Thesis Submitted for the Degree of
Doctor of Philosophy

Integrating Contour-Coupling with Spatio-Temporal Models in Multi-Dimensional Cardiac Image Segmentation

Author: Stephen O'Brien B.Eng.

Supervisor: Professor Paul F. Whelan

Dublin City University
School of Electronic Engineering
January 2011

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Doctor of Philosophy is entirely my own work, that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed:_____ ID No.:__53054756__

Candidate

Date: __17^{th} January 2011__

Acknowledgements

I wish to thank Prof. Paul F. Whelan for his continuous support and guidance throughout my three years of research. My gratitude is also extended to Dr. Ovidiu Ghita for many insightful discussions during the development of my research, which helped me mature as a researcher and as an individual.

Also, to the rest of my close colleagues; Aubrey Dunne, Sean Begley, Brendan Byrne and Tony Marrero for their friendship and shared experiences along the research path. Furthermore, thanks to the Irish Research Council for Science, Engineering and Technology, for generously funding my Ph.D. work.

Personal thanks, and deep gratitude go to Pat, Maria, David and Ashleigh for supporting my decision to continue my education and helping me through many difficult stages along the way. This thesis would not exist without your presence in my life.

Contents

Integrating Contour-Coupling with Spatio-Temporal Models in Multi-Dimensional Cardiac Image Segmentation

Stephen O'Brien

Abstract

Cardiovascular disease (CVD) is a major cause of death in the Western world. As such, timely and reliable diagnosis of CVD is a primary requirement in the clinical setting. Manual analysis of multi-dimensional cardiac images (usually regarding the left ventricle) is time-consuming and prone to inter/intra observer variability. Automatic segmentation algorithms are a promising solution to alleviate this issue.

Within the model-based segmentation domain, a popular strategy considers the entire segmentation target as a single entity. Although intuitive, this modus operandi suffers from significant practical limitations. One notable example is the requirement of significant training data, due to the difficulty of modelling 3D or 3D+time structures, that exhibit complex spatial and temporal deformation.

This thesis investigates an alternate modelling strategy that is adaptable to changes in structure, and scalable with respect to dimensionality. The major contributions presented in this thesis result from investigation into 3D+time cardiac left ventricle segmentation using the proposed approach. The first contribution explores whether all components of a segmentation target need to be explicitly and simultaneously modelled (contour coupling). The second investigates whether complex biological structures can be dimensionally subdivided for modelling and later unified for segmentation (spatio-temporal modelling). The final major contribution determines whether all training data, specifically in a multi-dimensional scenario, is categorically required to construct practical models for accurate segmentation (segmentation framework).

Comprehensive evaluation of the proposed method demonstrates that modelling only the crucial components of the segmentation target, while enforcing non-rigid *a priori* constraints at segmentation-time, allows the proposed method to adapt to configurations *outside* the training set. It is also illustrated that modelling dimensional variation separately alleviates excessive training requirements and aligning difficulties when compared to the standard unified-modelling approach.

In conclusion, this thesis presents a compelling argument for critically evaluating the physical and dimensional structure of the segmentation target to determine the best-suited modelling strategy. With respect to 3D+time cardiac left ventricle segmentation, the logic of sub-dividing the modelling task into component parts is soundly supported by theoretical and experimental evidence. Finally, a comprehensive segmentation framework is presented to accurately model, and segment, the complex spatial and temporal dynamics of the cardiac structure.

Publications

Journal Publications

Stephen P. O'Brien, Ovidiu Ghita, Paul F. Whelan (2010), **A Novel Model-Based 3D+time Left Ventricular Segmentation Technique**, IEEE Transactions on Medical Imaging, (In Press).

Conference Publications

Stephen O'Brien, Ovidiu Ghita, Paul F. Whelan (2009), **Segmenting the Left Ventricle in 3D Using a Coupled ASM and a Learned Non-Rigid Spatial Model**, MICCAI 3D Segmentation Challenge for Clinical Applications Workshop, September 2009, London, UK.

Stephen O'Brien, Ovidiu Ghita, Paul F. Whelan (2009), **Morphological-based 3D Segmentation Method of Cardiac MRI Data**, Swiss Society for Biomedical Engineering (SSBE), August 2009, Bern, Switzerland.

Glossary of Acronyms

Acronym	–	**Explanation**
CVD	–	Cardiovascular disease
MRI	–	Magnetic Resonance Imaging
NMR	–	Nuclear Magnetic Resonance
CMR	–	Cardiac MRI
PET	–	Positron Emission Tomography
SPECT	–	Single Photon Emission Computed Tomography
CT	–	Computed Tomography
CTA	–	Computed Tomography Angiography
CAD	–	Coronary Artery Disease
LV	–	Left Ventricle
CoD	–	Curse of Dimensionality
ASM	–	Active Shape Model
AAM	–	Active Appearance Model
2D	–	Two Dimensions
3D	–	Three Dimensions
SVD	–	Singular Value Decomposition
PCA	–	Principal Component Analysis
ICA	–	Independent Component Analysis
kPCA	–	Kernel Principal Component Analysis

List of Figures

List of Tables

Chapter 1

Introduction

Cardiovascular disease (CVD) is a major cause of death in the Western world. It is responsible for nearly half of all deaths in Europe [4], and approximately a third of all deaths in the USA [57]. With sedentary lifestyles and child obesity both rising trends, CVD risk factors will remain a central health-issue for many years to come. In fact, it is estimated that treating CVD costs the EU economy €192 billion annually [4]. Furthermore, it is known that over 80% of premature CVD cases are avoidable [35], therefore it is crucial that timely and reliable diagnosis of CVD is provided in a clinical setting.

Diagnosis of CVD requires the characterisation of the heart over a full cardiac cycle. The left ventricle is commonly used to determine the primary indicators of cardiac functionality. Of the four, this chamber is under the most workload, as it is required to pump the oxygen-rich blood supply around the entire body. As a consequence, pathologies can initially manifest themselves in the left ventricle. Therefore quantification of left ventricular spatial and temporal deformation is required in order to determine cardiac health in a patient.

Cardiac MRI is a popular imaging modality used to determine cardiac functionality. Its acquisition technology provides good contrast between blood and surrounding tissue structures. Datasets are commonly generated at high dimensions: 3D and, more recently, 3D+time. A 3D MRI dataset is presented to a physician as a stack of 2D cross-sectional short-axis images

(also known as 2D 'slices'). A 3D+time dataset[1] is simply an ordered set of 3D stacks, or volumes, where each volume is a snapshot of the heart at a temporal position in the cardiac cycle. In a 2D slice, the left ventricle is defined by two contours: the endocardium and epicardium, as shown in Fig. 1.1. These contours must be delineated in order to derive quantitative measurements of left ventricular function.

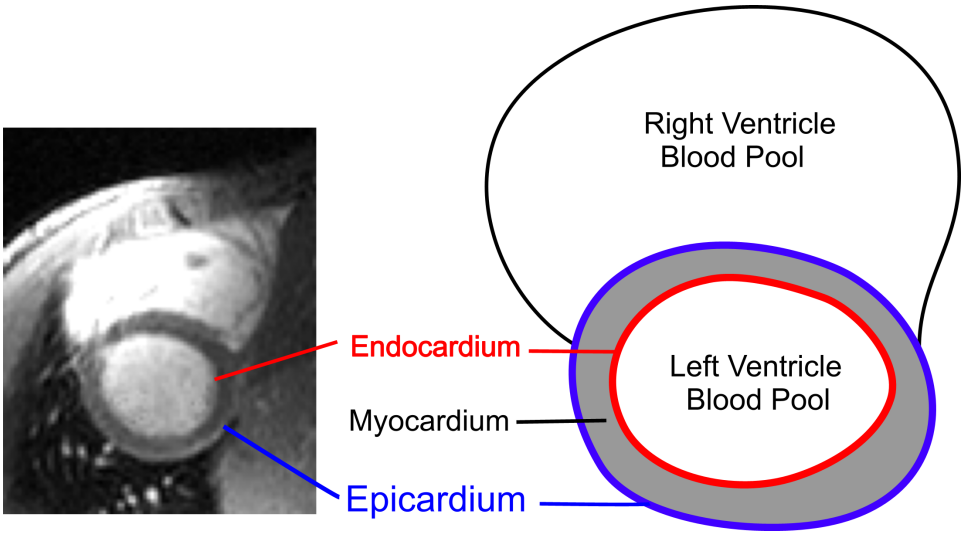


Figure 1.1: A sample 2D MRI image of the cardiac left ventricle is shown on the left. The endocardium is defined as the border between the bright blood-filled left ventricular cavity and the surrounding myocardium. This border is shown in red on the schematic representation on the right. The epicardium is defined as the outer edge of the left ventricle myocardium, shown in blue.

1.1 Motivation

Classically, this segmentation process is fully manual, carried out by a physician. For example, in a standard 3D+time cardiac image, with 10 and 20 slices in the spatial and temporal axes, respectively, 400 closed contours are required to be drawn. If each contour is defined with 32 landmarks, then a total of 12800 landmarks need to be manually placed by the physician to

[1]Throughout this thesis, the use of '3D' and '3D+time' act as a description of dimensionality. Of note, they do not refer to true 3D and 3D+time images. Therefore, as is common in the literature, the data used in this thesis contains spacing between slices. Please refer to Table 5.1 for full details of the datasets used in this thesis.

quantify LV functionality in a single patient scan. Clearly, this is an arduous task. The difficulty is further exacerbated by the fact that interpretation of cardiac MRI images is prone to inter- and intra-observer variability. This can be caused, for example, by low contrast between certain tissue structures or partial voluming effects, making contour definition somewhat subjective. Fig. 1.2 illustrates the range of image and anatomical variation seen in a population of left ventricle MRI datasets. These issues can have ramifications in terms of the efficacy of quantification accuracy and therefore on the accuracy of long-term patient studies.

Automatic segmentation algorithms offer a promising solution to this ongoing issue, with several distinct advantages over traditional manual segmentation. First, the rapid analysis and quantification of medical images removes the bottleneck of human interaction from the laborious manual segmentation process. Second, the continual increase in imaging resolution is no longer a burden on the physician. Instead, the usage of image information can be maximised, allowing for more accurate characterisations of patient scans to be achieved. Furthermore, more complex methods of image analysis can be employed, with the potential to create more subtle quantifiers of cardiac health, especially with respect to the temporal deformation of the heart (for example, stress and strain of the myocardial wall). Third, in medical research, where repeatable results are of paramount importance, automatic segmentation can be of great utility by reducing undesirable sources of variation in large studies. Finally, automatic segmentation algorithms afford a standard interpretation of medical images to be defined and implemented, facilitating accurate long-term observation of patients.

In the medical image segmentation domain, due to the complexity of biological structures, a popular segmentation strategy is to assume a model-based approach. Here, expert *a priori* knowledge is incorporated into the segmentation process, usually through a set of training samples. The model is constructed at the same dimensionality as the segmentation target. For example, modelling the left ventricle as a 3D object implies the segmentation of a 3D dataset. This is an intuitive approach, however, it is prone to significant limitations in practice due to several inter-related issues. First, as the dimensionality of the model rises so too does the required quantity of training samples. This is a direct result of the increased degrees of freedom

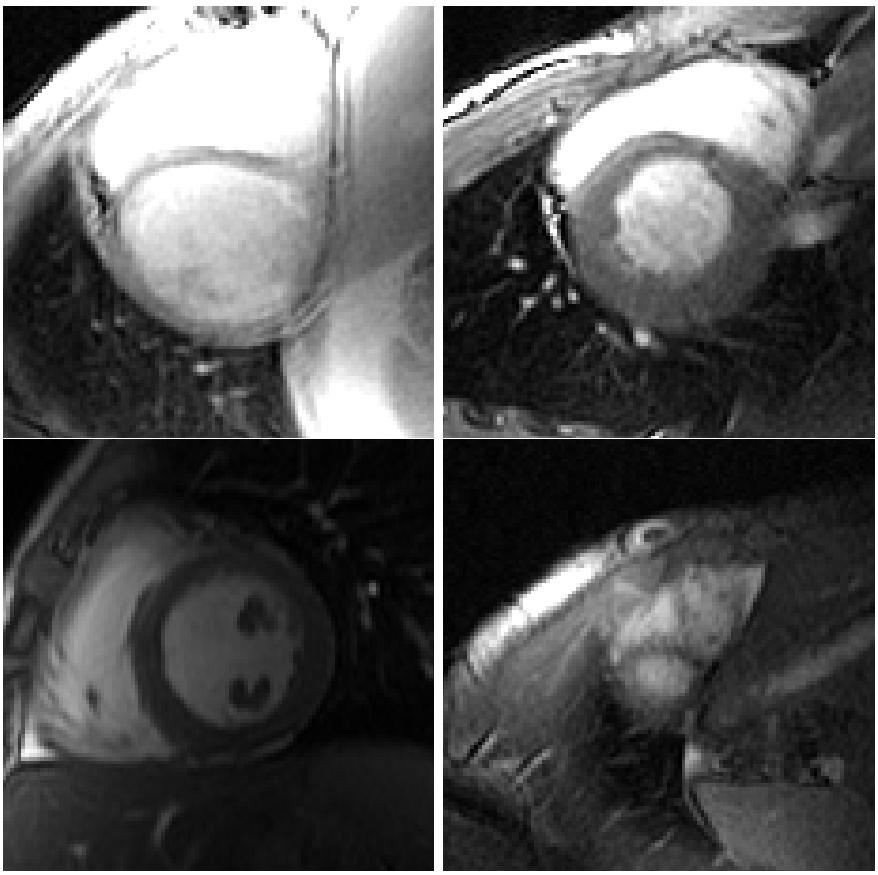


Figure 1.2: Four examples of the range of variation in cardiac left ventricular MRI images. The top-left image shows an enlarged left ventricle with a thin myocardium. The top-right image depicts a case of myocardium thickening (known as hypertrophy). The lower-left image illustrates the presence of the papillary muscles in the left ventricular cavity. The lower-right image shows an example slice towards the apex of the left ventricle.

in the model. Second, more model parameters are required to describe the extra level of model complexity, which can have an effect on convergence properties. Third, model complexity is also proportional to the number of local minima in the model, and also has negative effects on model convergence and accuracy. These first three issues are commonly referred to as the 'curse of dimensionality' (CoD) in the literature. Fourth, in the case of the left ventricle, two contours are needed to define the inner and outer borders of the left ventricular myocardium. Modelling these two contours simultaneously (i.e. considering them as a single entity) further exacerbates

the issues relating to the CoD. This is because the model must capture three different sources of variation simultaneously: the endocardium shape variation, the epicardium shape variation and the inter-contour distance seen at each training sample. Finally, to establish point-correspondence between training samples, complex registration procedures must be employed. This is a challenging problem in its own right.

1.2 Contributions

Motivated by these challenges, this thesis investigates an alternate modelling strategy, which is evaluated on 3D+time cardiac left ventricle segmentation. This segmentation framework can be used to automatically derive quantitative functional indicators of cardiac health from a patient scan. The proposed approach is shown to be: significantly less affected by the limitations caused by the CoD, adaptable to cases outside the training set and dimensionally scalable.

The first major contribution of this thesis investigates the necessity of explicitly modelling all the components of a higher dimensional segmentation target that exhibits complex spatial and temporal dynamics. Specifically, a novel contour coupling technique is developed where (in the case of left ventricle segmentation) the endo- and epicardium are not modelled as a single entity. Instead, they are considered separately at model-time, and, through the integration of expert *a priori* knowledge, non-rigid segmentation-time constraints enforce biologically-sound structural consistency into the segmentation process. A primary advantage of this contribution, compared to a combined model (which considers the endo- and epicardium as a single shape), is that it can adapt to contour configurations outside the training set. This is a significant advantage as the training set usually strictly defines the outer limits of allowable shape synthesis in model-based methods. This contribution is evaluated using two separate implementations: one data-driven and one model-driven. A detailed performance characterisation is given between the two, and the model-driven contour coupling method is experimentally proven to be superior in cardiac left ventricle segmentation.

The second contribution explores an alternate modelling strategy for high-

dimensional segmentation, whereby the spatial and temporal components of the 3D+time left ventricle are modelled separately, and combined later at segmentation-time. This exemplifies the modelling strategy proposed in this thesis. By modelling the shape, spatial and temporal components separately (while maintaining structural integrity through contour coupling constraints), the issues related to the CoD are greatly alleviated, allowing the complex cardiac 3D+time dynamics to be captured with reduced training set sizes. This has significant practical implications, since large annotated medical datasets containing both healthy and pathological cases are rare. Furthermore, this is a dimensionally scalable approach and can be applied in 2D+time, 3D and 3D+time scenarios.

The third contribution unifies contour coupling with the alternate modelling approach outlined above into a comprehensive segmentation framework, and evaluates the limits of required training to create practical models for segmentation. The proposed approach is evaluated on a publicly available dataset of 3D+time cardiac MRI images, which exhibit significant inter-dataset variation. It is proven experimentally that the proposed method requires only a fraction of the full 3D+time dataset for training in order to capture the complex spatial and temporal dynamics of the heart. This is a direct result of the novel modelling approach adopted in this thesis.

In essence, this thesis performs a critical analysis on the standard model-based approach to higher-dimensional modelling, details its limitations, and offers an alternative modelling strategy that solves many of the issues related to the adoption of the model-based approach in higher-dimensional segmentation problems. The result is a flexible, extensible segmentation framework that incorporates *a priori* biologically-based constraints in an adaptive manner.

1.3 Thesis Layout

The remainder of this thesis is organised as follows:

Chapter 2 provides a detailed introduction regarding the background material of this thesis. A description of the major imaging modalities is provided,

with special emphasis given to cardiac imaging applications. Next, a comprehensive overview of the current state-of-the-art in the literature is provided, where the main trends of research are analysed. This section gives context to the contributions presented in the remainder of this thesis and illustrates how the research complements and improves on the state-of-the-art.

Chapter 3 is concerned with the Active Shape Model (ASM). The ASM is the primary statistical modelling technique used in this thesis to learn shape, spatial and temporal variation of the left ventricle. As such, an in-depth description of its standard formulation is given, as well as the major extensions presented in the literature. The limitations of the standard ASM implementation are described when applied to left ventricle segmentation. Then, a novel optimisation technique is detailed that improves the robustness of ASM-fitting stage in complex imaging environments. This optimisation technique facilitates more significant contributions made in the remainder of this thesis.

Chapter 4 investigates modelling multiple contours using model-based approaches. Specifically, critical analysis is performed regarding the application of the ASM to simultaneously model the two contours that define the left ventricle. The limitations of this approach are examined and characterised. A novel contour coupling method is presented that replaces the standard combined modelling strategy. Here, non-rigid biologically-sound constraints are enforced at segmentation-time between separate shape models that expertly capture left ventricle shape variation. Two implementations of this technique are presented with direct comparison made against standard methods. It is shown that the proposed contour coupling approach can adapt to contour configurations outside the training set, usually a fundamental limitation of model-based techniques. This is shown to be of great benefit as the left ventricle exhibits significant anatomical variation, especially when pathologies are considered.

Chapter 5 extends the hypothesis developed in Chapter 4 to a more complex problem, whereby the task of modelling the cardiac left ventricle in 3D+time is decomposed into its component parts. Separate shape, spatial and temporal models are created at model-time. It is illustrated that the proposed technique avoids the prohibitive issues related to the curse of dimensionality

compared with a strategy that directly models the 3D+time structure. The separate models are combined at segmentation-time, unifying the spatial and temporal *a priori* learned variation. To evaluate the performance of the proposed modelling method, the contributions from Chapters 3 and 4 are combined, creating a practical segmentation framework that is adaptable to multiple contours and scalable to changes in dimensionality. The framework is applied to the segmentation of a real set of 3D+time cardiac MRI datasets that exhibit significant image and anatomical variation. The proposed segmentation method is directly compared against state-of-the-art techniques, and shows comparable segmentation accuracy while avoiding CoD-related issues. In fact, the proposed method can learn the same cardiac deformation knowledge from only minimal training sets, illustrating that a large portion of the full 3D+time annotation is redundant in its description of cardiac spatio-temporal dynamics.

Chapter 6 concludes this thesis. A summary is provided of the main arguments, research context and contributions presented throughout the thesis. Several avenues of future research are identified where the proposed contributions can be further improved and extended beyond their current state.

Chapter 2

Background

This chapter provides a detailed introduction to the heart, describing: anatomy, function, pathologies and indicators of cardiac health. Several cardiac pathologies that are relevant to this thesis are discussed, as well as the importance medical imaging plays in their diagnosis and management.

A full overview of the most prevalent medical imaging modalities is presented. Details are provided regarding the physical basis of each modality, the conditions where each modality is most suited, and the level of invasiveness of each imaging procedure. Special emphasis is given to the application of each modality to cardiac imaging. A note is made on the future direction of multiple modality fusion in medical imaging.

Finally, a detailed literature review is provided, where a general survey of the medical image segmentation field is presented, as well as a comprehensive description of the relevant research work presented in the literature, as it relates to this thesis. Analysis of the main research trends is provided, followed by a justification, based on the conclusions from the literature review, of the choice of model-based method assumed in this thesis. This gives context and relevancy to the contributions presented in the remainder of this thesis.

2.1 The Heart

The heart is responsible for the flow of blood around the body. Functionally, it can be described as a highly-optimised pump, responding to periodic electrical pulses from the brain, which in turn control its rhythmic pumping action. In fact, it is more accurate to describe the heart as a collection of pumps that operate in a synchronised, but independent, manner. The heart can be conceptually subdivided in several ways. First, it can be considered as two halves: the left and the right. The right half of the heart pumps de-oxygenated blood from the body to the lungs. The left half takes the freshly oxygenated blood and pumps it around the rest of the body. This dual action is cyclical and occurs approximately 70 times a minute in every healthy human being[1].

This conceptual model can be further decomposed, whereby each half of the heart is again divided in two. This is a closer representation of the anatomical level, where each half of the heart contains both an atrium and ventricle, and is illustrated in Fig. 2.1. Each of these four chambers is its own pump. The atria can be thought of as primer pumps for the ventricles, which are depended upon to execute the main pumping action; either to the lungs or the peripheral organs. The cardiac valves ensure the precise flow of blood from chamber to chamber, and prevent back-flow or other unwanted effects. The valves are placed under considerable pressure, and, as a result of their opening and closing action, produce the characteristic beating sound of the heart. The cardiac cycle itself can be broadly divided in the systolic and diastolic phases, although in reality it is quite complex. In diastole, the heart fills with blood, and in systole the heart contracts, pumping the blood from the cardiac chambers. Therefore, at the end-systolic point the heart is most contracted, and at the end-diastolic phase the heart is fully relaxed. This provides a general 'working-picture' of the operation of the heart. A complete medical description is obviously far more complex and outside the scope of this thesis, and is covered in detail in standard physiology books [1].

[1] This figure refers to the average heartbeat when a person is at rest.

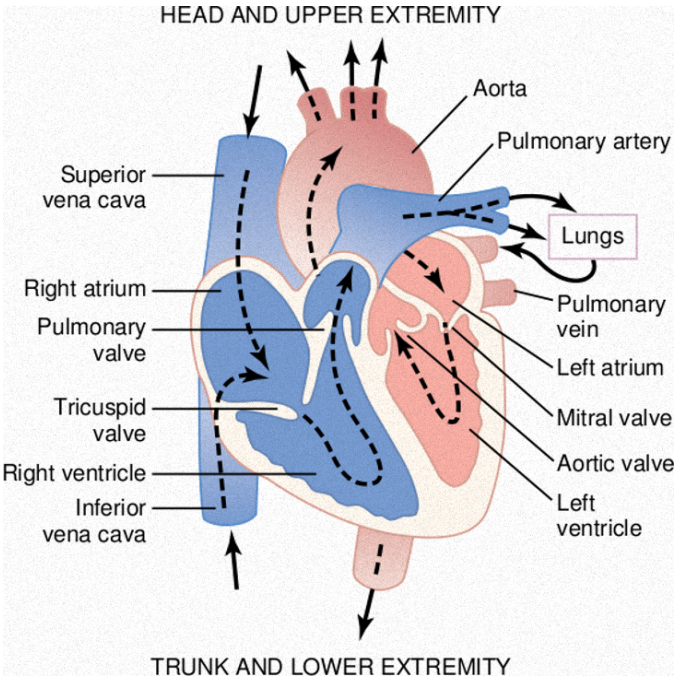


Figure 2.1: Schematic diagram of the heart structure, illustrating the relationship between the cardiac chambers and the flow of blood. Image taken from Guyton and Hall [1].

2.1.1 Anatomy

Working from the outside in, the cardiac structure is contained within the pericardium sac. Next, the epicardium defines the outer wall of the heart. The myocardium is the actual heart muscle, which contracts to pump blood to and from the chambers. Finally, the endocardium is the inner lining of the heart wall, which is in direct contact with the blood itself. In essence, each cardiac chamber has its own endo- and epicardium defining the inner and outer walls, respectively.

Since the left ventricle is the central segmentation target in this thesis, a more detailed description of its structure is warranted. In general, the left ventricle myocardium is approximately three times larger than the right ventricle myocardium. This is a consequence of the extra workload the left ventricle is under, as it has to pump blood around the entire body, as opposed to just the lungs. The left ventricle is also slightly larger than the right and is conical in shape, narrowing towards the apex of the heart.

Internally, the left ventricular cavity contains the papillary muscles and trabeculae carneae. These two sets of muscles lend support to the mitral valve when it closes under pressure to stop valvular regurgitation. It should be noted that there is no consensus in the medical community as to whether these muscles should be considered part of the blood pool or not when quantifying ventricular volume [8].

2.1.2 Functional Indicators

Assessment of cardiac health is determined through the derivation of several relatively simple measurements. More complex motion and deformation analysis can yield more subtle descriptions of cardiac health. Both categories will be discussed below.

Left ventricular volume is a simple measurement that indicates the volume of blood being pumped by the heart. For example, the difference between the end-diastolic volume (EDV) and the end-systolic volume (ESV) determine the Stroke Volume (SV) of the heart.

$$SV = EDV - ESV \tag{2.1}$$

The Ejection Fraction (EF) is usually one of the primary indicators used to determine cardiac health, and is derived from the stroke volume according to:

$$EF = \frac{SV}{EDV} \times 100\% \tag{2.2}$$

Cardiac output (CO) is a measure of the quantity of oxygenated blood being pumped by the heart and is also derived from the stroke volume, as well as from the heart rate (HR):

$$CO = SV \times HR \tag{2.3}$$

Typical measurements of EDV, ESV, SV, EF, HR and CO in a healthy patient are: $120ml$, $50ml$, $55-100ml$, 58%, 70 beats per minute and $4.9L/min$.

The thickness of the myocardium can be used to indicate several common cardiac pathologies. For example, excessive thickening of the myocardium (hypertrophy), can be identified through the LV mass (LVM). Assuming the LV myocardium volume can be approximated as the difference between the LV epicardium volume (V_{epi}) and the LV endocardium volume (V_{endo}), then the LVM is defined as (taking $1.05g/cm^3$ as muscle density):

$$LVM = 1.05 \times (V_{epi} - V_{endo}) \tag{2.4}$$

Note that all of these indicators are based on volumetric measurements of the ventricle. This highlights the dependence on robust and repeatable quantification of the left ventricle in a clinical setting. While the global functional indicators provide a simple description of the input/output characteristics of the heart, there are more advanced descriptors of cardiac health that analyse temporal deformation. These parameters aim to determine the dynamic behaviour of the myocardium over a cardiac cycle. Wall thickening for example, attempts to detect abnormal myocardial contraction. Strain analysis can also be used to describe complex deformation of the heart. These measurements can depend on more sophisticated image modalities such as tagged-MRI, and are not addressed in this thesis. Of course, these more complex measurements would also be extremely challenging to extract using a manual annotation process, and are thus precluded on the development of robust segmentation procedures.

2.2 Imaging Technologies

Medical imaging technologies are based on the principle of reconstruction of a controlled input signal after it has passed through an object. In the majority of cases, the signal comes from within the electromagnetic spectrum (the only exception being ultrasound, relying on mechanical waves, although the same discussion applies) and is emitted in a highly specified manner, so that after it has interacted with the object, the resulting signal can be analysed and used to reconstruct an accurate internal image of the object.

This is a fascinating idea. More fascinating, is the fact that the choice of

signal (i.e. the chosen frequency to emit from the electromagnetic frequency) can completely change the nature of the input signal and the interaction with the object itself. With such a broad electromagnetic spectrum, it is no surprise that several distinct medical imaging modalities have emerged since the first, X-rays, was discovered in 1895.

In general, higher imaging resolution and quality comes at the expense of higher cost and higher risk to the patient, which is in the form of ionising radiation. Lower-resolution modalities usually have little or no patient risk, but have relatively poor image quality. Moreover, each modality excels in a particular area. Therefore, the choice of modality depends on a wide range of factors, including the object under investigation, patient age and health, as well as economic considerations.

An overview of the major modern imaging modalities is presented in the following sections, covering: physical basis, level of invasiveness and the niche area(s) of each modality. Special emphasis is given to Magnetic Resonance Imaging, and its application to cardiac image acquisition, as it is the primary modality used in this thesis.

2.2.1 X-ray Radiography

The first diagnostic imaging modality, X-ray radiography, was discovered by Wilhelm Roentgen in 1895. The X-ray is useful because it exhibits very limited interaction with the human tissue structures. In fact, soft-tissue is, for all intents and purposes, invisible to the high-energy X-ray. This is a useful property, as it allows for high-resolution imaging of the otherwise hidden skeletal structure. However, X-rays expose the patient to ionising radiation, and so, have associated risk factors. Although these risks are relatively low for a one-shot imaging procedure, prolonged exposure to X-rays is dangerous. So, while the X-ray imaging modality provides excellent skeletal images, it cannot be used for long-term patient monitoring studies.

The modern manifestation of this technology is X-ray Computed Tomography, or CT. While the original X-ray method exposed the patient to a single burst of ionising radiation, the CT method increases the required radiation exposure level. This increased risk, however, comes at great benefit. CT al-

lows very high resolution 3D and 3D+time scans of the patient to be created from a large series of 2D slices. There has been a continual advancement in CT technology, leading to new imaging methods. For example, computed tomography angiography (CTA) is a non-invasive procedure that can be used to determine whether a patient has coronary artery disease (CAD). There are two main trends in the development of this modality that are opposed to each other but serve different purposes. On one hand, there is the demand for higher imaging quality, thus requiring higher doses of radiation. And on the other hand, the need for creating safer low-dose CT protocols to maximise patient safety.

2.2.2 Positron Emission Tomography

Positron Emission Tomography (PET) relies on the injection of a radioactive tracer substance into the patient before the scan can take place. The tracer undergoes beta decay (in effect, emitting photons from the body), which is detected by the external PET scanner. By contrast, a similar method, Single photon emission computed tomography (SPECT), also relies on an injected tracer but emits gamma radiation directly. SPECT can be used to create true 3D images and is useful in cardiac and neurological imaging. PET, however, exhibits greater spatial resolution and temporal resolution, allowing for rapidly changing biological processes to be accurately imaged.

PET has the unique ability to visualise metabolism in real-time and as a result has contributed greatly to the understanding of cardiac functionality, for example. This attribute has allowed its adoption in clinical practice in locating and diagnosing coronary stenosis [107].

2.2.3 Ultrasound

Compared to the other imaging modalities, ultrasound provides a low-cost imaging solution that is portable and versatile. Based on relatively simple principles of physics, ultrasound image formation relies on the propagation and reflection of mechanical waves in a medium. High-frequency sounds waves are produced by a piezoelectric transducer, which also acts as a receiver that collates reflected sound waves. Ultrasound is completely safe,

utilising no ionising radiation and requiring no contrast agents to function. In a sense, ultrasound is more interactive with the operator, since imaging systems are usually 'live': that is, the operator can image an area in real-time. Although this can be of great benefit in explorative imaging, it requires considerable operator skill and precision. This has led to the use of ultrasound in image-guided procedures. However, ultrasound is limited with respect to imaging depth, and cannot penetrate bone structures, precluding its use in neurological imaging.

Cardiac ultrasonography benefits greatly from Doppler ultrasound, where blood flow in the heart can be visualised in real-time. This plays a crucial role, for example, in the identification of valvular defects. Cardiac ultrasound also allows the derivation of cardiac output and ejection fraction. Although in 2D ultrasound this requires using geometrical assumptions about the left ventricle (which may not be accurate in pathological cases), the advent of 3D ultrasound has the potential to negate this requirement. Generally, derivation of functional indicators requires segmentation of the endocardium (i.e. the left ventricle blood pool/myocardium border). The epicardium needs to be segmented to asses wall motion characteristics. Like Magnetic Resonance Imaging, discussed in the next section, segmentation of the epicardium is significantly more challenging than the endocardium, especially towards the apical region. While this is a tractable problem in MRI segmentation, epicardium segmentation is very difficult with traditional ultrasound imaging systems due to poor contrast and image noise, leading to very little clinically validated related works presented in the literature. Despite this, much research has been presented towards automatic endocardium segmentation in 2D [13], 2D+time [17], 3D [69] and 3D+time [72]. Due to the noisy nature of ultrasound images, modelling the imaging physics is normally an integral part of segmentation techniques to reduce speckle noise and other image artefacts. Low spatial and image contrast, coupled with the high dependence on operator skill, means ultrasound is suited for non-repeatable, explorative images that do not require very accurate quantitative measurements. A full review of ultrasound segmentation techniques is covered in the recent review by Noble and Boukerroui [75].

2.2.4 Magnetic Resonance Imaging

Magnetic Resonance Imaging (MRI) is a widely used non-invasive imaging modality. Based on the principal of Nuclear Magnetic Resonance (NMR), MRI does not expose the patient to any ionising radiation, and is able to produce images with high spatial and temporal resolution. MRI is extensively used in cardiac imaging, and is currently the gold standard for functional assessment of cardiac health [9, 3, 85]. MRI image acquisition is very flexible, with many different imaging protocols in existence today, which can emphasise different tissue types.

Physical Operation

The primary means of interaction MRI has with organic tissue is through hydrogen atoms. Since hydrogen is abundant in water, which makes up $60 - 70\%$ of the human body, an MRI image is, in essence, an image of the relative presence, or absence, of water in a tissue sample.

The exact physical operation of MRI is based on quantum mechanical principles, and as such an in-depth discussion of its mathematical basis is beyond the scope of this thesis. However, a general technical discussion of its operation will be given and quantum mechanical results will be used where required.

Each molecule has an associated 'spin' value, characterised by its number of protons and neutrons. The hydrogen atom, along with carbon and nitrogen, has a spin value of $\frac{1}{2}$. Particles with this spin value, when placed in a time-independent magnetic field, can assume only two energy levels (from the Zeeman effect): $E_\uparrow$ and $E_\downarrow$. Both energy levels are equal in magnitude, but act in opposite directions, and are directly proportional to the strength of the applied magnetic field.

When a static magnetic field is applied to organic tissue, a potential imaging voxel (i.e. a three-dimensional imaging unit) obviously contains a very large number of protons. The distribution of energy levels in each voxel will always be such (by virtue of statistical quantum mechanics) that the overall polarization of a voxel will be in alignment with the direction of the ap-

plied magnetic field, corresponding to a majority $E_{\uparrow}$ proton spin. Crucially, the magnitude of the magnetisation vector, and the subsequent strength of the image signal (and hence image quality), is directly proportional to the strength of the applied field[2]. However, one cannot directly measure this magnetization vector. Instead, this state (referred to as dynamic equilibrium) must be disturbed, and the *response* to the disturbance is measured, and used to reconstruct the image.

A controlled disturbance of the dynamic equilibrium state is achieved by application of a separate electromagnetic signal, which is emitted at the resonant frequency of the protons. This is referred to as the *Larmor frequency* and is related, once again, to the applied static magnetic field. The effect of this field (actually an RF signal) is to disturb the net magnetized vector in each imaging voxel. Specifically, the RF signal introduces a transverse component into the vector, which is absent in dynamic equilibrium, and is the component that can be measured to form the MRI image. This effect is illustrated in Fig. 2.2.

After the RF pulse ceases, the deflected protons begin to realign with the original axis of the static magnetic field (like a disturbed compass needle) and release their own RF signal in the process. This second, induced, RF signal is captured, and forms the signal from which the MRI image is generated. The vector component parallel to the static field returns to its dynamic equilibrium state quite slowly, and its time to do so is called *T1 relaxation*. In contrast, the transverse component restores itself much faster, and its time to do so is called *T2 relaxation*. The MRI image can be weighted towards the T1 or T2 signal, depending on the required image characteristics.

The character and number of the first RF input pulse, used to disturb dynamic equilibrium, completely determines the kind of image produced, and is usually central in the description of the MRI imaging protocols. For example, spin echo sequences are used for high velocity blood imaging, where blood is dark in the image. In contrast, in gradient echo sequences blood appears bright.

[2]Therefore, to improve MRI image quality there has been a technological progression to increasingly higher applied magnetic fields. A standard MRI field seen in clinical practice can be either 1.5 or 3 Tesla. However, the latest MRI scanners produced are based on 7T magnetic fields.

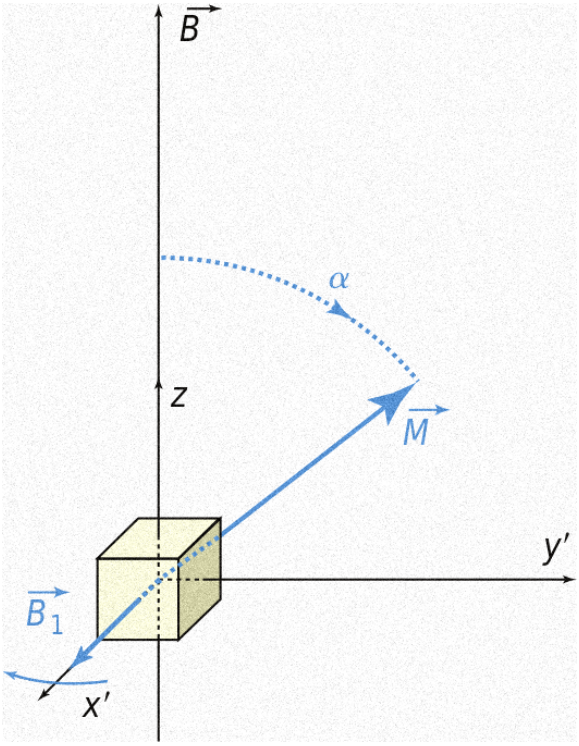


Figure 2.2: Initially, the application of the static magnetic field ($\vec{B}$) aligns the magnetisation vector associated with an imaging voxel (depicted as a box in this figure) in the z direction. Then, the application of the separate RF signal ($\vec{B_1}$) in the x direction, deflects the magnetization vector ($\vec{M}$) by an angle α. Image taken from Suetens [99].

Cardiac MRI

Cardiac MRI (CMR) is considered the gold standard for quantification of left ventricular function [9, 3, 85]. This is a result of the high spatial and temporal image resolution, as well as the repeatable and low-risk nature of the MRI modality. Using MRI to image the heart over a cardiac cycle requires Electrocardiographic (ECG)-gating. ECG-gating allows synchronisation between image acquisition and heart motion. The accuracy of ECG-gating is therefore crucial in Cardiac MRI, to avoid motion artefacts in a patient scan.

The bright-blood imaging technique is today the dominant approach for CMR. The standard MRI protocol is called steady-state free precession (SSFP) and is based on the gradient echo (bright blood) method. SSFP acquisition is quite fast, and can be achieved in a single breath-hold. Other

bright blood protocols include: standard non-breath hold and fast breath-hold segmented k-space. The datasets used in this thesis (full details are presented in Table 5.1 and discussed in Chapter 5) are acquired using the FIESTA protocol (also known as the Balanced Steady-State Free Precession or TrueFISP protocols). This protocol is well suited to cardiac cine imaging and produces good contrast between the blood pool and myocardium. It is also less sensitive to motion artefacts, making the technique useful in cases where the patient has difficulty holding their breath. The primary indicators of cardiac health can be derived from the short-axis slices (i.e. cross sectional views) of such an image. In normal clinical practice, where manual annotation is required, usually only the end-systolic and end-diastolic phase volumes are segmented to provide these quantitative measurements. Volume measurements are constructed using Simpson's rule. MRI images can be enhanced through the use of contrast agents, injected into the patient before image acquisition. This makes the procedure invasive, but this can be useful, for example, for the detection of myocardial infarction where the contrast agent will remain in the infarcted tissue for longer than normal myocardium [107].

There is a well defined method for physical image acquisition of the cardiac structure, regardless of imaging technique used, where a standard set of imaging planes are acquired to assess cardiac function. First, the sagittal images, taken parallel to the spine, are essential as they form the basis of all subsequent image planes. Usually the mitral valve plane and cardiac apex are acquired at this stage [68].

Next, the four chamber cardiac plane is localised from the sagittal images. The four chamber image requires the operator to bisect the plane from the cardiac apex to the mitral valve. This image allows the operator to determine the orientation required to take the multiple short-axis slices of the left ventricle, which are then used for left ventricle quantification. The short-axis plane is perpendicular to the cardiac septum, which separates the left and right half of the heart. Maintaining this plane, the left ventricle is imaged in multiple slices from basal-most to apical-most locations. This constitutes the standard imaging acquisition for left ventricular imaging.

If further imaging planes are required for diagnosis, several long-axis views

are acquired. As the name suggests, these are orthogonal to the short-axis images and taken at specific angles in order to image: the left ventricle and atrium (left ventricle long axis two chamber view), the three chamber view and the four chamber view. This process is illustrated graphically in Fig. 2.3. In most cases, the right ventricle is imaged using the same planes as for the left ventricle to create a similar set of images.

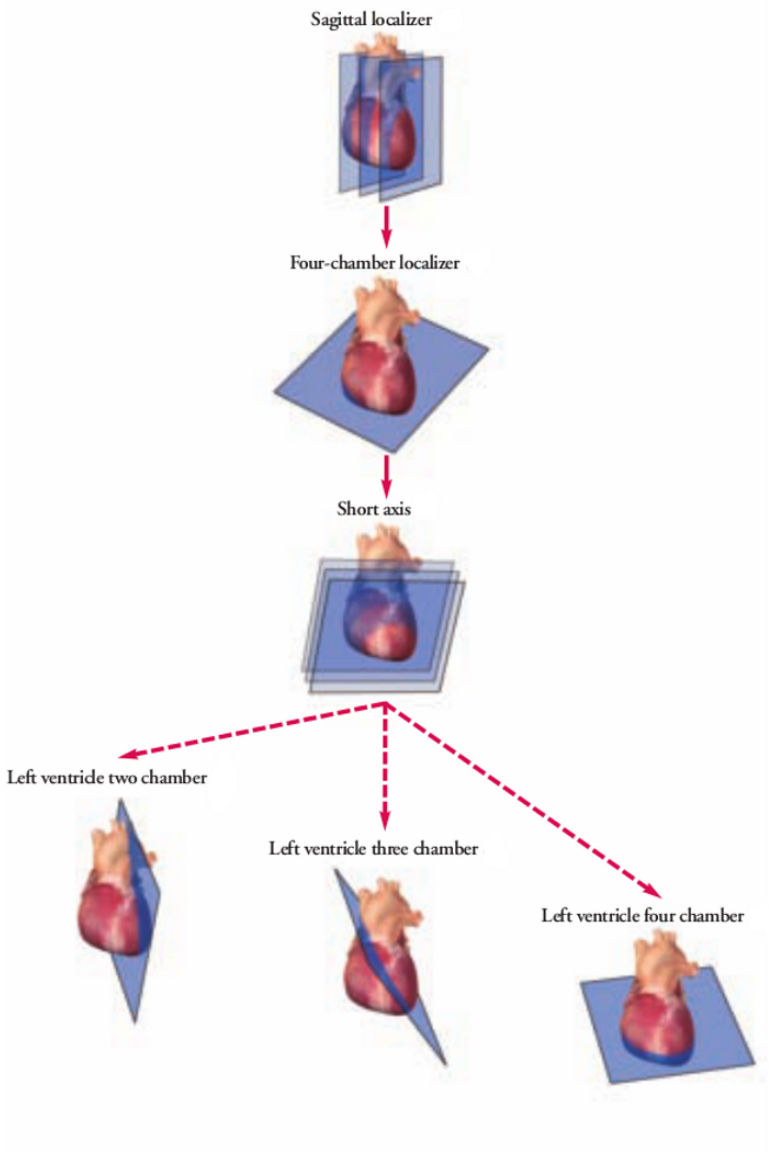


Figure 2.3: This image depicts the established image acquisition progression for cardiac datasets using MRI. This image is taken from McGee *et al.* [68].

Ultrasound and CMR are usually the primary imaging modalities used for investigation into patient cardiac pathologies. Ultrasound can be useful for initial observation of the patient, providing a qualitative assessment of cardiac health. CMR is used when accurate quantitative imaging of the cardiac anatomy is required to diagnose pathological cases and when sequential studies need to be carried out. For example, it is used to determine presence of tetralogy of Fallot, pulmonary atresia, tricuspid atresia and univentricular hearts [82].

CMR is also used to determine the presence of more common cardiac pathologies such as ventricular hypertrophy (thickening of the myocardium), myocardial infarction, valvular regurgitation, ventricular dilation, acute myocarditis and apical ballooning [68].

MRI Segmentation Challenges

In this thesis, the segmentation strategy developed is evaluated on 3D+time cine CMR images. These images are often challenging to segment. First, it is common, in practice, for MRI images to have anomalous inter-slice translations, due to patient movement. This can cause difficulty in methods that assume a proper 3D structure is present. Consequently, an image registration step is usually required before segmentation if this assumption is made [97, 98]. Registration is itself a very challenging topic, especially in the 3D+time case. Moreover, as highlighted in an earlier discussion, the correct imaging planes of the heart, while well defined, can be difficult to accurately acquire. For example, when acquiring short-axis slices of the ventricles it is difficult to determine the actual plane that is orthogonal to the inter-ventricular septum.

Second, while contrast between blood and muscle is excellent in MRI images, the contrast between tissue structures can often be minimal. For example, the contrast between the epicardium and the adjacent thoracic and hepatic regions can be low or even absent. As noted earlier, this leads to subjective interpretations in the manual annotation process. In the automatic case, segmentation of such regions is challenging, requiring maximum use of *a priori* knowledge to make an informed segmentation where image information is absent. Furthermore, towards the apex of the heart the inter-tissue contrast

becomes diffuse, again resulting in difficulties for segmentation methods that rely on image data alone. As MRI datasets are commonly built from short axis slices (i.e. without any long axis supplementary planes) the apex can often be difficult to determine with confidence and accuracy.

Third, the papillary muscles can cause significant ambiguity in the definition of the true endocardium border. There seems to be conflicting reports from medical organisations whether or not to include the papillary muscles as part of the left ventricular blood pool [8]. It appears somewhat more common to assume they are part of the blood pool and this assumption is followed throughout this thesis. Even so, in many short axis slices it can be very difficult to determine where the border between the endocardium and papillary muscles truly lies.

Finally, with the large inter-patient variation in anatomical form and temporal deformation in a population, especially when pathologies are included, the process required to generate practical models of the heart (discussed in Section 2.3.1) is quite challenging due to the lack of widely available training samples.

2.2.5 Image Modality Fusion

Recently, with the continual improvement of imaging technologies and the increasing integration of digital technology into the image analysis process, there has been a strong trend towards the fusion of image modalities. Image fusion combines the benefits of each modality, creating more detailed scans. It has been shown that fusion of multiple datasets is superior to a side-by-side comparison between modalities [5]. A popular combination of modalities is PET/CT [49], which, as a combined image, can be generated in a single imaging session. While SPECT has classically been used for the identification of cardiac pathologies such as Coronary Artery Disease (CAD), recent trends indicate a move towards the prediction and prevention of CAD [107]. To this end, PET can provide accurate quantification of myocardial blood flow and metabolism, and, when coupled with the highly accurate anatomical imaging CT, provides a hybrid image that can asses myocardial perfusion and coronary anatomy simultaneously [58].

Where modalities are not combined at acquisition time, image registration provides the basis for fusion of datasets that have been taken sequentially, or even at significantly different times. Indeed, there has arisen many clinical applications of various image modality combinations in the medical community, such as: PET/MRI [73], CT/MRI [45, 67] and Ultrasound/MRI fusion [93].

Although not the focus of this thesis, it is felt that the combination of different imaging modalities has the potential to establish an important trend in future research and could open up many new avenues of research.

2.3 Literature Review

The medical imaging field is composed of a set of diverse research topics, including: image and model registration, image modality fusion, robotic applications, statistical shape modelling, motion tracking, augmented reality applications and, of course, image segmentation. In general, the advancement in sophistication in all of these areas has progressed in step with the ever-present increase in computing power available to research institutions and individuals. Algorithms that are implemented today are likely to have been impossible to realise ten years ago. Currently, it is common to see research presented in the literature that takes advantage of modern Graphical Processing Unit (GPU) technology [76] and computing clusters [78].

Medical image segmentation greatly benefits from being an inter-disciplinary community where engineers, mathematicians, statisticians and clinicians closely cooperate in the development of integrated diagnosis solutions. For example, the development of general level-set methodology by mathematicians Osher and Sethian [79] was integrated into medical segmentation [66, 16] shortly after its development as a technique for the implicit definition of surfaces, and has achieved widespread usage in many scientific applications.

Image segmentation is concerned with robustly identifying a target structure in a single image, or a general set of images. The difficulty in the development of segmentation algorithms is determined by: the complexity of the target structure, the variation in image quality and the potential for the

target structure to be occluded in the image. Under controlled imaging environments (for example, in many industrial applications) these difficulties are minimised. In these cases, the segmentation algorithm can be relatively simple, relying perhaps, on a simple thresholding scheme or a rigid template matching technique. In more complex scenarios (for example, in medical image segmentation), very few assumptions can be made about the image quality, or level of definition of the target structure. In these cases, the sophistication of segmentation algorithms needs to be orders of magnitude higher. It is these kind of complex image segmentation challenges that have driven the interest of the medical image segmentation community.

2.3.1 Segmentation Method Categories

Modern segmentation algorithms can be broadly categorised into data-driven or model-based approaches. These two approaches have orthogonal philosophies. Data-driven methods rely purely on image data to determine the correct delineation of the segmentation target. By contrast, their model-based counterparts create learned models that are constrained based on expert *a priori* information. As a consequence, segmentations can be constrained to likely biologically-valid cases and are less likely to be influenced by image noise and image acquisition artefacts. The use of *a priori* knowledge is very useful in noisy images, where part of the target structure's boundary can be diffuse or even absent. This prevents over-segmentation and discontinuities, which may occur in the data-driven scenario.

It is natural that practical segmentation strategies tend to take elements from both approaches (in the model-based case this is a fundamental requirement, as segmentation requires some interaction with the image data!), however most methods can still be classified with a leaning towards a particular faction. However, it is widely recognised that maximising the use of *a priori* knowledge in a segmentation technique is a pre-requisite of robust segmentation algorithms in complex imaging environments. Consequently, model-based methods have enjoyed great popularity in the last ten years in the medical image segmentation research domain, and form the basis of the research carried out in this thesis.

To give context to the research presented in this thesis with respect to the

state of the art work, a detailed overview of each of the main model-based methods is given. Since the Active Shape Model is used in this thesis to capture the shape, spatial and temporal deformation of the cardiac left ventricle, special attention is given to the research developments in the statistical modelling area. For each segmentation technique, the major milestone papers are detailed, as well as the novel extensions that have been presented in the literature. The similarities or differences are indicated, where applicable, with the research presented in this thesis.

2.3.2 Data-driven Segmentation Techniques

Data-driven segmentation techniques attempt to use only the image information to partition the image into meaningful regions. While this is an intuitive approach (and is usually sufficient in simple image scenarios), the data-driven technique can be somewhat limited when applied to complex environments, such as medical image segmentation. These techniques are usually negatively affected by image noise, missing or poorly defined object boundaries, and object occlusion. Below, a brief survey of the historical data-driven methods are provided, with more detail regarding their modern manifestations. This section concludes with the evidence that the inclusion of *a priori* information into the segmentation process is a pre-requisite in complex imaging environments.

Threshold-based Methods

Image thresholding is one of the earliest segmentation techniques. It is based on the simple premise that the target region one wishes to segment is homogeneous and forms a clear and continuous boundary with surrounding structures. This is, of course, a very idealised set of assumptions, making the application of thresholding techniques on their own today rarely practical. For example, in medical image segmentation the anatomical structure to be segmented can be in the presence of significant noise. This led to the development of image filtering methods that remove, or minimise, image noise. Gradient information is also used to enhance the thresholding process, leading to the development of kernel-based edge-filtering techniques.

The most common of these are the Sobel, Laplacian and Canny [15] edge detectors.

Histogram analysis is a common means to improve the sophistication of thresholding techniques. In this approach, the histogram describing the image pixels is collated, where peaks, troughs, and other measurements can be employed to select the optimum point for one or more thresholds. Other methods focus on modelling the foreground and background of an image using, say, a Gaussian mixture model, whose parameters are iteratively updated to maximise the difference between models. Research has also been directed towards thresholding based on both the pixel intensity as well as the spatial attributes of a pixel. In this case, local neighbourhood information around a pixel can also be employed to determine its class. Multi-pass thresholding schemes have also been developed, where an initial global threshold is subsequently refined on a local basis. This style of thresholding is employed in document binarisation, before optical character recognition techniques are applied. For a full review and evaluation of modern thresholding techniques, see Sezgin and Sankur [92].

An example of a more advanced thresholding approach (based on morphology) is the watershed algorithm, which still enjoys application in medical image segmentation today [32, 42, 29]. Here, an image is considered as an elevation map, where gradient information corresponds to relative levels of height. The watershed algorithm applies various levels of 'flooding', via thresholding. Like most purely data-driven techniques, it is sensitive to noise and low contrast levels. Although, recent work by Cousty *et al.* [29] illustrates that an extra level of robustness can be included in this approach through the inclusion of different sources of *a priori* knowledge for left ventricle segmentation.

Clustering Methods

The natural advancement of the thresholding concept, is the set of clustering methods. Clustering aims to divide a target image into a set of homogeneous regions. This is normally an iterative process, requiring seeding in the region of interest. Clustering methods can be either unsupervised (i.e. employ no *a priori* information) or supervised (i.e. employ a training stage). K-

means clustering is a classical unsupervised method, whereby a predefined number of clusters is provided and the algorithm strives to provide k clusters with maximal inter-cluster variation and minimal intra-cluster variation. A natural extension to this in the medical imaging domain is fuzzy k-means clustering, where the segmentation is 'soft', in that it provides likely tissue classifications given an unseen image. This suits the medical imaging domain where there is inherent ambiguity in medical imaging acquisition leading to partial voluming effects [65]. Unsupervised neural networks have also been used in medical image segmentation, where the algorithm attempts to determine the correct weighting of internal feedback structures, given an input image. With respect to supervised clustering methods, techniques such as k-nearest neighbour classifiers and support vector machines are popular methods. The k-NN method trains a set of feature vectors and associated classifications from training samples. Then, given an unseen image, the k-NN evaluates the 'k' nearest neighbours and carries out a voting process to determine the most probable classification for that pixel.

Like thresholding methods, most clustering approaches will lead to discontinuities in noisy medical images, which exhibit incomplete boundary definitions. Therefore, the application of clustering techniques depends on the quality of the image, or indeed the quality of a pre-filtering stage before classification is carried out.

Deformable Models

Deformable models are a general categorisation for a set of closely related segmentation techniques. Some ambiguity exists in the literature regarding the exact grouping of methods defined by this term. In this review, deformable models refer to any free-form, or partially constrained, technique that relies on a set of internal and external forces to control the evolution of the solution contour. The use of deformable models originated with the seminal paper on active contours, or snakes, by Kass *et al.* [47]. Given initial parameters, the active contour evolves towards the segmentation target based on a combination of internal and external forces. The internal forces maintain physical properties of the contour, such as length and smoothness. External forces define the minimum energy level that the active contour will

attempt to reach. These forces are usually constructed in terms of pixel intensities or gradient information, and this allows very basic *a priori* information to be embedded into the segmentation process. Finally, due to the explicit nature of the contour definition in the active contour method, a regularization step must be included into the segmentation process so that control-points do not overlap and contour self-intersection does not occur.

After their introduction, Terzopoulos *et al.* [102] extended the original snake formulation to 3D, and named the method active surfaces. Both active contours and active surfaces achieved widespread use in medical image segmentation, illustrated by applications in the cardiac [71], pelvic region [28], neurological [22] and bone [39] domains.

However, the deformable model approach has several limitations. First, the technique is sensitive to initialisation parameters. Without quite accurate localisation of the segmentation target, the deformable model's external energy terms can become negatively influenced by spurious image gradients. This local minima issue during curve evolution precludes the method being used in noisy imaging environments or where object localisation is non-trivial. To address this shortcoming, Xu and Prince [111] presented a novel extension to the original snakes method, called the gradient vector field (GVF) snake. Here an image-wide gradient vector field is created before the snake is initialised. This GVF field prevents the snake from becoming stuck in local minima and addresses to some extent the issues related to over-sensitivity to initialisation parameters and image noise. Cohen [21] presented an alternate solution, whereby an extra ballooning force is added to the external energy. This extra term disturbs the contour when it reaches local minima by expanding or contracting the curve. This requires prior knowledge regarding whether the contour is inside or outside the target structure. Both the balloon force and GVF methods achieved significant adoption in the research community.

Based on the work of Xu and Prince, Santarelli *et al.* [90] proposed a 2D GVF snake technique to segment the endo- and epicardium in MRI images. Their method requires pre-filtering to maximise region homogeneity and their snake must be manually drawn inside the LV blood pool for initialisation. The epicardium is detected based on the segmentation of the

endocardium, but no *a priori* knowledge is enforced to constrain the solution. No spatial or temporal constraints are enforced for multidimensional datasets - instead the authors rely on a simple sequential segmentation. The authors note that their method is likely to fail towards apical slices where the LV blood pool and myocardium interface shows low contrast. Berbari *et al.* [37] also base their work on the GVF snake for left ventricle MRI segmentation. Here, the authors focus on pre-filtering the CMR images in order to remove the papillary muscles before the application of the GVF-snake. The authors note that their method is still susceptible to areas of noise and in cases where the papillary muscle is joined to the myocardial wall. Montagnat and Delingette [71] develop a 3D+time left ventricle segmentation method, using a set of 3D deformable active surface models represented with discrete simplex meshes. The authors include spatial and temporal constraints to improve the robustness of the method. They do not provide quantitative performance characteristics when applying their method to real 3D+time MRI data, but mention their method must compensate for images exhibiting low contrast with increasing the model rigidity parameter.

Another notable direction of extensions to the deformable model is regarding the representation of the model. The introduction of deformable surfaces brought with it a wide range of representations, which can be broadly classified as either discrete or continuous. In the discrete space, simplex meshes, spring-mass models and particle systems have been implemented. Continuous models are either explicit (as in the original formulation by Kass) or implicit, leading to more complex superquadratic, hyperquadratic and the well-known level-set techniques.

The most striking extension in the implicit representation space is that of level-sets. The level-set method, introduced by Osher and Sethian [79], solves many of the limitations related with explicitly defined deformable models. First, level-sets can account for topology changes during segmentation, unlike active contours, which must be initialised with the same topology as the segmentation target. The more robust mathematical basis of level-sets allows the segmentation of colour, texture and affords motion tracking, all of which are difficult to incorporate into the explicit deformable model.

Level-sets evolve a deformable curve in order to segment a target image,

based on an energy functional. When the segmentation is complete the level-set is at its minimal energy level. The desired energy level is controlled by the stopping term, which, in initial level-set applications, was defined purely in terms of image intensities and gradient information. Although more robust than active contours, image noise can still cause contour evolution to become stuck in local minima and areas of low, or absent, contrast on the border of the segmentation target can cause spilling of the solution curve.

To address this, Chan and Vese [18] introduced the region-based stopping term based on the Mumford-Shah energy model. This term aims to partition the image into homogeneous regions and has been shown to be more robust with respect to initialisation and spilling. This methodology has been incorporated into many medical image segmentation applications [56, 105]. Paragios [80] unifies both the edge-based and regional stopping terms to segment the endo- and epicardium using coupled level-sets. Shape priors are used to further constrain curve evolution.

The original level-set method is computationally expensive since contour evolution is determined by the entire image region. This becomes prohibitive in 3D and 3D+time segmentation problems. To address this practical limitation, the Narrow-band method was introduced by Adalsteinsson and Sethian [2] that calculates contour evolution based on a smaller section of the image information surrounding the contour. The Fast marching method, first described by Sethian [91], solves the myopic problem with initialisation that affects explicitly deformable models. Fast marching provides a fast initial approximate segmentation that is commonly used to initialise a full level-set evolution procedure.

Instead of relying purely on image data, significant research effort has been directed towards the inclusion of shape priors into the level-set process. Inclusion of shape priors allows an extra level of robustness, and is especially useful to prevent erroneous spilling during contour evolution in areas of low contrast. Leventon *et al.* [53] were the first to introduce shape priors into the level-set formulation. The authors applied PCA to a set of prior shapes and included an extra term into the evolution equation that encourages the contour towards the most likely shape, based on the image information. Tsai *et al.* [104] also employed a similar PCA-based approach. However,

the linearity assumptions of PCA are not strictly valid, with respect to the level-set formulation. Other shape representations have also been proposed in the literature, such as: embedding the harmonic description of a set of training shapes, as well as parametric and non-parametric probability maps of likely shape configurations.

The level-set method has been widely adopted as a robust segmentation method in the data-driven community with applications in: cardiac [63, 64, 10], brain and etc. Like Paragios [80], Lynch *et al.* [63] also presented a coupled approach to myocardium segmentation and incorporate both gradient and texture information into their stopping term. The coupling term is introduced so that the epicardium level-set is dependant on the endocardium. The evolution of both contours is further constrained with the inclusion of prior shape knowledge through PDFs. This work was later extended [64] to achieve 3D+time segmentation of the myocardium. The main contribution in this work is the inclusion of an *a priori* non-rigid temporal model, which encourages the level-set evolution towards learned volumetric deformation behaviour. More recently, Ayed *et al.* [10] proposes a promising method whereby overlap priors are used to constrain a left ventricle segmentation method. The authors do not depend on a training set, and so their method is more adaptable in pathological cases. Although their method shows robust experimental performance when evaluated on 10 MRI datasets, their method relies purely on image intensity information. In cases with low contrast, or where the left ventricle is affected by image artefacts due to patient motion, their method would be more likely to result in erroneous segmentations than a model-based method. A full review of modern level-set applications can be found in Suri *et al.* [101] and Cremers *et al.* [30].

The central limitation with deformable models, in general, is the difficulty in including *a priori* information in the segmentation process. Consequently, the deformable model can assume any smooth shape that fulfills the energy functional, even though the solution curve may be biologically invalid. In the case of level-set methods, although inclusion of shape priors is more feasible, there is still no guarantee that biologically-valid segmentations will occur. This, however, is a double edged sword. Compared to, say, the Active Shape Model technique where the range of synthesizable shapes is totally constrained by the training set, level-sets have no such limitation. While

this is a useful property in segmentation scenarios where the target object exhibits significant variation, it can also be a cause of error when the target object is poorly defined. This is a central differentiating factor between data-driven methods, like level-sets, and model-based techniques, like the Active Shape Model. In this thesis, which focuses on multi-dimensional cardiac left ventricle segmentation, it was determined that the segmentation technique employed should focus more on the level of control and constraint, due to the segmentation challenges posed by multi-dimensional cardiac MRI data. As such, it was determined that (despite valid arguments for both approaches) assuming a model-based strategy is more suitable.

2.3.3 Model-based Segmentation Techniques

Unlike deformable models, model-based segmentation techniques rely on manual annotations to construct models based on *a priori* knowledge. An example of this is atlas-based techniques. Here, shape and texture are captured from training data. Variation is learned through either: probability density functions (PDFs) [19], probabilistic atlases [41, 60, 81], non-parametric probability distributions [62], or by using point distribution models (PDMs) [88, 62]. The use of PDFs usually assumes Gaussian distributions of landmark and texture variation, and, as such, are computationally efficient. However, when the Gaussian assumption is not valid, non-parametric models or Gaussian mixture models [20] are employed. Unlike PDFs and PDMs, probabilistic atlases create one surface map and describe the likelihood of a shape given a trained atlas.

Atlas-based methods treat segmentation as a registration problem, where a mapping must be found between the atlas and the image being segmented. Bayesian techniques are a natural fit to estimation of probabilistic atlases in this manner, either relying on the maximum *a priori* (MAP) [81, 96] criteria, or similar. The Expectation Maximisation algorithm integrates well with this approach [41, 60, 87], where the image signal acts as the observed signal and the atlas parameters as the unknown data. It is common that a regularization step is required so that tissue classification is encouraged to form homogeneous units. Markov Random Fields have generally been used to achieve this purpose [60, 81], to encourage adjacent pixels or voxels to

share a common classification.

The atlas-based technique has achieved significant adoption by the imaging community, with applications in neurological [41, 86, 88], cardiac [60, 61, 62, 83, 114], lung [94] and abdomen [81] imaging.

The difficulty in creating 3D atlases of biological structures is illustrated in the work of Frangi *et al.* [38]. The authors dedicate the entire paper to the problem of automatically constructing an atlas of the left and right cardiac ventricles. The central contribution involves the automatic extraction of landmarks from an atlas, and their propagation, via a non-rigid registration algorithm, to a set of training samples. Lötjönen *et al.* [62] extend this approach and create atlases of the left and right ventricles and atria from short and long axis images. Lorenzo-Valdés *et al.* [60] present the only (to this author's knowledge) 3D+time left and right cardiac ventricle atlas in the literature. The authors construct probabilistic atlases of the ventricles as well as the background structures. A 3D intensity atlas is also created in order to align the atlas with an unseen dataset. Gaussian distributions are assumed and an extended 3D+time EM algorithm is employed to estimate the atlas parameters based on observed image data. Evaluation is only carried out on a limited number of datasets in a leave-out-scenario. The authors note that they must blur their atlas in order to account for myocardium hypertrophy (i.e. excessive thickening).

Atlas-based segmentation is a powerful technique but prone to some limitations. Although the atlas is constructed from manual segmentations, there is no guarantee that the resulting segmentation is a biologically valid shape. The number of model parameters and the dimensionality of the problem directly affect the computational complexity of registering the atlas to an image. Atlases can also be prone to local minima during segmentation, especially at higher dimensionality. While this method can be useful for constructing patient-specific atlases, extending this approach over, say, the left ventricle variability of a population is somewhat impractical. The result is that the atlas will be somewhat similar to the deformable model, as there is so much variation captured in the atlas that any smooth shape that approximates the target structure will be deemed valid. And, if pathologies are included in the modelling problem this issue is dramatically exacerbated.

Another related model-based approach (which, for comparison purposes, can be described as parametrically deformable atlases) is the statistical shape modelling technique. Like atlas-based methods, shape and texture variation is learned from a set of training samples. However, shapes synthesised using statistical models are guaranteed to be biologically valid as they are constrained within limits seen in training. This is of great benefit to segmentation accuracy, and statistical methods in general have been shown to perform well in noisy images where the segmentation target can be only partially defined. This unique property also makes them more suited to the construction of general anatomical models for use in medical image segmentation applications.

Statistical Models

The Active Shape Model (ASM), developed by Cootes *et al.* [24], was first applied to the segmentation of facial images. The ASM is based on learning statistical shape variation from a set of training samples. Local grey-scale information around each landmark is also modelled and used to assist in fitting the ASM to an unseen image. The flexibility of the ASM methodology is reflected in the number of extensions presented in the literature since its inception. The technique has been extended to 2D+time [6, 17] and 3D [110, 43]. The ASM-search phase has been replaced with more sophisticated methods, such as fuzzy inference methods [109]. Principal Component Analysis (PCA) has been replaced with Independent Component Analysis (ICA) [106] and Kernel PCA (kPCA) [17] in cases where the linearity of PCA is not suited to the application. Multi-scale segmentation approaches were proposed to address the ASM's sensitivity to initialisation. In fact, there have been so many alterations to the original formulation that the ASM can be thought of as a general segmentation approach, rather than strictly defined method in itself. As with all statistically-based techniques, the ASM is dependant of the quality and quantity of its training data - without this the model will be overly constrained and lack useful levels of synthesizable shape variation.

The Active Appearance Model (AAM) is a direct extension of the ASM that learns regional texture information as well as shape variation in the training

process. As a result, the AAM also has a more sophisticated optimisation process. Like the ASM, the AAM has been widely adopted in image segmentation applications. It is reported to be more robust than the ASM during fitting, however it also exhibits slower optimisation time, and is still prone to becoming stuck in local minima. Moreover, when creating 3D AAMs there have been some reports [98] that the texture component of the AAM has large memory requirements, and can make a practical implementation challenging. In this thesis, the ASM was chosen instead of the AAM for cardiac left ventricle segmentation, as it was determined that the myocardium and blood pool of the left ventricle exhibit minimal texture information, as they are both mainly homogeneous. Moreover, it has been noted that training a model directly on grey scale information limits the method to that imaging modality, reducing the generalisability of the model. Finally, MRI image acquisition errors can cause undesirable effects, such as smudging or spilling. Since the ASM is used in this thesis, no pre-processing stages are required to correct for such errors, as would be the case if a texture-based method like the AAM was employed. This is an advantage of using the approach described in this thesis.

2.3.4 Notable Model-based Approaches

As image acquisition technology progresses, datasets are now commonly presented in 3D or 3D+time. This has led to a corresponding drive to create model-based methods to model the segmentation target in multiple dimensions. The advancement of model-based methods from 2D to 3D has been largely successful, and resulted in a wide range of notable research works.

The vast majority of model-based approaches in this area have been developed using a *top-down* modelling strategy. In this approach, all of the desired modes of variation of the segmentation target is captured in one model, so that the model dimensionality matches that of the segmentation target. For clarity, this is referred to as the *single-model* strategy, as only one model is used to capture the desired variation. With respect to cardiac segmentation, the work by Mitchell *et al.* [69] extended the standard 2D Active Appearance Model (AAM) - first introduced by Cootes *et al.* [23] -

to a 3D implementation. Using a similar approach Bosch *et al.* [12] model temporal deformation in a 2D+time AAM-based segmentation of echocardiographic images. Kaus *et al.* [48], describe a 3D deformable model for LV myocardium segmentation. The authors present promising results but exclude datasets exhibiting large motion artefacts from their evaluation. This issue was later addressed by Stegmann *et al.* [97, 98] who use an unsupervised motion-correction scheme before segmentation. The authors concatenate the end-systolic and end-diastolic cardiac phases of the LV into a single AAM in an attempt to include temporal variation in their model. This technique is also used by Fritz *et al.* [40] who extended it further to consider the 3D surfaces of the left and right ventricles at both the end-systolic and end-diastolic cardiac phases as a single shape. The authors argue that if, for example, the right ventricle is poorly defined compared to the LV, then the fitting process will be guided by the stronger image feature, thus providing a more robust overall segmentation. However, they do not indicate whether they weight the landmark fitting process accordingly or if they rely purely on the ASM approximation of the landmarks to correctly guide the segmentation process. Another 3D-ASM LV segmentation technique is described by van Assen *et al.* [108] - also used in [110] and later developed in [109] - where the authors replace the standard landmark fitting stage in the ASM formulation with a fuzzy inference system. The authors report segmentation accuracy comparable to the state-of-the-art implementations evaluated in their study. Finally, to alleviate the requirement of comprehensive manual annotation of training samples, several papers have explored different approaches for the automatic generation of landmarks and point-correspondence [38, 88]. For a complete review of the application of model-based segmentation approaches see the recent work of Heimann *et al.* [44].

While largely successful as the research community transitioned from 2D to 3D segmentation, the *single-model* strategy is not easily generalisable to higher dimensional problems. This is reflected both by the notable lack of *single-model* 3D+time research and the criticisms recently expressed in a number of separate works [6, 46, 112, 34]. This raises the question regarding the specific problems that limit the adoption of model-based methods in 3D+time segmentation problems, and why those problems did not manifest in the 3D case. This issue will be dealt with in the following section of this

chapter.

2.3.5 Limiting Factors in 3D+time Model-based Segmentation

A model is only as good as its training set. Therefore, the quality, quantity and variation of these training samples are of paramount importance. The reality is that modelling complex 3D+time biological deformation requires a very large training database of manually segmented images. Such databases are simply not available. Without meeting the training requirements, models will not encompasses sufficient shape variation to be of utility in a general segmentation environment. This is one of the primary causes of the limited appearance of 3D+time model-based segmentation techniques in the medical image segmentation literature. In fact, the only notable *single-model* 3D+time work presented in the literature, by Lorenzo-Valdés *et al.* [60], is heavily reliant on pre-segmented data. As indicated earlier, the authors' technique requires not only the left and right ventricles but also the background structures to be pre-annotated to build their 3D+time atlas. Naturally, their training data exhibits insufficient variation and Gaussian blurring is used to artificially enlarge model variation. Koikkalainen *et al.* [50] also present a method for artificial enlargement of the variation in small training sets for construction of a 3D ASM of the four chambers of the heart. More recently, Zhang *et al.* [113] argue against the artificial enlargement of the training set to maintain structural consistency in the resulting models. Here, the authors extend the work of Mitchell *et al.* [70] using a hybrid ASM/AAM implementation, training both a 3D and a 3D+time model. The authors concede that they do not have enough training images to capture sufficient temporal variation in their 3D+time model, and so their 3D model must be used for a phase-by-phase final segmentation stage in order to avoid the over-constrained temporal variation in the 3D+time model. These works [113, 60] show the limitations, most notably with respect to training requirements, with the *single-model* approach for 3D+time segmentation.

The increased training requirement, however, is not the only difficulty facing 3D+time modelling. Registration between training samples also becomes very challenging. In fact, due to its challenging nature, registration is a

research domain in its own right. Like the training requirements, the difficulty in developing point-to-point correspondence between training samples is proportional to the dimensionality of the segmentation target.

Furthermore, as the dimensionality of the model rises so too does the number of model parameters. This can cause detrimental effects in model convergence, time to complete a segmentation, and can be prone to local minima. Manual initialisation, if required, can also be cumbersome and slow at such dimensions. The combination of these inter-related issues are commonly described as the 'curse of dimensionality' (CoD) in the literature [6]. It should be noted that all these issues were present in the 3D modelling case, but manifested to a lesser degree, and so were not addressed in literature, apart from some minor discussions [51].

The CoD precludes any practical 3D+time model-based method from being implemented without first overcoming considerable barriers. A survey of the recent literature shows that efforts have been made to address each issue related to CoD. Some papers present automatic landmarking techniques to alleviate the training requirements, while others present artificial enlargement of allowed model variation.

2.3.6 Related Research

A recent trend in the literature [6, 112, 34] has gone some way in reassessing the approach to higher dimensional modelling and segmentation problems. Instead of creating a single model, subsections of the overall variation are captured in separate models, which are then combined at segmentation-time. In this way, the full variation of the LV is still captured, but the dimensionality of each model is lower than an equivalent *single-model* approach and therefore has less training requirements, simpler aligning and is better able to encapsulate the statistical variation of the training set. Zambal *et al.* [112] describe a 3D left ventricle segmentation scheme, where the LV is modelled using a set of 2D Active Appearance Models (AAMs), each capturing a specific section of the ventricle. These models are interconnected with a separate 3D shape model that controls the global positioning and scale during segmentation. The authors report that their technique outperforms a standard 3D AAM by 11% in terms of segmentation accuracy. More impor-

tantly, the authors show that their method inherently copes with irregular translation between slices since each AAM can move freely in its own slice, while structural consistency is maintained using a separate 3D model. This is a benefit in cardiac segmentation where adjacent slices are often misaligned and this negates the need for a pre-aligning phase as required by the approach proposed by Stegmann *et al.* [97]. Andreopoulos *et al.* [6] recently presented a 3D+time cardiac segmentation method that relies on two components. The first captures 3D shape variation of the LV in a 3D AAM. The second component encompasses the temporal variation of the LV in a 2D+time ASM. The 2D+time ASM is used to refine the initial 3D AAM segmentation. Like Zambal *et al.* [112], the authors argue strongly against creating single 3D+time models due to the issues relating to the CoD. While their work does decompose the 3D+time problem into spatial and temporal models, the authors still build a *single-model* 3D AAM. Therefore, they are still faced with significant training requirements (and related issues) albeit to a lesser degree than Lorenzo-Valdés *et al.* [60]. To address this, the authors extend the wavelet-based hierarchical ASM presented by Davatzikos *et al.* [33] to a 3D implementation, a method shown to perform well with limited training samples.

2.3.7 Thesis Context

This thesis explores an alternate modelling strategy, whereby the segmentation target is decomposed into its constituent parts at model-time, which significantly simplifies the modelling process and returns 3D+time modelling and segmentation into a tractable problem domain. Then, at segmentation-time the separate models are combined, unifying the different sources of variation to use all the *a priori* information available. This thesis investigates this approach when applied to 3D+time cardiac left ventricle segmentation, as it exhibits complex spatial and temporal dynamics as well as MRI images being challenging in themselves. This image segmentation scenario is a true test of the validity, robustness and practicality of the proposed alternate modelling strategy. Furthermore, this general approach is also applied to the challenge of modelling multi-contour structures. Critical analysis is made regarding the implicit assumptions and consequences of considering a multi-contour structure as a single shape at model-time.

By training separate shape models, while enforcing structural consistency through biologically-valid constraints at segmentation-time, it is experimentally shown through myocardium segmentation that the proposed modelling approach yields superior results and provides enhanced flexibility compared to the combined modelling approach.

2.4 Summary

This chapter provides the background information to give context to the contributions presented in the remainder of this thesis. A description of the heart is given, providing anatomical, functional and pathological information. The main imaging modalities are described, with emphasis given to MRI imaging. The challenges associated with MRI are detailed to highlight the difficulties in creating robust 3D+time model-based segmentation algorithms. A directed literature review is then provided. This gives an overview of the main research trends seen today in the literature, as related to the research presented in this thesis. The major model-based methods are presented and described. This section concludes, based on the evidence of the literature review, with the most suited method for multi-dimensional cardiac left ventricle segmentation that will be examined in the remainder of this thesis.

Chapter 3

Active Shape Model

The Active Shape Model (ASM), first developed by Cootes *et al.* [24, 26], is a popular model-based segmentation technique. The technique can be decomposed into two separate parts: the model construction phase, and the model fitting phase. In model construction, the main modes of variation are extracted from a set of annotated training samples. During fitting, the ASM synthesizes shapes that are within a learned range of variation, defined by the set of training samples. The constrained nature of the ASM is its primary benefit over more data-driven methods, such as the active contour. In fact, the ASM was initially presented as a 'smart snake'. First applied to facial image segmentation, the ASM has enjoyed widespread adoption by the image segmentation community. Specifically in medical image segmentation, the ASM has been shown to be well suited for capturing biological shape variation from manual annotations. This allows *a priori* biological knowledge to be compactly captured, and guarantees synthesized shapes will be biologically valid, since they will be inferred from those seen in the training set.

The ASM formulation is quite flexible and has seen several significant extensions and alterations since its inception. The most notable is, of course, the AAM, also developed by Cootes *et al.* [23] that includes texture that is internal to the shape into the model. The ASM itself has been extended to 3D [103] and 2D+time [23] problems. its fitting algorithm has been replaced by other methods. Van Assen *et al.* [109] replace the ASM search

phase with a fuzzy inference system for cardiac left ventricle segmentation. Recently, Cristinacce and Cootes [31] create a hybrid between the ASM and AAM. In their approach, local texture patches are trained for each landmark (as opposed to the entire internal region for the AAM). The authors also use an optimisation method for shape and texture parameters, and present promising results.

Research has been exerted to reduce the effect of outliers in the fitting process [36]. The standard method to extract shape variation, Principal Component Analysis (PCA), has been replaced with Independent Component Analysis [100] (ICA) and kernel PCA [23] (kPCA) in the literature where the linear assumptions made by PCA are not valid.

The range of these extensions, therefore, show that the ASM is versatile and flexible, as opposed to a strictly defined method. In this chapter, the standard ASM formulation is presented, detailing the training and fitting stages as described by Cootes and Taylor [25]. Next, the alterations made to the standard method in this thesis are described and justification provided. Then, the limitations of the ASM fitting stage are described in detail, which drive the first contribution of this thesis: a novel ASM optimisation method. The proposed technique is experimentally compared and contrasted with the standard fitting method. Quantitative results illustrate that the proposed optimisation technique offers superior segmentation accuracy in complex medical imaging scenarios. Throughout this and following chapters, the presented contributions are evaluated on cardiac left ventricle MRI data.

3.1 ASM Training

The ASM is created from a set of training samples. The target object is usually described by a set of landmarks that are placed either at 'feature locations' (for example at corners, or areas of high curvature) or at regular intervals so that the contour can easily be described by fitting a spline or even a simple polygon to these landmarks. Therefore, a shape, $\mathbf{s}$, is described by a set of discrete landmarks according to Equation 3.1, where l_i is the i^{th} landmark describing the shape in the Euclidean space.

$$\mathbf{s} = [l_1, l_2, \ldots, l_i, \ldots, l_N] \tag{3.1}$$

It follows that the training set, $\mathbf{S}$, containing M shapes is defined as:

$$\mathbf{S} = [\mathbf{s_1}, \mathbf{s_2}, \ldots \mathbf{s_M}] \tag{3.2}$$

At this stage the set of training shapes must be aligned to remove scale, rotation and translation effects. Procrustes alignment is commonly employed for this task [25].

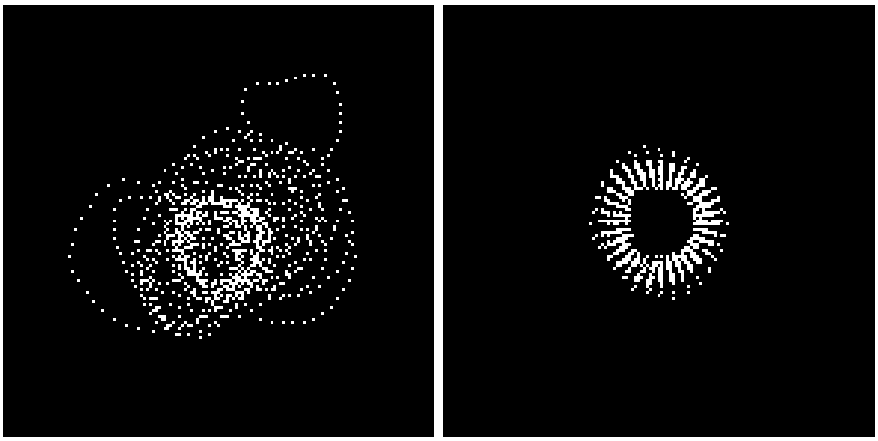


Figure 3.1: Example illustration of a set of unaligned endocardium training shapes (shown on the left) with translation removed and re-sampled to 32 equally-spaced landmarks (on the right), with scale included.

After the training set ($\mathbf{S}$) has been aligned ($\mathbf{S}'$) the ASM employs the Point Distribution Model (PDM) method to extract the modes of shape variation. If each shape in the training set is considered as a single point in Nd-D space (where N is the number of landmarks and d is the dimension of the shape, so that in this example $N = 32$ and $d = 2$), then the set of training shapes will form a point cloud in this space. In almost all practical applications, this will be a high-dimensional space. Principal Component Analysis is used to decompose this high dimensional description of shape via landmarking into a lower dimensional compact description. In essence, PCA extracts the main axes of variation from Nd-D space.

First,the mean of the training shapes is calculated according to Equation

(3.3). Using this, the covariance matrix, $\mathbf{C}$, can be derived using Equation (3.4).

$$\overline{\mathbf{s}} = \frac{1}{M}\sum_{i=1}^{M}\mathbf{s}_i \tag{3.3}$$

$$\mathbf{C} = \frac{1}{M-1}\sum_{i=1}^{M}(\mathbf{s}_i - \overline{\mathbf{s}})(\mathbf{s}_i - \overline{\mathbf{s}})^T \tag{3.4}$$

Next, the eigenvectors and eigenvalues of the covariance matrix are estimated (ϕ and λ, respectively). The eigenvectors are ordered from the vector corresponding to the largest eigenvalue to the smallest, where the largest represents the axis of most variance. The eigenvectors are, by definition, orthogonal, and coupled with their eigenvalues, and are directly used in the PDM formulation to synthesize new shapes for use in the ASM.

However, the set of eigenvectors are truncated before their employment. This allows the majority of shape variance to be captured with only a minimal number of parameters. The first R eigenvectors are kept so that the sum of their corresponding eigenvalues is a percentage (p_v) of the total variation. This is described in Eq (3.5).

$$\sum_{i=1}^{R}\lambda_i \geq p_v \sum \lambda_i \tag{3.5}$$

Usually p_v is chosen to be between 0.90 and 0.98 and, of the R eigenvectors that are kept, the first one or two are usually responsible for the majority of the variation. The truncated ordered set of eigenvectors is referred to as Φ. With the meanshape ($\overline{\mathbf{s}}$) and the truncated set of eigenvectors (Φ), new shapes can be synthesized by the well-known PDM equation:

$$\mathbf{s} \approx \overline{\mathbf{s}} + \Phi\,\beta \tag{3.6}$$

In Equation (3.6) the newly synthesized shape $\mathbf{s}$ is created from a linear weighted sum of the eigenvectors, where the β vector provides the relevant

weighting for each eigenvector. The β vector is the only variable in this equation, so, after the training process is complete, the range of synthesizable shapes is totally dependant on the value of β. As such, the control of β is the basis for synthesizing valid shapes using the ASM.

In the standard ASM formulation, shape synthesis using (3.6) is constrained according to:

$$\beta = \left\{ \beta_i, i = 1, \ldots, R \;\middle|\; |\beta_i| \leq \beta_i^{limit} \right\} \tag{3.7}$$

where $\beta_i^{limit} = 3\sqrt{\lambda_i}$ [25] and R are the number of truncated eigenvectors comprising Φ. β_{limit} is therefore the level of control placed on the allowable range of shape synthesis. However, this limit is not always adhered to - Koikkalainen *et al.* [50] greatly extend the allowed range of variation of the β vector, from the standard $\pm 3\sqrt{\lambda_i}$ to $\pm 5\sqrt{\lambda_i}$. This artificially enlarges the training set variation, which could have negative consequences with respect to generated biologically valid shapes.

However, the shape model is only one aspect of the ASM training. A set of landmark-specific gradient profile models also need to be trained, to assist in the fitting process. For each landmark, a symmetrical gradient profile (normal to the shape contour, see Fig. 3.2) k pixels either side of the landmark is recorded. This creates a set of M observations for each landmark in the shape. Each observation is normalised by dividing each element by the absolute sum of its gradient values. The j^{th} normalised observation for the i^{th} landmark is referred to as $\mathbf{g}_j^i$. Therefore, the i^{th} landmark will have set of observed gradient profiles from training, according to Equation (3.8).

$$\mathbf{G}_i = \left\{ \mathbf{g}_0^i, \mathbf{g}_1^i, \ldots, \mathbf{g}_j^i, \ldots, \mathbf{g}_M^i \right\} \tag{3.8}$$

Assuming a Gaussian distribution, $\mathbf{G}_i$ can be compactly described by its mean $\overline{\mathbf{g}}^i$ and covariance matrix $\mathbf{S}_g^i$. This process is repeated for each landmark, so that a model for each landmark is created, as in Equations (3.9) and (3.10) for mean and covariance matrices, respectively.

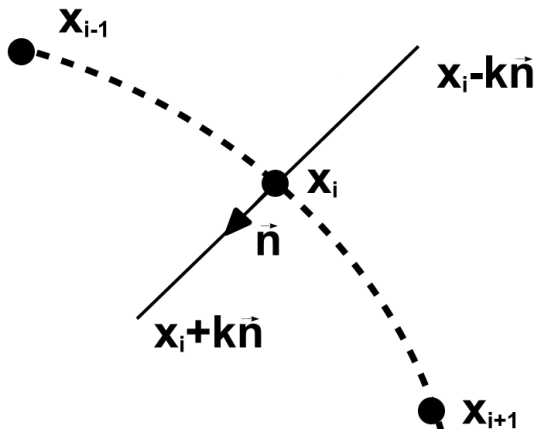


Figure 3.2: Illustration of the definition of an arbitrary landmark profile. The profile extends k pixels either side of the landmark so that each profile is $2k + 1$ pixels long. All profiles are calculated normal to the direction of the contour ($\vec{n}$).

$$\{\overline{\mathbf{g}}^0, \overline{\mathbf{g}}^1 \ldots, \overline{\mathbf{g}}^N\} \tag{3.9}$$

$$\{\mathbf{S}_g^0, \mathbf{S}_g^1, \ldots, \mathbf{S}_g^N\} \tag{3.10}$$

When fitting an ASM to a new image, as discussed in Section 3.2, these gradient profile models allow the quality of landmark fit to be determined. In many practical applications, however, the gradient model tends towards an ideal edge profile as the level of training increases (see Fig. 3.3). This is expected since contours of interest are usually defined by a strong image gradient. Therefore, when constructing gradient profile models in such a scenario, it may be possible to simply choose the strongest gradient location in the ASM search phase.

3.1.1 Alterations for LV Modelling

Section 3.1 describes the standard Cootes ASM modelling formulation. Like most other research, some alterations are made in this thesis to target the ASM to the application, in this case, left ventricle segmentation. With respect to the alignment phase, in this thesis, each training shape has an

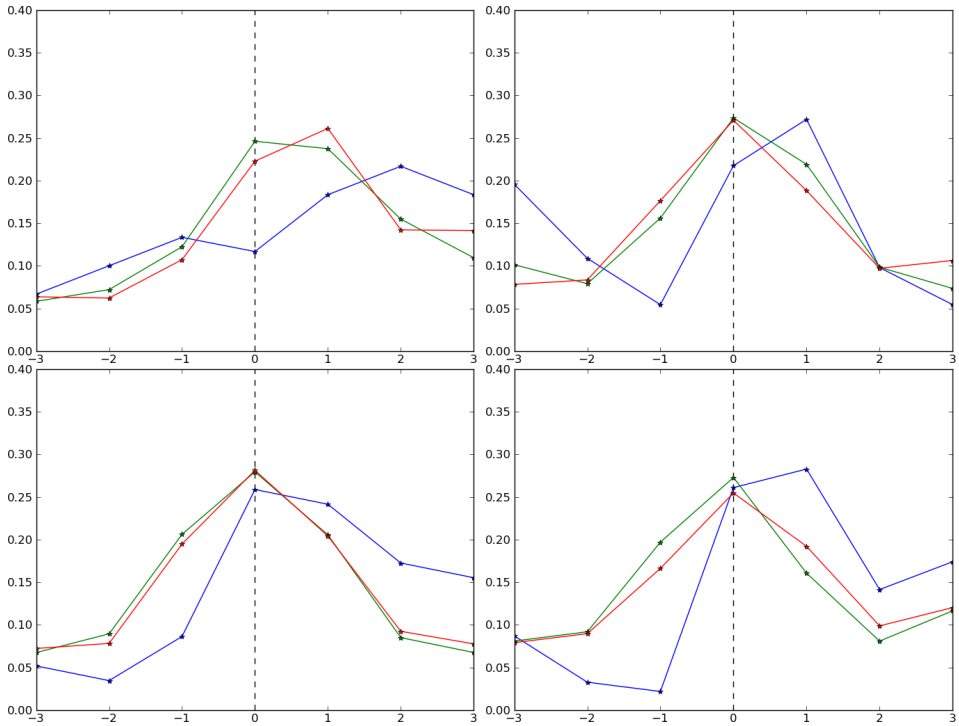


Figure 3.3: Four examples of cumulative edge profiles. In all sub-figures the blue, green and red lines correspond to the average profile over 1, 10 and 30 training shapes, respectively. As training increases in each case the learned gradient profile begins to approximate an ideal edge profile. The dotted vertical line indicates the centre of each profile. In each graph, the x-axis refers to the components of the profile, and the y-axis refers to the normalised gradient value at that profile location.

equal number of landmarks (32) that are evenly spaced (as illustrated in Fig. 3.1). As the contributions presented in this thesis are evaluated on cardiac left ventricle segmentation, some application specific alterations to the original alterations are implemented.

In this thesis, the scale variation is considered a valid part of the shape variation, and is thus not removed from the training shapes. In other research work in the literature, scale has also been considered part of the true shape variation [74, 54]. Indeed, when considering the temporal deformation of the left ventricle, one of the primary modes of variation is the scale change. With this in mind, scale does not have to be compensated for before PCA is applied to the training set. Therefore Fig. 3.1 is an accurate depiction of

a sample aligned training set used in this thesis.

In the general case rotation must also be removed, however, the data used in this thesis revealed that rotational effects are minimal relative to scale and shape variation, and as a result, point-correspondence is established directly with the arc-length re-sampling method (where landmarks are equally spaced) followed by the removal of translation. In summary, as there are an equal number of landmarks in each training shape and as scale is included and rotation is excluded from the training process, point-correspondence is established, in this thesis, by removing translation from each training shape.

Including scale into the ASM training stage, means that the first eigenvector/value in this scenario will normally account for a large quantity of the overall variation. This fact becomes useful in the novel optimisation method presented later in this chapter. To illustrate this fact, Fig. 3.4 illustrates two cases of training, one where scale included and one where it is excluded from the training process. The very strong first eigenvector/value corresponds to scale variation in the training set.

3.2 ASM Fitting

In the original formulation, the ASM is fitted to an unseen image according to an iterative two-stage process. As the ASM is model-based, some means of image interaction *must* be employed in order to deform the model to a target image. The use of 'smart snakes' in the title of the original ASM paper, indicated the direction Cootes *et al.* [26] took when implementing the ASM fitting process. Indeed, the mechanics of ASM fitting share many common properties with active contours: landmark search is carried out in the direction normal to the contour, image gradient information is used to determine likely landmark locations, the ASM is 'pulled' iteratively towards the global minimum, hopefully avoiding any local minima along the way and so on.

The specifics of the ASM searching stage are outlined by Cootes as follows. First, the meanshape is initialised on an image with an initial pose (i.e an initial scale, rotation and translation). Like active contours, the ASM was

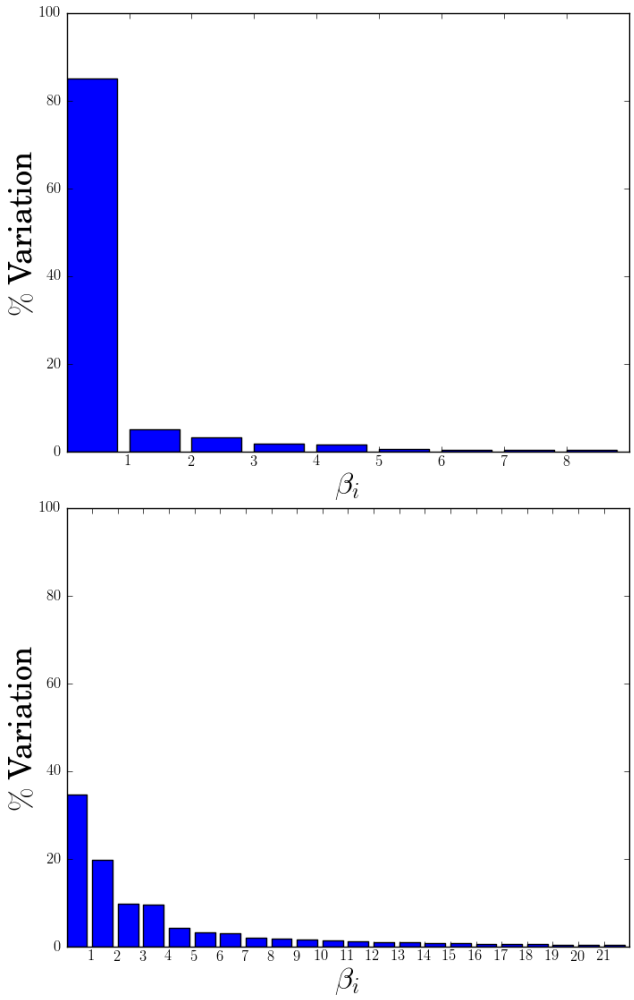


Figure 3.4: The upper figure shows the variation captured by each eigenvector when the training shapes do not have scale removed. The lower figure shows the variation captured by each eigenvector when scale is removed. In both cases 98% of training shape variation is kept.

initially designed as a semi-automatic method, although efforts have been made to make it fully automatic in application specific papers [27]. Then, each landmark in the ASM compares its learned gradient profile model, centred around each landmark on a 1D search profile, normal to the contour. The length of this search profile is pre-determined and is normally constant for all landmarks. The Malhabanois distance is used to estimate the quality of fit of each test location. A lower score represents a better quality fit between the model and the test profile. The location with the

minimal difference from the learned gradient profile model is chosen as the updated position for that landmark (see Fig. 3.6). Therefore, each landmark updates its location individually, using only very local image information and has no interaction with surrounding landmarks or the ASM itself (see Figs. 3.5,3.7). Equation (3.11) describes the scoring technique using the Malhabanois method, where $\mathbf{g}_t^i$ refers to the i^{th} landmark's test gradient profile and $\overline{\mathbf{g}}^i$ refers to the mean gradient profile model learned during ASM training.

$$M\left(\mathbf{g}_t^i\right) = \left(\mathbf{g}_t^i - \overline{\mathbf{g}}^i\right)^T \mathbf{S}_g^i \left(\mathbf{g}_t^i - \overline{\mathbf{g}}^i\right) \tag{3.11}$$

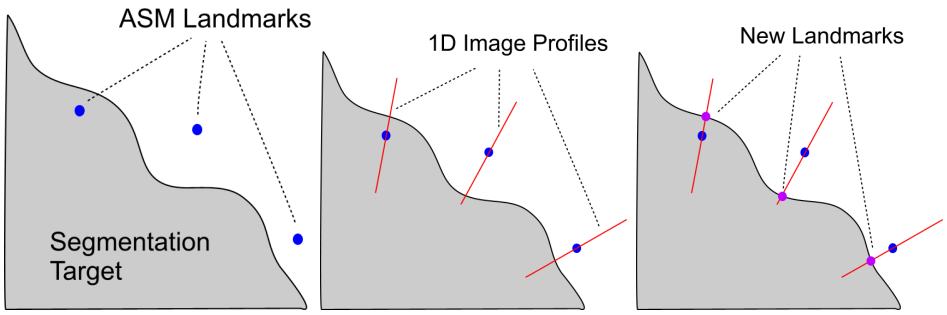


Figure 3.5: An overview of the progression of steps in the standard ASM landmark search phase. Given an initial set of contours (left) each landmark updates each location (right) based on 1D gradient profiles taken normal to the contour (centre).

This new set of landmarks (L') is then approximated by the ASM, according to Equations (3.12) and (3.13). This, in effect, coerces the set of landmarks to the closest shape within the realm of the shape model. Any outliers in L' can cause issues in the coercion process, and some research has been exerted to alleviate this issue [52, 36]. Much in the same way as the internal and external forces 'move' active contours to their minimum, the search-coerce process is the means by which the ASM adapts to a target image. The shape resulting from the coercion process, $\mathbf{s}'$, is used as the initialisation for the next search-coerce phase. The progression can then be illustrated in the following: $\overline{\mathbf{s}} \rightarrow L' \rightarrow \beta' \rightarrow \mathbf{s}' \rightarrow L'' \rightarrow \beta'' \rightarrow \mathbf{s}'' \rightarrow \ldots \rightarrow \mathbf{s}^{\mathbf{c}}$, where $\mathbf{s}^{\mathbf{c}}$ is the final converged segmentation. Part of this progression is depicted in Fig. 3.7.

$$\beta' \approx \Phi^T \left(L' - \bar{\mathbf{s}} \right) \tag{3.12}$$

$$\mathbf{s}' \approx \bar{\mathbf{s}} + \Phi \, \beta' \tag{3.13}$$

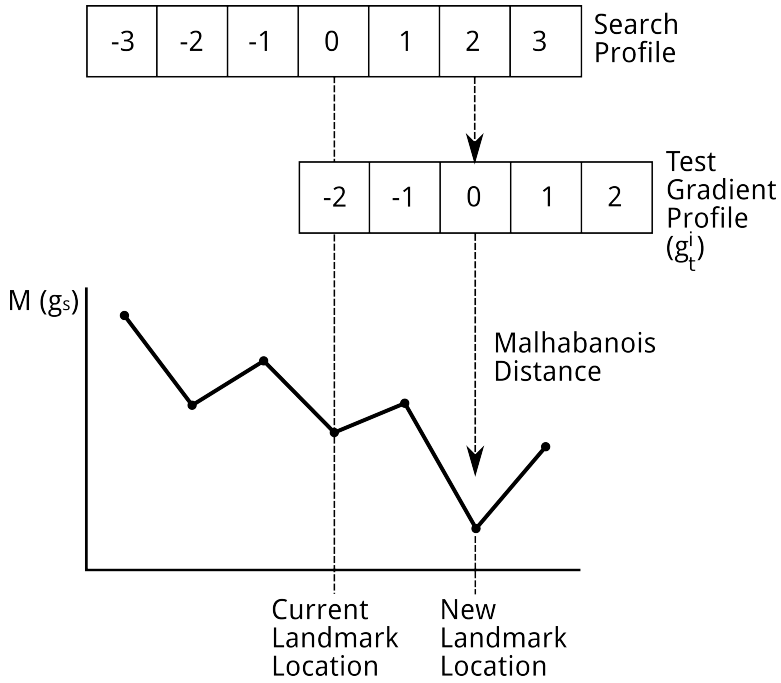


Figure 3.6: A schematic representation of the search phase in the standard ASM search phase. For each test location in the search profile the learned gradient profile is used to determine the quality of fit at that point. The location with the lowest score (i.e. the highest similarity to the gradient model) is chosen at the new landmark location and forms part of L'.

3.3 Limitations

Although the shape modelling aspect of the standard ASM formulation is useful for capturing statistical shape variation, the search phase of the ASM tends to exhibit significant limitations. When applied to practical complex imaging scenarios, such as cardiac left ventricle segmentation, the simplistic 1D search profiles can be negatively influenced by noise, and other misleading image artefacts. However, this is not the only limitation. The ASM

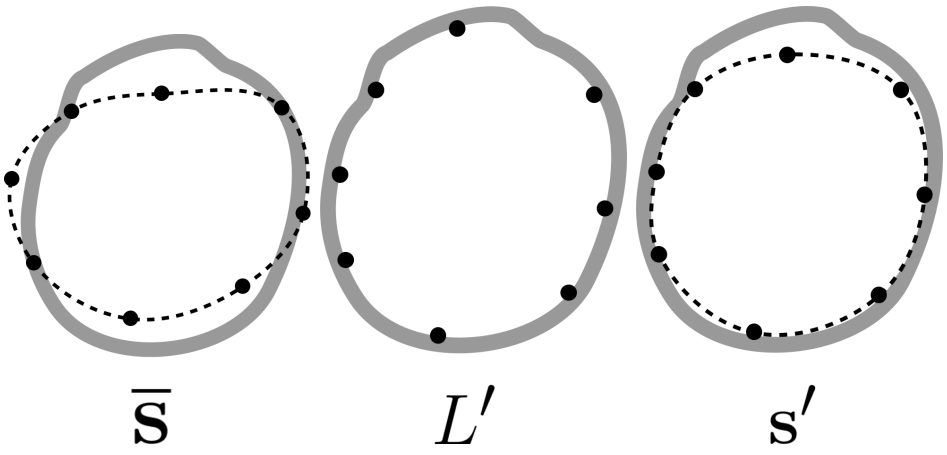


Figure 3.7: This figure shows a sample three stages of the standard ASM fitting process, beginning from the initialisation of the meanshape (left), $\overline{\mathbf{s}}$. The ASM landmark search phase (centre) produces the set of landmarks L' that are coerced by the ASM according to Equations (3.12) and (3.13) to produce the shape $\mathbf{s}'$ (right).

must be initialised with sufficient accuracy to be within the 'capture range' of the search profiles. Naturally, capture range is directly proportional to the length of the search profiles, however, longer search profiles are also susceptible to creating shapes (in the L' form) with significant outliers that can negatively influence the ASM coercion stage. A multi-resolution search technique was proposed by Cootes *et al.* [25] to alleviate this issue.

Furthermore, the 1D gradient profile models, which are constructed from the training set, are often insufficient in their description of local appearance around each landmark. The 1D profiles are the central weakness in ASM search phase, since they provide such limited interaction with the image data. There is a further limitation, which is detailed in the following chapter, regarding the application of the ASM technique to structures described with multiple contours.

3.4 ASM Optimisation

In order to address the limitations described above regarding the ASM search phase, a novel optimisation routine is introduced in this section. The op-

timisation technique is introduced with respect to the segmentation of the endocardium in left ventricle MRI images. In this thesis, it is felt that there is somewhat of a contradiction in the original ASM. The central benefit is that shapes are constrained within a learned range of variation. However, as highlighted earlier in this chapter, the ASM is fitted to an image on a landmark-specific basis, using only local image information to determine the likely landmark locations. In fact, there is usually a *dependence* on landmarks being drawn into unrealistic configurations while the ASM is 'pulled' towards the target image structure. Therefore, the 'search-coerce' progression appears to be somewhat ill-matched for the statistically constrained nature of the ASM.

In this thesis, the β vector is directly optimised during the ASM fitting stage, so that the search-phase *always* remains within known (and probable) limits of shape variation. Furthermore, the proposed method calculates the fitness score of a synthesized shape on the image information surrounding the *entire* contour, as opposed to limited 1D windows around each landmark. This in itself greatly enhances the robustness of the ASM search phase in noisy medical images.

According to Equation (3.6), the final shape $\mathbf{s^c}$ will always be some arrangement of the β vector and an associated pose. In the general case, assume there exists some fitness function f that, given an image I, a β vector (corresponding to the shape $\mathbf{s}_\beta$) and a pose $\mathbf{P}$ (comprised of scale, rotation and translation) will return a score describing the quality of fit of the shape, $\mathbf{s}_\beta$. In this scenario the ASM can be fitted to the image I according to the following optimisation routine:

$$\max_{|\beta| \leq \beta^{limit}} f\left(\beta, \mathbf{P}\right) \tag{3.14}$$

so that the shape score is maximised given an image. Equation (3.14) illustrates that the placement of a single constraint on the optimisation process elegantly maintains shape synthesis validity during the search phase. As stated earlier in this thesis, scale is included into the ASM itself (essentially scale is migrated from the pose vector into the first component of the β vector) and rotation is excluded, therefore when the ASM is applied to cardiac

data the general optimisation equation becomes:

$$\max_{|\beta| \leq \beta^{limit}} f(\beta, T) \tag{3.15}$$

To give a practical example, the objective function for the endocardium is detailed and its performance is characterised and compared with the standard ASM on real data. First, some descriptive notation must be defined. The objective function operates on the image information around the contour: defined as Ω_β^+ and Ω_β^- for the outer and inner regions, respectively. In practice, the boundaries of Ω^+ and Ω^- are defined as being three pixels outside and inside Ω, respectively. The limit of three pixels is chosen to provide a representative sample of the grey-scale values adjacent to a given Ω. This sample size also keeps the computational cost of evaluation of a contour score relatively low. This arrangement is graphically illustrated in Fig. 3.8. The function $I(\Omega)$ returns a vector of grey-scale values inside the image region Ω. It follows that $\overline{I}(\Omega)$ and $I^\sigma(\Omega)$ return the mean and standard deviation, respectively, of the grey-scale information in the region Ω.

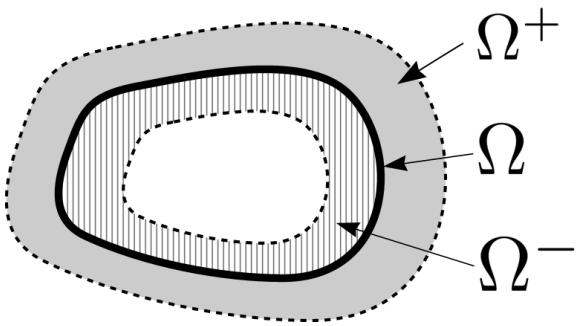


Figure 3.8: For any shape defining a region Ω, the proposed ASM optimisation method operates on the inner and outer regions of this contour, Ω^- and Ω^+, respectively.

The endocardium objective function is quite simple, since its contour is defined as the interface where the bright blood-filled left ventricular cavity meets the surrounding darker myocardium. And, as the objective function operates on the image information surrounding the entire contour, there is no need to add in any special constraints to deal with the influence of papillary

muscles. As indicated earlier, in the standard ASM search phase, each landmark is fitted based on a 1D profile. Naturally, the presence of papillary muscles can create misleading gradients that can be mis-interpreted as the endocardium boundary. This cannot happen, or is extremely unlikely to happen, in the proposed ASM optimisation technique because the fitting score is based on the entire contour. In the case of the endocardium objective function, any shape that tries to include the papillary muscle as part of the myocardium is likely to also have portions of the shape crossing the LV blood pool. Since there is no gradient information in the blood pool, a low contour score is generated, relative to the correct fit. The endocardium objective function (f_{endo}) is therefore defined as shown in Equation (3.16) and optimised according to Equation (3.17).

$$f_{endo}\left(\beta, T\right) = \left| \overline{I}\left(T\left(\Omega_{\beta}^{+}\right)\right) - \overline{I}\left(T\left(\Omega_{\beta}^{-}\right)\right) \right| \tag{3.16}$$

$$\max_{|\beta| \leq \beta^{limit}} f_{endo}\left(\beta, T\right) \tag{3.17}$$

Equation (3.17) states that the endocardium's energy is maximised when there is a maximum contrast between the inner and outer regions of Ω_{β}. This is expected since there is high contrast between the left ventricular blood pool and the surrounding darker myocardium. By operating on image regions, instead of individual 1D landmark profiles, the proposed optimisation method is not susceptible to outliers negatively influencing ASM convergence. This negates the need for outlier detection and removal schemes, as required in the standard ASM formulation [52].

Fig. 3.9 illustrates two cases where an image is associated with a search space created using the endocardium objective function described above. One can observe a shift in the peak in the search space (corresponding to the probable correct segmentation) between both cases.

Now, it can be seen that all shape possibilities can be considered by operating purely in the search space. This is in contrast to the standard ASM fitting method, where the initial contour must be 'pulled' towards the segmentation target, thereby exposing the method to local minima and negative influence

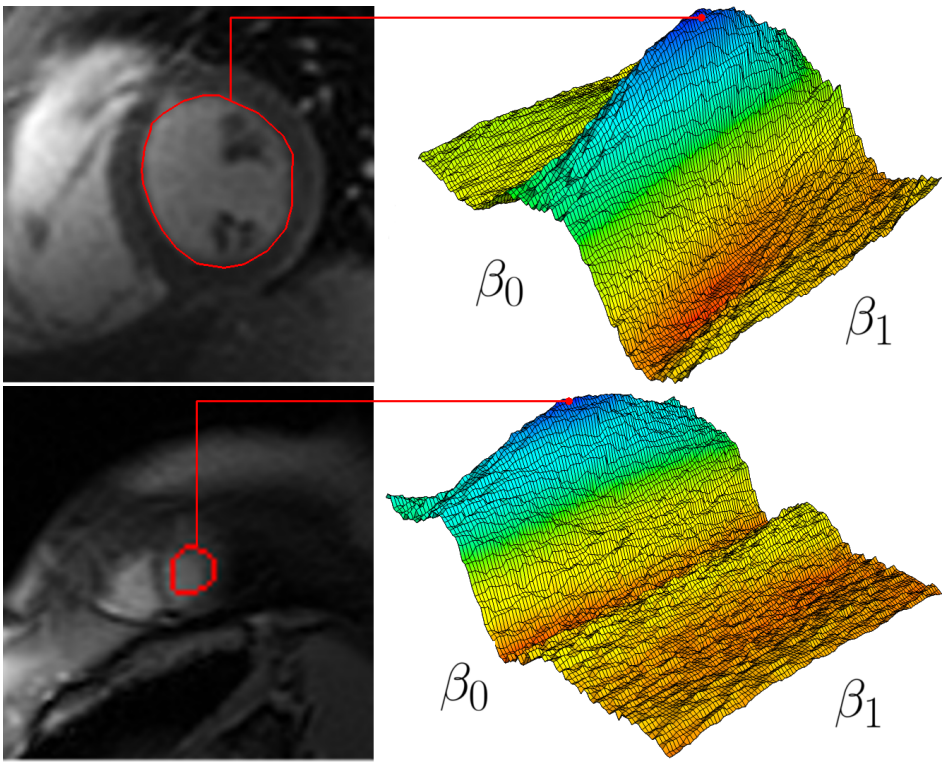


Figure 3.9: Two example endocardium search spaces using the objective function described in Equation (3.16). Blue peaks correspond to a better contour score. The search space is illustrated for the first two components of the beta vector. In both cases, ideal initial translation has been provided. The marked point in each search space corresponds to the contour drawn on the associated image.

from unrelated image features. The proposed optimisation approach holds no such limitation. In theory, all possible shape and pose configurations can be evaluated in a brute-force search. However, as the search space is quite smooth for cardiac left ventricle images (as depicted in Fig. 3.9), more targeted optimisation techniques can be employed to reduce the computational cost of fitting the ASM model to an image.

Since each β parameter is associated with an eigenvector, and since the eigenvectors Φ are ordered in terms of the percentage of variation they capture, the proposed method employs a coarse-to-fine optimisation scheme. Specifically, Brent's method [14] is used to maximise the contour score by optimising each β parameter sequentially, from largest eigenvector to small-

est. This process iterates until β converges, or until the difference between two successive iterations is below a small threshold.

The presented optimisation method is therefore applicable to any range of segmentation scenarios. The objective function is all that needs to be defined specific for an application. Although the objective function is pre-selected in this thesis, this step could be automated, as in the standard ASM implementation, by learning the expected region information from training samples. However, as noted earlier in this discussion (and illustrated in Fig.3.3) when the segmentation target is defined by a strong edge, as in the case of the endocardium, it is as, if not more, advantageous to create an 'ideal' objective function. In this way image noise, or poor training samples, cannot negatively influence the creation of the objective function. However in more complex scenarios where the definition of the segmentation target is not straightforward an objective function that is learned from training samples could be beneficial.

Very recently, Arámbula Cosío *et al.* [7] proposed a related method in the literature, that also directly optimises the shape parameters of the ASM. However the authors maintain the 1D gradient profile models for determining contour fitting. While this is robust enough for their presented application (hand segmentation) it can be argued that their method will exhibit similar weaknesses in noisy imaging environments, albeit to a lesser degree than the original ASM formulation. Moreover, the authors rely on the standard simplex optimisation routine to simultaneously optimise all the shape parameters, and thus do not use the sequential method proposed in this thesis.

3.5 Performance Evaluation

To highlight the difference in fitting accuracy a qualitative comparison between the standard Cootes ASM fitting formulation [26] and the proposed segmentation method is presented in Fig. 3.10. The Cootes fitting method (shown on the left hand side of Fig. 3.10) illustrates the disconnect between the landmark search phase and the corresponding ASM approximation. In the Cootes ASM image, one can observe several landmarks (shown as red

crosses) that have fitted to incorrect locations. Towards the bottom of the image a landmark fits to a papillary muscle, as it has a stronger gradient profile than the LV blood pool/myocardium interface. In the upper left, a landmark has found a gradient profile in the correct direction (from relative bright to dark) that is again stronger than the endocardium interface, even though is obviously in the wrong location relative to its neighbouring landmarks. These two examples show the weakness of the individual landmark fitting process in complex imaging scenarios, and how this negatively affects the resulting ASM approximation (shown as the continuous green contour). Naturally, this situation could be somewhat corrected by adding extra constraints to the standard ASM fitting process to prevent outliers and large relative differences between neighbouring landmarks. However, this does not address the underlying weakness, and based on the preceding discussion, it is clear that the proposed optimisation method (shown on the right in Fig. 3.10) is more suited to cardiac segmentation since it is not influenced by local features such as the papillary muscles, or locally noisy areas.

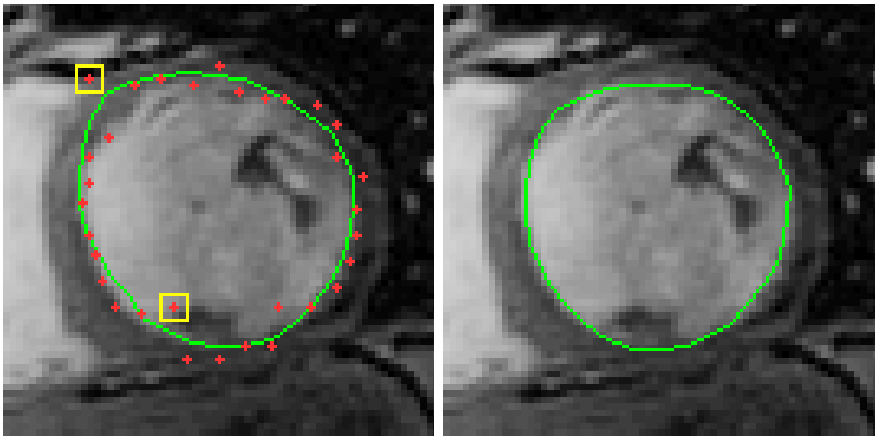


Figure 3.10: On the left is the standard Cootes ASM fitting formulation, with landmarks resulting from the ASM search phase shown as red crosses (problematic landmarks are highlighted with yellow boxes) and the corresponding ASM approximation shown as a continuous green contour. On the right is the result of the proposed ASM optimisation technique applied to the same image with the same initialisation parameters.

Next a quantitative comparison between the standard ASM segmentation accuracy and the proposed ASM optimisation method when applied to 2D endocardium segmentation is provided. Both ASM implementations are

applied to short-axis slices of the end-diastolic frame of 33 3D+time MRI datasets. In these datasets, the xy pixel resolution ranges between 0.93-1.64 mm. The effect of pixel resolution on segmentation performance is not evaluated in this thesis and is an artefact of dataset acquisition. In each volume the basal-most, apical-most and a mid-level slice are segmented by both ASMs. To determine the accuracy of the fitting process of each on its own, both ASMs are given an ideal initial translation from the manual segmentation to remove any bias in the results. Both ASMs are trained using a leave-one-out scenario to maximise the allowed synthesizable shape variation. Segmentation accuracy is determined through the use of point-to-curve error between the manual and segmented contours.

Table 3.1 illustrates the result of this experiment[1]. Figs. 3.11, 3.12 and 3.13 show three sets of side-by-side comparisons of final segmentations produced by both methods (the table and figures can be found at the end of this chapter). Examining the results in Table 3.1, one can observe that as the segmentations progress from the basal to apical slices the standard ASM produces more failed segmentations (taken as any mean error greater than 2 pixels). This is expected, as the endocardium border tends to become diffuse at the apex of the left ventricle. This causes the 1D search profiles to tend towards erroneous locations. These errors can accumulate as the ASM search progresses and lead to total segmentation failure. Fig. 3.13 contains two examples where the standard ASM method has completely missed the true endocardium border as these cases show very low contrast between the blood pool and the myocardium. In each case, the proposed optimisation method has correctly delineated the true endocardium border. Since initialisation parameters for both methods are identical, it can be said that the region-based contour score in the proposed method is superior to the standard ASM where gradient information is lacking.

The experimental evidence clearly indicates the superiority of the proposed method in terms of segmentation accuracy. As a general observation, the presented results for the standard ASM formulation indicate that the method is most susceptible to image noise, the influence of the papillary muscle, and

[1]While it is not within the scope of this thesis, an interesting avenue for future work is towards determining the effect of patient age on the performance of segmentation algorithms. It would be useful to determine what bias (if any) is introduced by the size of the patients heart.

diffuse endocardium borders in some images. By contrast, the proposed ASM optimisation method exhibits significant improvement with respect to all of these issues.

3.6 Summary

This chapter has dealt with the application of the ASM to cardiac left ventricle segmentation. First, a detailed description of the standard ASM modelling and fitting methodology is provided. Next, the limitations, specifically regarding the standard ASM fitting technique, are detailed. Discussion is also provided regarding application-specific alterations made to the ASM modelling phases for cardiac left ventricle shape modelling. Throughout this chapter references to related research presented in the literature are provided to give context to the discussion. The limitations regarding the ASM fitting drive the introduction of a novel ASM optimisation technique. The proposed optimisation method significantly improves the stability and segmentation accuracy of the ASM. The central difference is the fact that the proposed optimisation routine determines the fitting score based on the entire contour, as opposed to the local 1D profiles in the standard formulation. A detailed comparison of both approaches is presented, with the experimental evidence showing that the proposed optimisation method clearly outperforms the standard implementation. This contribution facilitates more significant contributions in the remainder of this thesis.

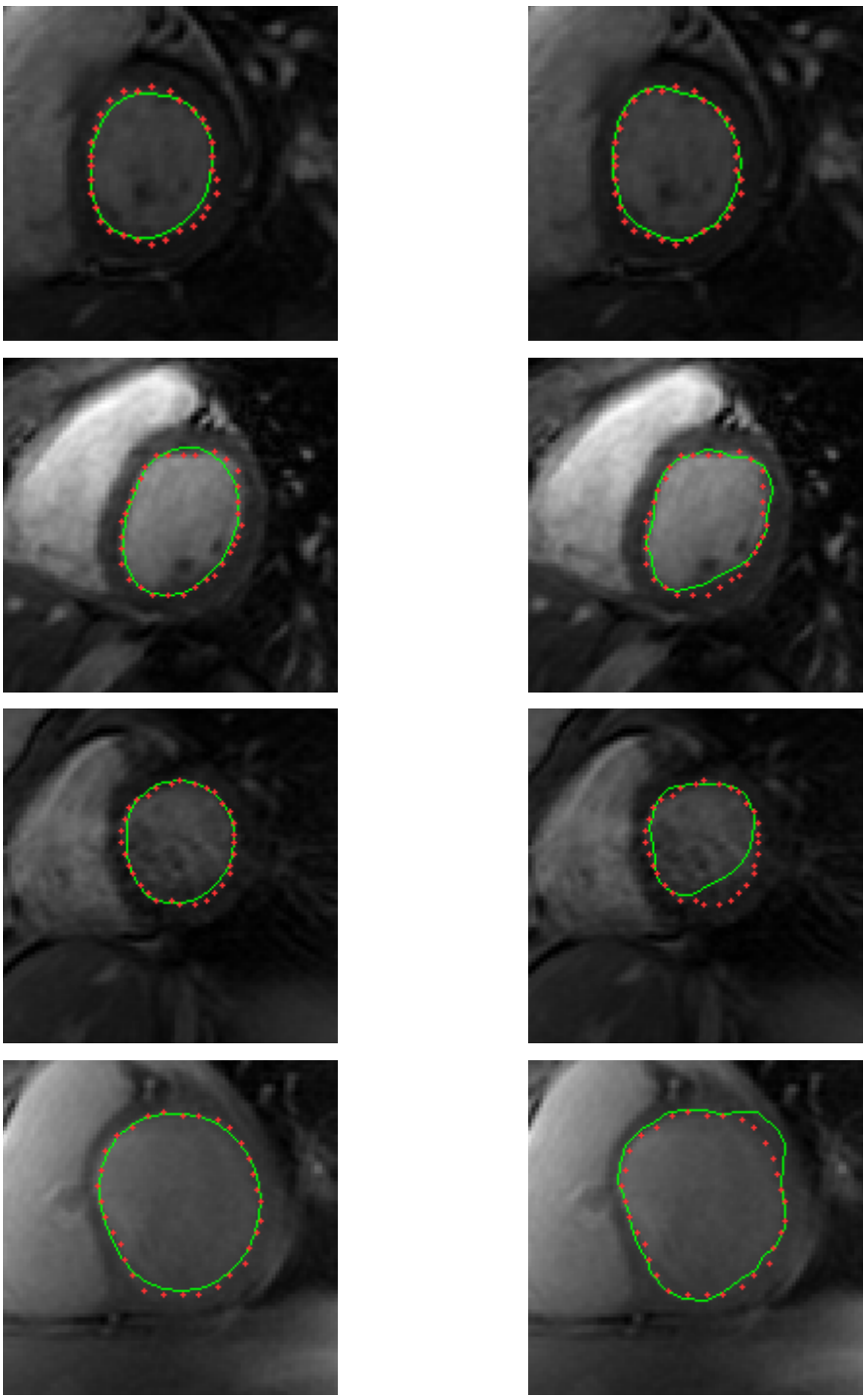


Figure 3.11: Example segmentations of basal-most slices. The left and right columns provide examples of the proposed optimisation technique and the standard ASM fitting method, respectively. Automatic segmentations are shown as continuous green contours, and manual annotations are shown as red crosses.

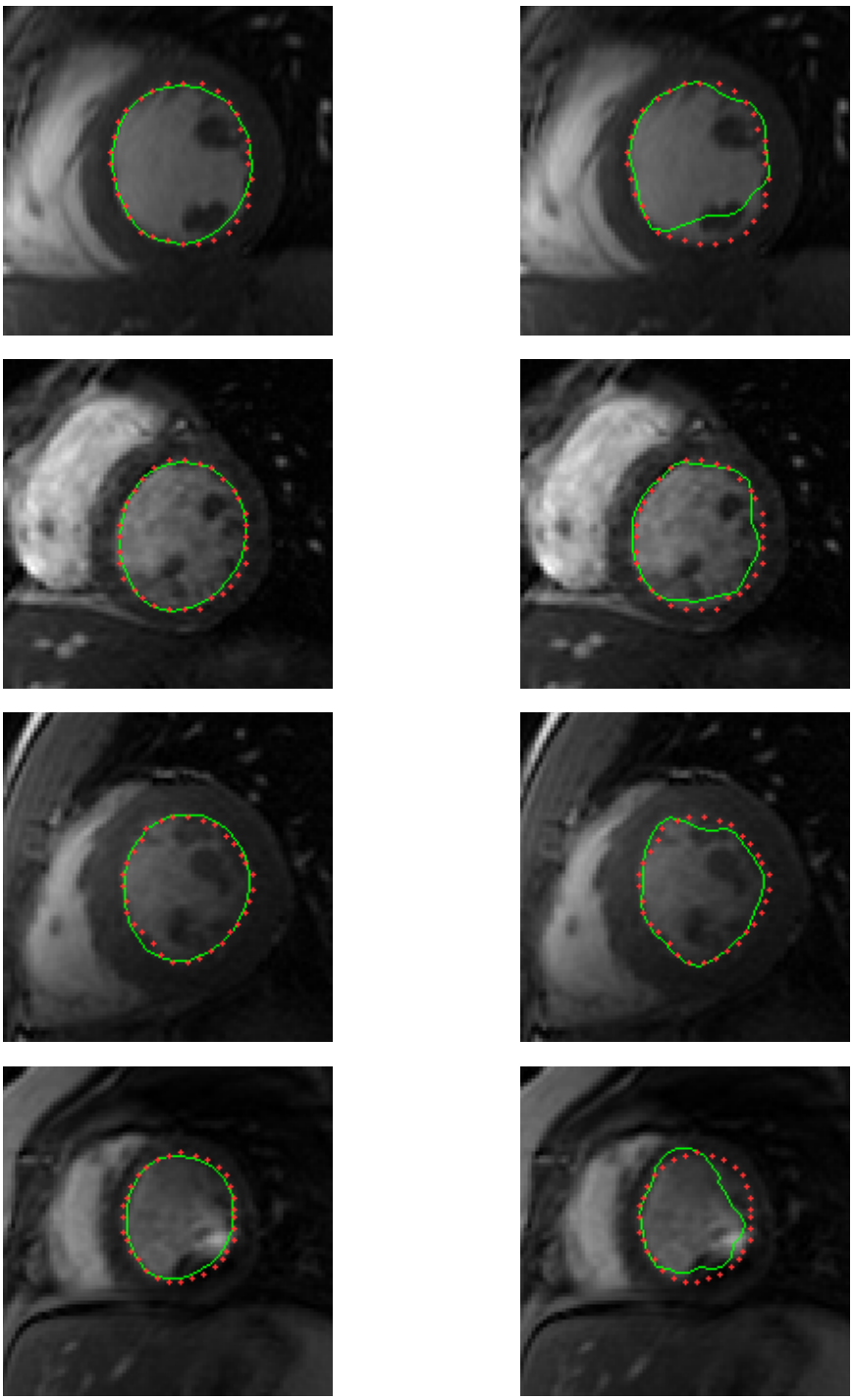


Figure 3.12: Example segmentations of mid-level slices. The left and right columns provide examples of the proposed optimisation technique and the standard ASM fitting method, respectively. Automatic segmentations are shown as continuous green contours, and manual annotations are shown as red crosses.

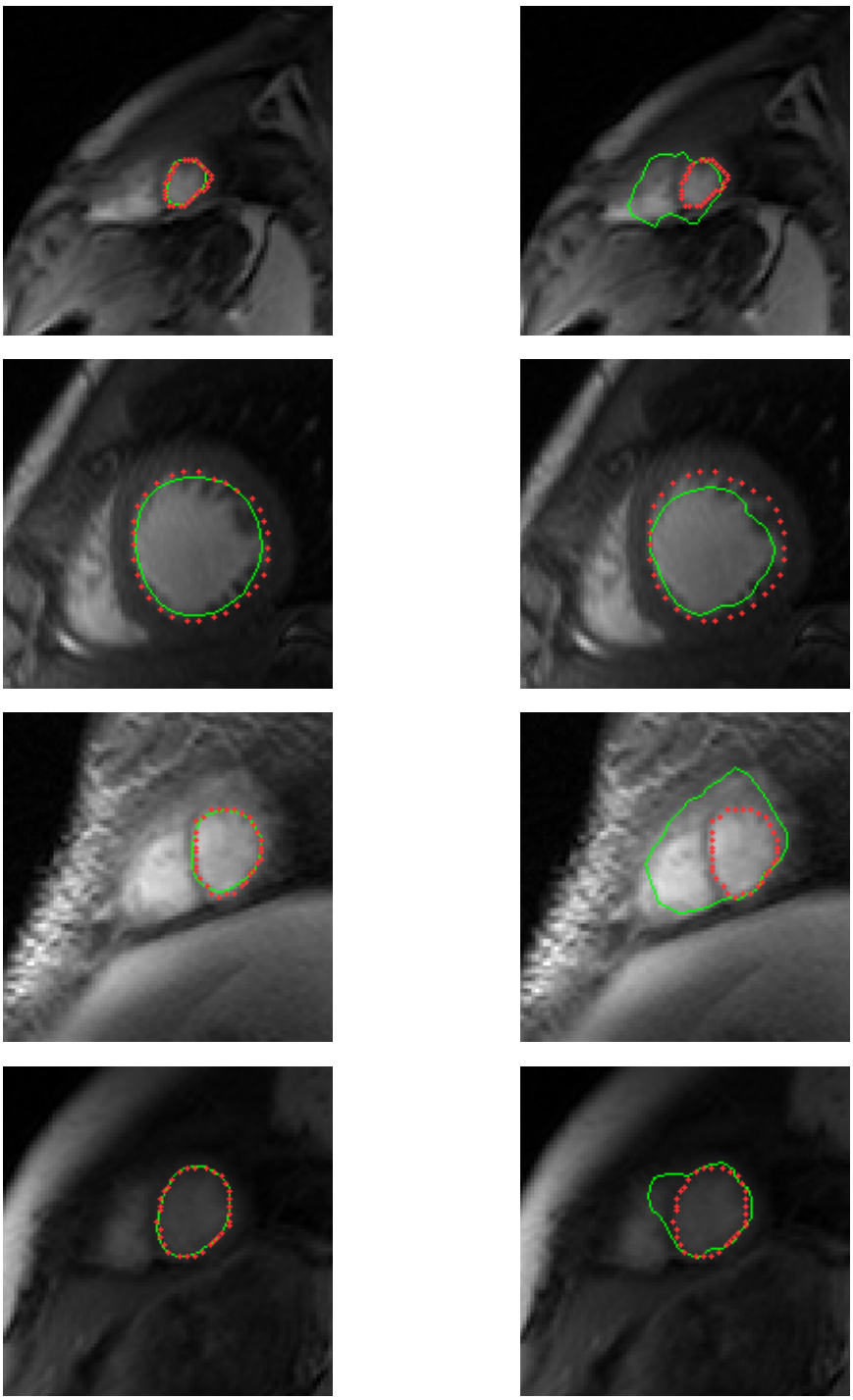


Figure 3.13: Example segmentations of apical-most slices. The left and right columns provide examples of the proposed optimisation technique and the standard ASM fitting method, respectively. Automatic segmentations are shown as continuous green contours, and manual annotations are shown as red crosses.

Table 3.1: ASM Comparison Results. Opt and Std refer to the proposed ASM optimisation method and the standard ASM fitting method, respectively. Each row, except the final two, present the average point-to-curve error between the manual and segmentation shapes of that slice. All measurements are in pixels.

Dataset	Basal		Mid		Apical	
	Opt	Std	Opt	Std	Opt	Std
1	0.56	0.91	1.19	1.59	0.97	1.19
2	0.97	0.97	0.56	1.34	0.59	2.25
3	0.75	1.06	0.69	1.38	0.69	2.38
4	0.75	0.84	1.06	1.09	0.91	0.62
5	0.63	1.03	0.66	1.97	0.72	5.47
6	0.59	1.78	1.09	1.59	0.69	1.06
7	0.66	0.91	0.81	2.16	0.66	1.41
8	0.56	0.75	0.63	1.06	1.28	2.03
9	0.47	0.56	0.63	1.53	1.12	2.09
10	0.56	0.78	0.72	1.03	1.47	2.25
11	0.56	0.84	0.56	3.16	0.75	1.19
12	0.59	1.12	0.84	1.62	0.62	7.16
13	0.69	0.56	0.47	0.75	0.56	0.44
14	0.63	0.75	0.88	1.66	0.66	0.66
15	0.84	0.88	0.88	1.12	0.91	2.72
16	0.63	0.88	0.56	0.81	0.66	1.06
17	0.72	0.72	1.38	1.31	0.66	0.78
18	0.81	1.16	0.94	1.59	0.78	1.00
19	0.81	1.06	0.81	2.00	0.56	5.28
20	0.41	1.78	0.63	1.06	1.25	1.59
21	0.81	0.50	0.81	1.19	0.66	1.59
22	0.94	0.97	1.16	2.28	0.94	1.16
23	0.66	1.03	0.91	1.88	0.69	1.75
24	0.69	1.47	0.66	1.38	0.97	1.25
25	0.59	1.59	0.56	1.75	1.22	1.41
26	0.63	1.06	0.91	1.16	1.09	0.91
27	0.88	2.00	0.63	2.69	0.62	1.34
28	0.53	1.25	1.25	1.25	0.47	6.62
29	1.19	1.03	1.03	1.09	0.53	0.75
30	0.72	0.97	0.63	1.22	0.97	1.47
31	0.72	1.19	0.78	0.69	1.00	2.16
32	0.56	0.81	0.59	1.19	0.66	2.34
33	0.97	0.78	0.56	1.00	0.50	1.81
Mean	0.70	1.03	0.80	1.47	0.81	2.04
Std Dev	0.16	0.35	0.23	0.54	0.25	1.67

Chapter 4

Contour Coupling

In many biological applications, only a single contour needs to be delineated. In these cases the ASM methodology is well suited and can, in general, successfully learn the shape variation of the target object given sufficient training samples. In the multi-contour scenario, however, directly applying the ASM methodology is not straightforward. In this chapter, the limitations of the multi-contour ASM will be outlined, and a novel contour coupling solution is presented as a solution. Two implementations of the proposed contour coupling technique are described. Both are evaluated against the standard multi-contour ASM and conclusions are drawn based on experimental evidence. In this thesis, the word 'coupling' is used to indicate that the relationship between two or more contours is enforced through biologically-based non-rigid constraints, as opposed to the contours being physically coupled.

The basis for creating multi-contour models, in general, is well grounded. In the model-based approach to segmentation, especially medical image segmentation, the maximisation of *a priori* knowledge is established as one of the primary motivation factors. As such, when the segmentation target involves two or more closely related structural borders, it is sensible to conclude that their relationship should also be modelled in order to create a model that reflects the biological structure as accurately as possible.

4.1 Multi-Contour ASMs

This chapter aims to illustrate that directly modelling multiple contours greatly limits the flexibility of the resulting model. The concrete example used throughout this chapter is modelling the endo- and epicardium. The common approach to modelling these related contours (in ASM-related research) is to considered both contours as a single shape [98, 109, 6]. In the case of the ASM, this means that the controlling β variable will vary both endo- and epicardium simultaneously. This technique will be referred to as the *endo+epi ASM*.

4.1.1 Search Profile Overlap

The *endo+epi ASM* methodology, when applied to myocardial segmentation, is prone to serious segmentation errors during the individual landmark fitting stage in the standard ASM implementation. During landmark search (outlined in Section 3.2), if the 1D search profiles of the endo- and epicardium landmarks overlap it is highly probable that either the endo- or epicardium landmarks will converge on the wrong image contour (i.e. endocardium landmarks may tend towards concavities associated with the papillary muscle, while epicardium landmarks may incorrectly be attracted to sections of the endocardial interface). This is the case with the standard ASM implementation where each landmark is given equal weighting during the fitting stage, and no inter-landmark knowledge is enforced. The ASM attempts to approximate this arrangement of landmarks and provides a 'mid-way' best-fit set of contours. Fig. 4.12(a) shows an example of this effect, where the lower right landmarks from both the endo- and epicardium (landmarks are shown as crosses, colour-coded for their respective contours) converge on the left ventricular blood pool/myocardium interface. The ASM still tries to approximate these landmarks, with obvious negative results. For clarity, a schematic of the erroneous process is visualised in Fig. 4.1.

A natural solution to this issue is one where the fitting process is weighted. For example, a possible arrangement is that the endocardium landmarks are fitted first, then the epicardium landmarks are constrained to fit outside the endocardium landmarks (or vice versa depending on the situation). Al-

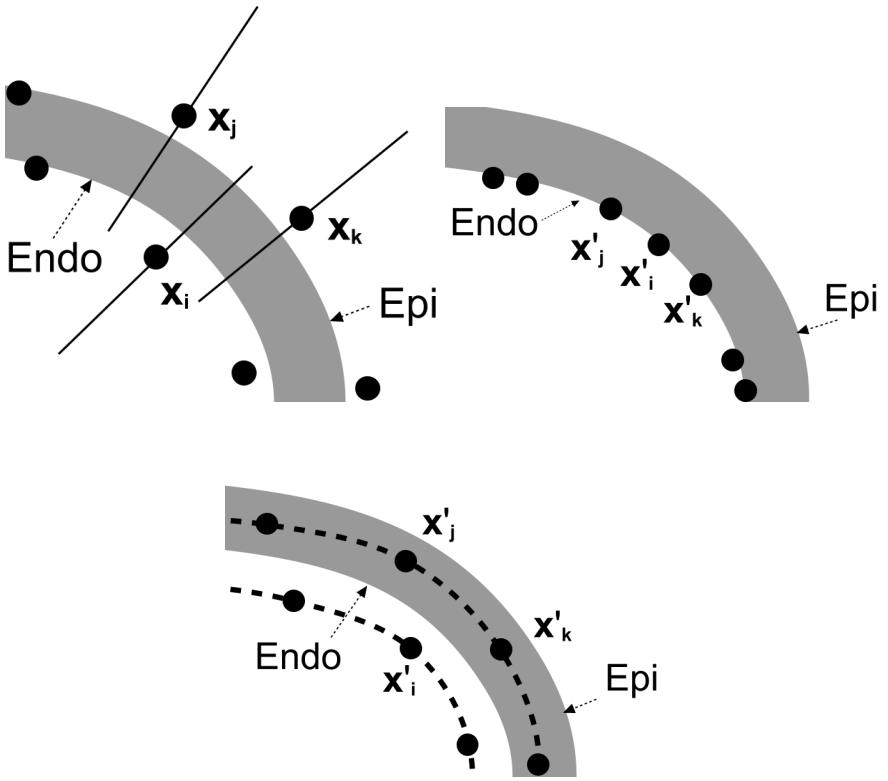


Figure 4.1: Schematic representation of the cause and effect of search profile overlap in the *endo+epi ASM* method. The upper-left image shows the initial landmark configuration, where the search profiles overlap both the endo- and epicardium interfaces. The upper-right image depicts outcome of landmark fitting. The endocardium is a stronger feature than the epicardium, and causes all landmarks to converge on its border. Finally, the lower image shows the result of the ASM when it tries to approximate this invalid arrangement of landmarks.

though better than the first case, since it prevents the search profile overlap problem, it is still prone to error since the endocardium fitting stage can be negatively influenced by the papillary muscles, (as discussed in Section 3.2), which may cause the epicardium landmarks to erroneously fit, resulting in an incorrect left ventricle segmentation. Of course, even in an ideal scenario where the ASM fitting method is fully robust, and there is no search-profile overlap, the variation of an *endo+epi ASM* will still be limited due to the theoretical reasons outlined in the following discussion.

4.1.2 Inter-Contour Distance Modelling

The *endo+epi ASM* captures three separate sources of variation. First, the shape variation of the endocardium is captured. Second, the epicardium shape variation. And finally third, the inter-contour distance is modelled, between the endo- and epicardium in each training sample.

Creating an *endo+epi ASM* relies on the assumption that certain endocardium shapes can *only* appear in the presence of associated epicardium shapes (and vice versa) and *only* at the observed distance between them in training. Obviously this is a restrictive assumption, for although the endo- and epicardium are closely related, it requires a significant number of training sets to sufficiently sample these sources of variation at the same time. Since large annotated training sets are uncommon, the variation in *endo+epi ASMs* will suffer as a result. Naturally, the severity of these difficulties are proportional to the number of contours considered as a single shape (for example, it requires yet more training if the LV and RV are considered as a single shape) as well as the dimensionality of the desired *endo+epi ASM*.

To illustrate this problem, an example scenario will be discussed, but first some descriptive notation needs to be introduced. Each shape, s, has an associated shape 'size' value, derived from its L^2-norm. The L^2 norm (also known as the $L2$-norm) is a common measurement used in the literature and in this thesis is used to quantify shape size, which, for a shape s, is defined as: $L^2(s) = \max(SVD(s))$ where $SVD(s)$ is the singular value decomposition of s. Therefore, an arbitrary endocardium shape size is denoted by L^2_{endo}. L^2_Δ is used to denote the difference in shape size between the endo- and epicardium in an arbitrary training sample. In a standard short-axis training sample, there is a single pairing of an endo- and an epicardium shape, denoted as: $(\mathbf{s}_{endo}, \mathbf{s}_{epi})$.

To experimentally illustrate the effect of modelling inter-contour distance in the *endo+epi ASM* method, a specific training and evaluation scenario is created. First, consider Fig. 4.2(a), which shows the shape size relationship between all short-axis slices in the datasets used throughout this thesis. Each point in this graph corresponds to a single endo- and epicardium pairing. This shows the range of observed variation between endo- and epicardium shape sizes from basal-most to apical-most short-axis slices. Of note, to-

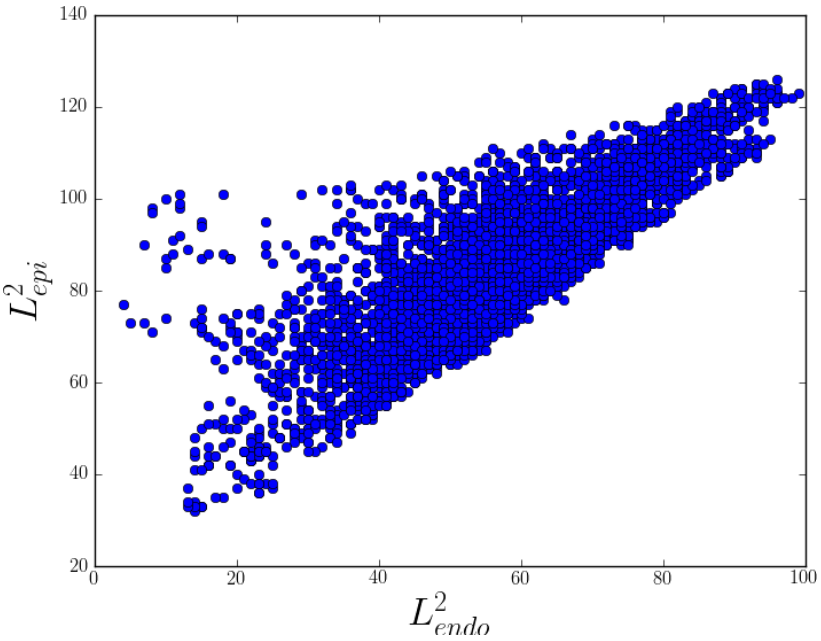
wards the left-hand side of the graph (corresponding to apical shapes), one can observe some cases where a very small endocardium (i.e. $L^2_{endo} < 20$) are associated with very large epicardium ($L^2_{epi} > 80$). Investigation of the data for these cases reveals that this occurs in a pathological dataset exhibiting significant ventricular hypertrophy (i.e. excessive myocardium thickening).

Fig. 4.2(b) presents a visualisation of the variation in inter-contour distance between each training sample, using L^2_Δ. For example, the region marked as A in this figure corresponds to cases where the epicardium is small, and the inter-contour distance between the endo- and epicardium is also small (i.e. the myocardium is thin). Region B, therefore, covers cases where the endocardium is larger and the myocardium is large. Finally, region C contains samples where the endocardium is as large as in region B, but here the myocardium is thin. To clarify the discussion, it can be said that since L^2_{endo} and L^2_Δ are both small in region A, it can be described as: $A = \{S, S\}$. Similarly, region B: $B = \{B, B\}$ (both big) and region C: $C = \{B, S\}$, since the endocardium is big, but the inter-contour distance is small.

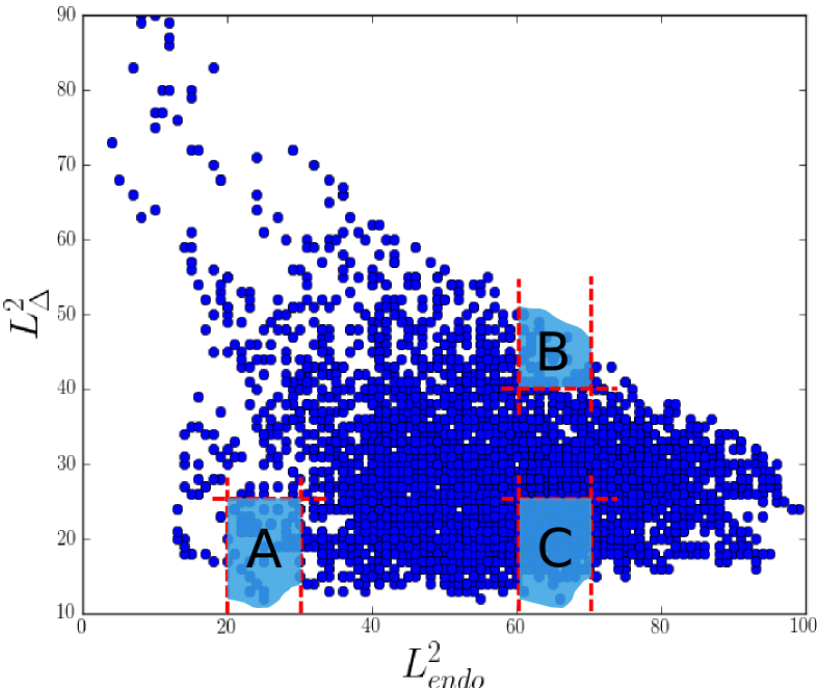
Due to the limitations of the approach, if an *endo+epi ASM* were to be trained using regions A and B, the resulting model would be unable to synthesize shapes to segment cases in region C, as well as a region defined by $\{S, B\}$[1]. This is considered as a short-coming in this thesis because, although the ASM has trained using samples where the endocardium is small (in region A) and the epicardium is large (region B), the ASM has never seen a case where a small endocardium has occurred *in the presence of* a large epicardium. This is clearly a weakness in the *endo+epi ASM* because there is no reason that one endocardium training sample cannot occur in the presence of another epicardium training sample. The only constraint on the pairing of shapes is that the epicardium must be larger than the endocardium.

To experimentally verify the presence of this limitation with the *endo+epi ASM* a training/evaluation scenario is defined, according to Fig.4.2(b). Specifically, the regions A, B and C:

[1] In a similar fashion, an *endo+epi ASM*, trained with regions $\{\{B, S\}, \{S, B\}\}$, will be unable to synthesize shapes in regions $\{\{B, B\}, \{S, S\}\}$.



(a) Relationship between L^2_{endo} and L^2_{epi} in the training datasets



(b) Illustrating training and evaluation regions with respect to inter-contour distance

Figure 4.2: Two figures illustrating the relationship between endo- and epicardium shape size, as well as the range of variation of inter-contour distance seen throughout the cardiac datasets used in this thesis.

$$A = \left\{(\mathbf{s}_{endo}, \mathbf{s}_{epi}) \mid 20 \leq L^2_{endo} \leq 30, L^2_{\Delta} \leq 25\right\} \tag{4.1}$$

$$B = \left\{(\mathbf{s}_{endo}, \mathbf{s}_{epi}) \mid 60 \leq L^2_{endo} \leq 70, L^2_{\Delta} \geq 40\right\} \tag{4.2}$$

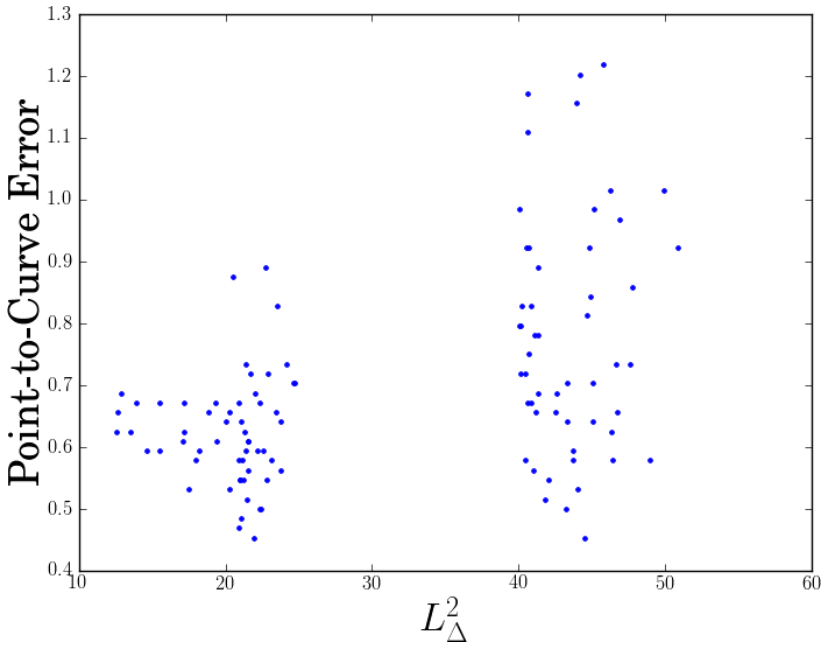
$$C = \left\{(\mathbf{s}_{endo}, \mathbf{s}_{epi}) \mid 60 \leq L^2_{endo} \leq 70, L^2_{\Delta} \leq 25\right\} \tag{4.3}$$

so that the training set is defined as: $\{A, B\}$ and the evaluation set is defined as: $\{C\}$. To evaluate the accuracy of ASM synthesis in this scenario, ASM segmentation is *not* applied. Instead, the manual segmentation for each evaluation dataset is directly approximated by the *endo+epi ASM*. This removes any segmentation error from the presented results, so that only the error in shape synthesis is displayed.

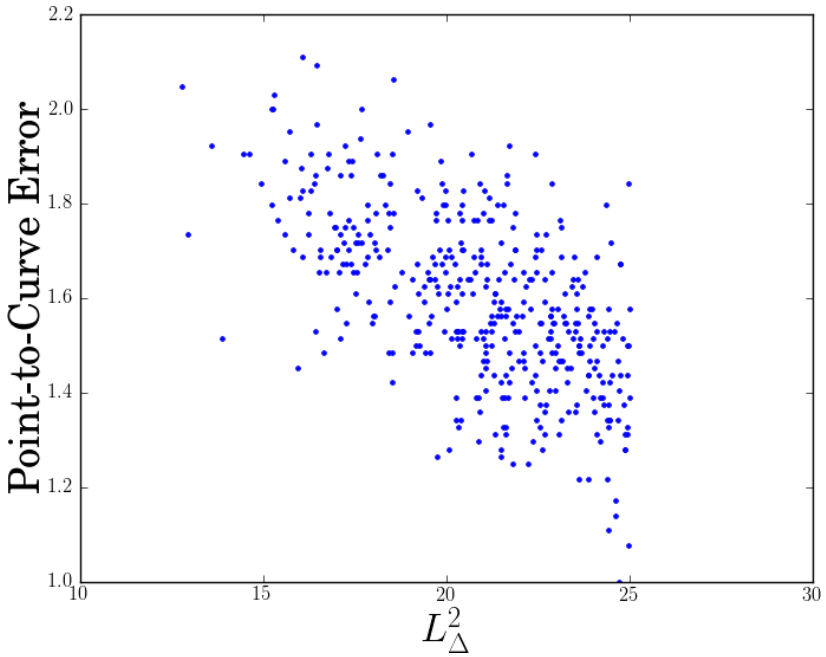
Before the evaluation results are discussed, a sanity check is first carried out by measuring the level of shape synthesis accuracy when the ASM is applied to its own training set. This is depicted in Fig. 4.3(a). As shown in this figure, the majority of cases show less than a single pixel of error on average. The handful of cases with slightly larger errors in the upper-right of this figure correspond to unusual shapes instances, which cause the decrease in accuracy, since they are towards the outer limits of allowed shape synthesis. Fig. 4.3(a) confirms that the trained *endo+epi ASM* can accurately synthesise shapes in its training set. By contrast, Fig.4.3(b) presents much larger shape synthesis errors when the same *endo+epi ASM* is applied to the evaluation set. Here, the notable increase in point-to-curve error indicates the inability of the *endo+epi ASM* to synthesize shapes from the region C evaluation set.

Fig.4.4 shows four example images from the evaluation scenario in Fig.4.3(b). In each image, the *endo+epi ASM* attempts a 'best-fit' between both the endo- and epicardium contours, however in each case it is unable to synthesize the correct inter-contour distance. This confirms the expected behaviour of the *endo+epi ASM* with respect to inter-contour distance modelling.

In summary, the two main limitations of the *endo+epi ASM* approach are search-profile overlap using the standard ASM fitting technique and the unnecessary limitation placed on allowable shape synthesis caused by considering both contours as a single shape. To address both of these short-comings a



(a) *Endo+epi ASM* shape synthesis accuracy on its training set (regions A and B).



(b) *Endo+epi ASM* shape synthesis accuracy on the evaluation set (region C).

Figure 4.3: Shape synthesis errors (measured in pixels) of the *endo+epi ASM*, when applied in two different evaluation scenarios.

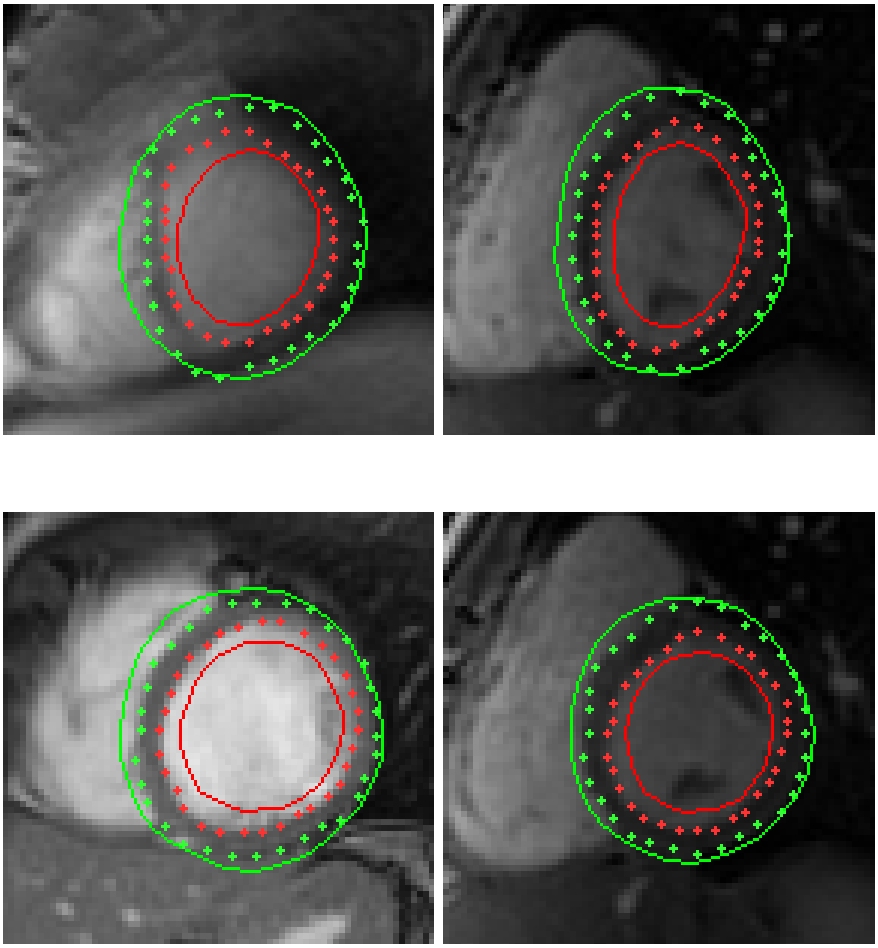


Figure 4.4: Four example images illustrating the effect of the *endo+epi ASM* when applied to cases outside its learned level of inter-contour variation. In each image, the manual annotation is shown as dotted contours and the shape synthesis result is shown with continuous contours. The colours red and green correspond to the endo- and epicardium, respectively.

novel contour coupling technique is developed in the following section, which is more flexible, even allowing contour configurations *outside* the training set to be correctly synthesized.

4.2 Contour Coupling

Unlike the *endo+epi ASM* approach, the contour coupling concept (referred to as *CC-ASM*) proposed in this chapter, allows extraneous limitations on the endo- and epicardium to be removed at model-construction time - instead enforcing coupling constraints at segmentation-time. Several benefits can be gleaned from this initial statement. Both contours are no longer considered as one shape, therefore two separate expert ASMs are trained for each contour. Moreover, there is no inherent dependence on learning one contour in the presence of another - i.e. *any* learned endocardium shape can appear with *any* other learned epicardium shape.

In essence, this means that the inter-contour distance between the endo- and epicardium is no longer inherently modelled at model-time (i.e. based on the variety of contour configurations seen in training). Instead, a rule-based technique can be developed to enforce *a priori* biological constraints.

This fact alone greatly enhances the 'capture range' of *CC-ASM* compared with the *endo+epi ASM* approach. Crucially however, by removing the inter-contour distance from the modelling process data-driven, or model-driven, segmentation-time constraints can be enforced, that ensure structural consistency is maintained during multi-contour segmentation. Combining these two benefits facilitates segmentation of unseen datasets that are *outside* the training set (normally an inherent constraint of *endo+epi ASM*-based methods) with increased flexibility in pathological cases, for example, where the myocardium is unusually thick (hypertrophy) or thin.

To illustrate the benefit of the contour coupling approach, the same experiment that was used earlier in the chapter to show the limitation of the *endo+epi ASM* approach with respect to inter-contour distance modelling will be repeated here. Therefore, two separate ASM models for the endo- and epicardium are created using the training sets defined by Equations

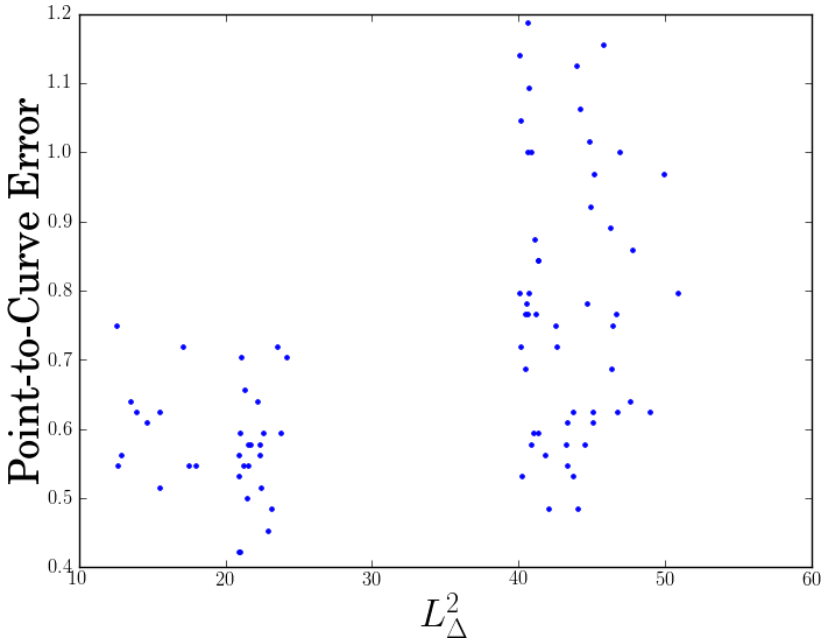
(4.1) and (4.2) and the evaluation set is defined in Equation (4.3). Again, this arrangement is illustrated graphically in Fig 4.2(b). Maintain

As illustrated in Fig 4.5(a), applying these ASMs to the evaluation set (again, as a sanity check) a similar level of performance between the contour coupling approach and the *endo+epi ASM* method is observed, as expected. Crucially however, when the contour coupling approach is applied to the evaluation set (Fig 4.5(b)), there is no loss in shape synthesis accuracy. This is because the contour coupling method can take take any endocardium training shape and pair it with any epicardium training shape to achieve segmentation. Compared to the equivalent evaluation results in Fig 4.3(b) for the *endo+epi ASM*, the benefit of the contour coupling is made clear. Figs. 4.6 and 4.7 show some side-by-side qualitative comparisons of the *endo+epi ASM* segmentations compared to segmentations generated using the proposed contour coupling approach.

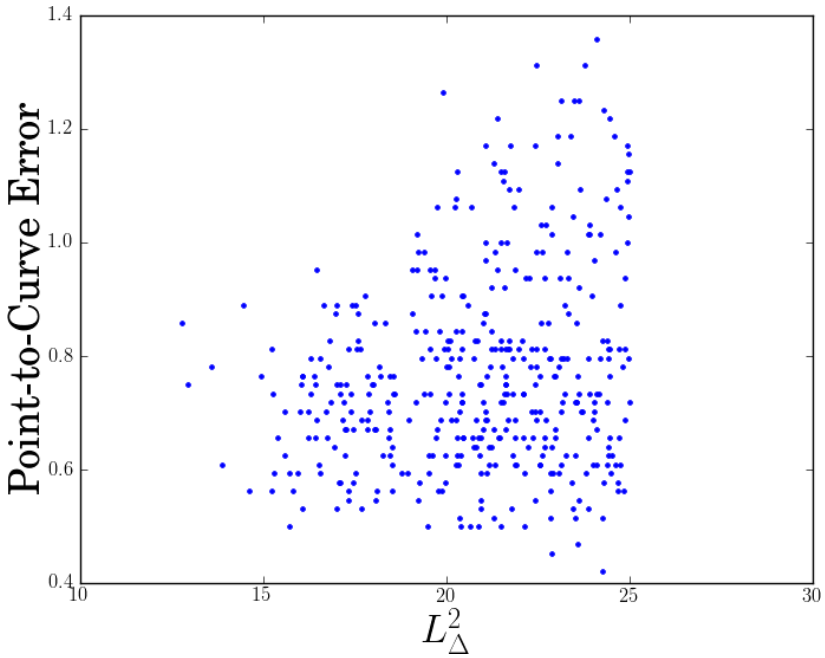
As indicated earlier, this experiment utilises the manual segmentations for each evaluation scenario so that segmentation errors are not included in the presented results. While this is useful to transparently illustrate the accuracy of shape synthesis alone, in a practical scenario an unseen evaluation image must be properly segmented. The method for endocardium segmentation, using the proposed ASM optimisation method presented in Chapter 3, has already been discussed. However, to achieve myocardial segmentation the epicardium's objective function must also be developed.

4.2.1 Epicardium Optimisation

The epicardium is optimised similarly to the endocardium, however the difference between the standard deviations of the inner and outer region also needs to be maximised since the area defined by the myocardium is relatively homogeneous compared to the region immediately outside the myocardium wall. Specifically, outside the myocardium there is significant intensity variation between the bright right ventricular blood pool and the very dark surrounding thoracic region. The addition of this extra term is required as the epicardium border is usually poorly defined in MRI images, therefore extra robustness is required at this segmentation stage. To highlight the fact that a different ASM is trained for the epicardium (this issue will be



(a) Shape synthesis accuracy using the contour coupling approach on its training set.



(b) Shape synthesis accuracy using the contour coupling approach on the evaluation set.

Figure 4.5: Shape synthesis errors (measured in pixels) of the contour coupling approach, when applied in two different evaluation scenarios.

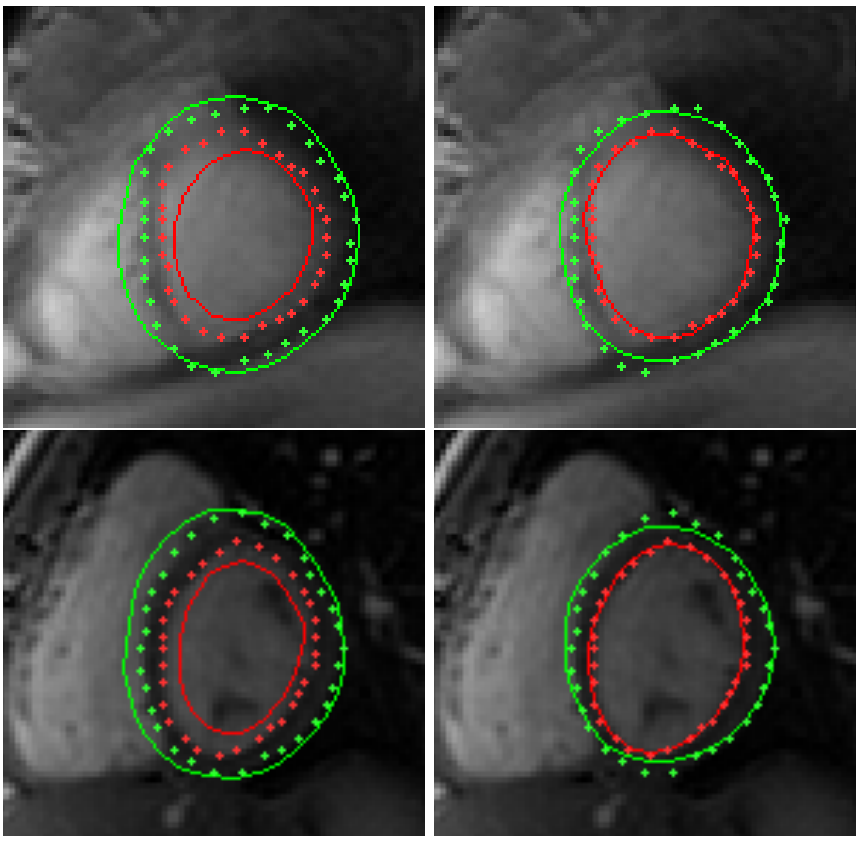


Figure 4.6: Side-by-side comparison of the segmentation results for the *endo+epi ASM* methodology (shown in the left column) and the contour coupling method applied to the same images (show in the right column).

discussed further in the next chapter on contour coupling), β_{epi} is used to refer to the epicardium shape parameters in the objective function:

$$\begin{aligned} f_{epi}\left(\beta_{epi}, T\right) &= \left[\left|\overline{I}\left(T\left(\Omega^{-}_{\beta_{epi}}\right)\right) - \overline{I}\left(T\left(\Omega^{+}_{\beta_{epi}}\right)\right)\right| \\ &+ \left|I^{\sigma}\left(T\left(\Omega^{-}_{\beta_{epi}}\right)\right) - I^{\sigma}\left(T\left(\Omega^{+}_{\beta_{epi}}\right)\right)\right|\right] \end{aligned} \tag{4.4}$$

Similarly to (3.17), the epicardium objective function is optimised according to:

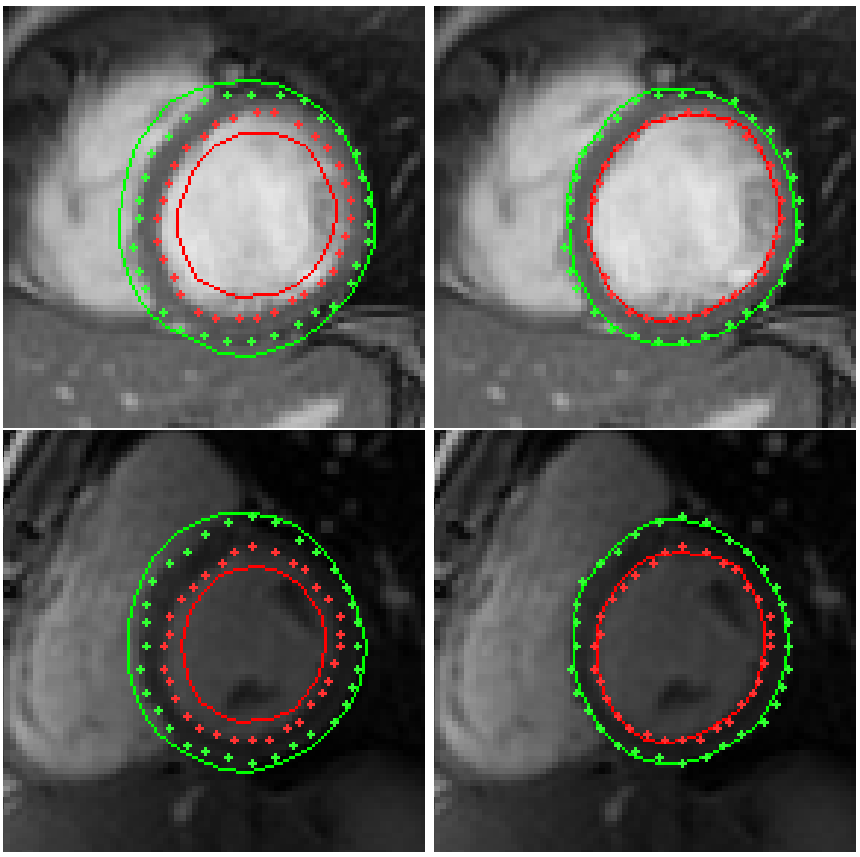


Figure 4.7: Side-by-side comparison of the segmentation results for the *endo+epi ASM* methodology (shown in the left column) and the contour coupling method applied to the same images (show in the right column).

$$\max_{|\beta_{epi}| \leq \beta_{epi}^{limit}} f_{epi}\left(\beta_{epi}, T\right) \tag{4.5}$$

Now that both of the objective functions have been introduced, the separate endo- and epicardium ASMs must be coupled. In this chapter two different contour coupling implementations are evaluated. The first involves a data-driven technique, where the epicardium region of interest (ROI) is derived from the image gradient information. The second approach derives the epicardium ROI from a model created from manual annotations during ASM training. The contour coupling technique is formulated as a multi-stage process. For both approaches, the endocardium is initially segmented in

the image. The endocardium region, Ω_{endo}, is used to drive the epicardium segmentation as it is the stronger feature of the two.

4.2.2 Data-driven Contour Coupling

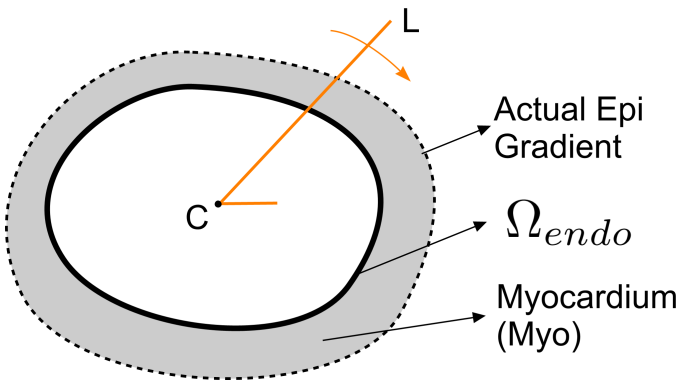


Figure 4.8: Radial search for gradient data, after endocardium segmentation.

In the data-driven approach (referred to as CC-DD), a radial search is performed, starting from the centre of the shape, $C = \frac{1}{N}\sum_{i=1}^{N} l_i$ (where l_i is the i^{th} landmark in the shape and N is the number of landmarks in the shape), by collating gradient information along a 1D profile of length L (Fig. 4.8). The result of the process used to construct the cumulated gradient profile is illustrated in Fig. 4.9. Using Ω_{endo}, a probability density function with parameters μ_{endo} and σ_{endo} can be derived, describing the mean distance and standard deviation, respectively, between the endocardium contour and C. Since the epicardium is approximately concentric with the endocardium, the probable distance of the epicardium from C, μ'_{epi}, can be *estimated* as the distance corresponding to the next peak in the 1D gradient profile after μ_{endo} (Fig. 4.9). To take account for irregularities in the shape of the epicardium and to allow for error in the estimate of the epicardium the region $\Omega_{\mu'_{epi}+3\sigma_{endo}}$ is defined as the outer limit of the ROI, where the epicardium ASM is allowed to optimise (see Fig. 4.10). Statistically, $3\sigma_{endo}$ is a robust indicator of the variation in the endocardium, as $3\sigma_{endo}$ accounts for 99.7% of the observed distribution in the collated endocardium gradient profile. $3\sigma_{epi}$ is not used since the epicardium gradient information is less reliable than the endocardium, and, since the two contours are biologically

linked, the endocardium variation reliably reflects the epicardium variation. Therefore the region of interest is defined as:

$$\Omega_{roi} = \Omega_{\mu'_{epi}+3\sigma_{endo}} \setminus \Omega_{endo} \tag{4.6}$$

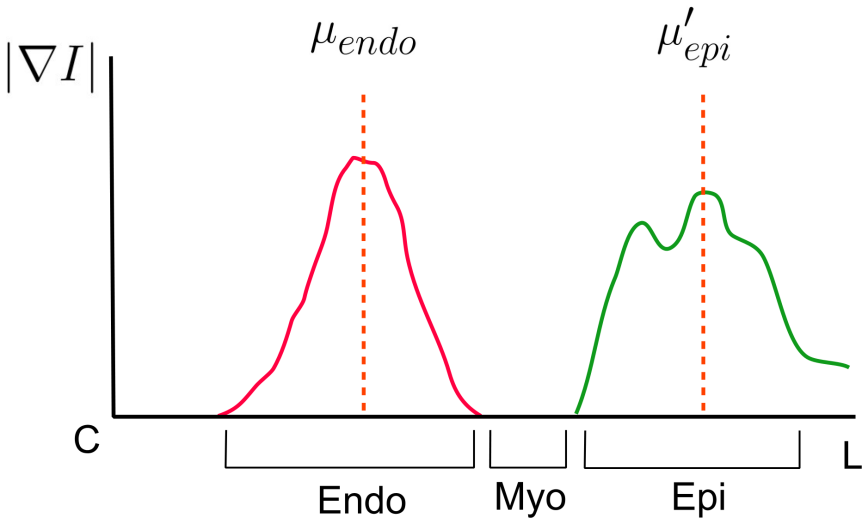


Figure 4.9: Representation of the epicardium estimation, derived from the known endocardium border.

The data-driven contour coupling constraint is applied so that Equation (4.5) is subject to the extra constraint:

$$\begin{aligned} \mathbf{s}_{\beta_{epi}} &\subseteq \Omega_{roi} \\ &\subseteq \Omega_{\mu'_{epi}+3\sigma_{endo}} \setminus \Omega_{endo} \end{aligned} \tag{4.7}$$

4.2.3 Model-driven Contour Coupling

Unlike CC-DD, the model-driven contour coupling implementation (referred to as CC-MD) does not derive Ω_{roi} from the image data. Instead, the manual annotations in the training data are used to construct a model that estimates the range of variation expected for the epicardium, given an endocardium shape. As stated earlier in this chapter, the L^2 norm is used to quantify

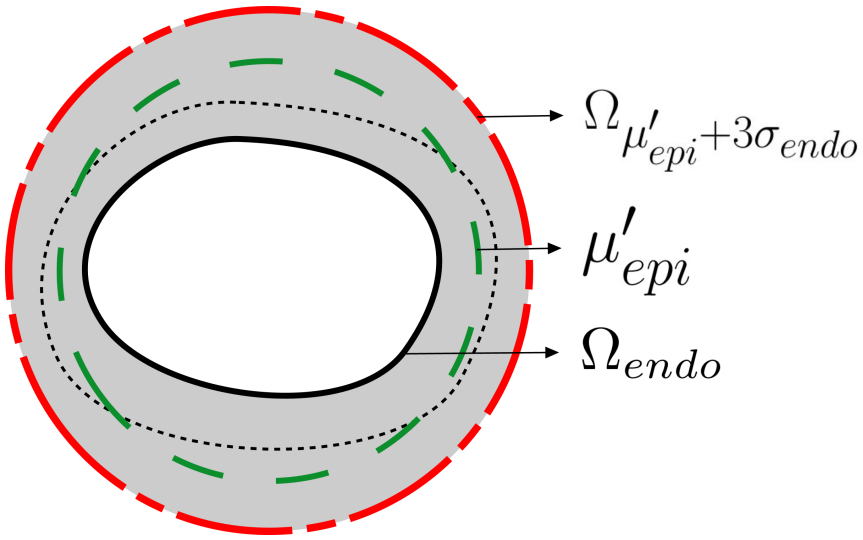


Figure 4.10: Using the estimated mean epicardium location μ'_{epi} to derive the outer limit of the region of interest, $\Omega_{\mu'_{epi}+3\sigma_{endo}}$.

shape size. While training the endocardium ASM, each endocardium's L^2 norm is collated with the corresponding epicardium L^2 norm. Fig. 4.11 shows a sample output of the shape size relationship between the endo- and epicardium from a normal training set. Since the endocardium is segmented first, it can be used to derive the outer limit of Ω_{roi} without the need to calculate image gradient information, as in the CC-DD. For example, given a segmented endocardium s_{endo} (with corresponding region, Ω_{endo}) its L^2 norm can be calculated and referred to as L^2_{endo} for clarity. Using the *a priori* model, the maximum epicardium shape size, $L^2_{epi'}$, is estimated according to:

$$L^2_{epi'} = \sup\{L_{var}\left(L^2_{endo}\right)\} \tag{4.8}$$

where $L_{var}(L^2_{endo})$ is the function that encodes the learned variability of the epicardium with respect to the L^2 norm of the endocardium (see Fig. 4.11) and $\sup\{L_{var}\left(L^2_{endo}\right)\}$ denotes the upper bound (supremum) of the epicardium shape variation for a given endocardium L^2 norm. Following from this, Ω_{roi} is defined for the CC-MD contour coupling method as:

$$\Omega_{roi} = \Omega_{L^2_{epi'}} \setminus \Omega_{endo} \tag{4.9}$$

where $\Omega_{L^2_{epi'}}$ is the region defined by the estimated maximum epicardium shape size, relative to the centre of s_{endo}. It is useful to note that shapes outside this region are not forbidden, but they are linearly weighted to reduce their score, in proportion to how much they deviate from the estimated limit (see Equation (4.10) and Fig. 4.11). On the other hand, shapes inside the endocardium region *are* forbidden, since it is known that the epicardium must be outside the endocardium. This encourages the optimisation routine to work within the ROI, but allows deviation outside, if the contour score is sufficient to overcome the adverse weighting. As such, the weighting function[2] is defined as:

$$w\left(\beta_{epi}\right) = \begin{cases} 0, \; L^2_{\beta_{epi}} \leq L^2_{endo} \\ 1, \quad L^2_{endo} < L^2_{\beta_{epi}} \leq L^2_{epi'} \\ \dfrac{L^2_{epi'}}{L^2_{\beta_{epi}}}, \; L^2_{epi'} < L^2_{\beta_{epi}} \end{cases} \tag{4.10}$$

It follows that the epicardium objective function must be amended for model-driven contour coupling according to:

$$f^{CC-MD}_{epi}\left(\beta_{epi}, T\right) = f_{epi}\left(\beta_{epi}, T\right) \times w\left(\beta_{epi}\right) \tag{4.11}$$

where $f_{epi}\left(\beta_{epi}, T\right)$ is defined in Equation (4.4) and $f^{CC-MD}_{epi}\left(\beta_{epi}, T\right)$ is maximised according to Equation (4.5).

The model-driven approach has a distinct advantage over the data-driven method presented in Section 4.2.2. Since there is no direct interaction with the image data in CC-MD, the estimation of Ω_{roi} cannot be negatively influenced by erroneous, or absent, gradient data. While the strong myocardium/right ventricle boundary is more than sufficient in CC-DD for calculating μ'_{epi}, this boundary is only present for the basal and mid long-

[2]This weighting function is chosen because of good performance in practice and the absence of manually selected weighting parameters. The use of other weighting functions is, of course, possible and should be driven by application-based constraints.

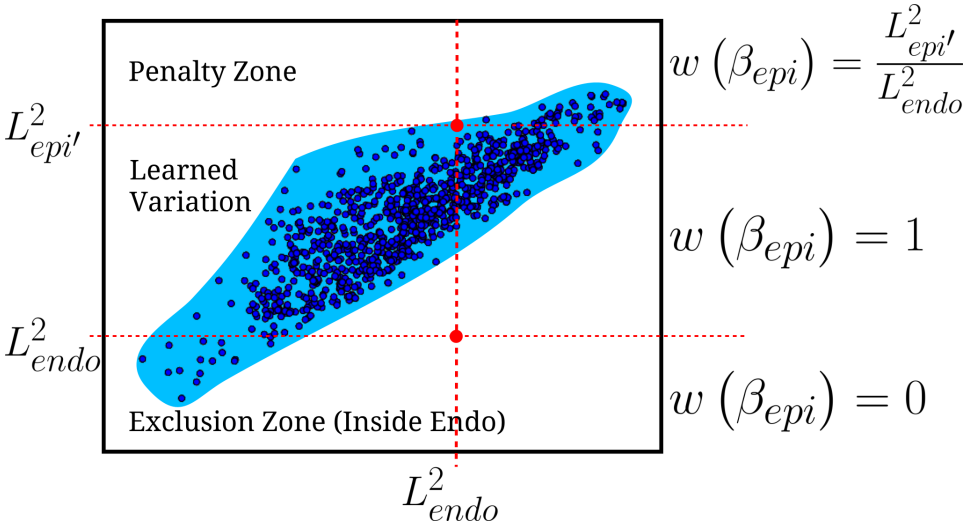


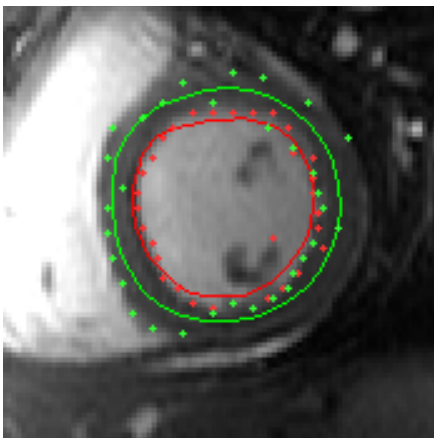
Figure 4.11: Example derivation of the epicardium region of interest using CC-MD. Given a segmentation endocardium shape s_{endo}, $L^2_{epi'}$ can be estimated according to Equation (4.8), which is directly used to define Ω_{roi} in Equation (4.9). Also, shown are the respective weighting constraints ($w(\beta)$) for epicardium optimisation that ensures structural consistency is maintained during segmentation. Note, that in the penalty zone: $w\left(\beta_{epi}\right) \in (0, 1)$. The x-axis and y-axis of this figure refer to endocardium and epicardium L^2 norm, respectively.

axis slices. Towards the apex, the myocardium can become extremely diffuse, which can affect derivation of the ROI from image data. Therefore, the model-driven method is expected to be more robust towards the apical slices.

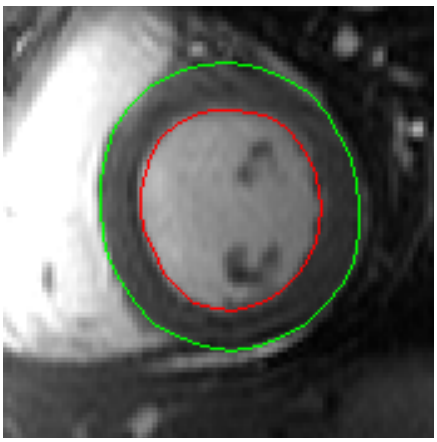
These two approaches are evaluated in Section 4.3.1 of this chapter. It will be shown in the evaluation that by avoiding image gradient issues the CC-MD implementation creates more accurate segmentations. Using either CC-DD or CC-MD, Ω_{roi} enforces structural consistency between the endo- and epicardium interfaces. The fact that the inherent relationship between the contours is enforced only at segmentation-time, through Ω_{roi}, allows many more potential unseen contour configurations, outside the training set variation, to be successfully segmented. To this end, the inner limit of the ROI in both CC-MD and CC-DD is constrained using only the *a priori* knowledge that the epicardium must simply be outside the endocardium.

Fig. 4.12 illustrates the benefit of the proposed coupling approach (created

using CC-MD) in contrast with an *endo+epi ASM*. As noted in the initial discussion regarding contour coupling, the *endo+epi ASM* fails to segment the left ventricle image because certain landmarks can converge where search profiles overlap between the endo- and epicardium landmarks (see the beginning of Section 4.2 for a detailed discussion), and produces a 'mid-way' set of contours as shown in Fig. 4.12(a). By contrast, the CC-MD approach (Fig. 4.12(b)) segments the endocardium first, derives the allowed ROI for the epicardium, which constrains the separate epicardium ASM optimisation. Note, that even if the proposed optimisation method is used with the *endo+epi ASM*, this approach would still be lacking in allowed variation as it captures only the inter-contour distances seen in training - so while it may generate an improved result compared with Fig. 4.12(a), its theoretical 'capture range' is still far more limited than the CC-ASM technique.



(a) Standard *endo+epi ASM* multi-contour segmentation result.



(b) The CC-MD contour coupling method applied to the same image.

Figure 4.12: Direct qualitative comparison between the standard ASM implementation when considering multiple contours as a single shape and the proposed contour coupling technique (CC-MD). The landmarks shown in Fig. 4.12(a) are a result of the ASM search phase.

4.3 Performance Evaluation

Earlier in this chapter, it has been shown that the *endo+epi* modelling strategy is not suited to cardiac left ventricle segmentation, due to search profile overlap (see Section 4.1.1) and inter-contour distance modelling (see

Section 4.1.2). Therefore, this approach will be unable to produce results comparable with the two novel contour coupling implementations presentation in Sections 4.2.2 and 4.2.3 and is therefore excluded from the evaluation section. As a result this evaluation section aims to determine which of the two contour coupling techniques is more robust for cardiac left ventricle segmentation and how the better technique performs in the presence of limited training data.

4.3.1 CC-DD vs. CC-MD

To evaluate the contour coupling methods, a leave-one-out training/evaluation scenario is employed. Specifically the CC-MD and CC-DD methods are tested on the basal-most, mid and apical-most short-axis slice of the first temporal frame in each dataset. This allows the full range of segmentation challenges associated with cardiac left ventricle segmentation to be applied to both methods. In each case the endocardium segmentation result needs only be presented once, as it is common to both methods. Table 4.1 contains the set of segmentation results for the entire evaluation scenario. The first observation that can be made of these results is that, overall, both the CC-DD and CC-MD methods have similar performance characteristics. However, upon closer examination it can be seen that the CC-DD is prone to significant segmentation errors in some cases (Basal: datasets 20 and 30, apical: dataset 28). Upon examination of the segmentation results in these cases it has been determined that the estimation of the likely epicardium location can be prone to miscalculation in cases where the epicardium interface is poorly defined, or when there is significant surrounding image noise. On one hand this is expected, since data-driven methods in general are commonly negatively affected by image artefacts. By contrast the CC-MD method exhibits no such instabilities since the method is, with respect to estimation of the epicardium characteristics, purely data driven. Therefore, it can be determined that the CC-MD is more suited to segmentation of complex imaging scenarios where the segmentation target can be poorly defined or partially obscured. For simpler image segmentation problems where the segmentation target is well defined, the CC-DD method is likely to have equivalent performance to CC-MD. Finally, a selection of segmentations from basal, mid and apical short-axis slices are presented in Fig. 4.13

from the CC-MD segmentation method. These images qualitatively show the range on contour configurations that can be successfully segmented using the proposed approach. Moreover these images show the accuracy with which the endo- and epicardium objective functions capture the target features, despite significant intensity variation and diffuse borders, especially towards apical slices.

4.4 Summary

In this chapter, the issues related with cardiac left ventricle myocardium segmentation have been addressed. Initially, the *endo+epi ASM* modelling strategy was discussed in depth. The central limitations that prohibit this method from use in this thesis are clearly described. The fundamental limitations are then analysed, which leads to the development of two novel contour coupling implementations that greatly alleviate the limitations of the former approach. Indeed, the proposed contour coupling technique can adapt to contour configurations that are *outside* the training set, normally an inherent limitation of model-based segmentation approaches. By considering the endo- and epicardium separately at model-time, which allows expert shape models to be created, and enforcing *a priori* biological constraints at segmentation-time, a robust and flexible multi-contour segmentation method is developed. Two separate contour coupling implementations are presented: one is based on image data at segmentation-time, and the other is model-based created at model-time. These two contour coupling implementations are thoroughly evaluated on real cardiac MRI datasets, and, after analysis of the quantitative results it is determined that the model-based contour coupling technique is more suited to complex image segmentation environments, as it is not influenced by image noise or other unwanted artefacts.

Table 4.1: Evaluation results for the CC-DD and CC-MD contour coupling methods. All results presented are average point-to-curve errors between the segmented contours and the manual segmentations. All measurements are made in pixels.

Dataset	Basal			Mid			Apical		
	Endo	CC-DD	CC-MD	Endo	CC-DD	CC-MD	Endo	CC-DD	CC-MD
1	0.62	1.19	0.97	1.28	1.19	1.00	0.97	1.12	0.88
2	0.91	0.97	0.94	0.53	0.69	0.91	0.62	1.66	0.94
3	0.66	0.84	0.59	0.53	0.66	0.69	0.81	0.69	0.66
4	0.81	1.75	0.97	1.22	1.94	0.47	0.56	1.06	0.81
5	0.62	1.06	0.56	0.78	0.91	0.72	0.59	1.06	1.28
6	0.62	0.75	0.59	0.72	0.59	0.59	1.31	1.03	1.62
7	1.31	1.34	1.38	0.62	0.88	1.16	1.25	1.09	1.31
8	0.72	0.47	0.41	0.72	0.81	0.84	0.97	0.91	0.97
9	0.47	0.66	0.66	0.59	0.72	1.19	1.34	1.38	1.22
10	0.50	0.75	0.69	0.75	0.69	0.78	1.28	0.72	0.94
11	0.50	1.41	1.41	0.47	1.00	0.81	1.00	0.81	0.75
12	0.47	0.59	0.59	0.56	1.50	1.38	0.62	1.50	1.09
13	0.91	0.91	0.75	0.59	0.75	0.66	0.59	1.22	0.59
14	0.62	0.66	0.56	0.81	0.50	0.81	0.72	0.78	0.47
15	0.84	0.72	0.88	0.72	0.81	0.59	0.62	1.88	0.91
16	0.69	1.00	0.75	0.56	0.62	0.59	0.59	1.12	0.78
17	0.81	0.66	0.50	0.81	1.19	0.81	0.75	1.69	0.59
18	0.66	0.62	0.88	0.84	0.72	0.81	0.81	0.91	0.88
19	0.94	0.97	0.69	0.69	0.75	0.53	0.75	1.88	0.91
20	0.53	5.16	1.06	0.56	0.94	1.69	0.75	0.50	0.62
21	0.56	0.91	1.44	0.53	0.75	1.47	0.56	1.19	1.34
22	0.69	0.56	0.88	0.91	1.12	1.69	1.06	1.34	1.56
23	0.59	1.00	0.50	0.81	0.72	0.94	0.66	0.84	0.81
24	0.91	1.44	1.06	0.53	1.69	1.16	0.84	1.69	1.16
25	0.56	0.72	0.62	0.84	0.78	1.03	1.00	0.66	1.19
26	0.84	1.53	1.16	1.19	1.00	0.81	0.78	0.97	1.03
27	1.00	0.56	0.62	1.25	0.97	1.06	0.72	0.72	0.66
28	0.75	0.84	0.88	1.25	1.19	1.31	0.59	7.81	0.88
29	1.56	0.72	1.00	0.72	0.88	1.09	0.59	0.72	1.12
30	0.78	3.84	2.66	0.91	1.09	1.03	0.69	0.78	0.69
31	0.81	0.91	1.06	0.81	0.94	0.97	1.09	1.03	1.72
32	0.72	1.62	1.59	0.50	1.19	1.28	0.38	1.03	1.81
33	1.03	1.16	0.91	0.59	0.47	0.75	0.81	1.31	0.91
Mean	0.76	1.16	0.91	0.76	0.93	0.96	0.81	1.31	1.00
Std Dev	0.24	0.94	0.43	0.24	0.32	0.31	0.25	1.22	0.34

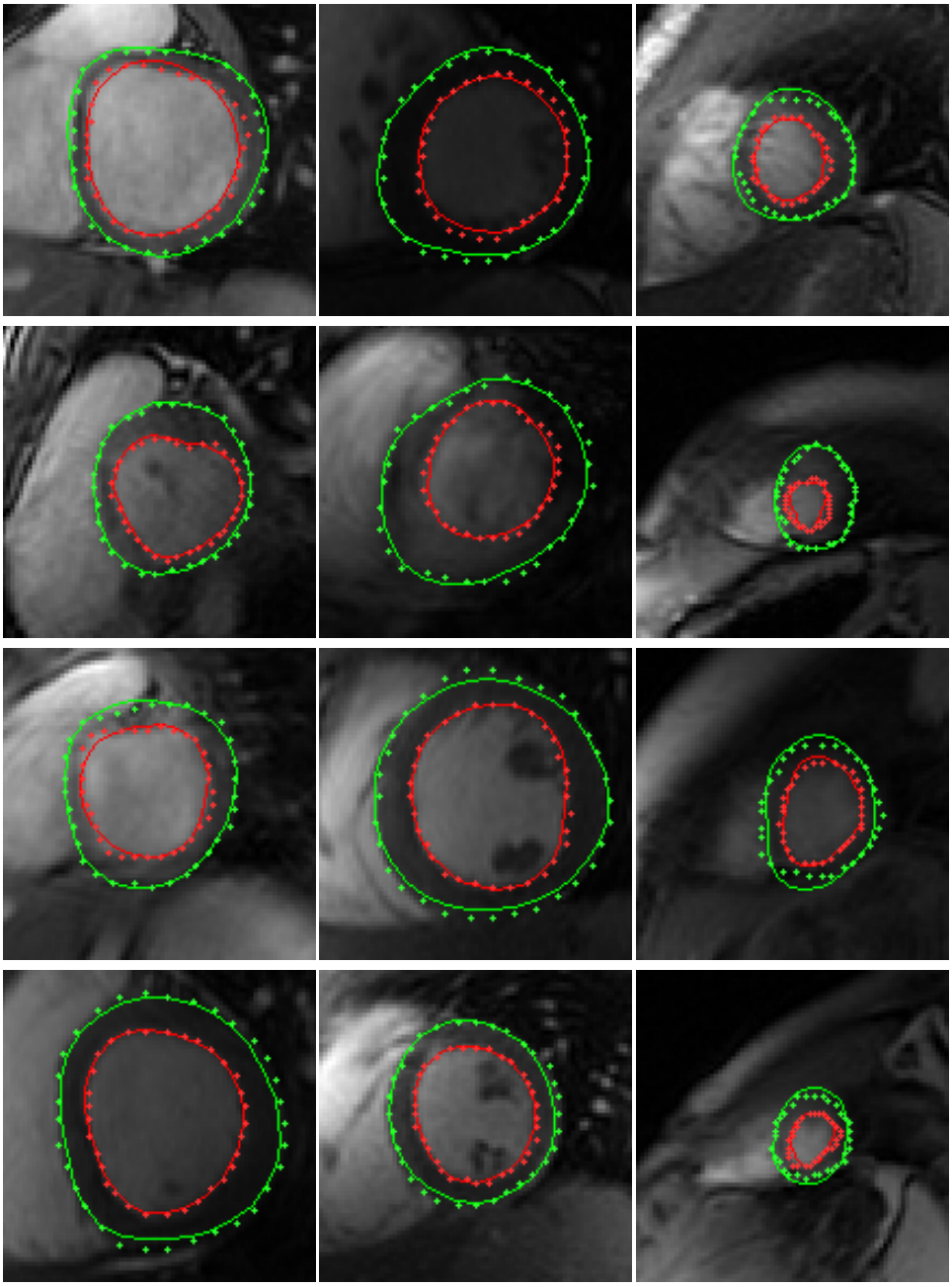


Figure 4.13: A random selection of myocardium segmentation examples using the CC-MD contour coupling method. The left-most column corresponds to basal-most slices, the centre column to mid-level slices and the right-most column to apical most-slices. In each image, the continuous contours represent the automatic segmentations using the proposed ASM optimisation and contour coupling techniques and the discrete contours represent the manual annotations for that short-axis slice.

Chapter 5

Spatio-Temporal Segmentation

Historically, model-based methods were first developed for 2D segmentation challenges. The modelling techniques developed during this period served as the platform for model-based research for the next two decades. The demand for 3D model-based techniques, with the arrival of 3D datasets, was met by extending then-current 2D methods to cope with the complexities of the extra dimension. The difficulties that emerged during this transition were regarding developing point-correspondence between training samples, initialisation challenges, local minima in models, new questions regarding the description of models (landmark-based, volume-based etc.) and the increased demand for training data.

More recently, 3D+time datasets have again raised the demands on model-based methods to capture complex spatio-temporal dynamics. Initially, some research focused on extending existing 3D methods to achieve 3D+time segmentation by using a sequential frame-based segmentation, or targeting a collection of 3D models, each focused on a sub-section of the temporal region. However, the development of true 3D+time models faces several significant practical limitations. Commonly, these issues are referred to as the 'curse of dimensionality'. In essence, (and to summarise the full discussion given in Section 2.3.5) as the dimensionality of the object to be modelled rises, there is a proportional rise in: the required size of training samples,

the complexity of developing point-to-point correspondence between training samples and the number of local minima in the model. Clearly, these are the same issues that faced the transition from 2D to 3D, the difference is that the magnitude of these problems increases dramatically as the dimensionality of the segmentation target grows.

Of these challenges, the increased training requirement is usually the first hurdle that 3D+time model-based methods cannot overcome, due to the scarcity of sufficiently large, and annotated, 3D+time datasets. This is further complicated since, as noted in Section 1.1, manual annotation of 3D+time datasets is very time consuming and prone to inter- and intra-observer variability. This will naturally affect the quality of the manual annotations, and hence the models derived from them. Some research effort has been dedicated towards the development of automatic annotation procedures to alleviate the dependence on manual annotation [38, 88], however the number of training samples required to construct useful models is still prohibitive.

In the case of 3D modelling and segmentation the magnitude of these problems were not prohibitive, but were still noticeable. To summarise Section 2.3.4, Kaus *et al.* [48] exclude datasets exhibiting large levels of variation from their evaluation method. Stegmann *et al.* [97, 98] developed a motion-correction scheme to remove inter-slice translation, so that assumptions regarding 3D geometry of datasets is valid. Both Lorenzo-Valdés *et al.* [60] and Koikkalainen *et al.* [50] are required to artificially enlarge their training sets so that their models encompass sufficient variation. Zhang *et al.* [113] note that their 3D+time model does not contain sufficient variation, and the authors resort to using a 3D model for the final stage in their segmentation pipeline.

Clearly, the difficulties faced by model-based methods has increased dramatically as the dimensionality of the segmentation problems has risen over time. The training requirements for 3D+time modelling, and the other factors related to the curse of dimensionality, hamper the development of 3D+time model-based methods. However, the implicit assumption in the preceding discussion is that models must be constructed at the same dimensionality of their segmentation target. That is to say, 2D models are built

for 2D segmentation, 3D models for 3D segmentation and 3D+time models for 3D+time segmentation. Although this is a natural course to take, it has been shown to be dimensionally limited (in a practical sense) and so alternative methods must be developed. The central focus of this chapter, therefore, is to readdress the *modelling strategy*, in an effort to develop a new approach to multi-dimensional modelling that is dimensionally scalable.

Using a strategy similar to that developed in the preceding chapter, the proposed approach captures shape, spatial and temporal variation in separate models. The hypothesis is that if the temporal variation is learned only from the basal-most slices and the spatial variation only from the end-diastolic volume the information from both models can be combined to segment any other slice in an unseen 3D+time dataset. As a result, only a fraction of a 3D+time dataset is required to be annotated for training compared with the standard *single-model* approach, described at the beginning of this chapter. For example, if a dataset has spatial and temporal dimensions z and t, respectively - the proposed method requires $z+t-1$ slices to be annotated as opposed to $z \times t$ in the single-model approach. This arrangement is shown graphically in Fig. 5.1.

5.1 Construction of Spatial and Temporal Models

This chapter unifies the contributions presented in the preceding two chapters. First, the novel ASM optimisation routine is utilised so that myocardial segmentation will be robust, especially towards apical slices. Second, contour coupling is enforced and separate endo- and epicardium models are created, so that the overall segmentation method will be much more flexible than an *endo+epi* model. Therefore, to achieve 3D+time segmentation of the cardiac left ventricle, six separate models are constructed: a 2D endocardium model, a 2D epicardium model, an endocardium spatial model, an endocardium temporal model, an epicardium spatial model and an epicardium temporal model. In this chapter, a 'spatial' model refers to a model constructed in the long-axis direction (i.e. in the vertical direction in Fig. 5.1). A 'temporal' model is constructed across the cardiac cycle (i.e. in the horizontal direction in Fig. 5.1). The necessity for the 2D endo- and epicardium models will be explained in the following discussion. Of the re-

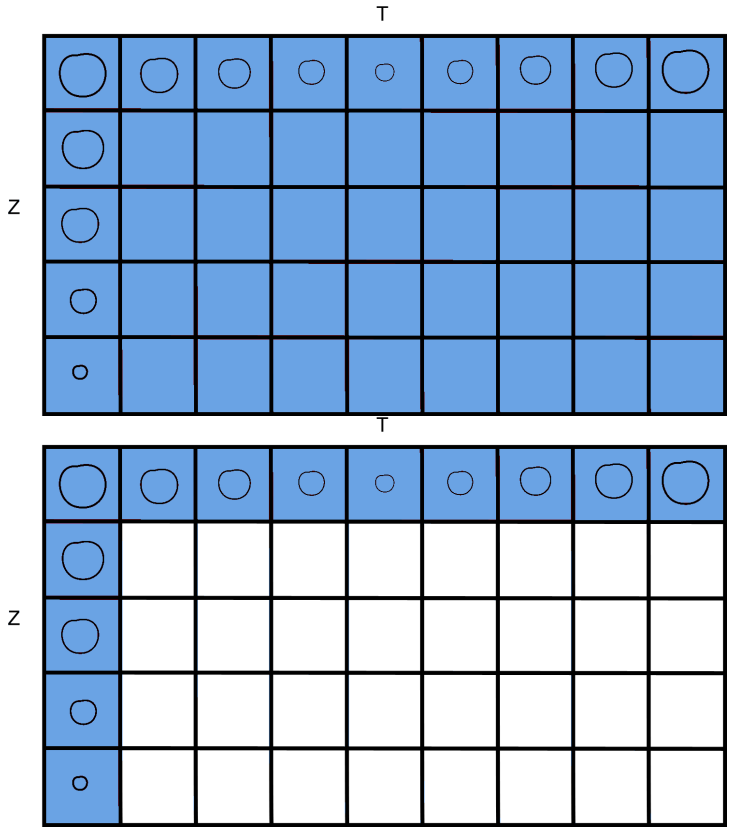


Figure 5.1: The upper image in this figure shows the required training level of a 3D+time dataset for training a *single-model* method. By contrast, the lower image illustrates the significant reduction the proposed segmentation technique affords by separating the spatial and temporal modelling process. In both images, the shaded region refers to the section of the dataset used for training.

maining spatial and temporal models enumerated above, the endocardium is segmented first using its spatial and temporal models, so that it can drive the epicardium segmentation using the proposed contour coupling technique described in the previous chapter. The construction of spatial and temporal models is described with respect to the creation of the endocardium temporal model, but a similar approach is used for the other three cases. Of course, since there are gaps between slices along the spatial axis in MRI images, the accuracy of left ventricle representation is dependant on the quality of MRI acquisition.

To capture the temporal variation, consider the basal-most slice of shape

annotations across the temporal axis as in Equation (5.1), where Y_i is the i^{th} training sample in the training set $\mathbf{Y}$.

$$Y_i = \left[\left[l_1^1, l_2^1, \ldots, l_N^1\right], \ldots, \left[l_1^M, l_2^M, \ldots, l_N^M\right]\right] \quad (5.1)$$

$$Y_i' = [\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_M] \quad (5.2)$$

$$Y_i'' = [\beta_1, \beta_2, \ldots, \beta_M] \quad (5.3)$$

where N is the number of landmarks in a shape, and M is the number of shapes in a single training sample. Since this is a set of 2D shapes, the alignment of each training sample is achieved by removing translation (Y_i'). Point correspondence is again enforced with arc-length re-sampling for each shape in the set (see Section 3.1.1 for details). Here, the 2D endocardium shape model is utilised. As the shape variation is already captured in this 2D ASM, it is redundant to relearn the shape variation described in landmark space in the temporal model, when the goal of the temporal model is to capture the 'macro' variation of the shapes - i.e. the larger sources of variation across the cardiac cycle. Therefore Equation (5.1) can be restated as shown in Equation (5.2) by considering each set of landmarks as a shape. Each shape in the temporal slice can now be mapped from landmark space into β space (as shown in Equation (5.3)), by employing: $\beta \approx \Phi^T (\mathbf{s} - \overline{\mathbf{s}})$.

This creates a compact and abstract slice description in β space that simplifies the description of each training sample. The abstract description of the temporal slice is now focused on the macro-variation of the shapes. This is repeated for each sample in the training set for the endocardium temporal model.

PCA can now be applied to the aligned training set ($\mathbf{Y}''$). This creates a new ASM modelling strategy that is described in Equation (5.4), where the α parameter controls the variation of LV shapes in the temporal axis - thus controlling the *temporal profile* of the model. From this progression it can be seen that the 2D endocardium ASM abstracts shape description from a set of landmarks to a single β vector, which is subsequently used by the endocardium temporal model to abstract temporal slice description from a set of set of landmarks, to a set of β vectors and finally to a single α vector.

A similar process is used to construct a spatial ASM (5.5) where the ξ parameters controls the *spatial profile* of the model. However, while the number of frames do not need normalisation for the temporal model (since all of the evaluation datasets are composed of 20 temporal volumes) there is variation in the number of spatial slices between datasets (ranging from 8-15). The interpolation stage used to normalise the number of spatial shapes is similar to the one used in [6], where an interpolated slice is created by bisecting line segments between corresponding landmarks in adjacent existing slices. Using this approach, all datasets are normalised to 15 long axis slices before training the spatial model.

$$\mathbf{t} \approx \bar{\mathbf{t}} + \Psi\,\alpha \tag{5.4}$$

$$\mathbf{z} \approx \bar{\mathbf{z}} + \Upsilon\,\xi \tag{5.5}$$

where α^{limit} and ξ^{limit} are constructed so that they are constrained according to their corresponding eigenvalues, in a method similar to that detailed in Section 3.1.

5.2 Spatial and Temporal Optimisation

The levels of abstraction used to create the Equations (5.4) and (5.5) can be reversed. As a result, each element of $\mathbf{t}$ and $\mathbf{z}$ (t^k and z^k) is itself a β vector and directly maps into shapes denoted by, α^k and ξ^k, respectively. This characteristic of the proposed segmentation approach allows the reuse of the endo- and epicardium objective functions developed in Sections 3.4 and 4.2.1, respectively, in the optimisation of the spatial and temporal models. The objective functions of $\mathbf{t}$ and $\mathbf{z}$ (g and h, respectively) can therefore be defined as a summation of the corresponding 2D endo- and epicardium (as α and ξ refer to the endocardium model parameters, α_{epi} and ξ_{epi} denote the equivalent epicardium parameters) objective functions, as follows:

$$
\begin{aligned}
g_{endo}\left(\alpha, \mathbf{T}\right) &= \sum_{k=1}^{P} \left| \overline{I}\left(T^k\left(\Omega^{+}_{\alpha^k}\right)\right) - \overline{I}\left(T^k\left(\Omega^{-}_{\alpha^k}\right)\right) \right| \\
&= \sum_{k=1}^{P} f_{endo}\left(\alpha^k, T^k\right) \qquad (5.6)
\end{aligned}
$$

$$
\begin{aligned}
h_{endo}\left(\xi, \mathbf{T}\right) &= \sum_{k=1}^{Q} \left| \overline{I}\left(T^k\left(\Omega^{+}_{\xi^k}\right)\right) - \overline{I}\left(T^k\left(\Omega^{-}_{\xi^k}\right)\right) \right| \\
&= \sum_{k=1}^{Q} f_{endo}\left(\xi^k, T^k\right) \qquad (5.7)
\end{aligned}
$$

$$
\begin{aligned}
&g_{epi}\left(\alpha_{epi}, \mathbf{T}\right) = \\
&\quad \sum_{k=1}^{P} \left[\left| \overline{I}\left(T^k\left(\Omega^{-}_{\alpha^k_{epi}}\right)\right) - \overline{I}\left(T^k\left(\Omega^{+}_{\alpha^k_{epi}}\right)\right) \right| \right. \\
&+ \quad \left. \left| I^{\sigma}\left(T^k\left(\Omega^{-}_{\alpha^k_{epi}}\right)\right) - I^{\sigma}\left(T^k\left(\Omega^{+}_{\alpha^k_{epi}}\right)\right) \right| \right] \\
&= \sum_{k=1}^{P} f_{epi}\left(\alpha^k_{epi}, T^k\right) \qquad (5.8)
\end{aligned}
$$

$$
\begin{aligned}
&h_{epi}\left(\xi_{epi}, \mathbf{T}\right) = \\
&\quad \sum_{k=1}^{Q} \left[\left| \overline{I}\left(T^k\left(\Omega^{-}_{\xi^k_{epi}}\right)\right) - \overline{I}\left(T^k\left(\Omega^{+}_{\xi^k_{epi}}\right)\right) \right| \right. \\
&+ \quad \left. \left| I^{\sigma}\left(T^k\left(\Omega^{-}_{\xi^k_{epi}}\right)\right) - I^{\sigma}\left(T^k\left(\Omega^{+}_{\xi^k_{epi}}\right)\right) \right| \right] \\
&= \sum_{k=1}^{Q} f_{epi}\left(\xi^k_{epi}, T^k\right) \qquad (5.9)
\end{aligned}
$$

where P and Q refer to the number of temporal volumes and spatial slices in the dataset, respectively. The objective functions for g and h are maximised

according to:

$$\max_{|\alpha| \leq \alpha^{limit}} g_{endo}\left(\alpha, \mathbf{T}\right) \tag{5.10}$$

$$\max_{|\xi| \leq \xi^{limit}} h_{endo}\left(\xi, \mathbf{T}\right) \tag{5.11}$$

$$\max_{|\alpha_{epi}| \leq \alpha_{epi}^{limit}} g_{epi}\left(\alpha_{epi}, \mathbf{T}\right) \tag{5.12}$$

$$\max_{|\xi_{epi}| \leq \xi_{epi}^{limit}} h_{epi}\left(\xi_{epi}, \mathbf{T}\right) \tag{5.13}$$

where Equations (5.12) and (5.13) are subject to the coupling constraints similar to (4.7) so that: $\mathbf{s}_{\alpha_{epi}} \subseteq \Omega_{roi}^{g}$ and $\mathbf{s}_{\xi_{epi}} \subseteq \Omega_{roi}^{h}$, respectively, for CC-DD and subject to weighting constraints $w\left(\alpha_{epi}\right)$ and $w\left(\xi_{epi}\right)$ for CC-MD.

Note that for each optimisation routine, the translation $\mathbf{T}$ is composed of individual translation components T_i for each slice (in the same way that t and k are composed of β vectors). This allows each synthesised shape to adapt to possible irregular translations between slices in the dataset.

In order to maximise the benefits of the coupling constraints when segmenting an unseen dataset, maximisation is performed between Equations (5.10) and (5.12), and between Equations (5.11) and (5.13). As indicated earlier in the chapter, the endocardium models are optimised first, which are used to derive the coupling constraints for the subsequent epicardium optimisation.

Fig. 5.2 illustrates the limits of the temporal model with respect to the endocardium scale variation (measured using the L^2 norm) in our evaluation datasets (further details can be found in Section 5.3). Since the temporal ASM is trained only on the basal-most slice of the training data, and as shape size (i.e. scale) is included as part of the ASM, it is necessary to illustrate that the resultant ASM is able to encapsulate the full range of 3D+time scale variation. The shaded area between the model limits (shown as the upper and lower lines) represent the range of shape size in the evaluation datasets from basal to apical regions. As shown in Fig. 5.2, it can be noted that with a reduction in the training level the temporal model still maintains sufficient scale adaptability to capture the full range of LV shape size variation from basal to apical slices. The spatial model exhibits similar characteristics.

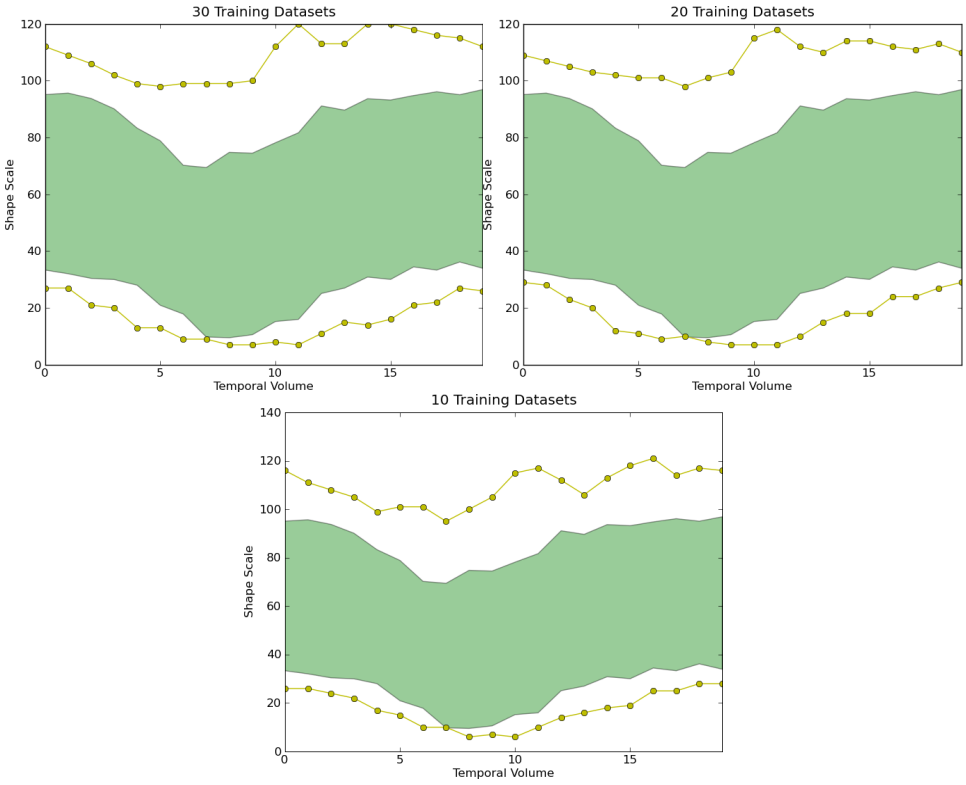


Figure 5.2: Scale variation of the endocardium in the evaluation datasets (shaded region) shown with respect to the upper and lower bound of the learned variation of the temporal model (upper and lower lines). The quantity of model training reduces from 30 to 20 to 10 datasets, from top to bottom.

5.3 Data Description

The datasets used for evaluation of the proposed 3D+time segmentation method have been kindly made public by Andreopoulos *et al.* [6]. The details of these 3D+time cardiac MRI datasets are presented in Table 5.1. All datasets have associated clinically validated annotations of the endo- and epicardium[1]. The papillary muscles and trabeculae carneae are treated as part of the LV blood pool in the endocardium ground truth. All contours were drawn on original short axis images. The majority of datasets contain pathologies (LV hypertrophy for example), however not all pertain to the

[1]It should be noted that ground truth annotations also involve some level of variability. Measurements from repeatability studies [59, 95, 89] indicate mean inter-observer variations of between 5%-8% for left ventricle volume estimation.

Table 5.1: Details of the 3D+time datasets used in evaluation of the method presented in this chapter.

Dataset Details	
Number	33
MR Scanner Type	GE Genesis Signa
Scanner Protocol	FIESTA
Patient Ages	2-17 years
Short axis resolution	256×256 pixels
Long axis resolution	8-15 slices
Temporal resolution	20 volumes
Pixel spacing	0.93-1.64 mm

left ventricle (for example, coarctation of the aorta and arrhythmogenic right ventricular dysplasia).

Direct quantitative comparison of the proposed method and to the method presented in [6] is presented in the following evaluation. As stated earlier, the authors go some way towards the proposed approach by separating spatial and temporal variation. However, since they train a *single-3D AAM* it gives an ideal point of comparison between the proposed approach and the *single-model* strategy. Therefore, while presenting these results it should be noted that the proposed method requires between 83% and 89% less training than the method detailed in [6]. It will be quantitatively shown that comparable or better segmentation accuracy is achieved despite the significant difference in training requirements.

5.4 Testing Protocol

The primary aim in this evaluation is to analyse the performance of the proposed method in scenarios with reduced training. To do this, a range of results is presented with varying levels of training and testing. Specifically, X datasets are used for training the spatial and temporal models models, while the remaining $33-X$ datasets are used for segmentation. This process is repeated 33 times for each value of X. Therefore, as the level of training

decreases the level of testing increases. This removes any bias in the choice of datasets for modelling at each stage. In total, the performance of the proposed algorithm is based on 2178 3D+time dataset segmentations - i.e. 43560 volume segmentations.

Segmentation accuracy is quantified using the volumetric error metrics as described in [6] so the methods can be directly compared. All volumetric measurements are calculated using Simpson's rule [11]. Both mean volumetric point-to-curve error and absolute volume error are employed. Volumetric point-to-curve error is calculated as the average of the minimum distance between each landmark and its corresponding ground truth contour, in a single temporal volume. The absolute volume error is the absolute volume difference between the segmented structure and the corresponding ground truth structure. The proposed approach does not require a complex manual initialisation procedure as in many *single-model* papers [113, 109] where multiple 3D model parameters need to be initialised or estimated from manual segmentations [6]. Instead, only a single xy seed point on the basal-most slice of the end-diastolic cardiac phase, providing an initial translation to localise the LV blood pool in that slice (in fact, this translation is used to initialise each T_i in the temporal translation vector $\mathbf{T}$). This stage could, of course, be automated using a method similar to that described in [46]. As stated in Zambal *et al.* [112], because there is no rigid connection between shapes in the proposed approach and since both β and translation T are optimised, the proposed approach can implicitly compensate for irregular translation between adjacent slices. As a result, no pre-registration phase is required before segmentation, as in [97].

5.5 Quantitative Results

Tables 5.2 and 5.3 provide the quantitative results from the evaluation of the method described in this chapter with reduced training sets - as well as comparison against other approaches presented in the literature. Please note that the endocardium results in Table 5.2 are common for both CC-DD and CC-MD methods in Table 5.3 since contour coupling only affects the epicardium optimisation stage of the segmentation. Primarily comparison is made against Andreopolous *et al.* [6] since the proposed method is eval-

uated on the same datasets. Two sets of results are presented from [6], the 'IC9' method, which uses an inverse compositional optimisation approach (one of their main contributions), and the results from the standard Gauss-Newton optimisation technique. Indicative comparisons are also provided against several contemporary related works [48, 64, 60, 113]. Details of the segmentation protocols used in these papers are presented below.

Kaus *et al.* [48] who use a 3D left ventricle deformable model method. They evaluate their technique on 121 3D cardiac MRI datasets using a leave-one-out scenario. The authors use mean surface-to-surface distance to evaluate the performance of their method. Papillary muscles are considered part of the LV blood pool in ground truth data, which is done on a slice-by-slice basis on the original short-axis images. Also, some datasets were excluded from the evaluation stage due to significant variation.

Lynch *et al.* [64] where the authors use a non-rigid model to encode *a priori* temporal deformations into a level-set formulation. The method is tested on six 3D+time cardiac datasets, and performance is quantified using point-to-curve errors between automatic segmentations and the manual tracings. The clinically-validated manual tracings do not consider the papillary muscles as part of the LV blood pool.

Lorenzo-Valdés *et al.* [60] builds a single 3D+time atlas of the left and right ventricles, as well as the myocardium and background structures. The method is evaluated on 14 3D+time datasets and tested using a leave-one-out scenario. There is no explicit reference to the inclusion, or exclusion, of the papillary muscles as part of the endocardium. Results are presented for the mean and standard deviation of the differences between the manual and automatic segmentations for the three middle slices of all time frames.

Zhang *et al.* [113] use a hybrid ASM/AAM technique to segment 50 3D+time cardiac MRI datasets. Manual annotations are performed on short-axis slices where papillary muscles are considered part of the LV blood

Table 5.2: Endocardium segmentation results. Mean refers to the average volumetric point-to-curve error between the proposed method and manual annotations. Volume refers to the absolute volume difference between the segmentation and manual annotations.

Endocardium Results		
Train - Test	Mean (mm)	Absolute Volume (cm^3)
32 - 1	1.42 ± 0.24	5.19 ± 2.65
30 - 3	1.47 ± 0.16	5.51 ± 1.79
25 - 8	1.50 ± 0.11	5.52 ± 0.77
20 - 13	1.55 ± 0.11	5.74 ± 1.02
15 - 18	1.63 ± 0.11	6.41 ± 1.10
10 - 23	1.87 ± 0.21	7.92 ± 1.68
IC9 [6]	1.43 ± 0.49	6.10 ± 6.44
Gauss-Newton [6]	1.40 ± 0.50	5.76 ± 5.35
Indicative Results		
Kaus [48]	2.28 ± 0.93	–
Lynch [64]	1.24 ± 1.33	–
Lorenzo-Valdés [60]	2.21 ± 2.22	–
Zhang [113]	1.69 ± 0.38	–

pool. Their method is evaluated using a hold-out strategy. Performance is quantified using mean surface-to-surface positioning errors between manual and automatic segmentations. In this chapter, the average of their reported results for normal and pathological datasets is used as a point of comparison.

Fig. 5.3 provides a qualitative example where the segmentation results differ between the CC-MD and the CC-DD methods. The evaluation also includes regression values for the endo- and epicardium. Figs. 5.4 and 5.5 show the regression graphs for the endo- and epicardium, respectively, as training reduces. Table 5.4 shows the regression values for the full range of testing. Comparison is provided against the recent variational approach of Ayed *et al.* [10] and as well as the previous work from Lynch *et al.* [63], in Table 5.5.

Fig. 5.6 illustrates the temporal performance of the proposed segmentation method. This figure shows comparison between both contour coupling methods and the endocardium segmentation accuracy with reducing training sets. Further to the results presented in Section 4.3, comparison is also given between the CC-DD and CC-MD methods with respect to regression characteristics in Table 5.6.

Finally, several qualitative examples of segmentations are shown using the proposed method, from a random selection of datasets and slices within those datasets. Three collections of images are shown (all using the CC-MD method) for 30, 20 and 10 training samples respectively (Figs. 5.7, 5.8 and 5.9). These images qualitatively illustrate the robustness of the proposed approach, with limited training, in the presence of significant anatomical variation, as well as variation in image quality and intensity.

5.6 Discussion

The initial observation that can be drawn from Tables 5.2 and 5.3 is that the proposed method shows excellent stability as training is reduced. This is a reflection on the flexibility of the method due to the use of separate endo- and epicardium models, as well as the application of the contour coupling technique. As the size of the training set is crucially important for *single-model* LV segmentation schemes, it is safe to say that these approaches will perform significantly worse than the proposed method as the amount of training data is reduced. It should be remembered that the datasets being segmented in this case exhibit significant variation, as they are captured from patients ranging from $2 - 17$ years of age.

With respect to the endocardium results, the proposed method achieves a mean volumetric point-to-curve error between the IC9 and Gauss-Newton methods from [6] but also maintain consistently tighter standard deviation measures. This is further reflected with the absolute volume errors in Table 5.2 that show more accurate volume segmentations with significantly less absolute volume deviation. This reduction in standard deviation in both the point-to-curve and volume errors illustrate that this approach generates more consistent segmentations and is less likely to fit in spurious image areas.

Table 5.3: Epicardium segmentation results. Mean refers to the average volumetric point-to-curve error between the proposed method and manual annotations. Volume refers to the absolute volume difference between the segmentation and manual annotations.

Epicardium Results		
CC-DD		
Train - Test	Mean (mm)	Absolute Volume (cm^3)
32 - 1	1.78 ± 0.45	10.78 ± 6.81
30 - 3	1.80 ± 0.20	11.48 ± 4.40
25 - 8	1.72 ± 0.13	10.10 ± 2.05
20 - 13	1.82 ± 0.11	10.84 ± 1.39
15 - 18	1.84 ± 0.09	11.37 ± 1.45
10 - 23	1.98 ± 0.13	12.26 ± 1.69
CC-MD		
Train - Test	Mean (mm)	Absolute Volume (cm^3)
32 - 1	1.63 ± 0.38	10.19 ± 7.23
30 - 3	1.66 ± 0.19	11.08 ± 4.15
25 - 8	1.69 ± 0.10	10.74 ± 1.76
20 - 13	1.72 ± 0.09	11.54 ± 1.82
15 - 18	1.76 ± 0.10	11.93 ± 1.53
10 - 23	1.91 ± 0.17	14.07 ± 3.04
IC9 [6]	1.51 ± 0.48	9.82 ± 8.97
Gauss-Newton [6]	1.62 ± 0.85	13.68 ± 15.19
Indicative Results		
Kaus [48]	2.62 ± 0.75	–
Lynch [64]	–	–
Lorenzo-Valdés [60]	2.99 ± 2.65	–
Zhang [113]	1.89 ± 0.49	–

It is also observed that the ASM optimisation method performs comparably with the advanced inverse compositional algorithm (IC9) presented in [6].

Epicardium segmentation is expected to be slightly less accurate than the endocardium since it is the more diffuse border of the two. Moreover trying

to globally optimise the epicardium can be difficult where there is little or no gradient between the myocardium and background structures. Although both the CC-MD and CC-DD coupling techniques have already been evaluated in Chapter 4, the segmentation scenario in that case was only regarding 2D segmentation, i.e. no spatial and temporal information was considered in the segmentation process. For this reason, it is deemed necessary to evaluate the performance of both methods in a 3D+time scenario to further investigate their relative segmentation behaviour. The CC-DD and CC-MD methods both show similar levels of stability as the training data is reduced, however it is observed that the CC-MD method is consistently more accurate than CC-DD. Examination of the results indicates that this is due to the difficulty in robustly estimating the probable location of the epicardium using gradient information, especially towards the apical slices. Fig. 5.3 illustrates two examples where there is a deviation between the CC-MD and CC-DD techniques. In both images (which are particularly challenging to segment) the epicardium gradient in the image is poorly defined, which negatively affects the CC-DD ROI derivation and resulting segmentation (shown in red). By contrast, the CC-MD segmentation correctly derives the ROI from the *a priori* model (see Fig. 4.11), allowing for a successful epicardium optimisation (shown in green), with manual annotations shown as blue crosses. From this, the evaluation section of the previous chapter, and the following experimental evidence, it can be concluded that the CC-MD method is more appropriate for LV segmentation. Consequently, in the remainder of this section all epicardium results are presented for the CC-MD method, unless otherwise indicated.

Our leave-one-out result for the epicardium is just outside the mean accuracy of the Gauss-Newton method, again however, the standard deviation of the proposed approach is considerably lower than both the IC9 and Gauss-Newton results, highlighting the consistency of segmentation using separate spatial and temporal models. The difference of $0.1mm$ in mean volumetric error between the IC9 result and the proposed method, over 660 volume segmentations, is considered to be very small. Again, based upon the discussion throughout this and previous chapters, it is expected, that in a scenario with reduced training, the proposed method would significantly out-perform *single-model* segmentation schemes. Attention is also brought to a compar-

Table 5.4: Correlation coefficients between manual and automatic volumes.

Volume Regression Results		
Train - Test	Endocardium	Epicardium
32 - 1	0.98	0.99
30 - 3	0.98	0.99
25 - 8	0.99	0.99
20 - 13	0.98	0.98
15 - 18	0.98	0.98
10 - 23	0.97	0.97
IC9 [6]	0.95	0.97

ison (albeit indicative) against the results presented by Lorenzo-Valdés *et al.* [60]. Their approach can be considered the antithesis of the modelling strategy described in this chapter. The experimental results indicate that the proposed method compares very favourable against theirs, especially considering the fact that they test purely on a leave-one-out basis to show their method in a best-case scenario. It should also be noted that their models require significant Gaussian blurring to improve model variation.

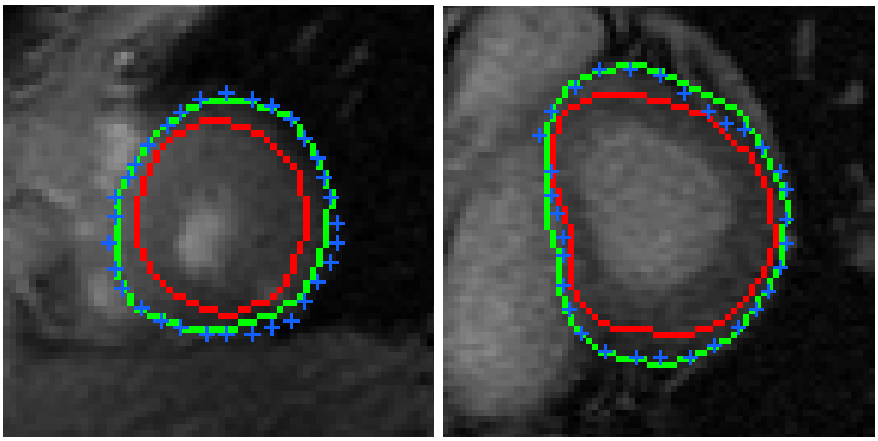


Figure 5.3: Two qualitative comparisons between the CC-DD and CC-MD coupling techniques. The epicardium optimisation result using CC-DD is shown as a continuous red contour. The equivalent contour using CC-MD is shown as a continuous green contour, and manual annotations shown as blue crosses.

The results in Table 5.5 show that the proposed method compares favourably

Table 5.5: Correlation coefficients between manual and automatic areas.

Area Regression Results		
Train - Test	Endocardium	Epicardium
32 - 1	0.97	0.95
30 - 3	0.96	0.94
25 - 8	0.96	0.94
20 - 13	0.95	0.93
15 - 18	0.95	0.92
10 - 23	0.91	0.88
Indicative Results		
Ayed [10]	0.94	0.96
Lynch [63]	0.82	0.89

with state of the art techniques and it should be noted that both papers [10, 63] aim for pixel-level segmentations, a scenario that favours the inclusion of the papillary muscles in the analysed 3D+time data.

In Fig. 5.6, with respect to endocardium performance, it is observed a general reduction in accuracy around the end-systolic phase. This trend is repeated in the CC-MD and CC-DD results. As expected, the errors associated with the end-systolic volume are larger since the LV volume is minimised at this cardiac phase. Consequently, the impact of errors induced during shape synthesis with respect to ground truth will have a greater impact on the overall segmentation accuracy than in situations close to the diastolic phase of the cardiac cycle. It is useful to note that the downgrade in performance for temporal volumes close to the systolic phase is a problem for all techniques that rely on *a priori* learned models, regardless of whether the segmentation is carried out using pixel-based (level-set) or shape-based (ASM) segmentation strategies. Unlike [6], no 'spikes' in errors are observed at certain temporal frames, as the proposed method does not switch between multiple temporal models during segmentation. The second observation is that for the three cases presented in Fig. 5.6 (i.e. endocardium, CC-DD and CC-MD segmentations) there is a slight degradation of accuracy as training reduces, corresponding to the reduction in ASM variation. At 10 training datasets however, there is a jump in error (also reflected in Tables 5.2 and 5.3), com-

mon to all three results, that suggests there is a minimum training limit, under which the trained ASMs are unable to robustly capture the full range of 3D+time cardiac deformation. It is also encouraging to observe the stability of the proposed method given only 15 training samples, especially since only $z + t - 1$ short-axis slices from each sample are required. It is believed that a *single-model* ASM would exhibit considerable instability as training reduces for the theoretical reasons outlined throughout this, and preceding, chapters.

The results in Table 5.6, in general, support the performance characterisation given by the point-to-curve volumetric errors and the temporal performance analysis. Of note is that the CC-MD method exhibits far more accurate intercept values compared with CC-DD, as well as corresponding R-values. Again, this is due to the apical errors in the CC-DD method where gradient information can be diffuse or absent. However, it can be noted that at 10 training sets, CC-MD performs slightly worse than CC-DD, which can be attributed to the reduced training affecting the sample size of the L^2-norm model for CC-MD.

5.7 Summary

This chapter has dealt with the 3D+time segmentation of cardiac left ventricle data. In a continuation of the modelling strategy presented in the preceding chapter, the shape, spatial and temporal sources of variation are captured separately in the proposed segmentation approach. One of the notable benefits of this approach is the significant reduction in required training data. Comprehensive quantitative evaluation against a *single-model* technique presented in the literature (and evaluated on the same datasets) proves that the proposed method does not suffer from any loss of segmentation accuracy despite requiring between 83% and 89% less training. The natural conclusion that can be drawn is that a large majority of the full $z \times t$ training set is redundant in its description of the spatio-temporal variation of the cardiac left ventricle. Moreover, the proposed method has made use of two previous contributions presented in this thesis, i.e. the ASM optimisation routine and the contour coupling method. The employment of these methods further enhances the flexibility and accuracy of the proposed

Table 5.6: Comparison of regression results between CC-MD and CC-DD

Epicardium Volume Regression Results			
CC-DD			
Train - Test	Gradient	Intercept	R-value
32 - 1	0.95	7.24	0.92
30 - 3	0.94	8.32	0.95
25 - 8	0.96	5.51	0.98
20 - 13	0.95	6.60	0.98
15 - 18	0.95	6.15	0.98
10 - 23	0.95	7.25	0.98
CC-MD			
Train - Test	Gradient	Intercept	R-value
32 - 1	0.96	0.51	0.99
30 - 3	0.95	0.23	0.99
25 - 8	0.96	0.89	0.99
20 - 13	0.94	3.06	0.98
15 - 18	0.93	5.66	0.98
10 - 23	0.89	8.97	0.97

3D+time segmentation technique. Finally, indicative comparison against several state-of-the-art works in the literature further solidifies the validity of the present contributions in this thesis for use in practical segmentation challenges.

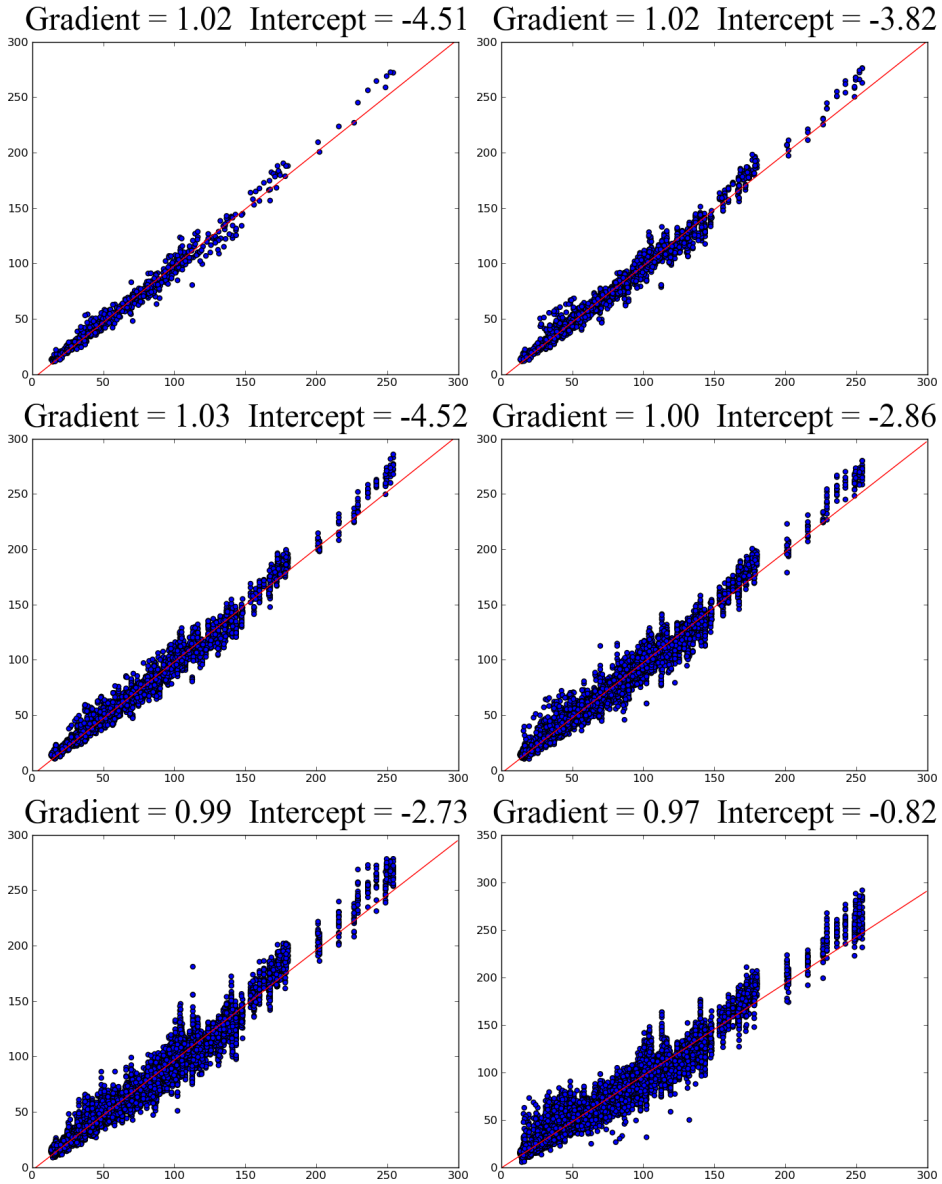


Figure 5.4: Volumetric Endocardium Regression Plots. All X and Y axes are manual volumes and automatic volumes, respectively. Both axes are measured in cm^3. Reading from left-to-right and from top-to-bottom the images correspond to evaluation using: 32, 30, 25, 20, 15 and 10 training datasets, respectively. Corresponding correlation coefficients are presented in Table 5.4.

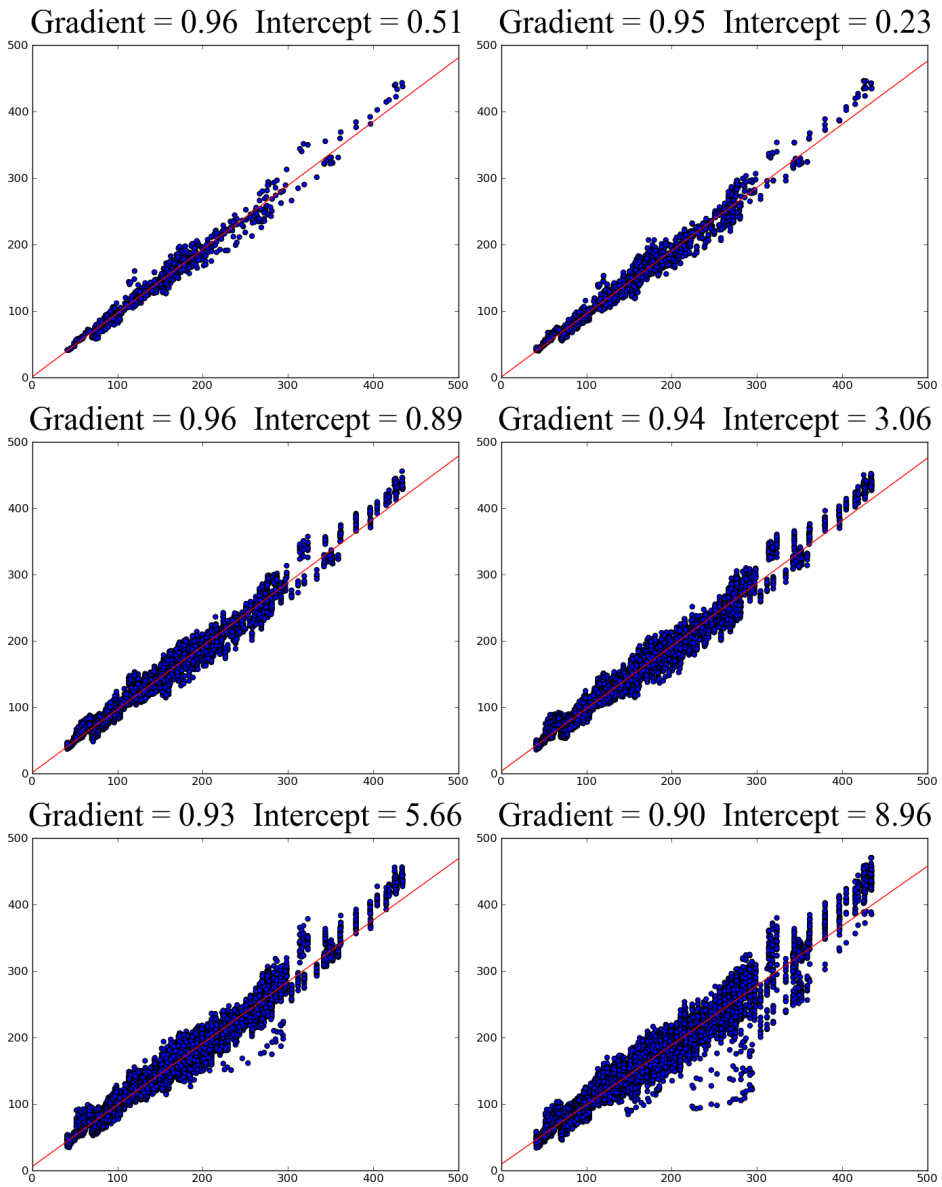


Figure 5.5: Volumetric Epicardium Regression Plots. All X and Y axes are manual volumes and automatic volumes, respectively. Both axes are measured in cm^3. Reading from left-to-right and from top-to-bottom the images correspond to evaluation using: 32, 30, 25, 20, 15 and 10 training datasets, respectively. Corresponding correlation coefficients are presented in Table 5.4.

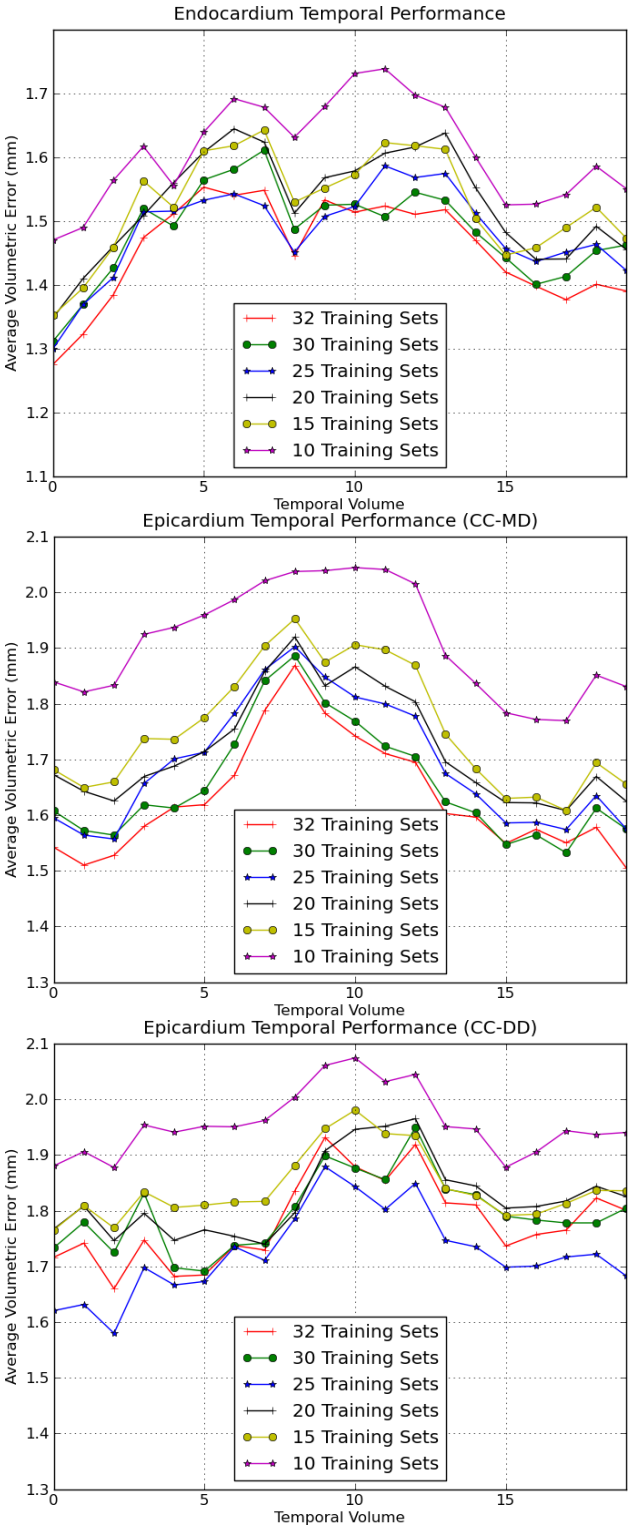


Figure 5.6: Mean Volumetric Error for both endocardium and epicardium shown with respect to temporal volume. Each graph shows results for 32, 30, 25, 20, 15, and 10 training sets, respectively.

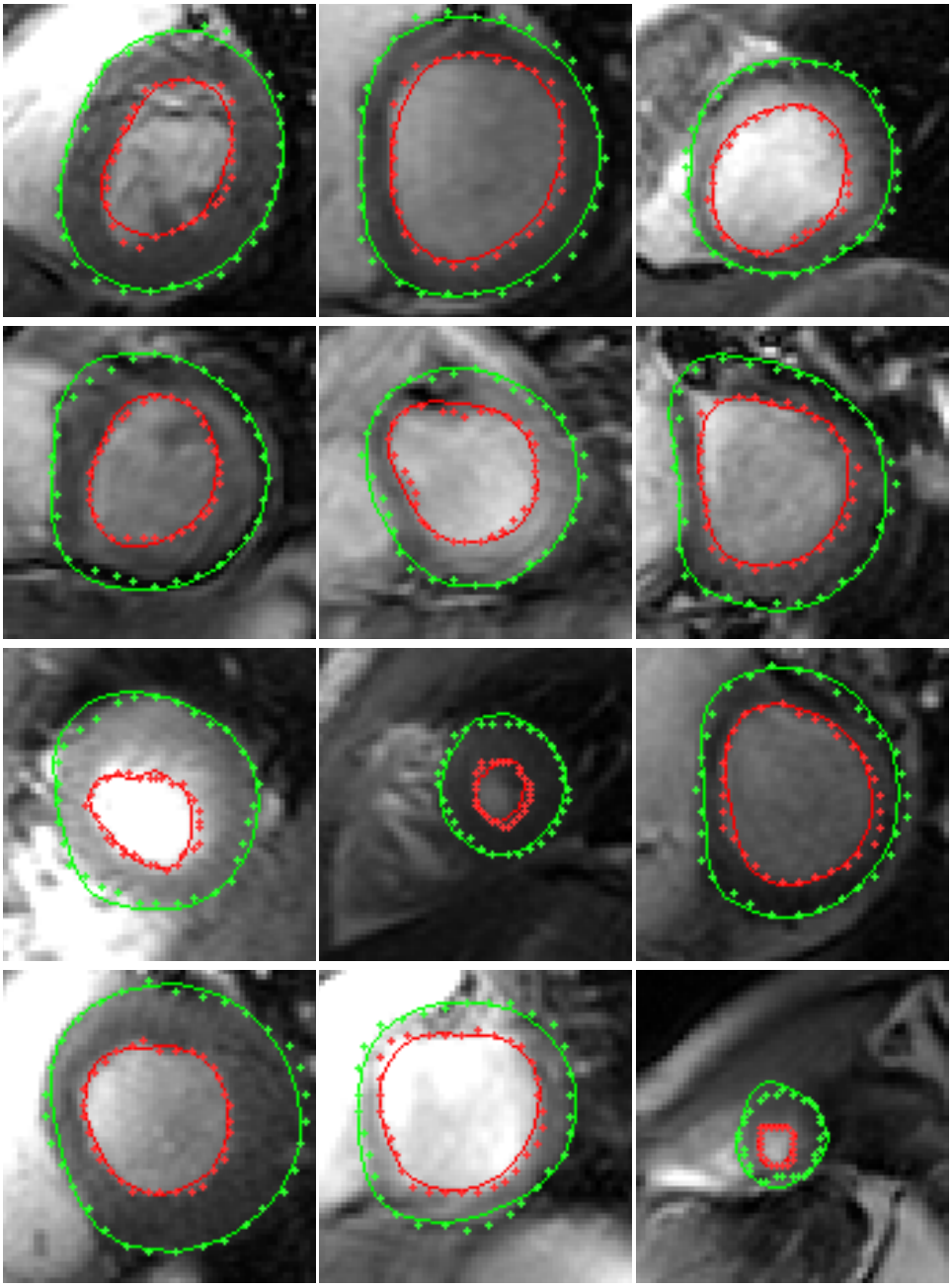


Figure 5.7: Sample segmentations using 30 training datasets. Manual annotations, shown as crosses, are colour-coded with respect to their associated contours.

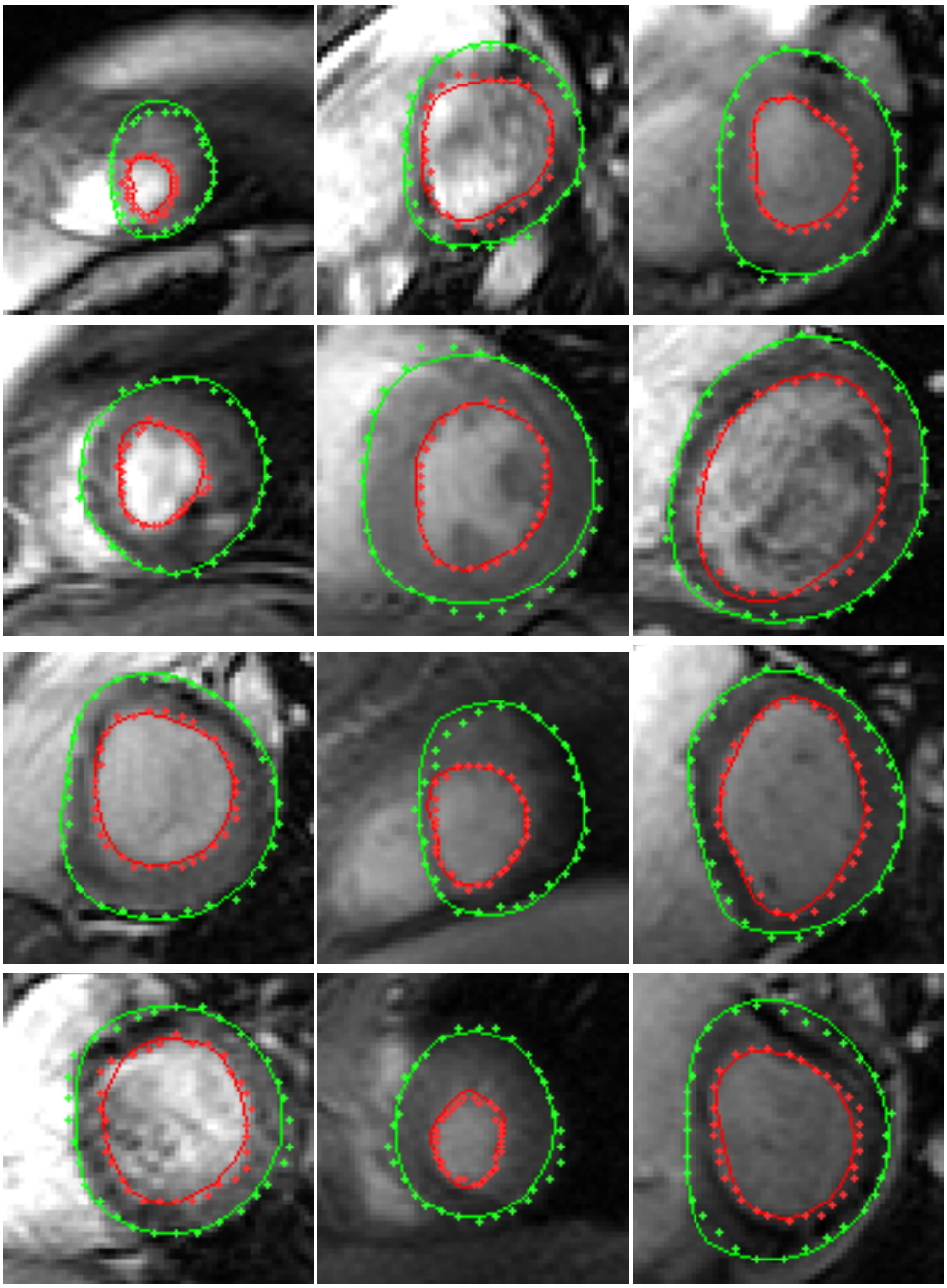


Figure 5.8: Sample segmentations using 20 training datasets. Manual annotations, shown as crosses, are colour-coded with respect to their associated contours.

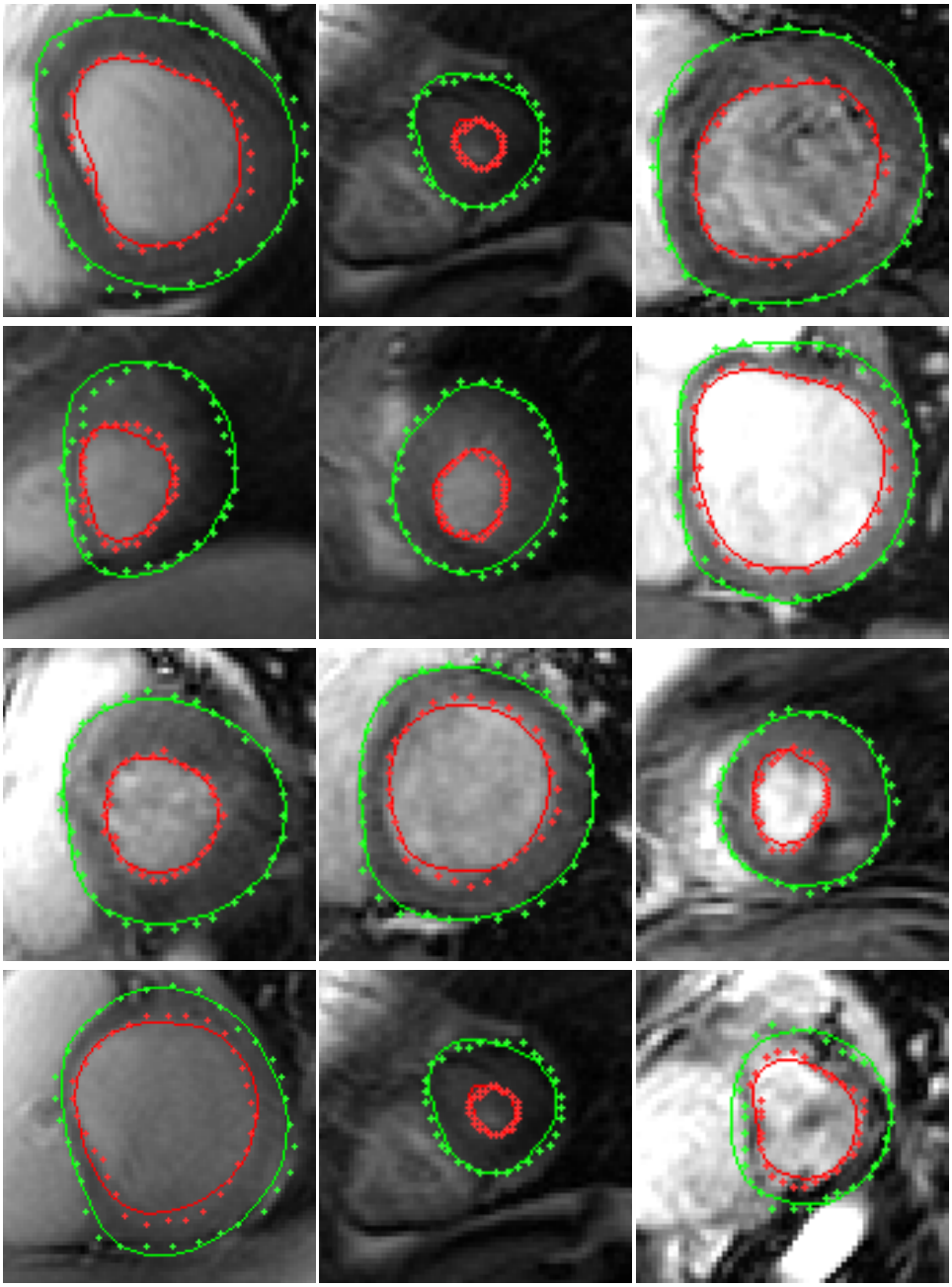


Figure 5.9: Sample segmentations using 10 training datasets. Manual annotations, shown as crosses, are colour-coded with respect to their associated contours.

Chapter 6

Conclusions and Future Work

This thesis has dealt with the application of model-based methods to multi-dimensional cardiac left ventricle segmentation. The motivation for this research is derived from the difficulty faced when attempting to manually annotate a multi-dimensional dataset in the clinical setting. Specifically, this task is complicated by: operating on a slice-by-slice basis, diffuse borders between target structures as well as inter- and intra-observer variability. Automatic segmentation techniques provide a promising solution to this issue. In general, automatic medical image segmentation methods can be categorised as either data-driven or model-based. That is, a segmentation technique can rely purely on the image-data provided, or it may include *a priori* knowledge about the segmentation target into the segmentation process. In general, data-driven methods are inclined to be more susceptible to over-segmentation and over-sensitivity to image noise in complex imaging environments. Although it is possible to include *a priori* information into modern data-driven methods, like the level-set technique, it is difficult to guarantee biologically-valid segmentation using data-driven methodologies.

In this thesis, a model-based method is chosen as the general framework to approach the problem of robust 3D+time cardiac left ventricle segmentation. It is believed that the maximisation of prior knowledge into the solution is beneficial since cardiac MRI datasets often contain challenging

image-artefacts that make segmentation non-trivial. Out of the range of model-based methods, the Active Shape Model is employed as the primary means of learning shape information from a training set of images. Detailed analysis of the literature reveals that the ASM, and other state-of-the-art model-based methods, suffer from limitations when constructing multi-contour 3D+time models. These limitations, collectively known as the curse of dimensionality, also affect 3D model-based methods, but to a lesser degree. One of the central limitations, identified in this thesis, to the adoption of 3D+time modelling techniques is the prohibitive number of training samples required to create general segmentation models with sufficient shape variation. To address this, and other issues, this thesis is structured in the logical order problems are dealt with when developing a model-based 3D+time left ventricle segmentation technique. In this way, the complexity of the problems being addressed increases as the thesis progresses, allowing analysis and contributions to be accumulated throughout the thesis. The first area of focus is on the robustness of the ASM as a segmentation technique. Detailed analysis of the performance of the standard ASM formulation is provided and its main limitations are described.

The first contribution, presented in Chapter 3, is the introduction of a novel ASM optimisation routine that replaces the standard ASM fitting method. Of note, the proposed optimisation routine is not negatively influenced by image noise or local features, such as the papillary muscles, when applied to the problem of 2D endocardium segmentation. This is because the contour fitting score is evaluated on the quality of fit of the entire contour, as opposed to sparse 1D profiles in the standard ASM formulation. As such, the optimisation routine relies of the construction of an objective function, which describes the maximisation criteria for the segmentation target. A full evaluation of the proposed optimisation technique compared to the standard ASM fitting formulation is performed. Experimental evidence supports the theoretical benefits of the proposed method, which exhibits significantly increased segmentation accuracy and stability in a complex imaging environment.

The next challenge addressed in Chapter 4 is that of myocardial segmentation. Again, in this chapter, the problem is addressed in the 2D domain. Here both the endo- and epicardial interfaces require delineation. Initially,

critical analysis is performed on the standard method of addressing myocardial segmentation. Usually, the endo- and epicardium are considered and modelled as a single shape. Although intuitive, this approach is shown to have far-reaching negative consequences on the flexibility of the method. In essence, the implicit assumption in this modelling approach is that each endocardium training sample can only appear with its associated epicardium training sample (and vice versa). This, of course, is not an accurate reflection of the relationship between the two contours and, as a result, the range of allowed variation using this approach is unnecessarily constrained.

The second contribution presented in this thesis comprehensively solves this issue. The central novelty is that the endo- and epicardium are modelled separately at model-time, but coupled using *a priori* constraints at segmentation-time. In contrast to the combined shape model, this allows any endocardium shape to appear with any epicardium shape seen in training (and vice versa) while the inter-contour structural integrity is maintained. A crucial benefit of this contour coupling approach is that contour configurations outside the training set can be correctly segmented. Two contour coupling implementations are described and evaluated on cardiac MRI datasets through myocardial segmentation: one data-driven and one model-driven. Since the proposed ASM optimisation routine is utilised in this evaluation process, the epicardium objective function is introduced. It is shown that the proposed contour-coupling technique shows excellent levels of stability and accuracy for myocardium segmentation. Of the two implementations proposed, the model-driven contour coupling technique is experimentally proven to be more robust. This is especially true towards apical slices, where the endo- and epicardium interfaces are commonly diffuse leading to incorrect segmentation using the data-driven method.

Using the contributions and insight developed in Chapters 3 and 4, the challenge of 3D+time modelling and segmentation of the cardiac left ventricle is addressed in Chapter 5. Here, an alternate modelling strategy is developed that avoids the issues related to the curse of dimensionality. Shape, spatial and temporal variation are modelled separately, but are unified at segmentation-time. The spatial and temporal models are constructed utilising the abstraction gained from the shape model (i.e moving from landmark-space to β-space), which allows for a compact description of the spatial and

temporal models. Moreover, the optimisation of the spatial and temporal models directly leverage the proposed ASM optimisation technique, which illustrates the modularity of the contributions presented in this thesis. As a result, the endo- and epicardium objective functions are re-used to deliver robust and accurate segmentation in a multi-dimensional scenario. This approach is shown to require significantly less training than a *single-model* method that attempts to model the 3D+time left ventricle structure as one entity. Contour coupling plays a significant role in maximising the flexibility and adaptability of the segmentation method to evaluation datasets that exhibit significant image and anatomical variation. This integration of the contributions presented throughout this thesis, represents the culmination of the alternate modelling strategy that has been developed through Chapters 2 to 5.

The proposed 3D+time segmentation method is then placed under a rigorous set of evaluation scenarios to determine its performance in ideal, and non-ideal training and segmentation environments. The experimental evidence strongly supports the initial hypothesis that it is indeed possible to capture the spatio-temporal deformation of the cardiac left ventricle in 3D+time using only the first temporal frame and basal-most slice from each training set. This supports the underlying belief that part of the full 3D+time training set is redundant in its description of the spatio-temporal deformation of the cardiac left ventricle. Moreover, the stability of the proposed segmentation method is conclusively shown through the repeated evaluation with reducing levels of training. This encouraging performance strongly supports the concepts put forward throughout this thesis: that the combination of separate targeted models can yield superior segmentation performance, increased practicality with respect to required training levels and even flexibility to adapt to cases outside the training set.

6.1 Future Work

The concepts developed in this thesis are experimentally evaluated exclusively on the application to cardiac left ventricle segmentation. Throughout the document, substantial effort has been directed towards the segregation of the underlying theory from the application domain to ensure that the con-

tributions outlined are not depicted as application-specific advancements. However, it is clear that the presented contributions require application in other segmentation domains to experimentally validate their generalisability. To this end, some current research into the applicability of the ASM optimisation and contour coupling contributions in other segmentation domains is presented in Appendix B.

From this author's point of view, the theory and application are equally important, and, as such, the avenues of future research can be classified into two categories: those directed towards the further advancement of the proposed multi-dimensional cardiac left ventricle solution and those directed towards the improvement in theoretical sophistication of the major contributions presented in this thesis. Of the two, further investigation into theoretical sophistication of the contributions is deemed to be the preferential option to explore after this thesis is complete, as this will naturally lead to superior segmentation performance in practice.

First, the future work directly related to cardiac left ventricle segmentation is presented. This is followed by broader goals, based on the contributions of this thesis.

6.1.1 Regarding Multi-Dimensional Left Ventricle Segmentation

Although significant progress has been made in this thesis towards a dimensionally-scalable model-based segmentation framework, several interesting avenues of future research remain to be explored and exploited. First, a reassessment of the inclusion of scale as part of the shape variation is required. This is especially relevant in the case of 3D+time modelling and segmentation, where learning scale only from training cases could over-constrain the resulting models. Exclusion of scale as part of direct shape modelling will afford more flexible models, with increased capture range.

Second, investigation into more sophisticated optimisation routines could greatly enhance the accuracy of the ASM segmentation method. Although sequential β parameter optimisation through Brent's method (further detailed in Appendix A) has been experimentally proven to be superior to the

standard ASM fitting technique, it is felt that more advanced techniques could be exploited for this task. One could also explore the application of objective functions that target different areas of the ASM contour. This could be useful, for example, in the improvement of the epicardium objective function, whose contour is generally composed of the myocardium/right ventricle interface and the interface between the myocardium and the background structures in the thoracic cavity. Some related work, such as that of Ordas *et al.* [77] could be of benefit in this approach. Another advantageous area is the application of the ASM approach to open contour segmentation. These segmentation problems are not feasible with more data-driven methods such as the level-set, active contour or active surface, as they depend on a closed contour in their formulation.

With respect to contour coupling, it is believed that investigation into the simultaneous optimisation of multiple contours could be greatly beneficial. This would remove the dependence on a successful endocardium segmentation to correctly segment the epicardium. Some related work in the literature could be exploited to this end, such as the work of Li *et al.* [55]. A natural extension of the research in this thesis would be the coupled segmentation of the left and right ventricles. It would be interesting to quantify the performance of the coupled method with one that considers the endocardium, epicardium and right ventricle interfaces as a single shape, such as the work presented by Fritz *et al.* [40].

Regarding the spatio-temporal modelling and segmentation framework for 3D+time segmentation, a primary area of future research could be direct toward the method of unification of the separate models during segmentation time. In this thesis, a maximisation is performed between the spatial and temporal segmentation of each slice. Furthermore, the flexibility of the spatial and temporal modelling process can be also redefined. While not crucial in left ventricle segmentation, where the number of temporal slices is constant between datasets and a simple interpolation scheme is used to normalise the number of long-axis slices, in other applications this approach may not be feasible. Therefore, re-evaluation of the implementation of the proposed modelling strategy for multi-dimensional modelling and segmentation would be a useful area of further research.

Finally, the application of further biological constraints on multi-dimensional segmentation could further enhance the stability of the proposed approach. For example, it is known that the left ventricle mass stays approximately constant across the cardiac cycle. Therefore, integration of this, and other, knowledge into the segmentation process can further enhance the stability and utility of the proposed multi-dimensional segmentation framework.

6.1.2 Further Advancement of Contributions

One of the primary methods of illustrating the generalisability of a theoretical contribution is in its application to a range of practical segmentation problems. As stated earlier in this chapter, the contributions in this thesis have been evaluated in only a single application domain. A primary motivating factor for future research is, therefore, the evaluation of the presented contributions to other medical and non-medical applications. As mentioned above, some initial research has already been directed towards this goal, and is detailed in Appendix B, where the ASM optimisation and contour coupling concepts are applied to hand segmentation. This appendix illustrates the applicability of the contour coupling approach to segmentation problems involving three and four objects in an image. Moreover, comparison is made against a combined modelling approach and the experimental evidence strongly supports the benefits of contour coupling previously highlighted in Chapter 4. In essence, this appendix supports the utility and flexibility of the contour coupling contribution in multi-object segmentation problems.

Continuing in this vein, future work will be directed towards framing the benefits of spatio-temporal segmentation and ASM optimisation in other application domains. The application of the proposed ASM optimisation routine on a wider range of segmentation challenges is a promising area of exploration. Since the definition of the objective function is the only application-specific component of the proposed approach, it is felt that this method could greatly enhance the performance of the ASM in a more general set of segmentation tasks. Indeed, as mentioned earlier in this thesis, it is potentially possible to learn the objective function automatically during the training process to automate this aspect, which could be useful in scenarios where the definition of segmentation target is not straightforward or

intuitive. With regards to spatio-temporal segmentation, other modalities for cardiac imaging, such as ultrasound could initially be explored. Here, the segmentation of the endo- and epicardial borders is significantly more challenging than MRI and would likely require increased sophistication of the objective function definition and as well as the optimisation routine. However, the benefit of global contour fitting using the ASM optimisation technique presented in this thesis, would certainly enhance segmentation robustness.

More generally, further investigation into the simultaneous optimisation of pose and shape parameters should be considered. For example, a coarse-to-fine search where initial pose and, say, the primary eigenvector of the shape model are used to localise the segmentation target in the image could be used before a more sensitive optimisation of shape parameters takes place. This could also alleviate some of the classical initialisation limitations associated with the ASM.

Towards a broader goal, the development of a technique where a model-based method controlling the allowable range of shape deformation that is intimately connected with a data-driven technique, say a level-set formulation, would be an ideal balance between the two opposing approaches. The dimensionally-scalable modelling approach developed in this thesis could constrain the evolution of a level-set method in a multi-dimensional environment. This would maximise the use of both the image data, through the data-driven method, and the *a priori* knowledge, through the model-based technique. For example, a dynamic weighting between the two methods could be controlled by the proximity of the currently synthesized shape to the limits of allowed variation. When the segmentation evolves outside these limits, preference could be given to the data-driven method.

Appendices

Appendix A

Brent's Optimisation Method

Brent's optimisation method [14] determines the minimum of a function, given a bracketed interval, in a robust manner. Briefly, the algorithm can be summarised in the follow steps:

1. Begin with a function f, and three points a, b and c such that $a < b < c$ and $f(a) > f(b) < f(c)$.
2. Interpolate the three points defined at $f(a)$, $f(b)$ and $f(c)$ using a parabola.
3. **IF** the minimum (or maximum, depending on the optimisation requirements) of the parabola is within the interval, use this point as an estimate for the global minimum, re-estimate the interval around this point, and re-iterate the algorithm from step 1 (or, if the interval is within the accepted tolerance level, accept this as the final minimum value).
4. **ELSE** use the Golden Section algorithm to estimate a new minimum estimate, re-estimate the interval around this estimate, and re-iterate the algorithm from step 1 (or, if the interval is within the accepted tolerance level, accept this as the final minimum value).

This appendix does not discuss the actual implementation details of the method, since some house-keeping parameters are required. Instead, focus is placed on the main concepts that give the optimisation routine its utility.

A.1 Parabola Estimation

Given a function f and the points r, s and t such that $r < s < t$ and $f(r) > f(s) < f(t)$ a parabola, described by Equation (A.1).

$$g(x) = a_2x^2 + a_1x + a_0 \tag{A.1}$$

Since each of the three points is on the parabola, Equations (A.2), (A.3) and (A.4), must be true.

$$\begin{aligned} f(r) &= a_2r^2 + a_1r + a_0 && \text{(A.2)} \\ f(s) &= a_2s^2 + a_1s + a_0 && \text{(A.3)} \\ f(t) &= a_2t^2 + a_1t + a_0 && \text{(A.4)} \end{aligned}$$

The minimum of the parabola is defined as the point where the first derivative of $g(x) = 0$, which leads to Equation (A.5).

$$\begin{aligned} g'(x) &= 2a_2x + a_1 = 0 \\ \implies x &= \frac{-a_1}{2a_2} \end{aligned} \tag{A.5}$$

Using substitution the value of x can be determined according to Equation (A.6).

$$x = s - \frac{1}{2}\frac{(s-r)^2[f(s)-f(t)] - (s-t)^2[f(s)-f(r)]}{(s-r)[f(s)-f(t)] - (s-t)[f(s)-f(r)]} \tag{A.6}$$

where x is the extremum of the parabola, which represents the estimated

minimum of the function f in the interval defined by r, s and t.

To be accepted the new point, x, estimated by the parabolic function must fulfill two criteria:

1. The estimated point must be within the interval defined by (r, t).
2. The movement from the new point x must be less than half the movement undergone in the step before last [84]. This ensures the optimisation routine is converging, as opposed to oscillating in a non-convergent manner.

A.2 Golden Section Algorithm

If the parabolic method described above fails, Brent's optimisation method reverts to the Golden Section Algorithm. This algorithm focuses on reducing the size of the current interval by employing a set of logical steps:

1. Begin with the function f and three evaluation points, r, s and t, such that $r < s < t$ and $f(r) > f(s) < f(t)$.
2. Choose a new point, q, such that $r < q < s < t$ (usually q is chosen to be between the larger of the two intervals: (r, s) and (s, t).
3. **IF** $f(q) > f(s)$ then the interval becomes (q, s, t).
4. **ELSE** if $f(q) < f(s)$ then the interval becomes (r, q, s).
5. Repeat this procedure until convergence is reached.

This progression of steps implies that the new interval at each step will either be (q, t) or (r, s). The algorithm requires these two intervals to be of equal length, which is achieved by choosing $q = r - s + t$. Finally, the distance between the initial interval (r, s, t) is chosen so that:

$$\frac{t-s}{s-r} = \varphi \tag{A.7}$$

where φ is the well-known golden ratio defined as $\frac{1+\sqrt{5}}{2}$. Constraining the interval spacing to these proportional rules controls the evolution of the

optimisation step, and ensures the progression is stable and predictable.

A.3 Evaluation on Endocardium Segmentation

The proposed ASM optimisation routine described in Chapter 3 directly manipulates the β vector, which is used to synthesize shapes. The β vector's range of allowed variation is constrained by the training set. The outer limit of variation is described by the vector, β_{limit}, so that for any synthesized shape, $|\beta| \leq \beta_{limit}$. It is straightforward to map this arrangement into the framework required for Brent's optimisation. The initial definition of the interval for the first β component, β^i is defined as: $\left(-\beta^i_{limit}, \beta^i_{limit}\right)$. From here the optimisation algorithm switches between parabola fitting and the Golden Section algorithm depending on the nature of the image being segmented (i.e. the shape of the function defined by the application of the endocardium objective function to the image).

To evaluate the efficacy of the optimisation routine, Brent's optimisation method is compared to a simple brute-force search. The evaluation scenario is similar to that developed in Chapter 3, where a 2D endocardium shape model is constructed and evaluated in a leave-one-out manner. The ASM is trained on the basal-most slice of the first temporal frame of each dataset. As in earlier chapters, each component of the β vector is optimised sequentially, and this process is repeated until the contour score converges. To evaluate the accuracy of the Brent's method and its robustness to local minima, the final converged contour score is compared with the global minimum for the entire search-space using a brute force optimisation routine. The results for this are presented in Fig. A.1. Examination of the results show that Brent's optimisation method is usually within 1% and 2% of the global minimum.

A.4 Noise Evaluation

With the stability of Brent's optimisation experimentally proven in the previous section, comparison is now made against several other relevant ASM fitting methods. Specifically, the Nelder-Mead method (also known as the downhill simplex method) and the standard Cootes ASM fitting method are

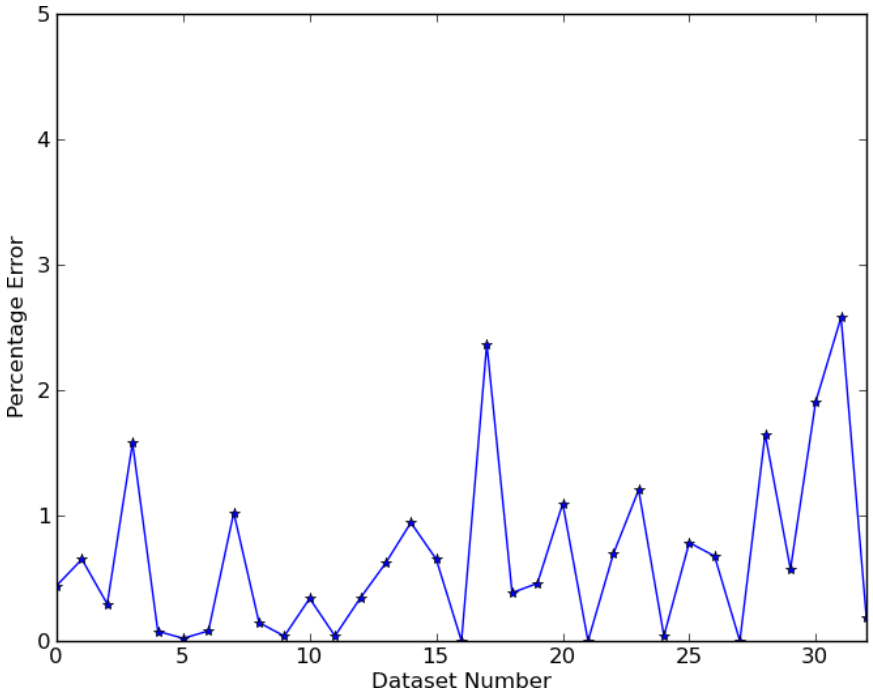


Figure A.1: This graph depicts the percentage error of the absolute difference between the final objective function score using the brute force method compared with Brent's optimisation method.

both compared against Brent's method for endocardium segmentation. The downhill simplex method is evaluation using the same objective function as the one developed in Chapter 3.

The three methods are evaluated on a single 2D short axis slice of a dataset that exhibits well defined contours and contains no papillary muscles or other image artefacts. This allows a 'base-level' of the segmentation accuracy of each method to be determined. Next, additive grey-level noise is applied to the evaluation image in increasing levels. At each level, each ASM fitting method is applied ten times to the evaluation image and the mean segmentation error is recorded. The overall mean error over ten segmentations is taken as the representative error for that method at that noise level. The additive grey-level noise level is applied by adding to each pixel a random grey-level value between zero and the maximum defined noise level.

Fig. A.2 depicts the performance of each method for maximum noise levels between 0 and 120. The initial observation that can be drawn from this

graph is that all three methods exhibit approximately equivalent segmentation accuracy when no noise is applied to the image. Next, the notable stability of Brent's optimisation method as the level of noise increases is clear compared to the other methods. Both the Cootes and Nelder-Mead methods exhibit similar performance characteristics, where the mean segmentation error rises in a linear fashion with the increase in applied noise level.

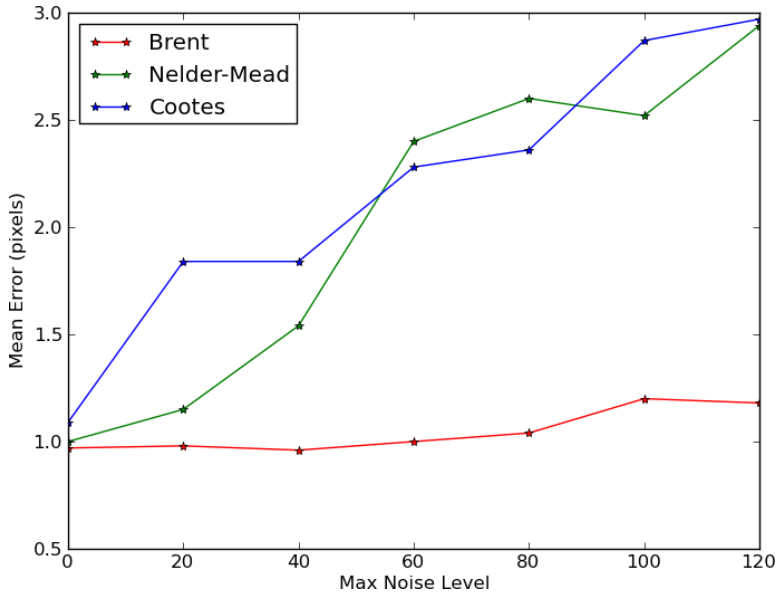


Figure A.2: This figure illustrates the relative performance between Brent's optimisation method, the Nelder-Mead optimisation method and the standard Cootes ASM fitting technique.

Naturally, these results give rise to the question of why Brent's method exhibits only a minimal rise in segmentation error when image noise increases. Fig. A.3 aims to illustrate the reason. In the left column are sample segmentation results from noise levels of 40 80 and 120, respectively. Opposite each image is its associated search space. As the noise level increases, the degradation of the search space is clearly visible. Because Brent's method optimises each eigenvector in a sequential manner (i.e. one dimension of the search space at a time) it is less likely than the Nelder-Mead method to be affected by image noise, since the Nelder-Mead method attempts to optimise the entire β vector simultaneously. The Nelder-Mead method is obviously prone to the increase in local minima in the search space as the noise level

increases. Cootes method shows similar performance to the Nelder-Mead method. The analysis performed in Chapter 3 of this thesis suggests that the degradation in performance is due to the individual nature of the landmark fitting process. In essence, each landmark can be negatively affected by the image noise and can lead to improbable contour configurations, causing segmentation errors.

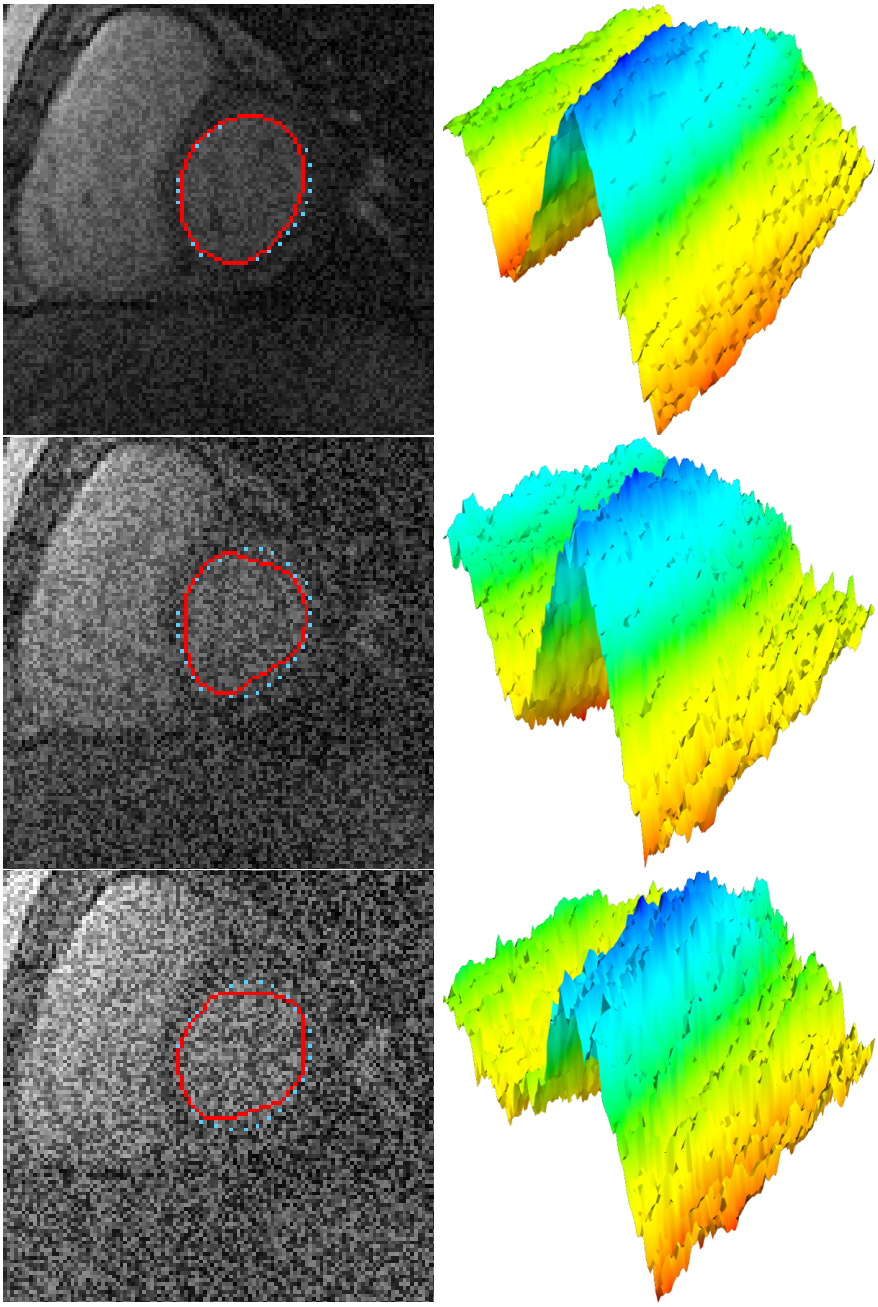


Figure A.3: Three qualitative examples of the proposed ASM optimisation routine using Brent's method. From top-to-bottom, the images correspond to a maximum level of 40, 80 and 120 additive grey-level noise. Opposite each segmented image (where manual annotations are shown as blue points, and automatic segmentations are continuous contours) is the corresponding search space, visualised for the first two eigenvectors.

Appendix B

Application to Hand Segmentation

The contributions presented in this thesis are not limited to the segmentation of cardiac left ventricle datasets. This appendix aims to illustrate the applicability of one of the presented contributions, contour coupling, to an alternate medical image segmentation problem. To illustrate the flexibility of the proposed contour coupling technique, hand segmentation is chosen as the problem domain in this appendix. The segmentation challenge is to extract the metacarpals as well as the phalanges bones from a radiograph of the hand. A sample dataset is shown in Fig. B.1. This appendix does not aim to develop a complete segmentation solution to extract all the bones from an image. Instead, two scenarios are investigated where a subset of the bones are selected for segmentation. The set of bones in each scenario are chosen to illustrate the benefits gained by assuming the proposed contour coupling approach, when compared with the more common combined modelling approach, where all segmentation objects are considered as a single shape.

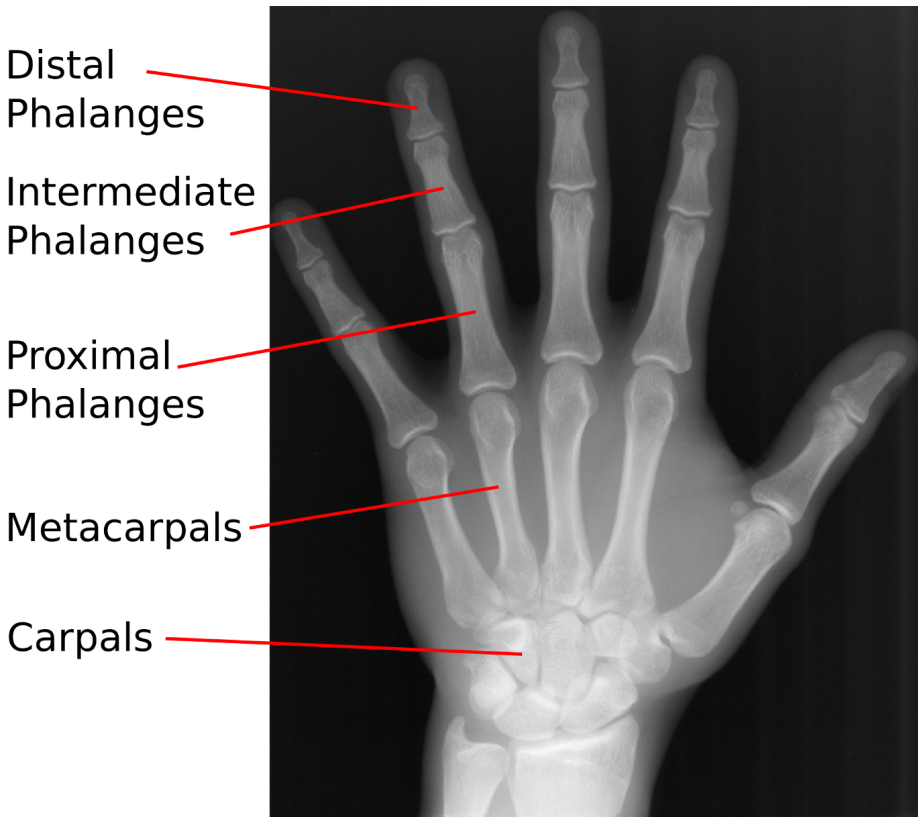


Figure B.1: Sample dataset showing an eighteen year old's hand. Labelled are the carpals, meta-carpals and phalanges.

B.1 Dataset

The dataset used in this appendix is a publicly available[1] set of radiographs of children and young adults up to the age of 18. The images in this dataset were all acquired from Childrens Hospital Los Angeles. The modality used is computed radiography (CR). CR is similar to standard X-ray imaging, except that the imaging plane is made from photostimulable storage phosphors. A primary benefit of this imaging plate is that the image can be digitally extracted from the plate after acquisition, reducing storage cost and allowing the plate to be reused in future examinations.

From this large dataset of approximately 1400 hand images, a selection of sixty datasets from patients aged 8 to 13 are used for evaluation. Each sub-

[1] `http://www.ipilab.org/BAAweb/`

structure of interest in each image has been annotated by the author of this thesis.

B.2 Problem Definition

The images in Fig. B.2 present a sample range of variation seen in this problem domain. As shown, there can be intensity variation between datasets, bones appear at different rotations and scales, relative rotation between fingers can change, and, in some cases, contrast between bone and surrounding soft tissue is quite low. The inter-carpal and inter-phalangial distance between bones, which are all in the same finger, varies significantly with respect to patient age. Finally, there can also be apparent overlap (albeit slight) between carpals and meta-carpals in some datasets. To aid clarity in the remainder of this appendix, Fig. B.3 depicts the notation used to label the bones of the hand in this segmentation application.

B.3 Evaluation Scenarios

The first evaluation scenario (referred to as S1) is a natural fit to the problem domain. Here, the set of bones to be segmented are: B3, M3, T3. These bones are, of course, inherently related, forming the structure of the index finger.

The second scenario (referred to as S2) consists of an arbitrary selection of bones: T4, M3, B3, and B1. The aim of this scenario is to show that the contour coupling technique can successfully adapt to challenging modelling and segmentation tasks.

In both of these evaluation scenarios, the contour coupling approach will be compared and contrasted against a combined modelling approach. The combined approach treats the set of bones in each scenario as a single shape.

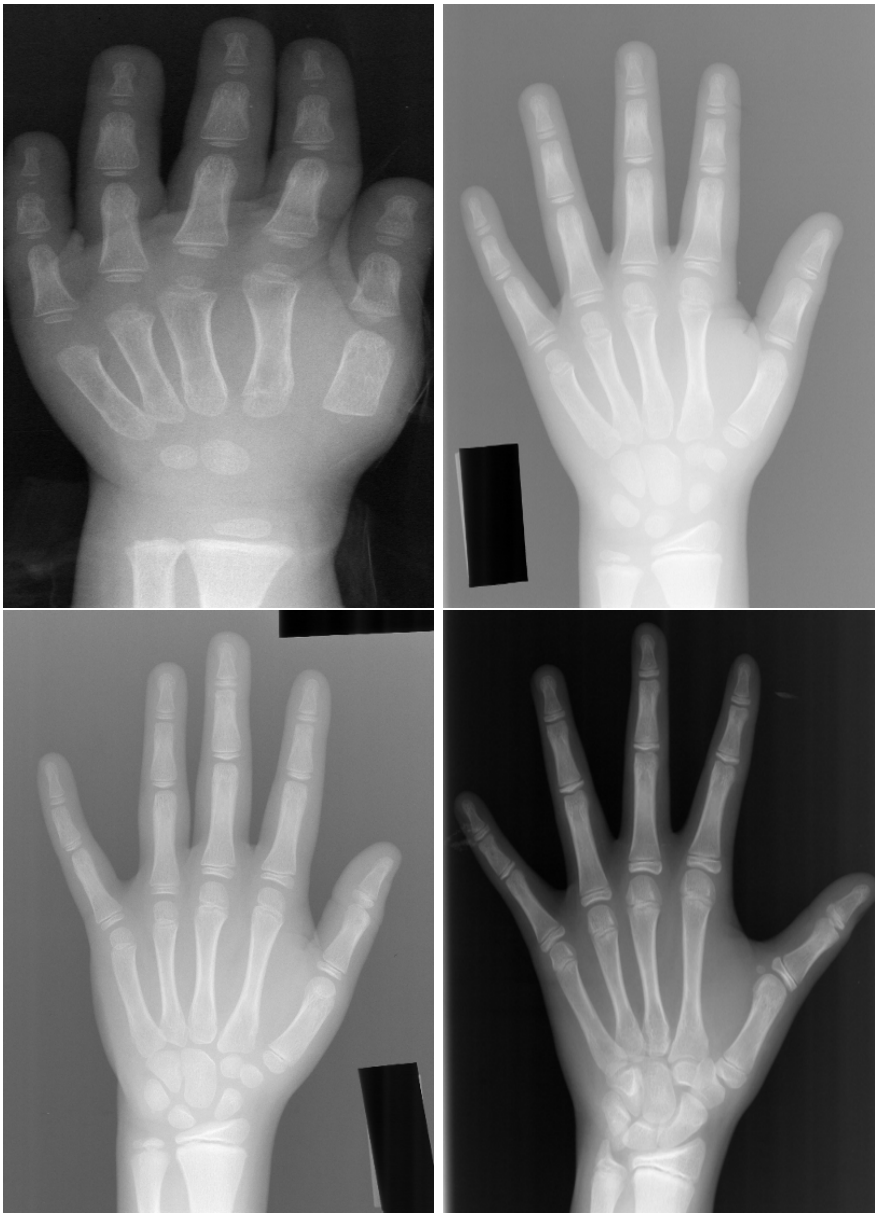


Figure B.2: Four example images illustrating the range of image and anatomical variation seen in hand radiographs. From left to right and from top to bottom the progression in images shows patients aged from newborn to eighteen.

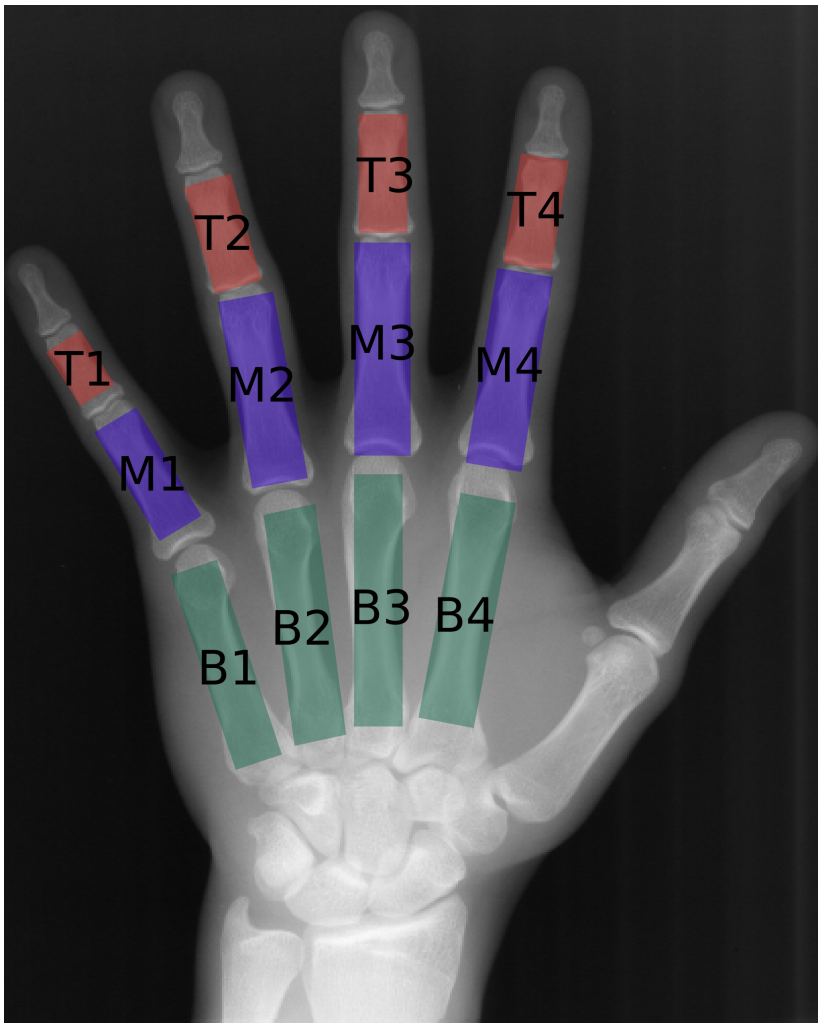


Figure B.3: This figure describes the notation used in the remainder of this appendix to refer to specific bones in the skeletal structure of the hand.

B.4 Model Construction

Following the approach employed throughout this thesis, the shape ASM for each bone is created separately. The number of landmarks that define each training sample is normalised to 33 landmarks using arc-length resampling. Next, translation, scale and rotation are removed by applying Procrustes analysis to the training shapes. At this point, PCA is applied to extract the eigenvectors and values that describe the modes of variation in the training set. The eigenvectors are truncated to capture 98% of shape variation in a method similar to that described in Section 3.1. In line with the notation put forward in Fig. B.3, a model trained for, say, the M2 bone will be referred as β_{M2}.

B.5 Coupling Constraints

The relationship between bone structures can be learned during the training process, similarly to the CC-MD technique developed in Chapter 4. In this appendix, the relationship between two bones, say B2 and M2, can be described as the relative translation, scale and rotational differences observed between each training sample. The translation and rotational difference between each sample is translated into polar coordinates for clarity, so that if the centre of gravity for B2 and M2 is defined as $(B2_x, B2_y)$ and $(M2_x, M2_y)$, respectively, then the magnitude and angular components of the polar are calculated from $r = \sqrt{(M2_y - B2_y)^2 + (M2_x - B2_x)^2}$ and $\theta = arcsin\left(\frac{M2_y - B2_y}{r}\right)$, respectively. The scale difference is measured as the difference in L^2-norm between B2 and M2.

For example, taking B2 as the reference point, the relationship to M2 is graphically depicted in Fig. B.4. The relationship, $R_{B2\rightarrow M2}$, is modelled as a Gaussian distribution based on the data in Fig. B.4. Of course, the inverse relationship, $R_{M2\rightarrow B2}$, can be derived by reversing each data point $\left(r, \theta, \Delta L^2_{M2\rightarrow B2}\right) \rightarrow \left(r, -\theta, -\Delta L^2_{M2\rightarrow B2}\right)$.

Given a segmented B2 shape, the relationship $R_{B2\rightarrow M2}$, allows coupling constraints to be applied so the subsequent segmentation of M2. Initial scale, rotation and translation can be applied as the maximum value of the

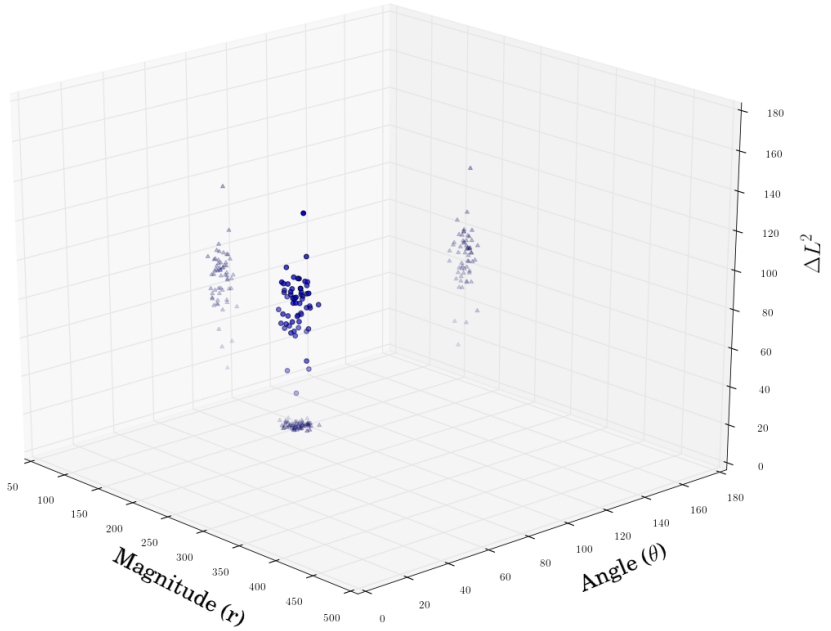


Figure B.4: Graphical depiction of the relationship between the B3 and M3 bones, using B3 as the reference shape. Each data point (shown in blue) represents the magnitude, θ and L^2 difference between the B3 and M3 bones in a training sample. For clarity, the blue triangles represent the projections of the point-cloud onto the relevant axes.

Gaussian associated with the relationship. The optimisation of β_{M2} is then constrained so that synthesized shapes are within three standard deviations of the Gaussian, allowing 99.7% of the learned variation to be employed in the application of the constraint.

Using this approach for the first evaluation scenario (B3, M3, and T3), the coupling constraints can be derived by calculating $R_{B3\rightarrow M3}$ and $R_{B3\rightarrow T3}$, by taking B3 to be the reference shape. These relationships are shown in Fig. B.5. Similarly, the second evaluation scenario (T4, M3, B3, and B1) requires the relationships $R_{M3\rightarrow T4}$, $R_{M3\rightarrow B3}$ and $R_{M3\rightarrow B1}$ to be calculated, taking M3 as the reference shape. Fig. B.6 illustrates these relationships.

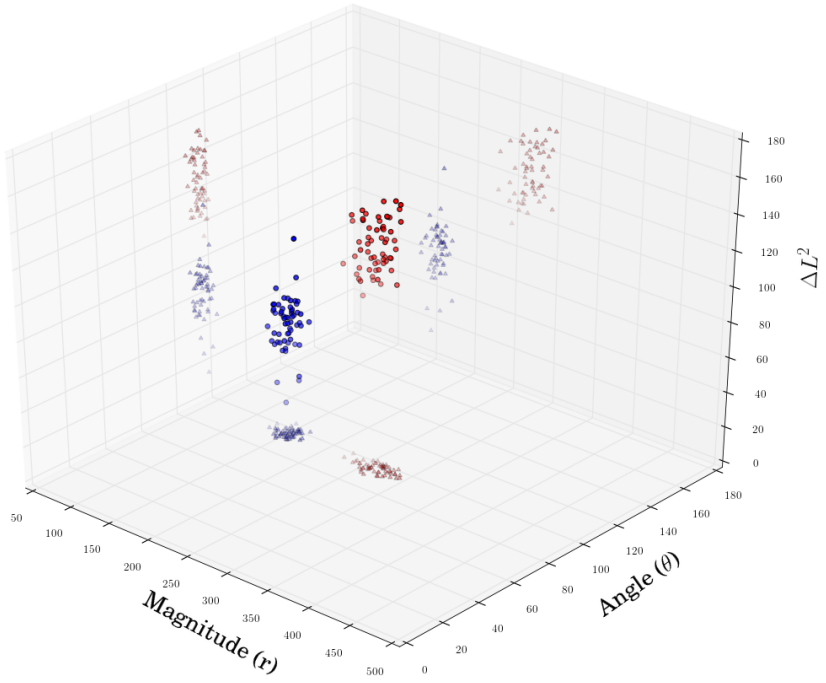


Figure B.5: $(r, \theta, \Delta L^2)$ relationship between B3 and M3 (shown in blue) and B3 and T3 (shown in red). Projections onto the surrounding axes are colour-coded with their associated point-cloud.

B.6 Results

Each scenario is evaluated in 30 trials. In each trial, 50 of the 60 annotated samples are randomly selected and used to train the required ASMs. One of the remaining 10 datasets is then segmented using the proposed approach. The is constrained so that a different dataset is guaranteed to be segmented in each trial. This process is repeated thirty times to ensure no bias is included in the presented results. A single xy seed point is used to initialise the reference shape in each scenario. When the reference shape is segmented, the coupling constraints developed in training are employed to segment the remainder of the required shapes.

Table B.1 contains the results for the first scenario where the B3, M3, and T3 bones are segmented, using B3 as the reference and the second scenario where B3, M3, T4 and B1 are segmented, using M3 as the reference shape. Each reference shape is denoted with a '*' in the table. Segmentation accuracy is quantified in these experiments using the Dice metric between the au-

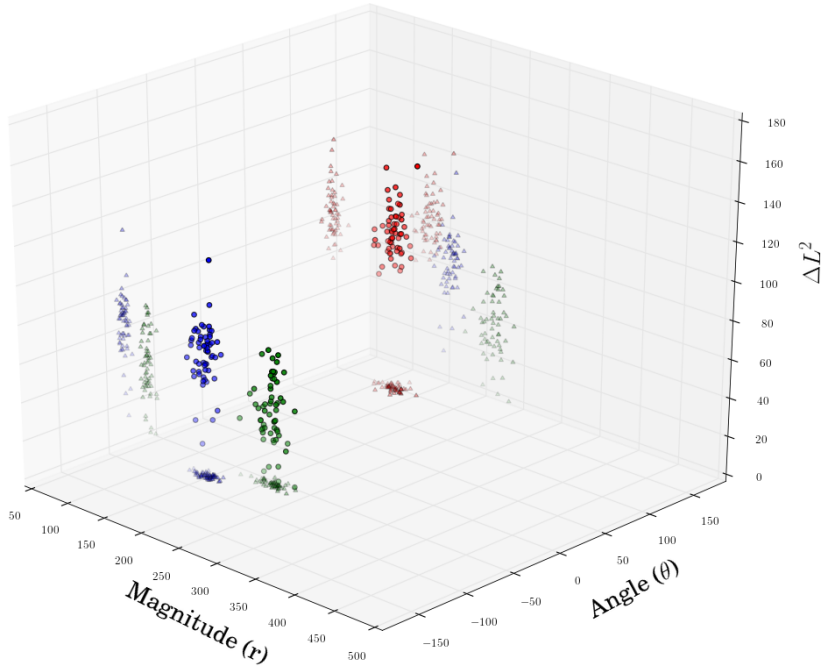


Figure B.6: $\left(r, \theta, \Delta L^2\right)$ relationship between M3 and B3 (shown in blue), M3 and T4 (shown in red) and M1 and B1 (shown in green). Projections onto the surrounding axes are colour-coded with their associated point-cloud.

tomatically segmented shape using the proposed contour coupling approach and the manual annotations. The Dice value quantifies the percentage overlap between two shapes, where 0% implies no overlap between shapes and 100% indicates that the two shapes are exactly aligned. Each row in the results table contains the Dice value for each bone segmented in that trial. The mean Dice value for B3, M3 and T3 is 91%, 93% and 94%. Figs. B.7 and B.8 show a selected of qualitative results from evaluation scenario one and two, respectively, using the proposed contour coupling approach.

The combined ASM method is evaluated in the same way as the contour coupling method. The same evaluation scenarios and testing protocol are employed. However, estimating the xy initialisation point of the combined shape is not intuitive. Therefore, to present comparable results, the xy seed point is determined from the manual annotations for the combined ASM evaluation. Table B.2 contains the results for both evaluation scenarios for the combined method. Although the bones in each case are considered as one shape, after segmentation this unified shape is decomposed into its

individual bone structures in order to determine the Dice value of each, and to provide direct comparison against the proposed contour coupling approach.

B.7 Discussion

Regrading the first evaluation scenario, the results for the combined ASM and contour coupling are approximately equivalent. This is expected, since the B3, M3 and T3 bones are very closely related, with minimal variation in rotation, scale and translation. As such, the combined ASM is able to synthesize the shapes with comparable accuracy to the contour coupling method.

By contrast, the second evaluation scenario illustrates significant differences between the combined ASM and the contour coupling segmentation results. The central difference between the evaluation scenarios is that, in the second, the bones are not closely related with each other. As a result, there are considerable rotation, scale and translation differences between training samples. This has two consequences that negatively affect the segmentation results for the combined ASM approach. First, developing point-to-point correspondence between the training samples is more difficult because of the increased complexity of the shape, which affects the allowed range of variation when calculating the eigenvectors and values. Second, fitting the combined ASM to an unseen dataset requires simultaneous segmentation of all the bones in that segmentation scenario, which is a very challenging task. Of course, a similar set of problems occur in left ventricle cardiac segmentation as detailed in Section 4.1.2.

Analysis of each row of Table B.2 for the second evaluation scenario depicts a noticeable pattern. In general, for each segmentation, two of the four bones are segmented relatively well, whereas the remaining two are segmented very poorly (see Fig. B.9). This highlights the limitation of the combined shape modelling approach when applied to complex segmentation problems.

Increasing the complexity of the segmentation problem has dramatic consequences for the combined approach. By contrast, the stability and segmentation accuracy of the contour coupling approach in the second evaluation

scenario shows no such degradation in performance. Indeed, visual comparison between the results present in Fig. B.8 and B.9 (for contour coupling and the combined approach, respectively) attests to this fact.

This appendix has achieved two primary goals. First, the applicability of the contour coupling concept outside the cardiac segmentation domain has been detailed. Second, the extensibility of contour coupling has been shown been comprehensively demonstrated with experiments in 3 and 4 object-modelling scenarios. Direct comparison against a combined modelling approach reinforces the benefits of assuming the contour coupling approach in complex modelling and segmentation domains.

Table B.1: Quantitative results from both evaluation scenarios using the proposed contour coupling method. All values presented are percentages. The reference shape in each scenario is denoted by a '*'.

	S1 Results			S2 Results			
Test No.	B3*	M3	T3	B3	M3*	T4	B1
1	93	96	96	93	94	95	86
2	94	90	95	95	96	90	74
3	93	89	96	87	95	95	95
4	85	94	96	85	96	93	90
5	86	94	90	90	95	92	84
6	92	97	90	92	92	94	90
7	94	97	93	89	83	92	82
8	89	92	93	92	90	91	93
9	82	93	95	83	94	95	85
10	82	94	96	75	95	92	88
11	92	94	95	91	95	95	86
12	93	91	91	90	95	93	93
13	93	92	96	89	93	95	90
14	92	96	95	83	91	93	88
15	90	94	93	87	85	93	88
16	94	91	93	92	96	94	86
17	94	90	93	84	94	93	85
18	92	82	95	95	87	92	94
19	87	88	85	85	95	93	95
20	89	95	94	86	90	90	89
21	89	95	93	92	83	93	85
22	93	94	94	87	94	93	86
23	91	96	95	78	94	91	84
24	86	96	95	89	85	87	90
25	91	95	92	94	83	87	94
26	92	95	93	93	94	96	92
27	91	96	97	90	96	91	94
28	94	89	93	84	88	97	90
29	94	95	94	89	96	93	92
30	91	95	94	90	95	92	89
Mean	90.60	93.17	93.67	88.30	91.97	92.67	88.56
Std-Dev	3.48	3.28	2.40	4.74	4.39	2.28	4.59

Table B.2: Quantitative results from both evaluation scenarios (S1 and S2) using the combined ASM approach. All values presented are percentages.

	S1 Results			S2 Results			
Test No.	B3	M3	T3	B3	M3	T4	B1
1	85	93	89	56	74	14	64
2	84	94	94	43	92	2	65
3	91	97	96	66	89	0	7
4	65	85	75	74	79	6	14
5	86	95	94	83	86	49	32
6	88	93	93	85	87	18	26
7	91	97	95	71	89	73	52
8	87	93	84	80	96	18	71
9	88	95	91	89	88	15	17
10	94	96	95	79	73	9	8
11	71	94	93	78	65	20	5
12	91	95	83	23	41	0	0
13	72	92	76	37	58	1	45
14	91	95	93	72	59	3	10
15	85	95	96	82	79	1	16
16	90	95	95	70	83	0	06
17	91	95	92	82	89	26	66
18	87	93	96	89	78	60	44
19	92	93	91	77	74	11	21
20	89	96	95	75	95	25	73
21	66	88	83	84	77	14	3
22	93	93	91	78	91	24	68
23	89	96	96	82	82	1	18
24	84	93	90	90	90	40	39
25	93	97	92	32	72	0	64
26	92	96	90	56	68	22	62
27	85	93	92	70	56	48	35
28	89	96	93	88	60	35	27
29	92	93	95	37	48	1	39
30	89	94	92	71	59	4	7
Mean	86.33	93.97	90.99	69.96	75.90	18.02	33.47
Std-Dev	7.73	2.60	5.54	18.53	14.72	19.50	24.34

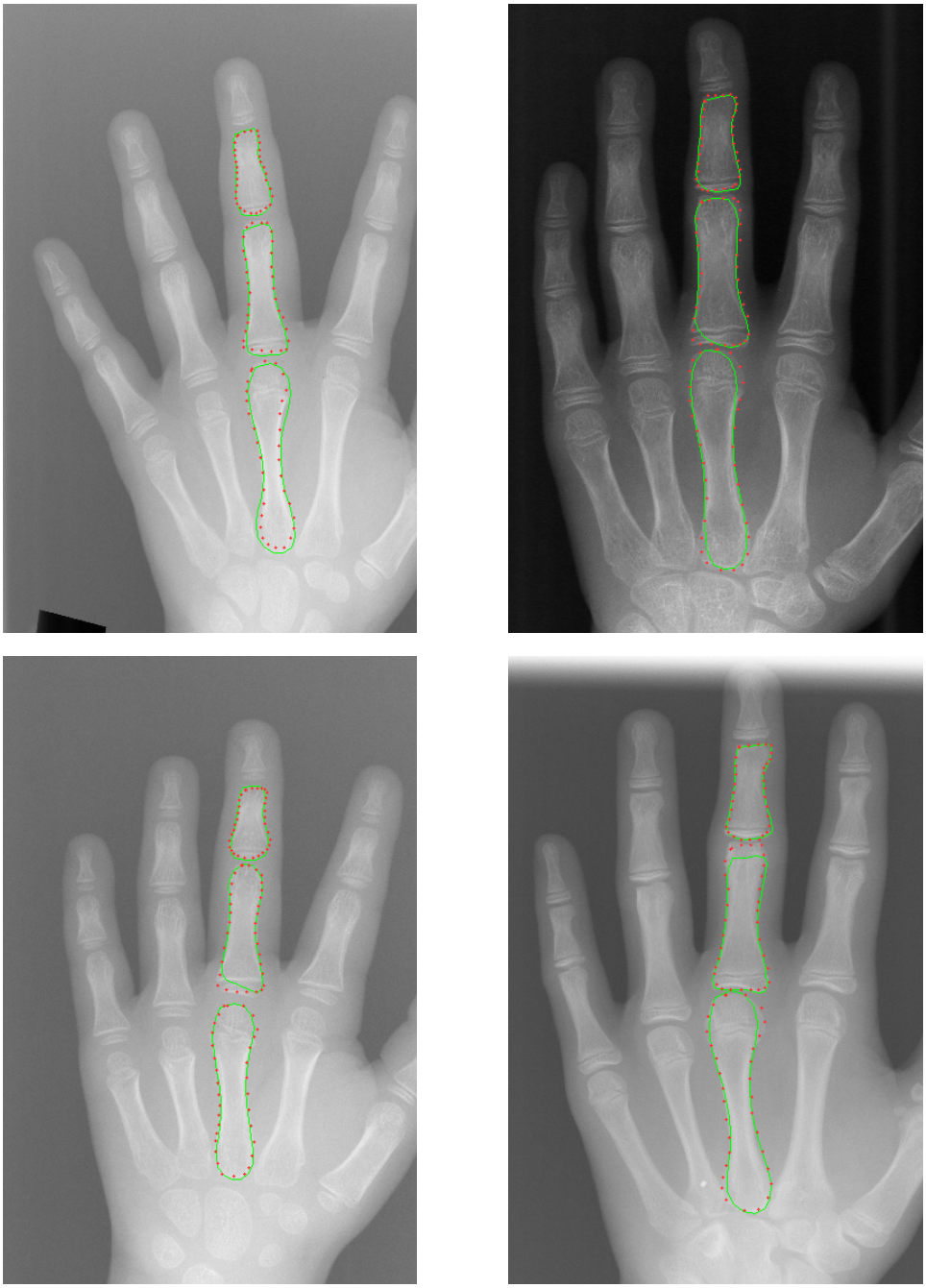


Figure B.7: Four randomly selected segmentation results (cropped for clarity) from the first evaluation scenario using the proposed contour coupling approach. Automatic segmentations are shown as continuous green contours and manual annotations are depicted with green crosses.

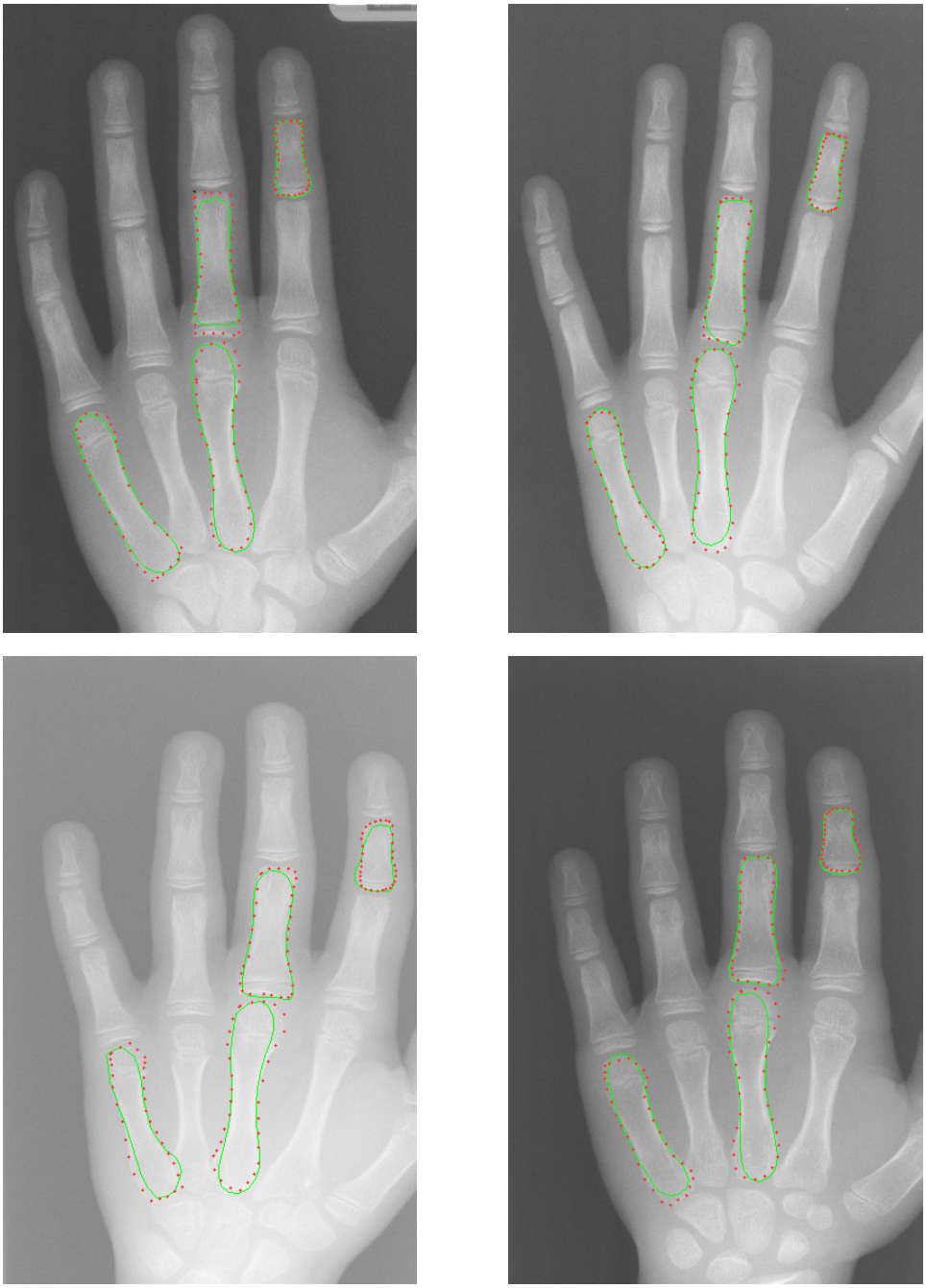


Figure B.8: Four randomly selected segmentation results (cropped for clarity) from the first evaluation scenario using the proposed contour coupling approach. Automatic segmentations are shown as continuous green contours and manual annotations are depicted with green crosses.

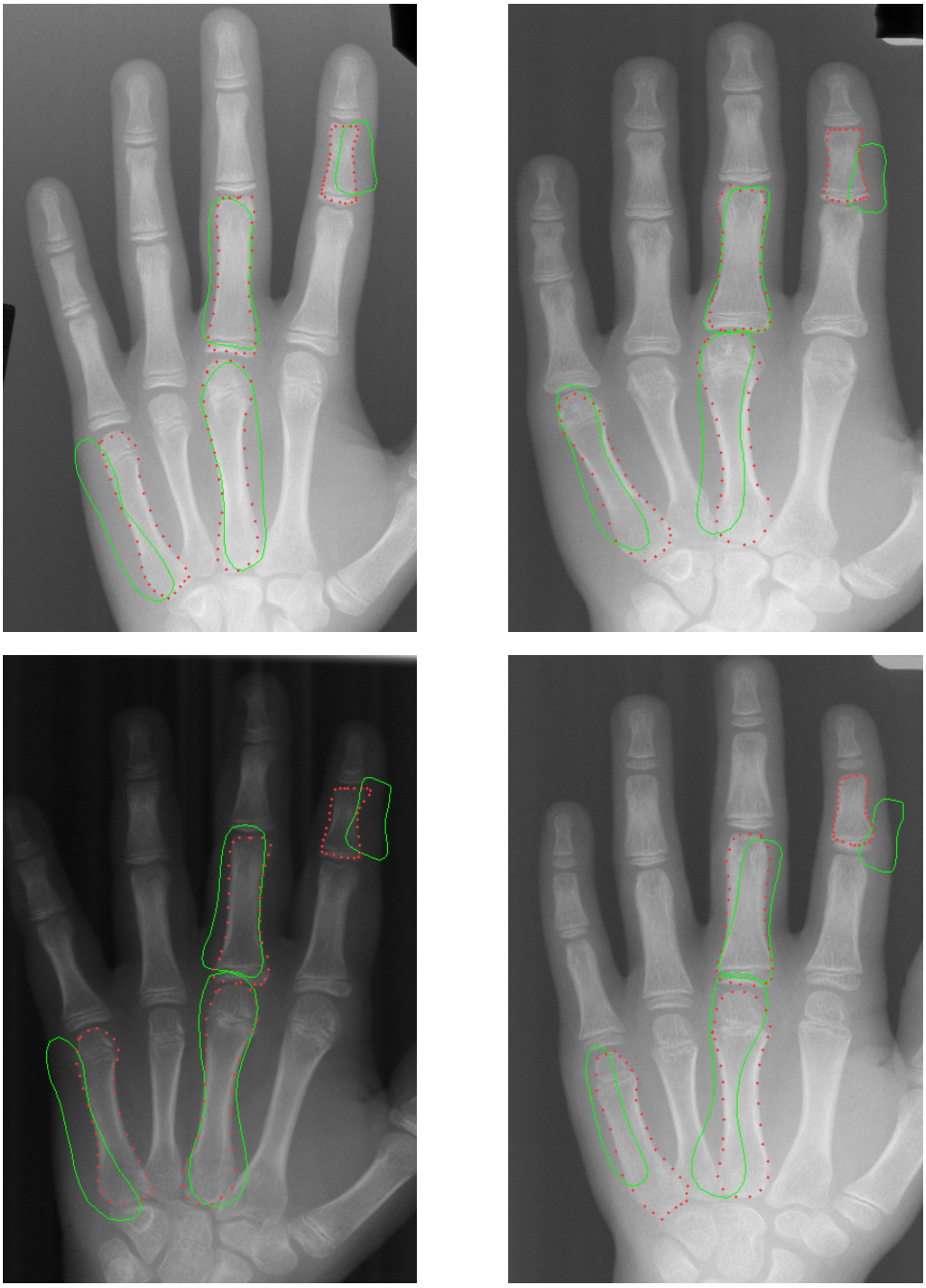


Figure B.9: Four sample segmentation results (cropped for clarity) from the second evaluation scenario using the combined ASM approach. The difficulty the combined ASM faces in correctly synthesizing all four shapes simultaneously is evident from the automatic segmentations (shown in green) compared to the manual annotations (show as green crosses).

Bibliography

[1] J.E. Hall A.C. Guyton. *Textbook of Medical Physiology*. Saunders/Elsevier, 11 edition, 2005.

[2] D. Adalsteinsson and J.A. Sethian. A fast level set method for propagating interfaces. *Journal of Computational Physics*, 118(2):269–277, 1995.

[3] K. Alfakih, S. Reid, T. Jones, and M. Sivananthan. Assessment of ventricular function and mass by cardiac magnetic resonance imaging. *European radiology*, 14(10):1813–1822, 2004.

[4] S. Allender, P. Scarborough, V. Peto, M. Rayner, J. Leal, R. Luengo-Fernandez, and A. Gray. European cardiovascular disease statistics. *Brussels: European Heart Network*, 2008.

[5] H. Amthauer, T. Denecke, T. Rohlfing, J. Ruf, M. Böhmig, M. Gutberlet, U. Plöckinger, R. Felix, and A.J. Lemke. Value of image fusion using single photon emission computed tomography with integrated low dose computed tomography in comparison with a retrospective voxel-based method in neuroendocrine tumours. *European radiology*, 15(7):1456–1462, 2005.

[6] A. Andreopoulos and J.K. Tsotsos. Efficient and generalizable statistical models of shape and appearance for analysis of cardiac MRI. *Medical Image Analysis*, 12(3):335–357, 2008.

[7] F. Arámbula Cosío, J.A. Flores, and M.A. Castaŋeda. Use of simplex search in active shape models for improved boundary segmentation. *Pattern Recognition Letters*, 2010.

[8] E.C. Barbier, L. Johansson, L. Lind, H. Ahlström, and T. Bjerner. The exactness of left ventricular segmentation in cine magnetic resonance imaging and its impact on systolic function values. *Acta Radiologica*, 48(3):285–291, 2007.

[9] A. Bayes de Luna, G. Wagner, Y. Birnbaum, K. Nikus, M. Fiol, A. Gorgels, J. Cinca, P.M. Clemmensen, O. Pahlm, S. Sclarovsky, et al. A new terminology for left ventricular walls and location of myocardial infarcts that present Q wave based on the standard of cardiac magnetic resonance imaging: a statement for healthcare professionals from a committee appointed by the International Society for Holter and Noninvasive Electrocardiography. *Circulation*, 114(16):1755, 2006.

[10] I. Ben Ayed, Li. Shuo, and I. Ross. Embedding overlap priors in variational left ventricle tracking. *IEEE Transactions on Medical Imaging*, 28(12):1902–1913, 2009.

[11] J. Bogaert, S. Dymarkowski, and A.M. Taylor. *Clinical cardiac MRI: with interactive CD-ROM*. Springer, 2005.

[12] J.G. Bosch, S.C. Mitchell, B.P.F. Lelieveldt, F. Nijland, O. Kamp, M. Sonka, and J.H.C. Reiber. Automatic segmentation of echocardiographic sequences by active appearance motion models. *IEEE Transactions on Medical Imaging*, 21(11):1374–1383, 2002.

[13] D. Boukerroui, A. Baskurt, J.A. Noble, and O. Basset. Segmentation of ultrasound images–multiresolution 2D and 3D algorithm based on global and local statistics. *Pattern Recognition Letters*, 24(4-5):779–790, 2003.

[14] R.P. Brent. *Algorithms for Minimization without Derivatives*. Prentice-Hall, 1973.

[15] J. Canny. A computational approach to edge detection. *Readings in computer vision: issues, problems, principles, and paradigms*, 184, 1987.

[16] V. Caselles, F. Catté, T. Coll, and F. Dibos. A geometric model for active contours in image processing. *Numerische Mathematik*, 66(1):1–31, 1993.

[17] R. Casero and J.A. Noble. A novel explicit 2D+t cyclic shape model applied to echocardiography. In *Medical image computing and computer-assisted intervention: MICCAI International Conference on Medical Image Computing and Computer-Assisted Intervention*, volume 11, pages 527–534. Springer, 2008.

[18] T.F. Chan and L.A. Vese. Active contours without edges. *IEEE Transactions on image processing*, 10(2):266–277, 2001.

[19] M. Chen, T. Kanade, D. Pomerleau, and J. Schneider. 3-D deformable registration of medical images using a statistical atlas. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI99*, pages 621–630. Springer, 1999.

[20] H. Chui, A. Rangarajan, J. Zhang, and C.M. Leonard. Unsupervised learning of an atlas from unlabeled point-sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 160–172, 2004.

[21] L.D. Cohen. On active contour models and balloons. *CVGIP: Image understanding*, 53(2):211–218, 1991.

[22] D.L. Collins, C.J. Holmes, T.M. Peters, and A.C. Evans. Automatic 3-D model-based neuroanatomical segmentation. *Human Brain Mapping*, 3(3):190–208, 1995.

[23] T.F. Cootes, G.J. Edwards, C.J. Taylor, et al. Active appearance models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6):681–685, 2001.

[24] T.F. Cootes and C.J. Taylor. Active shape models–smart snakes. In *Proc. British Machine Vision Conference*, pages 266–275, 1992.

[25] T.F. Cootes and C.J. Taylor. Statistical models of appearance for computer vision. *Technical Report*, 2001.

[26] T.F. Cootes, C.J. Taylor, D.H. Cooper, J. Graham, et al. Active shape models - their training and application. *Computer Vision and Image Understanding*, 61(1):38–59, 1995.

[27] F.A. Cosío. Automatic initialization of an active shape model of the prostate. *Medical Image Analysis*, 12(4):469–483, 2008.

[28] M. Costa, H. Delingette, S. Novellas, and N. Ayache. Automatic segmentation of bladder and prostate using coupled 3D deformable models. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007*, pages 252–260, 2007.

[29] J. Cousty, L. Najman, M. Couprie, S. Clément-Guinaudeau, T. Goissen, and J. Garot. Segmentation of 4D cardiac MRI: Automated method based on spatio-temporal watershed cuts. *Image and Vision Computing*, 2010.

[30] D. Cremers, M. Rousson, and R. Deriche. A review of statistical approaches to level set segmentation: integrating color, texture, motion and shape. *International Journal of Computer Vision*, 72(2):195–215, 2007.

[31] D. Cristinacce and T. Cootes. Automatic feature localisation with constrained local models. *Pattern Recognition*, 41(10):3054–3067, 2008.

[32] A. Cristoforetti, L. Faes, F. Ravelli, M. Centonze, M. Del Greco, R. Antolini, and G. Nollo. Isolation of the left atrial surface from cardiac multi-detector CT images based on marker controlled watershed segmentation. *Medical Engineering & Physics*, 30(1):48–58, 2008.

[33] C. Davatzikos, X. Tao, and D. Shen. Hierarchical active shape models, using the wavelet transform. *IEEE Transactions on Medical Imaging*, 22(3):414–423, 2003.

[34] M. De Bruijne, B. Van Ginneken, M.A. Viergever, and W.J. Niessen. Adapting active shape models for 3D segmentation of tubular structures in medical images. In *Proc. Information Processing in Medical Imaging*, pages 136–147. Springer, 2003.

[35] S.C. DSc and M. O'Flaherty. Trends in cardiovascular disease: Are we winning the war? *CMAJ*, 180:1285–1286, 2009.

[36] N. Duta and M. Sonka. An improved active shape model: Handling occlusion and outliers. *Lecture Notes in Computer Science*, pages 398–405, 1997.

[37] R. El Berbari, I. Bloch, A. Redheuil, E. Angelini, E. Mousseaux, F. Frouin, and A. Herment. Automated segmentation of the left ventricle including papillary muscles in cardiac magnetic resonance images. *Functional Imaging and Modeling of the Heart*, pages 453–462, 2007.

[38] A.F. Frangi, D. Rueckert, J.A. Schnabel, and W.J. Niessen. Automatic construction of multiple-object three-dimensional statistical shape models: application to cardiac modeling. *IEEE Transactions on Medical Imaging*, 21(9):1151–1166, 2002.

[39] J. Fripp, S. Crozier, S.K. Warfield, and S. Ourselin. Automatic segmentation and quantitative analysis of the articular cartilages from magnetic resonance images of the knee. *IEEE Transactions on Medical Imaging*, 29(1):55–64, 2009.

[40] D. Fritz, D. Rinck, R. Dillmann, and M. Scheuering. Segmentation of the left and right cardiac ventricle using a combined bi-temporal statistical model. *Medical Imaging 2006: Visualization, Image-Guided Procedures, and Display. Proceedings of the SPIE*, 6141:605–614, 2006.

[41] P.A. Habas, K. Kim, J.M. Corbett-Detig, F. Rousseau, O.A. Glenn, A.J. Barkovich, and C. Studholme. A spatiotemporal atlas of MR intensity, tissue probability and shape of the fetal brain with application to segmentation. *NeuroImage*, 2010.

[42] G. Hamarneh and X. Li. Watershed segmentation using prior shape and appearance knowledge. *Image and Vision Computing*, 27(1-2):59–68, 2009.

[43] J. Hansegard, F. Orderud, and S.I. Rabben. Real-Time Active Shape Models for Segmentation of 3D Cardiac Ultrasound. *Lecture Notes in Computer Science*, 4673:157, 2007.

[44] T. Heimann and H.P. Meinzer. Statistical shape models for 3D medical image segmentation: A review. *Medical Image Analysis*, 13:543–563, 2009.

[45] C. Jian-wu, P. Jian-ji, B. Peng-gang, Z. Wei, Z. Chun, C. Yun-bing, C. Shu-hua, L. Shao-jun, W. Jun-xin, N. Neoplasms, et al. Influ-

ence of CT-MRI fusion image on definition of the target volumes of nasopharyngeal carcinoma. *Tumor*, 2009.

[46] M.P. Jolly, H. Xue, L. Grady, and J. Guehring. Combining registration and minimum surfaces for the segmentation of the left ventricle in cardiac cine MR images. In *Proc. MICCAI2009*, volume 5762, pages 910–918, 2009.

[47] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331, 1988.

[48] M.R. Kaus, J. Berg, J. Weese, W. Niessen, and V. Pekar. Automated segmentation of the left ventricle in cardiac MRI. *Medical Image Analysis*, 8(3):245–254, 2004.

[49] P. Knaapen, S. De Haan, O.S. Hoekstra, R. Halbmeijer, Y.E. Appelman, J.G.J. Groothuis, E.F. Comans, M.R. Meijerink, A.A. Lammertsma, M. Lubberink, et al. Cardiac PET-CT: advanced hybrid imaging for the detection of coronary artery disease. 2010.

[50] J. Koikkalainen, T. Tolli, K. Lauerma, K. Antila, E. Mattila, M. Lilja, and J. Lotjonen. Methods of artificial enlargement of the training set for statistical shape models. *IEEE Transactions on Medical Imaging*, 27(11):1643–1654, 2008.

[51] H. Lamecker, M. Seebass, H.C. Hege, and P. Deuflhard. A 3d statistical shape model of the pelvic bone for segmentation. In *Proceedings of SPIE*, volume 5370, page 1341. Citeseer, 2004.

[52] K. Lekadir, R. Merrifield, and G.Z. Yang. Outlier detection and handling for robust 3-D active shape models search. *IEEE Transactions on Medical Imaging*, 26(2):212–222, 2007.

[53] M.E. Leventon, W.E.L. Grimson, and O. Faugeras. Statistical shape influence in geodesic active contours. In *Computer Vision and Pattern Recognition*, pages 316–323. Published by the IEEE Computer Society, 2000.

[54] B. Li and J.M. Reinhardt. Automatic generation of object shape models and their application to tomographic image segmentation. In *Pro-*

ceedings of SPIE, volume 4322, pages 311–322, 2001.

[55] K. Li, X. Wu, D.Z. Chen, and M. Sonka. Optimal surface segmentation in volumetric images-a graph-theoretic approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(1):119–134, 2006.

[56] J. Lie, M. Lysaker, and X.C. Tai. A binary level set model and some applications to Mumford–Shah image segmentation. *IEEE Transactions on Image Processing*, 15(5):1171, 2006.

[57] D. Lloyd-Jones, R. Adams, M. Carnethon, G. De Simone, T.B. Ferguson, K. Flegal, E. Ford, K. Furie, A. Go, K. Greenlund, et al. American Heart Association Statistics Committee and Stroke Statistics Subcommittee: Heart Disease and Stroke Statistics – 2009 Update: A Report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation*, 119(3):e21–e181, 2009.

[58] M.A. Lodge, H. Braess, F. Mahmoud, J. Suh, N. Englar, S. Geyser-Stoops, J. Jenkins, S.L. Bacharach, and V. Dilsizian. Developments in nuclear cardiology: transition from single photon emission computed tomography to positron emission tomography-computed tomography. *The Journal of Invasive Cardiology*, 17(9):491, 2005.

[59] C.H. Lorenz, E.S. Walker, V.L. Morgan, S.S. Klein, and T.P. Graham. Normal human right and left ventricular mass, systolic function, and gender differences by cine magnetic resonance imaging. *Journal of Cardiovascular Magnetic Resonance*, 1(1):7–21, 1999.

[60] M. Lorenzo-Valdés, G.I. Sanchez-Ortiz, A.G. Elkington, R.H. Mohiaddin, and D. Rueckert. Segmentation of 4D cardiac MR images using a probabilistic atlas and the EM algorithm. *Medical Image Analysis*, 8(3):255–265, 2004.

[61] M. Lorenzo-Valdés, G.I. Sanchez-Ortiz, R. Mohiaddin, and D. Rueckert. Atlas-based segmentation and tracking of 3D cardiac MR images using non-rigid registration. *Lecture Notes in Computer Science*, pages 642–650, 2002.

[62] J. Lötjönen, S. Kivistö, J. Koikkalainen, D. Smutek, and K. Lauerma. Statistical shape model of atria, ventricles and epicardium from short-

and long-axis MR images. *Medical Image Analysis*, 8(3):371–386, 2004.

[63] M. Lynch, O. Ghita, and P.F. Whelan. Left-ventricle myocardium segmentation using a coupled level-set with a priori knowledge. *Computerized Medical Imaging and Graphics*, 30(4):255–262, 2006.

[64] M. Lynch, O. Ghita, and P.F. Whelan. Segmentation of the left ventricle of the heart in 3-D+t MRI data using an optimized nonrigid temporal model. *IEEE Transactions on Medical Imaging*, 27(2):195–203, 2008.

[65] Z. Ma, J.M.R.S. Tavaresa, R.N. Jorgea, and T. Mascarenhasb. A review of algorithms for medical image segmentation and their applications to the female pelvic cavity. *Computer Methods in Biomechanics and Biomedical Engineering*, 13(2):235–246, 2010.

[66] R. Malladi, J.A. Sethian, and B.C. Vemuri. Shape modeling with front propagation: A level set approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(2), 1995.

[67] O. McArdle, G. Lee, M. Clemons, and R. Dinniwell. CT/MRI Fusion Significantly Reduces the Risk of Geographic Miss When Ablating the Ovaries by Radiotherapy. *International Journal of Radiation Oncology Biology Physics*, 72(1):S512, 2008.

[68] K.P. McGee, E.E. Williamson, and Julsrud P., editors. *Mayo Clinic Guide to Cardiac Magnetic Resonance Imaging.* Informa Healthcare, 2008.

[69] S.C. Mitchell, J.G. Bosch, B.P.F. Lelieveldt, R.J. van der Geest, J.H.C. Reiber, and M. Sonka. 3-D active appearance models: segmentation of cardiac MR and ultrasound images. *IEEE Transactions on Medical Imaging*, 21(9):1167–1178, 2002.

[70] S.C. Mitchell, B.P.F. Lelieveldt, R.J. van der Geest, H.G. Bosch, J.H.C. Reiver, and M. Sonka. Multistage hybrid active appearance model matching: segmentation of left and right ventricles in cardiac MR images. *IEEE Transactions on Medical Imaging*, 20(5):415–423, 2001.

[71] J. Montagnat and H. Delingette. 4D deformable models with temporal constraints: application to 4D cardiac image segmentation. *Medical Image Analysis*, 9(1):87–100, 2005.

[72] J. Montagnat, M. Sermesant, H. Delingette, G. Malandain, and N. Ayache. Anisotropic filtering for model-based segmentation of 4D cylindrical echocardiographic images. *Pattern Recognition Letters*, 24(4-5):815–828, 2003.

[73] K. Nakajo, M. Tatsumi, A. Inoue, K. Isohashi, I. Higuchi, H. Kato, E. Shimosegawa, and J. Hatazawa. Diagnostic accuracy in gynecological malignant tumors: A comparison between PET/CT and PET/MRI fusion images. In *Society of Nuclear Medicine Annual Meeting Abstracts*, volume 50, page 1706. Soc Nuclear Med, 2009.

[74] A. Neumann and C. Lorenz. Statistical shape model based segmentation of medical images. *Computerized Medical Imaging and Graphics*, 22(2):133–143, 1998.

[75] J.A. Noble and D. Boukerroui. Ultrasound image segmentation: a survey. *IEEE Transactions on Medical Imaging*, 25(8):987–1010, 2006.

[76] P.M. Novotny, J.A. Stoll, N.V. Vasilyev, P.J. Del Nido, P.E. Dupont, T.E. Zickler, and R.D. Howe. GPU based real-time instrument tracking with three-dimensional ultrasound. *Medical Image Analysis*, 11(5):458–464, 2007.

[77] S. Ordas, H.C. Assen, J. Puente, B.P.F. Lelieveldt, and A.F. Frangi. Parametric optimization of a model-based segmentation algorithm for cardiac MR image analysis: a grid-computing approach. *Studies in Health Technology and Informatics*, 112:146–156, 2005.

[78] S. Ordas, H.C. Van Assen, L. Boisrobert, M. Laucelli, J. Puente, B.P.F. Lelieveldt, and A.F. Frangi. Statistical modeling and segmentation in cardiac MRI using a grid computing approach. *Advances in Grid Computing-EGC 2005*, pages 6–15, 2005.

[79] S. Osher and J. Sethian. Fronts propagating with curvature-dependent speed- Algorithms based on Hamilton-Jacobi formulations. *Journal of Computational Physics*, 79(1):12–49, 1988.

[80] N. Paragios. A Variational Approach for the Segmentation of the Left Ventricle in Cardiac Image Analysis. *International Journal of Computer Vision*, 50(3):345–362, 2002.

[81] H. Park, P.H. Bland, and C.R. Meyer. Construction of an abdominal probabilistic atlas and its application in segmentation. *IEEE Transactions on Medical Imaging*, 22(4):483–492, 2003.

[82] D.J. Pennell, U.P. Sechtem, C.B. Higgins, W.J. Manning, G.M. Pohost, F.E. Rademakers, A.C. Van Rossum, L.J. Shaw, and E.K. Yucel. Clinical indications for cardiovascular magnetic resonance (CMR): Consensus Panel report. *European heart journal*, 25(21):1940, 2004.

[83] D. Perperidis, R. Mohiaddin, and D. Rueckert. Construction of a 4D statistical atlas of the cardiac anatomy and its use in classification. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2005*, pages 402–410, 2005.

[84] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling. *Numerical Recipes in Fortran 77: The Art of Scientific Computing*. Cambridge University Press, 2 edition, 1992.

[85] P. Rerkpattanapipat, W. Mazur, K.M. Link, and W.G. Hundley. Assessment of cardiac function with MR imaging. *Magnetic resonance imaging clinics of North America*, 11(1):67, 2003.

[86] T. Rohlfing, R. Brandt, R. Menzel, C.R. Maurer, et al. Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains. *NeuroImage*, 21(4):1428–1442, 2004.

[87] T. Rohlfing, D.B. Russakoff, and C.R. Maurer Jr. Performance-based classifier combination in atlas-based image segmentation using expectation-maximization parameter estimation. *IEEE Transactions on Medical Imaging*, 23(8):983, 2004.

[88] D. Rueckert, A.F. Frangi, and J.A. Schnabel. Automatic construction of 3D statistical deformation models using non-rigid registration. In *Proc. MICCAI 2001*, pages 77–84, 2001.

[89] J. Sandstede, C. Lipke, M. Beer, S. Hofmann, T. Pabst, W. Kenn, S. Neubauer, and D. Hahn. Age-and gender-specific differences in left and right ventricular cardiac function and mass determined by cine magnetic resonance imaging. *European radiology*, 10(3):438–442, 2000.

[90] M.F. Santarelli, V. Positano, C. Michelassi, M. Lombardi, and L. Landini. Automated cardiac MR image segmentation: theory and measurement evaluation. *Medical Engineering & Physics*, 25(2):149–159, 2003.

[91] J.A. Sethian. A fast marching level set method for monotonically advancing fronts. *Proceedings of the National Academy of Sciences of the United States of America*, 93(4):1591, 1996.

[92] M. Sezgin and B. Sankur. Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic imaging*, 13(1):146–168, 2004.

[93] A.K. Singh, J. Kruecker, S. Xu, N. Glossop, P. Guion, K. Ullman, P.L. Choyke, and B.J. Wood. Initial clinical experience with real-time transrectal ultrasonography-magnetic resonance imaging fusion-guided prostate biopsy. *BJU international*, 101(7):841–845, 2008.

[94] I. Sluimer, M. Prokop, and B. van Ginneken. Toward automated segmentation of the pathological lung in CT. *IEEE Transactions on Medical Imaging*, 24(8):1025, 2005.

[95] O.I.I. Soliman, S.W. Kirschbaum, B.M. van Dalen, H.B. van der Zwaan, B.M. Delavary, W.B. Vletter, R.J.M. van Geuns, F.J. Ten Cate, and M.L. Geleijnse. Accuracy and reproducibility of quantitation of left ventricular function by real-time three-dimensional echocardiography versus cardiac magnetic resonance. *The American journal of cardiology*, 102(6):778–783, 2008.

[96] L.H. Staib and J.S. Duncan. Model-based deformable surface finding for medical images. *IEEE Transactions on Medical Imaging*, 15(5):720–731, 1996.

[97] M.B. Stegmann, H. Olafsdottir, and H.B.W. Larsson. Unsupervised motion-compensation of multi-slice cardiac perfusion MRI. *Medical Image Analysis*, 9(4):394–410, 2005.

[98] M.B. Stegmann and D. Pedersen. Bi-temporal 3D active appearance models with applications to unsupervised ejection fraction estimation. *Proceedings of SPIE Medical Image Processing*, 5747:336–350, 2005.

[99] Paul Suetens. *Fundamentals of Medical Imaging*. Cambridge University Press, 2009.

[100] A. Suinesiaputra, A.F. Frangi, M. Uzumcu, J.H.C. Reiber, and B.P.F. Lelieveldt. Extraction of myocardial contractility patterns from short-axes MR images using independent component analysis. *Lecture Notes in Computer Science*, pages 75–86, 2004.

[101] J.S. Suri, K. Liu, S. Singh, S.N. Laxminarayan, X. Zeng, and L. Reden. Shape recovery algorithms using level sets in 2-D/3-D medical imagery: A state-of-the-art review. *IEEE Transactions on Information Technology in Biomedicine*, 6(1), 2002.

[102] D. Terzopoulos, A. Witkin, and M. Kass. Constraints on deformable models: Recovering 3D shape and nonrigid motion. *Artificial Intelligence*, 36(1):91–123, 1988.

[103] C. Tobon-Gomez, C. Butakoff, S. Ordas, S. Aguade, and A.F. Frangi. Comparative study of diverse model building strategies for 3D-ASM segmentation of dynamic gated SPECT data. *Medical Imaging 2007: Physiology, Function, and Structure from Medical Images. Edited by Manduca, Armando; Hu, Xiaoping*, 2007.

[104] A. Tsai, A. Yezzi Jr, W. Wells, C. Tempany, D. Tucker, A. Fan, W.E. Grimson, and A. Willsky. A shape-based approach to the segmentation of medical imagery using level sets. *IEEE Transactions on Medical Imaging*, 22(2):137, 2003.

[105] A. Tsai, A. Yezzi Jr, and A.S. Willsky. Curve Evolution Implementation of the Mumford–Shah Functional for Image Segmentation, Denoising, Interpolation, and Magnification. *IEEE Transactions On Image Processing*, 10(8):1169, 2001.

[106] M. Üzümcü, A.F. Frangi, J.H.C. Reiber, and B.P. Lelieveldt. Independent component analysis in statistical shape models. In *Proc. of SPIE*, volume 5032, pages 375–383. Citeseer, 2003.

[107] Dilsizian V. and Pohost G.M. *Cardiac CT, PET and MR*. Wiley-Blackwell, 2006.

[108] H.C. van Assen, M.G. Danilouchkine, F. Behloul, H.J. Lamb, R.J. Geest, J.H.C. Reiber, and B.P.F. Lelieveldt. Cardiac LV segmentation using a 3D active shape model driven by fuzzy inference. In *Proc. MICCAI 2003*, pages 533–540, 2003.

[109] H.C. van Assen, M.G. Danilouchkine, M.S. Dirksen, J.H. Reiber, and B.P. Lelieveldt. A 3-D active shape model driven by fuzzy inference: application to cardiac CT and MR. *IEEE Transactions on Information Technology in Biomedicine*, 12(5):595–605, 2008.

[110] H.C. van Assen, M.G. Danilouchkine, A.F. Frangi, S. Ordás, J.J.M. Westenberg, J.H.C. Reiber, and B.P.F. Lelieveldt. SPASM: A 3D-ASM for segmentation of sparse and arbitrarily oriented cardiac MRI data. *Medical Image Analysis*, 10(2):286–303, 2006.

[111] C. Xu and J.L. Prince. Snakes, shapes, and gradient vector flow. *IEEE Transactions on Image Processing*, 7(3):359–369, 1998.

[112] S. Zambal, J. Hladuvka, and K. Buhler. Improving segmentation of the left ventricle using a two-component statistical model. *Lecture Notes In Computer Science*, 4190:151–158, 2006.

[113] H. Zhang, A. Wahle, R. K. Johnson, T. D. Scholz, and M. Sonka. 4-D cardiac MR image analysis: left and right ventricular morphology and function. *IEEE Transactions on Medical Imaging*, 29(2):350–364, Feb 2010.

[114] X. Zhuang, K. Rhode, S. Arridge, R. Razavi, D. Hill, D. Hawkes, and S. Ourselin. An atlas-based segmentation propagation framework using locally affine registration–application to automatic whole heart segmentation. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2008*, pages 425–433, 2008.