.00 \pm 0.0310\%$	0.2865 ± 0.09574	3.7244 ± 1.4368	1.3149 ± 0.2421	1.4605 ± 0.2662

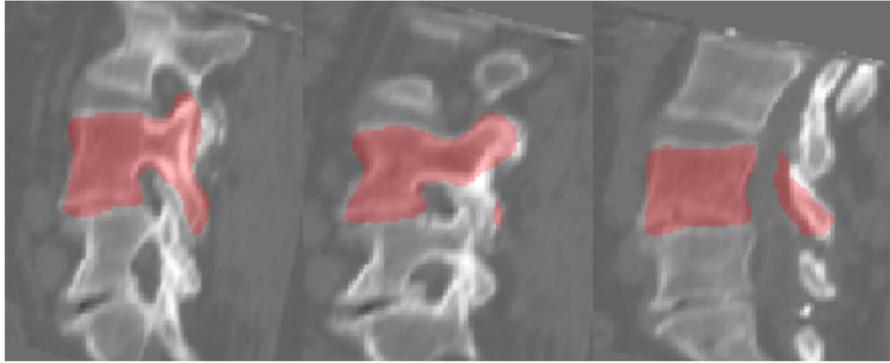


Figure 4.16: Sample slices shown for the shape-prior based segmentation procedure for a subject from Database 3.

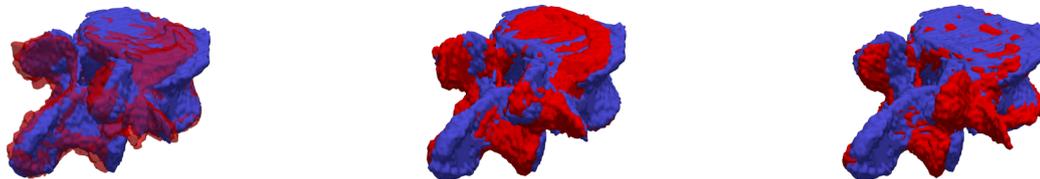


Figure 4.17: Three examples of segmentation (red) overlaid on the groundtruth (blue), for the shape prior based segmentation of L4 vertebra in Database 3.

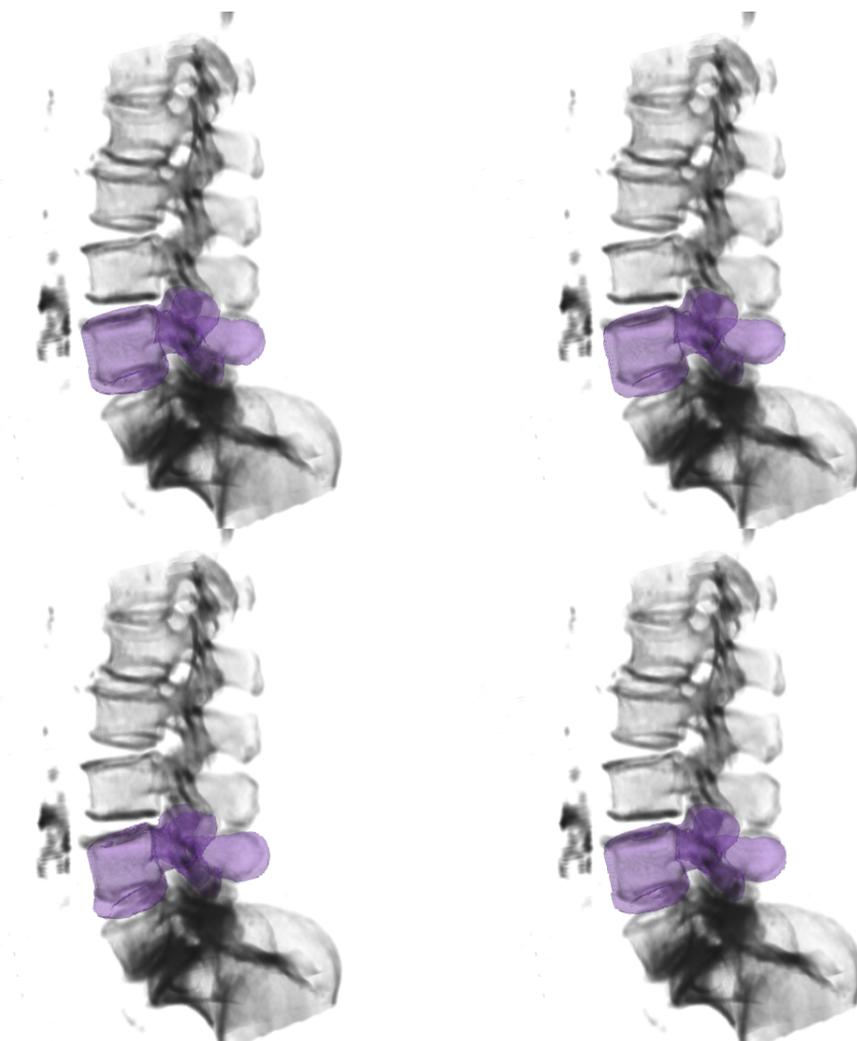


Figure 4.18: [Left to right and top to bottom] Shape-prior based segmentation process for a subject from Database 3. The entire evolution procedure is shown from the mean-shape initialization, to the final segmented volume.

Chapter 5

Conclusion

In this chapter, we comment on the major advantages and some limitations of our work. We then conclude this thesis with a brief discussion of possible extensions to this work.

5.1 Concluding remarks

In this thesis, we have proposed a two-phase segmentation pipeline which employs appropriate image analysis methods. To the best of our knowledge, the use of anisotropic diffusion filtering combined with flux maximizing flows for spine segmentation is new. In general, the previous studies have employed Gaussian filtering, which displaces the actual vertebral boundary locations. This is highly undesirable as it might lead to imprecise localization of the vertebral boundaries. We have overcome this issue by using edge-enhancing flows with the introduction of the edge-enhanced flux-map as discussed in Chapter 3. The proposed pre-processing and segmentation pipeline could be used for CT images of other organs. As an example, see Figure 5.1 for a segmentation of the CT femur bone. Qualitatively, edge-enhanced flux maps obey the piece-wise constant assumption that underlies the classic Chan-Vese binary region-based level-set segmentation. This makes the Chan-Vese bi-partition flow work well for the vertebral segmentation.

We have experimentally validated the performance of our region-based surface extraction

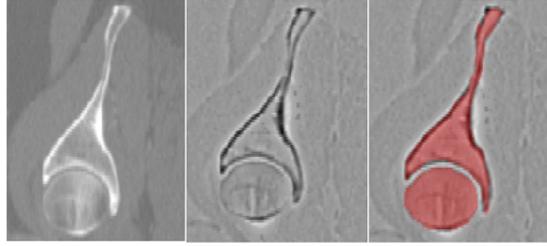


Figure 5.1: Leftmost image: sagittal slice of femur in CT image. The flux-map is shown in the middle, and the right most image shows the segmented region using the region-based surface extraction and shrink-wrapping process.

approach and the subsequent shrink-wrapping step on two publicly available databases. The shape-prior based method was evaluated on a database of spine cases with trauma. A detailed comparison via performance measures between our approach and previous methods in the literature was presented in Chapter 4. We do not require tracking or segmentation of nearby structures in the spinal column to segment the vertebrae. This again is a major advantage of the proposed method as it saves additional computational burden.

The recent data-driven deep learning methods rely heavily on large image databases. Our approach does not require any training images for the region-based spine segmentation and it still produces results close to the current state of the art. The other methods rely on training data and can thus, only segment the vertebrae for which the groundtruth segmentations are available. Deep learning based methods typically require days to train. [Lessmann et al., 2019] report that they trained their network for 4-5 days on Nvidia Titan X GPUs with 12 GB memory. We do not require any training and the entire process from pre-processing to shrink-wrapping takes around 9 minutes per lumbar region on a CPU. This, we believe is one of the key contributions of the surface extraction based method.

Our method works across different data distributions because it exploits information directly from the images based on concrete mathematical image analysis. It does not suffer from the lack of generalization, as is common with the deep learning methods. The major disadvantage of deep learning methods is that they are not interpretable. A clinician working with computer algorithms would want to know how the algorithm makes its decision,

but unfortunately this is not the case with the deep learning methods. Additionally, the deep networks do not handle the voxel spacing appropriately, and thus generally make the volume isotropic 1 mm^3 for tasks such as segmentation. This might lead to inappropriate segmentations.

As discussed previously, [Hammernik et al., 2015b] use similar energy functionals as proposed in our work. Their variational formulation is not implemented in a level-set framework unlike ours. Again, our method does not require any training data because we do not use any shape-prior for the region-based segmentation. Furthermore, they initialize the method using a mean-shape of the vertebrae. We eliminate the need for this by having a one-click initialization. The major advantage of their method is that the energy functional is convex and has a non-unique global optimal solution. However, the shape-prior based method proposed in this thesis might get stuck in a local minimum which is the major drawback of the level-set based method. We have not encountered any mis-segmentations, but we rely on manual-cropping of extraneous regions for the region-based segmentation method. Interestingly, we use shape variations (due to pose estimations) and are not limited to the mean-shape only. This way the proposed method in our work is better able to capture a wide variability in shapes.

In the second part of Chapter 3, we proposed a shape-prior based segmentation method. The method is quite fast and helps in the segmentation of the individual vertebrae having undergone trauma. A critique of the shape-prior based level-set method used in this thesis, is that these methods suffer from convergence issues. The algorithm is a local optimization method, which is sensitive to initialization. During our evaluations, we noticed that the method can sometimes get stuck in a local minimum, and as such, might provide inaccurate segmentation results. Furthermore, we have not used edge-based information in the shape-prior formulation. An edge-based term might help to better localize the vertebra. The segmented vertebrae using these methods can now be used for further downstream tasks such as registration with intraoperative ultrasound scans, and thus, aid neurosurgeons in planning spinal surgery. We have shown that with no training data (or little for shape-prior based segmentation), we achieve good vertebral

segmentations. It would be fruitful to see if this pipeline could be used in the preoperative surgical planning procedure as demonstrated by [Goulet, 2010].

5.2 Future Work

There is also some scope to improve the shrink-wrapping method used in this thesis. Instead of the direct application of geodesic active contours, the problem can be viewed as that of surface reconstruction. This would better help in filling-up the holes to obtain smoother vertebral surface. We are currently investigating a suitable method to use from the existing literature. Separately, we plan to develop a new shape-prior based level-set variational energy functional which incorporates boundary and region terms based on a set of features. This would include a structure tensor based stopping potential for the boundary term and a multivariate kernel density estimation to incorporate several features, leading to higher region based discriminatory power. The shape based term would again be determined by fitting a kernel. The last term would incorporate a feature vector to topologically adapt the evolving contour toward a desirable segmentation and thus, would help to preserve the complex topology of the vertebrae. We would try to eliminate the need for mean-shape initialization and, rather have a one-click initialization with a few voxels. It might be worthwhile to evaluate this new formulation on the publicly available dataset, xVert¹, which is a much more challenging database because the subjects have suffered spinal trauma. This dataset consists of 10 subjects with groundtruth reference segmentations for the lumbar vertebrae, thus comprising a total of 50 vertebrae.

As discussed above, the shape-prior based segmentation suffers from convergence guarantees as the solution obtained is often a local minimum. Thus, it would be beneficial to have sound mathematical properties for the energy functional which might then guarantee a globally optimal solution.

¹Database can obtained from: <http://lit.fe.uni-lj.si/xVertSeg>

Appendix A

Numerical Methods for Level-sets

We briefly review numerical methods to implement level-set methods for image segmentation. After the front is initialized and the embedding function ϕ has been computed, a set of schemes are used to evolve the front iteratively. Remember that the continuous formulation is now discretized. An excellent in-depth discussion can be found in the text by [Sethian, 1999].

A.1 Solution to the Hamilton-Jacobi Equation and Upwind scheme

Earlier we studied the formulation of the level-set based segmentation. Here, we shall look at how this scheme is implemented numerically. In particular, we summarize the numerical schemes for Hamilton-Jacobi equations which converge to the viscosity solution. The initial value problem of level-set equation $\phi_t + F |\nabla\phi| = 0$ can be interpreted as a general Hamilton-Jacobi equation,

$$\alpha U_t + H(U_x, U_y, U_z, x, y, z) = 0, \tag{A.1}$$

where H is known as the Hamiltonian and for our initial value problem level-set we have,

$$H(U_x, U_y, U_z, x, y, z) = F\sqrt{U_x^2 + U_y^2 + U_z^2}. \quad (\text{A.2})$$

Now, we can have a solution to the above equation by computing forward and backward differences in the domain of the function ϕ ,

$$U_{i,j,k}^{n+1} = U_{i,j,k}^n - \Delta t \cdot g \left(\begin{array}{l} \frac{U_{i,j,k}^n - U_{i-1,j,k}^n}{\Delta x}, \frac{U_{i+1,j,k}^n - U_{i,j,k}^n}{\Delta x}, \\ \frac{U_{i,j,k}^n - U_{i,j-1,k}^n}{\Delta y}, \frac{U_{i,j+1,k}^n - U_{i,j,k}^n}{\Delta y}, \\ \frac{U_{i,j,k}^n - U_{i,j,k-1}^n}{\Delta z}, \frac{U_{i,j,k+1}^n - U_{i,j,k}^n}{\Delta z} \end{array} \right). \quad (\text{A.3})$$

Thus, a first order convex solution for the initial value level-set equation can be given as,

$$\phi_{ijk}^{n+1} = \phi_{ijk}^n - \Delta t [\max(F_{ijk}, 0)\nabla^+ + \min(F_{ijk}, 0)\nabla^-], \quad (\text{A.4})$$

where

$$\begin{aligned} \nabla^+ = & [\max(D_{ijk}^{-x}, 0)^2 + \min(D_{ijk}^{+x}, 0)^2 + \\ & \max(D_{ijk}^{-y}, 0)^2 + \min(D_{ijk}^{+y}, 0)^2 + \\ & \max(D_{ijk}^{-z}, 0)^2 + \min(D_{ijk}^{+z}, 0)^2]^{1/2} \end{aligned} \quad (\text{A.5})$$

$$\begin{aligned} \nabla^- = & [\max(D_{ijk}^{+x}, 0)^2 + \min(D_{ijk}^{-x}, 0)^2 + \\ & \max(D_{ijk}^{+y}, 0)^2 + \min(D_{ijk}^{-y}, 0)^2 + \\ & \max(D_{ijk}^{+z}, 0)^2 + \min(D_{ijk}^{-z}, 0)^2]^{1/2}, \end{aligned} \quad (\text{A.6})$$

where D^{+x} and D^{-x} are the forward and backward derivatives. This entire formulation of using derivatives to arrive at the solution of the given partial differential equation is known as an *upwind*. The above scheme can also be extended to higher order and can have non-convex speed functions.

A.2 Narrow band implementation and CFL condition

To speed up the computation, the computation of the upwind scheme could be limited to a certain neighbourhood around the evolving front. We call this neighbourhood as the *narrowband*. Also, the evolving curve should follow the Courant Friedrichs Lewy (CFL) condition given by

$$\max_{\Omega} F \Delta t \leq \Delta x \quad (\text{A.7})$$

to avoid instability during the updates, where the symbols have the usual meaning.

A.3 Typical front evolution update equation

Summarizing the above numerical updates for curve evolution governed by three forces namely expansion motion, mean curvature and an advection term,

$$F = F_{prop} + F_{curv} + F_{adv}, \quad (\text{A.8})$$

where $F_{prop} = F_0$ is the propagation expansion speed, $F_{curv} = -\epsilon \kappa$ is the curvature dependent force and $F_{adv} = \vec{U}(x, y, t) \cdot \vec{n}$ where the $\vec{U}(x, y, t)$ is some velocity field which corresponds to advection.

The level-set variational formulation can be written simply as,

$$\phi_t + F_0 |\nabla \phi| + \vec{U}(x, y, t) \cdot \nabla \phi = \epsilon \operatorname{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) |\nabla \phi|. \quad (\text{A.9})$$

Thus, the numerical scheme for the above equation would be,

$$\begin{aligned}
\phi_{ij}^{n+1} = \phi_{ij}^n + \Delta t & \left[- [\max(F_{0ij}, 0)\nabla^+ + \min(F_{0ij}, 0)\nabla^-] \right. \\
& - \left\{ \max(u_{ij}^n, 0)D_{ij}^{-x} + \min(u_{ij}^n, 0)D_{ij}^{+x} \right. \\
& \left. \left. + \max(v_{ij}^n, 0)D_{ij}^{-y} + \min(v_{ij}^n, 0)D_{ij}^{+y} \right\} \right. \\
& \left. + [\epsilon K_{i,j}^n ((D_{ij}^{0x})^{1/2} + (D_{ij}^{0y})^{1/2})^{1/2}] \right], \tag{A.10}
\end{aligned}$$

where $K_{i,j}^n$ is the mean curvature computed using a central difference approximation.

Appendix B

Fast Marching Methods

Recall how the front propagation problem can be seen as a boundary value problem. Fast Marching Method (FFM) are solved by efficient use of the heap data structure, by keeping track of optimal ordering of the grid points (voxels). The goal is to design numerical methods to solve the Eikonal equation $|\nabla T|F = 1$. The upwind equation is given by,

$$[\max(D_{ijk}^{-x}T, D_{ijk}^{+x}T, 0) + \max(D_{ijk}^{-y}T, D_{ijk}^{+y}T, 0) + \max(D_{ijk}^{-z}T, D_{ijk}^{+z}T, 0)]^{1/2} = \frac{1}{F_{ijk}}. \quad (\text{B.1})$$

Solving this quadratic equation iteratively would require order $O(N^4)$ labour. To reduce time complexity, FFM exploits the causal relationship. FMM builds the solution of the above equation outward from the smallest T value, stepping away from the boundary condition in a downwind direction. We sweep the front along by considering points in a narrow band around the existing front and marching this front forward, freezing the values of the existing points and bringing new ones into the narrow band structure. The following is the FMM update procedure. Let us define the points in the boundary as *known* points, *trial* points are one grid point away and *far* as all other grid points. The following steps are performed iteratively,

- Let A be the *trial* point with the smallest T value.

-
- Add the point A to *known* and remove it from *trial*.
 - Tag as *trial* all the neighbours of A that are not *known*. If the neighbour is in *far*, remove, and add to the set *trial*.
 - Recompute the values of T at all *trial* neighbours of A by solving the above equation.

The initial conditions can be set as follows: label all the grid values points as *far away*, and assign them T values of ∞ . A signed distance function can be constructed around a narrow band from the initial hypersurface. The front can now be propagated in both forward and backward directions with speed F for a layer of grid points. All the points with negative crossing times are *known* points and with positive crossing times are the *trial* points with value T . Then we can begin the FMM algorithm described above.

The key to the efficient version is in the use of a min-heap data structure to have a fast way to locate the grid point in the narrow band with the smallest value. The smallest T value is thus at the root of the min-heap which takes $O(1)$ time to access it and then we also ensure that the values do not violate the heap property. Thus, the total time taken is $O(\log M)$ assuming there are M elements in the heap. A good in-depth discussion can be found in the text by [Sethian, 1999].

Appendix C

Additive Operator Splitting

As we have seen in Appendix A, the numerical updates are performed with Euler-forward explicit schemes based on simple finite discrete discretization which are stable only for very small time steps and often lead to poor efficiency in terms of implementation, making it of little practical use. [Weickert et al., 1998] present an implicit separable scheme based on the solution of a tridiagonal system of linear equations solved by *Thomas Algorithm* and *Additive Operator Splitting* methods that do not have any time step restrictions. This method is fast, stable, easy to implement and can accommodate large time steps. The authors introduced this numerical scheme for the case of non-linear diffusion filtering.

Problem Formulation

A filtered image $u(x,t)$ at a given time t of an image $f(x)$ can be given as a solution of the diffusion equation

$$\frac{\partial u}{\partial t} = \text{div}(g(|\nabla u_\sigma|^2)\nabla u) \quad (\text{C.1})$$

with the gradient of a smoothed version of u obtained by convolving u with a Gaussian of standard deviation (σ) given by $\nabla u_\sigma = \nabla K_\sigma * u$,

$$g(s) = \begin{cases} 1 & (s \leq 0) \\ 1 - \exp\left(\frac{-3.315}{(s/\lambda)^4}\right) & (s > 0). \end{cases}$$

Now, this diffusion equation can be discretized and written in an explicit setting and then in an implicit manner. Let's discretize the diffusion equation and see the explicit scheme.

$$\frac{u_i^{k+1} - u_i^k}{\tau} = \sum_{j \in N(i)} \frac{g_j^k + g_i^k}{2h^2} (u_j^k - u_i^k), \quad (\text{C.2})$$

where $N(i)$ is the neighbourhood of pixel i . We can write the above equation in a matrix vector notation,

$$\frac{u^{k+1} - u^k}{\tau} = A(u^k)u^k \quad (\text{C.3})$$

with $A(u^k) = (a_{ij}(u^k))$ and,

$$a_{ij}(u^k) := \begin{cases} \frac{g_j^k + g_i^k}{2h^2} & (j \in N(i)) \\ -\sum_{n \in N(i)} \frac{g_n^k + g_i^k}{2h^2} & (j = i) \\ 0 & \text{else} \end{cases}$$

or more compactly:

$$u^{k+1} = (I + \tau A(u^k))u^k, \quad (\text{C.4})$$

where $I \in R^N$ is the unit matrix. This is an explicit scheme as u^{k+1} can be directly calculated from u^k without solving a system of equations. This scheme suffers from a severe step-size restriction.

We now consider a slightly more complicated discretization,

$$u^k = (I - \tau A(u^k))u^{k+1}. \quad (\text{C.5})$$

This semi-implicit scheme does not suffer from any time-step restriction. But, this leads to solving a tridiagonal linear system which is solved using the famous Gaussian elimination algorithm, *Thomas Algorithm* as described in [Schwarz and Waldvogel, 1989] comprising of LR decomposition, forward and backward substitutions.

Now, let us state the diffusion equation for the m-dimensional case,

$$\frac{\partial u}{\partial t} = \sum_{l=1}^m \partial_{x_l} (g(|\nabla u_\sigma|^2) \partial_{x_l} u). \quad (\text{C.6})$$

The m-dim explicit scheme and its semi-implicit counterpart is given by the following two equations:

$$u^{k+1} = (I + \tau \sum_{l=1}^m A_l(u^k)) u^k, \quad (\text{C.7})$$

$$u^{k+1} = (I - \tau \sum_{l=1}^m A_l(u^k))^{-1} u^k. \quad (\text{C.8})$$

The explicit scheme again suffers from having to deal with small time-steps but the implicit scheme suffers from the destruction of the zeros in the band if we directly apply Gaussian elimination method and would lead to a massive computational burden. Iterative methods are required and hence a new method called Additive Operator Splitting (AOS) is proposed,

$$u^{k+1} = \frac{1}{m} (I - m\tau \sum_{l=1}^m A_l(u^k))^{-1} u^k. \quad (\text{C.9})$$

This method has several advantages, since all the coordinate axes are treated in exactly the same manner. These schemes satisfy all criteria for discrete nonlinear diffusion scale-spaces and are easy to implement in any dimension. Both computational and storage effort is linear in the number of pixels.

Appendix D

Sparse Field Methods

The narrow band algorithm reduces the computation time by restricting the computations to a small band near the evolving interface. The narrow band should however be wide enough to justify the updates. The sparse-field algorithm [Whitaker, 1998] takes narrow band to an extreme by computing updates on a band of grid points that is only one point wide. The key strategy is to create a neighbourhood around the active points that is precisely the width needed to calculate the derivatives for the next time step. The approach has several advantages: the algorithm does not require the recalculation of the positions of the level-sets and their distance transforms, thus performing the exact number of calculations needed to compute the next position of the level-set. Linked-list data structure is used to keep track of the number of points being computed. Thus, at each iteration only the points adjacent to the k -level curve are visited. Also, this approach helps in accurate computation of external forces as each level set is identified with a specific set of points. Another major advantage is that this algorithm allows for larger time steps in the update equations whereas in the narrow band method, the time step is limited by the speed of the fastest moving level curve.

Formulation We briefly look at the algorithm. A more detailed and intuitive explanation can be found in [Lankton, 2009]. The sparse-field method uses a list of points (via a doubly linked-list) to represent the zero level-set (called the *active set*). The neighbourhoods of

the active sets are defined in layers such as L_{+1}, \dots, L_{+N} and L_{-1}, \dots, L_{-N} . The active set have grid point values in the range $[-0.5, +0.5]$ and the rest of the layers have values in the range $[i - 1/2, i + 1/2]$.

- For each point x_j in the active set, compute the local geometry and the change in u_{x_j} based on the forces using some upwind equations.
- Now, using the updated $u_{x_j}(t + \Delta t)$ falls outside the range $[i - 1/2, i + 1/2]$ or not. If it does, then put the point x_j into the *status list*: S_{-1} or S_{+1} if its value is less than $-1/2$ or greater than $+1/2$ respectively.
- Next, visit each of the layers L_i in the following: $\pm 1, \pm 2, \dots$ and so on, and update the grid point values based on the values of the next inner layers $L_{\mp 1}$. If the grid point has no $L_{\mp 1}$ neighbours, then it gets demoted to $L_{\pm 1}$ i.e. the next level away from the active set.
- Lastly, for each status list do the following,
 - For each point x_j on the list S_i , remove x_j from the list $L_{\mp 1}$ and add it to the list L_i
 - Add $L_{\mp 1}$ neighbours to the list $S_{i\pm 1}$

Appendix E

Segmentation Evaluation Metrics

For the sake of completeness, we provide a short note on the evaluation metrics. For the segmentation evaluations in this report we use one spatial overlap measure and four surface distance based measures. Distance-based measures take into consideration the spatial position of voxels. The Hausdorff distance is generally sensitive to outliers due to presence of noise. It is thus recommended to use the Average Hausdorff distance or its 95th percentile.

E.1 Dice Coefficient

Let S_{sg} be the segmentation obtained from a segmentation algorithm and S_{gt} be the groundtruth segmentation. The Dice score can then be obtained as,

$$DICE = \frac{2|S_{sg} \cap S_{gt}|}{|S_{sg}| + |S_{gt}|} \quad (\text{E.1})$$

E.2 Average Symmetric Surface Distance (ASSD)

ASSD is the average of all the distances from points on the boundary of one surface to the boundary of the other surface and vice versa. Let B_{sg} and B_{gt} be the points on the surface of the segmentation and its groundtruth.

$$ASSD = \frac{1}{\|B_{sg}\| + \|B_{gt}\|} \sum_{x \in B_{sg}} d(x, B_{gt}) + \sum_{y \in B_{gt}} d(y, B_{sg}) \quad (\text{E.2})$$

E.3 Hausdorff Distance (HD)

The Hausdorff distance between two finite points sets P and Q is defined by,

$$HD(P, Q) = \max(h(P, Q), h(Q, P)) \quad (\text{E.3})$$

where $h(P, Q)$ is called the directed Hausdorff distance,

$$h(P, Q) = \max_{p \in P} \min_{q \in Q} \|p - q\| \quad (\text{E.4})$$

where $\max_{p \in P} \min_{q \in Q} \|p - q\|$ is some norm such as the Euclidean norm.

E.4 Average Hausdorff Distance (AHD)

The Hausdorff distance between two finite points sets P and Q is defined by,

$$HD(P, Q) = \max(h(P, Q), h(Q, P)) \quad (\text{E.5})$$

where $h(P, Q)$ is called the directed Average Hausdorff distance,

$$h(P, Q) = \frac{1}{N} \sum_{p \in P} \min_{q \in Q} \|p - q\| \quad (\text{E.6})$$

where N is the total number of points in set P or Q .

E.5 Hausdorff Distance 95th percentile (HD95)

This computes the 95th percentile of the Hausdorff Distance between the two surfaces B_{sg} and B_{gt} . Compared to the Hausdorff distance, this metric is slightly more stable to outliers.

Appendix F

Software used in this work

In this section, we would like to acknowledge some of the open-source software and code repositories used in this thesis.

- ITK-SNAP [Yushkevich et al., 2006] for placement of seeds, bounding-box and cropping out extraneous regions, and for conversion between different file formats.
- Paraview [Ayachit, 2015] for visualization. Most of the figures in thesis are based on the features of this software.
- 3DSlicer [Kikinis et al., 2014] for visualization.
- Seaborn library [Waskom et al., 2017] for visualization.
- MedPy [Maier et al., 2019] for computing the metrics.
- EvaluateSegmentation [Taha and Hanbury, 2015] for computing the metrics.
- FreeSurfer [Dale et al., 1999] for registration of all the images to a common space.
- Inkscape [Bah, 2011] for image editing and making the flowchart.

Bibliography

- [Aslan et al., 2010] Aslan, M. S., Ali, A., Farag, A. A., Rara, H., Arnold, B., and Xiang, P. (2010). 3d vertebral body segmentation using shape based graph cuts. In *2010 20th International Conference on Pattern Recognition*, pages 3951–3954.
- [Ayachit, 2015] Ayachit, U. (2015). *The ParaView Guide: A Parallel Visualization Application*. Kitware, Inc., USA.
- [Badino et al., 2012] Badino, H., Huber, D., and Kanade, T. (2012). Real-time topometric localization. In *International Conference on Robotics and Automation*.
- [Bah, 2011] Bah, T. (2011). *Inkscape: Guide to a Vector Drawing Program*. Prentice Hall Press, Upper Saddle River, NJ, USA, 4th edition.
- [Bresson et al., 2006] Bresson, X., Vandergheynst, P., and Thiran, J.-P. (2006). A variational model for object segmentation using boundary information and shape prior driven by the mumford-shah functional. *International Journal of Computer Vision*, 68(2):145–162.
- [Caselles et al., 1995] Caselles, V., Kimmel, R., and Sapiro, G. (1995). Geodesic active contours. In *Proceedings of IEEE International Conference on Computer Vision*, pages 694–699.
- [Castro-Mateos et al., 2015] Castro-Mateos, I., Pozo, J. M., Perea, M., Lekadir, K., Lazary, A., and Frangi, A. F. (2015). Statistical interspace models (sims): Application to robust 3d spine segmentation. *IEEE Transactions on Medical Imaging*, 34(8):1663–1675.

- [Chan and Zhu, 2005] Chan, T. and Zhu, W. (2005). Level set based shape prior segmentation. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 2, pages 1164–1170. IEEE.
- [Chan and Vese, 2001] Chan, T. F. and Vese, L. A. (2001). Active contours without edges. *IEEE Transactions on image processing*, 10(2):266–277.
- [Chu et al., 2015a] Chu, C., Belavý, D. L., Armbrecht, G., Bansmann, M., Felsenberg, D., and Zheng, G. (2015a). Fully automatic localization and segmentation of 3d vertebral bodies from ct/mr images via a learning-based method. *PloS one*, 10(11):e0143327.
- [Chu et al., 2015b] Chu, C., Belav, D. L., Armbrecht, G., Bansmann, M., Felsenberg, D., and Zheng, G. (2015b). Fully automatic localization and segmentation of 3d vertebral bodies from ct/mr images via a learning-based method. *PLOS ONE*, 10(11):1–22.
- [Cootes et al., 1995] Cootes, T. F., Taylor, C. J., Cooper, D. H., and Graham, J. (1995). Active shape models-their training and application. *Computer vision and image understanding*, 61(1):38–59.
- [Cremers et al., 2007] Cremers, D., Rousson, M., and Deriche, R. (2007). A review of statistical approaches to level set segmentation: Integrating color, texture, motion and shape. *International Journal of Computer Vision*, 72(2):195–215.
- [Dalal and Triggs, 2005] Dalal, N. and Triggs, B. (2005). Histograms of oriented gradients for human detection. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 1, pages 886–893 vol. 1.
- [Dale et al., 1999] Dale, A., Fischl, B., and Sereno, M. I. (1999). Cortical surface-based analysis: I. segmentation and surface reconstruction. *NeuroImage*, 9(2):179 – 194.
- [Dambreville et al., 2007] Dambreville, S., Niethammer, M., Yezzi, A., and Tannenbaum, A. R. (2007). A variational framework combining level-sets and thresholding. Georgia Institute of Technology.

- [Drouin et al., 2017] Drouin, S., Kochanowska, A., Kersten-Oertel, M., Gerard, I. J., Zelmann, R., De Nigris, D., Bériault, S., Arbel, T., Sirhan, D., Sadikot, A. F., Hall, J. A., Sinclair, D. S., Petrecca, K., DelMaestro, R. F., and Collins, D. L. (2017). Ibis: an or ready open-source platform for image-guided neurosurgery. *International Journal of Computer Assisted Radiology and Surgery*, 12(3):363–378.
- [Forsberg, 2015] Forsberg, D. (2015). *Atlas-Based Registration for Accurate Segmentation of Thoracic and Lumbar Vertebrae in CT Data*, pages 49–59. Springer International Publishing, Cham.
- [Frangakis and Hegerl, 2001] Frangakis, A. S. and Hegerl, R. (2001). Noise reduction in electron tomographic reconstructions using nonlinear anisotropic diffusion. *Journal of structural biology*, 135(3):239–250.
- [Gill et al., 2012] Gill, S., Abolmaesumi, P., Fichtinger, G., Boisvert, J., Pichora, D., Borshneck, D., and Mousavi, P. (2012). Biomechanically constrained groupwise ultrasound to ct registration of the lumbar spine. *Medical Image Analysis*, 16(3):662 – 674. Computer Assisted Interventions.
- [Goulet, 2010] Goulet, B. G. (2010). Lumbar-sacral pedicle screw insertion with preoperative ct-based navigation.
- [Hammernik, 2015] Hammernik, K. (2015). *Convex framework for 2D & 3D image segmentation using shape constraints*. PhD thesis, Masters thesis, Graz University of Technology, Austria.
- [Hammernik et al., 2015a] Hammernik, K., Ebner, T., Stern, D., Urschler, M., and Pock, T. (2015a). *Vertebrae Segmentation in 3D CT Images Based on a Variational Framework*, pages 227–233. Springer International Publishing, Cham.
- [Hammernik et al., 2015b] Hammernik, K., Ebner, T., Stern, D., Urschler, M., and Pock, T. (2015b). Vertebrae segmentation in 3d ct images based on a variational framework. In *Recent advances in computational methods and clinical applications for spine imaging*, pages 227–233. Springer.

- [Huang, 2013] Huang (2013). An improved level set method for vertebra ct image segmentation. *Biomed Eng Online*, 12:48.
- [Ibragimov et al., 2015] Ibragimov, B., Korez, R., Likar, B., Pernuš, F., and Vrtovec, T. (2015). *Interpolation-Based Detection of Lumbar Vertebrae in CT Spine Images*, pages 73–84. Springer International Publishing, Cham.
- [Ibragimov et al., 2014] Ibragimov, B., Likar, B., Pernuš, F., and Vrtovec, T. (2014). Segmentation of vertebrae from 3d spine images by applying concepts from transportation and game theories. In Yao, J., Klinder, T., and Li, S., editors, *Computational Methods and Clinical Applications for Spine Imaging*, pages 3–14, Cham. Springer International Publishing.
- [Ibragimov et al., 2014] Ibragimov, B., Likar, B., Pernu, F., and Vrtovec, T. (2014). Shape representation for efficient landmark-based segmentation in 3-d. *IEEE Transactions on Medical Imaging*, 33(4):861–874.
- [Janssens et al., 2018] Janssens, R., Zeng, G., and Zheng, G. (2018). Fully automatic segmentation of lumbar vertebrae from ct images using cascaded 3d fully convolutional networks. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pages 893–897.
- [Kadoury and Paragios, 2010] Kadoury, S. and Paragios, N. (2010). Nonlinear embedding towards articulated spine shape inference using higher-order mrfs. In Jiang, T., Navab, N., Pluim, J. P. W., and Viergever, M. A., editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2010*, pages 579–586, Berlin, Heidelberg. Springer Berlin Heidelberg.
- [Kaminsky et al., 2004] Kaminsky, J., Klinge, P., Rodt, T., Bokemeyer, M., Luedemann, W., and Samii, M. (2004). Specially adapted interactive tools for an improved 3d-segmentation of the spine. *Computerized Medical Imaging and Graphics*, 28(3):119 – 127.

- [Kass et al., 1988] Kass, M., Witkin, A., and Terzopoulos, D. (1988). Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331.
- [Kelm et al., 2013] Kelm, B. M., Wels, M., Zhou, S. K., Seifert, S., Suehling, M., Zheng, Y., and Comaniciu, D. (2013). Spine detection in ct and mr using iterated marginal space learning. *Medical Image Analysis*, 17(8):1283 – 1292.
- [Kikinis et al., 2014] Kikinis, R., Pieper, S. D., and Vosburgh, K. G. (2014). *3D Slicer: A Platform for Subject-Specific Image Analysis, Visualization, and Clinical Support*, pages 277–289. Springer New York, New York, NY.
- [Kim et al., 2005] Kim, J., Fisher, J. W., Yezzi, A., Çetin, M., and Willsky, A. S. (2005). A nonparametric statistical method for image segmentation using information theory and curve evolution. *IEEE Transactions on Image processing*, 14(10):1486–1502.
- [Kim and Kim, 2009] Kim, Y. and Kim, D. (2009). A fully automatic vertebra segmentation method using 3d deformable fences. *Computerized Medical Imaging and Graphics*, 33(5):343 – 352.
- [Klinder et al., 2009] Klinder, T., Ostermann, J., Ehm, M., Franz, A., Kneser, R., and Lorenz, C. (2009). Automated model-based vertebra detection, identification, and segmentation in ct images. *Medical Image Analysis*, 13(3):471 – 482.
- [Korez et al., 2015] Korez, R., Ibragimov, B., Likar, B., Pernuš, F., and Vrtovec, T. (2015). *An Improved Shape-Constrained Deformable Model for Segmentation of Vertebrae from CT Lumbar Spine Images*, pages 85–94. Springer International Publishing, Cham.
- [Korez et al., 2015] Korez, R., Ibragimov, B., Likar, B., Pernu, F., and Vrtovec, T. (2015). A framework for automated spine and vertebrae interpolation-based detection and model-based segmentation. *IEEE Transactions on Medical Imaging*, 34(8):1649–1662.
- [Korez et al., 2016] Korez, R., Likar, B., Pernuš, F., and Vrtovec, T. (2016). Model-based segmentation of vertebral bodies from mr images with 3d cnns. In Ourselin, S., Joskowicz, L., Sabuncu, M. R., Unal, G., and Wells, W., editors, *Medical Image Computing*

- and Computer-Assisted Intervention – MICCAI 2016*, pages 433–441, Cham. Springer International Publishing.
- [Kroon and Slump, 2009] Kroon, D. and Slump, C. H. (2009). Coherence filtering to enhance the mandibular canal in cone-beam ct data. In *IEEE-EMBS Benelux Chapter Symposium*, volume 11, pages 40–41.
- [Kroon et al., 2010] Kroon, D.-J., Slump, C. H., and Maal, T. J. (2010). Optimized anisotropic rotational invariant diffusion scheme on cone-beam ct. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 221–228. Springer.
- [Lankton, 2009] Lankton, S. (2009). Sparse field methods - technical report. <http://www.shawnlankton.com/wp-content/uploads/articles/lankton-sfm-TR-2009.pdf>.
- [Law and Chung, 2009] Law, M. W. and Chung, A. C. (2009). Efficient implementation for spherical flux computation and its application to vascular segmentation. *IEEE transactions on image processing*, 18(3):596–612.
- [Lessmann et al., 2019] Lessmann, N., van Ginneken, B., de Jong, P. A., and Igum, I. (2019). Iterative fully convolutional neural networks for automatic vertebra segmentation and identification. *Medical Image Analysis*, 53:142 – 155.
- [Leventon et al., 2002] Leventon, M. E., Grimson, W. E. L., and Faugeras, O. (2002). Statistical shape influence in geodesic active contours. In *5th IEEE EMBS International Summer School on Biomedical Imaging, 2002.*, pages 8–pp. IEEE.
- [Lim et al., 2013] Lim, P. H., Bagci, U., and Bai, L. (2013). Introducing willmore flow into level set segmentation of spinal vertebrae. *IEEE Transactions on Biomedical Engineering*, 60(1):115–122.
- [Maier et al., 2019] Maier, O., Rothberg, A., Raamana, P. R., Bges, R., Isensee, F., Ahern, M., mamrehn, VincentXWD, and Joshi, J. (2019). loli/medpy: Medpy 0.4.0.

- [Malladi et al., 1995] Malladi, R., Sethian, J. A., and Vemuri, B. C. (1995). Shape modeling with front propagation: a level set approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(2):158–175.
- [Marquez-Neila et al., 2013] Marquez-Neila, P., Baumela, L., and Alvarez, L. (2013). A morphological approach to curvature-based evolution of curves and surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(1):2–17.
- [Michailovich et al., 2007] Michailovich, O., Rathi, Y., and Tannenbaum, A. (2007). Image segmentation using active contours driven by the bhattacharyya gradient flow. *IEEE Transactions on Image Processing*, 16(11):2787–2801.
- [Mukherjee and Acton, 2015] Mukherjee, S. and Acton, S. T. (2015). Region based segmentation in presence of intensity inhomogeneity using legendre polynomials. *IEEE Signal Processing Letters*, 22(3):298–302.
- [Oliva and Torralba, 2001] Oliva, A. and Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. *International Journal of Computer Vision*, 42(3):145–175.
- [Osher and Sethian, 1988] Osher, S. and Sethian, J. A. (1988). Fronts propagating with curvature-dependent speed: Algorithms based on hamilton-jacobi formulations. *Journal of Computational Physics*, 79(1):12 – 49.
- [Otsu, 1979] Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66.
- [Pereañez et al., 2015] Pereañez, M., Lekadir, K., Hoogendoorn, C., Castro-Mateos, I., and Frangi, A. (2015). *Detailed Vertebral Segmentation Using Part-Based Decomposition and Conditional Shape Models*, pages 95–103. Springer International Publishing, Cham.
- [Perona and Malik, 1990] Perona, P. and Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on pattern analysis and machine intelligence*, 12(7):629–639.

- [Rasoulilian et al., 2013] Rasoulilian, A., Rohling, R., and Abolmaesumi, P. (2013). Lumbar spine segmentation using a statistical multi-vertebrae anatomical shape+pose model. *IEEE Transactions on Medical Imaging*, 32(10):1890–1900.
- [Rathi et al., 2006] Rathi, Y., Michailovich, O., Malcolm, J., and Tannenbaum, A. (2006). Seeing the unseen: Segmenting with distributions. In *Intl. Conf. Signal and Image Processing*, volume 534.
- [Rousson, 2004] Rousson, M. (2004). *Cue Integration and Front Evolution in Image Segmentation*. Theses, Université Nice Sophia Antipolis.
- [Rousson et al., 2003] Rousson, M., Brox, T., and Deriche, R. (2003). Active unsupervised texture segmentation on a diffusion based feature space. In *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings.*, volume 2, pages II–699. IEEE.
- [Rudin et al., 1992] Rudin, L. I., Osher, S., and Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D: nonlinear phenomena*, 60(1-4):259–268.
- [Sandhu et al., 2008] Sandhu, R., Georgiou, T., and Tannenbaum, A. (2008). A new distribution metric for image segmentation. In *Medical Imaging 2008: Image Processing*, volume 6914, page 691404. International Society for Optics and Photonics.
- [Schwarz and Waldvogel, 1989] Schwarz, H. R. and Waldvogel, J. (1989). *Numerical analysis: a comprehensive introduction*. John Wiley & Sons.
- [Sethian, 1996] Sethian, J. A. (1996). A fast marching level set method for monotonically advancing fronts. *Proceedings of the National Academy of Sciences*, 93(4):1591–1595.
- [Sethian, 1999] Sethian, J. A. (1999). *Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science*, volume 3. Cambridge university press.
- [Shalaby et al., 2014] Shalaby, A., Farag, A., and Aslan, M. (2014). 2d-pca based tensor level set framework for vertebral body segmentation. In Yao, J., Klinder, T., and Li,

- S., editors, *Computational Methods and Clinical Applications for Spine Imaging*, pages 35–48, Cham. Springer International Publishing.
- [Stephansen, 2012] Stephansen, U. L. (2012). Level-set appearance modeling for segmentation of anatomical structures in 3-d medical images. *Technical Report*.
- [Stern et al., 2011] Stern, D., Likar, B., Pernu, F., and Vrtovec, T. (2011). Parametric modelling and segmentation of vertebral bodies in 3d ct and mr spine images. *Physics in Medicine Biology*, 56(23):7505.
- [Taha and Hanbury, 2015] Taha, A. A. and Hanbury, A. (2015). Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC Medical Imaging*, 15:29.
- [Tan et al., 2008] Tan, S., Yao, J., Ward, M. M., Yao, L., and Summers, R. M. (2008). Computer aided evaluation of ankylosing spondylitis using high-resolution ct. *IEEE Transactions on Medical Imaging*, 27(9):1252–1267.
- [Tsai et al., 2003] Tsai, A., Yezzi, A., Wells III, W., Tempany, C., Tucker, D., Fan, A., Grimson, W. E., and Willsky, A. S. (2003). A shape-based approach to the segmentation of medical imagery using level sets.
- [Turner, 1986] Turner, M. R. (1986). Texture discrimination by gabor functions. *Biological cybernetics*, 55(2-3):71–82.
- [Vasilevskiy and Siddiqi, 2002] Vasilevskiy, A. and Siddiqi, K. (2002). Flux maximizing geometric flows. *IEEE transactions on pattern analysis and machine intelligence*, 24(12):1565–1578.
- [Wang et al., 2015] Wang, Z., Zhen, X., Tay, K., Osman, S., Romano, W., and Li, S. (2015). Regression segmentation for m^3 spinal images. *IEEE Transactions on Medical Imaging*, 34(8):1640–1648.
- [Waskom et al., 2017] Waskom, M., Botvinnik, O., O’Kane, D., Hobson, P., Lukauskas, S., Gemperline, D. C., Augspurger, T., Halchenko, Y., Cole, J. B., Warmenhoven, J.,

- de Ruiter, J., Pye, C., Hoyer, S., Vanderplas, J., Villalba, S., Kunter, G., Quintero, E., Bachant, P., Martin, M., Meyer, K., Miles, A., Ram, Y., Yarkoni, T., Williams, M. L., Evans, C., Fitzgerald, C., Brian, Fonnesbeck, C., Lee, A., and Qalieh, A. (2017). `mwaskom/seaborn: v0.8.1` (september 2017).
- [Weickert et al., 1998] Weickert, J., Romeny, B. M. T. H., and Viergever, M. A. (1998). Efficient and reliable schemes for nonlinear diffusion filtering. *IEEE Transactions on Image Processing*, 7(3):398–410.
- [Weickert and Schar, 2002] Weickert, J. and Schar, H. (2002). A scheme for coherence-enhancing diffusion filtering with optimized rotation invariance. *Journal of Visual Communication and Image Representation*, 13(1-2):103–118.
- [Whitaker, 1998] Whitaker, R. T. (1998). A level-set approach to 3d reconstruction from range data. *International Journal of Computer Vision*, 29(3):203–231.
- [Wikipedia, 2015] Wikipedia (2015). Anatomy of a vertebra.
- [Xu and Prince, 1998] Xu, C. and Prince, J. L. (1998). Snakes, shapes, and gradient vector flow. *IEEE Transactions on Image Processing*, 7(3):359–369.
- [Yushkevich et al., 2006] Yushkevich, P. A., Piven, J., Cody Hazlett, H., Gimpel Smith, R., Ho, S., Gee, J. C., and Gerig, G. (2006). User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage*, 31(3):1116–1128.
- [Zhang et al., 2010] Zhang, J., Yan, C.-H., Chui, C.-K., and Ong, S.-H. (2010). Fast segmentation of bone in ct images using 3d adaptive thresholding. *Computers in biology and medicine*, 40(2):231–236.
- [Zhao et al., 2001] Zhao, Osher, S., and Fedkiw, R. (2001). Fast surface reconstruction using the level set method. In *Proceedings IEEE Workshop on Variational and Level Set Methods in Computer Vision*, pages 194–201.