Automated Cardiac MR Image Analysis for Population Imaging

Xènia Albà

TESI DOCTORAL UPF / 2017
Thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Directors de la tesi
Prof. Alejandro F. Frangi
Dr. Karim Lekadir
Dra. Rosa M. Figueras

Department of Information and Communication Technologies
This work was carried out at the Center for Computational Imaging and Simulation Technologies in Biomedicine (CISTIB), Universitat Pompeu Fabra, Barcelona.

This thesis was supported by a departmental scholarship from the Department of Information and Communications Technologies at Universitat Pompeu Fabra (DTIC-UPF) and by an FPU grant from the Spanish Ministry of Education Culture and Sport. Moreover, we acknowledge funding support from the Spanish Ministry of Science and Innovation and the Spanish Ministry of Health, under the grant agreement cvREMOD (CEN-20091044). We also acknowledge funding support from the European Commission Seventh Framework Programme under grant agreements euHeart (FP7-ICT-2007-224495).
# Contents

Abstract xi  
Resum xiii  

1 General Introduction 1  
1.1 The Era of Big Data 2  
1.1.1 Big Data in Health-care and Biomedicine 2  
1.1.2 Big Data in Medical Imaging 3  
1.1.3 Estimating New Biomarkers from Large-scale Studies 5  
1.1.4 The Case of Cardiovascular Diseases and Cardiac Imaging 6  
1.2 Aim and Objectives of This Thesis 7  
1.3 Contributions 8  
1.4 Academic Context and Funding 9  

2 Automatic Cardiac LV Segmentation in MRI using Modified Graph Cut 11  
2.1 Introduction 13  
2.2 Methods 14  
2.2.1 Overview 14  
2.2.2 LV Initial Localization 14  
2.2.3 LV Segmentation 18  
2.3 Results 22  
2.3.1 Data Description 22  
2.3.2 Reference Technique: 3D-ASM Segmentation 22  
2.3.3 Evaluation Process 23  
2.3.4 LE-MRI Segmentation 23  
2.3.5 Cine-MRI Segmentation 24  
2.3.6 Parameter Influence 24  
2.4 Discussion 24  
2.5 Conclusion 28  

3 An Algorithm for the Segmentation of Highly Abnormal Hearts using a Generic SSM 29  
3.1 Introduction 31  
3.2 Methods 33  
3.2.1 Overview 33  
3.2.2 Initialization 35  
3.2.3 Feature Points Search 36  
3.2.4 Estimation of the Remodeling Transformation 36  
3.2.5 Model Fitting 37  
3.3 Results 38
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.3.1</td>
<td>Data Description</td>
<td>38</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Experimental Setup</td>
<td>39</td>
</tr>
<tr>
<td>3.3.3</td>
<td>Results for PH Patients</td>
<td>40</td>
</tr>
<tr>
<td>3.3.4</td>
<td>Results for HCM Patients</td>
<td>41</td>
</tr>
<tr>
<td>3.3.5</td>
<td>Statistical Significance</td>
<td>42</td>
</tr>
<tr>
<td>3.3.6</td>
<td>Effect of Mapping</td>
<td>42</td>
</tr>
<tr>
<td>3.3.7</td>
<td>Illustrative Results</td>
<td>46</td>
</tr>
<tr>
<td>3.4</td>
<td>Discussion and Conclusion</td>
<td>49</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>55</td>
</tr>
<tr>
<td>4.2</td>
<td>Method</td>
<td>58</td>
</tr>
<tr>
<td>4.2.1</td>
<td>Overview</td>
<td>58</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Automatic Key Landmark Detection</td>
<td>60</td>
</tr>
<tr>
<td>4.2.3</td>
<td>Three-dimensional Model-based Cardiac Segmentation</td>
<td>64</td>
</tr>
<tr>
<td>4.2.4</td>
<td>Segmentation Quality Control</td>
<td>65</td>
</tr>
<tr>
<td>4.3</td>
<td>Data Description</td>
<td>68</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Images</td>
<td>68</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Ground Truth</td>
<td>68</td>
</tr>
<tr>
<td>4.3.3</td>
<td>Data Preprocessing</td>
<td>68</td>
</tr>
<tr>
<td>4.4</td>
<td>Results</td>
<td>69</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Landmark Detection Evaluation</td>
<td>69</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Quality Control Evaluation</td>
<td>70</td>
</tr>
<tr>
<td>4.4.3</td>
<td>Fully Automatic Segmentation Errors</td>
<td>72</td>
</tr>
<tr>
<td>4.4.4</td>
<td>Illustrative Results</td>
<td>73</td>
</tr>
<tr>
<td>4.5</td>
<td>Discussion and Conclusion</td>
<td>74</td>
</tr>
<tr>
<td>5.1</td>
<td>Overview</td>
<td>77</td>
</tr>
<tr>
<td>5.2</td>
<td>Outlook and Further Research Directions</td>
<td>78</td>
</tr>
<tr>
<td>References</td>
<td></td>
<td>81</td>
</tr>
<tr>
<td>List of Publications</td>
<td></td>
<td>xv</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td></td>
<td>xvii</td>
</tr>
<tr>
<td>Biography</td>
<td></td>
<td>xix</td>
</tr>
</tbody>
</table>
List of Figures

1.1 The three V’s in Big Data. Adapted from [35]. ........................................ 3
1.2 Cardiac short-axis MR images at end-diastole from normal, MI, PH and HCM subjects .......................................................... 7

2.1 Examples of myocardial segmentations in LE-MRI .................................. 16
2.2 Template used in the initialization step .................................................. 16
2.3 Schematic illustration of the proposed initialization .................................. 17
2.4 Schema of the connectivity for a control point with its neighbors, the control points in the within-slice direction and in the interslice direction .......... 19
2.5 LE-MRI SA slices from two patients with the ground truth, the automatic initialization step, our automatic final segmentation and 3D-ASM segmentation . 25
2.6 Cine-MRI SA slices from two patients of the Sunnybrook dataset: ground truth and automatic segmentation ................................................. 26

3.1 Cardiac SA MRI at end-diastole from normal, PH and HCM subjects ........ 32
3.2 Schematic workflow of our previous implementation in [4] ..................... 33
3.3 Schematic workflow of the proposed method for the segmentation of abnormal hearts using a reference PDM ............................................ 35
3.4 P2S errors for the segmentation of the PH cases ..................................... 40
3.5 P2S errors for the segmentation of the HCM cases .................................. 42
3.6 Average segmentation errors as a function of different proportions of feature points used for the estimation of the transformation ..................... 43
3.7 Histogram of average shape changes due to the remodeling mapping, for normal and abnormal cases ....................................................... 44
3.8 Mahalanobis distances of each of the cases (normal, PH, and HCM) with respect to the normal shape model ............................................. 45
3.9 The main steps involved in the proposed algorithm for the segmentation of a PH heart ................................................................. 46
3.10 Three examples showing myocardial segmentations on PH patients using manual delineation, the reference PDM, the PH-specific model, and the proposed technique ....................................................... 47
3.11 Three examples showing myocardial segmentations on HCM patients using manual delineation, the reference PDM, the PH-specific model, and the proposed technique .................................................. 48
3.12 3D illustrations showing average P2S segmentation errors for PH cases using the reference PDM and the proposed technique ....................... 49
3.13 3D illustrations showing average P2S segmentation errors for HCM cases using the reference PDM and the proposed technique ....................... 50

4.1 Pipeline of the LV automatic segmentation ........................................... 59
4.2 The intersection of the LA and SA images provides rough positions for the LV . 62
4.3 Pipeline of the training step for landmark detection .................................. 62
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.4</td>
<td>Performance results of increasing the number of trees in the forest for landmark detection</td>
<td>70</td>
</tr>
<tr>
<td>4.5</td>
<td>Cumulative error distribution curves for each landmark detection</td>
<td>71</td>
</tr>
<tr>
<td>4.6</td>
<td>Cumulative error distribution curves for myocardial segmentation error</td>
<td>72</td>
</tr>
<tr>
<td>4.7</td>
<td>Visual examples of landmark detection and segmentation results for AV datasets</td>
<td>73</td>
</tr>
<tr>
<td>4.8</td>
<td>Visual examples of segmentation results for MI datasets</td>
<td>74</td>
</tr>
<tr>
<td>4.9</td>
<td>Visual examples of discarded segmentation results</td>
<td>74</td>
</tr>
</tbody>
</table>
## List of Tables

2.1 List of LV properties used as rules to constraint myocardial initialization and segmentation in multi-sequence MRI ........................................... 15
2.2 Average P2P distance errors and Dice similarity coefficients for the LE-MRI and cine-MRI dataset ................................................................. 23

3.1 P2S segmentation error statistics for the PH cases .................................. 40
3.2 P2S segmentation error statistics for the HCM Cases ................................. 41
3.3 Statistical significance of the differences between the results using the three segmentation strategies ................................................................. 43
3.4 Statistical significance of the differences between the Mahalanobis distances with respect to the normal shape model of the normal population and the abnormal shapes ................................................................. 44
3.5 P2S segmentation error statistics for the PH and HCM ............................. 44

4.1 Summary of state of the art methods for automatic and semi-automatic LV segmentation with large databases .............................................. 56
4.2 Automatic landmark detection methods in cardiac MRI ............................ 57
4.3 Summary of the data available for evaluation ........................................... 69
4.4 P2P distance error statistics in landmark detection (subset of group AV) ...... 70
4.5 Performance of segmentation quality control .......................................... 70
4.6 P2P error statistics for automatic image parsing ...................................... 71
4.7 P2P error statistics for automatic image parsing (manual vs automatic) .... 71
## Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASM</td>
<td>Active Shape Model</td>
</tr>
<tr>
<td>CAD</td>
<td>Coronary Arterial Disease</td>
</tr>
<tr>
<td>CAP</td>
<td>Cardiac Atlas Project</td>
</tr>
<tr>
<td>CRT</td>
<td>Cardiac Resynchronization Therapy</td>
</tr>
<tr>
<td>CT</td>
<td>Computed Tomography</td>
</tr>
<tr>
<td>CVD</td>
<td>Cardiovascular Disease</td>
</tr>
<tr>
<td>DSC</td>
<td>Dice Similarity Coefficient</td>
</tr>
<tr>
<td>ED</td>
<td>End of Diastole</td>
</tr>
<tr>
<td>EMR</td>
<td>Electronic Medical Records</td>
</tr>
<tr>
<td>ES</td>
<td>End of Systole</td>
</tr>
<tr>
<td>GC</td>
<td>Graph Cut</td>
</tr>
<tr>
<td>GF</td>
<td>Gabor Filter</td>
</tr>
<tr>
<td>HCM</td>
<td>Hypertrophic Cardiomyopathy</td>
</tr>
<tr>
<td>HF</td>
<td>Heart Failure</td>
</tr>
<tr>
<td>HLA</td>
<td>Horizontal Long Axis</td>
</tr>
<tr>
<td>HOG</td>
<td>Histogram of Oriented Gradients</td>
</tr>
<tr>
<td>IAM</td>
<td>Intensity Appearance Model</td>
</tr>
<tr>
<td>LA</td>
<td>Long Axis</td>
</tr>
<tr>
<td>LE-MRI</td>
<td>Late-Enhancement MRI</td>
</tr>
<tr>
<td>LV</td>
<td>Left Ventricle/Left Ventricular</td>
</tr>
<tr>
<td>MD</td>
<td>Mahalanobis Distance</td>
</tr>
<tr>
<td>MI</td>
<td>Myocardial Infarction</td>
</tr>
<tr>
<td>PH</td>
<td>Pulmonary Hypertension</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic Resonance Imaging</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal Component Analysis</td>
</tr>
<tr>
<td>PDM</td>
<td>Point Distribution Model</td>
</tr>
<tr>
<td>P2P</td>
<td>Point-to-point</td>
</tr>
<tr>
<td>P2S</td>
<td>Point-to-surface</td>
</tr>
<tr>
<td>RF</td>
<td>Random Forest</td>
</tr>
<tr>
<td>RV</td>
<td>Right Ventricle/Right Ventricular</td>
</tr>
<tr>
<td>SA</td>
<td>Short Axis</td>
</tr>
<tr>
<td>SD</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>SPASM</td>
<td>Sparse ASM</td>
</tr>
<tr>
<td>SSM</td>
<td>Statistical Shape Model</td>
</tr>
<tr>
<td>SSFP</td>
<td>Steady State Free Precession</td>
</tr>
<tr>
<td>TPS</td>
<td>Thin-plate Spline</td>
</tr>
<tr>
<td>VLA</td>
<td>Vertical Long Axis</td>
</tr>
</tbody>
</table>
Abstract

Despite advances in cardiac analysis research, cardiovascular diseases (CVDs) remain the leading cause of deaths worldwide. It is vital to improve our understanding of what causes CVDs and how to prevent and treat them. Recently, we have seen an increasing volume of medical image data and annotations being collected and stored. The exploitation of large-scale population image data is expected to have the potential to improve care, save lives and lower costs by discovering associations and understanding patterns and trends within this data. There is a need to reach its potential and to translate it into knowledge for improved healthcare. Exceptional technologies, frameworks and algorithms are thus needed to efficiently process these large quantities of data to obtain useful information.

To analyze large-scale cardiovascular studies, an important pre-requisite is to automatically and robustly process the image data and extract information about the morphology and function. The cardiac function can be quantified through several indices such as ventricle volumes, masses and ejection fraction, by previously segmenting the left (LV) and right (RV) ventricles from magnetic resonance images (MRI). The aim of this thesis is to develop techniques capable of segmenting large-scale, highly variable cardiac MRI without the need of expert assistance. These methods have to be able to deal with the inherent inconsistencies and changes in image sequences. Moreover, the approaches have to be robust to different groups of individuals and classes of disorders. And, finally, regarding the large amount of data to be processed, the methods need to be fully automatic. Particularly, the idea is to develop methods that automatically load data, initialize a process, run it and, subsequently, score the quality of the results. The main aim has been divided into three objectives, each of them resulting in a scientific contribution published or submitted to a peer-review journal.

In the first contribution, we developed an automatic approach, which can be applied across MRI sequences without the need for specific tuning. A generic rule-based framework was used to segment the LV boundaries using intensity information, and including shape and inter-slice smoothness constraints. The segmentation was evaluated on two MRI sequences, LE-MRI and cine-MRI, showing a good balance between flexibility and robustness. Our second contribution proposed an algorithm for the segmentation of severely deformed pathological hearts in MRI. The algorithm was highly flexible, as it did not require a priori knowledge of the involved pathology. The idea of the algorithm was to approximate the abnormality effect with a virtual remodeling transformation between the patient-specific geometry and the average shape of a reference model. The algorithm was validated with patients diagnosed with two different cardiac pathologies showing the effectiveness of the technique. Finally, in the third contribution, a fully automatic framework was presented, which parsed images to perform LV detection, LV segmentation and quality control for large-scale cardiac MRI databases. The method included a regression-based approach to predict the initial location of the heart and a quality control based on anatomical-justified descriptors. Validation based on two public large cohorts showed the promise of the approach for user-independent analysis of such populations.
Resum

Malgrat els avanços en la investigació d’anàlisi cardíaca, les malalties cardiovasculears (MCV) continuen sent una de les principals causes de mort a tot el món. És de vital importància millorar la nostra comprensió en les causes de les MCV i la forma de prevenir-les i tractar-les. En els últims temps hi ha hagut un increment en el volum d’imatges mèdiques i dades clíniques emmagatzemades. S’espera que l’ús d’aquestes imatges a gran escala tingui el potencial per millorar l’atenció clínica, salvar vides i reduir els costos gràcies a la detecció de relacions entre elles i la comprensió de patrons i tendències dins d’aquestes dades. Es necessiten eines per abordar plenament el potencial de les dades en el món de la salut i traduir-les en coneixement per a la millora de l’assistència sanitària. Especialment en el cas d’imatges biomèdiques, això requereix del desenvolupament de nous algoritmes que puguin fer front de manera eficient a la mida, la complexitat i la variabilitat de grans bases de dades.

Per analitzar grans bases de dades cardíiques, primer cal processar-les de manera eficient i extreure informació sobre la funció cardíica i la morfologia. La funció cardíica es pot quantificar a partir del càlcul dels volums, les àrees i la fracció d’ejecció dels ventricles, segmentant prèviament el ventricle dret (VD) i l’esquerre (VE) en imatges de ressonància magnètica (RM). L’objectiu d’aquesta tesi és desenvolupar tècniques que, tot i la variabilitat de les imatges cardíiques, siguin capaces de segmentar-ne grans quantitats sense la necessitat de cap expert. Aquests mètodes han de ser capaços de tractar els diferents tipus de seqüències d’imatges i les seves inconsistències inherents. D’altra banda, les tècniques desenvolupades han de mantenir la validesa aplicades independentment dels diferents individus i desordres. Finalment, pel que fa als grans volums de dades que s’han de processar, els mètodes han de ser completament automàtics. En concret, la idea és desenvolupar mètodes que de manera automàtica carreguin les dades, inicialitzin el procés, l’executin i, posteriorment, avaluin la qualitat dels resultats. L’objectiu principal s’ha dividit en tres subobjectius, cadascun dels quals ha donat lloc a una contribució científica publicada o enviada a una revista tècnica internacional.

En la primera de les contribucions, vam desenvolupar un mètode automàtic que es podia utilitzar amb diferents seqüències de RM sense necessitat de cap modificació. Per segmentar els límits del VD utilitzàvem un esquema basat en regles que feien ús de la intensitat i que inclouien restriccions de forma i d’homogeneïtat entre els diferents talls de la RM. Per avaluar el procés vam utilitzar dues seqüències de RM, la de contrast amb realçament tardà (RT-RM) i la funcional amb la tècnica gradient-echo (cine-RM), la qual cosa va demostrar el bon comportament de la tècnica pel que fa a la flexibilitat i l’eficiència. En la nostra segona contribució vam proposar un algoritme per segmentar cors amb formes anormals en imatges de RM. L’algoritme era molt flexible i no era necessari el coneixement del model de patologia diagnosticada. La idea era aproximar l’efecte d’aquesta anormalitat per mitjà d’una transformació virtual del remodelat entre la geometria del pacient i la mitjana d’un model de referència. L’algoritme es va avaluar amb imatges de pacients diagnosticats amb dues patologies cardíiques diferents, mostrant l’efectivitat del mètode. Finalment, en la tercera de les contribucions, es va presentar un mètode per detectar i segmentar el VD en grans bases de dades de manera automàtica i avaluar-ne el resultat sense interacció manual. El mètode incloïa un sistema de regressió per predir la posició inicial del cor, i el control de qualitat estava basat en descriptors de textura. La validació, basada en dos grans cohorts d’accés públic, va demostrar que el mètode era adequat per a l’anàlisi automàtica de grans bases de dades.
CHAPTER 1

General Introduction
1.1 The Era of Big Data

The thesis emerges in the context of the so-called big data revolution, which continues to attract significant attention in research and beyond. According to a description by IBM [1], “Big data is being generated by everything around us at all times. Every digital process and social media exchange produces it. Systems, sensors and mobile devices transmit it. Big data is arriving from multiple sources at an alarming velocity, volume and variety. To extract meaningful value from big data, you need optimal processing power, analytics capabilities and skills”. The general characteristics of big data have been widely defined by three major features, commonly known as the 3Vs (see Fig. 1.1): volume, variety, and velocity [11, 60]. In effect, the current statistics show that approximately 2.5 quintillion bytes of data are created every day and that 90% of all of the world’s data were produced in the past two years alone [1]. Furthermore, the volume of global data overall is increasing exponentially, from 130 exabytes (an exabyte is $10^{18}$ bytes of data) in 2005 to 7.910 exabytes in 2015. By 2020, there will be 35 zettabytes ($10^{21}$ bytes) of digital data. However, only 20% of the world’s data is structured (suitable for statistical processing), with unstructured data (e.g., untagged text, image, audio and video files) growing at 15 times the rate of structured data [40]. The processing of such large amounts of data is expected to revolutionize our lives in many areas. For this, it is important to develop new techniques to process both structured and unstructured data, as well as to extract new information and knowledge from such data to impact our modern societies. As put forward by Prof. G. King, Albert J. Weatherhead III University Professor at Harvard University, the revolution does not lie so much in the data but more in the development of new statistical and computational methods that will change our ability to process and exploit these newly available large-scale data [58].

1.1.1 Big Data in Health-care and Biomedicine

Extremely large data volumes at high velocities were originally the realm of supercomputers, nuclear physics, military simulations and space travel. Starting in 1990, the Human Genome Project was the moon launch of big data in health-care, a data-intensive research effort that pushed the limits of available data processing technology. Increasingly powerful hardware and software, improvements in data management and integration, new analytic tools, and accumulating experience, are building a foundation for the increasing use of big data and analytics in health-care [40].

The health-care industry is currently generating large amounts of data, driven by record keeping, compliance and regulatory requirements, and patient care [102]. While most data is stored in hard copy form, the trend of the last decades is toward rapid digitization of these large amounts of data. Driven by mandatory requirements and the potential to improve the quality of health-care delivery meanwhile reducing the costs, these massive quantities of data hold the promise of supporting a wide range of medical and health-care functions, including, among others, clinical decision support, disease surveillance, and population health management.

By definition, big data in health-care refers to electronic health data sets so large and complex that they are difficult (or impossible) to manage with traditional software and/or hardware; or they can not be easily managed with traditional or common data management tools and methods [2]. With technological advances, the tendency to adopt digital platforms for health information, and the computerization of business
operations, it is easy to understand why a flood of big data is inundating the healthcare sector. This amount of data is overwhelming not only because of its volume but also because of the diversity of data types and the speed at which it must be managed [2]. There are two major types of digital data in healthcare: clinical records and health research records.

Clinical records include electronic medical records (EMRs), digital images, and information-sensing wireless medical devices. EMRs and longitudinal patient medical records containing multiple file types are the center of computerized health information systems. Medical images are now created digitally and stored in Picture Archive and Communication Systems (PACS). They are the largest contributor to the expanding volume of big data in healthcare [39]. And, as medical imaging devices improve in digital resolution capabilities, the data files dramatically increase in size. Moreover, information-sensing wireless medical devices have also the potential for being one of the major contributors to the flood of big data. With advances in sensor technology, there has been rapid growth in the number of wireless medical devices that continuously monitor patients and send reports to providers.

Health research, from drug development to biotechnology to public health, has long been data-intensive, and today its output is growing exponentially as well. Additionally, research often generates large amounts of raw data that remains unstructured and unlabeled, effectively making it useless to other researchers who then must expend time and resources collecting and storing similar data.

1.1.2 Big Data in Medical Imaging

Medical imaging data is, as we pointed out before, the largest contributor to the expanding volume of big data in healthcare. Millions of medical images are generated each year. The following are some eye-opening statistics [39, 116]:

- Medical image archives are increasing by 20-40 percent each year. In the U.S., for example, by 2012 there was 1 billion medical images stored.
It is estimated that medical imaging information storage constitutes one-third of global storage demand, which, in 2007, was the equivalent of 1.2 billion average hard drives.

By 2015, the average hospital in the U.S. had two-thirds of a petabyte (665 terabytes) of patient data, 80% of which was unstructured image data like CT and MR scans.

This trend has been accelerated by standard medical practice, which is moving from relatively subjective decision making to evidence-based health-care, making use of complex diagnosis technologies, such as comprehensive laboratory analysis and powerful imaging techniques. Moreover, in countries such as in the Western World, there are incentives to professionals/hospitals to use EMR technology [88]. On the other hand, the developed capturing devices and mobile applications are also having a significant impact on the increase of available imaging data.

Big medical imaging data is a new paradigm that contributes to a transformation of case-based studies to large-scale, data-driven research [78]. In this context, the above-mentioned three Vs of big data (see Fig. 1.1) could be understood as follows:

First and most significantly, the volume of data is particular high in the biomedical fields [86]. Data from millions of patients have already been collected and stored in an electronic format, and these accumulated data could potentially enhance health-care services and increase research opportunities. In addition, the advances in medical imaging (e.g., MRI, CT scans), which mean vast amounts of data with even more complex features and broader dimensions, are produced routinely in clinical practice, as well as in clinical research. One such example is the UK Biobank project [95], which is a large population cohort study with 500,000 participants, developed to study a wide range of poorly understood diseases such as cardiovascular diseases (CVDs) in order to improve prevention and therapy in the context of personalized medicine.

The second feature of big data is the variety of data types and structures. The structures of biomedical big data comprise many different levels of data sources to create a rich array of data for clinicians and researchers. In the case of medical imaging, for example, this variety refers to the numerous of imaging modalities and sequences that exist, as well as to the fact that the image data are gathered from different sources (e.g., from different clinical centers) or using different protocols. As an example, a widely used modality such as magnetic resonance imaging (MRI), which is the imaging modality used in this thesis, can be used to scan individuals in various protocols (e.g., cine-MRI, Late-Enhancement MRI, perfusion MRI), with different clinical assessments [59].

The third characteristic, velocity, refers both to the acquisition and processing of the imaging data. While the developments in biomedical hardware and devices has accelerated data acquisition, fast processing of big data in health-care remains a significant challenge. For instance, the processing of the image scans have been done for a long time through expert radiologists, or by using semi-automated software to extract quantities and biomarkers from the images for clinical diagnostics. However, the use of humans or semi-automated software becomes impractical when dealing with huge amounts of imaging scans.
Consequently, to fully exploit large-scale biomedical data and to realize the promise of big data in health-care practice, it is important to develop a new generation of computational techniques that are capable to handle the large volume of biomedical imaging data, to deal with their variability in terms of sources and formats, and to process them rapidly and automatically.

### 1.1.3 Estimating New Biomarkers from Large-scale Studies

Among large-scale studies, a colossal pool of knowledge remains largely frozen and unexplored, and consequently, research innovations in big image data analysis are expected to impact and transform the future of biomedicine [103] both in clinical practice (how to diagnose and treat patients based on historical image data) and in medical research (how to understand pathologies and their progressions based on novel imaging biomarkers, how to improve and develop treatments).

Currently, under the one-size-fits-all medical practice, patients with similar symptoms all too often receive the same treatments, which limits their efficacy. Furthermore, for a lot of major diseases such cardiovascular and neurological disorders, diagnosis is often made at late symptomatic stages, which leads to late interventions and decreased efficacy of medical care. In this context, one promising avenue is the quantification of new biomarkers from large-scale databases of clinical and biological data. The term *biomarker*, a portmanteau of ‘biological marker’, refers to a broad subcategory of medical indices (objective indications of medical state observed from outside the patient), which can be measured accurately and reproducibly [111]. From a clinical perspective, biomarkers have a variety of functions, which correspond to different stages in the development of a disease, from early to late stage. Biomarkers can assist in the care of patients who have no apparent disease (screening biomarkers), those who have the risk of having an illness (antecedent biomarkers), those who are suspected to have disease (diagnostic biomarkers), and those with overt disease (prognostic biomarkers) [122]. For example, diagnostic and prognostic cardiovascular biomarkers are available, but there are no widely accepted biomarkers for screening. This has been an active area of investigation, because preventing events, for example in those at risk of cardiovascular disease, is likely to have a substantial impact on the overall public-health burden [46]. There are barriers to identifying new biomarkers, particularly for screening or prognostic uses. One difficulty is the requirement for large, adequately powered clinical studies. Large studies are necessary because the predictive effects of new biomarkers might be smaller than those observed with classic risk factors and because multiple biomarkers are often studied concurrently. Meta-analysis of individual participant data from multiple cohorts provide a potentially valuable tool for circumventing sample-size limitations in single cohorts [46].

Applying big data strategies to better inform decision making is particularly relevant to the practice of personalized medicine that aims to individualize the diagnosis of a disease and therapy according to the individual patient characteristics (e.g., clinical co-morbidities and genetics), as opposed to decisions based on evidences and guidelines derived from case-based studies and clinical trials. A major emphasis of personalized medicine is to match the right drug with the right dosage to the right patient at the right time. The process of personalized medicine could also be facilitated with the comparison of a new patient to patients with similar characteristics. This could lead to faster and more accurate diagnoses and consideration of therapeutic options [90]. In this context, the extraction of accurate and reproducible
biomarkers from large-scale databases to enable early diagnosis and individualized treatments is one of the most important goals of personalized medicine research. The extraction of new biomedical knowledge using large-scale medical image data will be achieved through data analytics, but with assistance from other fields including applied mathematical methods, machine learning, data visualization, and medical image computing. This is because images constitute essentially of pixels and voxels, resulting in extremely high-dimensional feature spaces of low semantic value. Fundamentally, the goals of medical image computing are to process, aggregate, and reduce these raw intensity image signals, i.e. to transform the images into higher-level representations (e.g., clusters, labels, shapes, biomarkers, etc.).

1.1.4 The Case of Cardiovascular Diseases and Cardiac Imaging

Despite the advances in cardiovascular image analysis and in cardiovascular research over the last few decades, CVDs still remain the leading causes of deaths worldwide, according to a report of the World Health Organization (WHO) [128], including in low and middle-income countries. A lot of CVDs, such as coronary heart disease and cardiomyopathy, are often diagnosed at late stages or after serious events occur. It is therefore of paramount importance to identify new cardiovascular phenotypes and biomarkers from large-scale studies for stratified diagnosis of CVDs at early stages. To analyse large-scale cardiovascular studies, an important pre-requisite is to automatically and robustly process the image data and extract information about the cardiovascular morphology and function using segmentation techniques. The aim of this thesis is to develop such techniques capable of segmenting highly numerous and highly variable cardiovascular image data without the need for the assistance of expert clinicians, which indeed becomes impractical in the case of large studies. To enable high throughput analysis of cardiac imaging data, a first step is the delineation automatically and reliably of the myocardial boundaries.

To provide application focus and concrete value to the thesis, we will address the specific case of population studies in cardiac MRI, which provides non-invasive and radiation-free means to evaluate, among other things, the cardiac function and dynamics or the myocardial infarction [106]. When applied in the appropriate clinical context, the information obtained from MRI can offer diagnostic and prognostic information otherwise unavailable to the clinician. Moreover, cardiac MRI provides high spatial resolution images of the heart in every potential spatial plane without being limited by acoustic windows, as is echocardiography, and without the need to expose the patient to ionizing radiation or iodinated contrast, as in computed tomography [16, 62]. Many advances have led to the effective utilization of cardiac MRI in evaluation of myocardial morphology, ventricular function, and myocardial tissue composition; structural and functional assessment of the coronary vessels; evaluation of arterial and venous vasculature; and assessment of flow dynamics to quantify valvular competence, regurgitant volumes, and intra- and extracardiac shunting [16, 106].

Cardiac MRI allows the quantitative evaluation of cardiac function through several indices such as ventricular volumes, masses and ejection fraction [93], which involve the previous delineation of the myocardial wall. Specifically, the first step of the cardiac analysis is the segmentation of the endo- and epicardial contours on every slice of the MR volume, which is the main goal of the methods developed in this thesis, in order to perform the subsequent computation of clinical parameters: ejection fraction, systolic and diastolic volumes, myocardial mass, etc.
1.2. Aim and Objectives of This Thesis

The main aim of this thesis is to present versatile and automatic tools to tackle some of the challenges derived from the analysis of large amount of cardiac MRI sequences. Ultimately, these tools should be useful for delineating the ventricular boundaries for a subsequent cardiac analysis. These methods ideally should be able to deal with the inherent inconsistencies and changes in image sequences. Moreover, the approaches should be robust to different cardiac pathologies and patient groups. Finally, when
dealing with large-scale studies, one would expect the methods to be fully automatic. This thesis has been structured around the following specific objectives:

- To provide a method to segment the ventricular boundaries independently of the MRI sequence at hand (e.g., cine-MRI vs. LE-MRI).
- To develop a method to segment the ventricular boundaries independently of the severity of the pathology.
- To develop a method to segment and assess the ventricular boundaries without the need of any user interaction.

1.3 Contributions

The proposed solutions to each of the three objectives are presented in the next chapters, which represent three specific contributions, with one chapter dedicated to each of them:

- Chapter 2 describes an automatic approach, which can be applied across MRI sequences without the need for sequence-specific tuning. Furthermore, the type of constraints used allows the method to be able to adapt to the inherent inconsistencies in image sequences. A generic rule-based framework is used to automatically segment the left ventricular (LV) endocardial (inner) and epicardial (outer) boundaries. This is accomplished using intensity information, and including shape and inter-slice smoothness constraints. The segmentation algorithm uses a decoupled, modified graph cut approach with control points. The segmentation was evaluated on two MR sequences, late-enhancement (LE) MRI and cine-MRI, showing a good balance between flexibility and robustness.

- Chapter 3 proposes a new algorithm for the segmentation of severely deformed pathological hearts in MRI. The algorithm is highly flexible, as it does not require a priori knowledge of the involved pathology or any specific parameter tuning to be applied to the cardiac image under analysis. The idea of the algorithm is to approximate the abnormality with a virtual remodeling transformation between the patient-specific geometry and the average shape of a reference model. To define this mapping, a set of landmark points are automatically identified during boundary point search, using the reliability of the candidate points. With the obtained transformation, the feature points extracted are then projected onto the space of the reference, where the model is used to effectively constrain and guide the segmentation process. The obtained shape is finally propagated back to the original image space. The algorithm was validated with patients diagnosed with Pulmonary Hypertension (PH) and Hypertrophic Cardiomyopathy (HCM) showing the effectiveness of the technique for the segmentation of pathological hearts with gross morphological variability.

- Chapter 4 presents a fully automatic framework, which parses images to perform LV detection, LV segmentation and quality control without any user interaction for large-scale cardiac MRI samples. The method includes a regression-based approach to predict the initial location of the heart and of its key anatomical landmarks using synthetic training samples to learn and predict displacements based on a set of imaging features and random forests. Furthermore,
anatomically-justified texture descriptors are used at the end of the segmentation to verify the quality of the result for failure detection and removal from the subsequent population analysis. Detailed validation based on two public large databases from the Cardiac Atlas Project showed the promise of the approach for user-independent analysis of such population.

Each of these three chapters are self-contained, and are an adaptation of the articles that were submitted or published in a peer-reviewed journal. Finally, Chapter 5 concludes the thesis and discusses the outlook and future work.

1.4 Academic Context and Funding

This research work has been carried out within the Center for Computational Imaging and Simulation Technologies in Biomedicine (CISTIB), at the Universitat Pompeu Fabra (UPF), Barcelona, Spain. This group is directed by Prof. Alejandro F. Frangi. This thesis is associated with the work of Dr. Rosa M. Figueras i Ventura and Dr. Karim Lekadir. Additionally, thanks to the active collaboration of CISTIB with clinical centers and other research groups in Spain, UK and New Zealand, we had access to unique image databases for validation, as well as to fruitful feedback on the value of the techniques and results.

This thesis was supported by a departmental scholarship from the Department of Information and Communications Technologies at UPF and by an FPU grant from the Spanish Ministry of Education Culture and Sport.
Automatic Cardiac LV Segmentation in MRI using Modified Graph Cut with Smoothness and Inter-slice Constraints

The content of this chapter is based on the publication:

Abstract

Magnetic resonance imaging (MRI), specifically Late-Enhancement MRI (LE-MRI), is the standard clinical imaging protocol to assess cardiac viability. Segmentation of myocardial walls is a prerequisite for this assessment. Automatic and robust multi-sequence segmentation is required to support processing massive quantities of data. A generic rule-based framework to automatically segment the left ventricle (LV) myocardium is presented here. We use intensity information, and include shape and inter-slice smoothness constraints, providing robustness to subject- and study-specific changes. Our automatic initialization considers the geometrical and appearance properties of the LV, as well as inter-slice information. The segmentation algorithm uses a decoupled, modified graph cut approach with control points, providing a good balance between flexibility and robustness. The method was evaluated on LE-MRI images from a 20 patient in-house database, and on cine-MRI images from a 15 patient open access database, both using as reference manually delineated contours. Segmentation agreement, measured using the Dice coefficient, was $0.81 \pm 0.05$ and $0.92 \pm 0.04$ for LE-MRI and cine-MRI, respectively. The method was also compared favorably to a 3D Active Shape Model approach. The experimental validation with two MR sequences demonstrates increased accuracy and versatility.
2.1 Introduction

Multisequence magnetic resonance imaging (MRI) plays an increasingly important role in the assessment of myocardial function and viability [7, 57, 124]. MRI is a non-invasive medical imaging technique used in radiology to visualize detailed internal structure and limited function of the body. The imaging protocol consists of the acquisition of different MRI sequences, including cine steady-state free precession (cine-MRI) and Late-Enhancement MRI (LE-MRI). Cine-MRI is an anatomical image, which highlights the soft tissue, and LE-MRI, a delay gadolinium enhancement image, highlights the necrotic tissue. Areas of nonviable myocardium retain extremely high signal intensity, due to the accumulation of gadolinium in the extracellular space, whereas black areas show normal tissue. However, the myocardial boundaries are hardly enhanced, whereas in the cine-MRI the myocardium is clearly visible without tissue viability information. To enable automatic and reliable interpretation in clinical practice of such data, a number of challenges must be addressed. Computational imaging tools for tackling these challenges need to be versatile, adapt to the inherent inconsistencies and changes in image protocols and must be capable of handling data- and subject-specific characteristics.

In the analysis of MR images by segmenting the myocardial walls, significant inconsistencies can be observed, which can roughly be classified in three categories. First, different segmentation tools are generally used for distinct MRI techniques and sequence designs (e.g., cine- and LE-MRI), which leads to inconsistencies among segmentations. Many approaches have been published for the automatic segmentation of cine-MRI (see review in [96]) but most of them are inadequate for the more challenging LE-MRI case. This has motivated the development of techniques specifically for LE-MRI data [3, 25], which means multiple segmentation tools have to be used for a single multi-sequence patient study. Alternatively, it has been proposed that LV boundaries segmented from the cine-MRI are propagated into the LE-MR images [13, 34, 126, 127]. However, spatial and temporal variations are inevitable between the two acquisitions, while the application of an image registration technique is complicated by the differences in the number of short axis (SA) slices and appearance properties.

Moreover, the morphological variations across individuals and groups of individuals also cause inconsistencies between segmentations. For example, the statistically based Active Shape Model (ASM) [27] has been extensively adapted for myocardial segmentation [65, 121, 129] but it has been shown to generalize poorly to abnormal cases that differ from the data used for training, a situation that is common in practice [64, 125].

Finally, a third source of segmentation variability comes from the fact that most techniques require a certain degree of user interaction, generally through an initialization stage [19, 119], which to a high extent determines the accuracy and convergence of subsequent boundary localization. The use of an automatic and reliable initialization technique would therefore significantly reduce the intra- and inter-observer variability.

In this chapter, we present a generic rule-based framework for consistent LV myocardial segmentation sequence-independent MRI (in SA view), with fully automatic initialization. With this work, a single segmentation tool can be applied across MRI sequences without the need for subject-specific tuning or for user initialization. Fur-
thermore, the type of constraints used in this approach introduces a good balance between robustness and flexibility, which makes our technique promising for clinical practice. The validation is carried out based on cine- and LE-MRI datasets including normal and pathological cases.

2.2 Methods

2.2.1 Overview

The algorithm consists of a two-stage initialization to robustly localize the LV, and a two-stage segmentation using slice-by-slice graph cuts (GC) with inter-slice and shape constraints. Both parts are based on a rule-based formulation developed to achieve two key objectives:

- to impose morphological constraints that are common across MRI sequences;
- to achieve robustness to variations in grey-level appearance and to image inhomogeneities.

To achieve this, we define a set of rules to reflect appearance, geometrical and topological LV characteristics to constrain myocardial initialization and segmentation. The six rules are listed in Table 2.1, along with the rationale behind their use in this work for multi-sequence MRI segmentation. To obtain generic morphological constraints, it is important to note that these features are common, or can be generalized, to all MRI sequences and classes of LV shapes. Although conceptually simple, we will show that with these rules used in the right sequence and/or combinations, a reliable initialization of the LV and a consistent segmentation of the myocardium across MRI sequences and subjects can be obtained, as detailed in subsequent sections and illustrated in Fig. 2.1.

2.2.2 LV Initial Localization

Automatic initialization is an important part of this work, in order to reduce user interaction, thus improving repeatability and feasibility to automatically process large multi-sequence MRI databases. The few existing techniques [22, 66, 91] are computationally expensive due to the large search space spanned by the image volumes and the number of potential solutions they consider. To simplify this task and to reduce the search space, the proposed framework uses an algorithm that takes advantage of some key characteristics of the LV (in particular R1 and R2) to derive an efficient and reliable automatic initialization.

2D Template-Based Initialization

The first step of the automatic initialization is to find an approximate location for the LV in each SA image. The idea is to scan each slice with a 2D template that describes the constraints R1, R2 and R4. In other words, the template will describe the appearance and shape of the blood pool (white circle) and of the myocardium (dark ring) as illustrated in Fig. 2.2(a). Note that in this case we use a simplified version of R1, which considers a brighter blood pool appearance. Mathematically,
### Table 2.1: List of LV properties used as rules to constraint myocardial initialization and segmentation in multi-sequence MRI

<table>
<thead>
<tr>
<th>Rule</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>R1</strong></td>
<td>The blood pool has a bright contrast appearance. Used for 2D initialization of the SA images, as well as for endocardial segmentation.</td>
</tr>
<tr>
<td><strong>R2</strong></td>
<td>The blood pool has a pseudo-circular shape at each SA. Used to prevent incoherent changes along in the endocardial wall (e.g., due to papillary muscle or scar tissue, see Fig. 2.2).</td>
</tr>
<tr>
<td><strong>R3</strong></td>
<td>The blood pool is continuous in the longitudinal direction. Used to prevent incorrect slice initializations that can be obtained during 3D initialization.</td>
</tr>
<tr>
<td><strong>R4</strong></td>
<td>The myocardium could include healthy (dark voxels) and diseased tissue (bright voxels) (myocardial intensity not homogeneous). Used to adapt to the presence of heterogeneous texture appearance in the myocardium.</td>
</tr>
<tr>
<td><strong>R5</strong></td>
<td>The myocardial thickness changes smoothly. Used to prevent incoherent changes in myocardial thickness (e.g., due to poor epicardial wall appearance or surrounding artifacts).</td>
</tr>
<tr>
<td><strong>R6</strong></td>
<td>The 3D global shape of the myocardium is smooth. Used to ensure consistency in myocardial wall definition between the SAs.</td>
</tr>
</tbody>
</table>

\[ [z_0, d_0, w_0] = \arg \max_{d,w,z} (T_{d,w} * P)(z), \] (2.1)

where \( d \) and \( w \) are the radius and thickness of the kernel respectively, and \( z \) is the location of the center in the slice, which is completely scanned. The values of the thickness \( w \) go from 5 to 20 mm and the radius \( d \) from 17 to 35 mm, both ranges have been defined using anatomical heart information reported in the literature [61]. The pixel resolution of the image is taken into account to adapt the values to pixels.

### Initialization Refinement

Due to the presence of surrounding structures and image artifacts, incorrect slice initializations might occur during the previous step. Mainly, the initialization at the most apical and basal image slices can be complicated due to the structural definition. In order to improve the robustness of the initialization, the constraint R3 is now used to identify incoherent initializations by comparing the results from all slices. More specifically, since the blood pool is a continuous structure in the direction of the...
Figure 2.1: An example of myocardial segmentation in a LE-MR image (a). In (c), the incorrect segmentation due to the scar being mistaken for a region of the blood pool is corrected (d) by the application of the pseudo-circular shape constraint (R2). In (e), an incorrect homogeneous myocardium is localized, while the heterogeneous constraint (R4) enables to obtain a segmentation that contains both healthy and unhealthy tissues (f), similar to the manual delineation in (b).

Figure 2.2: In the initialization step, a Laplacian of Gaussian template (a) is used to estimate the position, radius and width of the LV myocardium (b).
2.2. Methods

Figure 2.3: Schematic illustration of the proposed initialization. (a) shows the results of the 2D template-based initialization, with a few incorrect initial contours, which are detected in (b) based on the refined search window obtained from the median center (dashed line). (c) shows the new round of 2D initializations, leading to reliable result consistency along all SA images.

longitudinal axis, it is expected that the correct 2D initializations overlap each other, while erroneous initializations would be located further apart in the image volume, as illustrated in Fig. 2.3(a). Consequently, after the first round of 2D initializations described in the previous section, the median xy-location of the obtained centers for all the short axis (SA) images is computed. Using this median, incorrect slice initializations could be detected as illustrated in Fig. 2.3(b). Subsequently, these are corrected by applying a new 2D initialization procedure to the corresponding SA images that focuses on a window (15 mm of radius) around the obtained median center, thus allowing for a more accurate detection of the LV location consistently along the blood pool axis (Fig. 2.3(c)). Algorithm 2.1 summarizes the listing for the proposed LV initialization technique.

The initialization step provides an initial set of labeled points, used in the posterior graph cut process, and an initial circular shape $O_o$, defined by a center $z_o$, a radius $d_o$ and a thickness $w_o$, which will be used as a prior in the following steps (see Fig. 2.2).

Next section shows how this initial LV estimate is used in combination with the remaining LV constraints of Table 2.1 to search for the myocardial boundary walls.

Algorithm 2.1 Automatic 3D LV initialization

**INPUT:** $M$ SA image $P_k$, $k = 0, ..., M$

for each SA image $P_k$ do

Calculate image response for various templates $T_{d_i,w_i}$

Choose optimal template $T_{d_k,w_k}$ and corresponding LV center $z_k$ according to Eq. (2.1)

end for

$z_m = median(z)$ (median of the LV centers for all slices)

for each SA image $P_k$ do

Calculate a region of interest $R_k$ around the median center $z_m$

if $z_k$ is not in $R_k$ then

Update $T_{d_k,w_k}$ and $z_k$ according to Eq. (2.1) for $z_k \in R_k$

end if

end for

**OUTPUT:** Endocardial initial contours $O_o(d_o,w_o,z_o)$ for each slice
2.2.3 LV Segmentation

The aim of this section is to present a formulation that allows to adequately combine the various LV constraints into an efficient image matching framework. To this end, we adapt the graph cut (GC) formulation from [18, 19] taking advantage of the flexibility it offers [5, 9, 44, 110] for adding constraints of different nature (region, edge, geometric, structural, see Table 2.1) for robust and flexible LV segmentation. At the same time, we obtain good convergence properties for distinct MRI sequences and classes of LV shapes. Algorithm 2.2 summarizes the steps for the LV segmentation technique.

Graph Cut Formulation

In the GC paradigm, an image is represented by a graph \( G = (V, E) \), consisting of a set of nodes, \( V \), and a set of directed edges, \( E \), that connect them. The nodes are \( V = \{s, t\} \cup P \), where \( P \) is the set of nodes that represent all the pixels of the slice, and \( s \) and \( t \) are auxiliary nodes, called terminals, corresponding to the set of labels that can be assigned to pixels. Each node \( p \in P \) is connected to both \( s \) and \( t \) via so-called t-links and each node \( p \) is also connected to neighbors using a neighborhood system \( N \), through n-links. The algorithm finds the cut with the minimum cost (the sum of link weights of the cut), which corresponds to the global energy minimum of the segmentation, assuming the submodularity condition is satisfied [18].

The cost of n-links corresponds to a penalty for discontinuity between the pixels, which is usually derived from the pixel interaction term \( I_{p,q} \). The cost of a t-link connecting a pixel and a terminal corresponds to a penalty for assigning the corresponding label to the pixel. This cost is normally derived from the data term \( D_p \) in the energy equation. The energies of the min-cut algorithm, which will be minimized, can be represented as

\[
E(L) = \sum_{p \in P} D_p(L_p) + \lambda \sum_{(p,q) \in N} I_{p,q}(L_p, L_q),
\]

where \( L = \{L_p| p \in P\} \) is a labeling of the image \( P \) and \( L_p \in \{0, 1\} \) is the segmentation label of pixel \( p \), where 0 and 1 correspond to the outside and the inside of the region, respectively. Moreover, as described above, \( D_p(L_p) \) is a data term (which measures the penalty of assigning label \( L_p \) to pixel \( p \)) and \( I_{p,q}(L_p, L_q) \) is an interaction term (which measures the penalty of assigning labels \( L_p \) and \( L_q \) to the neighboring control points \( p, q \), respectively). Finally, \( \lambda \) is the ratio of importance between both terms.

In our implementation, we have resampled the object contours equidistantly to obtain a set of control points, which will be special nodes of the graph used to for one of the energy terms. In that way, \( C = C_h \cup C_v \) will be the set of neighboring control points that includes two subsets: \( C_h \), control points in within-slice direction, and \( C_v \), control points in inter-slice direction (see Fig. 2.4). The control points, which are recomputed at every iteration, have an associated label corresponding to the object region, i.e., \( \forall p \in C, L_p = 1 \). Their use limits the amount of operations that need to be performed at each slice, making computations more efficient.
2.2. Methods

Blood Pool Segmentation

The boundary between the blood pool and the background is obtained by finding the minimum cost cut on the graph $G$, which minimizes an energy function. Since the shape of the blood pool is approximately circular (property R2 from Table 2.1), a smoothness shape constraint is incorporated into the energy function, which is formulated as

$$E(L) = \sum_{p \in P} D_B^p(L_p)$$

$$+ \lambda \sum_{p, c \in N_c, r, u \in C_h, (c, r, u) \in C_v} Y_c(p, r, u, L_N_c)$$

$$+ \mu \sum_{p \in P, (p, q) \in N} S_{p, q}(L_p, L_q, O_o),$$  \hspace{1cm} (2.3)

where $D_B^p$ is the blood pool $B$ data term, $Y_c$ is the inter-slice constraint term and $S_{p, q}$ is the shape constraint term.

The data term $D_B^p$ is expressed as follows:

$$D_B^p(L_p) = -|\nabla J(p)|^2 \cdot (J(p) - J_o) \cdot L_p,$$ \hspace{1cm} (2.4)

where $|\nabla J(p)|^2$ is the magnitude of the intensity gradient in pixel $p$, $J(p)$ is the intensity of the pixel, and $J_o$ is the initial mean intensity of the object, which is a good compromise between speed and performance. This term imposes a penalty based on gradient and intensity of the pixels. The contour of the region is attracted by the local maxima of the gradient of the pixels with higher intensity value, based on R1 in Table 2.1.
The term $Y_c$ in Eq. 2.3 includes inter-slice smoothness (R3 in Table 2.1) and is computed as:

$$Y_c(p, r, u, L_{N_c}) = \left| p - \frac{r + u}{2} \right| \cdot B(p, L_{N_c}),$$  

(2.5)

where $B(p, L_{N_c})$ is a function, which is 1 if $p$ is in the boundary or 0 otherwise, according to the labels of the neighborhood $L_{N_c}$. This term controls the variability of the myocardium in the inter-slice direction, where $p$ is part of $N_c$, which is the set of points belonging to the neighborhood in the radial direction of each control point $c$ in $C_h$ of the current slice. $r$ and $u$ are the two neighbor vertical control points from the slices above and below the current slice (see Fig. 2.4). This term is computed for some nodes and not for the whole image to speed up the process. The control points $C_h$ and $C_v$ are recomputed at every stage according to the segmentation obtained, as detailed in Algorithm 2.2.

The circular shape constraint term (R1 from Table 2.1) is defined as the unsigned point-to-curve distance from a boundary point (defined as the midpoint $(p + q)/2$) to the initial circular shape $O_o$,

$$S_{p,q}(L_p, L_q, O_o) = \text{dist} \left( \frac{p + q}{2}, O_o \right) \cdot F(L_p \neq L_q),$$  

(2.6)

where $F(L_p \neq L_q)$ is 1 if $L_p \neq L_q$ and 0 otherwise. For a pair of neighboring pixels $p$ and $q$, one belonging to the object and the other to the background, if they are close to the initial circular shape, it would satisfy

$$\text{dist}((p + q)/2, O_o) \approx 0,$$

(2.7)

and the energy would be low. Consequently, under the same conditions, the cost of a cut near the circle would be smaller than a cut farther away from the circle.

**Myocardium Segmentation**

With the robust localization of the blood pool achieved in the previous step, an initialization of the epicardium segmentation is available by dilating the obtained blood pool boundary. The algorithm uses the basic morphological operation of dilation with a rounded structural element of width $w_o$ obtained in the initialization step.

Our energies for this region are similar to the ones used in the blood pool segmentation

$$E(L) = \sum_{p \in P} D^M_p(L_p)$$

$$+ \lambda \sum_{p, c \in N_c, \ c \in C_h, \ (c, r, u) \in C_v} Y_c(p, r, u, L_{N_c})$$

$$+ \mu \sum_{p \in P, \ (p, q) \in N} W_{p,q}(L_p, L_q, O_b).$$  

(2.8)

The term $Y_c$ is also defined as described in Eq. 2.5. Referring to the data term $D^M_p$ of the myocardial region $M$, we employ the following equation:

$$D^M_p(L_p) = -|\nabla J(p)|^2 \cdot L_p,$$

(2.9)
2.2. Methods

**Algorithm 2.2 Automatic LV segmentation**

**INPUT:** \(M\) SA images \(P_k, k = 0, \ldots, M\) and initial blood pool region labeling \(L_k = \{L_p | p \in P_k\}\)

```plaintext
while Not converged do
  for each SA image \(P_k\) with a labeling \(L_k\) do
    Extract endocardial boundaries
      ↯ Find the minimum cost cut using Eq. (2.3)
      Update image labeling \(L_k\)
      Update control points \(C_v\) and \(C_h\)
  end for
end while

while Not converged do
  for each SA image \(P_k\) with a labeling \(L_k\) do
    if First iteration then
      Initial LV region labeling
      ↯ Dilation of the blood pool boundary using \(w_o\)
    end if
    Extract epicardial boundaries
      ↯ Find the minimum cost cut using Eq. (2.8)
      Update image labeling \(L_k\)
      Update control points \(C_v\) and \(C_h\)
  end for
end while

**OUTPUT:** Final image labeling \(L_k\) for each \(P_k, k = 0, \ldots, M\) (endocardial and epicardial final contours).
```

where \(\|\nabla J(p)\|^2\) is the magnitude of the intensity gradient in pixel \(p\). This term imposes a penalty based on gradient of the pixels; disregarding the pixel intensity allows the inclusion of both dark and bright pixels, as listed as R4 in Table 2.1.

Since the thickness of the myocardium is expected to have low variations (R5 in Table 2.1), we use the shape of the blood pool as a constraint. The \(W(\cdot)\) is defined as the unsigned distance from the centroid of \(p\) and \(q\) to the blood pool contour \(O_b\) (obtained in the previous step) and the estimated thickness \(w_o\) (obtained in the initialization step):

\[
W_{p,q}(L_p, L_q, O_b) = \left| \text{dist} \left( \frac{p + q}{2}, O_b \right) - w_o \right| \cdot F(L_p \neq L_q), \tag{2.10}
\]

penalizing a different labeling of a pair of neighboring pixels \(p\) and \(q\) if they are far from the blood pool shape.

Note that each stage of the algorithm uses two weighting factors during image search. This simplifies significantly the definition of the optimal value of the parameters and improves the independence from the database of the method’s performance. The values of the weighting were found empirically and fixed \((\lambda = 0.25\) and \(\mu = 0.8\)) for the LE-MRI dataset. Our method is implemented in Matlab based on a mincut-maxflow algorithm as described in [20] to compute minimal cost in the graph.
2.3 Results

2.3.1 Data Description

LE-MRI Dataset

We have evaluated our approach in an in-house LE-MRI database of 20 patients. The studies were acquired using a GE Signa CVi-HDx, 1.5T scanner (General Electric, Milwaukee, USA). Short axis (SA) images were acquired with a late gadolinium enhancement (LGE) inversion recovery sequence (after IV administration of 0.2 mmol/kg of gadopentate dimeglumine contrast). Each dataset contained between 8 and 12 SA images, covering the LV from the atrioventricular ring to the apex. The slice thickness was 8 mm with an in-plane pixel resolution of 1.56 mm × 1.56 mm and spacing between slices between 8 and 10 mm. The left ventricle was manually segmented by two experienced researchers on all slices of the 20 datasets.

Cine-MRI Dataset

We have also tested the method using 15 patients from the Sunnybrook Cardiac Database [101], also known as the 2009 Cardiac MR Left Ventricle Segmentation Challenge data. The 15 patients correspond to the portion used as the test set group. The database consists of cine-MRI images from a mix of healthy subjects and patients with hypertrophy, heart failure with infarction and heart failure without infarction. Cine-MR SA images were obtained with a 1.5T GE Signa MRI (General Electric, Milwaukee, USA). Between 6 to 12 SA images were obtained from the atrioventricular ring to the apex, the slice thickness was between 8 mm and 10 mm, with an in-plane pixel resolution of 1.36 mm × 1.36 mm and spacing between slices of 8 mm. Endocardial and epicardial contours were drawn by an experienced cardiologist in all slices at the end-diastolic and end-systolic phases and confirmed by another cardiologist.

2.3.2 Reference Technique: 3D-ASM Segmentation

To evaluate the performance of our method, we compare it to a well-established segmentation approach: the 3D-ASM. The essential components of a (3D-)ASM are a shape model, an intensity model, and a matching algorithm (full description in [27]). The shape model is a template of the organ of interest represented as a distribution of landmarks or Point Distribution Model (PDM). Our shape model is a surface mesh representing the left ventricle. Our matching algorithm is based on the sparse fitting method SPASM [121].

To deal with multimodal data, the shape and intensity components of the ASM were decoupled. The shape information was trained from a large CT database [89] and the intensity information was trained from MRI datasets. The CT database used for shape model training had higher resolution, thus, the shape model is more reliable. This approach has proven to obtain accurate segmentation results on cine-MRI datasets [119]. In our study we trained the intensity model with the current LE-MRI data using a leave-one-out strategy for LE-MRI segmentation, and with the training set of the Sunnybrook data for the cine-MRI segmentation.

To initialize the shape model, the model’s mean shape is fitted to three anatomical landmarks: the aortic valve and mitral valve centroids, and the LV endocardial apex.
2.3. Results

Table 2.2: Average point-to-point distance errors and Dice similarity coefficients for the LE-MRI and cine-MRI dataset (final segmentation of our approach and the 3D-ASM segmentation).

<table>
<thead>
<tr>
<th>Initialization</th>
<th>Our Method</th>
<th>3D-ASM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Endo Epi DSC</td>
<td>Endo Epi DSC</td>
</tr>
<tr>
<td>Mean mm</td>
<td>2.34 2.75 0.68</td>
<td>1.83 2.38 0.81</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.80 0.75 0.08</td>
<td>0.50 0.53 0.05</td>
</tr>
<tr>
<td>Maximum value mm</td>
<td>4.28 5.15 0.84</td>
<td>2.58 3.41 0.86</td>
</tr>
<tr>
<td>Minimum value mm</td>
<td>1.34 1.58 0.52</td>
<td>1.09 1.59 0.62</td>
</tr>
</tbody>
</table>

LE-MRI

| Mean mm        | 3.15 2.78 0.88 | 2.76 2.58 0.92 | 2.79 2.56 n/a |
| Standard Deviation | 0.78 0.63 0.06 | 0.53 0.39 0.04 | 1.15 1.23 n/a |
| Maximum value mm | 4.34 4.05 0.92 | 4.03 3.82 0.94 | 6.29 6.59 n/a |
| Minimum value mm | 2.40 1.98 0.83 | 2.23 2.18 0.85 | 1.41 1.54 n/a |

cine-MRI

Endo=Endocardial contour; Epi=Epicardial contour; n/a = not available

The landmarks are extracted automatically using the same initialization step as for the GC-based method, to ensure a consistent starting point for both segmentation methods. We consider the centroid of the mitral valve as the center of the base of the ventricle, the aortic valve centroid is the control point $C_h$ closest to the real position of the aortic valve in our dataset. We determine as the third landmark, the endocardial apex, the center of the template in the last slice.

2.3.3 Evaluation Process

The quantitative results are summarized in Table 2.2 for our approach and for 3D-ASM segmentation, together with results obtained in the initialization step. The accuracy of the extracted contours has been evaluated by calculating the average point-to-point perpendicular distance and the Dice similarity coefficient (denoted as DSC) with each manual segmentation. The average perpendicular distance measures the distance from the automatically segmented contour to the corresponding manually drawn expert contour, averaged over all contour points. The DSC is a measure of overlap utilizing the contoured areas automatically segmented, manually segmented, and their intersection. DSC is always between 0 and 1, with higher DSC indicating better match between automatic and manual segmentations. The labeling of the 3D-ASM are extracted finding the intersection of the segmentation meshes with the slice plane.

2.3.4 LE-MRI Segmentation

The mean slice-wise point-to-point distance and the DSC were evaluated for the 20 subjects and are shown in the first part of Table 2.2. It can be observed that the initialization errors already outperform the obtained with the ASM tool, due to the restricted shape domain of the 3D-ASM, which is affected by the misalignments and the inhomogeneities of the images. The low error of the initial contours demonstrates that the robust and automatic initialization step is reliable to estimate the initial position of the ventricle and to adapt well to slice misalignments, common in MRI. Regarding the final segmentation, the average point-to-point distance is equal to 1.83 mm and 2.38 mm and the standard deviation is of 0.50 mm and 0.53 mm for
endocardium and epicardium contours, which shows the good performance of the technique, in particular in the presence of myocardial abnormalities.

Fig. 2.5 shows two examples for visual evaluation, including LE-MRI SA slices from the same patient together with the ground truth segmentation (two observers), the automatic initialization step, our automatic final segmentation and 3D-ASM segmentation. It can be seen the good initialization provided by our automatic LV location and the good match with the observers contours.

2.3.5 Cine-MRI Segmentation

We have demonstrated the accuracy of the technique for the challenging task of LE-MRI. In this section, we illustrate the flexibility as well as the consistency of the rule-based approach presented in this chapter with the segmentation of cine-MRI datasets. The numerical results are summarized in bottom rows of Table 2.2, the slice-wise point-to-point distance and the DSC were computed from all 15 subjects. The average DSC is equal to 0.92 and the standard deviation is of only 0.04, which demonstrates the high consistency of the technique. Furthermore, all DSC results are above 0.85, which demonstrates robustness to imaging or morphological inconsistencies that are common in such datasets, in particular in the presence of subject-specific abnormalities.

To evaluate the results visually, Fig. 2.6 shows slices from two patients of the Sunnybrook data with the automatically obtained myocardial contours and the ground truth provided. The automatic approach provides a good match with the observer delineation.

2.3.6 Parameter Influence

We fixed the two parameters of the graph cut, $\lambda$ and $\mu$, at the value of 0.25 and 0.8, giving more weight to the data term than to the inter-slice constraint. The errors reported in Table 2.2 are the ones obtained with that combination. However, we have run different experiments to study the influence of the weights to the final segmentation results in the LE-MRI database. With $\lambda \in (0.1, 1)$ and $\mu \in (0.1, 1)$ and a step of 0.05, the mean point-to-point error obtained for the LE-MRI database is $1.89 \pm 0.03$ mm for the endocardium, with a maximum error of 1.93 mm. For the epicardium, the mean error is $2.43 \pm 0.06$ mm, with a maximum error of 2.51 mm.

If we set the $\lambda$ to 0 the values obtained are $2.32 \pm 0.58$ mm (endo) and $2.69 \pm 0.68$ mm (epi). We have also run the experiment setting $\mu$ to 0, the obtained results are $2.11 \pm 0.70$ mm (endo) and $2.61 \pm 0.62$ mm (epi).

2.4 Discussion

The purpose of the presented work was to design a fully automatic approach to provide accurate and consistent LV myocardial segmentation in different sequences of MRI. Although automatic LV myocardial segmentation in MRI has been studied extensively, few works have focused on its full segmentation in LE-MRI and, to the best of our knowledge, none of them has been used to segment at the same time two sequences (cine-MRI and LE-MRI).

Regarding LE-MRI segmentation, unlike all related works ([25, 34, 126, 127]), which use cine-MRI segmentation to perform the LE-MRI segmentation, our approach uses
Figure 2.5: LE-MRI SA slices from two patients. From left to right: ground truth segmentation (two observers), the automatic initialization step, our automatic final segmentation, 3D-ASM segmentation. (cropped for better viewing)
only LE-MRI to directly segment the myocardium there. This characteristic helps speeding up the segmentation process and avoiding cumulative segmentation errors and variability. Furthermore, unlike Wei et al. [127] (where cine-MRI delineations were obtained using a semi-automatic method with manual corrections), our method works in a fully independent and automatic way. While the work of Dikici et al. [34] and Wei et al. [126] is done in a 2D environment without taking into account the longitudinal axis, thus not necessarily keeping a consistent 3D shape, we take into account inter-slice interactions. On the other hand, Ciofolo et al. [25] work deforming a 3D mesh, but misalignment correction appears to be omitted, whereas Wei et al. [127], who work within a 3D framework, use long axis images and their segmentations to first realign SA slices. Our method, while conceptually simple, makes use of both the longitudinal and within-slice information to deal with potential misalignments and avoid possible segmentation errors without the need to segment any long axis or short axis cine images. We have shown the importance of the inter-slice and the shape constraints by segmenting without one of these constraints (see Results). Quantitatively, a comparison across studies is difficult to perform, since datasets differ in number of acquisitions, image quality, image resolution, etc., and the gold standard and error measures are distinct or not defined. Nevertheless, our results are in the range of the reported techniques, and in some cases even better, while having the benefit of an automatic technique.

The obtained segmentation errors in cine-MRI are in the range of the MICCAI challenge [96, 101], which is quite challenging if we take into account that our technique has been designed to be as generic as possible, without taking advantage of cine-MRI specificities.

Our numerical validation results are shown on Table 2.2. The errors reported on
the first column of the table demonstrate that the initial location of the LV is properly and consistently found in both databases, with small error means and standard deviations. This indicates that reliable results can be obtained in practice without the need for user interaction. The rest of the validation results in Table 2.2 show that our approach clearly improves upon the results obtained with 3D-ASM. This is because the existing statistical-based technique is affected by the presence of abnormalities that are typically found in MRI sequences, especially for the segmentation of LE-MRI data. In the case of the cine-MRI data, the results are similar to those obtained with the 3D-ASM technique, however our approach provides greater consistency. Additionally, the proposed algorithm provides more flexibility and, at the same time, robust constraints that enable it to adapt well to subject-specific morphologies. Comparing both results, it can be observed in Table 2.2 that the average error of the cine-MRI database is higher than the LE-MRI, mainly because the Sunnybrook is a challenging database with many artifacts. However, the high mean DSC indicates good agreement with the expert segmentation.

Qualitative results obtained on the LE-MRI dataset are shown in Fig. 2.5(a). It can be seen that the inevitable presence of short axis misalignments affects the segmentation of 3D-ASM (see arrows), which means the LV myocardium can lie outside the allowable shape domain of the 3D-ASM. Our proposed algorithm, by using an inter-slice constraint and an alignment process, can explicitly handle such scan-related inconsistencies. Our method is robust to imaging artifacts particularly due to the decoupled approach that allows robust segmentation of the endocardium and effective subsequent constraining of the epicardial search. This is well illustrated in Fig. 2.5(b), where the boundaries are consistently well identified throughout the myocardium, including in error prone areas with scar tissue (see arrows).

Results from two different patients of the Sunnybrook database are shown in Fig. 2.6, giving an indication of the accuracy of the technique. The errors are small, and mostly concentrated in areas where different observers may give different segmentations. These results illustrate the generic nature of the technique, since no changes in the implementation were required for this type of images. In contrast, many statistical-based approaches call for additional training and algorithm tuning depending on the modality, application, and clinical population.

As seen both quantitatively and qualitatively, both databases are segmented with high accuracy, and more importantly, the exact same technique has been used. The validation shows that the proposed approach adapts well to different types of imaging sequences. This is a key property of the proposed approach: the same initialization stage, segmentation procedure, and geometric constraints are used for the different imaging sequences. Unlike with existing techniques, no sequence- or subject-specific tuning is required. The same tool is used consistently across individuals, groups of individuals and types of images, without the need for any user input. This leads to more consistent boundary delineations across patient examinations and clinical studies, which allows obtaining reproducible and statistically meaningful results. We have proven the robustness of the method by studying the sensitivity of the weights $\lambda$ and $\mu$. We have seen that these weights slightly modify the segmentation results, without a considerable effect. In addition to being fully automatic, the computational time of our segmentation is short: less than 2 minutes per volume in a Matlab (The Mathworks, inc., Nantucket MA., USA) implementation executed on an Intel(R) Core(TM)2 Quad CPU Q6600 @ 2.40GHz. All these features give to our method a
clinical value, particularly for multi-sequence viability assessment, which is the norm in practice.

2.5 Conclusion

We have presented a fully automatic approach that provides accurate and consistent left ventricle myocardial segmentation in MRI. The introduced initialization is reliable and it adapts well to appearance inconsistencies and slice misalignments. Our method also allows working with different imaging sequences without the need for prior registration or training, as shown with the results presented here using LE-MRI and cine-MRI. All these characteristics provide to the method a high clinical potential in terms of multi-sequence viability assessment.
An Algorithm for the Segmentation of Highly Abnormal Hearts using a Generic Statistical Shape Model

The content of this chapter is based on the publication:

ABSTRACT

Statistical shape models (SSMs) have been widely employed in cardiac image segmentation. However, in conditions that induce severe shape abnormality and remodeling, such as in the case of pulmonary hypertension (PH) or hypertrophic cardiomyopathy (HCM), a single SSM is rarely capable of capturing the anatomical variability in the extremes of the distribution. This work presents a new algorithm for the segmentation of severely abnormal hearts. The algorithm is highly flexible, as it does not require a priori knowledge of the involved pathology or any specific parameter tuning to be applied to the cardiac image under analysis. The fundamental idea is to approximate the gross effect of the abnormality with a virtual remodeling transformation between the patient-specific geometry and the average shape of the reference model (e.g., average normal morphology). To define this mapping, a set of landmark points are automatically identified during boundary point search, by estimating the reliability of the candidate points. With the obtained transformation, the feature points extracted from the patient image volume are then projected onto the space of the reference SSM, where the model is used to effectively constrain and guide the segmentation process. The extracted shape in the reference space is finally propagated back to the original image of the abnormal heart to obtain the final segmentation. Detailed validation with patients diagnosed with PH and HCM shows the robustness and flexibility of the technique for the segmentation of highly abnormal hearts of different pathologies.
3.1 Introduction

Cardiac MRI segmentation has received considerable attention in medical image analysis yet it remains challenging due to the geometrical complexity and high variability of the whole heart. Over the last decade alone a wide range of approaches have been investigated [96, 118], including image-based classification [29, 49, 53, 56, 73], pixel classification [79, 87, 92], deformable models [12, 14], cardiac atlases [10, 70, 71], and statistical models [48, 85, 89, 121, 129, 133]. Most of these techniques (about 70% of those listed in the recent review by Petitjean et al. [96]) have focused on the left ventricle (LV). More recently, the more challenging task of segmenting the right ventricle (RV) has received increased attention [23, 97], because of new findings that confirm its impact on several cardiac diseases [82], as well as to enable cardiac assessments that take into account the coupled biventricular motion of the heart.

In order to allow the translation of algorithms to the clinics, one major research challenge in cardiac segmentation is the development of approaches that are robust to different groups of individuals and classes of disorders. In the existing literature, however, most of the techniques have been mainly developed and validated with normal cases [65, 121], and in some exceptional cases with mildly abnormal subjects [12, 64, 112]. Furthermore, to the best of our knowledge, only few regional septal defects have been considered, such as hypertrophic cardiomyopathy (HCM) [6, 38, 70, 109]. These techniques are developed in a generic form for both normal and abnormal cases and do not have a mechanism to handle explicitly large remodeling effects owing to cardiac diseases.

In clinical practice, however, routine cardiac MRI quantification is concerned mostly with diseased subjects or subjects suspected to be diseased. The hearts under investigation vary from normal, to regionally abnormal, to severely remodeled cases. Some pathological hearts are particularly difficult to segment automatically, especially for the RV [23], as they undergo significant remodeling. This is the case in pulmonary hypertension (PH) [83, 117], which is a common disease of the pulmonary arteries. It is characterized by vascular proliferation, which results in a progressive increase in pulmonary vascular resistance and complex shape remodeling of both the left and right ventricles. As shown in Fig. 3.1, in a healthy heart, the RV generally shows as a crescent-like shape and is smaller than the more circular LV. In contrast, for a PH patient, the RV becomes very dilated, pushing onto the LV, which deforms and loses its roundness [32, 123]. To assess the severity of the PH and the patient response to treatment, accurate segmentation of the MR images is critical [24, 32, 37].

A potential approach for the segmentation of such pathological hearts is through the use of prior knowledge, such as in the form of statistical shape models (SSMs) or cardiac atlases [51]. SSMs have been used extensively to guide cardiac image segmentation based on active shape models (ASMs) [121, 129]. However, the existing models are generally derived from a relatively limited number of training datasets, which might not represent well enough the wide range of variability between classes of individuals and abnormalities. Alternatively, one can construct multiple statistical shape models for different cardiac pathologies. In practice, this is also unrealistic considering the challenges of each model construction, which requires the collection of a sufficiently large and representative training dataset, as well as the detailed delineation of the boundaries with point correspondence. The use of cardiac atlases combined with nonrigid image registration has also been widely investigated for cardiac image segmentation [10, 69, 70, 108, 109, 134, 135]. But the approach
Segmentation of Highly Abnormal Hearts using a Generic SSM

Normal cases

PH cases

HCM cases

Figure 3.1: Cardiac short axis MR images at end-diastole from normal (top), PH (middle) and HCM (bottom) subjects.

suffers from similar limitations when dealing with highly remodeled hearts, as existing cardiac atlases are generally constructed from populations of normal subjects or mildly pathological hearts. This means the required image registration can fail to recover accurately the large and complex deformations due to severe pathologies such as PH. Also, similarly to the SSM case, the construction of multiple atlases to account for the various cardiac pathologies assessed in the clinic is highly challenging in practice. These difficulties mean the applications of statistical models and atlases for the segmentation of highly abnormal hearts remain limited, or through the use of user interaction as in [125].

In this chapter, we present a novel semi-automatic method for accurate 3D segmentation of severely abnormal hearts using a unique statistical shape model constructed from a generic image population. More specifically, for any new case within a pathological shape class (e.g., PH and HCM), a geometrical transformation is estimated during the image search from automatically identified landmarks such that it approximates the gross anatomical departure and remodeling from the reference population (e.g., normal hearts). Based on this mapping between the two spaces, while the boundary search process takes place in the original image (corresponding to the abnormal anatomy), the actual model fitting and shape constraints are applied in the space of the reference statistical model. With this approach, any class of abnormal hearts can be segmented from a unique reference cardiac model without the need for
3.2. Methods

3.2.1 Overview

The aim of this work is to automatically segment severely abnormal hearts using a statistical shape modeling approach. More specifically, the proposed method is an extension of the well-known ASM technique [27], such that a unique reference point distribution model (PDM) can be re-used to segment different types of abnormal hearts without the need for model re-training. To illustrate the rationale of the proposed technique, we will perform automatic segmentation of both HCM and PH datasets, which display large yet distinct shape remodeling as illustrated in Fig. 3.1, by using a single PDM built from a seemingly normal population. The reference PDM is defined as

$$PDM_{ref} = (\bar{x}_{ref}, \Phi_{ref}, \Lambda_{ref}),$$

where $\bar{x}_{ref}$ is the reference mean shape and $\Phi_{ref}$ and $\Lambda_{ref}$ are the reference eigenvectors and eigenvalues, respectively.

An abnormal heart can be regarded as a previously healthy heart with a normal morphology that undergoes a series of local (i.e., deformation) and global (i.e., dilation) shape changes. Let us define $T_{remod}$ the geometrical transformation that corresponds to this virtual remodeling (from normal to abnormal morphology). In this chapter, by remodeling we refer to shape and size deviation/changes of the ventricles due to abnormality. The fundamental idea behind the proposed technique is that the remodeling transformation, if known, could be used to transform the image

Figure 3.2: Schematic workflow of our previous implementation of the method, using image registration and transformation of the original patient image data [4].
I_{A} to I_{N}

\[ I_{N} = T_{\text{remod}}^{-1}(I_{A}), \quad (3.2) \]

which can then be segmented through an ASM procedure by using as constraints the reference point distribution model \( \text{PDM}_{\text{ref}} \) as

\[ x_{N} = T_{\text{rigid}}(\bar{x}_{\text{ref}} + \Phi_{\text{ref}}b), \quad (3.3) \]

where \( T_{\text{rigid}} \) is the rigid transformation that aligns the new shape to the model, and \( b \) is the shape parameter vector describing the new deviation from the mean shape. Subsequently, the abnormal shape \( x_{A} \) can then be derived by projecting the obtained shape \( x_{N} \) back to the image space of the abnormal heart, i.e.,

\[ x_{A} = T_{\text{remod}}(x_{N}). \quad (3.4) \]

This was the idea behind our previous conference paper [4], which is schematically illustrated in Fig. 3.2. While the presented results at the Statistical Atlases and Computational Modeling of the Heart (STACOM) workshop [4] demonstrated a proof-of-concept, this approach has three drawbacks that we address in this chapter of the thesis, i.e.,

1. It required the manual definition of landmarks to calculate the remodeling transformation \( T_{\text{remod}} \) from the abnormal to the reference space.
2. The deformation of the abnormal image can produce distortions of the appearance, in particular due to the large space between the short axis images in MRI.
3. The transformation of the image volume renders the segmentation method computationally demanding.

In this chapter, we propose instead a new approach that addresses these issues using the following two main steps:

1. First, the remodeling transformation \( T_{\text{remod}} \) is estimated automatically. To this end, we estimate the reliability of each landmark during feature point search and we fit a thin plate spline model to a set of reliable landmarks with material correspondence in both the image and the reference model, as detailed in Sec. 3.2.3.
2. Second, instead of deforming the image, an algorithm is introduced (see Algorithm 3.1) that performs feature point search in the image of the abnormal heart, feature point transformation onto the reference model space, model fitting using the reference PDM, and finally shape propagation back onto the image space using the inverse transform of the estimated remodeling. As a result, no distortion of the image intensities is produced and the final segmentation result of the abnormal heart is obtained faster.

Note that the steps (1) and (2) are carried out within an iterative segmentation procedure, which means the estimation of \( T_{\text{remod}} \) is improved at each iteration until convergence, as schematically represented in Fig. 3.3. In the subsequent sections, we will describe in detail the main stages and the implementation details of the proposed algorithm for ASM segmentation of abnormal hearts.
3.2. Methods

Figure 3.3: Schematic workflow of the main steps involved in the proposed method for the segmentation of abnormal hearts using a reference PDM. The main stages of the proposed technique are (a) the search for boundary feature points in the patient image, (b) selection of the most reliable landmarks for the estimation of the virtual remodeling from the reference mean to the abnormal heart, (c) estimation of the transform using TPS and the corresponding landmarks in the reference mean shape, (d) transformation of the feature points onto the reference space, (e) application of the shape constraints using the reference PDM, and finally, (f) transformation of the obtained shape to its original abnormal space using the inverse remodeling mapping. The process is then (g) iterated until it converges.

3.2.2 Initialization

An ASM segmentation requires an initialization step, which is independent of the class of heart under investigation, in order to determine the initial translation, rotation and scaling of the model. To achieve this in the proposed technique, we define a number of key anatomical landmarks that include:

- Aortic valve centroid
- Centroid of the tricuspid valve
- Centroid of the mitral valve
- LV endocardial apex
- RV endocardial apex

These landmarks are needed (i) to suitably place the initial shape inside the image volume (translation), (ii) to scale the initial shape to the current image dataset (scaling) based on the LV apex-to-base distance along the axis of the heart; and (iii) to define the initial orientation of the heart based on the relative positions of the valves (rotation). We thus estimate these initial pose parameters by registering the
key anatomical landmarks listed above to their corresponding points on the mean shape $\bar{x}_{\text{ref}}$. We then obtain the initial shape $x_\text{A}^{(0)}$.

### 3.2.3 Feature Points Search

The second stage in ASM segmentation is the feature point search, i.e., finding for each landmark in the shape the most plausible boundary point amongst a set of candidates usually along a perpendicular search line [27].

An intensity appearance model $\text{IAM}_{\text{ref}}$ is required in this step for all landmarks, which is obtained at training. It aims to hold the intensity distribution typically found near the object’s boundaries. We do so by sampling a 1D profile of gray-level values with size 15 pixels for each landmark along the perpendiculars to the boundary. During training, from voxels sampled for each landmark $i$, a mean grey level profile $\bar{g}_i$ and its corresponding covariance matrix $S_{g_i}$ are estimated and stored. In our case, we have used a region-based approach for appearance modeling, i.e., all grey-level profiles of the landmarks of the same regions (using a 17-segment model) are combined, which leads to models that encode larger variability.

During the search for feature points at the segmentation stage, the algorithm first finds the intersections of the shape with the image planes. These intersections make a stack of 2D contours and each landmark in the contour will be positioned in a new location. We obtain this new location, or feature point, $y_i$ for each landmark by searching for a profile (sampled along the surface normals), which best resembles the one stored during training. In our case, the best feature point location is selected by minimizing the Mahalanobis distance between the sampled profile $g_i(y_i)$ and the mean profile $\bar{g}_i$ as:

$$
\arg\min_{y_i} \left( d_2^2(\bar{g}_i, g_i(y_i)) \right).
$$

The Mahalanobis distance is computed as follows:

$$
d_2^2(g_i, \bar{g}_i) = (g_i - \bar{g}_i)^T S_{g_i}^{-1}(g_i - \bar{g}_i),
$$

where $g_i$ represents the grey-level profile for landmark $i$.

The feature points $(y_1, ..., y_n)$ corresponding to each landmark of the model are then used in the next steps to calculate the remodeling transformation (Sec. 3.2.4), as well as to obtain a valid shape based on the reference shape model (Sec. 3.2.5).

### 3.2.4 Estimation of the Remodeling Transformation

In this section, we describe the method used to automatically estimate the remodeling transformation $T_{\text{remod}}$. To this end, we use a point-based registration approach between the reference (i.e., normal) and the abnormal hearts. For the reference shape, we simply use the landmarks of the mean shape $(\bar{x}_1, ..., \bar{x}_n)$ as the source shape in the registration task. Defining the destination set of landmarks in the registration, which correspond to the abnormal heart is more challenging as the target shape is unknown at this stage. Fortunately, we have at our disposal the feature points $(y_1, ..., y_n)$ as extracted at the previous section. Evidently, we cannot use all of these points in the registration as they inevitably include outliers amongst them, i.e., some of the feature points are located on incorrect boundary positions due to the presence of image artifacts and inhomogeneities. The inclusion of outliers in the point-based
registration procedure can lead to significant errors in the value of \( T_{\text{remod}} \) and in the subsequent cardiac segmentation. To address this issue, a potential solution is to detect and disregard the outliers amongst the \( n \) feature points, by using an outlier detection algorithm \([65]\), or by defining a reliability threshold at the training [115] step to distinguish between the reliable and unreliable feature points. Instead, in this work, a simpler and faster approach is used by assuming that the outliers constitute the minority in the feature points, i.e., we assume that the feature point search produces at least 50\% of correctly located points. By ranking the feature points \((y_1, ..., y_n)\) based on their reliability measures \( r = (r_1, ..., r_n) \), computed in this work using the defined Mahalanobis distance in Eq. 3.6, we then select the \( n/2 \) most reliable feature points to serve as the basis for the landmark-based registration. By doing so, we eliminate the most significant outliers, as well as some other correctly placed feature points. However, this does not affect the segmentation results as using half of the feature points is sufficient to derive a suitable approximation of \( T_{\text{remod}} \).

Finally, we need to choose a deformation model for the point-based registration between the selected reliable feature points and the corresponding landmarks in the mean shape. To this end, we use the well-known thin-plate spline (TPS) technique \([17]\) based on the selected \( n/2 \) landmarks, as follows:

\[
T_{\text{remod}} = \text{TPS}(\bar{x}, y, r). \tag{3.7}
\]

Correspondence away from the landmark points is defined by interpolating the transformation with the TPS deformation model. This produces a smooth transformation, which is guaranteed to be invertible when the selected feature points are free of self-foldings \([17]\).

### 3.2.5 Model Fitting

The ASM model fitting for the segmentation of the abnormal image \( I_A \) using the reference statistical shape model \( \text{PDM}_{\text{ref}} \) follows three main steps.

First, the remodeling transformation calculated at the previous section is used to project the feature points \( y \) onto the space of the reference PDM as follows:

\[
y_N = T_{\text{remod}}^{-1}(y_A). \tag{3.8}
\]

This is equivalent to deforming the abnormal image \( I_A \) to a reference image \( I_N \), followed by feature point search in \( I_N \), but in a much faster fashion and without grey-level intensity distortions.

Second, the ASM segmentation is continued in the reference space by applying model fitting to the transformed feature points \( T(y_A) \) based on the reference model \( \text{PDM}_{\text{ref}} \). This step is performed as in a standard ASM procedure to find the shape and pose parameters \((\bar{x}_{\text{ref}}, \Phi_{\text{ref}}, \Lambda_{\text{ref}}, b)\) that best fit the feature points \( y_N \) to the PDM. We thus obtain the following plausible shape in the reference model space:

\[
x_N = T_{\text{rigid}}(\bar{x}_{\text{ref}} + \Phi_{\text{ref}}b). \tag{3.9}
\]

Finally, the obtained shape \( x_N \) is propagated back to the image space of the abnormal heart to obtain the estimate of the target abnormal shape as:

\[
x_A = T_{\text{remod}}(x_N) = T_{\text{remod}}(T_{\text{rigid}}(\bar{x}_{\text{ref}} + \Phi_{\text{ref}}b)). \tag{3.10}
\]
Algorithm 3.1 Proposed segmentation of abnormal hearts

**INPUT:**
- Image of abnormal case $I_A$ (short axis MRI images)
- Reference point distribution model $PDM_{ref}$
- Reference mean shape: $\bar{x}_{ref}$
- Reference intensity appearance model $IAM_{ref}$

Initial shape $x_A^{(0)}$

$k=1$

Do

1. Search for feature points
   \[ y_A^{(k)} = \text{GreylevelAppearanceSearch} \left( I_A, x_A^{(k-1)}, IAM_{ref} \right) \]

2. Rank and select $n/2$ most reliable points
   \[ r_A^{(k)} = \text{ReliablePointsSelection} \left( y_A^{(k)}, I_A, IAM_{ref} \right) \]

3. Calculate remodeling transformation
   \[ T_{\text{remod}}^{(k)} = \text{ThinPlateSpline} \left( \bar{x}_{ref}, y_A^{(k)}, r_A^{(k)} \right) \]

4. Project feature points onto the reference space
   \[ y_N^{(k)} = \text{ApplyTransform} \left( (T_{\text{remod}}^{(k)})^{-1}, y_A^{(k)} \right) \]

5. Calculate shape parameters in reference space
   \[ x_N^{(k)} = \text{ShapeModelFitting} \left( y_N^{(k)}, PDM_{ref} \right) \]

6. Project result onto the abnormal image space
   \[ x_A^{(k)} = \text{ApplyTransform} \left( T_{\text{remod}}^{(k)}, x_N^{(k)} \right) \]

While \((\text{Distance} (x_A^{(k)}, x_A^{(k-1)}) > \text{Threshold}) \) or \((k < \text{IterationLimit})\)

**OUTPUT:** Final Shape $x_A$

The main stages of the proposed method, i.e. 1) feature point search in $I_A$, 2) $T_{\text{remod}}$ estimation, 3) $x_N$ constraining, and 4) abnormal shape estimation, are repeated until convergence of the segmentation result as illustrated in Fig. 3.3. This iterative process enables to improve, iteration after iteration, the feature point search, as well as the estimation of the virtual remodeling transformation. A pseudo-code representation of the proposed algorithm for abnormal image segmentation can be found in the listing of Algorithm 3.1.

### 3.3 Results

#### 3.3.1 Data Description

In order to assess the strength of the proposed technique, a total of 60 datasets were acquired from three different groups, i.e., 1) 20 cases from a reference population, 2) 20 PH cases, and 3) 20 HCM cases.

### Reference population

To build the reference statistical shape model, 20 MRI datasets from a general image population (with normal cardiac parameters such as ejection fraction) were acquired.
3.3. Results

using a 1.5T Philips Achieva System (Philips Healthcare, Best, The Netherlands). The short axis images were obtained using a cine SSFP sequence (TR/TE 2.9/1.5 ms, matrix 256 × 256). The slice thickness is 8 mm, a slice spacing of 10 mm and an in-plane pixel resolution of 1.42 × 1.42 mm.

PH Database

In order to test the robustness of the proposed technique with highly abnormal hearts, we use a dataset of 20 MR datasets corresponding to patients with PH with a significant shape remodeling of the ventricles. MR imaging was performed on a 1.5T whole body scanner GE HDx (General Electrics Healthcare, Milwaukee, USA) based on a balanced Steady State Free Precession (bSSFP) image sequence (TR/TE 3.7/1.6 ms, matrix 256 × 256). In this dataset, the slice thickness is 8 mm and a slice spacing of 10 mm with an in-plane resolution of 0.9375 × 0.9375 mm.

HCM Databases

Furthermore, to test the flexibility of the technique with respect to different types of abnormalities, in particular in the presence of regional shape changes and remodeling, we also consider in this validation an MRI dataset of 20 subjects diagnosed with hypertrophic cardiomyopathy (HCM). The MR image data were acquired using a 1.5T whole body scanner GE (General Electrics, Milwaukee, USA). The short axis image stack was obtained using a bSSFP sequence (TR/TE 3.7/1.6 ms, matrix 224 × 224), with a slice thickness of 8 mm, a slice spacing of 8 mm and a pixel resolution of 1.56 × 1.56 mm.

For all 60 images, both the left and right ventricles were manually segmented by using an interactive cardiac imaging software [63].

3.3.2 Experimental Setup

The goal of this study is to assess whether a new statistical shape model needs to be trained for each cardiac abnormality, which is a very tedious task, or whether we can re-use a pre-existing generic reference model for the segmentation of any highly abnormal heart based on the proposed algorithm that adapts to the characteristics of the new patient/disease. In the evaluation, we will quantify in particular whether the accuracy of the proposed technique is at least as good as the one obtained with a class-specific PDM created from pathological cases (e.g., HCM or PH), in which case the proposed technique would have the advantage of being more flexible and not requiring re-training for each new disease class. To achieve this, we segment each of the abnormal datasets using the following three methods:

1. ASMs with a PDM built from the 20 normal cases (reference PDM),
2. ASMs with a PDM built from the abnormal cases (PH or HCM) (leave-one-out basis),
3. the proposed method (combining feature-based transformation and the reference PDM built from the normal cases).

In the experiments, all the PDMs were constructed by using the standard modeling approach by Cootes et al. [27]. In order to establish point correspondence between all
training shapes before the construction of the models, the main stages of the shape extraction included the labeling of the MRI slices (LV endocardium, RV endocardium, epicardium) to obtain binary masks, the extraction of LV and RV surface meshes from the binary images using the Marching Cubes algorithm [68]. Subsequently, we used the surface currents projection method [36] to establish point correspondence across the extracted surfaces using a barycentric mapping as described in [94]. At the end, all cardiac shapes were represented using a set of 4000 corresponding landmarks. For all the approaches in Sec. 3.3.5 used for comparison, we used the ASM segmentation method in [121] and we calculated for accuracy evaluation the point-to-surface (P2S) errors between the manual and the automatic segmentations for both the LV and RV.

Table 3.1: P2S segmentation error statistics for the PH cases using the normal PDM, the PH-specific PDM, and the proposed technique.

<table>
<thead>
<tr>
<th>Percentage of datasets</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
<th>&lt;2.5 mm</th>
<th>&lt;3 mm</th>
<th>&lt;4 mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference PDM</td>
<td>3.10</td>
<td>0.69</td>
<td>2.18</td>
<td>4.21</td>
<td>25</td>
<td>40</td>
<td>85</td>
</tr>
<tr>
<td>PH-specific PDM</td>
<td>3.02</td>
<td>0.51</td>
<td>2.16</td>
<td>4.04</td>
<td>20</td>
<td>45</td>
<td>90</td>
</tr>
<tr>
<td>Proposed technique</td>
<td>2.60</td>
<td>0.34</td>
<td>1.89</td>
<td>3.02</td>
<td>40</td>
<td>95</td>
<td>100</td>
</tr>
</tbody>
</table>

SD=Standard Deviation

Figure 3.4: P2S errors (in mm) for the segmentation of the PH cases as obtained by using the reference PDM (light grey), the PH model (grey), and the proposed technique (dark grey).

3.3.3 Results for PH Patients

The segmentation errors for the 20 PH datasets are summarized in Table 3.1, which includes the average, standard deviation, maximum, and minimum errors. Furthermore, to have an indication of the relative distribution of the errors in the sample,
cumulative percentages were also added to the table for the three approaches. It can be seen that the direct use of both the reference and PH models induce average errors over 3 mm, with almost half of the datasets segmented with errors over 3 mm. In contrast, the proposed technique with feature-based transformation reduces the errors to an average of 2.60 mm and 95% of the cases are now segmented with P2S errors with less than 3 mm P2S error. The proposed technique outperforms the results of the standard ASM with a reference PDM by an average of 16.0%, with improvements reaching in some cases 29.8%. Additionally, there is also improvement over the results of the PH model by 13.1%, which is significant given the fact that the proposed technique has the additional advantage of preventing any tedious collection/delineation of abnormal cases and re-training of the PDM.

The individual results for the 20 cases detailed in Fig. 3.4 show that the proposed technique outperforms the ASM segmentation with the PH model in 90% of the cases, with minimal differences in the remaining two cases (#3 and #14). As expected, the reference PDM on its own generally lacks the flexibility required to segment the highly remodeled PH hearts. It is worth noting that there are cases, in which the reference PDM performs better than the re-trained PH model, illustrating the difficulty to obtain a representative statistical model in the presence of complex and severe pathologies. From Fig. 3.4 and Table 3.1, it can also be seen that the maximal segmentation errors are of 4.21 mm and 4.04 mm for the reference and PH models, while it is reduced to 3.02 mm for the proposed technique, which illustrates its consistency throughout the data sample. This demonstrates the strength of the estimated remodeling transformation, which together with the ASM search with a normal PDM, overcomes the inevitable lack of the representability in the training set of substantially abnormal morphologies.

Table 3.2: P2S segmentation error statistics for the HCM cases by using the three segmentation strategies

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
<th>&lt;2.5 mm</th>
<th>&lt;3 mm</th>
<th>&lt;4 mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference PDM</td>
<td>3.44</td>
<td>0.58</td>
<td>2.86</td>
<td>4.55</td>
<td>5</td>
<td>40</td>
<td>80</td>
</tr>
<tr>
<td>HCM-specific PDM</td>
<td>3.23</td>
<td>0.66</td>
<td>2.37</td>
<td>4.70</td>
<td>10</td>
<td>45</td>
<td>85</td>
</tr>
<tr>
<td>Proposed technique</td>
<td>2.57</td>
<td>0.46</td>
<td>1.99</td>
<td>3.61</td>
<td>50</td>
<td>80</td>
<td>100</td>
</tr>
</tbody>
</table>

SD=Standard Deviation

3.3.4 Results for HCM Patients

In order to show the flexibility and robustness of the technique to various types of abnormalities, we also evaluated our method with HCM cases, which induce more localized anomalies, generally in the septal region. We report the results of segmenting the 20 HCM cases described in Sec. 3.3.1, by using the reference PDM, the trained HCM model, and the reference PDM together with the proposed algorithm. The segmentation error statistics summarized in Table 3.2 show that the proposed method has the highest segmentation accuracy, with a mean error of 2.57 mm, a maximum of 3.61 mm, and a standard deviation below 0.5 mm. Furthermore, the direct use of the reference and PH models leads to maximal errors of 4.55 mm and 4.70 mm for the reference and HCM models, respectively.
The P2S segmentation errors for the 20 HCM datasets are plotted in Fig. 3.5. It can be seen that the proposed technique improves significantly and consistently upon the reference and the re-trained HCM models, except in very minor instances (e.g., Case #1) where the re-trained HCM model outperforms slightly the proposed technique. In 50% of the cases, the segmentation errors are below 2.5 mm by using the proposed technique, while this number decreases to only 5% and 10% for the reference and HCM models, respectively. All these results show the consistency of the technique not only across datasets, but also across pathology type (severely remodeled PH or regionally abnormal HCM cases).

3.3.5 Statistical Significance

To compare the performance of the proposed method with respect to the other two segmentation approaches (with normal and pathology-specific shape models) and to test whether the results are statistically significant, two-tailed paired t-test were performed and the obtained p-values for all pairwise comparisons are given in Table 3.3. The statistical significance is considered to be obtained for p-values < 0.01. The improvement by the proposed technique is statistically significant for both the PH and HCM cases. In contrast, the improvements achieved by using pathology-specific models are not statistically significant. The limited performance of the pathology-specific PDMs can be explained by the fact that even with model re-training, it is easy to find cases that remain extreme exemplars within the abnormal shape space distribution, thus leading to over-constraining and poor generalization during the segmentation of such highly abnormal hearts.

3.3.6 Effect of Mapping

In this section, we present a number of experiments carried out to assess in more detail the importance of the estimated remodeling transformation on the segmen-
3.3. Results

Table 3.3: Statistical significance (p-Values) of the differences between the results using the three segmentation strategies.

<table>
<thead>
<tr>
<th></th>
<th>PH</th>
<th>HCM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Our technique vs. reference PDM</td>
<td>0.007</td>
<td>0.005</td>
</tr>
<tr>
<td>Our technique vs. pathology-specific PDM</td>
<td>0.003</td>
<td>0.001</td>
</tr>
<tr>
<td>Reference PDM vs. pathology-specific PDM</td>
<td>0.773</td>
<td>0.667</td>
</tr>
</tbody>
</table>

Figure 3.6: Average segmentation errors (in mm) as a function of different proportions of feature points used for the estimation of the transformation (Sec. 3.2.4).

mentation of abnormal hearts. Firstly, we evaluated our hypothesis of using the 50% most reliable feature points for the estimation of the mapping. Fig. 3.6 shows the average segmentation error as a function of the proportion of feature points used to estimate the transformation. It can be seen that the average segmentation error indeed decreases progressively until about 50-60%, after which it increases again.

Secondly, we have quantified the extent of the changes due to the remodeling transformation by computing the point-to-surface average distances between the shapes before and after mapping. Fig. 3.7 shows a histogram of average shape changes due to the remodeling mapping, for normal (light grey) and abnormal (dark grey) cases. We can observe that the changes due to the mapping are relatively uniform, varying from 5 to 10 mm for the abnormal cases depending on the severity of the abnormality, indicating the consistency of the proposed technique.

In another experiment, we have calculated the Mahalanobis distances (MD) of each of the cases (normals, PH and HCM) with respect to the normal shape model (see Fig. 3.8). The cases, which are out of the ranges of normality based on the maximum value of the MD in the normal population, are plotted in red. It can be seen that before the application of the mapping in Fig. 3.8 (a), most of the PH and HCM shapes are out of the normality ranges. However, after the transformation of the shapes using the estimated mapping, all the MD values of the PH and HCM cases become
Figure 3.7: Histogram of average shape changes (in mm) due to the remodeling mapping, for normal (light grey) and abnormal (dark grey) cases.

Table 3.4: Statistical significance of the differences between the Mahalanobis distances with respect to the normal shape model of the normal population and the abnormal shapes.

<table>
<thead>
<tr>
<th></th>
<th>PH</th>
<th>HCM</th>
</tr>
</thead>
<tbody>
<tr>
<td>p-values before mapping</td>
<td>$1.2 \cdot 10^{-8}$</td>
<td>$1.4 \cdot 10^{-9}$</td>
</tr>
<tr>
<td>p-values after mapping</td>
<td>0.9649</td>
<td>0.223</td>
</tr>
</tbody>
</table>

Table 3.5: P2S segmentation error statistics for the PH and HCM cases using the initial mapping as the final segmentation compared to the final segmentation with PDM constraining in the reference space.

<table>
<thead>
<tr>
<th></th>
<th>PH</th>
<th>HCM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean mm</td>
<td>3.14</td>
<td>3.22</td>
</tr>
<tr>
<td>SD mm</td>
<td>0.59</td>
<td>0.48</td>
</tr>
<tr>
<td>Min mm</td>
<td>2.40</td>
<td>2.68</td>
</tr>
<tr>
<td>Max mm</td>
<td>4.46</td>
<td>4.10</td>
</tr>
<tr>
<td>Initial Mapping w/o ASM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final Segmentation w/ reference PDM constraint</td>
<td>2.60</td>
<td>2.57</td>
</tr>
<tr>
<td></td>
<td>0.34</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td>1.89</td>
<td>1.99</td>
</tr>
<tr>
<td></td>
<td>3.02</td>
<td>3.61</td>
</tr>
<tr>
<td>Improvement</td>
<td>16.6%</td>
<td>21.1%</td>
</tr>
<tr>
<td></td>
<td>6.1%</td>
<td>6.8%</td>
</tr>
<tr>
<td></td>
<td>6.5%</td>
<td>8.1%</td>
</tr>
<tr>
<td></td>
<td>28.8%</td>
<td>30.1%</td>
</tr>
</tbody>
</table>

$SD$=Standard Deviation

within the normality ranges as shown in Fig. 3.8 (b). This clearly indicates the role of the remodelling mapping in the proposed technique, allowing the subsequent use of the reference PDM to constraint the segmentation in a unique reference space for all cases and abnormalities.

Furthermore, we have added a statistical evaluation of the calculated MD values using a two tailed paired $t$-test in Table 3.4. The obtained $p$-values show how the differences between the normal population and the abnormal shapes are statistically
3.3. Results

Figure 3.8: Mahalanobis distances (MD) of each of the 60 cases (normal, PH, and HCM) with respect to the normal shape model: (a) without the application of the mapping and (b) after mapping. The red points indicate that the cases are out of the ranges of normality, while the dashed line show the maximum value of the MD based on the normal population. We can observe how most of the PH and HCM shapes are out the normality ranges before the application of the mapping (a) and how they are transformed into within the normal range in (b) by using the estimated mapping.

significant before the application of the mapping. However, this is clearly altered after the transformation of the unseen shapes as indicated by the new p-values in the table.

Finally, we have calculated the segmentation error obtained by using the initial mapping as the final result, with comparison to the actual segmentation results obtained with the proposed technique including shape constraining with the normal shape model. The results in Table 3.5 showed marked improvement in the results.
Figure 3.9: The main steps involved in the proposed algorithm for the segmentation of a PH heart. For visualization purposes the example is divided into LV endocardium, RV endocardium, and epicardium. The main stages of the proposed technique include (a) the search for the feature points in the patient image space, (b) the transformation of the feature points onto the reference PDM space, (c) the application of the shape constraints using the reference shape model, and finally, (d) the back-transformation of the obtained shape to the abnormal image space.

This indicates that the mapping plays an important role in the proposed algorithm but that it is insufficient on its own. It is its combination with model fitting in the reference space followed by back-propagation to the original abnormal image space that leads to the actual performance of the proposed technique for the segmentation of highly abnormal hearts such as PH and HCM.

3.3.7 Illustrative Results

The main steps of the proposed algorithm are illustrated in the examples shown in Fig. 3.9 for a PH dataset. For the LV endocardium (top row), it can be seen how the originally deformed LV in (a) recovers its normal roundness after the transformation onto the normal PDM space (see (b)-(c)), before the result is transformed back to its original abnormal morphology (d). Similarly, the abnormally deformed RV (middle row (a)) is transformed onto the reference space with a more usual crescent-like shape and an apparent decreased RV volume ((b)-(c)). The final RV result in (d) shows a dilated abnormal RV due to the PH pathology. Finally, the last row shows how the mapping onto the normal space leads to an LV/RV epicardium with the LV and
3.3. Results

(a) Manual (b) Reference (c) PH-specific (d) Proposed PDM PDM Technique

Case #2

Case #18

Figure 3.10: Three examples (case numbers refer to those of Fig. 3.4) showing myocardial segmentations on PH patients using (a) manual delineation, (b) the reference PDM, (c) the PH-specific model, and (d) the proposed technique.

RV in normal proportions (see (b)-(c)), while the final segmentation result obtained with the proposed technique (d) displays an RV that is abnormally much larger than the LV due to PH.

To illustrate the strength of the proposed approach, 2D visual examples for the segmentation of PH and HCM cases are displayed in Fig. 3.10 and Fig. 3.11, respectively. The manual segmentations are plotted together with the results obtained with the three ASM searches that we used in the experiments, i.e., reference PDM, abnormal PDM, and the proposed technique.

Regarding the PH examples in Fig. 3.10, it can be seen how both the reference and PH models produce errors at the RV and the septal wall, illustrating the difficulty for a single PDM to correctly capture the large RV variability in PH patients. The first row shows an example with a less pronounced abnormality, where the PH-specific model is not capable of localizing the RV variability of the patient. The second row corresponds to a severely abnormal heart (Case #6 in Fig. 3.4), with a LV that is deformed and abnormally smaller than the RV. The reference PDM naturally fails due to the change in relative ventricle size and, also, due to the lost of roundness of the LV, which is not valid according the reference statistical model. In this case, the PH-specific PDM presents a similar results than by using our segmentation approach. However, the estimation of the virtual remodeling between the normal PDM and the abnormal hearts eliminates the need for re-training and construction of new PDMs for every abnormality or class of individuals. Finally, the last row of Fig. 3.10 shows a mildly abnormal case. In this situation, the reference model outperforms the PH-specific model, as the latter one considers such morphology as implausible with regard to most PH cases in the training set. In contrast, the proposed technique maintains
Figure 3.11: Three examples showing myocardial segmentations on HCM patients using (a) manual delineation, (b) the reference PDM, (c) the HCM model, and (d) the proposed technique.

the same accuracy is this case as well.

With respect to the HCM 2D examples in Fig. 3.11, the main challenge is the abnormal variability of the myocardial thickness. As a result, in the first column, one can see how the reference PDM fails to accurately segment such cases that suffers from thickening in the septal region. The first and second examples correspond to a pronounced thickening of the myocardium, concretely in the basal anterior wall, which affects the application of the reference PDM, as it tends to generalize poorly to such localized changes in the width. In the last example (last row), the reference PDM rejects the loss of the crescent shape of the RV, which it considers as implausible. On other hand, the HCM-specific PDM enables to produce relatively accurate segmentations for the examples 2 and 3, but fails in the more challenging example 1. This illustrates the difficulty of pathology-specific PDMs to consistently generalize to extreme variants of the class. Furthermore, unlike our technique, this has the disadvantage of requiring a training of the PDM, which can be particularly tedious for morphologies with localized complex changes due to pathologies such as HCM.

Finally, to provide an understanding of the distribution of the segmentation errors, the average errors over all the datasets are color-encoded on the mean surfaces in Fig. 3.12 and 3.13. For the PH segmentations in Fig. 3.12, as expected, it can be seen how the errors of the reference PDM are localized mainly in the RV, both for the epicardium and endocardium, as well as in the septal region of the LV. As the RV becomes very dilated, its shape falls outside of the acceptable range of the reference PDM, resulting in large errors. For the segmentation of HCM cases, we can observe from the color maps in Fig. 3.13 that the errors are located at the RV endocardium and at the septal region, where the reference PDM naturally fails due to the change in
3.4 Discussion and Conclusion

In this chapter, we present a pragmatic approach to the segmentation of cardiac MRI of substantially diseased hearts based on statistical shape modeling and where the statistical model does only rely on a database of a generic patient population. The main benefits of the proposed approach are:

- it prevents tedious model re-training to accommodate different gross pathologies,
- it produces more accurate segmentations than those that would be obtained through re-training.

The central idea of the approach is the iterative estimation of the shape remodeling of largely abnormal hearts from MRI, which is then used to build a mapping between the image space of the abnormal heart and that of the reference PDM. The data-driven estimated transformation is specific not only to the pathology but also to the patient, which contributes to obtaining both an accurate and robust segmentation across complex anatomical configurations. As a result, the need for costly disease-specific PDM building is eliminated. At the same time, we preserve, and even improve,
the segmentation accuracy obtained with a disease-specific PDMs, which generally cannot fully cover the large anatomical variability related to severe pathology.

In terms of time complexity, our algorithm significantly improves upon our previous implementation [4] due to the new point-based mapping between image space and the reference model. The average running time for the full 3D segmentations is 1.46 min per subject, as opposed to 4.20 min in our previous conference paper [4]. All experiments were performed using MATLAB (The Mathworks, Inc., Nantucket, MA) with non-optimized code running on a 2.4GHz Windows PC with 8GB of RAM.

It is important to note that the estimated geometrical transformation represents a virtual remodeling between abnormal hearts and a generic (normal) PDM, and thus does not correspond necessarily to a real cardiac remodeling. This means the technique can also be applied to other types of cardiac abnormalities such as due to congenital diseases. Furthermore, the estimated transformation is likely to differ depending on the degree of abnormality across pathologies, and thus this could be used for classification purposes. However, this would require deriving suitable discriminatory metrics based on the transforms.

There are a number of limitations that are worth noting. First, we mostly focused on the adaptation of the PDM and did not address potential changes in the appearance between reference and abnormal scans. However, by using local intensity profiles in the feature point search, we found that the appearance-related boundary models are preserved across datasets. For example, the gray-level gradient between the blood pool and the myocardium is locally consistent across healthy and diseased datasets. Second, the selection of the most reliable features points was performed based on appearance criteria only, which can potentially lead to an uneven distribution of the landmarks used for the estimation of the remodeling transformation. One potential solution is to select the reliable feature points iteratively, such that the selection...
of a feature point takes into account not only its reliability measure but also its location with respect to the previously selected reliable points. Another limitation, which derives from the use of the ASM technique, is that there are no valves in our 3D model. The presence of the valve was shown to provide better localization and segmentation in the basal region of the heart. Therefore, one possible solution is the use of a biventricular shape model which includes valves [47]. Finally, our implementation works under the assumption that the TPS control points do not lead to self-folding, which was valid in our case as our search profiles do not intersect during feature point search. A potential improvement is to explicitly identify self-foldings during feature point search and to remove the corresponding landmarks from the subsequent mapping estimation onto the normal shape space.

In summary, the proposed approach shows promise for the segmentation of severely and/or regionally abnormal hearts, thus increasing the potential and applicability of the statistical shape modeling approach to a wider set of clinical cases.
Fully Automatic Image Parsing and Quality Control for Population Cardiac MRI

The content of this chapter is based on the submitted publication:

Xènia Albà, Karim Lekadir, Marco Pereañez, Pau Medrano-Gracia, Alistair A. Young, and Alejandro F. Frangi. Fully Automatic Image Parsing and Quality Control for Population Cardiac MRI. Submitted
Abstract

Continuous advances in imaging technologies enable ever more comprehensive phenotype on human anatomy and physiology. Concomitant reduction of imaging costs has thus resulted in widespread use of imaging in large clinical trials and population imaging studies. Magnetic Resonance Imaging (MRI), in particular, offers one-stop-shop multidimensional biomarkers of cardiovascular physiology and pathology. A wide range of analysis methods offer sophisticated cardiac image assessment for clinical and research studies. However, most of these methods have only been evaluated on relatively small databases and their translation and scalability to large clinical trials or population imaging cohorts is uncertain. In this chapter, we present a fully automatic parsing scheme that performs LV detection, segmentation, and overall quality control, and is applicable to large-scale cardiac MRI databases. Our method uses image features to predict an accurate initial position of the heart and key anatomical landmarks in an MRI volume using random forests regression. A generative model of positioning errors is used to synthesize an extensive training set to learn optimal regressor parameters. In processing a full imaging database, for each image volume, we predict the optimal corrective displacements from a standardized initial heart position and relative salient landmarks. Subsequently, we illustrate the pipeline with a previously published model-based cardiac segmentation. Finally, texture descriptors are used to control the quality of the results and detect failures to be corrected or removed from subsequent statistical analysis. The two steps introduced here, initialization and quality control, overcome tedious manual intervention or visual assessment and enable automatic image analysis. Detailed validation based on more than 1200 cases obtained from the Cardiac Atlas Project shows the promise of the approach.
4.1 Introduction

Continuous advances in imaging technology, which enable ever more comprehensive phenotype on human anatomy and physiology and concomitant reduction of imaging costs, have resulted in widespread use of imaging in large clinical trials and population imaging studies [86]. Moreover, there has been an emergence of large-scale population imaging databases [104], opening up challenges and opportunities for the understanding of disease phenotypes, and for the delivery of precision imaging [43]. In the area of cardiovascular imaging, for example, the exploitation of large-scale population image data is expected to impact the characterization of cardiovascular phenotypes like never before [114]. This, however, calls for the development of new techniques for cardiac image analysis that can handle the scale and variability associated with large imaging studies [84].

Among existing imaging techniques, cardiovascular Magnetic Resonance Imaging (MRI) has established itself as the one-stop-shop approach for non-invasive examination of cardiac morphology and function [99]. To enable high throughput analysis of imaging data automatically and reliably, a first step is the delineation of the myocardial boundaries followed by the estimation of various cardiac functional indices [93]. However, practical problems arise when dealing with large MRI studies. In particular, manual expert input becomes unfeasible and there is a need for fast and scalable methods to parse image data under large variability of anatomy, physiology and image quality.

In the existing literature, a wide range of approaches for automatic and semi-automatic cardiac MRI segmentation methods have been proposed (some reviews have been published [42, 93, 96, 118]). These methods include a wide range of techniques including image-based classification [29, 53, 56, 77], pixel classification [79, 87, 92], deformable models [12, 14, 28, 100], cardiac atlases [10, 70, 71], statistical models [48, 75, 89, 121, 129, 133], and learning-based approaches [8, 38, 81, 120].

Most published methods have been developed and validated typically using a few dozen image datasets and image databases that are not openly accessible. Consequently, published performance indexes are not directly comparable across studies and, more importantly, they have questionable translation and scalability as a measure of usability and performance in large clinical trials or population imaging cohorts. In particular, most existing techniques still rely on considerable manual intervention of some sort either during initialization or correction of cardiac image segmentation. The quality control of the segmentation process and results is also manual or visual, which becomes prohibitive when dealing with hundreds or thousands of image volumes. One of these problems is solved by segmentation challenges on publicly available image databases and evaluation protocols and measures [84]. Two relevant challenges for LV segmentation, for instance, are the 2009 challenge [101] organized during the MICCAI Conference; or the 2011 challenge [112] organized as part of the STACOM Workshop, also during MICCAI. These two challenges, however, were tested only on 30 and 100 cardiac MRI datasets for which manual annotations were provided as ground-truth. Given the small size of the datasets used for testing, whether the reported results can be generalized to larger cohorts remains highly questionable. A summary of the methods for semi-automatic and automatic LV segmentation in MRI using a large database (more than 100 datasets) of the last decade can be found in Table 4.1.
### Table 4.1: Summary of state of the art methods for automatic and semi-automatic LV segmentation with large databases.

<table>
<thead>
<tr>
<th>Reference (year)</th>
<th>Views</th>
<th>Datasets</th>
<th>Phases</th>
<th>Pathologies</th>
<th>Interaction Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lin et al. (2006) [67]</td>
<td>SA</td>
<td>Cine</td>
<td>330</td>
<td>25 CAD, PAD, MI</td>
<td>Combination of temporal Fourier analysis with contour detection</td>
</tr>
<tr>
<td>Zhu et al. (2007) [132]</td>
<td>SA</td>
<td>Cine</td>
<td>225</td>
<td>n/a</td>
<td>Deformable model with intensity and volume constraint of myocardium</td>
</tr>
<tr>
<td>Lu et al. (2011) [75]</td>
<td>SA</td>
<td>Cine</td>
<td>100</td>
<td>20 CAD</td>
<td>Joint LV-RV model combined both spatial and temporal context</td>
</tr>
<tr>
<td>Jolly et al. (2012) [54]</td>
<td>SA</td>
<td>Cine</td>
<td>100</td>
<td>20 CAD</td>
<td>Combined deformable registration method with gray level based shortest path segmentation algorithm</td>
</tr>
<tr>
<td>Margeta et al. (2012) [81]</td>
<td>SA</td>
<td>Cine</td>
<td>100</td>
<td>20 CAD</td>
<td>Supervised voxel-wise classification with layered spatio-temporal forests</td>
</tr>
<tr>
<td>Eslami et al. (2013) [38]</td>
<td>SA</td>
<td>Cine</td>
<td>104</td>
<td>18-25 N, DCM, HCM, MI, HF</td>
<td>Segmentation by retrieval with guided random walks</td>
</tr>
<tr>
<td>Tsadok et al. (2013) [120]</td>
<td>LA</td>
<td>Cine</td>
<td>126</td>
<td>1 CAD</td>
<td>Combination of a shortest path algorithm and a non-rigid registration</td>
</tr>
<tr>
<td>Lu et al. (2013) [76]</td>
<td>SA</td>
<td>Cine</td>
<td>133</td>
<td>20 N, HCM, MI, HF</td>
<td>Combination of optimal thresholding, fast Fourier transform, and multiple seeds region growing</td>
</tr>
</tbody>
</table>

n/a = not available; CAD = Coronary Artery Disease, PAD = Peripheral Artery Disease, MI = Myocardial Infarction, N = Normal, DCM = Dilated Cardiomyopathy, HCM = Hypertrophic Cardiomyopathy, HF = Heart Failure.
<table>
<thead>
<tr>
<th>Reference (year)</th>
<th>Views</th>
<th># Datasets</th>
<th># Phases</th>
<th>Landmark</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lu et al. (2009) [73]</td>
<td>LA Cine</td>
<td>116</td>
<td>1</td>
<td>Apex and MV</td>
<td>Joint context based approach under a learning-based object detection framework</td>
</tr>
<tr>
<td>Zheng et al. (2009) [131]</td>
<td>LA Cine</td>
<td>163</td>
<td>1</td>
<td>Apex and MV</td>
<td>Marginal Space Learning and Component-Based Voting</td>
</tr>
<tr>
<td>Lu et al. (2010) [72]</td>
<td>SA / LA Cine</td>
<td>188 (LA) 338 (SA)</td>
<td>1</td>
<td>RVI, MV, Apex and RVL</td>
<td>Context Modeling</td>
</tr>
<tr>
<td>Mahapatra et al. (2012) [80]</td>
<td>SA / LA Cine</td>
<td>80†</td>
<td>1</td>
<td>RVI, BACA and MV</td>
<td>Morphological operators and Random Forests with low level features</td>
</tr>
<tr>
<td>Lu et al. (2012) [74]</td>
<td>SA / LA Cine</td>
<td>100†</td>
<td>1</td>
<td>RVI, BACA and MV</td>
<td>Joint context based approach</td>
</tr>
</tbody>
</table>

MV = Mitral valve points ; RVI = RV insert points ; RVL= RV lateral point; BACA = Base-to-apex central axis points
† Data from CAP as part of the STACOM Challenge 2012 [41]
Existing tools often rely on manual user intervention for the initialization [38] of image segmentation or definition of key anatomical landmarks [120]. This approach becomes unfeasible when dealing with hundreds or thousands of MRI studies. Very few works deal with the automatic detection of key ventricular landmarks in large size databases [84]. A list of automatic landmark detection methods is provided in Table 4.2, with sizes varying between 80 and 338 datasets. Zheng et al. [131], who restrict their work to long axis (LA) images, evaluate their landmark detection method with 163 cases by using a marginal space learning approach. As an indication of the difficulty of the task, only two participants submitted their results in the Landmark Detection Challenge of STACOM 2012. The first technique by Mahapatra [80] required prior image segmentation, which is generally unavailable since it is precisely the purpose we seek to address. On the other hand, Lu et al. [74] proposed context modeling for LV landmark detection based on [73] and [72]. Both are supervised methods relying on large datasets with manually annotated ground truth for training [84].

A final limitation of existing approaches to analyze cardiac MRI data is their reliance on visual verification of the segmentation results. It is important to identify/remove grossly incorrect segmentations before they are used to derive anatomical or functional image biomarkers (e.g., ejection fraction, left ventricular volume, wall thickness, etc.). This also becomes impractical in large population imaging studies and needs to be automated. Ideally, one would want a method that automatically initializes the segmentation process and that subsequently scores the quality of the segmentation results.

In this chapter, we present a fully automatic parsing scheme that performs LV detection, segmentation, and overall quality control, and is applicable to large-scale cardiac MRI databases. Our method uses image features to predict an accurate initial position of the heart and key anatomical landmarks in an MRI volume using random forests regression. A generative model of positioning errors is used to synthesize an extensive training set to learn optimal regressor parameters. In processing a full imaging database, for each image volume, we predict the optimal corrective displacements from a standardized initial heart position and relative salient landmarks. Subsequently, we illustrate the pipeline with a previously published Active Shape Model (ASM) cardiac segmentation method. Finally, texture descriptors are used to control the quality of the results and detect failures so that they can be corrected or removed from subsequent statistical analysis. The two steps introduced, initialization and quality control, overcome tedious manual intervention or visual assessment and enable automatic image analysis.

In order to evaluate the proposed framework, we use MRI datasets obtained from the Cardiac Atlas Project (CAP) [41]. To the best of our knowledge, this is the first attempt to parse more than 1200 cardiac MRI datasets from both normal and abnormal cases using a fully automatic pipeline.

4.2 Method

4.2.1 Overview

Our automatic MRI parsing framework has three main components: regression-based organ detection, model-based organ segmentation, and classification-based quality
4.2. Method

Figure 4.1: Pipeline of the LV automatic segmentation. The first step is a key landmark detector, which is used to obtain an initial shape. A 3D-ASM segmentation step provides the final delineation of the ventricular boundaries. A quality control step indicates the confidence of the obtained results.
control. We illustrate them schematically in Fig. 4.1.

The goal of the proposed initialization is to estimate, without the need for any user interaction, the location of the left ventricle (LV) as well as key anatomical landmarks (basal and apical points). From these points it is possible to compute the initial position, height and size of the heart that feed the subsequent segmentation algorithm. The basic idea behind the proposed technique is to use a machine learning regressor to model the displacements from arbitrary incorrect positions to the correct localization of ventricular axes and anatomical landmarks. More specifically, we start with a rough estimation of the pose of the heart based on the positions of the short and long axis images. We then refine this position by predicting through regression the displacements of the ventricular axis at each short axis slice to a more suitable position. The same strategy is used for the detection of other landmarks such as the apex and mitral valve points.

To build the pose correction regression model, the training phase is illustrated in Fig. 4.3. The idea is to generate synthetic displacements by sampling arbitrary positions around the ground truth (i.e., ventricle’s center and the correct positions of the salient anatomical landmarks). Therefore, instead of learning the appearance around a particular landmark, we learn the features related to the corrections from arbitrary locations. Based on the characteristic appearance around an initial position (e.g., the middle of the image or a location that is known to fall within the blood pool), the regression model will propose a positional correction.

After the initialization, a segmentation method is applied to obtain the final delineation of the ventricular boundaries. The segmentation is performed based on a sparse ASM paradigm (SPASM) [121], which uses statistical shape models (SSMs) of the heart to encode prior statistical knowledge about cardiac shape and MRI intensity in a training set.

In view of using the proposed image parsing methodology in population imaging, visual inspection of the segmentation results is as unpractical as the segmentation itself. Hence, we wish to endow the segmentation method with a mechanism of self-verification. This is highly desirable and could open the way to self-correcting segmentation methods. The training set is built by creating synthetic displacements and deformations from the ground truth left ventricle segmentation so as to learn a mapping between the appearance of correct and failed segmentations.

All the steps are explained in the next section and shown in Fig. 4.1.

4.2.2 Automatic Key Landmark Detection

Landmark Set Definition

To enable automatic initialization, an accurate identification of key anatomical landmarks is an essential step in cardiac MRI segmentation, as these points will estimate the initial pose, scale and rotation of the model. Generally, as shown in Table 4.2, relevant landmark points include the mitral valve hinge points (defining the base line position), the apex and the central axis position, which are often used to initialize the segmentation process [98, 119, 120]. These points are usually defined manually, which is impractical for large population studies.

In cardiac MRI, the mitral valve (located between the left atrium and the LV) is shown in long axis (LA) view, because it contains both the atrium and LV. The
4.2. Method

line connecting the mitral valve hinge points is referred to as the base plane, which
defines the border of LV at base. The left ventricular apex is located at the bottom
of the LV, opposite the base of the heart. The central axis location, which is not
an actual anatomical feature, is defined at the centroid of the LV cavity on each SA
slice. The axis location will be exploited to define the LV position for the models.

In this work, all these key anatomical landmarks, as well as the LV axis points on
each SA slice are identified automatically as described in the following subsections.

Coarse LV Localization

As an initial step in the automatic initialization, we identify the location of the LV
through a rough estimation of its axis. We do this by estimating the intersection
between the long axis (LA) and short axis (SA) slices of each CMR study. The
main MRI planes used for cardiac function [59] typically include three LA views
(horizontal long axis − HLA, vertical long axis − VLA, and left ventricular outflow
tract − LVOT), and a complete SA stack covering the LV and RV. The series of SA
views are generated starting from the LV base at the level of the mitral valve, all the
way down to the apex. The VLA is orthogonal to the SA and passes approximately
through the apex and center of the mitral valve. The HLA is aligned orthogonally
to the VLA, also passing through the apex and center of the mitral valve. The
intersection between the HLA, the VLA and the SA images is exploited to obtain a
rough estimation of LV position. Fig. 4.2 illustrates the intersection of the two LA
images and the SA images, providing a point on each SA image and a set of points
on the LA images. However, the result is approximate due to scanning inaccuracies
and changes in breath-hold positions, which lead to mis-registration between slices.
The goal of the next step is to refine the pose estimation and to obtain a more
accurate estimation of the LV axis to increase the robustness of the subsequent LV
segmentation.

Key Landmarks Detection

To refine the location of the LV axis, as well as to identify the positions of the key
anatomical landmarks of the LV, we learn a regressor that refines the localization of
key LV landmark points using image features that characterize the local appearance
of these landmarks as detected in SA/LA views. The landmark regressors are trained
by creating a large synthetic training set of random displacements that emulate
localization errors from the ground-truth. The regressors learn the mapping between
these offsets and the known ground-truth positions using a Random Forests (RF).
The main steps in the landmark detection are illustrated in Fig. 4.1.

Training Set Generation

To train the landmark regressors, we first emulate displacements errors as random displacements from the ground-truth landmark positions that are available from expert annotations. Specifically, each landmark is perturbed randomly from its correct position denoted as the vector \( \mathbf{d}_i \). To each of these displacements, we associate image features \( \mathbf{f}_i \) sampled from their neighborhood of the simulated erroneous position. Let us denote all these pairs in the training sets as \( \mathbf{P}_i = (\mathbf{f}_i, \mathbf{d}_i) \). In our case, the displacements are generated as random vectors with a norm between 2 and 15 mm with uniform angular distribution. Subsequently, the goal is to learn a mapping function \( M \) from \( \mathbf{f}_i \), the image feature space (around
The intersection of the two LA and the SA images (a) provides rough positions for the LV, which is shown as a point on each SA image (b) and a set of points on the LA images (c).

In each training dataset, we sample a patch around the landmark (e.g., centroid of the LV) and we produce arbitrary displacements from the real position provided by the manual annotation of the expert. We then calculate image features from the patches and we learn a mapping function from the feature space to the displacement space.

Random Forests Regressor In this chapter, we use the well-known Random Forest (RF) [21] to compute a regression model for landmark detection. RFs are widely considered as one of the most robust classifiers and regressors. The algorithm is based on decision trees ensembles exploiting two mechanisms: building a tree ensemble via bagging with replacement, whereby any example selected from the training set can be used again; and random feature selection at each tree node, i.e., random selection of a small fraction of features and further splitting using the best feature from the current set. During training, the RF model calculates a response variable by creating many different regression trees (the forest of trees) and then processing each object to be modeled down each of the trees. Once the RF is trained, it is treated as a complete ensemble of base learning trees. Each leaf contains a distribution for the continuous output variable/s. Leaves have associated different degrees of confidence. To perform the regression, each learner produces a prediction individually and then predictions of all learners in the ensemble are combined to generate a prediction of the ensemble as a whole. The number of trees can be adapted to find the desired trade-off between accuracy and computational efficiency of the detection process. We constructed a RF with 50 trees with a maximum decision depth of 15, the values were selected based on experiments (see Results section).
Predictive Landmark Correction  Once the displacement regression model is trained, given a new slice, we sample a patch around the initial rough position $z_s$ obtained as described in Sec. 4.2.2 from the intersection of the LA and SA planes. From this patch, we calculate an image feature vector $f_s$. Subsequently, through the trained mapping $M$, we can calculate a predicted displacement $d'_s = M(f'_s)$, and then $z_s + d'_s$ becomes the prediction of the correct landmark position. We use a distinct displacement regression model for each landmark described in Sec. 4.2.2.

Feature Extraction  The last step of the automatic initialization defines the imaging features $f_i$. Using raw image intensities directly, as in deep learning, is an option. However, we have chosen to design more specialized image features that explicitly relate to the orientation and extent of deformation required to correct landmark positions. While deep learning overcomes the need for feature handcrafting, they do so at the expense of handcrafting the network architecture that could be equally arbitrary and less directly associated to the nature of the problem. We choose two complementary descriptors, i.e., the Histogram of Oriented Gradients (HOG) and Gabor Filters (GFs). This choice is motivated by the need for descriptors that can predict both the extent and orientation of the displacements.

- **Histogram of Oriented Gradients (HOG)** – The HOG descriptor [31] is a local statistic of the orientations of the image gradients, thus describing the local appearance. It is characterized by its invariance to rotation and illumination changes. Moreover, its computation technique is simple and fast. The main idea behind this descriptor is that local object appearance and shape can often be characterized rather well by the distribution of local intensity gradients or edge directions. The HOG feature divides the patch into many cells, with each cell location $(x, y)$, represented through its unsigned gradient orientation angle $\theta$ and gradient magnitude $\rho$. A histogram counts the occurrences of pixels orientations given by their gradients, i.e., each cell $(x, y)$ votes according to its magnitude into the bin corresponding to its gradient orientation. The final HOG descriptor is then built as a combination of these histograms. The local histograms are contrast-normalized by calculating a measure of the intensity across different regions and then using this value to normalize all cells.

- **Gabor Filters (GFs)** – The Gabor filter [45] is an established texture descriptor in image analysis. It is applied to extract features by analyzing the frequency domain of the image. A GF is basically a Gaussian function modulated by complex sinusoidal of frequency and orientation. It has the ability to filter both in the spatial and frequency domains and can be extended to any number of dimensions. These filters are more desirable since they provide for finer characterization of different textures. We apply two-dimensional (2D) Gabor filter since the data sets collected are sparse collections of 2D images. Two-dimensional Gabor filter banks decompose an image into feature maps related to different scales and orientations, thus capturing visual properties such as spatial localization, orientation selectivity, and spatial frequency. The 2D GF consists of a complex exponential centered at a given frequency and modulated by a Gaussian envelope. To obtain the Gabor-filtered image $I_g$ of a given image $I$ the 2D convolution operation is performed:

$$I_g(x, y) = G(x, y, \sigma_x, \sigma_y, \omega, \theta) \ast I(x, y),$$

(4.1)
where \( G \) is the Gabor function, and \( \sigma_x \) and \( \sigma_y \) are the standard deviations of the Gaussian envelope along the \( x \) and \( y \) axes. The parameters \( \omega \) and \( \theta \) are, respectively, the central frequency and the rotation of the Gabor filter. The values of these parameters \( \sigma_x, \sigma_y, \omega, \theta \) are chosen empirically based on tests. In our approach, the Gabor features are extracted by applying Gabor filters with three different scales and four orientations.

In summary, the feature variables for each patch are thus obtained as \( f_i = \{f_{\text{HOG}}, f_{\text{GF}}\} \), denoting HOG and GF, respectively.

### 4.2.3 Three-dimensional Model-based Cardiac Segmentation

The cardiac segmentation of the LV is performed with a modified 3D-SPASM segmentation method [121]. The essential components of a standard ASM [27] scheme are a point distribution model (PDM), an intensity appearance model (IAM), and a model matching algorithm. We use Principal Component Analysis (PCA) for the PDM, normalized intensity gradient, like in standard ASM, for the IAM, and an adapted fitting algorithm, SPASM, for 3D model to sparse image matching.

**Shape Model**

The point distribution model (PDM) is a statistical template of the organ of interest represented by corresponding anatomical points or nodes. In our case, the PDM is a surface mesh representing the LV, which includes endocardial and epicardial surfaces. The PDM is built during training by PCA to a set of aligned shapes [27] and maintaining eigenvectors corresponding to a predefined percentage of shape variability. The learned shape variability can be modeled as:

\[
\hat{x} = \bar{x} + \Phi b, \quad (4.2)
\]

where \( \mathbf{x} \) is a vector representing the shape, \( \bar{x} \) is the mean shape, \( \Phi \) is the eigenvector matrix and \( \mathbf{b} \) is a vector of scaling values for each principal component. By modifying \( \mathbf{b} \), different shapes can be defined.

**Intensity Model**

The intensity appearance model (IAM) learns the appearance around the boundaries of the target organ. The IAM is created at each heart mesh node by sampling the intensity in the neighborhood of each node over the image training set. In our case, the IAM models the intensity distribution that characterizes the myocardial boundaries. This is done by sampling a 1D profile of gray-level values for each node along the perpendiculars to the boundary. During training, from voxels sampled for each node \( i \), a mean grey level profile \( \bar{g}_i \) and its corresponding covariance matrix \( \mathbf{S}_{g_i} \) are estimated and stored. We have used a region-based approach for appearance modeling, i.e., all grey-level profiles of the nodes of the same regions (using a 17-segment model) are combined, which leads to models that encode larger variability.

**Model Initialization**

In order to determine the initial translation, rotation and scaling of the shape model, we rely on the anatomical landmarks computed in Sec. 4.2.2.
4.2. Method

These landmarks are used (i) to suitably place the initial shape inside the image volume (translation), (ii) to scale the initial shape to the current image dataset (scaling) based on the LV apex-to-base distance along the axis of the heart; and (iii) to define the initial orientation of the heart based on the relative position of the valve (rotation). We thus estimate these initial pose parameters by registering the obtained landmarks to their corresponding points on the mean shape \( \bar{x} \) and we obtain the initial shape \( x^0 \).

Matching Algorithm

The last element of the segmentation process is the matching algorithm, which has the role of deforming the mesh to match the image data \( I \). Our approach is based on the sparse fitting method SPASM [121]. The ASM segmentation searches for optimal boundary points, which has the best match between candidate boundary points and a model of intensity features subject to global anatomical statistical constraints [27].

Our algorithm, which is based on the SPASM technique [121], first finds the intersections of the shape with the imaging planes, which could be in any arbitrary orientation in 3D space. These model-plane intersections generate a stack of 2D contours. Each node in these contours is used to compute the closest boundary point in the image plane it belongs to. The candidate boundary point, or feature point, \( y_i \) is estimated by searching along a normal profile to the original model surface at the current contour point. Our algorithm uses an intensity model where each candidate point is obtained by selecting the minimal Mahalanobis distance between the sampled profiles \( g_i(y_i) \) and the mean profiles \( \bar{g}_i \) of the intensity model IAM as:

\[
\arg\min_{y_i} \left( d^2_i(\bar{g}_i, g_i(y_i)) \right). \tag{4.3}
\]

The Mahalanobis distance is computed as follows:

\[
d^2_i(\bar{g}_i, g_i) = (g_i - \bar{g}_i)^T S^{-1}_g (g_i - \bar{g}_i), \tag{4.4}
\]

where \( g_i \) represents the grey-level profile for node \( i \).

The obtained the feature points \( (y_1, ..., y_n) \) operate as deformation forces propagated to neighboring nodes with a weighting function, normalized by a Gaussian kernel as

\[
w(\lambda, \omega) = e^{-\frac{||\lambda - \omega||^2}{2\sigma^2}}, \tag{4.5}
\]

where \( (||\lambda - \omega||^2) \) is the geodesic distance between source and receiving nodes, and \( \sigma_p \) is the width of the normalizing Gaussian kernel.

Finally, the deformed mesh obtained after propagations is used to obtain a valid shape based on the reference shape model PDM. The parameter vector \( b \) controlling model deformation is calculated by computing an adjustment with respect to the previous iteration that best fit the current shape to the PDM. The allowed shape instances are limited by the statistical shape description from the training set.

The mesh is deformed for several iterations until the best-fit location is found.

4.2.4 Segmentation Quality Control

Incorrect segmentation results sometimes are inevitable when processing large population studies due to varying image quality, sub-optimal segmentation parameters or
other algorithmic failures. Therefore, when dealing with large amounts of data, one would desire to endow the image parsing pipeline with self-verification capabilities so as to automatically detect incorrect results either to reprocess them, or to disregard them altogether from subsequent data analysis.

Our quality control uses an anatomical-motivated definition of correct vs. incorrect LV delineations. More specifically, independently of its shape and size, the LV in our CMR studies is expected to have a set of geometrical and appearance characteristics as follows:

- The blood pool has bright MRI contrast.
- The myocardium has dark MRI contrast.
- The blood pool has a quasi-circular shape in SA.
- The myocardial wall thickness changes smoothly.

Similar to Sec. 4.2.2, we use a database of manual ground-truth segmentations alongside a set of incorrect segmentations generated under controlled conditions to develop a quality control system based on a statistical model of the appearance and geometry of the LV. We learn intensity features associated to these anatomical regions (blood pool and myocardium), thus distinguishing between incorrect and accurate segmentations using the ground-truth segmentations as a reference.

**Blood Pool and Myocardial Feature Description**

Based on the previous premises, we define some statistical features computed directly on image intensity values of the blood pool and myocardium. In the MRI sequences of the databases used in our experiments, the average intensity is expected to be higher in the blood pool and lower in the myocardial region. The dispersion of the intensity values shown by the standard and absolute deviations is related to texture regularity. Other statistical moment-based features we consider are skewness and kurtosis, which account for higher order information of the intensity probability distribution inside the LV regions. For the statistical descriptors, image intensity ranges were normalized to enable comparison across subjects using a max-min normalization strategy.

Moreover, in order to study the gray-scale spatial patterns, we compute gray level co-occurrence matrices (GLCM’s) [50]. GLCM’s characterize the texture of an image by calculating how often pairs of pixel with specific values and in a specified spatial relationship occur in an image, creating a GLCM, and then extracting statistical measures from this matrix. We generate several GLCM’s using four orientations with 1 pixel of distance between the pixel of interest and its neighbor, comprising the 4-connected pixels. Statistics on the summation of these four matrices are then used to provide information about the regularity of patterns occurring (as the energy or the entropy) and some information about the types of the patterns themselves (contrast and homogeneity). Inhomogeneous patches have low first order entropy, while a homogeneous one has a high entropy.

The final measures of patterns complexity used are based on the fractal dimension (FD) of the image, which measures a ratio of the change in detail to the change in scale. While the run length features could work in only one dimension at a time, the
4.2. Method

FD works in two dimensions. Using a differential box-counting approach [15], the FD at each pixel in a slice is computed, resulting in the FD image, and these are aggregated in the mean, standard deviation, and lacunarity (how densely the fractal fills the space).

We extract these texture features from the blood pool $f_{BP}^i$ and the myocardium $f_{M}^i$ regions separately. The feature vector is composed by both vectors, as $f_i = \{f_{BP}^i, f_{M}^i\}$. These features are expected to differ between correct and incorrect segmentations according to the previous assumptions, and thus they can be used to automatically and robustly evaluate the quality of the end segmentations in population samples.

Training Set Generation

In order to create a training set to learn to discriminate between correct and incorrect segmentation results, we proceed similarly to Sec. 4.2.2. We generate training samples by producing correct and incorrect segmentations as perturbations about the manual ground truth.

In general, there are two main types of possible failures in the LV segmentation: an incorrect positioning of the initial shape due to a failure in the initialization and large contour errors due to a failure in the segmentation. To generate training samples, we try to reproduce these two types of failures. On the one hand, we translate the contours to reproduce incorrect initializations and, on the other hand, we deform the contours to reproduce incorrect segmentations.

To generate exemplars of correct segmentations, we translate the contours arbitrarily using a random vector $v$ with uniform distribution and norm under 3 mm. To generate exemplars of incorrect segmentations, we do the same but with random perturbations with norms in the range between 6 and 10 mm. Moreover, incorrect segmentations also include deformed contours. To deform the segmentation contours and create incorrect segmentation contours, we fit a spline to the gold-standard contours and then we randomly perturb the control points introducing a displacement with a norm in the range between 3 and 6 mm. These values were chosen empirically based on observations in the experiments. Once the generation of the training samples is finalized, we estimate the feature descriptors for the original and generated exemplars.

Classification of Correct and Incorrect Segmentations

Similar to the automatic initialization, we use a RF classifier [21] to perform the detection of incorrect segmentations. Given the training labeled database, the RF classifier builds a set of classification trees, which predicts the class label (correct or incorrect) of a sample using a set of features. During training, each branch node of a tree learns features and threshold that results in the best split of the training samples into its child nodes. The splitting process continues recursively until the maximum tree depth is reached. At this time, a leaf node is created and the class distribution of the training samples reaching the leaf node is used to predict the class label of unseen samples. The final classification is based on the majority votes from individually developed tree classifiers in the forest. More details on RF can be found in [21]. The random forest for quality control were constructed with 100 trees, with each tree having a maximum depth of 15.
4.3 Data Description

4.3.1 Images

Validation is performed using more than 1200 MRI datasets obtained from the Cardiac Atlas Project (CAP) [41] (see Table 4.3). CAP is a web-accessible resource (www.cardiacatlas.org), which provides a resource for cardiac image data sharing and atlas-based shape analysis for population studies. The datasets used in this chapter are part of two cohorts: asymptomatic volunteers (AV) and patients with myocardial infarction (MI). Imaging protocols are described in [130].

**Population Set 1 - Asymptomatic Volunteers (AV)**

The AV group used fast gradient-recalled echo (GRE) imaging with 10-12 SA slices and 2-3 LA slices with typical parameters: 6 mm thickness, 4 mm gap, field of view 360-400 mm, $256 \times 160$ matrix, flip angle 20°, TE 3-5 ms, TR 8-10 ms, and pixel size from 1.4 to 2.5 mm/pixel depending on the size of the patient.

**Population Set 2 - Patients with Myocardial Infarction (MI)**

The MI group used retrospectively gated steady-state free precession (SSFP) cine imaging with 10-12 SA slices and 2-3 LA slices. The typical SA slice parameters were either a 6 mm slice thickness with 4 mm gap or 8 mm slice thickness with 2 mm gap and a field of view 360-400 mm. Image size was ranging from $138 \times 192$ to $512 \times 512$ pixels and pixel size from 0.7-2.5 mm/pixel depending on the patient.

For both groups, sufficient SA slices were expected to be acquired to cover the whole heart, as well as LA slices in the 4-chamber (HLA) and 2-chamber (VLA) views.

4.3.2 Ground Truth

To evaluate the segmentation results, manual contours were also provided by the CAP. Moreover, to evaluate the landmark detection algorithm, manual landmarks were manually marked for 500 patients by using the interactive medical image processing software GIMIAS v.1.6 (www.gimias.org) [63].

4.3.3 Data Preprocessing

Prior to any analysis, from the provided by the Cardiac Atlas Project, we have discarded incomplete datasets, which included:

- sets with incomplete or inconsistent DICOM tags,
- sets with inconsistent SA image geometry (different image dimensions of slices in a stack),
- sets with missing data (SA or LA series),
- sets with missing manual contours.

A summary of the data used in this chapter is shown in Table 4.3.
4.4 Results

<table>
<thead>
<tr>
<th>Table 4.3: Data available for evaluation.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Complete Datasets</strong> with Manual Contours</td>
</tr>
<tr>
<td>Asymptomatic (AV)</td>
</tr>
<tr>
<td>(855 test + 200 train)</td>
</tr>
<tr>
<td>Myocardial infarction (MI)</td>
</tr>
<tr>
<td>(all test)</td>
</tr>
</tbody>
</table>

*Complete datasets include SA slices covering the whole heart, 2-chamber and 4-chamber LA view slices with all required DICOM tags.

4.4 Results

The performance of the landmark detection framework has been evaluated on a subset of group AV, described in Table 4.3. The segmentation framework has been evaluated in two groups. To study the robustness of the presented method, we did the training exclusively on images from group AV and tested them on part of AV and, also, on a completely independent dataset, the group MI.

4.4.1 Landmark Detection Evaluation

To enable automatic initialization, we have detected four landmarks including the mitral valve hinge points (left and right), the apex and the central axis position, which are described in Sec. 4.2.2. There are 500 cases with manual landmarks from the AV group, as described in Sec. 4.3.2. In order to evaluate the landmark detection, we conducted Monte Carlo Cross-Validation (MCCV), also known as repeated random sub-sampling cross-validation, which consists of repeated rounds of validation conducted on a fixed dataset [107]. In each validation round, we randomly split the dataset into training (200 sets) and validation (300 sets) data. A total of ten validation rounds were performed and the automatic landmarks were compared with the manual ones using the Euclidean distance.

We tested the performance of the Random Forest (RF) regression as a function of the number of trees based on a fixed depth of 15. Our results are shown in Fig. 4.4, which indicate that stable error rates for the RF are achieved from 50 trees, which is the value fixed in the remaining experiments.

The statistics of the landmark detection error for each landmark are summarized in Table 4.4, which includes the median, average, and standard deviation (SD) values. In the case of the LV central axis, the LV centroid position is computed on each SA slice defining the axis, the result on Table 4.4 is the average distance error. It can be seen that the median errors vary between 2.19 mm for the ventricular axis to 6.69 mm for the more challenging mitral valve points. Furthermore, to have an indication of the relative distribution of the errors in the sample, cumulative percentages corresponding to the percentage of test images for which the error is less than a specific value are produced and shown in Fig 4.5. The percentage values show that more than 90 percent of the landmarks are identified within 1.5 cm error. Specifically, the central axis is detected with an error less than a 5 mm in more than 90% of the cases. In these experiments, we have not discarded the cases finally detected as ‘fails’ by the final quality control.
Figure 4.4: Performance results of increasing the number of trees in the forest for landmark detection.

Table 4.4: Point-to-point (P2P) distance error statistics in landmark detection (subset of group AV).

<table>
<thead>
<tr>
<th></th>
<th>Median (mm)</th>
<th>Mean (mm)</th>
<th>SD (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LV Central axis</td>
<td>2.19</td>
<td>2.47</td>
<td>1.48</td>
</tr>
<tr>
<td>Apex</td>
<td>4.56</td>
<td>5.60</td>
<td>4.13</td>
</tr>
<tr>
<td>Mitral Valve Right</td>
<td>5.19</td>
<td>6.57</td>
<td>4.90</td>
</tr>
<tr>
<td>Mitral Valve Left</td>
<td>6.69</td>
<td>7.76</td>
<td>5.08</td>
</tr>
</tbody>
</table>

SD=Standard Deviation

Table 4.5: Performance of segmentation quality control.

<table>
<thead>
<tr>
<th></th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group AV</td>
<td>0.96</td>
<td>0.96</td>
<td>0.95</td>
</tr>
<tr>
<td>Group MI</td>
<td>0.96</td>
<td>0.98</td>
<td>0.93</td>
</tr>
</tbody>
</table>

4.4.2 Quality Control Evaluation

The performance of the proposed quality control method is evaluated using both data groups, AV and MI. For the purpose of quantitative evaluation, the accuracy, sensitivity and specificity of the detection were calculated. The performance measurement metrics used is defined as follows:

\[
\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN}
\]

\[
\text{Sensitivity} = \frac{TP}{TP + FP}
\]

\[
\text{Specificity} = \frac{TN}{TN + FN}
\]
4.4. Results

Figure 4.5: Cumulative error distribution curves for each landmark detection.

Table 4.6: P2P error statistics for automatic image parsing for groups AV and MI.

<table>
<thead>
<tr>
<th></th>
<th>Before Automatic Quality Control</th>
<th>After Automatic Quality Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Initialization</td>
<td>Segmentation</td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>Mean</td>
</tr>
<tr>
<td>Endo</td>
<td>3.81</td>
<td>4.32</td>
</tr>
<tr>
<td>Epi</td>
<td>4.16</td>
<td>4.75</td>
</tr>
<tr>
<td>MI Endo</td>
<td>3.92</td>
<td>4.42</td>
</tr>
<tr>
<td>Epi</td>
<td>4.23</td>
<td>4.63</td>
</tr>
</tbody>
</table>

Endo=Endocardial contour; Epi=Epicardial contour; SD=Standard Deviation

Table 4.7: P2P error statistics for automatic image parsing in the subset of group AV (manual versus automatic initialization).

<table>
<thead>
<tr>
<th></th>
<th>Before Automatic Quality Control</th>
<th>After Automatic Quality Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Segmentation with manual init</td>
<td>Segmentation with auto init</td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>Mean</td>
</tr>
<tr>
<td>Endo</td>
<td>2.62</td>
<td>2.87</td>
</tr>
<tr>
<td>Epi</td>
<td>3.03</td>
<td>3.27</td>
</tr>
</tbody>
</table>

Endo=Endocardial contour; Epi=Epicardial contour; SD=Standard Deviation

where TP (true positive) represents the number of times the classifier guessed a correct positive value; TN (true negative) is the number of times the classifier correctly guessed a negative value; FP (false positive) is the number of times the classifier predicted a positive value, but the correct value was negative; and FN (false negative) is the inverse of FP. In this failure detection problem, TP has been defined as the case with less than a certain segmentation error and FN with a larger error. We tested the performance of the classifier as a function of the threshold defined as
a wrong segmentation. We have considered as failed segmentations those with an error over 4 mm with respect to the ground truth. The results are summarized in Table 4.5, indicating over 96% accuracy for the quality control. The obtained results are consistent for both groups.

4.4.3 Fully Automatic Segmentation Errors

The anatomical landmarks detected in the previous section are then used to initialize the 3D SPASM model, which provides a 3D final segmentation mesh. For evaluation, 2D contours are obtained for each slice by using the intersection between the mesh and the SA slices. In order to show the flexibility and robustness of the technique to different databases, we have evaluated our method using the AV and MI cases. We calculated for accuracy evaluation the point-to-point (P2P) errors between the manual and the automatic segmentations. We report the results of segmenting the cases described in Sec. 4.3, which are summarized in Table 4.3.

![Figure 4.6: Cumulative error distribution curves for myocardial segmentation error, before and after the quality control](image)

To assess the validity of our approach, the results of the segmentation using our algorithm are compared against the manual contours of the LV. Point-to-point segmentation error statistics are given in Table 4.6, including the median, average, and standard deviation errors for the initialization step and the final segmentation (with and without quality control of segmentation). Over the full test set, Table 4.6 shows that the proposed method has a final median error of 2.51 mm and 2.82 mm, respectively, for endocardial and epicardial segmentations for group AV and 2.41 mm and 2.47 mm for group MI. The mean and the median of the segmentation with the failures included are quite different, probably due to some failed initializations. Thanks to the automatic quality control, these cases are identified and discarded. We have also included in the table the errors for the initialization step, indicating a median error around 4 mm for both databases, subsequently refined by the segmentation step. Fig. 4.6 shows the effect of the quality control, most of the cases have less than 4 mm error after the failures have been discarded.
4.4. Results

Figure 4.7: Examples of landmark detection and segmentation results for AV datasets: landmark detection, (a) mitral valve points, apex and (b) the central axis and (c) the automatic segmentation contours together with manual contours. Our method is marked as red and the ground truth as green.

Furthermore, to specifically evaluate the relevance of the automatic initialization, Table 4.7 presents the point-to-point segmentation errors obtained by using (1) manual initialization (available for 300 cases) vs. (2) the proposed automatic initialization. The errors are relatively similar, and slightly better by using automatic initialization after segmentation quality control, showing the accuracy and relevance of our method. The mean and median of the results for manual initialization are even closer, which shows that the maximum errors have less influence than in the case of the automatic initialization, proving the importance of an automatic quality control in an image parsing approach.

4.4.4 Illustrative Results

Fig. 4.7 shows three examples of segmentation and landmark detection results for AV group. First and second columns show the landmarks detected by our method (red) together with the manually defined points (green). The final segmentation results are also shown, both the automatic (green) and the manual (red) delineations. The satisfactory quantitative analysis (Table 4.4 and 4.6) are confirmed with our visual analysis. It can be seen how the larger errors are produced in the mitral valve and
the central axis is detected with lower error. The segmentation of the endocardium is performed slightly better than the epicardium. This can be explained by the high variability of the surrounding tissues and the lack of contrast on some parts of the boundary, while the endocardium only interfaces with the myocardium. The errors on the endocardium are produced mainly around the papillary muscles.

Fig. 4.8 shows four examples of segmentation results for MI database. First thing to observe is the high variability of the images in appearance and shape between them and, also, compared to the AV examples. However, although we have trained our method with sets from group AV, similar results have been obtained.

Fig. 4.9 shows four examples of wrong segmentation results for both databases. These cases were detected as failures by the quality control and discarded from the final evaluation results. Note that the main problem is that the RV is considered as part of the LV.

4.5 Discussion and Conclusion

We described and validated a fully automatic approach to perform the analysis of cardiac MRI in large MRI studies, which cannot be processed through user interaction. The technique combines an RF-based landmark detector, a 3D-SPASM segmentation algorithm and an anatomical-driven classifier for quality control. Different to existing techniques, our method predicts the correction of the initial ventricular axis and anatomical landmarks by using a pose correction regression model. Subsequently,
the landmarks are used to initialize a three-dimensional LV segmentation based on ASM search. After segmentation, a quality control enables to score the quality of the segmentations and discard failures from subsequent data analytics.

The novelty of the technique lies in the use of a regression/classification approach based on large-scale synthetic examples of correct and incorrect LV characteristics. As such, it is easy to implement by using RFs and feature descriptors that are widely used in medical image analysis, and thus to be used in cardiovascular population studies. A complete validation based on two large-scale public databases, groups AV and MI, demonstrates the potential of our fully automatic segmentation for large-scale populations. To the best of our knowledge, this is the first attempt to parse more than 1200 MRI datasets from both normal and abnormal cases using a fully automatic pipeline.

The algorithm demonstrated good computational performance, the average running time to produce all the anatomical landmarks is 6.87 s per subject and the full segmentation procedure with the score takes on average 28 s per subject. All experiments were performed using MATLAB (The Mathworks, Inc., Nantucket, MA) with non-optimized code running on a 2.4 GHz Windows PC with 8 GB of RAM.

Future perspectives for this work include (i) applying it to other public large databases (e.g., the UK Biobank), which would further state the scalability and robustness of the method and (ii) use of the results to compute and analyze cardiac measures as EF or volume and extract new clinical knowledge of MRI-based phenotypes of cardiovascular health and disease.
General Conclusions

5.1 Overview

The motivation of this thesis was the development of tools to overcome some of the challenges associated with the processing of large amounts of cardiac image data. Concretely, we have presented algorithms for the segmentation of ventricular boundaries in large-scale MR studies. In the following paragraphs, we outline the main contributions of the dissertation.

Firstly, we have presented an automatic method to segment multiple sequences of MRI such as cine-MRI and LE-MRI without the need for any sequence-specific tuning (Chapter 2). A generic rule-based framework was used to automatically detect and segment the LV boundaries, which was accomplished using intensity information, as well as by including two constraints: shape and inter-slice smoothness. The segmentation algorithm used a decoupled graph cut approach with control points. The evaluation using two MRI different sequences (i.e. LE-MRI and cine-MRI) showed flexibility and robustness. Moreover, the characteristics of the method provided a high clinical potential in terms of cardiac viability assessment using multi-sequence cardiac MRI.

Secondly, in Chapter 3, we proposed a new algorithm for the segmentation of cardiac MRI with severely deformed pathological hearts. The algorithm did not require a priori knowledge of the involved pathology, or of any specific parameter tuning. The algorithm was validated with cases from two different pathologies, PH and HCM, showing the effectiveness of the technique for the segmentation of highly abnormal hearts of different groups. Moreover, a detailed validation was performed by comparing the proposed technique with the results of segmenting the cases with the reference model and, also, with a pathology-specific trained model. The proposed approach showed improvement for the segmentation of all cases including regionally abnormal hearts, thus increasing the potential and applicability of the statistical shape modeling approach to a wider set of clinical cases.

Finally, Chapter 4 presented a fully automatic framework, which parsed images to perform LV detection, LV segmentation and quality control without any user interaction for large-scale cardiac MRI samples. The method included a regression-based approach to predict the initial location of the heart and of its key anatomical landmarks using a synthetic training sample to learn and predict displacements to correct positions. Furthermore, anatomically-justified texture descriptors were used at the
General Conclusions

end of the segmentation to verify the quality of the end result for failure detection and removal of failed segmentations from the subsequent population analysis. Detailed validation based on two public large databases showed the promise of the approach for user-independent analysis of such large-scale populations.

To summarize, this thesis focused on the development of new tools to segment the ventricular boundaries in cardiac MRI for sequences with different appearances and for patients with varying pathologies. Furthermore, to facilitate the possibility of their usage in large-scale studies, we have also developed a fully automatic framework for segmentation as well as for failure detection without the need for any tedious user initialization or user quality control.

5.2 Outlook and Further Research Directions

The last few years have seen a significant increase in the amount of medical records clinicians routinely generate and collect, as well as in their ability to use technology to analyze and understand the data. However, unexplored knowledge remains currently frozen in the massive volumes of data that could impact health-care delivery. In this thesis, we presented tools that can contribute to extract such knowledge in the case of cardiac research and care, by enabling the processing of cardiac images at large scale.

A potential future direction for research based on the results of this thesis is patient stratification. By enabling the extraction of morphological and functional information of a large population, one could envisage approaches where new patient data would be compared to those of a baseline population for decision making. This, of course, will need to be done not only based on image-driven information but also in combination with other clinical, biochemical, environmental, and biomolecular information. This will allow clinicians to quantify risks and to predict outcomes of the treatments through the analysis of past data from other patients with the same condition and profile. Moreover, they will be able to stratify care so as to improve efficiency and reduce costs.

The segmentation of the anatomy of the heart, using the techniques developed in this thesis, is a crucial first step to extract the meaningful indices to be able to predict and classify patient evolution. However, to fully exploit the benefits of the described techniques, they need to be placed in the context of full data analysis pipelines, which includes several other tasks such as data acquisition, data organization, information extraction, data integration, finally followed by interpretation and presentation of the results [52]. Thus far, these tasks have been investigated in isolation and significant work will need to be carried out to integrate the techniques within user-friendly and clinically useful pipelines. In this section, we aim to enumerate some of the challenges and adaptations that require further attention by researchers in the field of population studies.

Data harmonization. Frequently, the datasets collected by researchers and clinicians, in particular when carried out at different sites, are recorded in varying formats and protocols. This means that they cannot be processed directly by the computational techniques, such as the ones developed in this thesis, without previous re-organization. However, achieving this consistently and completely is a technical challenge. For example, in the last chapter of this thesis, we semi-automatically cleaned and re-organized the data from two databases to enable the image input to
be read by the segmentation pipeline. Even by working with two public databases that have been widely used by researchers, several challenges had to be overcome. As a consequence, for effective large-scale analysis of the data, the cleaning and organization is an important point that will need to be investigated in a more generic and automated approach to enable to exploit the computational techniques that are being developed for data analysis of imaging studies.

**Missing data.** Closely related to the previous point is the challenge of the incompleteness of data in large-scale studies and, specifically, in medical imaging. Incomplete data, which refers to the missing of data field values for some samples, creates uncertainties during data analysis, which must be suitably managed during data analysis. These missing values are the consequence of different practical limitations, such as errors in the recording or data acquisition, as well as some systematic policies to intentionally skip some values (such as to reduce scanning time). While most algorithms have inbuilt solutions to handle missing values (such as ignoring data fields with missing values), data imputation for large-scale studies in the specific case of image data will need to be investigated to produce improved models and quantifications.

**Metadata generation.** Another important challenge is to automatically generate the right metadata to describe information about the way data is recorded. Metadata is a powerful tool to annotate and exploit image-related information for clinical and research purposes and to organize and retrieve images and associated data into archives. Metadata organizing systems can minimize the human burden and will enable to standardize the formats to facilitate subsequent data analysis.

**Multi-source data integration.** In this thesis, we developed techniques to extract organ-level anatomical information from large-scale image studies. However, such information describes one angle of health and disease. Other important information, such as environment, lifestyle, biochemistry, and genetics provide important complementary information. As population studies become established, an important research direction will consist of integrating these different sources of data and scales of biology and physiology to provide a systemic description of health and disease.

**Knowledge extraction and interpretation.** Once the image data is suitably accessed by the algorithms (with the assistance of data harmonization algorithms and the use of standard metadata formats), processed to extract anatomical information, integrated with other sources of data, the next natural challenge for researchers will be the interpretation of the data for clinical purposes. The analysis of patient data towards knowledge extraction will deliver better health-care with regard to the disease diagnosis and prognosis, as well as for treatment stratification. More efforts have to be done on information processing and knowledge extraction and on understanding relationships between data and the patterns behind information. The transformation from data to information seeks answers to the questions of ‘who?’, ‘what?’, ‘when?’ and ‘where?’, and hence delivers useful, organized and structured information [43]. In the case of cardiac research, tools are needed to extract more advanced and more complex phenotypes of cardiac health from large imaging studies [114]. For example, many works have been published proving the importance of left ventricular ejection fraction in predicting prognosis [30] or on the relationship between left ventricular wall thickness with a higher mortality [33, 105]. The study of new cardiac function indicators and multi-source biomarkers thanks to big data analysis will enable early identification of patients at risk of cardiovascular events,
which remain a great source of mortality and morbidity in the developed countries.

In summary, big data analytics is an emerging field that is creating new opportunities, as well as challenges. In the future, we will see the rapid, widespread implementation and use of big data analytics across the health-care organization and the health-care industry. To that end, the several challenges highlighted above must be addressed in a holistic manner. As big data analytics becomes more widespread, establishing standards and continually improving the tools and technologies will garner attention.

In this context, proper analysis of big amounts of medical image data is an important step toward the ability of processing and extracting the information that can be derived from this data. Thus, the segmentation and analysis tools that have been presented in this thesis are an important step toward the capability of achieving big data analytics of all medical related information. Big data analytics and applications in health-care are at a nascent stage of development, but new advances in tools can accelerate their maturing process.
References


References


References


[63] Larrabide, I., Omedas, P., Martelli, Y., Planes, X., Nieber, M., Moya, J., Butakoff, C., Sebastian, R., Camara, O., De Craene, M., Bijnen, B., and Frangi, A. F. GIMIAS:


References


[91] Pednekar, A., Kurkure, U., Muthupillai, R., Flamm, S., and Kakadiaris, I. Automated left ventricular segmentation in cardiac MRI. IEEE Transactions on Biomedical En-
References


[108] Shi, W., Lambeart, H., Bai, W., Ledig, C., Zhuang, X., Marvao, A., Dawes, T., ORe-


List of Publications

Peer reviewed papers in international journals


Peer reviewed papers in conferences proceedings


The completion of my Ph.D. has been a long journey, which would not have been possible without a lot of people who have been by my side all these years...

Firstly, I would like to express my sincere gratitude to my advisor Prof. Alejandro F. Frangi for providing me with the unique opportunity to work in his research group and for his help and support during my Ph.D. studies. I also want to thank Dr. Rosa M. Figueras, who was the first to encourage me to pursue a Ph.D. and helped me at the hard beginnings. And, finally, I would like to thank Dr. Karim Lekadir, his day-to-day guidance and constant encouragement was vital throughout all the stages of my dissertation. He spent endless hours reviewing my research papers and giving me excellent suggestions, which always resulted in improved versions of the documents.

Without any doubt, the environment of a Ph.D. student has enormous impact on the quality of the time spent, and surely on the quality of the thesis itself. I would like to thank my colleagues in the UPF, the “old” and the “new” ones, Corné, Marco, Hernán, Hrvoje, Arjan, Patri, Carlos, and many others... for being there and sharing our adventures together, you have made my life easier. A special mention to the girls, Cata, Martha and Issis, you brightened up my days even when you were far away, gracias bonitas. And also, to Veronika, Antonio and Nerea, I feel really fortunate to have you, you have been my refuge so many times and I know this could not be done without you and those coffees.

I would like to thank my friends in Terrassa who, even without understanding why I was doing a Ph.D. or what I was doing, supported and encouraged me always, especially this last year. No us anomenaré a tots perquè ompliria aquesta pàgina, vosaltres sabeu qui sou... Gràcies pel vostre suport, per alegrar-vos per mi en els bons moments i per donar-me ànims en els no tan bons. Moltissimes gràcies! I en especial al Raül, gràcies per demostrar-me com n’estaves d’orgullós de mi i de la meva feina.

I also want to mention my friends Lorena, Bea, Ester and Anna, les meves “majes”, moltes gràcies pels vostres ànims, sobretot aquest últim any. No sé si us ho dit mai, però aproveito ara, tenir-vos al costat aquells anys d’estudis va ser el millor de la universitat i el millor que em vaig emportar.

I would like to mention my other life, my political life, and the ones with whom I have worked with over these years. It could seem naïve but having this other life has saved my Ph.D. many times.

Finally, thanks to my whole family, they are an essential part of my life. Tinc una família única, Santi, Marc, Meritxell, David, Júlia, Gerard i Georgina, estic orgullosa i feliç de tenir-vos a la meva vida. Specially to my father, who has demonstrated he is so proud of me, gràcies per les teves paraules i pel teu suport incondicional. To my mother, this could not have been possible without her... et diria moltes coses i me’n deixaria moltes més, així que ho resumiré amb una frase: ets la millor mare que podria tenir. Estic convençuda que no hauria arribat fins aquí sense vosaltres al meu costat. Us estimo moltíssim.
Biography

Xènia Albà was born on July 5th, 1981 in Terrassa (Catalunya), where she grew up, a nice city 30 km of Barcelona. She received a BSc in Telecommunication Technical Engineering from Universitat Politècnica de Catalunya (UPC) in 2005. During her studies, she did a 6-month stay at the Fontys University of Applied Sciences in Eindhoven, The Netherlands. In 2009 and 2010, respectively, she obtained a postgraduate qualification in Telecommunication Engineering and a Master degree in Computer Vision and Artificial Intelligence from the Universitat Autònoma de Barcelona (UAB). In the meanwhile, she worked as part-time assistant professor at the UPC teaching in pre-graduate official studies.

After her graduation, she joined the Centre for Computational Imaging and Simulation Technologies in Biomedicine (CISTIB) at the Universitat Pompeu Fabra (UPF) as a researcher in training and, in 2011, she enrolled the PhD Programme of the Departament de Tecnologies de la Informació i les Comunicacions in the UPF. During her time at CISTIB, she taught Calculus to first-year students at UPF’s Escola Superior Politècnica. She obtained a FPU fellowship from the Spanish Ministry in 2012. During her PhD, her main research line has been cardiac image analysis.