Left Ventricle Myocardium Segmentation from 3D Cardiac MR Images using Combined Probabilistic Atlas and Graph Cut-based Approaches

Aditya Daryanani
ad1154@rit.edu

Follow this and additional works at: https://scholarworks.rit.edu/theses

Recommended Citation

This Thesis is brought to you for free and open access by RIT Scholar Works. It has been accepted for inclusion in Theses by an authorized administrator of RIT Scholar Works. For more information, please contact ritscholarworks@rit.edu.
Left Ventricle Myocardium Segmentation from 3D Cardiac MR Images using Combined Probabilistic Atlas and Graph Cut-based Approaches

by

Aditya Daryanani

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Computer Engineering

Supervised by

Dr. Cristian A. Linte
Department of Biomedical Engineering
Kate Gleason College of Engineering
Rochester Institute of Technology
Rochester, New York
December 2015

Approved by:

Dr. Cristian A. Linte, Assistant Professor
Thesis Advisor, Department of Biomedical Engineering

Dr. Raymond Ptucha, Assistant Professor
Committee Member, Department of Computer Engineering

Dr. Andreas Savakis, Professor
Committee Member, Department of Computer Engineering
© Copyright 2015 by Aditya Daryanani
All Rights Reserved
Dedicated to my family, for being a pillar of strength.
Acknowledgments

I wish to thank my thesis advisor Dr. Cristian A. Linte, for his valuable advice, understanding, the freedom he gave me to experiment, explore and learn, his patience, and for the countless hours spent reviewing and critiquing this work. A special thanks to Dr. Raymond Ptucha for his consistent support, encouragement, helpfulness and mentorship throughout the course of my time at this university. Thanks to Dr. Andreas Savakis for his excellent teaching and agreeing to be on my committee, and the rest of the computer engineering department and staff. I would also like to thank my friends and colleagues Shusil Dangi and Yehuda K. Ben-Zikri for their valuable inputs that have had a significant impact on this work.
Abstract

Left Ventricle Myocardium Segmentation from 3D Cardiac MR Images using Combined Probabilistic Atlas and Graph Cut-based Approaches

Aditya Daryanani

Medical imaging modalities, including Computed Tomography (CT) Magnetic Resonance Imaging (MRI) and Ultrasound (US) are critical for the diagnosis and progress monitoring of many cardiac conditions, planning, visualization and delivery of therapy via minimally invasive intervention procedures, as well as for teaching, training and simulation applications.

Image segmentation is a processing technique that allows the user to extract the necessary information from an image dataset, in the form of a surface model of the region of interest from the anatomy. A wide variety of segmentation techniques have been developed and implemented for cardiac MR images. Despite their complexity and performance, many of them are intended for specific image datasets or are too specific to be employed for segmenting classical clinical quality Magnetic Resonance (MR) images.

Graph Cut based segmentation algorithms have been shown to work well in regards to medical image segmentation. In addition, they are computationally efficient, which scales well to real time applications. While the basic graph cuts algorithms use lower-order statistics, combining this segmentation approach with atlas-based methods may help improve segmentation accuracy at a lower computational cost.

The proposed technique will be tested at each step during the development by assessing the segmentation results against the available ground truth segmentation. Several metrics will be used to quantify the performance of the proposed technique, including computational performance, segmentation accuracy and fidelity assessed via the Sørensen-Dice Coefficient (DSC), Mean Absolute Distance (MAD) and Hausdorff Distance (HD) metrics.
Contents

Dedication .......................................................... iii

Acknowledgments ................................................... iv

Abstract ............................................................. v

1 Introduction .................................................. 1
   1.1 Opening Remarks and Generic Statement .............. 1
   1.2 Cardiac Anatomy and Physiology ...................... 3
   1.3 Cardiac Imaging Modalities .......................... 5
      1.3.1 Magnetic Resonance Imaging ................... 6
      1.3.2 Computed Tomography ......................... 8
      1.3.3 Ultrasound Imaging .......................... 10
   1.4 Imaging Applications ................................. 11
      1.4.1 Computer-aided diagnosis ................... 11
      1.4.2 Image Guided Interventions .................. 11
      1.4.3 Teaching, Simulation and Training .......... 12
   1.5 Overview of segmentation methods .................. 12
      1.5.1 Classification and Clustering ............... 14
      1.5.2 Active Contours ............................ 15
      1.5.3 Level Set Segmentation ..................... 16
         1.5.3.1 Conventional Level Set ............... 16
         1.5.3.2 GeoDesic Level Sets ............... 17
      1.5.4 Shape and Appearance Models ............. 17
1.5.5 Atlas-based Methods ........................................ 17
1.5.6 Graph cut-based Energy Minimization Methods ... 18
1.6 Image Registration ............................................. 18
  1.6.1 Geometric Transformations .............................. 19
    1.6.1.1 Translation ........................................ 19
    1.6.1.2 Scaling ............................................. 20
    1.6.1.3 Shearing ........................................... 21
    1.6.1.4 Rotation ........................................... 21
    1.6.1.5 Rigid, Similarity and Affine Transforma-
            tions .............................................. 22
  1.6.2 Global Registration ....................................... 22
  1.6.3 Local Registration ....................................... 22
    1.6.3.1 Piecewise Rigid .................................. 23
    1.6.3.2 Basis Spline (B-spline) ........................... 23
    1.6.3.3 Linear Elastic Models ............................. 23
    1.6.3.4 Finite Element Modelling (FEM) ................. 23
    1.6.3.5 Viscous Fluid Models ............................ 24
  1.6.4 Similarity measures ..................................... 24
    1.6.4.1 Sum of Squared Differences ....................... 24
    1.6.4.2 Pattern Intensity ................................ 25
    1.6.4.3 Normalized Cross-correlation ..................... 25
    1.6.4.4 Mutual information ................................ 26
  1.6.5 Optimizers ............................................... 26
    1.6.5.1 The Simplex Optimizer ............................ 26
    1.6.5.2 The Quasi-Newton Optimizer Family ............ 26
  1.7 Segmentation Evaluation .................................. 27
    1.7.1 Sørensen-Dice Coefficient (DSC) .................... 27
    1.7.2 Mean Absolute Distance (MAD) ....................... 28
1.7.3 Hausdorff Distance (HD) \hspace{1cm} 28
1.8 Segmentation Challenges \hspace{1cm} 30
1.9 Proposed Solution and Thesis Outline \hspace{1cm} 31

2 Methodology \hspace{1cm} 33
2.1 Overview \hspace{1cm} 33
2.2 Imaging Datasets \hspace{1cm} 34
2.3 Data preparation and Pre-processing \hspace{1cm} 35
2.4 Atlas Construction via Registration \hspace{1cm} 37
  2.4.1 Image Registration Implementation \hspace{1cm} 38
  2.4.2 Averaging transformations \hspace{1cm} 38
    2.4.2.1 Global (Affine) \hspace{1cm} 38
    2.4.2.2 Local (B-spline non-rigid) \hspace{1cm} 39
  2.4.3 Global Atlas \hspace{1cm} 40
  2.4.4 Global + Local Atlas \hspace{1cm} 41
2.5 Atlas Based Initial Label Propagation \hspace{1cm} 42
2.6 Graph cuts based segmentation implementation \hspace{1cm} 43
  2.6.1 Basic graph cut segmentation operation \hspace{1cm} 43
  2.6.2 Continuous Max Flow approach \hspace{1cm} 46
2.7 Atlas + Continuous Max-Flow (CMF) Segmentation Integration \hspace{1cm} 46
  2.7.1 Combined Atlas and Continuous max Flow Segmentation by Weighted Sum \hspace{1cm} 46
  2.7.2 Combined Atlas and Continuous max Flow Segmentation by Comparison Selection \hspace{1cm} 48
2.8 Evaluation \hspace{1cm} 49

3 Results \hspace{1cm} 50
3.1 Atlas-based Label Propagation \hspace{1cm} 50
3.1.1 Reference Patient-based Atlas ...................... 51
  3.1.1.1 Affine Label Propagation .................. 51
    3.1.1.1.1 3D Evaluation ....................... 51
    3.1.1.1.2 2D Evaluation ....................... 52
3.1.2 Global Atlas (GA) .................. 53
  3.1.2.1 Affine Label Propagation .................. 53
    3.1.2.1.1 3D Evaluation ....................... 53
    3.1.2.1.2 2D Evaluation ....................... 53
3.1.3 Global + Local Atlas (GLA) .................. 53
  3.1.3.1 Affine + Non-Rigid Label Propagation ... 53
    3.1.3.1.1 3D Evaluation ....................... 53
    3.1.3.1.2 2D Evaluation ....................... 54
  3.1.3.2 Affine Label Propagation .................. 55
    3.1.3.2.1 3D Evaluation ....................... 55
    3.1.3.2.2 2D Evaluation ....................... 55
3.2 Global + Local Atlas with Continuous Max Flow .... 58
  3.2.1 Affine Label Propagation and Weighted Sum CMF Refinement .................. 58
    3.2.1.1 3D Evaluation ....................... 58
    3.2.1.2 2D Evaluation ....................... 58
  3.2.2 Affine Label Propagation and CMF Comparison Selection .................. 58
    3.2.2.1 3D Evaluation ....................... 58
    3.2.2.2 2D Evaluation ....................... 59
3.3 Summary ........................................... 61

4 Summary, Discussion and Future Work .................. 64
  4.1 Summary ........................................... 64
  4.2 Discussion ........................................ 65
4.3 Future directions .................................................. 69
  4.3.1 Multi-phase Scaling through Shape Modeling .... 69
  4.3.2 Integration with Other Low-cost Segmentation Meth-
       ods .......................................................... 70
  4.3.3 Alternate Approaches to Combine Labels from Dif-
        ferent algorithms ........................................ 70
  4.3.4 Potential Scaling to Real-time Performance ....... 70

Bibliography .......................................................... 72
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>3D Single reference patient-based affine label propagation evaluation</td>
<td>51</td>
</tr>
<tr>
<td>3.2</td>
<td>3D global atlas and affine label propagation evaluation</td>
<td>53</td>
</tr>
<tr>
<td>3.3</td>
<td>3D global + local atlas and affine + non-rigid label propagation evaluation</td>
<td>55</td>
</tr>
<tr>
<td>3.4</td>
<td>3D global + local atlas and affine label propagation evaluation</td>
<td>55</td>
</tr>
<tr>
<td>3.5</td>
<td>3D evaluation of global + local atlas with affine label propagation and subsequent weighted-sum CMF refinement</td>
<td>58</td>
</tr>
<tr>
<td>3.6</td>
<td>3D evaluation of global + local atlas with affine label propagation and comparison selection-based CMF refinement</td>
<td>60</td>
</tr>
<tr>
<td>3.7</td>
<td>Results summary</td>
<td>62</td>
</tr>
<tr>
<td>4.1</td>
<td>Label propagation performance</td>
<td>69</td>
</tr>
</tbody>
</table>
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Cardiac anatomy</td>
<td>3</td>
</tr>
<tr>
<td>1.2</td>
<td>Whole heart geometric view</td>
<td>5</td>
</tr>
<tr>
<td>1.3</td>
<td>Examples of cardiac images acquires using different medical imaging modalities</td>
<td>6</td>
</tr>
<tr>
<td>1.4</td>
<td>Example of short- and long-axis views through the left ventricle</td>
<td>9</td>
</tr>
<tr>
<td>1.5</td>
<td>Schematic representation of various geometric transformations on a simple image</td>
<td>20</td>
</tr>
<tr>
<td>1.6</td>
<td>Sørensen-Dice Coefficient (DSC)</td>
<td>27</td>
</tr>
<tr>
<td>1.7</td>
<td>Schematic representation of the Hausdorff Distance (HD) metric</td>
<td>29</td>
</tr>
<tr>
<td>2.1</td>
<td>Training dataset variability with respect to reference patient prior to global alignment</td>
<td>36</td>
</tr>
<tr>
<td>2.2</td>
<td>Registration visualization</td>
<td>39</td>
</tr>
<tr>
<td>2.3</td>
<td>Global atlas construction</td>
<td>40</td>
</tr>
<tr>
<td>2.4</td>
<td>Training dataset variability with respect to reference patient after affine global alignment</td>
<td>42</td>
</tr>
<tr>
<td>2.5</td>
<td>Global + Local atlas construction</td>
<td>43</td>
</tr>
<tr>
<td>2.6</td>
<td>Non-Rigid Atlas Construction</td>
<td>44</td>
</tr>
<tr>
<td>2.7</td>
<td>Basic graph cuts</td>
<td>45</td>
</tr>
<tr>
<td>2.8</td>
<td>Weighted Sum label combination method</td>
<td>47</td>
</tr>
<tr>
<td>2.9</td>
<td>Comparison Selection label combination method</td>
<td>48</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>3.1</td>
<td>Slice-wise 2D reference patient-based affine label propagation evaluation</td>
<td>52</td>
</tr>
<tr>
<td>3.2</td>
<td>Slice-wise 2D global atlas and affine label propagation evaluation</td>
<td>54</td>
</tr>
<tr>
<td>3.3</td>
<td>Slice-wise 2D global + local atlas with affine + non-rigid label propagation evaluation</td>
<td>56</td>
</tr>
<tr>
<td>3.4</td>
<td>Slice-wise 2D global + local atlas with affine label propagation evaluation</td>
<td>57</td>
</tr>
<tr>
<td>3.5</td>
<td>Slice-wise 2D evaluation of the global + local atlas with affine label propagation and weighted sum CMF refinement</td>
<td>59</td>
</tr>
<tr>
<td>3.6</td>
<td>Slice-wise 2D evaluation of the comparison selection CMF refinement and affine label propagation of the global + local atlas</td>
<td>61</td>
</tr>
<tr>
<td>3.7</td>
<td>Visual display of the segmentation results shown across the same LV mid-slice across all 30 testing patients</td>
<td>63</td>
</tr>
<tr>
<td>4.1</td>
<td>Effect of training on patient dataset variability</td>
<td>66</td>
</tr>
<tr>
<td>4.2</td>
<td>Atlas-to-patient misregistrations and effect on label propagation</td>
<td>67</td>
</tr>
<tr>
<td>4.3</td>
<td>CMF performance visualization</td>
<td>68</td>
</tr>
</tbody>
</table>
Acronyms

**B-spline**  Basis Spline. viii, ix, 23, 38, 39

**CMF**  Continuous Max-Flow. ix, x, xii, xiv, 45, 46, 49, 50, 58, 60, 64, 67, 68, 69, 70

**CT**  Computed Tomography. v, 1, 5, 8, 9

**DSC**  Sørensen-Dice Coefficient. v, viii, xiii, 2, 27, 32, 34, 49, 50, 51, 52, 53, 55, 58, 60, 67

**FEM**  Finite Element Modelling. viii, 23

**FOV**  Field Of View. 34

**GA**  Global Atlas. x, 53, 55, 67, 70

**GLA**  Global + Local Atlas. x, 53, 55, 58, 60, 67, 70

**GT**  Ground Truth. 40, 41, 64, 65

**HD**  Hausdorff Distance. v, viii, xiii, 2, 27, 28, 29, 32, 34, 50, 51, 52, 53, 55, 58, 60, 67, 68

**kNN**  k-Nearest Neighbour. 14

**L-BFGS**  Limited Memory Broyden Fletcher Goldfarb Shanno. 38

**LV**  Left Ventricle. 10, 34, 50, 60

**MAD**  Mean Absolute Distance. v, viii, 2, 27, 28, 32, 34, 50, 51, 52, 53, 55, 58, 60, 67, 68

**MR**  Magnetic Resonance. v, 1, 2, 5, 7, 8, 34, 35, 69
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRI</td>
<td>Magnetic Resonance Imaging</td>
<td>v, 1, 2, 5, 7, 8, 13, 30, 31, 34</td>
</tr>
<tr>
<td>NCC</td>
<td>Normalized Cross Correlation</td>
<td>25, 38, 41</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear Magnetic Resonance</td>
<td>5</td>
</tr>
<tr>
<td>NMV</td>
<td>Net Magnetization Vector</td>
<td>6</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal Component Analysis</td>
<td>69</td>
</tr>
<tr>
<td>PDF</td>
<td>Probability Density Function</td>
<td>25, 26</td>
</tr>
<tr>
<td>PI</td>
<td>Pattern Intensity</td>
<td>24, 25</td>
</tr>
<tr>
<td>ROI</td>
<td>Region of Interest</td>
<td>36</td>
</tr>
<tr>
<td>SONAR</td>
<td>Sound Navigation and Ranging</td>
<td>10</td>
</tr>
<tr>
<td>SSD</td>
<td>Sum of Squared Differences</td>
<td>24, 25</td>
</tr>
<tr>
<td>SSFP</td>
<td>Steady State Free Precession</td>
<td>34</td>
</tr>
<tr>
<td>TE</td>
<td>Echo Time</td>
<td>34</td>
</tr>
<tr>
<td>TR</td>
<td>Repetition Time</td>
<td>34</td>
</tr>
<tr>
<td>US</td>
<td>Ultrasound</td>
<td>v, 1, 5, 10</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 Opening Remarks and Generic Statement

Medical imaging modalities, including CT, MRI and US are critical for the diagnosis and progress monitoring of many cardiac conditions, planning, visualization and delivery of therapy via minimally invasive intervention procedures, as well as for teaching, training and simulation applications.

MRI is the preferred modality for acquiring high-quality, high-resolution images of the heart, not only due to its non-invasive aspects (i.e., no radiation dose delivered to the patient, as it is the case with CT imaging), but also thanks to its superior soft tissue contrast achieved without the use of contrast agents.

For most clinical applications described above, not all imaging data is necessary to achieve the desired, specific task. As an example, for assessment of the left ventricular function, the ejection fraction is computed by estimating the endocardial blood pool volume enclosed by the left ventricle during the diastole and systole phase. Hence, a computer-aided diagnosis technique intended to assist with estimating the left ventricle ejection fraction should focus on the left ventricle anatomy.

Image segmentation is a processing technique that allows the user to extract the necessary information from an image dataset 3D or 4D (3D + time), in the form of a surface model of the region of interest from the anatomy. Such virtual models of the anatomy could be used to perform quantitative physiological measurements based on the patient-specific data, visualize the internal anatomy to help plan or guide a therapeutic intervention, or even
teach anatomy concepts using virtual animation tools that replace the traditional textbook approach. A wide variety of segmentation techniques have been developed and implemented for cardiac MR images. Despite their complexity and performance, many of them are intended for specific image datasets or are too specific to be employed for segmenting classical, clinical quality MR images.

In this project the combination of several traditional segmentation techniques (i.e., graph cut based approach) with the use of non-rigid image registration to improve the extraction of cardiac features from MR images of a specific cardiac chamber are used, while maintaining a balanced trade-off between user interaction, computational performance, and segmentation accuracy. Graph cut based methods have been shown to work well in regards to medical image segmentation; moreover, they are computationally efficient which, scales well to real time applications. While the basic graph cut algorithms use lower-order statistics, the incorporation of segmentation estimates from another source such as that obtained through registration with an average “atlas” of the heart might help improve segmentation accuracy.

At first a global atlas-based segmentation will be implemented and its performance tested on the already available STACOM 2011 [24] MRI datasets of both healthy volunteers and patients. As a next step, the use of non-rigid image registration will be explored to determine if the performance can be improved over an affine only atlas. Lastly, we will rely on an open-source database of cardiac images (healthy volunteers and patients) to identify several parameters of variation across the cardiac anatomy and use machine learning to train an atlas which then constrains the estimate from our graph cut-based segmentation algorithm to incorporate these variations into the segmentation of new images.

We will test the proposed technique at each step during the development by assessing the segmentation results against those obtained using traditional segmentation techniques previously developed by other groups. To ensure consistency, we will rely once again on a database of available heart images from the STACOM 2011 Left Ventricle Segmentation Challenge [24], that have already been segmented using previous approaches, as
well as manually segmented by experts (expert manual segmentation is considered closest to gold-standard). Several metrics will be used to quantify the performance of the proposed technique, including computational performance, segmentation accuracy (estimated via area and volume measurements in 2D and 3D respectively), as well as segmentation fidelity assessed via the DSC correlation metric, MAD and HD and as compared to the gold standard.

1.2 Cardiac Anatomy and Physiology

![Cardiac Anatomy](https://via.placeholder.com/150)

The heart is a critical organ, whose function is to pump oxygenated blood throughout the body. It sits within a fluid filled pericardial cavity, which is surrounded by a special membrane called the pericardium. The chief function of this cavity is to act as a fluid filled lubricating membrane in order to prevent friction [61] between the moving heart and the surrounding fixed organs, to hold the heart in position.
The heart comprises four chambers — two ventricles and two atria. Deoxygenated blood returns to the right side of the heart from the venous circulation system [46] and is pumped into the pulmonary circulation for gas exchange in the lungs. After oxygenation, blood travels through the pulmonary veins into the left atrium, then into the left ventricle. The left ventricle then contracts to eject the blood into the systemic circulation through the aorta, then relaxes and refills with a fresh blood supply (diastolic phase), preparing for the next contraction (systolic) cycle. The atrio-ventricular septum separates the two sides of the heart [46], allowing no direct communication between the two sides of a normal healthy heart. The only way for blood to travel across the two sides of the heart is through the lungs, whose chief function is gaseous exchange. In spite of no direct communication, the two chambers, right and left atrium, and the right and left ventricle, contract and relax in unison.

The cardiac muscle, also known as the myocardium, is a very specialized class of muscle. The cells of the myocardium, contract through a process of depolarization which happens when the cells of the heart contract by a process of rise in intracellular calcium. Due to the electrical conductivity between neighbouring cells, depolarization of one cell leads to a domino effect, inducing in effect a wave across the myocardium that ultimately results in spontaneous contraction. This systolic phase lasts approximately a quarter of a second [62], but is highly heart rate dependent. All the timing activity in the heart is controlled by a group of cells at the sino-atrial node, which is located in the right atrium. A diagrammatic representation of the heart can be seen in Figure 1.1.

Since the left ventricle plays a major role in the cardiac function, a great deal of research has been dedicated to evaluating its function function and modeling its contractile and electrical activity. To enable the development of such models, the processing of imaging data is critical to extract the left ventricle geometry and parametrize it as needed. Hence, in this work we focus on the development of techniques to extract the left ventricle anatomy from cardiac images. Similar techniques can be applied to extract the geometry of the right ventricle, an equally important task; however, for the purpose of technique evaluation and in light of the available data on the assessment of
left ventricle segmentation, we resort to the left ventricle anatomy.

### 1.3 Cardiac Imaging Modalities

The three most utilized imaging modalities used to acquire ‘snapshots” of the heart minimally or non-invasively are Ultrasound (US), Computed Tomography (CT) and Magnetic Resonance Imaging (MRI). Each modality is based on different image acquisition principles: US builds images based on the amplitude and travel time of sound waves transmitted into and reflected from the tissue; CT images tissue density and builds an image based on the extent of attenuation of the X-ray transmitted through the tissue; lastly, MRI leverages the response of the tissue to externally applied and manipulated magnetic fields and their effect on the precession frequency of phase of the protons that make up the tissue and the magnetic properties of their surrounding environment. The typical imaging axes relevant to cardiac imaging can be seen in figure 1.2.
1.3.1 Magnetic Resonance Imaging

MRI is based on Nuclear Magnetic Resonance (NMR) [31], which is in turn based on the physical quantum mechanical property of spin. A moving electric charge produces a magnetic field. A proton in an atomic nucleus has a certain spin associated with it, and a spinning proton gives rise to a certain magnetic field. Since the human body contains a relatively large proportion of water, MRI for medical applications uses this signal from hydrogen atoms within the body to generate an image [64].

Under the influence of an external magnetic field ($B_0$), the proton’s magnetic moments align both in parallel and anti-parallel, with a slight majority being parallely aligned along the direction of the magnetic field and undergo precession at a characteristic speed called Larmor frequency in direct proportion to the strength of the external magnetic field. This is given by the equation

$$\omega_0 = \gamma_0 \cdot B_0,$$

where $\omega_0$ is the Larmor frequency in MHz (Megahertz), $\gamma_0$ is the gyromagnetic ratio (42.58 MHz/T for protons) and $B_0$ is the magnetic field strength measured in T (Tesla) [31].

Under the influence of this external magnetic field, the whole system eventually moves into a steady state, but as the system settles into a stable
state, magnetization along the Z-axis, also known as longitudinal magnetization ($M_z$) builds up because of the aligning parallel individual magnetic moments, which are only slightly more than the anti-parallel aligned moments, and this small energy difference is known as the Net Magnetization Vector (NMV) [64]. As the external field strength increases so does the NMV.

An electromagnetic wave of the same frequency as the Larmor frequency is applied to this system in a stable condition to induce a resonance condition. This results in the excitation of the system and results in the longitudinal magnetization to eventually get tipped perpendicular to the direction of the external magnetic field because of the synchronized precession of protons. [21] The resulting magnetization ($M_{xy}$) vector now lies in the xy-plane instead of along the z-axis. This time the perpendicular tipped magnetization starts to precess about the z-axis, which in turn induces an alternating potential across the receiver coil at the Larmor frequency. This is the MR signal that is then processed, first by digitization and then further processed into an image [30]. The main components of an MRI scanner are the magnet, the radio-frequency coils and gradient coils, which are used to vary the magnetic fields and allow spatial encoding of location through variations in frequency and phase. The external magnetic field is typically 1.5T or higher, roughly 30,000 times greater than earth’s magnetic field [48].

Several advantages including high image quality, non-invasiveness, soft tissue contrast, versatility, and lack of exposure to ionizing radiation have clinically established MRI as a preferred imaging modality for cardiac diagnosis, as well as subsequent planning and recently intra-operative guidance of cardiac interventions. MRI is commonly employed in clinical practice for deriving important Left Ventricular (LV) parameters such as mass, volume, and ejection fraction, as well as to generate high quality anatomical models of the anatomy for use in the planning and guidance of minimally invasive interventions under the assistance of image guidance. One example of such intervention that requires detailed knowledge of the LV muscle anatomy is endomyocardial biopsy. During this procedure, the clinicians use image guidance to extract samples of the LV muscle using a biopsy catheter, and therefore require precise knowledge of the diseased region within the global
LV representation. Typically these diseased regions are identified using cardiac MRI, and hence the need for fast and robust segmentation approaches for extraction of the LV geometry is critical for the success of the procedure. Moreover, since the MR image acquisition technology has evolved to the extent that it enables the acquisition of peri-operative (just in time) images of the patient within the interventional suite moments before the procedure, the need for fast and accurate segmentation approaches has been increasing.

The standard imaging plane for cardiac MRI is generally perpendicular to the long axis (apex-to-base direction of the heart). Depending on the acquisition protocol, the blood pool may appear is bright, whereas the myocardium appears dark, hence the name “white-blood MRI”, similar to the datasets used for analysis in this thesis. There is a large amount of variability between patients when it comes to the size, shape, image contrast, and brightness in each image. There are artifacts caused by respiratory motion as well. Figure 1.4 shows the various labeled regions within a typical short-axis cardiac MR image. The right ventricle, left ventricle, endocardium, myocardium, papillary muscles and epicardium are labeled in this view. Long axis views are not typically used for diagnostic purposes, other than for calibration purposes. Such a long axis view of the heart an be seen in Figure 1.3b.

1.3.2 Computed Tomography

Cardiac CT is a non-invasive imaging method that acquires minimally invasive images of the internal anatomy via the use of X-rays [6]. The principle behind CT relies on the the detection of the X-rays intensity passing through the anatomy to measure the “material density” and acquire the image of the anatomy. Electrons are beamed onto a metallic surface, where they rapidly decelerate upon encountering the metal object, causing rapid deceleration and the emission of radiation. The electron source is basically a point source from which the x-ray beam is emitted in the form of a cone type beam [48].

In CT, this emitted beam is collimated such that it “passes” through the patient in the form of a slice. Absorption occurs via the Compton effect [14].
The intensity of the X-ray beams is measured before patient entry and upon patient exit and the attenuation coefficient is measured. This attenuation is based on the Beer-Lambert law, which states that amount of scattered beam elements (photons or electrons) is proportional to the intensity of the beam for a particular “slice” expressed using the equation below

\[ E(x) = E_0 \cdot e^{-\mu x}, \]  

(1.2)

where \( E_0 \) is the initial number of electrons in the beam, and \( x \) is the thickness of the target object, in this case a slice of the patient, and \( \mu \) is the linear attenuation coefficient, which is dependent on both the energy in the beam and the material being imaged [1].

The main factor that differentiates CT from X-ray imaging is that the acquisition X-ray beam source rotates around the patient in a gantry, thus enabling image capture at different angles with respect to the patient. The rotation enables the receiver on the other side of the x-ray source to detect absorption at different angles around the patient. The 3D X-ray projection information is then used to construct a 3D depth image for that volume. This reconstruction requires computational power, and can be performed
using backpropagation [29, 38, 48, 56]. A large number of attenuation measurements are computed at different acquisition angles using the equation 1.2 and are subsequently processed to create an image [34]. An example of an axial CT slice can be seen in Figure 4.1a.

### 1.3.3 Ultrasound Imaging

US is an acoustics-based imaging modality. When used for the specific purpose of cardiac imaging it is also called *echocardiography*. The principle behind ultrasound is similar to that of echolocation used by bats and whales, or Sound Navigation and Ranging (SONAR) in submarines, where reflected sound is used to estimate the presence of surfaces. The bio-effects of SONAR on marine creatures constituted the idea that inspired the use of ultrasound for medical applications [60]. Sound is reflected back to varying degrees from the interfaces between two media. Similarly, sound transmitted through tissue is reflected at the interfaces between different tissue types. By measuring the time it takes for the sound to be reflected, the depth of tissue can be estimated. US uses frequencies far beyond the human range of hearing typically between 2 - 15Mhz [17].

Ultrasound does not transmit well through bone and air, on account of the extreme low and high densities associated with air and bone, respectively. As such, care must be taken to direct the transducer to an appropriate location, so as to avoid these structures, which results in specific locations where the transducer must be placed in order to get an accurate view of the organs under scrutiny. On account of this several organs such as the brain cannot be imaged using US, on account of the presence of high density bone in those areas. However, there are some ways to try to image soft tissue through bone [8].

Ultrasound is particularly well suited to vascular, cardiac, circulatory, acquiring images of moving objects — the heart muscle and valves, blood flow and turbulence because of the easy flow measurements possible on account of the Doppler effect from reflected sound waves [48]. An example of an echocardiogram can be seen in figure 1.3c.
1.4 Imaging Applications

1.4.1 Computer-aided diagnosis

One of the most common applications that relies on the accurate segmentation of the left ventricle is computer-aided diagnosis of cardiac function. The assessment entails knowledge of both the blood pool, as well as the myocardial mass, as it was demonstrated that typically abnormal cardiac function exhibits low ejection fractions (ratio of the stroke volume to end-diastolic volume), low end-diastolic volumes, and also quite large myocardial volume (mass). This behavior is explained by the fact that the heart tends to compensate for the low ejection by thickening the muscle. As such, accurate segmentation of the left ventricle muscle and inherently the blood pool will provide a quantitative measure of the end-systolic and end-diastolic blood pool volume, ejection fraction, and Left Ventricle (LV) myocardial mass, providing the clinicians with more consistent, less variable and less subjective parameters for cardiac function assessment.

1.4.2 Image Guided Interventions

Another class of applications that entail the use of the extracted cardiac anatomy from high quality pre-operative images are image-guided procedures. These procedures focus on the minimization of the access inside the chest to perform cardiac procedures, and, to counteract for the lack of direct visualization, most procedures rely on the combination of pre- and intra-operative data for achieving the necessary visualization. Moreover, since the clinicians do not need access to the whole extent of the heart portrayed in the pre-procedural, diagnostic images, the extraction of relevant information pertaining to the appropriate part of the cardiac anatomy is critical. Moreover, during procedures, the pre-procedural images and models generated via segmentation, are typically registered to the patient and fused with images acquired during the intervention, such as 2D and 3D ultrasound or 2D X-ray images. As a result, the pre-procedural models of the heart chambers are used to provide the big picture for the surgical instrument to target navigation; once on target, the intra-operative images are
used for accurate instrument localization on target for accurate therapy delivery. Since the pre-procedural models are not directly used to guide the procedure, but rather to aid the clinician with the navigation stage, their accuracy is not the most critical for the success of the procedure and will most likely not compromise the procedure outcome, but if well-constructed and well-registered to the intra-operative environment, these models have the potential to streamline procedure workflow by providing more intuitive visualization.

1.4.3 Teaching, Simulation and Training

In addition to their utilization for diagnosis and procedure guidance, 2D, 3D and 4D (3D + time) medical images, since widely available and routinely acquired, could be employed to generate didactic materials employed for teaching, simulation and training. The visualization and interactive manipulation of 3D medical image models would provide a more immersive access to the biomedical and anatomical data than currently available didactic materials, such as traditional books or atlases. From a simulation perspective, virtual and augmented anatomical environments can be developed and utilized as simulation platforms for training of residents and fellows. In addition, the virtual models can be translated into real, life-size physical models thanks to the current available 3D printing technologies, providing clinicians with access to the pre-procedural diagnostic and planning data prior to the intervention, allowing them to identify the best course of therapy and trajectory access to the surgical targets.

1.5 Overview of segmentation methods

Various segmentation techniques have been proposed in the literature [49] including some based on weak priors or no priors at all such as thresholding, edge-based and region based approaches or pixel-based classification, and others based on strong prior such as shape prior based deformation models, active shape and appearance models, and atlas based approaches. The image-based approaches require user interaction for proper segmentation of
the ill-defined regions, whereas the constraints imposed using the strong priors can overcome these problems, yet at the expense of manually building a training set. The atlas-based segmentation models constructed using a training set that is sufficiently representative of all possible heart shapes can nevertheless yield highly accurate segmentation. Image Segmentation is an important research area in biomedical image analysis.

In medical imaging, the extraction of three-dimensional static or four-dimensional (3D + time) information is important to facilitate visualization tailored to a particular task in diagnosis, quantitative measurements, for diagnostic purposes, as well as for surgical intervention, monitoring and therapy planning.

A large body of work has been conducted on medical image segmentation for MRI. Overall, these methods can be divided into two classes: the first class comprises low level image segmentation techniques, which use features from the image to define regions and classify voxels into different tissue types. The various methods that fall within this category are thresholding, region growing, clustering methods, classifiers, and morphology based segmentation.

Model Fitting Techniques incorporate prior information into the segmentation approach, the accuracy and robustness of the segmentation procedure is improved. There are different types of prior information such as those based on shapes, and topology, object appearance and topology, and from expert annotators. The major difference between most model fitting techniques is that they differ in the way prior knowledge is integrated.

The second class consists of model-based approaches. These approaches aim to fit a model to the given image data. A few examples include active contours, level sets, active shape and appearance models, and registration-based approaches that require the use of an atlas to instantiate new image datasets.

The current direction of recent research seems to favor hybrid approaches that make use of an ensemble of techniques to perform the required segmentation tasks. Model based approaches are still relevant and competitive. Since there exists a large volume of literature on medical segmentation, this
section will only touch upon a brief overview of popular types of segmentation, relevant to cardiac imaging.

1.5.1 Classification and Clustering

Image classification is a pattern recognition technique often used in segmentation tasks. Specifically, voxel classification is one of the low level segmentation techniques that can be used for image segmentation.

Features for classification may include information based on appearance, as well as location of the voxel. This space feature is used to train a classifier. The classifier may be supervised or unsupervised. For supervised classification labelled imaging data are used to learn the distribution of various features in the spatial feature.

Unsupervised classification lacks the presence of a spatial feature. These methods group the voxels based on a predetermined number of clusters by analyzing the distribution of voxels in the spatial feature. Clusters can be found by minimizing the intraclass distances and maximizing the interclass distance.

The advantage of supervised methods is that prior information specific to that particular method can be incorporated into the segmentation task. The eventual segmentation depends on the quality of the training set, which in turn, depends upon the annotator who labelled the image dataset. The training stage is tedious and each time the imaging data changes, retraining is needed.

The classification methods can either be subdivided into parametric and non-parametric approaches. The gaussian mixture model is a frequently used parametric model that makes use of the parameterized probability density functions of the different classes. This model assumes the classes are distributed in feature space. The parameters describing the gaussian mixture model can be determined in a supervised or unsupervised manner. An example of an unsupervised method is the k-Nearest Neighbour (kNN). In the kNN method voxels are classified in feature space depending on the location of its k-nearest voxels in feature space. The voxel is then assigned to the most frequently occurring label among the neighbors. This approach
also lends itself naturally to probabilistic image segmentation.

The important issues for the classification performance of these systems are feature extraction, feature selection, and the choice of a classifier, which are all active areas of research in pattern recognition and medical image segmentation. Further references on recent low level techniques can be found in the works of Bishop [2] and Duda et al. [20].

1.5.2 Active Contours

The active contours segmentation method, also known as the snakes method, belongs to a category of methods called deformable models [36]. It allows a curve defined in the image domain to evolve under the influence of internal and external forces. Generally, the internal force is nothing but a method for smoothness control of the contour, while the external force is derived from the image and drives deformations of the contour. A function that detects the external edge generally constitutes the external force as well. The active contour model uses an energy minimization approach i.e, it tries to seek the lowest energy of an objective function. The total energy can be represented as the sum of two energies [35]

\[ E_t = E_i + E_e, \]  

where \( E_i \) denotes the internal energy incorporating prior knowledge such as smoothness or a particular shape. \( E_e \) denotes the external energy and is a measure of how well the formed curve is close to the image data, and \( E_t \) is the total energy.

The curve itself can be denoted by \( c(s) \) and can be represented by the \( x \) and \( y \) measures as

\[ c(s) = [x(s), y(s)], \quad 0 \leq s \leq 1. \] 

The internal and external energies can be mathematically denoted using

\[ E_i = \int_0^1 E_i[c(s)]ds \]

\[ E_e = \int_0^1 E_e[c(s)]ds \]
and

\[ E_{in}(v(s)) = \alpha(s) \left\| \frac{dv}{ds} \right\| (Elasticity) + \beta(s) \left\| \frac{d^2v}{ds^2} \right\|^2 (Stiffness). \]  (1.6)

Although the active contour model is an important work in segmentation, it has its inherent limitations. It is sensitive to the initialization, because if the initialization is incorrect it may get stuck in a local minimum and fail to converge to a global minimum solution. However, there are ways in which the snakes can be made less sensitive to initialization, particularly the work by Cohen et al [13]. The performance in the presence of noise is degraded causing the contour to pass through the field of view of the image when the image has high amounts of noise. In addition, the accuracy of the active contour model depends upon the convergence criteria employed in the minimization technique.

1.5.3 Level Set Segmentation

1.5.3.1 Conventional Level Set

Level sets was introduced by Malladi et al. [42] and Caselles et al. [9], based on the work by Osher and Sethian [47]. The evolving contour is defined as the zero level set, generally of a higher dimensional function. The level set method can update topology while the changes in the evolution embedded function remains well defined. They can be readily deployed in discrete grids and arbitrary dimensions.

The underlying idea of level set-based segmentation is to evolve the embedding function such that its zero level set captures the object of interest. Updating the embedding function is a costly procedure. The performance of the algorithm can be improved by using a narrow band implementation, in which the update occurs in a small region around the zero level set. The narrow band needs to be reinitialized if the level set approaches the boundaries of the narrow band.
1.5.3.2 GeoDesic Level Sets

The major limitation of this formulation compared to the active contour approach, lies in the fact that there is no energy term to minimize the evolution process. The energy minimization approach can be used to form an implicit deformable model, by evolving the zero level set such that a geodesic curve or surface in image space is found, where the distances im image space are defined on the basis of image content in the image to be segmented [9].

A modified metric is introduced, such that the state of minimal energy corresponds to the desired object segmentation. Distances along the zero level sets decrease at higher gradient magnitude, so that the states of minimum energy tend to occur at high values of the gradient.

1.5.4 Shape and Appearance Models

During the last decade, model-based segmentation techniques that incorporate prior statistical information such as shape, appearance, motion and deformation have experienced enhanced interest and popularity. The \textit{a priori} knowledge is “learned” from a set of “training examples” and is typically expressed as an average and a set of characteristic variations. These models greatly improve robustness to noise, artifacts and missing data, and hence the reliability of image segmentation is improved.

There are many different types of statistical shape models that have been proposed. Some of the most popular are active shape models (ASM) [16, 39], active appearance models (AAM) [15, 43], models based on Fourier parameterization [50], and statistical deformation models [55].

1.5.5 Atlas-based Methods

Atlas-based approaches transform the segmentation problem into a registration problem [7] and can be categorized into four strategies, depending on the atlas selection [52]. The simplest approach is to select a single individual atlas from the training set and propagate the label to the test image via registration. The segmentation can be improved by selecting the training image
most similar to the test image (e.g., based on similarity following affine registration) to serve as atlas. Furthermore, an average atlas, which is similar to most of the individuals comprising the training dataset population, can be created from the training set such that the test image requires very little deformation to achieve a good match to the atlas, hence leading to improved segmentation accuracy. Lastly, multi-atlas based segmentation approaches can be combined using a classifier approach by trading off a more accurate segmentation for multiple computationally expensive registrations and label fusion.

1.5.6 Graph cut-based Energy Minimization Methods

Graph cuts based energy minimization techniques are popular in computer vision [3]. The segmentation problem is approached by an energy minimization approach, that helps to find a globally optimized segmentation. The graph is constructed by representing pixels or voxels as edges and some cost associated with them as the edges for optimization. The edges can only be assigned non-negative costs. The edge costs are decided based on some similarity or dissimilarity criteria between the voxels or pixels, or can be user input in the form of seeds in case of an interactive graph cut method. A minimum cut is computed through this graph representation that renders the graph into two disjoint sets which represents the segmentation of the image. The specific type of functions that can be optimized via graph cuts have been studied in detail by Kolmogorov et al [37]. Graph cuts is discussed in more detail in 2.6.1.

1.6 Image Registration

Although it may seem counter-intuitive to dedicate a section to image registration given the focus of the work on segmentation, image registration is an integral component of image segmentation, especially for atlas-based techniques, as well as for temporal any applications that entail the segmentation of the heart at a single cardiac phase and its propagation to subsequent
phases using motion information extracted by registering the multi-phase image frames.

Image registration was used to align two different views of an object in an image [41]. The process entails the deformation of one image to be similar to another image, such that the dissimilarity between them is minimized. The image being is referred to as the moving image, while the target image towards which the moving image is deformed is referred to as the fixed image. The process of minimizing the distance between the moving and fixed images is done through the use of transformations, which mathematically transform points in the moving image towards the fixed image [22]. Depending on the deformations they induce, these transformations can be either global or local, and operate either at the intensity level (i.e., align images such that the dissimilarity between the registered moving and fixed image is minimized in terms of the image voxel intensity) or at feature level (i.e., the distance between features in the registered moving and fixed image is minimized).

The registration process can be broken down into several components [48],

1. Some form of metric to measure the disparity between two images.
2. The way in which the moving image is transformed to match the target - rigid, affine, non-rigid, etc.
3. Some kind of optimization with regularization that tries to minimize the dissimilarity between two images.

1.6.1 Geometric Transformations

In the cartesian co-ordinate system a point is represented by its position $P(x, y)$ in a plane or $P(x, y, z)$ in space.

1.6.1.1 Translation

Translation is simply the displacement of a point to a new location by a certain amount in each coordinate system direction denoted by $P(x+T_x, y+$
Figure 1.5: From left to right: the original image, rigidly transformed image, similarity-transformed image, affinely transformed image, and non-rigidly-transformed image.

$T_y, z + T_z$ in three dimensions, where a point in space is said to have been translated by $T_x, T_y$ and $T_z$ respectively

$$
\begin{bmatrix}
1 & 0 & 0 & T_x \\
0 & 1 & 0 & T_y \\
0 & 0 & 1 & T_z \\
0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
z \\
1
\end{bmatrix}
= 
\begin{bmatrix}
x + T_x \\
y + T_y \\
z + T_z \\
1
\end{bmatrix}
$$

(1.7)

1.6.1.2 Scaling

An image or object in space can be scaled to a different size when its coordinates are multiplied by a suitable constant for each dimension, i.e, scaling along $x$, $y$ and $z$.

$$
\begin{bmatrix}
S_x & 0 & 0 & 0 \\
0 & S_y & 0 & 0 \\
0 & 0 & S_z & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
z \\
1
\end{bmatrix}
= 
\begin{bmatrix}
S_x \cdot x \\
S_y \cdot y \\
S_z \cdot z \\
1
\end{bmatrix}
$$

(1.8)
1.6.1.3 Shearing

Similarly, a shearing operation can be equated to an example such as the conversion of a cube shape to a parallelepiped [18].

\[
\begin{bmatrix}
1 & Sh_{xy} & Sh_{xz} & 0 \\
Sh_{yx} & 1 & Sh_{yz} & 0 \\
Sh_{zx} & Sh_{zy} & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
z \\
1
\end{bmatrix}
= \\
\begin{bmatrix}
x + Sh_{xy} \cdot y + Sh_{xz} \cdot z \\
Sh_{yx} \cdot x + y + Sh_{yz} \cdot z \\
Sh_{zx} \cdot x + Sh_{zy} \cdot y + z \\
1
\end{bmatrix}
\]  

(1.9)

1.6.1.4 Rotation

Let us assume that \( R_x, R_y \) and \( R_z \) represent the rotation of an object or image about the x, y and z axes respectively. Therefore the rotation about the three axes can be computed using the transformation matrices listed below, where \( R_1, R_2 \) and \( R_3 \) denote the rotation transformations along the x, y and z axis, respectively.

\[
R_1 = \\
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & \cos(R_x) & -\sin(R_x) & 0 \\
0 & \sin(R_x) & \cos(R_x) & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\]  

(1.10)

\[
R_2 = \\
\begin{bmatrix}
\cos(R_y) & 0 & \sin(R_y) & 0 \\
0 & 1 & 0 & 0 \\
-\sin(R_y) & 0 & \cos(R_y) & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\]  

(1.11)

\[
R_3 = \\
\begin{bmatrix}
\cos(R_z) & -\sin(R_z) & 0 & 0 \\
\sin(R_z) & \cos(R_z) & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\]  

(1.12)

The global rotation matrix \( R_m \) can then be concatenated into its final form denoted by

\[
R_m = R_1 \cdot R_2 \cdot R_3
\]  

(1.13)
1.6.1.5 Rigid, Similarity and Affine Transformations

3D Rigid Transformations have typically six degrees of freedom, three translation degrees of freedom, and three for rotation. By adding a scaling component, 3D rigid Transformations have nine degrees of freedom, three from translation, three from rotation and three additional ones from the scaling. Lastly, by allowing shearing along the three axes, affine transformations can be constructed, which involve twelve degrees of freedom, nine of which arise from the similarity transformation and the additional three arise from the shearing.

1.6.2 Global Registration

Global registration is typically used as a precursor to non-rigid registration. The most common form of global registration utilized is the affine transformation, which attempts to give a more coarser overall alignment between two objects of interest in images. Additionally, non-rigid or local deformation takes care of the residual local deformations required to complete alignment between two objects of interest in an image [1].

1.6.3 Local Registration

Local registration, also called non-rigid registration is usually carried out as part of a multi-step process with affine or global registration., wherein a global registration is performed at first for general alignment and then a more finer non-rigid registration performed, wherein the non-rigid registration may or may not have the same optimizers. Non-rigid registration works through deformation, where we need to specify how each image unit moves. A deformation field defines where image elements move, usually denoted by a vector that defines direction and amplitude of displacement. It can be described as a model that defines deformation behavior [1].

The image may be divided into a grid wherein each local region is matched to the other region, based on an optimization function. Some of the non-rigid registration methods are introduced below.
1.6.3.1 Piecewise Rigid

Piecewise rigid registration entails the sub-division of the image into sub-regions, which are separately registered rigidly, with the resulting differences resolved through volume regularization. Piecewise rigid are also called featurelet based approaches.

1.6.3.2 B-spline

The images are divided into a grid or mesh, where each point in the grid can be thought of as a point for a smooth function. This kind of model for a smooth function is called a spline. A optimization function determines how much force is applied to each spline, and it models deformation in a smooth manner, but huge local deformation will affect the overall deformation in other areas of the image. This is analogous to a wooden spline, wherein the way the pressure is applied on the spline is equally important to where the spline is pivoted [1]

1.6.3.3 Linear Elastic Models

Elastic models [44] were initially studied by Broit [4]. Linear elastic models are only suitable for a very small local deformation. Given a very small deformation, the surface undergoing deformation, can be said to be similar to an elastic solid. A force is applied on the said analogous elastic solid until stability is achieved. There is an underlying assumption of linearity because of the small deformations involved. The advantage of this linear model lies in the fact that non-local effects should not occur like in the case of the B-spline deformation model.

1.6.3.4 Finite Element Modelling (FEM)

This type of deformation modeling more closely tries to mimic the structure of biological organs. Soft tissue could me modeled as some kind of viscous fluid or elastic model, and harder objects such as bone could be treated as rigid bodies. This could be very close to a realistic model, taking unto account well defined properties on the interfaces of such regions. The
drawbacks are intensive computation required for solving partial differential equations and defining a complex mesh, which require the object of interest to be fully segmented.

1.6.3.5 Viscous Fluid Models

A viscous fluid model is similar to a linear elastic model, in that it is less constrained than an elastic model. Strong local deformations are permitted in a viscous model, similar to what one might expect when a fluid is deformed. However such large deformations could result in a catastrophic misalignment in registration because of its relatively uncontrolled nature, if the required parameters are not set correctly.

1.6.4 Similarity measures

As mentioned earlier, when two images are registered, a similarity metric is being optimized, such that the differences between the registered moving and fixed images are minimized. Depending on the images at hand, as well as the modality with which they were acquired (i.e., intra- vs. inter-modality), different similarity metrics are utilized.

1.6.4.1 Sum of Squared Differences

Sum of Squared Differences (SSD) is an intensity based image similarity measure, which is rather simple. The SSD is given by,

\[
SSD = \frac{1}{N} \sum_{N} (I_{fixed}(x, y, z) - I_{moving}(x, y, z))^2,
\]

where \(N\) is the number of voxels, \(I_{fixed}\) is the fixed image, \(I_{moving}\) is the moving image. The equation 1.14 shows the difference over each voxel location in the images, therefore the greater the misalignment between two images the greater will be the total sum of squared differences between the two images. This metric works best for images which belong to the same modalities, and have similar histograms and intensity.
1.6.4.2 Pattern Intensity

For intramodal images, another metric is defined as Pattern Intensity (PI). This requires the use of a difference image, denoted $I_d$.

$$PI = \sum_{x,y,z}^{N} \sum_{d^2 \leq \sigma^2} \frac{\sigma^2}{\sigma^2 + (I_d(x, y, z) - I_d(p, q, r))^2}$$ \hspace{1cm} (1.15)

where,

$$d = (x - p)^2 + (y - q)^2 + (z - r)^2.$$ \hspace{1cm} (1.16)

In the following equation,

$$\sigma = \text{Scaling Factor (Internal)}$$ \hspace{1cm} (1.17)

d is the diameter surrounding each voxel, therefore a similar surrounding implies a good match. PI provides a measure of local differences in intensity values of the difference image, within a given radius $r$ [1]. The internal scaling factor $\sigma$ is used for gradient control. Like SSD, it also requires a good match in image histogram content and intensity similarity. PI is suitable when the images being matched are already relatively close enough.

1.6.4.3 Normalized Cross-correlation

Normalized Cross Correlation (NCC) is an intramodality similarity measure derived from Pearson’s correlation coefficient [1], and can be defined as

$$NCC = \sum_{i=1}^{N} \frac{(I_i^{fixed} - \overline{I^{fixed}})(I_i^{moving} - \overline{I^{moving}})}{\sum_{i=1}^{N}(I_i^{fixed} - \overline{I^{fixed}})^2 \sum_{i=1}^{N}(I_i^{moving} - \overline{I^{moving}})^2}$$ \hspace{1cm} (1.18)

where, $\overline{I^{fixed}}$ is the mean (expected value) of the fixed image, likewise, $\overline{I^{moving}}$ is the mean (expected value) of the moving image.
1.6.4.4 Mutual information

Mutual Information (MI) and Normalized Mutual Information (NMI) are typically used for multi-modality registration, however they can also be used for intramodal registration. It is designed to measure statistical dependence between two random variables, and is not affected by the tint or the intensity level of the pixels. A Probability Density Function (PDF) is computed for each image [1]. Mutual information is defined as

\[
MI = E(P(I_{moving}, I_{fixed})) \ln \frac{P(I_{moving}, I_{fixed})}{P(I_{moving}), P(I_{fixed})},
\]

(1.19)

where,

\(E(P(I_{moving}, I_{fixed}))\), is the expected value of the joint PDF - \(P(I_{moving}, I_{fixed})\), \(P(I_{moving})\) and \(P(I_{fixed})\) represent the PDF’s for the intensity values in the moving and fixed image respectively [1].

1.6.5 Optimizers

Two popular optimization methods relevant to medical image registration are the simplex [45] and quasi-Newton family of optimizers.

1.6.5.1 The Simplex Optimizer

The simplex or Nelder-Mead method for function minimization was proposed by John Nelder and Roger Mead in 1965 [45]. A simplex is analogous to an n-dimensional triangle [1]. This is also called the Nelder-Mead algorithm and is good at minimizing a nonlinear function [58].

1.6.5.2 The Quasi-Newton Optimizer Family

Quasi-Newton methods use the gradient measurement at each iteration to build a quadratic model of the objective function such that superlinear convergence is achieved. Unlike Newton methods, quasi-Newton methods do not require the computation of Hessians which makes this less expensive
in certain situations [32]. One of the popular quasi-Newton methods is the Broyden-Fletcher-Goldfarb-Shanno [57] [26] [23] [5] method.

1.7 Segmentation Evaluation

Segmentation results need to be compared to the gold standard or expert annotated segmentation in order to evaluate the performance. Some of the three widely used and reliable metrics to assess segmentation performance are the DSC, HD and MAD measures. DSC is a region based measure, whereas HD and MAD are contour based measures.

1.7.1 Sørensen-Dice Coefficient (DSC)

The DSC measure evaluates the result of segmentation against the gold standard by measuring the overlap between two sets of pixels or voxels. The DSC score is denoted by the following formula,

\[ D = \frac{2|S \cap G|}{|S| + |G|} \]  

(1.20)

where the absolute number of intersecting voxels are denoted by \(|S \cap G|\), and \(|S|\) and \(|G|\) are the sum of the number of voxels in the segmentation and gold standard respectively. The graphical representation of the terms in equation 1.20, can be seen in Figure 1.6.

![Figure 1.6: Representation of DSC](image)
1.7.2 Mean Absolute Distance (MAD)

The mean absolute distance denotes the mean of differences between two surfaces — typically how much on average two surfaces differ from one another. A point-to-point correspondence between two surfaces is needed to compute the mean absolute distance [25].

Assuming that we have two contours $R$ and $Q$, and given two sets of points representing the points on two contours, $R = \{ r_1, \ldots, r_j \}$ and $Q = \{ q_1, \ldots, q_w \}$, the $d(q_w, R)$ can be stated to be

$$d(q_w, R) = \min_j \| r_j - q_w \|,$$  \hspace{1cm} (1.21)

where, $r_j$ is the point on contour $R$, $q_w$ is a point on contour $Q$. These distances are computed for all points on both curves and are averaged to result in the MAD between the two curves, given in the equation below [11],

$$MAD = \frac{1}{2} \left\{ \frac{1}{n} \sum_{n=1}^{w} d(q_w, R) + \frac{1}{n} \sum_{n=1}^{j} d(r_j, Q) \right\} \hspace{1cm} (1.22)$$

MAD can also be called an $L^1$ norm that integrates both over and under estimation of a contour [10].

1.7.3 Hausdorff Distance (HD)

The HD is another metric used to assess the match between the contours of the segmentation and the gold standard. It can be defined as the largest of shortest distances between two contours. Given two sets of points representing the points on two contours, $R = \{ r_1, \ldots, r_p \}$ and $Q = \{ q_1, \ldots, q_p \}$, the HD can be stated to be [33],

$$D\{R, Q\} = \max (d(R, Q), d(Q, R)), \hspace{1cm} (1.23)$$

where

$$d(R, Q) = \max_{r \in R} \min_{q \in Q} \| r - q \|; \hspace{1cm} (1.24)$$
where the distance \(d(R, Q)\) represents the largest of the minimum distances from points in set \(R\) to points in set \(Q\). Eventually the larger of the two distances between \(d(R, Q)\) and \(d(Q, R)\) becomes the HD.

**Figure 1.7** shows that the smallest distance from the inner circle to the outer contour can be represented by the line going towards \(B\) on the outer contour, and the largest of such minimum distances would be the line going from the inner contour towards point \(E\) on the outer contour. On the other hand, the largest of the minimum distances in a set of all points in the outer contour towards the inner circle can be seen by the line going from point \(C\) to \(A\).

The \(\| \cdot \|\) in equation 1.24 could therefore be said to represent some form of norm on the points in the sets \(R\) and \(Q\), for example the \(L2\) or Euclidean norm [33]. **Figure 1.7** visually illustrates the asymmetric nature of the HD, where the minimum distance from point \(B\) leads to \(A\), and the minimum distance from the point \(A\) leads to \(C\) [1].

![Figure 1.7: Graphical representation of HD and its asymmetry.](image_url)
1.8 Segmentation Challenges

As described in the previous sub-sections, despite the variety of available segmentation approaches, ranging from low-level image manipulation techniques, such as thresholding, to region growing methods and other high level approaches, such as level sets, graph cuts and atlas- and registration-based segmentation, no single segmentation technique can be labeled as the preferred, most appropriate method for automatically extracting anatomy of interest from a given image or set of patient images.

Low-level techniques such as thresholding and region growing are known to work well when the intensity range of the region of interest if sufficiently different from that of the surrounding tissues. As this scenario is not the case for most traditional cardiac (and not only) MRI images, these techniques may yield better results if applied following some semi-automated pre-processing. Often semi-automated image manipulation helps better identify and narrow down the range on intensities subsequently used for tissue classification in a manual rather than fully automated fashion, rather than relying on fully automated voxel classification.

Similarly, the performance of atlas-based segmentation approaches are usually limited to the variability of the datasets employed during the training stage and are highly unlikely to show satisfactory performance for datasets that are sufficiently different from the training population. In addition, another limitation of atlas-based segmentation techniques lies within the accuracy with which the mean atlas label can be propagated to a new patient image via registration. While on one hand a non-rigid deformable registration approach may be rendered optimal given the soft tissue presence, global affine-based registration techniques may provide better overall alignment between the atlas and new patient image, without the risk of getting trapped in a local minimum, resulting in incorrect label propagation from the atlas to the new dataset.

Another limitation of atlas-based segmentation and registration-based label propagation is the computational performance associated with the use of non-rigid, deformable registration. Most algorithms under-perform, requiring lengthy computing times, which may not be suitable for applications
that require to be conducted in a timely fashion, such as segmentation of structures and anatomy for intra-procedural applications.

High-level segmentation techniques such as the level sets and graph cut methods may show optimal performance, however they require laborious and significant parameter adjustment, given their high sensitivity and dataset dependency.

In addition to the uncertainties introduced by the utilized segmentation approach, a great deal of difference between the so-called “ground truth and segmented dataset arise from the inherent variability and subjectivity associated with manual tracing and segmentation of images when performed by expert clinicians. Despite the expert manual segmentation of medical images being established as the “gold standard segmentation technique, the ground truth is often “stained by the user experience and, when compared to the result of a new, proposed segmentation technique, it may render the latter as under-performing, when in fact the former suffers from inherent inaccuracies. Nevertheless, the variability associated with manual segmentation is widely accepted, along with its subjectivity, and it has become common practice to still treat the manual segmentation results as the reference against which new segmentation techniques are evaluated.

In essence, we acknowledge the uncertainties associates with both the available segmentation techniques as well as the gold standard data, and as art of this work, we propose to investigate a combination of atlas-based segmentation tools with high-level graph cut-based segmentation and assess their joint performance on a set of available cardiac MRI image datasets.

1.9 Proposed Solution and Thesis Outline

In response to the challenges outlined above, here we propose to investigate the utilization of atlases of different complexity (single-patient reference image atlas, multiple patient atlas built based on original intensity image volumes, multiple patient atlas built based on segmented ground truth volumes using either global affine and global affine + local non-rigid registration) followed by affine and non-rigid registration-based label propagation
and coupled with graph-cut based segmentation subsequently applied to the atlas initialization.

Chapter 2 details the construction of the various atlases utilized in this work, including the cardiac image database employed, the image pre-processing steps, the registration techniques employed, and the label propagation approaches. In addition, we also describe the adaptation and formulation of the graph cut-based technique, the voxel classification approach used as input to the graph cut, as well as the coupling of the graph cut with the atlas-based initialization.

Chapter 3 focuses on the evaluation of the proposed pipeline at different stages according to several metrics the Sørensen-Dice Coefficient (DSC), the Mean Absolute Distance (MAD) error and Hausdorff Distance (HD) error computed at both the epicardial and endocardial borders, computed both on a 3D-to-3D basis, as well as a slice-by-slice 2D basis and classified according to the slice location: apical slices, mid-slices or basal slices.

Chapter 4 summarizes the work, identifies some of its limitations and outlines some future direction for improvement of the current results in the context of other concurrent developments in the field.
Chapter 2

Methodology

2.1 Overview

Atlas based segmentation techniques require the construction of an atlas — whether that is a selected reference patient which could be any one of the patients from the dataset, or the use of an “average” model based on statistics acquired from all the patients in the dataset.

Average atlases have been demonstrated to better represent the variability across all the patients in the dataset, whereas the selection of any patient at random from the dataset is likely to result in an atlas that is heavily biased towards the patients most similar to the selected patient. Since an average atlas construction requires the incorporation of statistics from different patient datasets often referred to as training datasets, it is more likely that such an average atlas is uniformly closer to all the training patients and thus more likely to represent the likelihood of the presence of the left ventricle anatomy.

For most atlas-based segmentation approaches, registration is usually a two step process. First, a global registration (i.e., affine transformation) compensates for size, orientation, and translation differences between two volumes and yields an initial alignment using with a low computation. Subsequently, a local, non-rigid registration compensates for residual local variation in shape remaining over after and anatomy at much higher computational cost.

Although this two-stage strategy is known to produce good segmentation results [54], the high computation cost associated with most non-rigid registration methods is a motivating factor to investigate an alternative approach
wherein the test volume is aligned to an affine atlas using global registration and subsequently use this initial segmentation result is an input into a graph cut-based segmentation algorithm to further refine the segmentation.

The objective of this work is to compare the segmentation results of the above-mentioned two approaches against expert annotated ground truth segmentation available along with the image dataset. In support of this overall objective, two average atlas models were built using a training dataset comprising 70 images from both healthy volunteers and patients. The first atlas was built using only global affine registration, while the second was constructed using both global and local non-rigid registration. The segmentation results were tested against the expert annotated gold standard for the remaining 30 test patients.

The assessment was carried out on the basis of several metrics including DSC, HD, and MAD, in order to assess the global vs. global + local atlas-based segmentation, with the intent to further improve the global atlas-based segmentation over global + local atlas based segmentation to better replicate the ground truth using the subsequent refinement by graph cuts based segmentation methods from the initial atlas based segmentations.

### 2.2 Imaging Datasets

This study employed 100 de-identified cardiac image datasets from healthy volunteers and patients suffering from myocardial infarction and impaired Left Ventricle (LV) contraction available as a part of the STACOM Cardiac Atlas Segmentation Challenge project [24] database\(^1\). Cine-MRI images in short-axis and long-axis views are available for each case. The images were acquired using the Steady State Free Precession (SSFP) MR imaging protocol with typical thickness \(\leq 10\text{mm}\), gap \(\leq 2\text{mm}\), Repetition Time (TR) \(30 - 50\text{ms}\), Echo Time (TE) \(1.6\text{ms}\), flip angle \(60^0\), Field Of View (FOV) \(360\text{mm}\) and \(256 \times 256\text{mm}\) image matrix using multiple scanners from various manufacturers. The collated results from those submitted for this challenge are available in [59].

\(^1\) [http://www.cardiacatlas.org](http://www.cardiacatlas.org)
Expertly-annotated ground truth data is available for all 100 cases in the training set. Since the ground truth was obtained using 3D surface finite element model analyzed by experts, partial myocardium can be observed in the basal slices due to partial intersection between the model and the image slices. For the purpose of this experiment, the 100 cases were randomly divided into 70 training and 30 test cases. Hence, we evaluate our results on the provided ground truth segmentation for the 30 test cases, after optimizing the training part on the training patient dataset.

This dataset is particularly challenging because of the wide variability in the dataset between all of the patients in terms of contrast and intensity, noise artifacts, orientation, voxel resolution, and apex-base or base-apex slice arrangements. This wide amount of variability can be visualized in Figure 2.1, which shows variability with respect to $T_x$, $T_y$ and $T_z$, which represent the translation, whereas $R_x$, $R_y$ and $R_z$ represent the rotation and $S_x$, $S_y$ and $S_z$ represents the scaling along the the x, y and z-axes respectively, with the last six parameters representation shear along different directions. It can be seen that the maximum variability can be found along the z axis rotation, where the rotation along the axis is highly variable and is not centered about the reference patient. This variability is with respect to the reference patient selected at random from the training set.

Each patient had cine MR images acquired at different phases, with the varying number of phases for each patient. The cine MR images were acquired from full relaxation (diastole) to (systole) and then back to diastole. Therefore the end slices in the time series phase represent maximum size caused by diastole and the middle slices represent full contraction (systole). The slice thickness between different patients was found to be varying between 6 mm and 10 mm, accompanied by different in-plane image resolutions.

### 2.3 Data preparation and Pre-processing

Prior to atlas construction and the first registration step, all patient datasets were re-interpolated to a uniform voxel resolution of $1.5mm \times 1.5mm \times$
Figure 2.1: Training dataset variability with respect to reference patient prior to global alignment.

6mm, along the x, y and z axes, respectively, to ensure correct registration as required by the subsequent steps.

In addition, all image datasets were then cropped into a Region of Interest (ROI) centered approximately around the left ventricle, but also including some parts from the right ventricle and surrounding tissue regions. Each image volume was then normalized into a uniform volume size of 100 × 100 × 25, and cropped in a cylindrical region, so as to accommodate the varying heart sizes and to help the image registration method perform better, when having to deal with uniform image sizes throughout the range. This work does not focus on ROI extraction, but is concerned with the question of atlas generation and then using this atlas to propagate labels through registration, and determine whether the atlas-based segmentation may be improved upon by some other low cost segmentation algorithm such as a
2.4 Atlas Construction via Registration

Since the average atlas-based approach needs a single registration to obtain a segmentation of an unknown test patient and yields good segmentation results [52], here an approach is proposed that follows a similar strategy, but rather exploits the computational efficiency of global- rather than local-based atlas construction and label propagation, with the objective of minimizing computational complexity and improving efficiency without compromising segmentation accuracy.

A reference patient with good contrast was chosen from the training set. To minimize the bias of the average atlas towards the reference patient, the reference image was affinely registered to each image in the training dataset. The parameters of all the registration transforms were averaged and subsequently applied to the reference patient to obtain a globally-centered average atlas.

In the next step, all training datasets were registered to the global average atlas. The local residual deformation post-registration was computed for each patient dataset between the global average atlas and the globally registered patient datasets using the B-spline free-form deformation non-rigid registration approach proposed by by Reuckert et al. [54].

After computing the local residual deformation between the global average atlas and the globally registered training datasets, the average residual deformation field was applied to the global average atlas, yielding the average anatomical shape of the training images [27].

This registration pipeline was used to generate two atlases: the former global atlas only relied on global, affine subsequent registration between the training datasets and the iteratively-updated average global atlas. The latter atlas was generated by further refining the global atlas by incorporating the local residual deformations between the average atlas and individual training images not accounted for by the global affine transforms. As such, the combined global and local atlas entailed a series of non-rigid registration
transformations between the global atlas and the individual training datasets, with the average deformation field applied to the iteratively-updated atlas until convergence.

2.4.1 Image Registration Implementation

For this work, the registration technique based on that proposed by Reuckert et al. [54] is used. The registration was done using affine and non-rigid registration using grid points set at $29 \times 29 \times 16$, with a spacing of about $4 \times 4 \times 2$. The registration employed the Limited Memory Broyden Fletcher Goldfarb Shanno (L-BFGS) [40] optimizer and Normalized Cross Correlation (NCC) as the similarity metric.

Figure 2.2a illustrates the moving image in the left panel, and the moving image is overlayed onto the fixed image and a coarse grid is initialized on the moving image in the right panel. The moving image is affinely registered to the fixed image and the result can be observed in Figure 2.2b. This affinely transformed moving image is now globally aligned to the reference patient.

A finer grid ($29 \times 29 \times 16$) is now overlayed onto this affinely transformed moving volume, and it is non-rigidly deformed using B-spline based registration. This non-rigidly deformed volume is illustrated in Figure 2.2c. This is the final product of registration and is illustrated alongside the fixed image in the right panel of Figure 2.2d.

2.4.2 Averaging transformations

2.4.2.1 Global (Affine)

The affine components from each registration were obtained, and all affine components other than rotation are averaged. Rotation angles cannot be averaged linearly and therefore they were averaged based on circular mean which gives both the mean and variance for the angles [2] according to the

Figure 2.2: (a) Moving Image with initialized grid (b) Affinely Registered to Fixed volume (c) Non-Rigidly Registered to fixed volume at a finer grid size (d) The resulting image from non-rigid registration after affine registration

relationship below

\[ \bar{\theta} = \text{atan2} \left( \frac{\sum_{j=1}^{n} \sin \theta_j}{n}, \frac{\sum_{j=1}^{n} \cos \theta_j}{n} \right). \]  \hspace{1cm} (2.1)

After the averaging of these components, a matrix is constructed from the averaged components as explained in section 1.6.1, representing the average transformation matrix.

2.4.2.2 Local (B-spline non-rigid)

Averaging a set of local transformations is rather straightforward as compared to global transformations and is simply performed by linearly averaging the x-, y- and z-coordinates of the deformation field extracted by non-rigidly registering each image dataset to the most recently updated global average.
2.4.3  Global Atlas

A first global average atlas model was generated using the intensity based global affine transformation as follows:

1. The reference patient image Ground Truth (GT) label $A_1$ was globally aligned to all the training images $P_i$ by a global affine transformation $T_i$ from the reference patient $A_1$ to each training patient. This procedure is illustrated in Figure 2.3, where the transformations for the global atlas are denoted by $T$.

2. The average transform computed based on each dataset registration to the reference image $T_{Avg1-2}$ was then applied to reference patient. This step moves the reference patient closer to all the other patients in the dataset in
parametric space. The same transformation was also applied to the intensity volume images for the update step.

3. Next all the patients $P_i$ were aligned to the second average where the second average $A_2$, which became the fixed image. The affine transformations from this process transform voxels $x_i$ in $P_i$ as $x_i' = T_i(x_i)$, such that $P_i$ was transformed into $P_i'$ so as to maximize NCC between the two images.

4. The affinely transformed training images $P_i'$ were averaged to produce a mean intensity average image $A_3$ with the global average shape from $A_2$, which gives $A_3$, as shown in Figure 2.5.

5. The latest updated average was the final average used, representing mean shape in $A_2$ and mean shape and variance in $A_3$.

Note that following global alignment, the variability of all training datasets with respect to the average atlas is now centered about the parameters of the average atlas, indicating that the average atlas provides a realistic ‘mean” representation of the dataset in the affine parameter space spanning all training datasets (Figure 2.4).

2.4.4 Global + Local Atlas

A second incremental global + local average atlas model was generated using the GT based global affine plus local non-rigid transformations using B-spline based free form deformation as follows:

1. The reference patient image GT label $A_1$ was globally + locally aligned to all the training images $P_i$ by a deformation grid $D_i$ from the reference patient $A_1$ to each training patient. This can be seen in Figure 2.3, where the deformation grids are denoted by $D$.

2. The average deformation from all these registrations $D_{Avg1–2}$ was then applied to reference patient. This process moved the reference patient closer to all the other patients in the dataset in parametric space. The same deformation was applied to the intensity volume images for the update step.

3. The affinely transformed training images $P_i'$ were non-rigidly aligned to the second average $A_2$, yielding a non-rigidly transformed image $P_i''$.

4. The affine plus non-rigidly transformed training images $P_i''$ were averaged to produce a mean intensity average image $A_3$ with the shape of the
first average $A_2$, as shown in **Figure 2.5**.

5. The last updated global + local average represented the final average representing mean intensity, variance and deformation from the training set.

### 2.5 Atlas Based Initial Label Propagation

The generated affine atlas was registered to each of the 30 test images using affine transformations. The obtained transformations were applied to the ground truth segmentation of the affine atlas to propagate the label and obtain an approximate segmentation for all the test images.

In a similar fashion, a homologous global + local based segmentation of each of the 30 datasets was obtained by using both affine plus non-rigid
Figure 2.5: Schematic diagram illustrating the subsequent steps applied to the global atlas to generate the non-rigid atlas using deformable registration to account for residual non-global deformations.

transformation to propagate the label corresponding to the global + local-based atlas to the test images.

## 2.6 Graph cuts based segmentation implementation

### 2.6.1 Basic graph cut segmentation operation

The basic premise of graph cuts is that the input volume can be represented as a graph which each voxel within this image as a node in the said graph. The number of classes that the said volume needs to be segmented into will be the nodes on each of these graphs.

For this work, the three classes of voxels are the myocardium, blood pool and background. The problem can be addressed as finding a labeling $l$ as
Figure 2.6: Final Non-Rigid Atlas with the ground truth overlayed on the mean volume as visualized in ITKSnap

Addressed in the below equation [19]:

\[ E(l) = \sum_{\{a,b\} \in \mathcal{N}} V_{a,b}(f_a, f_b) + \sum_{a \in \mathcal{P}} D_a(i_a, l_a), \]  

(2.2)

where the first term represents the smoothness energy which forces pixels \( a \) and \( b \), defined by the set of interacting pair of pixels \( \mathcal{N} \), towards same label, and the second term represents the data energy that reduces the disagreement between the labeling \( l \) and the observed data \( i_a \) [19].

The data term weights the link between each pixel and the terminals, known as \( t \)-links, and is formulated as a negative natural logarithm of the normal distribution [19]

\[ D_a(i_a, l_a) = -\ln \left( \frac{1}{\sigma \sqrt{2\pi}} e^{\exp \left( -\frac{(i_a - \mu)^2}{2\sigma^2} \right)} \right), \]  

(2.3)
where $\mu$ and $\sigma$ are the mean and standard deviation (SD) for different classes, respectively. For this application, we define three classes: background, myocardium and blood pool. The mean for background region are taken as 0, respectively, whereas that for the blood pool and muscle region are obtained from the reference patient image.
2.6.2 Continuous Max Flow approach

In this work a graph cut-based algorithm following the work of Yuan et al. [67] [66] is used for multi-label segmentation. A multi region and multi node labelling for graph cuts can be addressed using the Continuous Max-Flow (CMF) approach and is similar to the continuous Potts model approach — analogous to “min-cut” graph theory. This approach proposed by Yuan et al. is computationally efficient in enforcing the simplex constraints and consequently requires a reduced number of iterations compared to the previous state of the art convex methods for the continuous Potts model [67].

2.7 Atlas + CMF Segmentation Integration

From the initial label obtained from the affine only probabilistic atlas propagated using global-only affine based registration techniques, we obtain a segmentation estimate that was subsequently used to obtain the initialization statistics for the max flow based graph cuts based algorithm, which in turn uses these statistics to generate an independent segmentation estimate. This segmentation estimate is then limited by the area covered by the two segmentations; hence, a weighted non-common sum of the graph cut-based segmentation combined with the affine-only atlas-based segmentation provides a final result that is superior to both methods individually. The multi label graph cuts algorithm is based on the continuous max-flow Potts model approach described by Yuan et al. [67].

2.7.1 Combined Atlas and Continuous max Flow Segmentation by Weighted Sum

1. All of the unknown test patients datasets were histogram equalized with the reference patient.
2. The mean intensity statistics from the reference patient — (0 - background, 0.1499 for the myocardium and 0.5899 for the blood pool) were

used to segment the intensity volume.

3. The atlas-based segmentation mask estimate was used to constrain the CMF segmentation into the probable region of interest by multiplying the intensity image volume by the segmentation mask obtained using the atlas label propagation.

4. To further constrain the segmentation, the CMF estimate was then multiplied by the mean of the myocardium intensity estimated from the reference patient.

5. The estimate from the atlas and the region constrained by the CMF estimate were added together and thresholded at 25%.

6. As a result, this approach boosts the myocardial regions that are double-emphasized in both the atlas estimate and the graph cut estimate, leading to a higher likelihood of true voxels labeled as myocardium to cross above the threshold. This procedure is schematically illustrated in Figure 2.8.
2.7.2 Combined Atlas and Continuous max Flow Segmentation by Comparison Selection

The use of label combination via Weighted Sum versus Comparison Selection was dictated by the similarity between the atlas-based label and graph cut-based label. Several numerical experiments conducted on the training datasets suggested that for similarities of 75% and above between the atlas- and graph cut-based labels, the label estimated via the graph cut-based approach was consistently more accurate than the atlas-based label, and was therefore selected as the preferred label following their comparison. However, for the cases in which the similarity between the atlas- and graph cut-based labels was less than 75% the weighted sum approach for combining the two labels and enhancing the common regions was utilized.

Figure 2.9: Combination of atlas-based and CMF-based labels using the Comparison Selection approach, resulting in the selection of the CMF-constrained atlas-propagated label as the preferred segmentation, provided a 75% or higher agreement between the atlas-based and CMF-refined labels.
The Comparison Selection approach proceeded according to the following steps:

1. The process followed a similar protocol as the Weighted Sum approach until reaching Step 3.
2. The constrained CMF segmentation estimate was compared to the atlas estimate, and if there was more than 75% DSC agreement between the two, then the atlas constrained CMF estimate was chosen as the final estimate.
3. If there was less than 75% DSC agreement between the two labels, then the weighted sum estimate from the previous method is chosen as the final segmentation estimate. This protocol is schematically illustrated in Figure 2.9.

2.8 Evaluation

As further reported in Chapter 3, we assessed the accuracy of the global- and global + local-based segmentation according to several parameters. Moreover, while we recognize that the latter may yield superior segmentation results, the premise of this work is to demonstrate how similar segmentation fidelity may be achieved by refining the initial global atlas-based segmentations using the graph cut based-approaches in a more efficient manner, reducing the computational time and induced registration uncertainties associated with the non-rigid registration.
Chapter 3

Results

This chapter summarizes the results associated with each of the steps in the implemented segmentation pipeline described in Chapter 2 to the ground truth segmentations of all test datasets. Each segmentation result is evaluated according to several well-established metrics described in Chapter 1, specifically the Sørensen-Dice Coefficient (DSC), Hausdorff Distance (HD) and Mean Absolute Distance (MAD) error computed at both the endocardial and epicardial borders.

The segmentation results are quantified according to their atlas construction (i.e., single reference patient, multi-patient global and multi-patient global + local atlas), label propagation (i.e., affine or non-rigid registration), as well as graph-cut based methods subsequently applied to the atlas-based segmentation results.

Lastly, the evaluation is conducted by assessing the segmentation evaluation metrics in three dimensions, by comparing the resulting Left Ventricle (LV) surface volume to the ground truth, as well as in two dimensions, by comparing the achieved segmentation with the ground truth segmentation on a slice-by-slice method. The latter is the approach typically employed for assessing segmentation performance, as it truly evaluates the 2D metrics in a 2D fashion, as they were initially designed.

3.1 Atlas-based Label Propagation

In this section segmentation performance will be assessed following label propagation from the reference patient atlas, multi-patient global (i.e.,
affine-based) atlas, and multi-patient global + local (non-rigid-based) atlas. In addition, we also show the segmentation results in response to different label propagation approaches (affine- or non-rigid registration based label propagation), and subsequently investigate further segmentation refinement beyond the atlas-based results using a variant of the Continuous Max-Flow (CMF) segmentation approach [67].

3.1.1 Reference Patient-based Atlas

3.1.1.1 Affine Label Propagation

3.1.1.1.1 3D Evaluation For the reference patient-based atlas, there is no training involved, since a patient is randomly selected from the training dataset and is treated as an atlas from which we propagate the label to the remaining training and testing datasets using affine registration. In other words, the label propagation is performed in a similar fashion as with the multi-patient atlases, however without performing any training, but rather selecting random patients and their corresponding ground truth segmentations to serve as the “atlas”.

Table 3.1 shows the mean 3D performance across all the training and test datasets, although there is no real need to separate the two categories, since no training was involved. The segmentation was assessed according to the 3D DSC, HD, and MAD at a threshold of 25%.

Table 3.1: Reference patient-based affinely propagated segmentation of the diastolic phase compared against the ground truth segmentation for all $N_{Train}^i = 70$ training datasets and $N_{Test}^i = 30$ testing datasets.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>DSC (%)</th>
<th>Endocardium MAD (mm)</th>
<th>Epicardium MAD (mm)</th>
<th>Endocardium HD (mm)</th>
<th>Epicardium HD (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{Train}^i$</td>
<td>$44.83 \pm 13.35$</td>
<td>$5.33 \pm 1.61$</td>
<td>$6.35 \pm 1.72$</td>
<td>$17.93 \pm 5.0$</td>
<td>$21.11 \pm 6.13$</td>
</tr>
<tr>
<td>$N_{Test}^i$</td>
<td>$43.67 \pm 12.20$</td>
<td>$5.32 \pm 1.46$</td>
<td>$6.49 \pm 1.97$</td>
<td>$18.81 \pm 7.35$</td>
<td>$21.56 \pm 8.87$</td>
</tr>
</tbody>
</table>
3.1.1.1.2 2D Evaluation  Figure 3.5 shows the 2D slice-wise segmentation evaluation. All segmentation parameters are reported as “per slice” averages computed over all test patients.

Figure 3.1: Slice-wise 2D evaluation of the reference patient-based affine label propagation for all training and testing datasets showing parameter (DSC, endocardial and epicardial HD, and endocardial MAD and epicardial MAD) variation according to slice location (apical, middle and basal slices).
3.1.2 Global Atlas (GA)

3.1.2.1 Affine Label Propagation

3.1.2.1.1 3D Evaluation  Table 3.2 shows the performance of the Global Atlas (GA) over all the training and test patients following affine label propagation. The segmentation was assessed in 3D according to the DSC, HD, and MAD, at a threshold of 25%. We observe that the GA performs significantly better than selecting a single reference patient from the training dataset and designating it as an atlas, as demonstrated in the previous section.

Table 3.2: Global atlas-based affinely propagated segmentation for the diastolic phase compared to the ground truth segmentation for $N_{Train}^i = 70$ training patient datasets and $N_{Test}^i = 30$ testing datasets.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>DSC (%)</th>
<th>Endocardium MAD (mm)</th>
<th>Epicardium MAD (mm)</th>
<th>Endocardium HD (mm)</th>
<th>Epicardium HD (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{Train}^i$</td>
<td>62.06 ± 9.29</td>
<td>4.31 ± 1.41</td>
<td>4.09 ± 1.17</td>
<td>16.58 ± 5.26</td>
<td>18.84 ± 6.32</td>
</tr>
<tr>
<td>$N_{Test}^i$</td>
<td>58.54 ± 10.35</td>
<td>4.67 ± 1.76</td>
<td>4.37 ± 1.77</td>
<td>16.52 ± 6.99</td>
<td>17.9 ± 9.05</td>
</tr>
</tbody>
</table>

3.1.2.1.2 2D Evaluation  Figure 3.2 shows the 2D slice-wise performance, once again estimated across all testing datasets and averages at each slice.

3.1.3 Global + Local Atlas (GLA)

3.1.3.1 Affine + Non-Rigid Label Propagation

3.1.3.1.1 3D Evaluation  For the results shown in Table 3.3, the label was propagated from the Global + Local Atlas (GLA) first using affine registration, followed by a non-rigid alignment. Segmentation was assessed in 3D according to the DSC, HD, and MAD, at a threshold of 25%.
Figure 3.2: Slice-wise 2D evaluation of the global atlas followed by affine label propagation for all training and testing datasets showing parameter (DSC, endocardial and epicardial HD, and endocardial MAD and epicardial MAD) variation according to slice location (apical, middle and basal slices).

3.1.3.1.2 2D Evaluation  Figure 3.3 shows the 2D slice-wise performance for the global + local atlas followed by affine + non-rigid label propagation.
Table 3.3: Global + local atlas-based non-rigidly propagated segmentation for the diastolic phase compared to the ground truth segmentation for N = 70 training patient datasets.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>DSC (%)</th>
<th>Endocardium MAD (mm)</th>
<th>Epicardium MAD (mm)</th>
<th>Endocardium HD (mm)</th>
<th>Epicardium HD (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N^i_{Train} )</td>
<td>62.52 ± 9.38</td>
<td>4.19 ± 1.37</td>
<td>4.07 ± 1.16</td>
<td>16.57 ± 5.22</td>
<td>17.94 ± 5.44</td>
</tr>
<tr>
<td>( N^i_{Test} )</td>
<td>58.92 ± 10.11</td>
<td>4.62 ± 1.76</td>
<td>4.36 ± 1.70</td>
<td>17.28 ± 7.68</td>
<td>17.97 ± 8.79</td>
</tr>
</tbody>
</table>

3.1.3.2 Affine Label Propagation

3.1.3.2.1 3D Evaluation Table 3.4 summarizes the segmentation results for the case when the affine + local atlas label was propagated affinely from the GLA. It can be seen that the performance difference between the GA label propagated using affine registration vs. the same label propagated via affine + non-rigid registration are minimal.

Table 3.4: Global + local atlas-based affinely propagated segmentation for the diastolic phase against the ground truth segmentation for N = 70 training patient datasets. Segmentation was assessed in 3D according to the Dice correlation, Hausdorff distance, and Mean Absolute Distance error at a threshold of 0.25

<table>
<thead>
<tr>
<th>Dataset</th>
<th>DSC (%)</th>
<th>Endocardium MAD (mm)</th>
<th>Epicardium MAD (mm)</th>
<th>Endocardium HD (mm)</th>
<th>Epicardium HD (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>62.29 ± 9.45</td>
<td>4.21 ± 1.38</td>
<td>4.11 ± 1.18</td>
<td>16.71 ± 5.29</td>
<td>18.32 ± 5.94</td>
</tr>
<tr>
<td>Test</td>
<td>58.57 ± 10.21</td>
<td>4.68 ± 1.79</td>
<td>4.40 ± 1.76</td>
<td>17.46 ± 7.82</td>
<td>17.94 ± 9.26</td>
</tr>
</tbody>
</table>

3.1.3.2.2 2D Evaluation Figure 3.4 shows the 2D slice-wise performance evaluation for the affine propagation of the global + local atlas label.
Figure 3.3: Slice-wise 2D evaluation of the global + local atlas followed by affine + non-rigid label propagation for all training and testing datasets showing parameter (DSC, endocardial and epicardial HD, and endocardial MAD and epicardial MAD) variation according to slice location (apical, middle and basal slices).
Figure 3.4: 2D slicewise averages for testing patients propagated from the GLA affinely and showing parameter (DSC, endocardial and epicardial HD, and endocardial MAD and epicardial MAD) variation according to slice location (apical, middle and basal slices).
3.2 Global + Local Atlas with Continuous Max Flow

3.2.1 Affine Label Propagation and Weighted Sum CMF Refinement

3.2.1.1 3D Evaluation

Table 3.5 summarizes the segmentation assessment following affine label propagation and subsequent via weighted sum-based CMF refinement.

Table 3.5: Global + local atlas-based affine label propagation and weighted sum CMF refined segmentation for the diastolic phase compared against the ground truth segmentation for $N_{Train} = 70$ training patient datasets and $N_{Test} = 30$ testing datasets.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>DSC (%)</th>
<th>Endocardium MAD (mm)</th>
<th>Epicardium MAD (mm)</th>
<th>Endocardium HD (mm)</th>
<th>Epicardium HD (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{Train}^i$</td>
<td>63.62 ± 8.77</td>
<td>4.24 ± 1.50</td>
<td>4.08 ± 1.12</td>
<td>16.47 ± 5.23</td>
<td>18.58 ± 6.44</td>
</tr>
<tr>
<td>$N_{Test}^i$</td>
<td>59.87 ± 9.53</td>
<td>4.62 ± 1.84</td>
<td>4.37 ± 1.76</td>
<td>16.77 ± 6.97</td>
<td>17.63 ± 9</td>
</tr>
</tbody>
</table>

3.2.1.2 2D Evaluation

Figure 3.5 shows the 2D slice-wise performance for the global + local atlas following affine label propagation an weighted-sum CMF refinement.

3.2.2 Affine Label Propagation and CMF Comparison Selection

3.2.2.1 3D Evaluation

Table 3.6 summarizes the segmentation assessment results achieved via the combination of atlas label propagation followed by comparison selection. The reader is reminded that for this approach, the continuous max-flow results was used as the optimal segmentation should the agreement between the CMF the atlas-propagated label exceeded 75%.
Figure 3.5: 2D slice-wise segmentation evaluation parameter averages for testing patients propagated from the GLA using affine registration followed by weighted sum CMF refinement and showing parameter (DSC, endocardial and epicardial HD, and endocardial MAD and epicardial MAD) variation according to slice location (apical, middle and basal slices).

3.2.2.2 2D Evaluation

Figure 3.6 shows the 2D slice-wise performance for the comparison selection CMF refinement combined with the affine label propagation of the
Table 3.6: Global + local atlas-based affine with CMF propagated segmentation by comparison selection for the diastolic phase compared to the ground truth segmentation for $N_{Train} = 70$ training patient datasets and $N_{Test} = 30$ testing datasets

<table>
<thead>
<tr>
<th>Dataset</th>
<th>DSC (%)</th>
<th>Endocardium MAD (mm)</th>
<th>Epicardium MAD (mm)</th>
<th>Endocardium HD (mm)</th>
<th>Epicardium HD (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{N}_{Train}$</td>
<td>66.54 ± 8.85</td>
<td>5.85 ± 5.34</td>
<td>3.97 ± 1.49</td>
<td>29.67 ± 17.58</td>
<td>19.91 ± 8</td>
</tr>
<tr>
<td>$\bar{N}_{Test}$</td>
<td>63.55 ± 10.11</td>
<td>4.84 ± 2.62</td>
<td>4.01 ± 1.28</td>
<td>24.97 ± 12.68</td>
<td>19.69 ± 7.56</td>
</tr>
</tbody>
</table>

global + local atlas.

Lastly, **Figure 3.7** provides an example for the visual qualitative assessment of the 2D slice-wise evaluation of the segmentation against the ground truth. These diagrams are typical for visually displaying and comparing the results of a segmentation technique vis-a-vis the ground truth segmentation of the image dataset. The white regions represent areas that were depicted by both the tested and ground truth segmentation, the red regions identify areas that belong to the ground truth, but not depicted by the tested method, while the blue regions identify areas depicted by the tested method, but not included in the ground truth.

As expected, and also demonstrated by the 2D slice-based evaluations, most mid-slices were sufficiently accurately segmented, while the apical and basal slices presented some challenges, mainly due to their “incomplete segmentation pattern either in the ground truth data or the tested segmentation techniques. In addition, both the apical and basal regions are challenging to segment, simply because the LV structure does not feature the “ring-like” appearance and it’s rather difficult to identify the optimal start and end points of the anatomical structure, hence the lower DSC correlation and higher HD and MAD errors, as consistently shown in all 2D slice-wise evaluation plots.
Figure 3.6: 2D slice-wise segmentation evaluation parameter averages for testing patients propagated from the GLA using affine registration followed by comparison selection CMF refinement and showing parameter (DSC, endocardial and epicardial HD, and endocardial MAD and epicardial MAD) variation according to slice location (apical, middle and basal slices).

### 3.3 Summary

Table 3.7 displays the performance of all the test patient datasets from all the methods.
Table 3.7: Assessment summary of various techniques against the ground truth segmentation of $N = 30$ unknown test patient datasets.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Label Propagation</td>
<td>Affine</td>
<td>Affine</td>
<td>Affine</td>
<td>Affine + Non-Rigid</td>
<td>Weighted Sum</td>
<td>Comparison Selection</td>
</tr>
<tr>
<td>DICE</td>
<td>43.67 ± 13.35</td>
<td>58.54 ± 10.35</td>
<td>58.57 ± 12.6</td>
<td>58.92 ± 12.6</td>
<td>59.87 ± 9.53</td>
<td>63.55 ± 10.11</td>
</tr>
<tr>
<td>Endocardium MAD</td>
<td>5.32 ± 1.46</td>
<td>4.67 ± 1.76</td>
<td>4.68 ± 2.83</td>
<td>4.62 ± 1.76</td>
<td>4.62 ± 1.84</td>
<td>4.84 ± 2.62</td>
</tr>
<tr>
<td>Epicardium MAD</td>
<td>6.49 ± 1.97</td>
<td>4.37 ± 1.77</td>
<td>4.40 ± 1.76</td>
<td>4.36 ± 1.70</td>
<td>4.37 ± 1.76</td>
<td>4.01 ± 1.28</td>
</tr>
<tr>
<td>Endocardium HD</td>
<td>18.81 ± 7.35</td>
<td>16.52 ± 6.99</td>
<td>17.46 ± 7.82</td>
<td>17.28 ± 7.68</td>
<td>16.77 ± 6.97</td>
<td>24.97 ± 12.68</td>
</tr>
<tr>
<td>Epicardium HD</td>
<td>21.56 ± 8.87</td>
<td>17.9 ± 9.05</td>
<td>17.94 ± 9.26</td>
<td>17.97 ± 8.79</td>
<td>17.63 ± 9</td>
<td>19.69 ± 7.56</td>
</tr>
</tbody>
</table>
Figure 3.7: Segmentation for 30 unknown test patients from the global + local atlas following affine label propagation and graph cut-based refinement using weighted sum. Note that the blue area identifies regions segmented by the proposed method, but not included in the ground truth; the red areas identify regions that belong to the ground truth, but were not identified in the tested segmentation, while the white areas represent ground truth regions accurately identified by the tested method.
Chapter 4

Summary, Discussion and Future Work

4.1 Summary

The work presented in this thesis investigates the use of various atlases of different complexity — single reference patient-based atlas, multi-patient global affine atlas and multi-patient global (affine) + local (non-rigid) atlas — as a means to generate segmentations of the left ventricle myocardium from 3D cardiac MR images. Starting from the premise that most atlases require computationally intensive non-rigid registration and deformable label propagation to achieve optimal accuracy, this work is aimed at exploring the feasibility of combining atlas-based segmentation labels with patient-specific atlas propagated labels used as precursors for high level graph cut-based segmentation refinement.

In support of this theory, we have explored the use of a Continuous Max-Flow (CMF) [67] [66] based approach that operates on the patient-specific atlas-propagated label to perform the voxel classification with region constraints provided from the atlas to CMF algorithm. In addition, the resulting CMF segmentation refinements were combined with the atlas-based labels via either a weighted sum or comparison selection approach, depending on the level of agreement between the atlas-derived and CMF-based labels.

The subsequent sections focus on summarizing the methodologies and results and also outline challenges and limitations of the proposed technique vis-a-vis the inherent limitations imposed by the datasets, the image variability, and the inherent uncertainties of the ground truth segmented data annotated by expert clinicians.
4.2 Discussion

In section 3.1.1, we summarized the results associated with the use of a single reference patient serving as atlas. To obtain a segmentation of a new test patient image, the reference image ground truth was propagated to the new patient by using the registration transform Ground Truth (GT) identified via intensity-based registration of the reference patient to the new patient test image. According to both the 2D and 3D segmentation metrics used for assessment, the overall segmentation performance was relatively poor.

These findings can be explained by the fact that patients that are similar to the reference patient in parametric space, as well as shape and intensity, would perform better, while those patients that are largely dissimilar to the reference patient would perform poorly, simply to the gross difference in size and shape and the inability of the registration algorithm to account to a sufficient extent for such large shape and size variations (Figure 4.1). Nevertheless, following global alignment and registration of all patient datasets into the same parameter space following global affine registration, the variability of the patient dataset was minimized with respect to the average global atlas (Figure 4.1).

To visualize why some of the patients perform well in registration while the others perform worse, we must understand how different the images being registered are from each other. Let us consider the worse performing patient for our reference patient atlas shown in Figure 4.2. The patient shown in the left panel features a poor quality volume with severe distortion and step artifacts caused most likely by motion of the patient while being imaged. These artifacts resulted in poor registration to the reference patient shown on in the right panel. In response to the poor test patient-to-reference patient registration, the affine transformation obtained is not optimal and therefore the segmentation label associated with the reference patient is not accurately propagated to the new test patient, resulting in segmentation uncertainties present in the new test patient image.

If observed more closely, the segmented label propagated to the fixed patient GT shows similar step artifacts along the z-axis, further contributing to poor overlap between the obtained segmentation volumes and the original
Figure 4.1: Illustration of the patient image dataset variability with respect to the reference patient (a) before global alignment and (b) relative to the global atlas
Figure 4.2: Visual example of a test image dataset featuring poor atlas label propagation performance: (a) Poor image quality of the in a test patient (b) Reference patient

ground truth, also reflected by the quantitative performance metrics. needless to mention, this aparticular test patient also features a largely different orientation for that of the patient dataset chosen as reference.

Despite the poor image quality for some patients, both the Global Atlas (GA) (refer section 3.1.2) and Global + Local Atlas (GLA) (refer section 3.1.3) atlas performed significantly better than the reference patient atlas, which demonstrates that global atlas construction is vastly better with regards to its performance than the reference patient atlas, and also comparable in performance to the Global + Local Atlas (GLA). In Figure 3.2 we can observe that the global atlas performs rather poorly at the basal (slices at the upper end) and apical slices (slices at the lower end), but performs reasonably well across all the mid-slices. The performance for GLA is similar.

We can observe that the proposed hybrid label propagation approach that combines the atlas-based affinely propagated label refined using Continuous Max-Flow (CMF)-based segmentation demonstrates higher 3D segmentation accuracy than either the traditional non-rigid atlas with non-rigidly propagated labels or the graph cut segmentations alone. The computational time, while higher than graph cuts alone, provides much greater accuracy than either of the baseline methods at a marginally higher computational time than graph cuts alone and significantly lower computational time than traditional non-rigid registration based label propagation (Table 4.1).
For the comparison selection method, although the performance quantified via the Sørensen-Dice Coefficient (DSC) is the highest, which is an area based measure, the other parameters such as Mean Absolute Distance (MAD) and Hausdorff Distance (HD) are lower, simply because for some patients the CMF segmentation does not return a complete ring, thus preventing full surface closure for some slices, which significantly impacts the contour based evaluation metrics HD and MAD.

![Segmentation performance visualization](image)

**Figure 4.3:** Segmentation performance visualization for the affine + Continuous Max-Flow (CMF) comparison selection method for a test patient where the CMF estimate was chosen as the final estimate. Incomplete slice rings are observed, leading to good area overlap performance, but poor contour-to-contour distance performance (HD and MAD)

The transformation of binary masks using affine or non-rigid grid based transforms leads to interpolation errors. These interpolation errors add noise to the binary image by adding values not at a 0 or 1 intensity level, wherever a voxel is mapped to a location that corresponds to a sub-voxel location. To reduce the impact of interpolation-induced noise, the binary mask needs to be re-thresholded to an acceptable level to eliminate the intensity values that
correspond to sub-voxel locations. However, a threshold too low results in too much noise, rendering image regions as myocardium when in fact they are not, whereas a threshold too high leads to information loss mostly in the apical slices, thus reducing performance accuracy in those regions and reducing the overall average of the quantitative segmentation metrics. These limitations emphasize the need for an optimal threshold that provides a balanced trade-off between information loss and interpolation-induced noise reduction.

For the cases under consideration and through a process of careful iterative, empirical numerical experimentation given the intensity range used throughout this work, an optimal threshold was empirically determined to be at 25%, However, thresholding itself is an example of user bias that affects segmentation performance.

Table 4.1: Average label propagation time per patient for different methods, as can be observed affine label propagation is faster and results in nearly the same accuracy as non-rigid label propagation, but with the addition of Continuous Max-Flow (CMF) results in better performance and multiple times lower computation time than traditional non rigid label propagation

<table>
<thead>
<tr>
<th>Label Propagation</th>
<th>Affine</th>
<th>Affine + Non-Rigid</th>
<th>CMF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time in seconds</td>
<td>1.2940 ± 0.31</td>
<td>88.91 ± 12.13</td>
<td>0.7433 ± 0.08</td>
</tr>
</tbody>
</table>

4.3 Future directions

4.3.1 Multi-phase Scaling through Shape Modeling

In the future, the atlas trained for one cardiac phase (diastolic) can be used to propagate the label to another phase. The optimal frame-to-frame transforms between the successive frames of a 4D cardiac MR dataset will be used to determine the principal modes of variation of the cardiac features during the cardiac cycle. This information will then be used to propagate the segmented 3D model from the diastolic phase throughout the remaining
phases of the dataset, and optimize the segmentation to better fit each individual phase. Additionally, the probabilistic atlas could be converted into a full statistical shape model where the modes of variation in the training data are obtained through Principal Component Analysis (PCA).

4.3.2 Integration with Other Low-cost Segmentation Methods

Also other segmentation methods such as level sets using the atlas as prior can be implemented, as this work has demonstrated that the low cost affine registration combined with another low cost method such as graph cuts or level sets could be a faster and more accurate way of propagating labels from an atlas in comparison to traditional non-rigid registration.

4.3.3 Alternate Approaches to Combine Labels from Different algorithms

More efficient methods to combine the segmentation from two sources similar to the STAPLE [63] algorithm, or a combination of two segmentations based on some expectation maximization [53] techniques, might be one area to explore, in terms of efficient combination of labels obtained from different segmentation techniques.

4.3.4 Potential Scaling to Real-time Performance

The difference in time for atlas construction between the Global Atlas (GA) and the Global + Local Atlas (GLA) is on the order of several hours. Therefore the fast construction time taken for the global atlas, fast label propagation time, and further integration with another fast segmentation algorithm (as has been demonstrated in this work through CMF) might be the basis for scaling to real time surgical interventions and applications. This approach will enable the use of atlas-based segmentation techniques to be considered for such applications, particularly with parallelization of the registration (CMF can already be parallelized), as in the past such registration
based techniques might have been considered computationally expensive, for mission critical applications.
Bibliography


[46] The University of Nottingham. A beginners guide to normal heart function, sinus rhythm and common cardiac arrhythmias.


[61] Tim Taylor. Human heart diagram and anatomy of the heart.


