UNIVERSITY OF CALIFORNIA

Los Angeles

Unified Deterministic/Statistical Deformable Models for Cardiac Image Analysis

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Computer Science

by

Sharath Kumar Gopal

2016

ABSTRACT OF THE DISSERTATION

Unified Deterministic/Statistical Deformable Models for Cardiac Image Analysis

by

Sharath Kumar Gopal

Doctor of Philosophy in Computer Science

University of California, Los Angeles, 2016

Professor Demetri Terzopoulos, Chair

This thesis proposes to fully automate the shape and motion reconstruction of non-rigid objects from visual information using a unified deterministic/statistical deformable model. The model enhances the global control of a statistical deformable model with local control, based on assumptions of the material properties of the non-rigid object being modeled. A Histogram of Oriented Gradients (HoG) based object detector for a 3D volume is proposed to compute initial model estimates that are crucial for automation. This thesis also develops a unified variational method for 4D (3D+time) non-rigid shape reconstruction with anatomical and temporal smoothness constraints. The proposed unified model and method are combined in a fully automated Computer Vision and Machine Learning based framework for the clinically important application of segmenting the myocardium in cardiac cine Magnetic Resonance images.

The dissertation of Sharath Kumar Gopal is approved.

Piotr J. Slomka

Luminita Aura Vese

Alan Loddon Yuille

Song-Chun Zhu

Demetri Terzopoulos, Committee Chair

University of California, Los Angeles

2016

To Suma and Gopal . . .

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

2001–2005	B.E. (Computer Science), R.V. College Engineering, Bengaluru, Karnataka, India
2005–2008	Staff Software Engineer, National Instruments, Bengaluru.
2010-2012	Visiting Graduate Student, Artificial Intelligence in Medicine, Cedars-Sinai Medical Center
2010-2012	Graduate Student Researcher, UCLA Computer Science Department.
2012-2015	Teaching Fellow, Computer Science Department, UCLA.
2013-2014	Visiting Research Assistant, USC Institute for Creative Technologies.

PUBLICATIONS

Sharath Gopal, Demetri Terzopoulos, "Automated Model-Based Left Ventricle Segmentation in Cardiac MR Images," *Proc. 6th International Workshop on Statistical Atlases and Computational Models of the Heart (STACOM 2015)*, Munich, Germany, October, 2015, in *Statistical Atlases and Computational Models of the Heart: Imaging and Modelling Challenges*, Lecture Notes in Computer Science, Vol. 9534, O. Camara, T. Mansi, M. Pop, K. Rhode, M. Sermesant, A. Young (eds.), Springer-Verlag, Berlin, 2016, 3–12).

Sharath Gopal, Demetri Terzopoulos, "A Unified Statistical/Deterministic Deformable Model for LV Segmentation in Cardiac MRI," *Proc. 4th International Workshop on Statistical Atlases and Computational Models of the Heart (STACOM 2013)*, Nagoya, Japan, September, 2013, in *Statis-*

tical Atlases and Computational Models of the Heart: Imaging and Modelling Challenges, Lecture Notes in Computer Science, Vol. 8330, O. Camara, T. Mansi, M. Pop, K. Rhode, M. Sermesant, A. Young (eds.), Springer-Verlag, Berlin, 2014, 180–187.

Sharath Gopal, Yuka Otaki, Reza Arsanjani, Daniel Berman, Demetri Terzopoulos, Piotr Slomka, "Combining Active Appearance and Deformable Superquadric Models for LV Segmentation in Cardiac MRI," *Medical Imaging 2013: Image Processing*, N. Karssemeijer, E. Samei (eds.), Proc. SPIE 8669, Lake Buena Vista, FL, February 2013, 8669-15:1–8.

CHAPTER 1

Introduction

Advances in computer vision, image processing, machine learning and statistics have played a crucial role in improving the performance of automated software systems for medical image analysis. Greater computational resources, in the form of faster multi-core CPUs and GPUs, have enabled the deployment of resource-heavy automated software. All these developments contribute to the exciting prospect of having an Artificial Intelligence system without any human assistance accurately predict the condition of a patient's health.

High level modeling problems in medical image analysis include segmentation and shape reconstruction as well as registration and motion/deformation analysis of anatomical structures. All these problems have an underlying clinical motivation which would require the computation of indexes/parameters that might be used further to assess an individual's health. Such problems have benefited from pioneering research efforts in deformable models and the statistical analysis of shape.

Deformable models have played an important role in addressing the challenging problem of non-rigid shape and motion reconstruction from noisy image data. Both, deterministic and statistical deformable models assist in the incorporation of prior knowledge about an anatomical structure's shape, its temporal characteristics, and its appearance in a specific type of image (MR, CT, Ultrasound, etc). The incorporation of such prior knowledge increases the robustness of model-based image analysis algorithms, and it also potentially enables the full automation of systems using these algorithms.

The incorporation of useful priors into deformable models is one of the underlying themes of this thesis. In particular, a fully automated system is proposed and developed for the task of cardiac Left Ventricle (LV) reconstruction from cine Magnetic Resonance (MR) images.

1.1 Problem Statement

The myocardial wall in the LV is the main pumping structure of the heart and its function is important in the assessment of cardiovascular disease. By accurately segmenting the LV in cine MR images, cardiac contractile function can be quantified according to LV volumes and ejection fractions. Manual segmentation in MR images is a tedious process performed by clinicians, which is subject to inter-observer and intra-observer variability that can lead to inconsistent diagnosis. These issues have motivated researchers to develop automated methods that aspire to match the ability of expert clinicians to segment LV shape.

The recent survey by Petitjean and Dacher [2011] provides an overview of various methods that have been applied to LV segmentation in MR images. An earlier survey can be found in [Frangi et al., 2001]. The great amount of effort that has been invested by researchers on this problem in the past two decades clearly indicates that this is an important problem of interest to the medical imaging and cardiology communities. A number of challenges conducted at workshops and conferences [Radau et al., 2009; Suinesiaputra et al., 2014] shows that the published, state-of-the-art solutions to this difficult problem still cannot compete with the ability of human experts to interpret cardiac MR images. This thesis addresses the challenges involved in closing this gap.

In this thesis, Statistical and Deterministic deformable models are employed to fully automate the segmentation and reconstruction of the human cardiac Left Ventricle (LV) in cardiac cine Magnetic Resonance (MR) images.
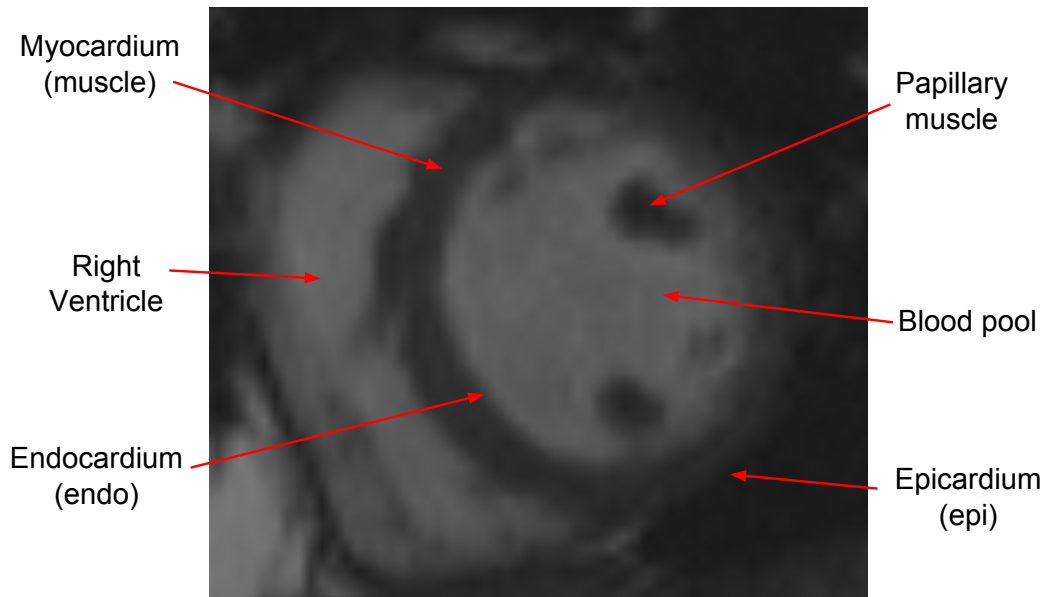
Figure 1.1: *Cross section of the LV*

1.1.1   The Cardiac Left Ventricle

The human heart is a muscular organ that functions as a pump for the circulation of blood in the body. It consists of the following four chambers: right atrium, left atrium, right ventricle (RV) and left ventricle (LV). While the two atria receive blood flowing to the heart, the two ventricles pump blood out of the heart. Among the two ventricles, the LV is the thickest muscle, which is in the shape of a bullet. This is due to the fact that, while the RV is responsible for pumping deoxygenated blood to the nearby lungs (pulmonary circulation), the LV must supply oxygenated blood to distant parts of the human body (systemic circulation). Hence, the study of the structure and function of LV is considered highly relevant to cardiovascular health.

A cross section of the LV (as captured in MR images) is shown in Figure 1.1. Myocardium forms the muscular wall of the the LV. The inner lining of this wall is called the Endocardium (endo) and the outer lining is called the Epicardium (epi). The myocardium encloses a (bright) LV blood pool. It includes papillary muscles and trabeculae which can appear in the form of dark spots in MR images.
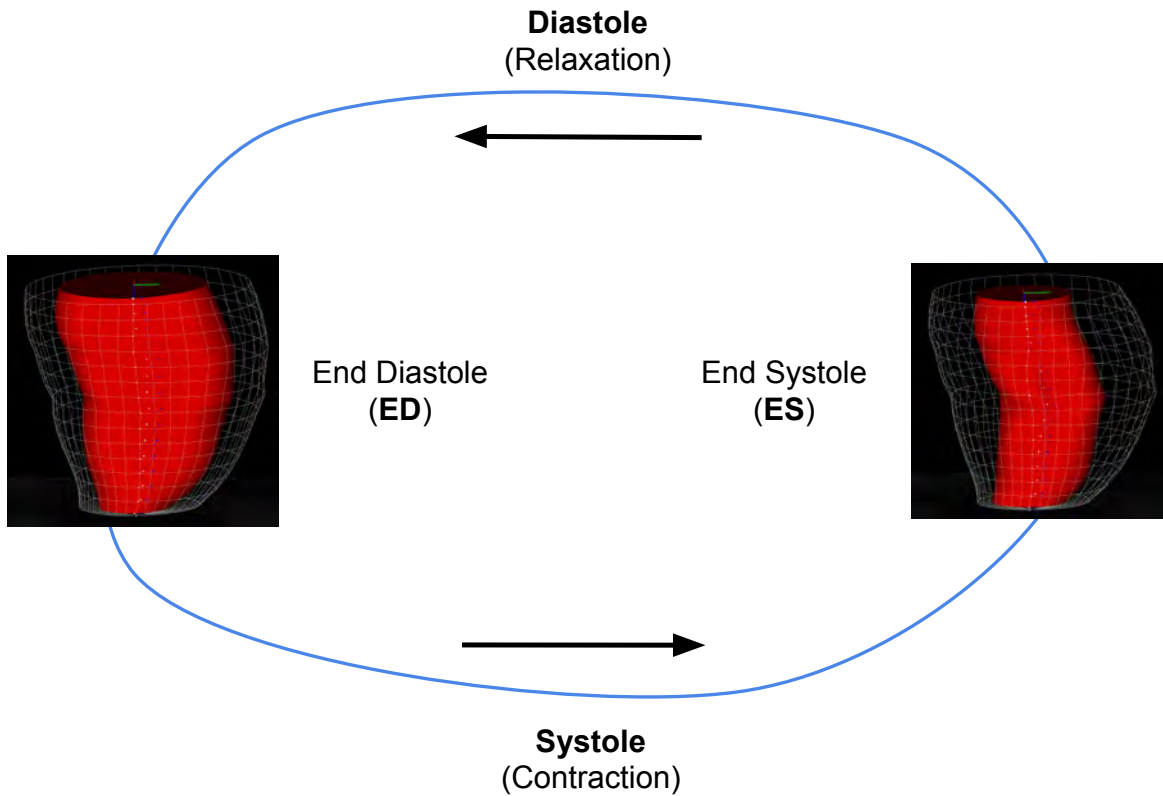
Figure 1.2: *A model of the LV during the Cardiac Cycle*

The LV is a dynamic structure that works continuously to maintain blood flow throughout the life of an individual. As illustrated in Figure 1.2, the LV goes through two distinct phases during a single cardiac cycle:

- **Diastole** — This is the phase when the LV is relaxed and its cavity is being filled with blood. End-Diastole (ED) is the end of this filling phase.

- **Systole** — This is the phase when the LV is contracting and blood is being pushed into the aorta. End-Systole (ES) is the end of this emptying phase.

1.1.2 Clinical Practice

Cardiovascular diseases (CVD) are a set of heart and blood vessel disorders that are the leading cause of death globally. The World Health Organization (WHO) has estimated that 17.5 million people died (31% of all global deaths) due to CVDs in the year 2012 [World Health Organization, 2016]. Studying LV function is important in clinical cardiology for the detection and management of CVDs in patients with risk factors (obesity, hypertension, diabetes, etc). Some of the main global indexes of LV function [Frangi et al., 2001] are as follows:

- End-Diastolic Volume (EDV) — Blood volume in the ED phase.

- End-Systolic Volume (ESV) — Blood volume in the ES phase.

- Stroke Volume (SV) — Volume of blood ejected between the ED and ES phases:

$$SV = EDV - ESV.$$

- Ejection Fraction (EF) — Percentage of blood ejected during systole:

$$EF = \frac{EDV - ESV}{EDV} \times 100.$$

- Myocardium Mass — Mass of the LV muscle.

1.1.3 Cardiac Cine MR Imaging

Modern techniques to diagnose, assess and treat cardiovascular diseases rely on cardiac imaging modalities such as Computerized Tomography (CT), Coronary Angiography, Ultrasound, and Magnetic Resonance (MR) Imaging. Cardiac MR is a common imaging technique that can be used to study cardiac anatomy, global function, motion/deformation, and perfusion [Frangi et al.,
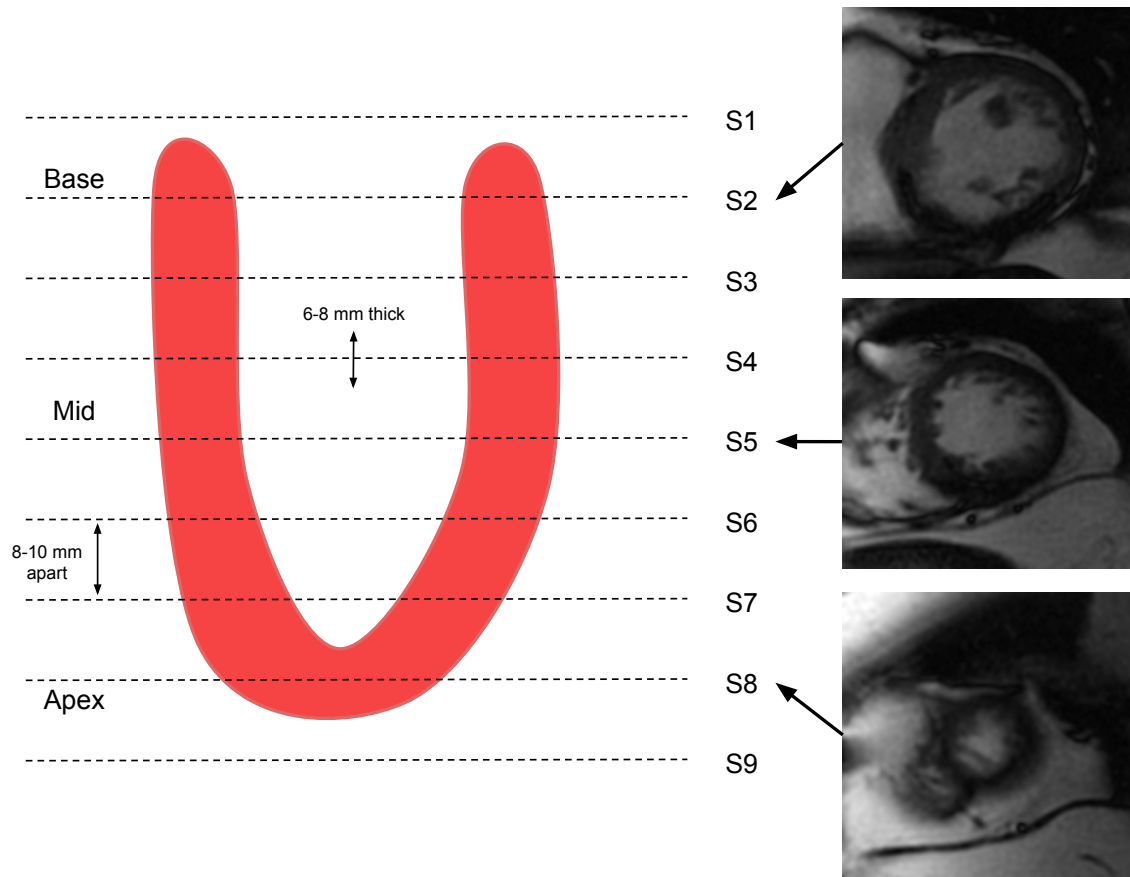
Figure 1.3: *Short Axis Cardiac MRI*

2001; Suinesiaputra et al., 2014]. It has the following advantages: (i) it is non-invasive, (ii) it uses non-ionizing radiation, (iii) it is capable of imaging at arbitrary and multiple planes, and (iv) it has high soft-tissue contrast between blood and myocardium without requiring a contrast medium.

Figure 1.3 shows short-axis cine/anatomical cardiac MR consisting of a stack of 7–12 slices from base to apex of the LV, where the slices are 8–10 mm apart and 6–8 mm thick. These images have a spatial resolution of about 1.1–1.4 mm. Each of the slices has good temporal resolution with 15–30 frames that cover the entire cardiac cycle. The frames are obtained by echocardiogram (ECG) gating.

Due to its non-invasive nature and accuracy, cardiac cine MR imaging is a widely accepted standard for computing indexes of global LV function in clinical practice. The myocardium mass, blood

(a) Partial Volume Effect (b) Intensity Inhomogeneity (c) Weak Epi Gradients

(d) Slice Misregistration
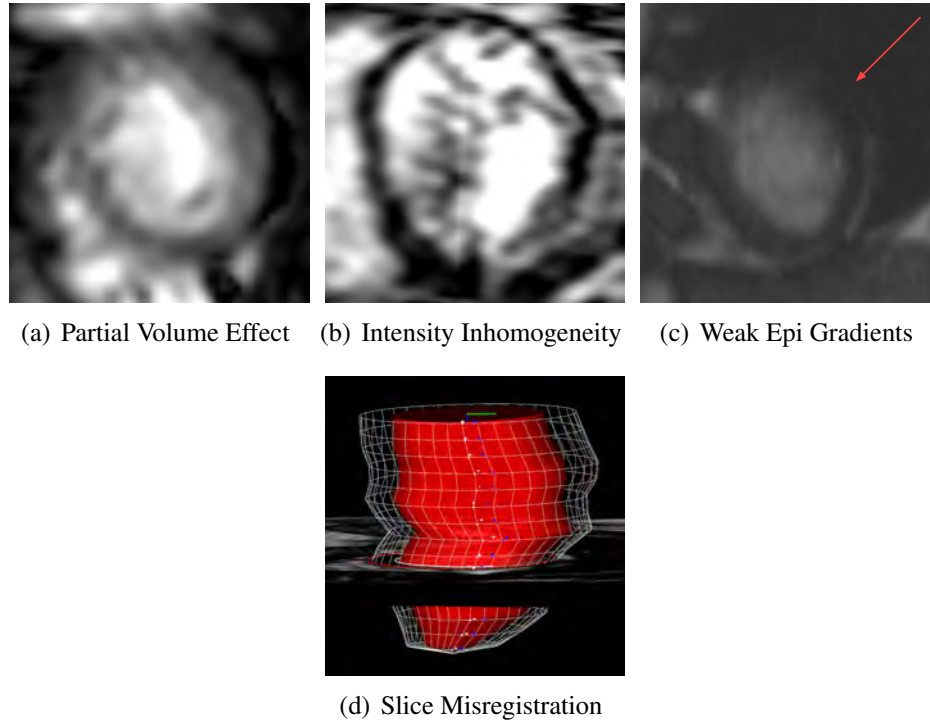
Figure 1.4: *Challenges in cardiac MRI*

volumes (EDV, ESV), and ejection fraction (EF) can be computed by delineating the inner (endo) and outer (epi) walls of the myocardium in the ED and ES phases. A clinician typically takes 10–15 minutes to perform this task. Manual delineation of the myocardium boundaries for all frames (15–30) is very time consuming, and it is subject to inter/intra observer variability. This serves as strong motivation for research efforts to fully automate the LV shape reconstruction process [Petitjean and Dacher, 2011; Suinesiaputra et al., 2014; Radau et al., 2009]. The main focus of this thesis is the reconstruction of LV shape and the tracking of its nonrigid motion throughout the cardiac cycle from visual information in cardiac cine MR images.

1.1.4 Challenges

Cardiac MR imaging and acquisition artifacts can pose hard challenges to image analysis algorithms (Figure 1.4). Due to the thickness of the image slices, partial volume effects manifest themselves in the form of fuzzy myocardium boundaries (especially towards the apex) and blood flow

artifacts (Figures 1.4(a) and 1.4(b)). Intensity inhomogeneity and noise can mislead algorithms that make assumptions about the range of intensity values for the myocardium pixels. Weak gradients at the myocardium-lung interface (Figure 1.4(c)) make it hard to separate the myocardium pixels with certainty. Patient motion and the acquisition of image slices over separate breathholds can cause the slices to be misregistered (Figure 1.4(d)), yielding a non-smooth LV shape reconstruction. Due to a distance of 8–10 mm between successive image slices, intensity values for voxels lying between two slices must be computed by interpolation. Finally, as demonstrated in [Suinesiaputra et al., 2014], the presence of papillary muscles and trabeculae structures cause disagreements among expert clinicians, yielding different ground truth myocardium labels.

A fully automated system should be robust enough to deal with these challenges, produce good LV shape reconstruction that does not deviate significantly from a consensus ground truth [Suinesiaputra et al., 2014] and, finally, be able to compute the global clinical indexes.

1.2 Contributions of the Thesis

Fully automated systems are developed in this thesis for the static (ED only) and dynamic (4D; i.e., 3D+time) reconstruction of the cardiac LV from cardiac cine MR images. To our knowledge, the three novel contributions (Figure 1.5) are as follows:

1. A Unified Deformable Model is proposed that combines the strengths of both deterministic and statistical deformable models.

2. A Histogram-of-Oriented-Gradients (HoG) based LV detector is proposed for estimating the position and scale of the LV in a given MR volume.

3. A Unified Variational Method is proposed for volume-constrained LV reconstruction in 4D cardiac images. Along the way, the Simultaneous Inverse Compositional (SIC) and Project-out Inverse Compositional (POIC) variants [Matthews and Baker, 2004] of the 2D Active Appearance Model are extended and applied to the 3D LV shape.
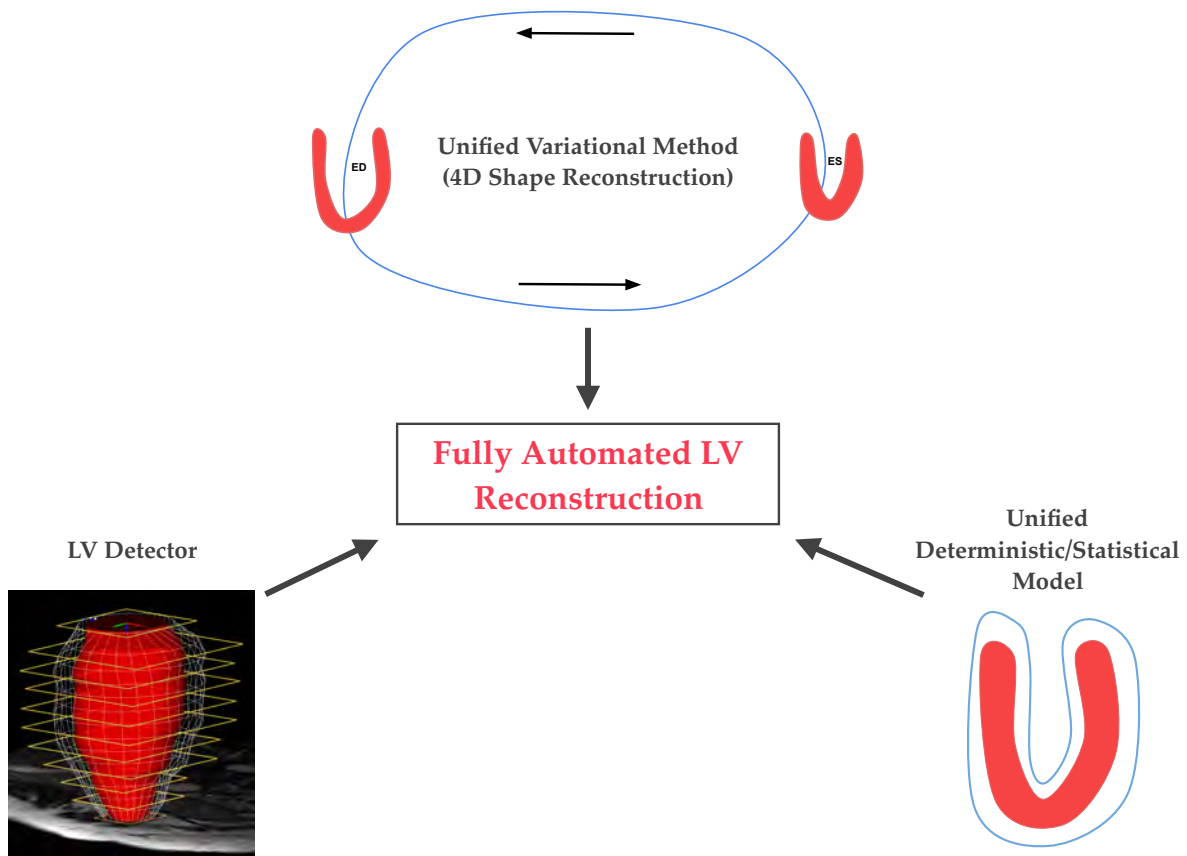
Figure 1.5: *Contribution*

We also develop UniMR, a software framework for interacting with and visualizing cardiac MR volumes. UniMR also facilitates the easy prototyping of new models.

The details of each of these contributions are provided in the subsequent chapters.

1.3   Thesis Overview

The focus of this work is on the computer vision and machine learning based algorithms and models, and not on the clinical aspects of the problem. In the remainder of this dissertation, the terms "reconstruction" and "segmentation" are used interchangeably, and they both refer tp the

recovery of LV shape from cardiac image data.

This organization of the remainder of the dissertation is as follows:

Chapter 2 discusses some of the relevant works found in the LV reconstruction research literature. It provides the context for the main contributions mentioned above.

Chapter 3 describes deterministic and statistical models of shape that are used in this work.

Chapter 4 proposes a unified model that combines the deterministic and statistical modes of deformation. This has been adapted from [Gopal and Terzopoulos, 2014].

Chapter 5 proposes a Histogram of Oriented Gradients (HoG) based LV detector that is used to obtain accurate initial model estimates. This is adapted from [Gopal and Terzopoulos, 2016].

Chapter 6 describes the automation of static and dynamic LV reconstruction. Section 6.1 proposes a fully automated system for static reconstruction of the LV using HoG detectors and unified deformable models. This is adapted from [Gopal and Terzopoulos, 2016]. Section 6.2 develops a unified variational method for 4D LV reconstruction with volume and smoothness constraints on the shape.

Chapter 7 provides details about the experiments conducted and the metrics used for evaluating the performance of the unified model, the HoG detector, the automated static reconstruction method, and the automated dynamic reconstruction method.

Chapter 8 concludes the document by summarizing the work done and discussing promising directions for future research.

Appendix A provides a description of all the features implemented in the UniMR visualization interface.

Appendix B describes mathematical tools that are frequently used in this thesis.

Finally, Appendix C provides the details of the variational method described in Section 6.2.

CHAPTER 2

Literature Review

Many methods have been developed in the past two decades for reconstructing the left ventricle (LV) in cardiac MR images. Enumerating and describing each of these methods is beyond the scope of this thesis, and the interested reader is referred to the detailed surveys by Petitjean and Dacher [2011] and Frangi et al. [2001]. Instead, some of the methods that most closely directly related to the work in this thesis are briefly reviewed.

The amount of prior knowledge incorporated during the reconstruction process serves as the main criterion in [Petitjean and Dacher, 2011] for grouping the different methods that have been reported in the literature. Prior information can be incorporated in the form of statistical models of LV shape and appearance in MR images, spatial assumptions specific to the image type (left ventricle is to the right side of the right ventricle in short-axis MR images), anatomical assumptions (cross section shows a circular and smooth LV), and knowledge about heart biomechanics (myocardium that is smooth and volume-preserving across time).

Techniques that incorporate no prior or some form of weak prior are those that rely, almost exclusively, on visual cues in the form of image intensity/gray value, image features (corners, edges), and local region statistics. Intensity thresholding, k-means clustering, and pixel classification using Gaussian Mixture Models (GMM) are some of the algorithms used by these methods to label a given pixel's tissue type (myocardium, blood, fat, or the background). On the other hand, methods that incorporate strong priors are the ones that use cardiac MR atlases (describing the different organs/structures present in the images), or use statistical models for LV shape and texture. Such methods tend to be more robust, but their main disadvantage is that they require manually seg-

mented images for an off-line training procedure.

The flexibility and versatility of deformable models [Terzopoulos, 1986; Kass et al., 1988; Terzopoulos and Metaxas, 1991] allow them to be used under both the weak-prior and strong-prior categories [McInerney and Terzopoulos, 1996]. Their physically-based variational framework can be augmented to incorporate strong priors of LV shape and myocardium material properties. They can deform into a variety of shapes, which makes them well suited for reconstructing the non-rigid LV. Such physically-based models can be naturally extended to include LV temporal dynamics priors for the purpose of 4D shape reconstruction. Hence, a single model for the LV has the ability to incorporate a rich set of priors and deform under the influence of external image forces and internal biomechanical constraint forces, yielding physically-consistent solutions.

The focus for the rest of the chapter is on briefly describing some methods that are similar to or directly related to what is proposed in this thesis, a fully automated, model-based approach for 4D shape reconstruction of the LV from cardiac MR. Prior such methods achieve one or more of the following: (i) use deterministic or statistical deformable models, (ii) reconstruct the LV across the entire cardiac cycle (2D+time or 3D+time), and (iii) achieve full automation.

2.1   Related Methods

Active Shape Models (ASM) [Cootes et al., 1995] and Active Appearance Models (AAM) [Cootes et al., 1998; Matthews and Baker, 2004] are statistical deformable models that encode strong priors of shape and appearance of a non-rigid object using Principal Components Analysis (PCA) (Appendix B.3). The AAM has been used by Mitchell et al. [2002] to train static 3D shape and appearance models of the LV for the ED phase, and eventually used to reconstruct the LV shape from a new patient's MR image volume. Although not a full 4D reconstruction method, Stegmann and Pedersen [2005] extended the static 3D AAM model of Mitchell et al. [2002] to a bi-temporal model by combining the ED and ES shape and appearance vectors while applying PCA.

To combine the robustnuss of global image matching in AAMs with the local image feature search of ASMs, a multistage hybrid AAM/ASM matching procedure is proposed by Mitchell et al. [2001] to improve the LV boundary localization accuracy. With similar motivation, a combination of deterministic and statistical deformable models is proposed to segment the ED and ES phases by [Gopal et al., 2013]. While a trained 3D AAM is used to segment the ED phase, a deformable superquadric model [Terzopoulos and Metaxas, 1991] is used to track the LV from the ED phase to the ES phase with the help of a learned ED-to-ES correspondence model.

Hybrid models are also employed by Zhang et al. [2010], who propose a two-step approach to segment the LV and the RV. The first step requires manual initialization for a 4D model that fits to image data using a variant of the multistage hybrid ASM/AAM procedure [Mitchell et al., 2001]. This segmentation is used as an initial estimate for the second step, which uses a single 3D AAM model (trained on all phases) to segment each phase separately. Lelieveldt et al. [2001] extend the AAM framework to a 2D+time Active Appearance Motion Model (AAMM) (similar to the spatio-temporal models of [Hamarneh and Gustavsson, 2004]), by considering an entire phase-normalized image sequence as one shape/intensity sample while applying PCA.

Kaus et al. [2004] apply a triangular mesh-based deformable surface model to fully-automated LV reconstruction in the ED and ES phases. A Point Distibution Model (PDM), built using PCA, is used to constrain the solution to the set of variations seen in the training shapes. A learned deterministic feature-model based external energy term drives the triangular mesh towards interesting image features. Montagnat and Delingette [2005] used a simplex-mesh in a classical Newtonian deformable model framework. The model could include weak spatial constraints by having the vertex's mean curvature tend towards the weighted average mean curvature of the neighborhood, or it could include strong spatial constraints by making the deformable surface tend towards a reference shape. Similarly, weak temporal constraints were included by attracting a vertex to the centroid of the neighboring vertices, or it could include strong temporal constraints by making use of reference vertex trajectory priors.

Lynch et al. [2008] model the temporal movement over the cardiac cycle by estimating inverted

Gaussians at each voxel in the narrowband of a level set. A two step update procedure, similar to Expectation Maximization (EM) [Dempster et al., 1977], is used to iteratively refine the initial model estimates that are obtained via fast marching method. The authors also mention the use of coupled level sets to delineate both the endo and the epi by using two inverted Gaussians. Lorenzo-Valdés et al. [2004] used Expectation Maximization (EM) and Markov Random Fields (MRF) in an iterative scheme to estimate the Gaussian distribution parameters of tissue classes corresponding to LV, RV, myocardium, and background. The 4D MRF incorporates spatial and temporal contexts, and a 4D probabilistic atlas is used as a priori information to further constrain the solution across space and time. Automation was achieved by registering a 3D intensity template to the test volume and computing the initial model parameters.

Variational methods have frequently been used for solving shape and motion recovery problems in computer vision (examples include [Terzopoulos, 1986], [Kass et al., 1988], [Chan and Vese, 2001], and [Horn and Schunck, 1981]). Paragios [2002] proposes a variational method that uses a 2D level-set implementation of the Geodesic Active Contour [Caselles et al., 1997] that incorporates a curvature term, a boundary term based on Gradient Vector Flow (GVF) [Xu and Prince, 1997], a region term [Zhu and Yuille, 1996], and an anatomical constraint term that couples the endo and epi contours.

SCR [Jolly et al., 2012] and INR [Margeta et al., 2012] are two fully automated methods that have been reported on the Cardiac Atlas Project (CAP) consensus database [Suinesiaputra et al., 2014]. SCR is a registration-based method for 4D LV segmentation. Before segmenting the LV, the first harmonic of a temporal Fourier transform is examined to detect the blood pool and a machine learning algorithm is used to detect the mitral valve anchor points. The segmentation technique is based on an inverse consistent deformable registration approach that is executed on each slice in a sequential manner. The other method, INR, is a machine-learning-based 4D LV segmentation approach that makes no use of prior LV shape knowledge. Instead, voxels are classified as myocardium or background by two layers of spatio-temporal decision forests.

2.2 Summary

This chapter discussed the related research literature that has been reported for LV reconstruction from cardiac cine MR images. The benefits of using a physically-based deformable model framework was discussed in the context of prior information-based classification of the reported methods. Finally, some of the reported methods that are directly related to this thesis were described.

CHAPTER 3

Deformable Models

This chapter introduces the unique features of deterministic and statistical deformable models and highlights the important differences between them. The formulation for each model is described with examples to illustrate the modes of deformation and the range of shapes that can be produced.

3.1 Deterministic Deformable Models

Snakes [Kass et al., 1988] and Deformable Superquadrics [Terzopoulos and Metaxas, 1991] are energy-minimizing deterministic models that provide a versatile framework for recovering the geometry and dynamics of non-rigid objects from noisy visual or range data. These are physics-based models due to the physical meaning of the energy they try to minimize. For instance, the models by Terzopoulos and Metaxas [1991] and Terzopoulos and Witkin [1988] use a spline deformation energy to mimic the properties of an elastic material. These models have a Lagrangian dynamics formulation, wherein the degrees of freedom (DOFs) of the model are made functions of time and become generalized coordinates. The stationary point of the action integral's functional derivative yields the following governing equations of motion:

$$\mathbf{M}\ddot{\mathbf{q}} + \mathbf{C}\dot{\mathbf{q}} + \mathbf{K}\mathbf{q} = \mathbf{f}_q + \mathbf{g}_q, \tag{3.1}$$

where $\mathbf{q}$ is the vector of the generalized coordinates, with overstruck dots denoting time derivatives, $\mathbf{M}$ is the mass matrix, $\mathbf{C}$ is the damping matrix, $\mathbf{K}$ is the stiffness matrix, $\mathbf{g}_q$ are inertial forces,

(a) FEM skin (white) + 2D reference (red)

(b) Stiffness matrix: Red (< 0), Green (> 0), Black ($= 0$)
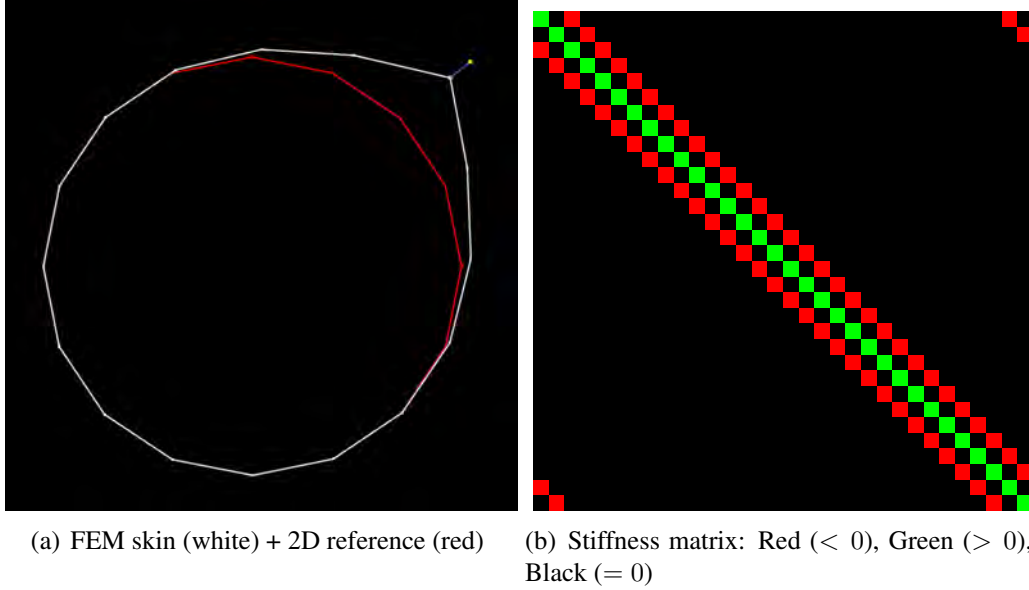
Figure 3.1: *Deterministic Deformable Model*

and $\mathbf{f}_q$ are the generalized forces associated with the degrees of freedom of the model.

The external forces applied on the model are converted to generalized forces using the principle of virtual work [Shabana, 2013], and finally, the above equations are discretized (using finite element or finite difference methods). A minimum energy configuration is attained by stepping the discrete equations in time, while simulating the dynamics under the influence of inertial, damping, deformation, and generalized forces.

Figure 3.1(a) illustrates a simple deterministic model that has a 2D reference shape (red circle with radius $r(t)$) representing its natural shape at a certain time. Similar to the formulation in [Terzopoulos and Metaxas, 1991], a finite element deformable "skin" is superimposed on the reference shape. A spline deformation energy that penalizes total displacement from the reference shape plus the total squared first partial derivative of the skin's shape, is imposed on the skin. Figure 3.1(b) shows the symmetric sparse banded (tridiagonal) structure of the stiffness matrix $\mathbf{K}$ associated with this deformation energy.

Interaction with a similar (higher resolution) model is illustrated with the pull and release sequence in Figure 3.2. When spring forces pull on the model's skin (Figures 3.2(a) and 3.2(c)), the refer-
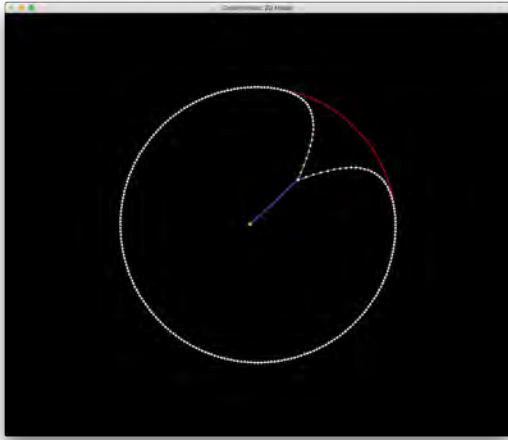
(a) Pull
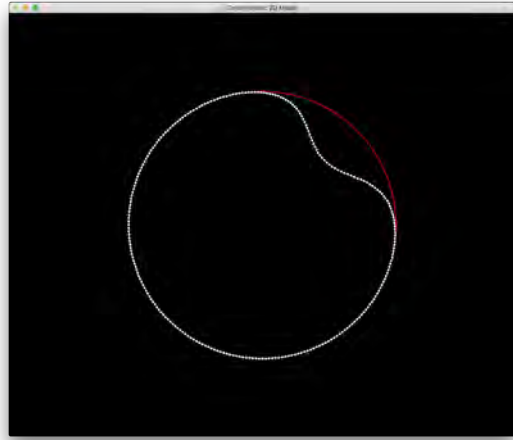
(b) Release
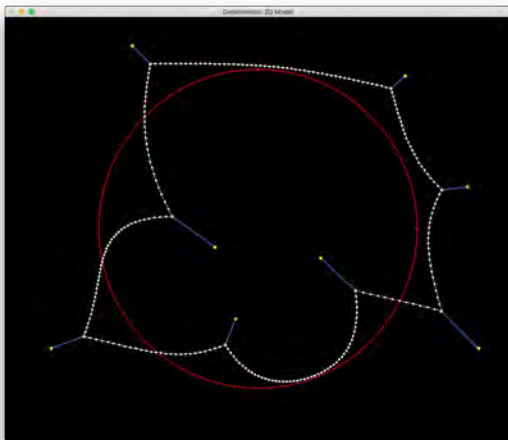
(c) Pull
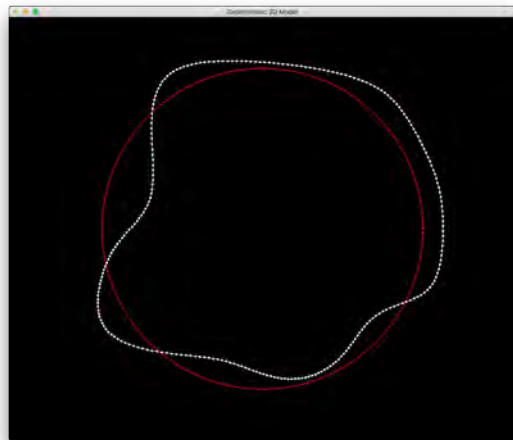
(d) Release
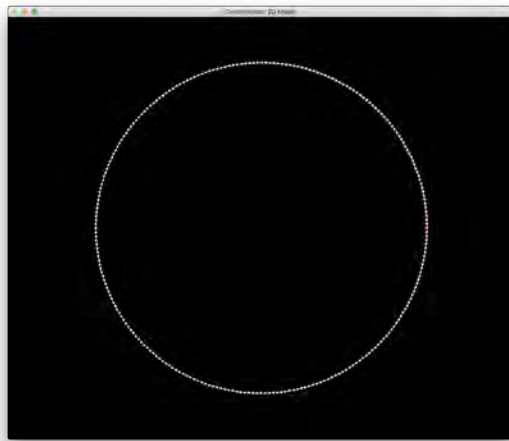
(e) Return to reference shape

Figure 3.2: *Interaction with a Deterministic Deformable Model*
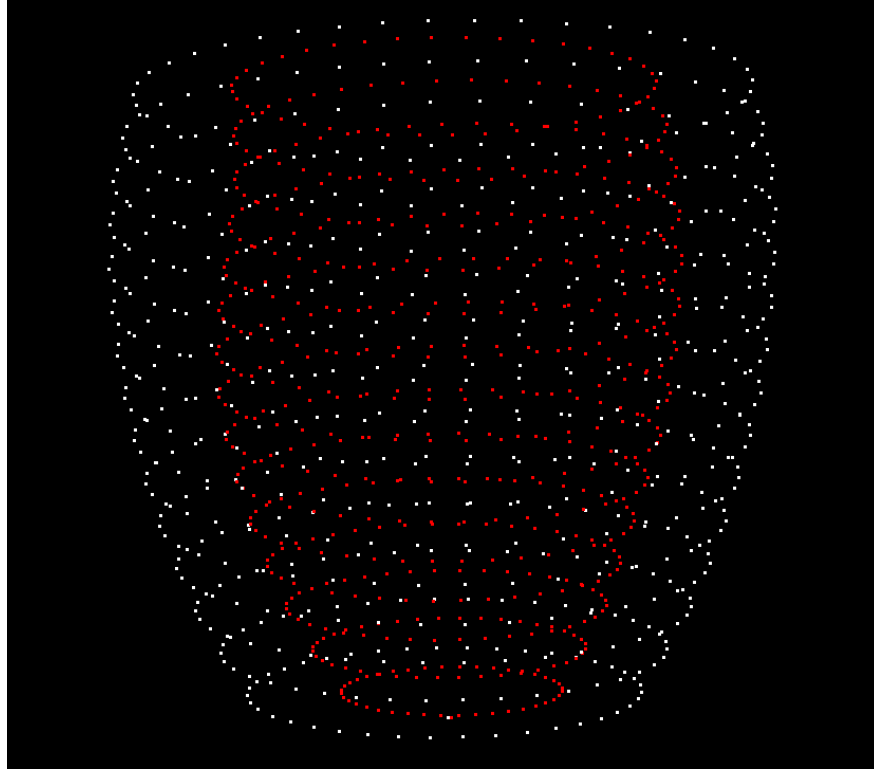
Figure 3.3: *Point Distribution Model of cardiac left ventricle*

ence shape evolves accordingly by increasing or decreasing the radius. Due to the displacement-penalizing deformation energy, the skin returns to the current reference shape (Figure 3.2(e)) when the springs are released. Due to the powerful combination of global (reference shape) and local (FE skin) deformation degrees of freedom, these models are able to generate a rich variety of shapes. For such purely deterministic models, the modes of deformation are decided by the geometric reference shape and the elastic properties of the skin encoded in the deformation energy.

3.2 Statistical Deformable Models

Complementary to deterministic models, statistical deformable models inherit their modes of deformation from the statistical variations present in a population of shapes. Building a Point Distribution Model (PDM) is one way to capture the statistical modes of deformation. Active Shape Models (ASM) [Cootes et al., 1995] and Active Appearance Models (AAM) [Cootes et al., 1998]

are examples of statistical deformable models that build a PDM from a training set of shapes, and use it to synthesize new shapes while fitting to visual data. The main motivation for such approaches is to constrain solutions to be valid instances of the object being modeled.

PDMs require that shapes in the population be represented by corresponding landmarks. Figure 3.3 shows the 3D landmarks marked on a cardiac left ventricle (LV). Each LV shape is represented by a vector of p 3D landmarks. In order to capture only the non-rigid variations, the shapes are aligned using Orthogonal Procrustes Analysis (OPA) (Appendix B.2). The final step involves applying Principal Components Analysis (PCA) (Appendix B.3) to the aligned shapes. Consequently, any shape $\mathbf{s} \in \mathcal{R}^{3p}$ can be represented as

$$\mathbf{s} = \bar{\mathbf{s}} + \mathbf{P}_s \mathbf{q}_s, \tag{3.2}$$

where $\bar{\mathbf{s}}$ is the average shape and columns of $\mathbf{P}_s \in \mathcal{R}^{3p \times m}$ represent the top m uncorrelated modes of variation in the sample population. The first mode is the direction of maximum data variance in the population. The top 4 modes are visualized on either side of the average shape in table 3.1. The first two modes show the z and xy scaling that exists across the population of shapes. It is interesting to note that since the sample population contains LV shapes from all phases of the cardiac cycle, the fourth mode represents the thickening of the myocardium wall. The statistical modes of deformation can be used to synthesize new shapes.

Since PCA fits an ellipsoid to the population, each synthesized shape can be assigned a probability based on its distance to the average shape. This probability is a function of the shape coefficients $\mathbf{q}_s$ and the variances (λ_i) along the axes of the ellipsoid (eigenvalues and eigenvectors of the covariance matrix), and it is proportional to

$$p(\mathbf{q}_s) \sim \exp\left[-\sum_{i=1}^{m} \frac{\mathbf{q}_s(i)^2}{\lambda_i} \right]. \tag{3.3}$$

The average shape $\bar{\mathbf{s}}$ has the highest probability, since it is the maximum likelihood estimate with respect to the sum of squared errors for the given population of shapes. As one moves away from the population average, the probability drops according to the above exponential function.
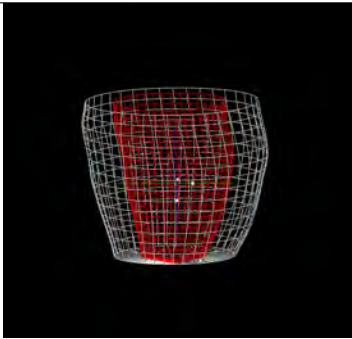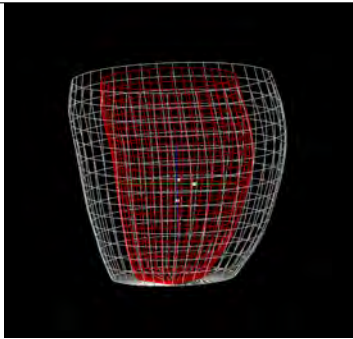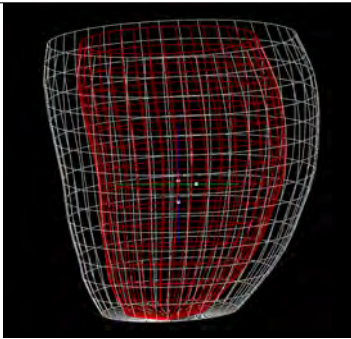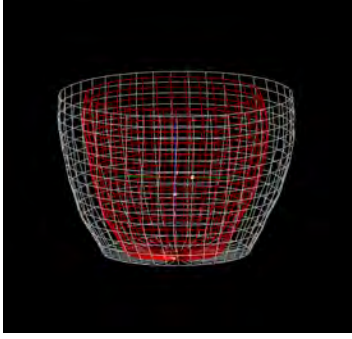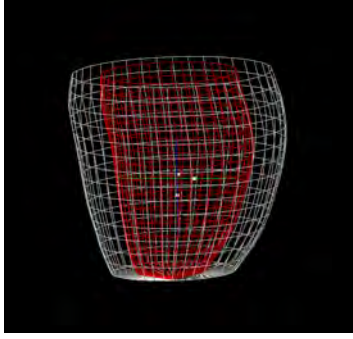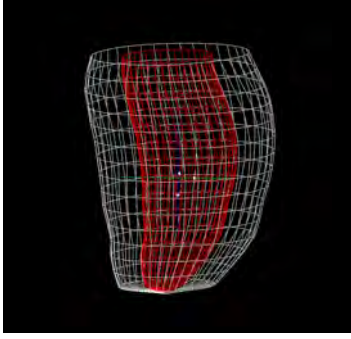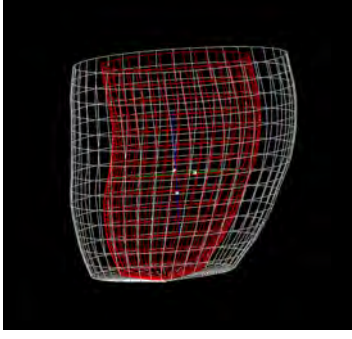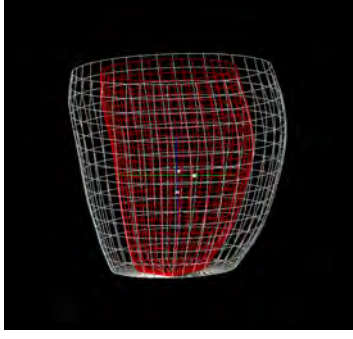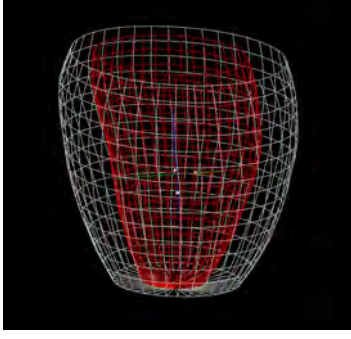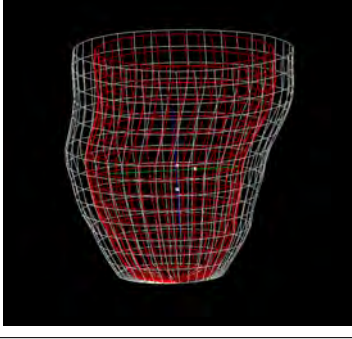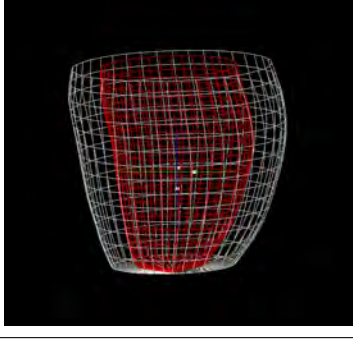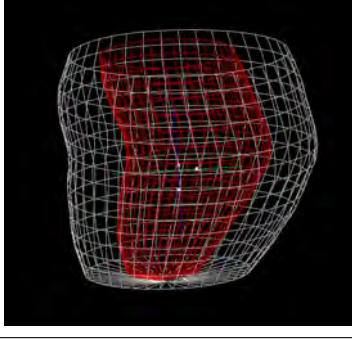
	Mean$-1.5\lambda_i$	Mean	Mean$+1.5\lambda_i$
Mode 1			
Mode 2			
Mode 3			
Mode 4			

Table 3.1: *Statistical Modes of Deformation*

3.3 Summary

In this chapter, we reviewed deterministic and statistical deformable models. The physically-based energy minimizing framework of deterministic models was described. A simple 2D model was used to illustrate the global and local deterministic modes of deformation as well as its behavior under the influence of external forces. Then, statistical deformable models were described using a PDM of the cardiac LV. The statistical modes of deformation learned from a population of LV shapes was illustrated. Finally, a probabilistic interpretation of the ellipsoid data fit from PCA was provided.

CHAPTER 4

A Unified Deterministic/Statistical Deformable Model

In this chapter, a novel deformable model is proposed that has both statistical and deterministic components suitable for LV segmentation in cardiac cine MR images. The statistical deformable component learns a global reference model of the LV using Principal Components Analysis (PCA), which serves as a strong shape prior, while the deterministic deformable component consists of a finite-element deformable surface coupled to the reference model. The statistical model accounts for the global variations in shape learned from the training set, while the deterministic deformable skin accounts for the local deformations consistent with the detailed image features when applying the model to new images.

4.1 Model Formulation

The formulation of our model mainly involves embedding the PCA reference as illustrated in Figure 4.1.

4.1.1 Geometry

The model is a closed surface that has $\mathbf{u} = (u, v)$ as its material coordinates. Principal Components Analysis (PCA) is applied to a set of aligned 3D LV shapes to obtain a discrete reference shape $\mathbf{s}(\mathbf{u})$ as

$$\mathbf{s}(\mathbf{u}) = \bar{\mathbf{s}}(\mathbf{u}) + \mathbf{P}_s(\mathbf{u})\mathbf{q}_s(\mathbf{u})^T, \tag{4.1}$$
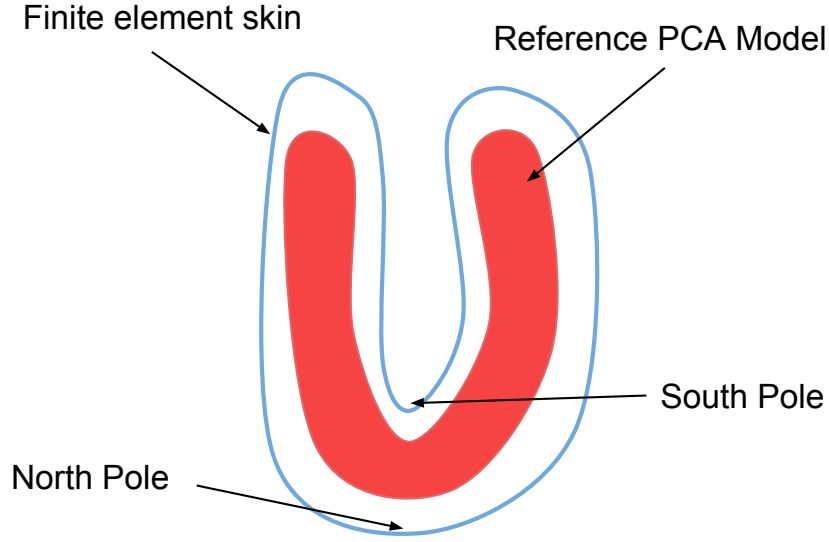
Figure 4.1: *Model Geometry*

where $\bar{\mathbf{s}}$ is the mean shape, the columns of matrix $\mathbf{P}_s$ are the modes of variation, and $\mathbf{q}_s$ are the the PCA shape parameters. The translational offsets across the training shapes are removed by translating the respective centroids to the origin, and the rotational offsets are removed using Ordinary Procrustes Analysis (Appendix B.2). The Jacobian of the PCA reference shape $\mathbf{s}$ is given by

$$\mathbf{J}(\mathbf{u}) = \frac{\partial \mathbf{s}(\mathbf{u})}{\partial \mathbf{q}_s(\mathbf{u})} = \mathbf{P}_s(\mathbf{u}), \tag{4.2}$$

thus characterizing how the shape changes when the parameters $\mathbf{q}_s$ change. The Jacobian is key to the interaction of external forces with the model dynamics described later.

A finite element deformable skin is superimposed on the reference shape to account for local deformations (Figure 4.2). The local displacements $\mathbf{d}(\mathbf{u})$ are expressed as a linear combination of finite element basis functions $\mathbf{b}_i(\mathbf{u})$ as follows:

$$\mathbf{d}(\mathbf{u}) = \sum_i \text{diag}(\mathbf{b}_i(\mathbf{u}))\mathbf{q}_i = \mathbf{S}\mathbf{q}_d, \tag{4.3}$$

where $\mathbf{q}_d = (..., \mathbf{q}_i, ...)^T$ is a set of local displacements $\mathbf{q}_i$ at each mesh node i, and $\mathbf{S}$ holds the

24
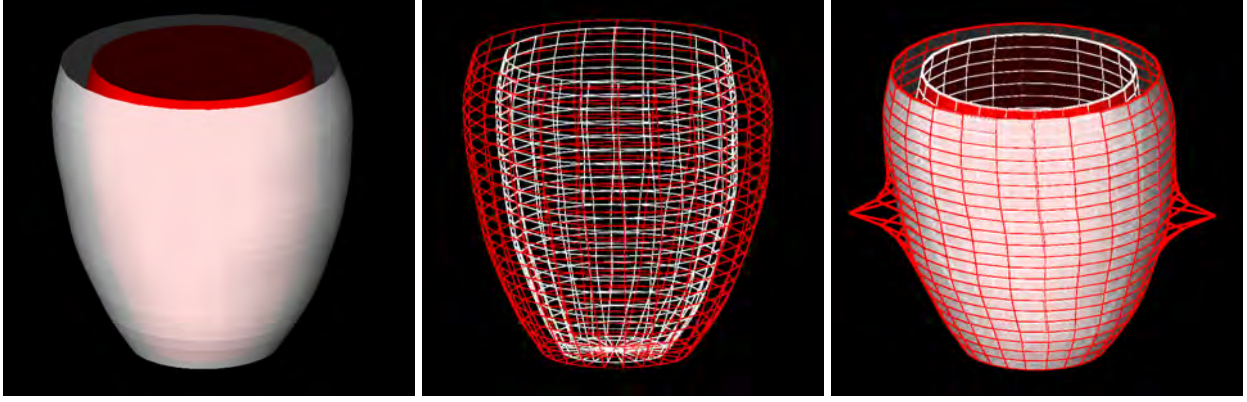
Figure 4.2: *PCA reference (left). Finite Element skin (center). Skin pulled away from the reference (right).*

basis functions. In addition to the PCA parameters $\mathbf{q}_s$ and the local displacement parameters $\mathbf{q}_d$, the unified model also has global translation and rotation parameters $\mathbf{q}_c$ and $\mathbf{q}_\theta$. All the degrees of freedom (DOFs) for the model are collected in a single vector

$$\mathbf{q} = (\mathbf{q}_c^T, \mathbf{q}_\theta^T, \mathbf{q}_s^T, \mathbf{q}_d^T)^T. \tag{4.4}$$

4.1.2 Dynamics

Given a new set of MR image slices for a patient, the vector $\mathbf{q}$ yielding a model that best fits the images must be computed. Applying Lagrangian dynamics, the model is made dynamic in $\mathbf{q}$, thus characterizing the evolution of $\mathbf{q}$ under the influence of external forces. The equations of motion are given as

$$\mathbf{C}\dot{\mathbf{q}} + \mathbf{K}\mathbf{q} = \mathbf{f}_q, \tag{4.5}$$

where $\dot{\mathbf{q}}$ is the time derivative of the DOF, $\mathbf{C}\dot{\mathbf{q}}$ are damping forces, $\mathbf{K}\mathbf{q}$ are elastic forces and $\mathbf{f}_q$ are external forces applied to the model. The stiffness matrix $\mathbf{K}$ determines the material/elastic properties of the finite element skin.

We impose a spline deformation energy on the local displacements $\mathbf{q}_d$ as follows:

$$E(\mathbf{q}_d) = \int w_1(\mathbf{u}) \left(\left(\frac{\partial \mathbf{q}_d}{\partial u} \right)^2 + \left(\frac{\partial \mathbf{q}_d}{\partial v} \right)^2 \right) + w_0(\mathbf{u}) \mathbf{q}_d^2 \, du \, dv, \tag{4.6}$$

where $w_0(\mathbf{u})$ controls the magnitude of the local deformation and $w_1(\mathbf{u})$ controls its variation across adjacent nodes on the skin.

The equations of motion in (4.5) are integrated through time using an explicit Euler method. The degrees of freedom in the vector $\mathbf{q}$ are updated from time t to time $t + \Delta t$ as follows:

$$\mathbf{q}^{(t+\Delta t)} = \mathbf{q}^{(t)} + \Delta t \left(\mathbf{C}^{(t)} \right)^{-1} \left(\mathbf{f}_q^{(t)} - \mathbf{K}\mathbf{q}^{(t)} \right). \tag{4.7}$$

Such a system will come to rest when the internal (damping and elastic) and external (image) forces equilibrate. Additional background details about the formulation and implementation are provided in [Terzopoulos and Metaxas, 1991].

4.2 Model Behavior

To illustrate the versatility of the model, each degree of freedom in $\mathbf{q}$ can be activated individually by applying external spring forces. Figures 4.3(a) and 4.3(b) show the rigid body DOFs when a single spring is pulling on one of the surface nodes. The non-rigid statistical modes of deformation of the model are shown in Figures 4.4(a) and 4.4(b), where only the first 2 and first 7 columns, respectively, of $\mathbf{P}_s$ in (4.1) are used. Similarly, Figure 4.4(c) shows the non-rigid deterministic modes of deformation of the finite element skin as described by the stiffness matrix $\mathbf{K}$. To illustrate the model's ability to cope with multiple body forces, springs are attached at different surface nodes as shown in Figure 4.5. Due to the combination of deterministic and statistical modes, the model has a richer shape subspace; hence, it can produce complex shapes that are not possible using just the statistical modes of deformation.
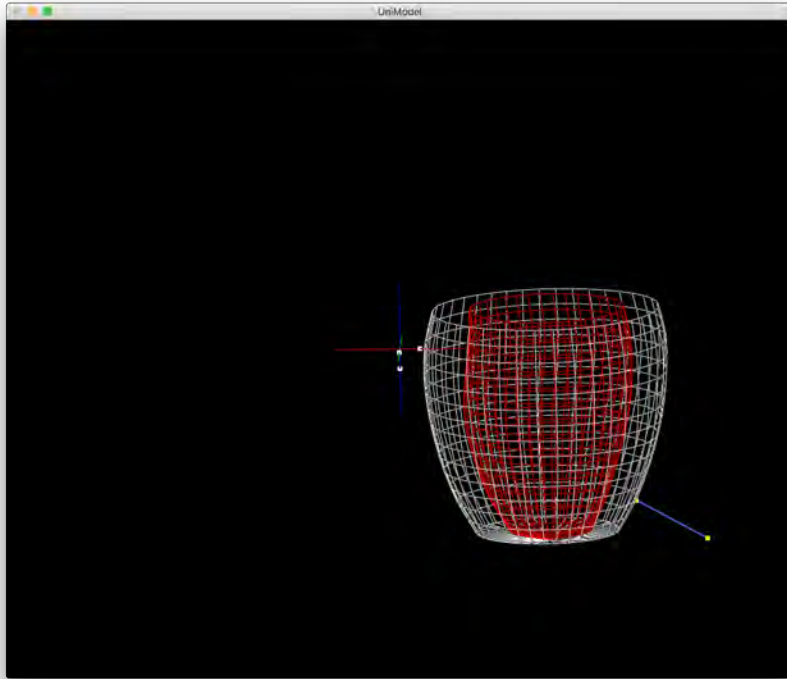
(a) Translation
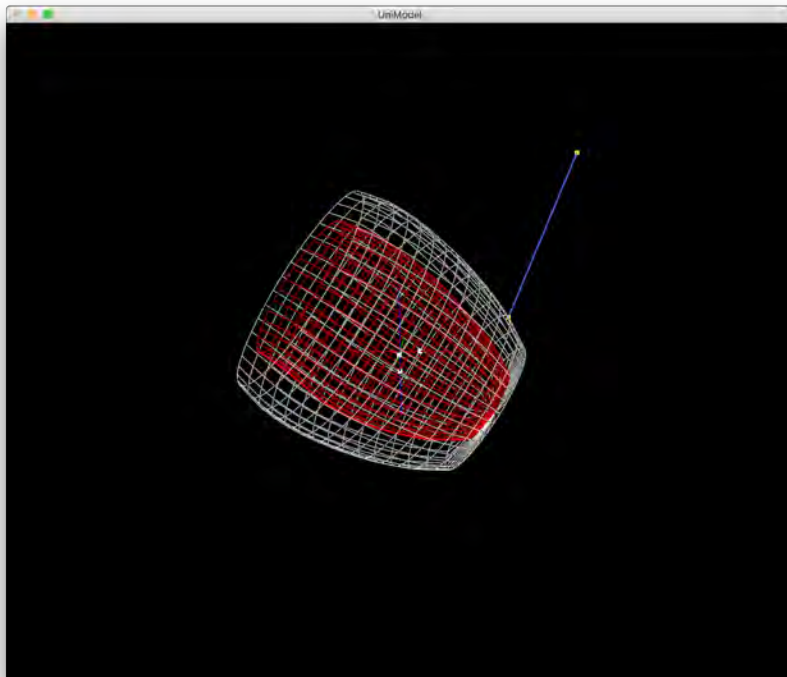
(b) Rotation

Figure 4.3: *Unified Model — Rigid Degrees of Freedom*
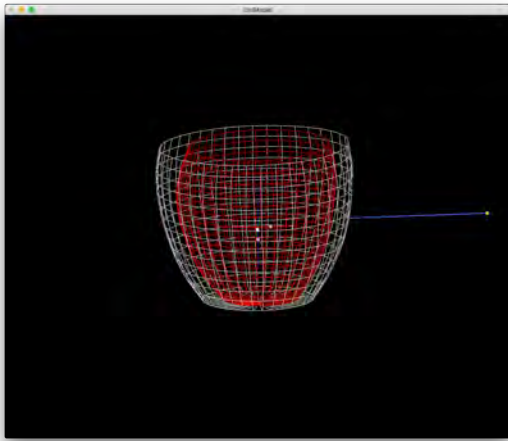
(a) Statistical Modes - 2 Active Modes
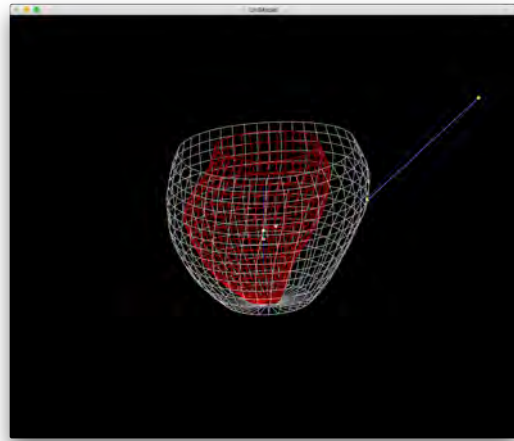
(b) Statistical Modes - 7 Active Modes
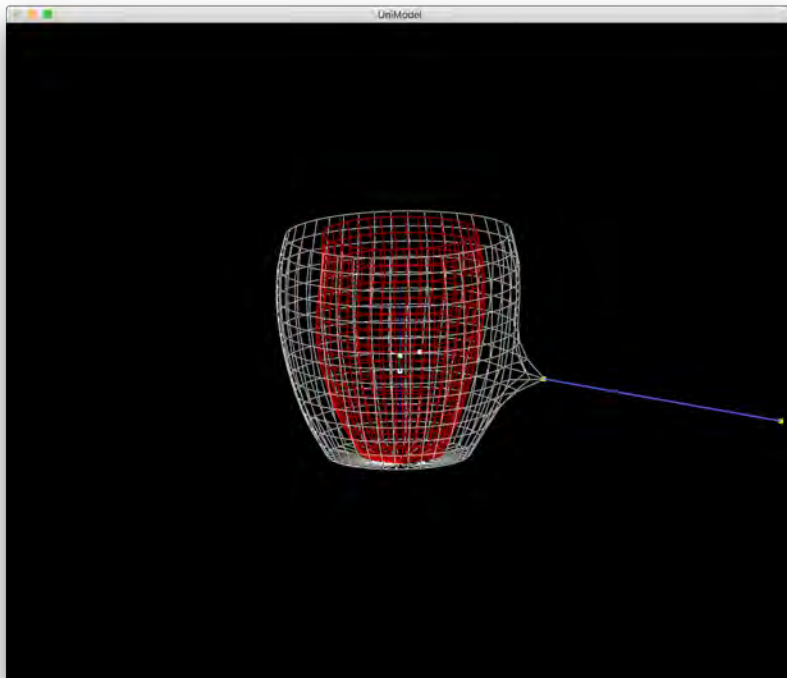
(c) FEM Skin

Figure 4.4: *Unified Model — Non-Rigid Degrees of Freedom*
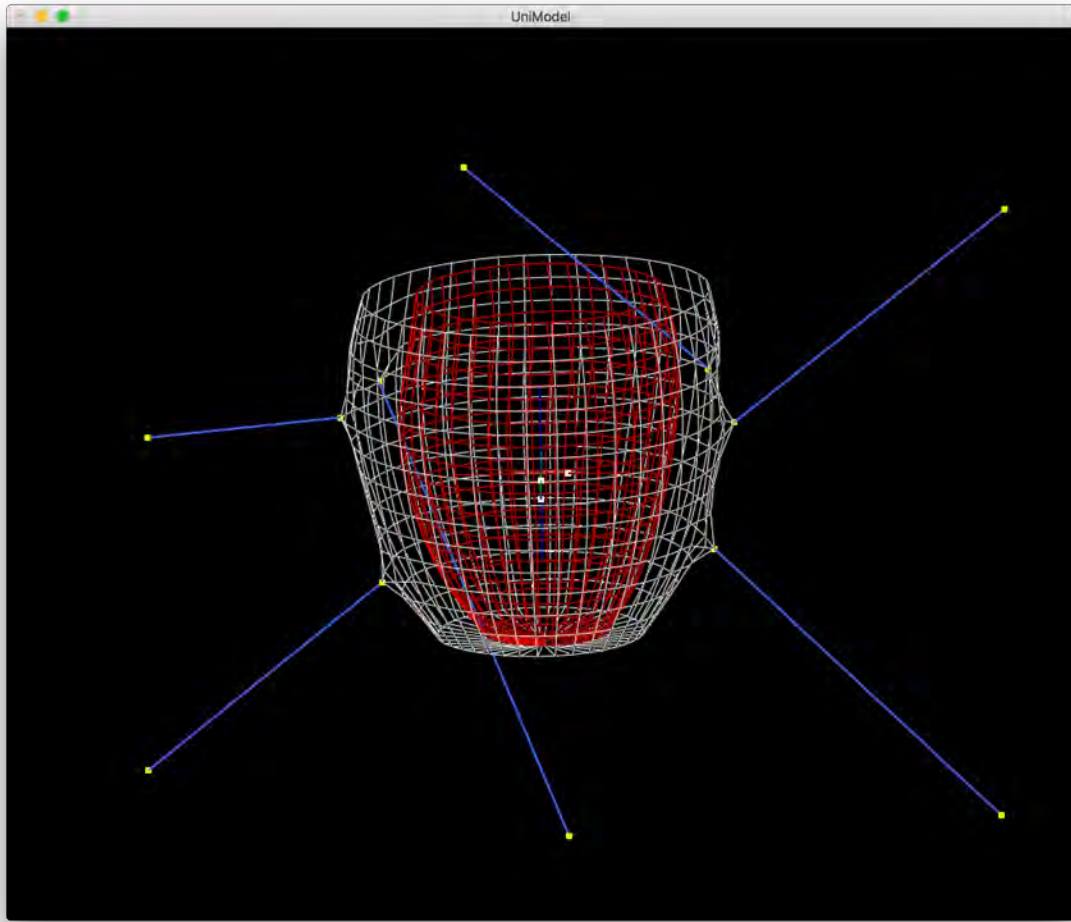
Figure 4.5: *Unified Model — Multiple spring forces acting on the model*

4.2.1 Image Forces

The model, initialized with a mean PCA reference shape, is placed in the 3D volume formed by the MR slice stack, and deformed under the influence of image forces. The image forces are designed to attract the surface of the model towards the respective myocardial boundaries (endo and epi). The first step in designing forces is to define a gradient-based potential function on an image I as

$$P = \|\nabla(G_\sigma * I)\|, \tag{4.8}$$

where the Gaussian smoothing width σ determines the range of influence of the forces. Multiple smoothing widths (Figure 4.6) are used to attain equilibrium faster. Such a potential function presents image forces that attract the surface of the model towards image intensity edges. The force distribution is the gradient of the potential function:

$$\mathbf{f} = \beta \nabla P, \tag{4.9}$$

where β controls the scale of the force. The values for β, the stiffness parameters $w_1(\mathbf{u})$ and $w_0(\mathbf{u})$, and the time step Δt are carefully selected to maintain stability. In our implementation, we have used constant values $w_1 = 4 \times 10^{-3}$ and $w_0 = 2 \times 10^{-6}$, $\beta = 30$, and $\Delta t = 1$. The image smoothing and normalization methods affect the choice of these values.

We apply two different kinds of forces to the inner and outer walls of our model to differentiate between the endocardial wall (LV blood pool-myocardium interface) and the epicardial wall (myocardium/right ventricle (RV) and myocardium/outer organs interfaces). Since the blood pool and the pericardial fat appear bright and the myocardium appears dark in cine MR images, we can make use of the information present in the direction of the image gradients. At the endocardial border, the image gradients are oriented towards the LV blood pool, whereas at the epicardial border, the image gradients are oriented away from the LV blood pool. Thus, the endocardial forces $\mathbf{f}_\mathrm{I}$ and the epicardial forces $\mathbf{f}_\mathrm{o}$ are given as

$$\mathbf{f}_\mathrm{I}(\mathbf{u}) = \begin{cases} \mathbf{f}(\mathbf{u}), & \text{if } \nabla I \cdot \mathbf{x}(\mathbf{u}) < 0 \\ 0, & \text{otherwise,} \end{cases} \tag{4.10}$$

$$\mathbf{f}_\mathrm{o}(\mathbf{u}) = \begin{cases} \mathbf{f}(\mathbf{u}), & \text{if } \nabla I \cdot \mathbf{x}(\mathbf{u}) > 0 \\ 0, & \text{otherwise,} \end{cases} \tag{4.11}$$

where the tests involve the projection of the image gradient on the position vectors $\mathbf{x}(\mathbf{u})$ of the points on the model surface whose centroid is at the origin.
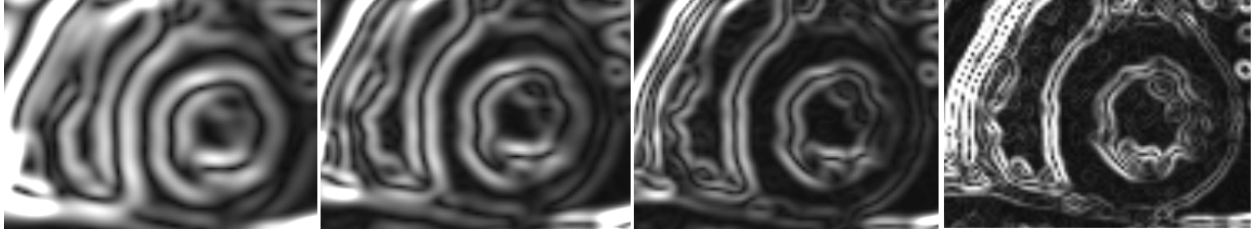
Figure 4.6: *Image potentials at multiple smoothing widths (4, 3, 2, 0 mm)*

4.3 Summary

A unified model was proposed that combines statistical and deterministic modes of deformation. A deterministic finite element interpolated skin is superimposed on a statistical reference shape prior that has been learned from a population of shapes via PCA. The rigid and non-rigid DoFs of the model were illustrated by applying external spring forces on the surface nodes. Finally, we designed separate image gradient-based external forces for the endo and epi surface nodes of our model in order that it be useful in reconstructing the LV shape from MR image data.

CHAPTER 5

Model Initialization

This chapter proposes a multi-scale detector, based on the Histogram of Oriented Gradients (HoG) [Dalal and Triggs, 2005] that produces initial estimates of LV position and scale in cardiac cine MR volumes.

5.1 Histogram of Oriented Gradients Detector

Due to the sensitivity of most optimization-based segmentation algorithms to model initialization, a good initial estimate of LV position and scale can help avoid convergence on sub-optimal local extrema. To provide good estimates, we propose a new multi-region LV detector based on the Histogram of Oriented Gradients (HoG) [Dalal and Triggs, 2005].

Multi-region HoG detectors are trained to estimate the position and scale of the LV in image slices from the base to the apex of the heart. The change in size and appearance of the LV myocardium from the base to the apex is the main reason for having 3 different detectors (of 3 different sizes) that are trained on basal, mid, and apical regions (Figure 5.1).

Given a patient's image volume as a training case, it is first divided into 3 sets of slices from base to apex. For each of the 3 slice sets, a HoG feature vector (L_2 block normalization) is computed on every slice in that set and fed to a Support Vector Machine (SVM) classifier (linear, $C = 1$) [Chang and Lin, 2011] as positive training cases. Negative training cases are obtained by computing the HoG feature vector in the background of the slice in the respective set. The details of each HoG

32
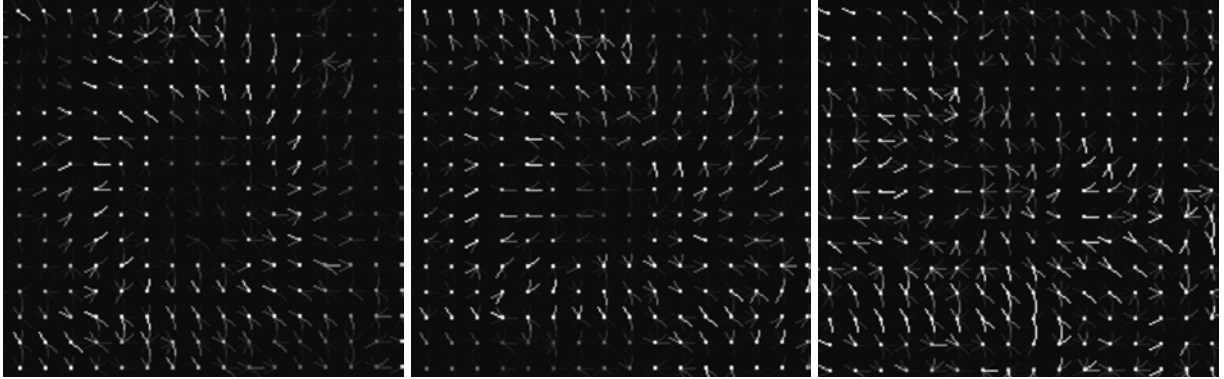
Figure 5.1: *HoGs visualized for Base (left), Mid (middle), and Apex (right) regions*

Table 5.1: *HoG Feature Descriptors*

Detector	Region	Box Size	Block Size	Cell Size	Block Stride	Bins
D_b	Base	80x80	10x10	5x5	5	12
D_m	Mid	60x60	6x6	3x3	3	12
D_a	Apex	42x42	6x6	3x3	3	12

descriptor are given in Table 5.1.

Given a new patient's MR slice stack, a standard sliding window procedure, followed by Non-Maximum Suppression (NMS), can be used to obtain a set of candidate bounding boxes (BBs) for each slice. As shown in Figure 5.2 and Figure 5.4(a), due to neighboring structures that appear very similar to the LV (bright pixels surrounded by a dark ring), false positives still exist after NMS. Therefore, we propose additional post-processing steps in Procedure 1, that exploit the 3D structure of the LV in the MR image volume to yield just one BB for every slice. For example, step 4 of this procedure chooses a combination of BBs that are better "aligned" along the LV longitudinal axis. As shown in Figure 5.3 and Figure 5.4(b), the final output of this procedure is a set of BBs, from the basal slice p to the apical slice q, which have been registered to remove any slice mis-registration due to patient motion or breathing. The volume defined by the BBs provides an estimate of the position and scale of the LV in the images.

To construct an initial LV shape from the detected BBs, we have to first guess the position of inner and outer walls of the myocardium in each slice. The inscribed circle for a BB is used as the guess
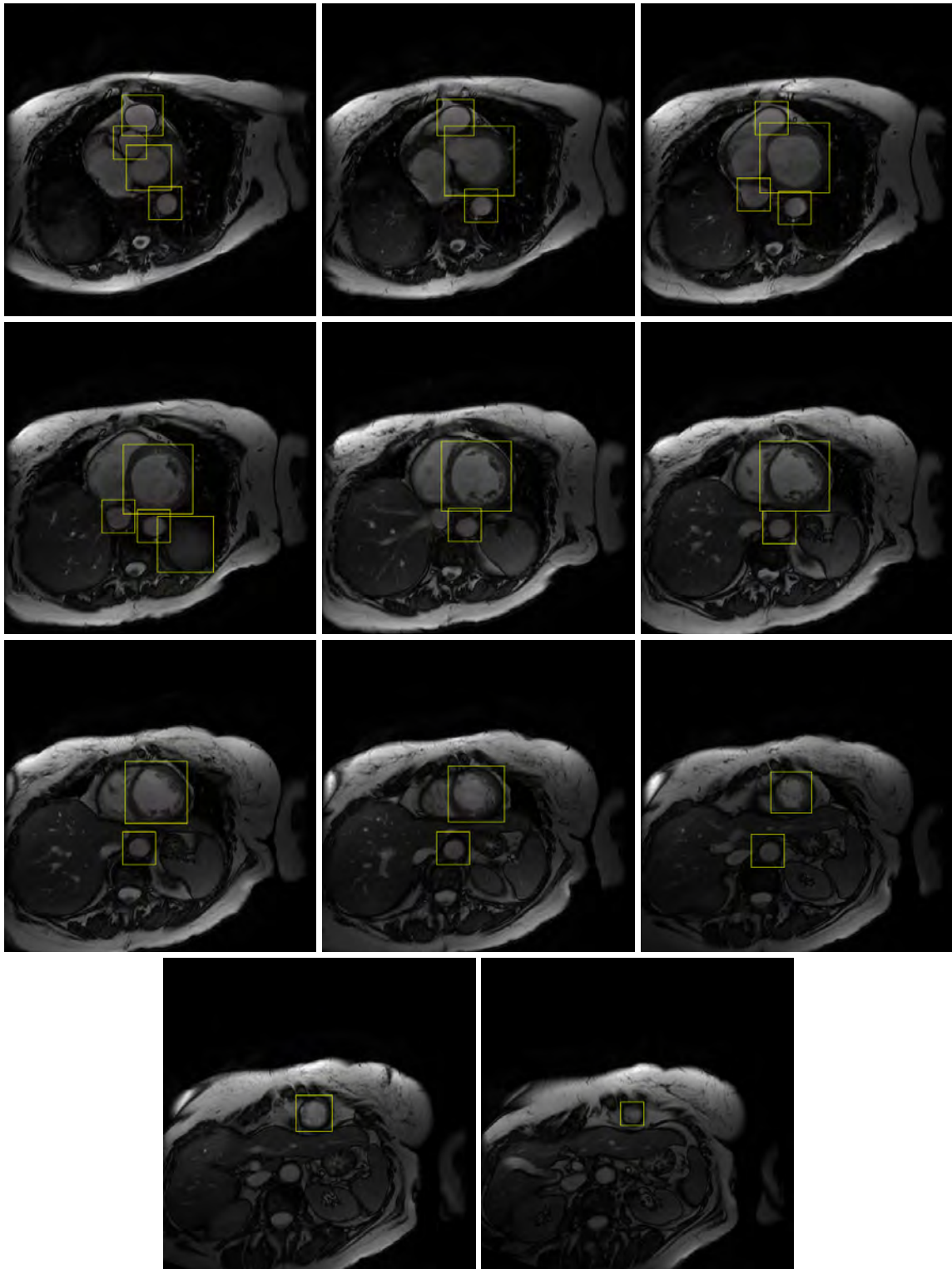
Figure 5.2: *Detector results - False positives*

Procedure 1: Left Ventricle Detector

Input: Image slices $1, 2, \ldots, N$

Output: Bounding box for each slice in range $p - q$, ($p < q$ and $1 \leq p$ and $q \leq N$), where p and q are the estimated basal and apical slices.

1 Compute SVM score masks $M_D = (m_1, m_2, \ldots, m_N)$ using sliding window detection for all detectors $D \in (D_b, D_m, D_a)$ across all N slices.

2 Compute a candidate BB set $C = (c_1, c_2, \ldots, c_N)$, where $c_i : i \in (1, 2, \ldots, N)$ is the candidate set of BBs for slice i, by performing NMS on M_{Db}, M_{Dm}, and M_{Da} score masks using an overlap threshold ϵ.

3 Set p to the smallest slice number such that candidate sets $c_1, c_2, \ldots, c_{p-1}$ are empty. Similarly, set q to the largest slice number such that candidate sets $c_{q+1}, c_{q+2}, \ldots, c_N$ are empty. If any set from c_p to c_q is empty, then fill them with averages of the BBs from neighboring slices.

4 Compute "scatter" scores for all possible combinations of BBs in the sets $c_p, c_{p+1}, \ldots, c_q$. For a given combination of BBs, $B_p, B_{p+1}, \ldots, B_q$, the scatter score is the sum of the 2 eigenvalues of the covariance matrix of the Gaussian fit to the 2D centers of the BBs. Select the combination with the least scatter score.

5 Register the slices p to q such that the centers for the final BBs line up along the longitudinal axis of the volume.

---

for the outer wall (epi). The endo must be guessed from the quantity of pixels that belong to the blood pool. To this end, we first collect the intensity values for all the pixels inside the BBs and fit a bimodal Gaussian Mixture Model (GMM) using Expectation Maximization (EM) [Dempster et al., 1977]. The GMM is subsequently used to classify the pixels as either blood pool or myocardium (Figure 5.4(c)). Finally, the ratio of endo and epi contour radii $\frac{r_{\text{endo}}}{r_{\text{epi}}} = \sqrt{\frac{A_{\text{bp}}}{A_{\text{total}}}}$ (where A_{bp} is the number of pixels labeled as blood pool, and A_{total} is the total number of pixels inside the BB) is used to produce the inner circle (Figure 5.4(d)). The stack of two concentric circles is used as the initial guess for the LV shape (Figure 5.4(e)).

Using this approximate LV shape $\mathbf{t} \in R^n$ as the target shape, and given another LV shape $\mathbf{m} \in R^n$ that needs to be scaled by a factor k, we can minimize the following cost function:

$$\arg \min_k \|\mathbf{t} - k\mathbf{m}\|^2. \tag{5.1}$$

The minimizer is $k = \mathbf{m}^T \mathbf{t} / \|\mathbf{m}\|^2$. These initial estimates of position and scale are crucial for a segmentation system to be fully automated and less susceptible to suboptimal local minima.
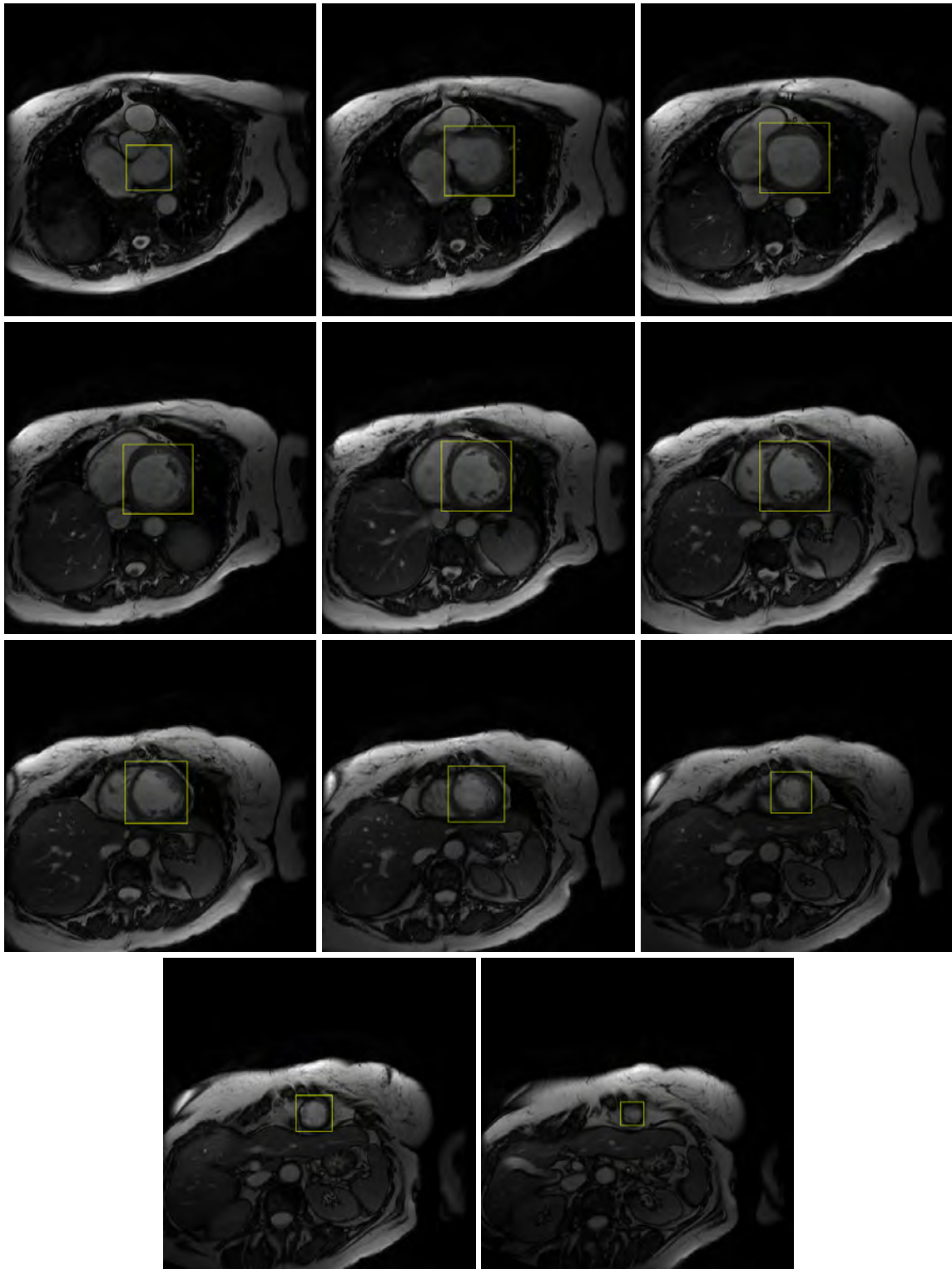
Figure 5.3: *Detector results - No False positives*
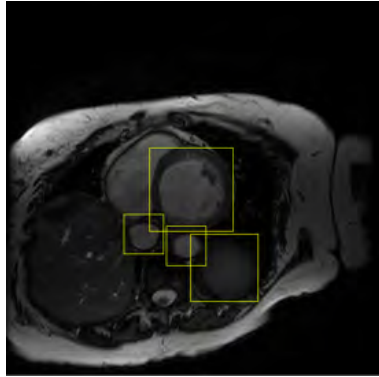
(a) Sliding window HoG Detector
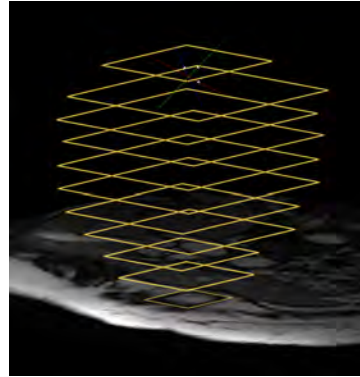
(b) Apply procedure 1 to remove false positives
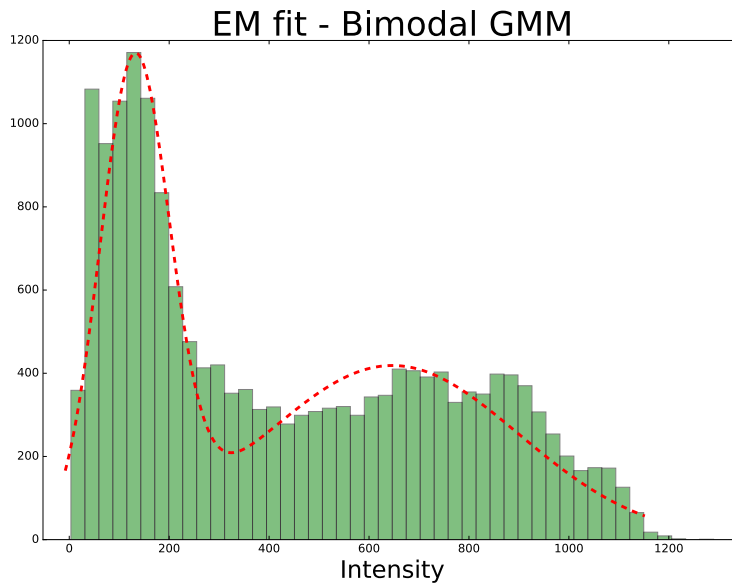
EM fit - Bimodal GMM

(c) Collect pixels inside bounding boxes and fit bi-modal Gaussian Mixture Model
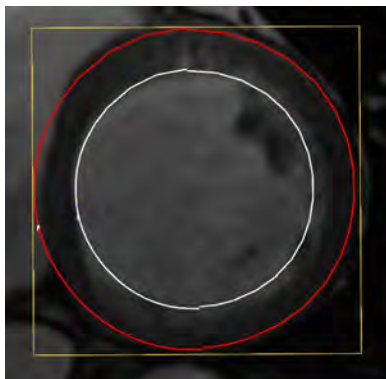
(d) Compute concentric circle from GMM labels
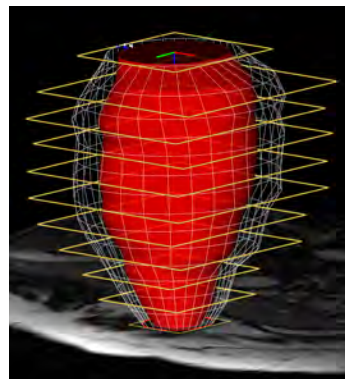
(e) Use 3D LV shape as initial guess

Figure 5.4: *Steps to obtain an initial guess with the HoG detector*

5.2 Summary

In this chapter we proposed a HoG based detector for cardiac cine MR volumes, which can be used to initialize our unified statistical/deterministic deformable model. Three distinct SVM classifiers were trained on base, mid, and apex HoG descriptors. A sliding window procedure followed by NMS produce candidate bounding boxes for the slices. A post-processing procedure was proposed to remove the false positives from each slice. Finally, a BB volume was obtained to provide estimates of position and scale of the LV in an MR volume.

CHAPTER 6

Automated Reconstruction

In this chapter we present fully automated systems for segmenting the LV in cardiac MR images based on statistical/deformable models, using both static (3D) cardiac MR image analysis and dynamic (3D + time; i.e., 4D) cardiac cine MR image sequence analysis. In the static case, a Project-Out Inverse Compositional Active Appearance Model (POIC-AAM) of 3D LV shape produces segmentations that are refined using our unified statistical/deterministic deformable model. In the dynamic case, a 3D statistical deformable model (SIC-AAM) is used in a novel smoothness and volume constrained variational optimization scheme to provide temporally consistent results.

6.1   Automated Static Reconstruction

The key algorithms/components employed by our system and our main motivations for including them are as follows: A Project-Out Inverse Compositional (POIC) Active Appearance Model (AAM) [Matthews and Baker, 2004] underlies a model-based segmentation procedure and a unified statistical/deterministic deformable model refines the performance of the AAM. Due to the sensitivity of most optimization-based segmentation algorithms to model initialization, a good initial estimate of LV position and scale can help avoid convergence on sub-optimal local extrema. To provide good estimates, we use the Histogram of Oriented Gradients (HoG) [Dalal and Triggs, 2005] based LV detector proposed in the previous chapter.

6.1.1 Project Out Inverse Compositional 3D AAM

AAMs learn linear models of shape and appearance variation. They are classified as Combined AAMs [Cootes et al., 1998] or Independent AAMs [Matthews and Baker, 2004] based, respectively, on whether or not they have a common set of parameters that control both shape and appearance. The POIC is an Independent AAM, which we use to train linear models of shape and appearance on aligned (by Ordinary Procrustes Analysis as described in Appendix B.2) 3D shapes of the LV, and shape-normalized appearance sampled cardiac MR image information within the 3D LV shape. PCA on such training data yields the following generative models of shape $\mathbf{s}$ and appearance $\mathbf{g}$:

$$\mathbf{s} = \bar{\mathbf{s}} + \mathbf{P}_s\mathbf{b}_s, \tag{6.1}$$

$$\mathbf{g} = \bar{\mathbf{g}} + \mathbf{P}_g\mathbf{b}_g. \tag{6.2}$$

The appearance model is trained by warping the training shapes to the base/mean shape $\bar{\mathbf{s}}$ using piecewise affine warping on the tetrahedral 3D tessellation.

Given a new set of images $I(\mathbf{x})$, the fitting process involves a Gauss-Newton minimization of the following non-linear cost function (error image), with respect to the parameters $\mathbf{b}_s$ and $\mathbf{b}_g$ (with reduced dimensions):

$$\|\bar{\mathbf{g}} + \mathbf{P}_g\mathbf{b}_g - I(\mathbf{W}(\mathbf{x}; \mathbf{b}_s))\|^2. \tag{6.3}$$

The ability to precompute the steepest descent images and the Hessian, makes POIC an efficient algorithm.

6.1.2 Unified Statistical/Deformable Model

The above AAM provides an estimate of the LV shape by linearly combining deformation modes with coefficients $\mathbf{b}_s$. Since this is a limited set of orthonormal modes, the actual solution might not be present in the subspace spanned by them. It can be challenging to obtain accurate segmentation

estimates with just an AAM. Therefore, to further refine the segmentation, we employ a variant of our Unified Statistical/Deterministic deformable model that was developed in Chapter 4.

The model's geometry consists of a 3D deformable "skin" superimposed on a reference PCA shape (as in the POIC (6.1)). The skin is composed of spring-edged cuboids (12 springs) that impose elastic/structural constraints on neighboring nodes present on the endo and epi surfaces. Applying Lagrangian dynamics, the model is then made dynamic with respect to its parameters, thus yielding the following (massless) equations of motion:

$$\mathbf{C}\dot{\mathbf{q}} + \mathbf{K}\mathbf{q} = \mathbf{f}_q, \tag{6.4}$$

where $\dot{\mathbf{q}}$ is the time derivative of the vector $\mathbf{q}$ of degrees of freedom (DOFs) of the model, which includes the pose parameters, PCA parameters, and displacement parameters of the skin from the PCA reference shape, $\mathbf{C}\dot{\mathbf{q}}$ are damping forces, $\mathbf{K}\mathbf{q}$ are internal elastic forces, and $\mathbf{f}_q$ are image gradient-based external forces applied to the model. The stiffness matrix $\mathbf{K}$ can be assembled "element-wise" by using the spring element matrix $\begin{bmatrix} k & -k \\ -k & k \end{bmatrix}$, where k is the spring stiffness, which can be tuned to control the elasticity of the skin that deforms away from the reference shape under the influence of external forces. Finally, we initialize the PCA parameters in $\mathbf{q}$ to the values obtained from the POIC AAM fitting algorithm and explicitly time-integrate the above equations of motion under the influence of the image forces that pull the skin towards nearby image intensity edges.

6.2 Automated Dynamic Reconstruction

A method to reconstruct the 3D LV shape across time must consider the temporal properties of the image features and myocardium biomechanics. Due to the presence of papillary muscles and trabeculae, the inner wall (endo) of the myocardium becomes hard to localize in frames close to the ES frame. Even human experts must look at the nearby image frames for cues on the myocardium

41

boundaries. Hence, temporal reconstruction methods must emulate this action by incorporating it as prior knowledge. This is indirectly achieved by imposing temporal smoothness constraints on the parameters of the Simultaneous Inverse Compositional AAM (SIC-AAM). A sequential application of SIC-AAM for the successive frames will not achieve temporal consistency, because the model would drift and accumulate error in frames close to the ES frame. Moreover, since this is not a real-time sequential tracking application, recursive estimation algorithms, neither particle filters [Isard and Blake, 1998] nor kalman filters [Terzopoulos and Szeliski, 1993; Metaxas and Terzopoulos, 1993] are required. All the frames can be considered simultaneously to achieve temporal consistency.

It is important to note that the temporal constraints are solely used to achieve temporal consistency and not to recover the displacement field, as is done in tagged MRI. These constraints are there to help achieve better myocardium segmentation across time. To this end, the following unified variational method is proposed to solve for the shape and appearance parameters of the SIC-AAM as functions of time.

6.2.1   Unified Variational Method

This section describes the terms used in the unified variational method for 4D LV shape reconstruction in cardiac MRI. The method involves the minimization of the following energy (continuous form):

$$E = \int_{\mathcal{T}} \alpha E_{\text{data}} + E_{\text{smooth}} + \gamma E_{\text{vol}} \, dt \qquad (6.5)$$

where $\mathcal{T}$ represents a periodic cardiac cycle in an MR volume. Each of the terms inside the integral, and the motivation for having it, is described in the following three sections.

6.2.1.1 E_{data} — Statistical Deformable Models

This is an energy term that is dependent on the visual data (voxel intensities) in the cardiac MR volume. Due to the unique shape and appearance characteristics of the cardiac LV in MR images, a data term that encodes prior knowledge of its shape and appearance would be a good choice. Hence, we choose Active Appearance Models (AAMs) which are statistical deformable models that build generative linear models of shape and appearance from a training set using Principal Components Analysis (PCA) (see Appendix B.3 for more details). Many variants of AAMs have been proposed, and they can differ in the way they synthesize new model instances (Combined vs Independent), learn to fit to image data (Multivariate Regression vs Gradient Descent), or update parameters (Additive vs Compositional, Inverse vs Forward). We choose the Simultaneous Inverse Compositional (SIC) algorithm due to its better accuracy. More details of the different AAMs can be found in [Matthews and Baker, 2004].

SIC is an Independent AAM that trains linear models of shape and appearance on registered/aligned (OPA; see Appendix B.2 for more details) 3D shapes of LV, and shape-normalized appearance sampled on cardiac MR visual information within the 3D LV shape. PCA on such training data yields the following generative models of shape $\mathbf{s} \in \mathcal{R}^{3p}$ (with p 3D vertices) and appearance $\mathbf{a} \in \mathcal{R}^{v}$ (with v voxels):

$$\mathbf{s} = \bar{\mathbf{s}} + \mathbf{P}_s \mathbf{q}_s, \tag{6.6}$$

$$\mathbf{a} = \bar{\mathbf{a}} + \mathbf{P}_a \mathbf{q}_a, \tag{6.7}$$

where $\bar{\mathbf{s}} \in \mathcal{R}^{3p}$ is the mean shape, $\mathbf{P}_s \in \mathcal{R}^{3p \times n}$ are the modes of shape variation, $\mathbf{q}_s \in \mathcal{R}^n$ are the n shape coefficients, $\bar{\mathbf{a}} \in \mathcal{R}^v$ is the mean appearance, $\mathbf{P}_a \in \mathcal{R}^{v \times m}$ are the modes of appearance variation, and $\mathbf{q}_a \in \mathcal{R}^m$ are the m appearance coefficients. The shape mean $\bar{\mathbf{s}}$ and appearance mean $\bar{\mathbf{a}}$ are shown in Figure 6.1.

The appearance model is trained by warping the training shapes to the base/mean shape $\bar{\mathbf{s}}$ using piecewise affine warps on the tetrahedral 3D tessellation (see Appendix C for more details). The above two models can be used to synthesize new shapes and appearances by linearly combining
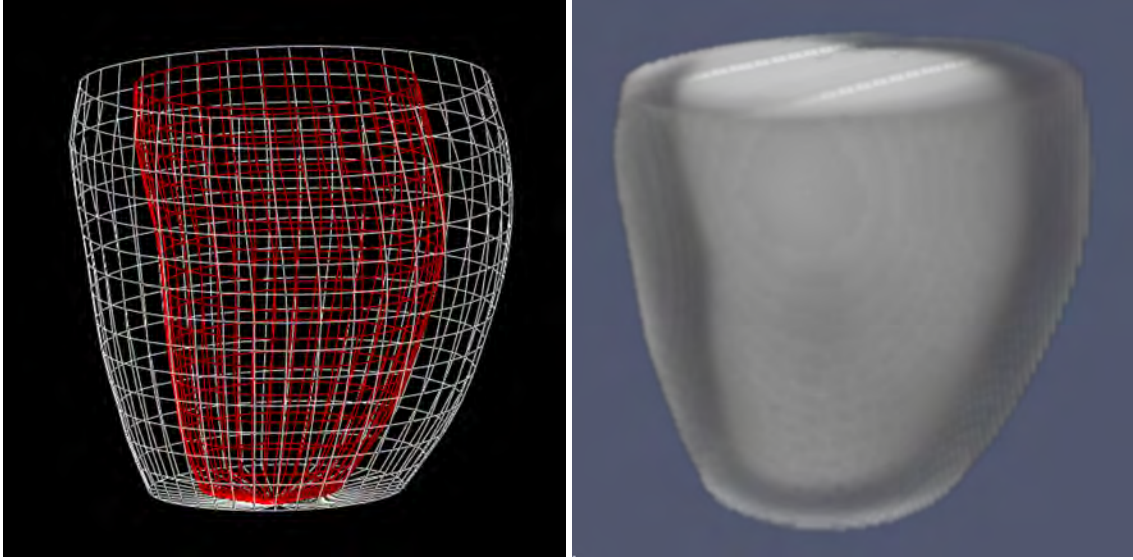
Figure 6.1: *Shape and Appearance Means*

the respective modes of variation.

The SIC-AAM fitting algorithm involves a gradient descent procedure to synthesize the shape model and appearance model instances, so as to minimize the L_2 norm of the error between the appearance model and the test image voxels $I(\mathbf{x})$ (as defined by the current shape model),

$$E_{\text{data}} = \sum_{\mathbf{x}} [\bar{\mathbf{a}}(\mathbf{x}) + \mathbf{P}_a(\mathbf{x})\mathbf{q}_a - I(\mathbf{N}(\mathbf{x}; \mathbf{q}_c, \mathbf{q}_\theta, \mathbf{q}_s))]^2, \tag{6.8}$$

where the summation over $\mathbf{x} \in \mathcal{R}^3$ is for all the N voxels inside the $\bar{\mathbf{s}}$ mean shape, $\mathbf{q}_c, \mathbf{q}_\theta$ are the rigid transformation parameters (3 translation + 3 rotation), and $\mathbf{N} : \mathcal{R}^3 \rightarrow \mathcal{R}^3$, defined as $\mathbf{N}(\mathbf{x}; \mathbf{q}_c, \mathbf{q}_\theta, \mathbf{q}_s) = \mathbf{R}(\mathbf{q}_\theta)\mathbf{W}(\mathbf{x}; \mathbf{q}_s) + \mathbf{T}(\mathbf{q}_c)$, is a coordinate transform that applies a rigid transform on the piecewise affine warp $\mathbf{W} : \mathcal{R}^3 \rightarrow \mathcal{R}^3$. Henceforth, we combine the unknowns into a single vector $\mathbf{q} = [\mathbf{q}_c^T \mathbf{q}_\theta^T \mathbf{q}_s^T \mathbf{q}_a^T]^T \in \mathcal{R}^{n+m+6}$.

6.2.1.2 E_{smooth} — Smoothness

Deterministic deformable model formulations ([Terzopoulos et al., 1988; Terzopoulos and Metaxas, 1991]) have used first-order and second-order spatial derivative terms to regularize and obtain smooth solutions. Since the cardiac LV shape's motion is smooth across the cardiac cycle [Glass et al., 1991], we penalize high values for first and second order temporal derivatives:

$$E_{\text{smooth}} = \mathbf{q}_t^T \mathbf{B}_1 \mathbf{q}_t + \mathbf{q}_{tt}^T \mathbf{B}_2 \mathbf{q}_{tt}, \tag{6.9}$$

where $\mathbf{q}_t = \frac{\partial \mathbf{q}(t)}{\partial t}$, $\mathbf{q}_{tt} = \frac{\partial^2 \mathbf{q}(t)}{\partial t^2}$, and $\mathbf{B}_1 \in \mathcal{R}^{n+m+6}$ and $\mathbf{B}_2 \in \mathcal{R}^{n+m+6}$ are diagonal weighting matrices that can be used to control the penalty. It can be shown that the temporal smoothness of $\mathbf{q}$ implies temporal smoothness of shape $\mathbf{s}$ and appearance $\mathbf{a}$ because of the linearity of the models in (6.6) and (6.7).

6.2.1.3 E_{vol} — Incompressible Motion

Due to its composition (mainly water) and the less than 4% change in total volume in a single cardiac cycle due to perfusion [Song and Leahy, 1991; Bistoquet et al., 2007; Wang and Amini, 2012], the myocardium is considered to be nearly incompressible. We impose the following volume change penalty on the tessellated geometry of the myocardium:

$$E_{\text{vol}} = \left[\sum_i V_i(\mathbf{q}) - V_{tar} \right]^2, \tag{6.10}$$

where the summation is over the volume of all the geometric elements (tetrahedrons in our case) $V_i(\mathbf{q})$ composing the myocardium, and V_{tar} is the constant target volume to be achieved.

6.2.2 Optimization

Since we must solve for shape and appearance parameters for each time frame across a periodic cardiac cycle, we make the vector of unknowns $\mathbf{q}$ a function of time and rewrite the energy functional (6.5) in the following way:

$$E(\mathbf{q}(t)) = \int_{\mathcal{T}} \alpha E_{\text{data}}(\mathbf{q}(t)) + E_{\text{smooth}}(\mathbf{q}(t)) + \gamma E_{\text{vol}}(\mathbf{q}(t))dt. \tag{6.11}$$

Finding the optimal value of the function $\mathbf{q}(t)$ that minimizes the total energy is a nonlinear least squares problem because, although the energy is linear with respect to appearance parameters $\mathbf{q}_a$, it is non-linear with respect to the shape parameters $\mathbf{q}_c$, $\mathbf{q}_\theta$, and $\mathbf{q}_s$. The calculus of variations is used to obtain the fourth-order Euler-Lagrange equations that yield necessary conditions that have to be satisfied by the extremal $\mathbf{q}(t)$. These equations are discretized and stepped in an iterative Gauss-Newton scheme [Nocedal and Wright, 2006] to minimize the energy. We provide more formulation details in Appendix C.

6.3 Summary

This chapter proposed fully automated techniques for static and dynamic LV reconstruction. Using the initial model estimates from the HoG based LV detector described in Chapter 5, a combination of POIC-AAM and the Unified Deterministic/Statistical Model was used to reconstruct and refine the LV shape given a single phase volume (static). Next, a variational method was proposed to reconstruct the LV across the entire cardiac cycle (dynamic) with smoothness and volume constraints.

CHAPTER 7

Experiments and Results

This chapter presents the results of our experiments for the Unified Deterministic/Statistical Model, the 3D LV detector, the automated static LV reconstruction method (POIC-AAM + Unified Model), and the automated dynamic LV reconstruction that were proposed in Chapters 4, 5, 6.1, and 6.2, respectively.

The UniMR interface (see Appendix A for more details) has been used to run the experiments and to compute the performance metrics for all the models and methods. The two cardiac cine MR image databases that were used for measuring performance are (i) the MICCAI LV Segmentation challenge database [Radau et al., 2009] and (ii) the Cardiac Atlas Project (CAP) consensus database [Suinesiaputra et al., 2014]. In the remainder of the chapter, we will refer to these databases as MICCAI and CAP for brevity.

7.1 Unified Deterministic/Statistical Deformable Model

We validated the segmentation ability of the Unified Deterministic/Statistical Deformable Model (proposed in Chapter 4) using leave-one-out training on end-diastolic (ED) images of the 45 cases in the MICCAI database. The leave-one-out validation was fully automated and the mean reference model was initialized in the volume such that the centroid of the model coincided with the center of the mid-slice. Initially, the model was subject only to translational forces designed using optical flow potentials across phases. Such forces approximately localize the myocardium and help in

47

MPD	ED-Epi	ED-Endo		Dice	ED-Epi	ED-Endo
MEAN	3.6	3.7		MEAN	0.88	0.84
STD	0.52	0.62		STD	0.02	0.04
MAX	4.89	4.68		MAX	0.93	0.91
MIN	2.11	2.16		MIN	0.81	0.75

Table 7.1: *Mean Perpendicular Distance (mm) and Dice coefficient (45 cases)*

moving the initial mean reference closer to the actual solution. Subsequently, all the parameters (rigid and non-rigid) were stepped forward in time. The PCA parameters $\mathbf{q}_s$ are restricted within $+2$ and -2 standard deviations (which can be obtained from the corresponding eigenvalues) from the mean in order to prevent unlikely shapes. The constants that control the deformation of the skin are set to $w_1 = 4 \times 10^{-3}$ and $w_0 = 2 \times 10^{-6}$. The model is stepped forward across multiple Gaussian smoothing widths (4, 3, 2, 0 mm), finally converging at the myocardial boundaries.

Due to the ambiguous gradient information at the myocardium interface with lungs and other organs, the epicardial boundary is harder to localize. By virtue of the model having the MICCAI database trained statistical reference shape, the papillary muscles were included in the blood pool (Figure 7.1). We used the Mean Perpendicular Distance (MPD) and Dice coefficients (Table 7.1) to compare the positioning errors of automated contours with respect to the expert-delineated contours. The average Dice coefficient and the average MPD for the ED segmentation are 0.86 and 3.65 mm respectively, and these are close to the results presented in [Jolly, 2009], [Huang et al., 2009], and [Lu et al., 2009]. The automated contours for the mid-slices are more accurate than those for the slices towards the apex due to partial volume effects.

7.2 HoG Detector

We now discuss the results for the HoG detector proposed in Chapter 5. The C++ SVM implementation in LibSVM [Chang and Lin, 2011] was used for the detector's HoG feature training, and the MLPACK [Curtin et al., 2013] implementation of GMM was used for labeling the BB pixels.
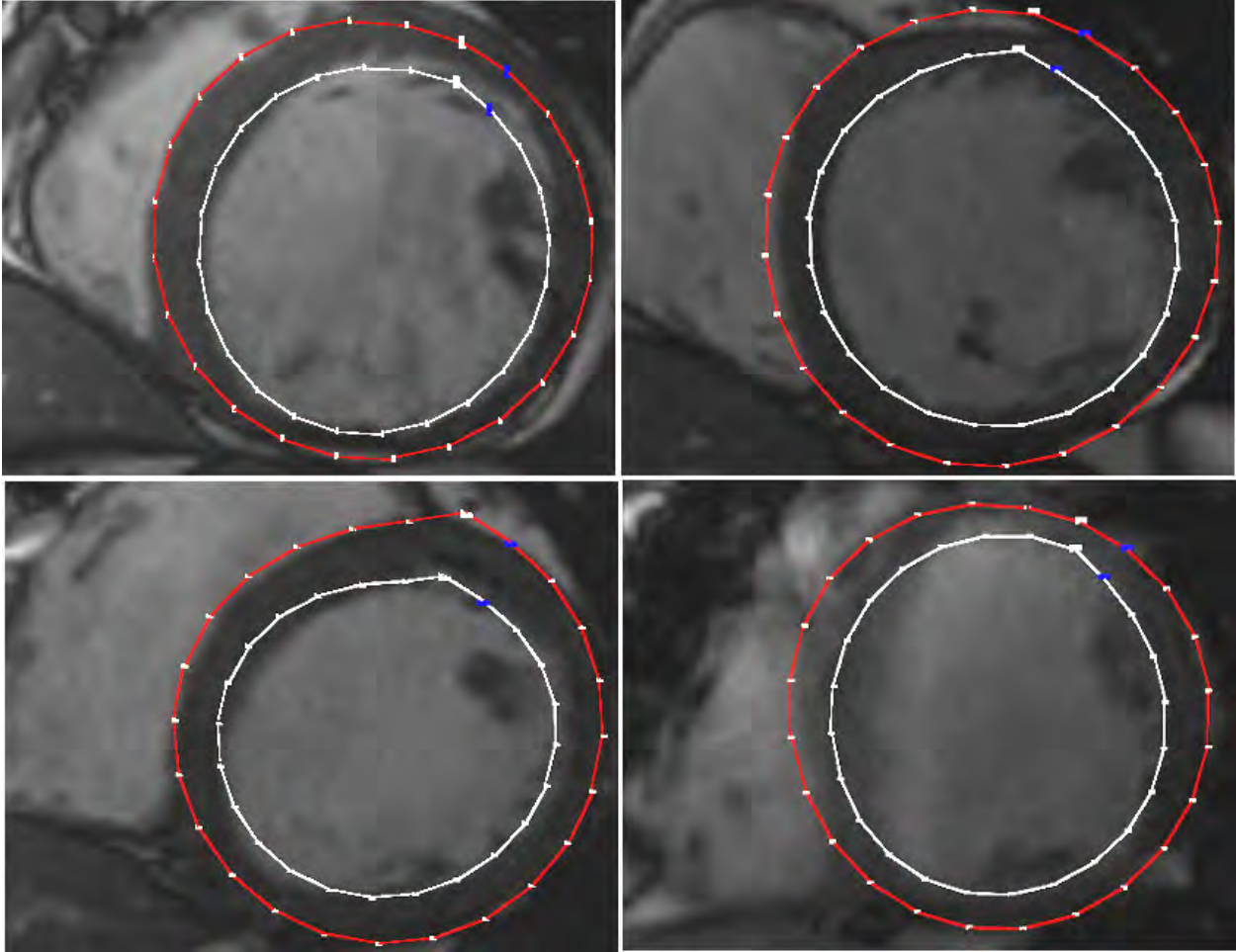
Figure 7.1: *Examples of automated contour segmentation for four cases*

The overlap threshold for the NMS step was set to $\epsilon = 0.5$. The performance of the detector was evaluated on the MICCAI (15 training and 30 testing cases) and CAP (100 training and 95 testing cases) databases separately.

For each of these testing sets, the precision and recall between the set of ground truth BB_{gt} and the estimated BB_{est} bounding boxes are computed (for the whole volume) as below. The f-score (the harmonic mean of precision and recall) is also computed for convenience.

Table 7.2: *Detector performance on MICCAI and CAP databases*

	MICCAI			CAP		
	Precision	Recall	F-Score	Precision	Recall	F-Score
Avg	0.85	0.94	**0.89**	0.92	0.91	**0.91**
STD	0.07	0.04	**0.04**	0.09	0.01	**0.07**

$$Precision(P) = \frac{BB_{\text{gt}} \cap BB_{\text{est}}}{BB_{\text{est}}} \tag{7.1}$$

$$Recall(R) = \frac{BB_{\text{gt}} \cap BB_{\text{est}}}{BB_{\text{gt}}} \tag{7.2}$$

$$F - Score(F) = \frac{2 * P * R}{P + R} \tag{7.3}$$

The average and standard deviation for the 3 metrics are reported for both databases in Table 7.2. Additionally, these metrics are listed individually for each of the 30 MICCAI cases in Table 7.3. The f-score averages for the MICCAI and CAP databases are 0.89 and 0.91 respectively. This shows that the detection system, as described in Figure 5.4 and Procedure 1, is able to obtain good estimates of position and scale of the LV in the 3D MR volume.

7.3 Automated Static Reconstruction

This section describes the results obtained while statically reconstructing the LV using the deformable-model-based (POIC-AAM + Unified Model) approach proposed in Section 6.1. The MICCAI database consisting of short-axis (SA) image slices for 15 training and 30 test cases was used to train and test the deformable-model-based segmentation components. Each of the 45 cases contains ground truth epicardial (epi) and endocardial (endo) contours for the end-diastolic (ED) phase, and just the endo contours for the end-systolic (ES) phase. We evaluate our system on the ED phase only. The HoG detector proposed in Chapter 5 is used for initialization. We used 8 shape modes for the POIC-AAM and a spring constant $k = 0.5$ for the unified model. Except

Table 7.3: *Precision (P), Recall (R), and F-Score (F) for the LV detector*

ID	Case	Detector		
		P	R	F
1	SC-HF-I-05	0.82	0.99	0.90
2	SC-HF-I-06	0.91	0.96	0.93
3	SC-HF-I-07	0.93	0.95	0.94
4	SC-HF-I-08	0.89	0.98	0.93
5	SC-HF-NI-07	0.87	0.94	0.91
6	SC-HF-NI-11	0.93	0.95	0.94
7	SC-HF-NI-31	0.88	0.98	0.93
8	SC-HF-NI-33	0.93	0.91	0.92
9	SC-HYP-06	0.89	0.95	0.92
10	SC-HYP-07	0.90	0.90	0.90
11	SC-HYP-08	0.83	0.88	0.86
12	SC-HYP-37	0.89	0.84	0.86
13	SC-N-05	0.88	0.89	0.89
14	SC-N-06	0.93	0.85	0.89
15	SC-N-07	0.77	0.98	0.86
16	SC-HF-I-09	0.87	0.98	0.93
17	SC-HF-I-10	0.85	0.98	0.91
18	SC-HF-I-11	0.84	0.97	0.90
19	SC-HF-I-12	0.85	0.94	0.90
20	SC-HF-NI-12	0.87	0.98	0.92
21	SC-HF-NI-13	0.75	0.93	0.83
22	SC-HF-NI-14	0.82	0.97	0.89
23	SC-HF-NI-15	0.62	0.98	0.76
24	SC-HYP-09	0.78	1.00	0.88
25	SC-HYP-10	0.94	0.91	0.93
26	SC-HYP-11	0.89	0.91	0.90
27	SC-HYP-12	0.83	0.93	0.88
28	SC-N-09	0.87	0.97	0.92
29	SC-N-10	0.74	0.93	0.82
30	SC-N-11	0.75	0.98	0.85
	Average	0.85	0.94	0.89

for assuming that the ED phase is known and the slices $1, 2, \ldots, N$ go from base to the apex, all components of the system are fully automated.

For each of these 30 cases, the precision and recall between the set of ground truth M_{gt} and the estimated M_{est} masks (for the whole volume) are reported in Table 7.4. These are computed for the overlap of areas for EPI (pixels inside the epi contour), ENDO (pixels inside the endo contour), and MYO (pixels in-between epi and endo contours), individually. The f-score is also reported for convenience. Precision, recall and f-score are evaluated as follows:

$$Precision(P) = \frac{M_{\text{gt}} \cap M_{\text{est}}}{M_{\text{est}}} \tag{7.4}$$

$$Recall(R) = \frac{M_{\text{gt}} \cap M_{\text{est}}}{M_{\text{gt}}} \tag{7.5}$$

$$F - Score(F) = \frac{2 * P * R}{P + R}. \tag{7.6}$$

The average f-scores for contour overlap are 0.80 for endo, 0.82 for epi, and 0.46 for myocardium. The potential refinement of the unified deformable model can be seen in Figure 7.2 for the case HC-HF-I-05, where POIC-AAM only estimates are compared with those that have been further refined. We also show in Figure 7.3, the myocardial overlap masks for case SC-HF-I-09 (with f-score - 0.91 for endo, 0.93 for epi, 0.64 for myo). These images have been color-coded to show the True Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN) myocardial labels obtained by overlaying the ground truth contours with the estimated contours. The myocardial f-score is a good indicator of closeness between the estimated LV shape and the ground truth LV shape. It can be very challenging to achieve a high MYO f-score due to the high degree of accuracy expected of a segmentation system on the testing cases.

To further investigate the performance, we measure the region-wise (base, mid, apex) contour overlap with respect to the ground truth contours (Recall). The average Recall for the base, mid and apex regions are as follows: endo (0.91, 0.89, 0.77), epi (0.90, 0.91, 0.81), and myo (0.50, 0.56, 0.43). These scores show that our automated system performs better at the base and mid LV

Table 7.4: *Precision (P), Recall (R), and F-Score (F) for the LV contours*

ID	Case	ENDO			EPI			MYO		
		P	R	F	P	R	F	P	R	F
1	SC-HF-I-05	0.81	0.88	0.84	0.81	0.89	0.85	0.55	0.63	0.59
2	SC-HF-I-06	0.98	0.86	0.92	0.99	0.83	0.90	0.72	0.55	0.62
3	SC-HF-I-07	0.96	0.69	0.80	0.98	0.64	0.77	0.30	0.17	0.22
4	SC-HF-I-08	0.97	0.84	0.90	0.96	0.89	0.92	0.62	0.64	0.63
5	SC-HF-NI-07	0.97	0.82	0.89	0.95	0.89	0.92	0.56	0.62	0.59
6	SC-HF-NI-11	0.77	0.84	0.80	0.77	0.88	0.83	0.52	0.65	0.58
7	SC-HF-NI-31	0.73	0.93	0.82	0.70	0.98	0.82	0.32	0.53	0.40
8	SC-HF-NI-33	0.84	0.80	0.82	0.85	0.84	0.85	0.52	0.54	0.53
9	SC-HYP-06	0.79	0.82	0.81	0.81	0.78	0.80	0.53	0.46	0.49
10	SC-HYP-07	0.60	0.97	0.74	0.71	0.97	0.82	0.35	0.38	0.37
11	SC-HYP-08	0.59	0.86	0.70	0.71	0.77	0.74	0.54	0.41	0.47
12	SC-HYP-37	0.50	0.90	0.64	0.64	0.79	0.71	0.47	0.37	0.42
13	SC-N-05	0.89	0.83	0.86	0.96	0.78	0.86	0.62	0.41	0.49
14	SC-N-06	0.70	0.89	0.78	0.69	0.96	0.80	0.39	0.61	0.48
15	SC-N-07	0.66	0.99	0.79	0.64	0.99	0.78	0.19	0.31	0.24
16	SC-HF-I-09	0.96	0.87	0.91	0.93	0.93	0.93	0.58	0.72	0.64
17	SC-HF-I-10	0.91	0.89	0.90	0.88	0.92	0.90	0.54	0.65	0.59
18	SC-HF-I-11	0.93	0.90	0.91	0.85	0.95	0.90	0.50	0.74	0.59
19	SC-HF-I-12	0.95	0.84	0.89	0.97	0.83	0.89	0.67	0.53	0.59
20	SC-HF-NI-12	0.92	0.90	0.91	0.88	0.96	0.92	0.52	0.69	0.59
21	SC-HF-NI-13	0.69	0.86	0.77	0.65	0.94	0.77	0.19	0.38	0.25
22	SC-HF-NI-14	0.52	0.85	0.65	0.57	0.85	0.68	0.36	0.48	0.41
23	SC-HF-NI-15	0.39	0.79	0.52	0.41	0.85	0.55	0.17	0.37	0.23
24	SC-HYP-09	0.58	0.95	0.72	0.52	0.96	0.68	0.21	0.47	0.29
25	SC-HYP-10	1.00	0.70	0.82	1.00	0.71	0.83	0.36	0.26	0.30
26	SC-HYP-11	0.42	0.95	0.59	0.50	0.97	0.66	0.25	0.40	0.31
27	SC-HYP-12	1.00	0.72	0.84	0.99	0.78	0.87	0.46	0.42	0.44
28	SC-N-09	0.86	0.94	0.90	0.84	0.95	0.89	0.59	0.73	0.65
29	SC-N-10	0.61	0.92	0.74	0.81	0.91	0.86	0.52	0.41	0.45
30	SC-N-11	0.70	0.97	0.81	0.80	0.95	0.87	0.46	0.44	0.45
	Average	0.77	0.87	0.80	0.79	0.88	0.82	0.45	0.50	0.46
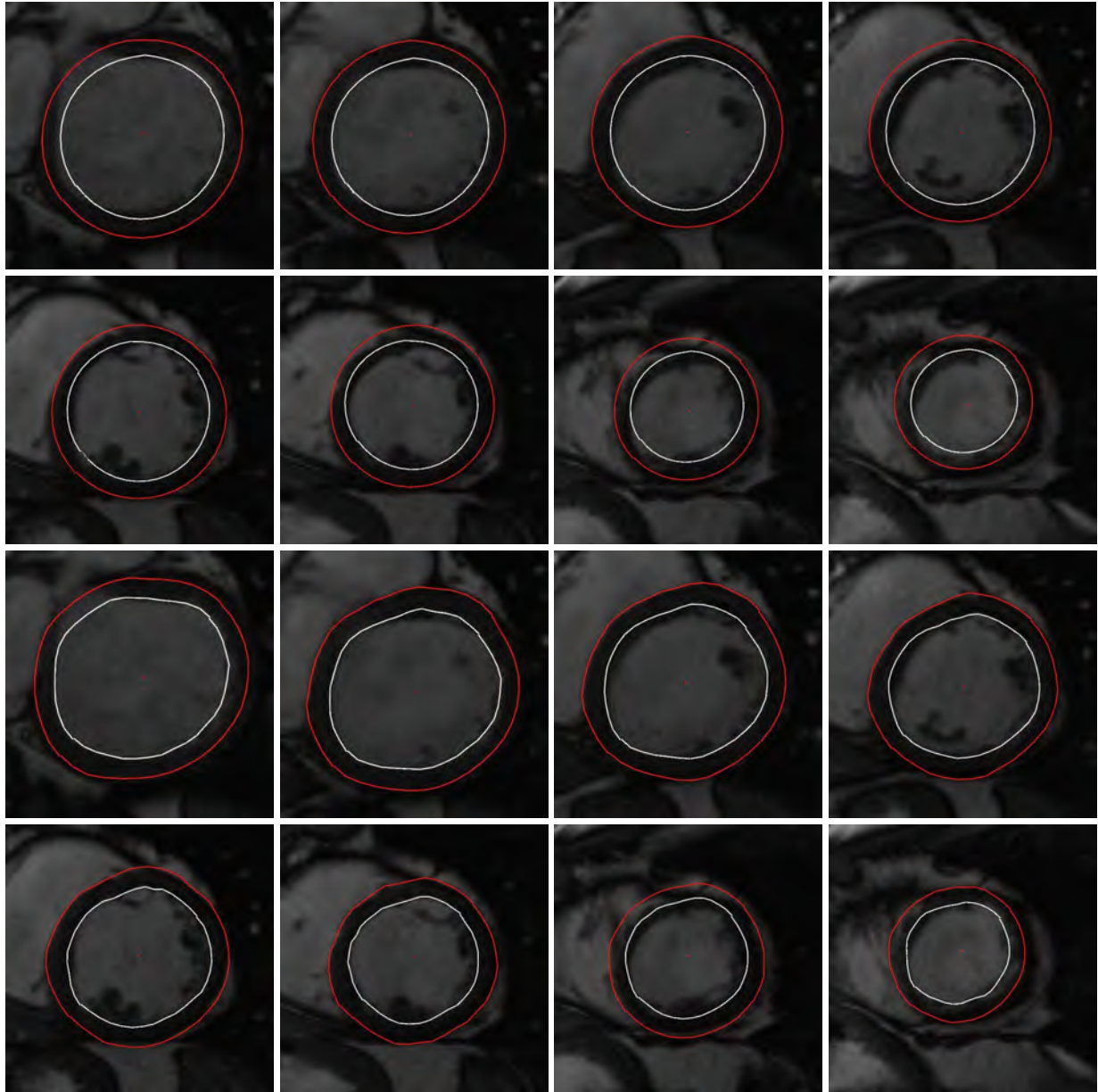
Figure 7.2: *POIC-AAM without and with refinement. Top 2 rows: POIC-AAM only. Bottom 2 rows: POIC-AAM + Refinement*
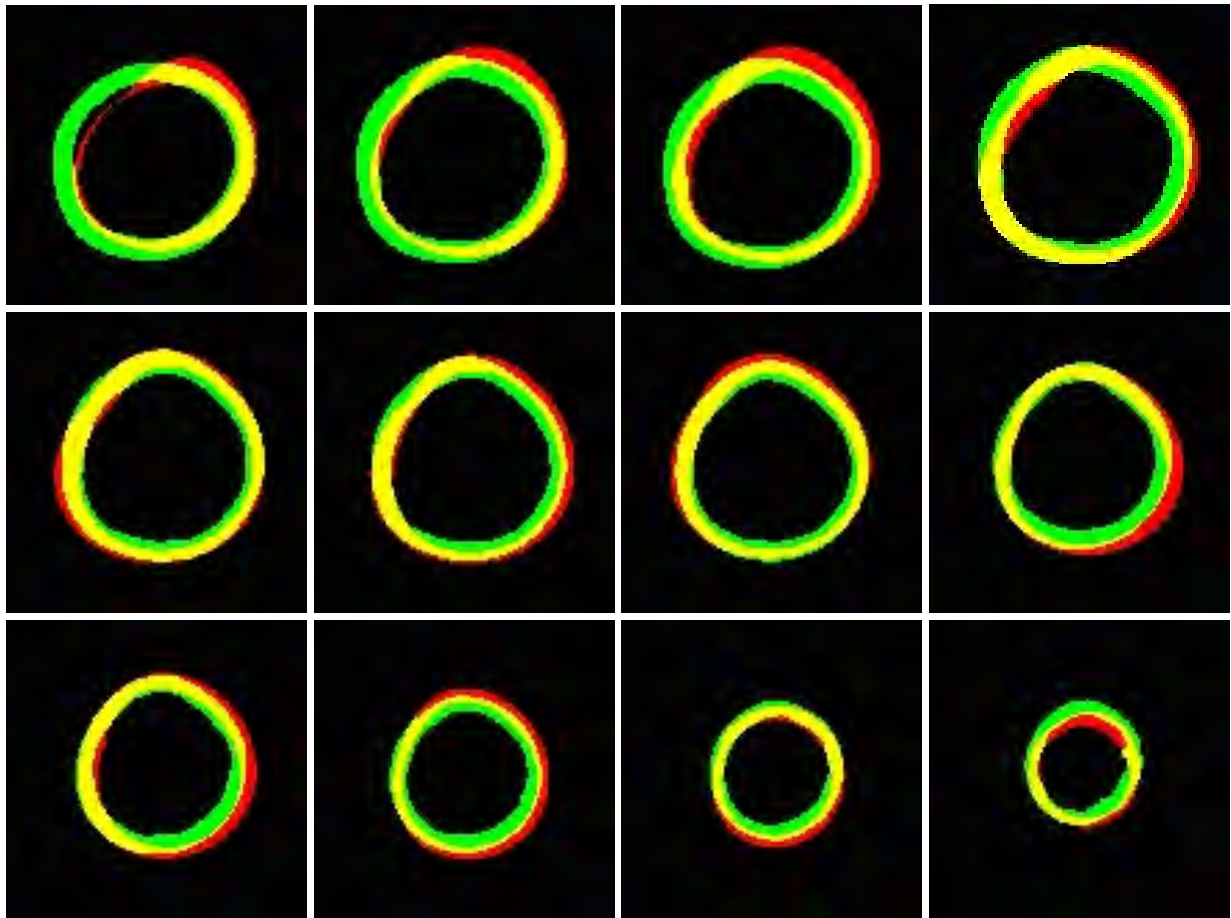
Figure 7.3: *Overlap of ground truth and estimated myocardial pixels (MYO) across slices. TP (yellow), FP (green), FN (red), and TN (black).*

regions when compared to the apex region. Dice coefficients and Average Perpendicular Distance (APD) [Radau et al., 2009] were also measured for the test cases. The dice coefficients were 0.87 and 0.89, and APD were 3.06 and 3.17 mm for endo and epi, respectively. Results by other works (such as [Feng et al., 2013; Huang et al., 2009; Lu et al., 2013]) are either reported on private databases or they include final averages for both ED and ES phases on the MICCAI database. Since our method has been applied only to the ED phase, an objective comparison of the final performance metrics will be hard.

7.4 Automated Dynamic Reconstruction

This section describes the results for the automated statistical-deformable-model based 4D reconstruction method proposed in Section 6.2. The CAP database consisting of cine MR images for 100 training and 95 validation cases was used to train and test the proposed unified variational method. This database contains 4D consensus ground truth myocardium labels for the 100 training cases. These training cases were used to build the statistical shape and appearance models for SIC-AAM.

The shape model is a 3D model (not 4D) which was trained by applying PCA (see Appendix B.3 for more details) to aligned 3D LV shapes from all phases of the cardiac cycle. For each of the 100 training cases, we used 10 3D LV shapes that were equally spaced across the cardiac cycle. The endo and epi borders were sampled at equal distances in the clockwise direction, starting from the anterior RV insertion point. Landmark correspondences were established across the slices and frames using the symmetric nearest neighbor method [Papademetris et al., 2002] (for both space and time) and the resulting shapes were aligned using Orthogonal Procrustes Analysis (OPA; see Appendix B.2 for more details).

During the iterative Gauss-Newton procedure for model fitting, 8 non-rigid shape modes and 30 appearance modes were used. The weights used for the energy terms are $\alpha = 1$ and $\gamma = 10^{-2}$, and the diagonal elements of $\mathbf{B}_1$ and $\mathbf{B}_2$ were all set to 0.05. Since the target volume V_{tar} is unknown, the volume constraint E_{vol} was turned off for the first run until convergence. The average myocardium volume for all the converged shapes is used as the target volume in the second run. Figure 7.4 shows the volume differences in the two runs for two cases. In the second run, the volume difference across phases falls to very low values as compared to the differences in the first run.

To speed up the convergence of model fitting, a multi-resolution approach is employed. Two appearance models, one at full resolution and the other at half the resolution, were trained and used during fitting. To fully automate the method, an initial guess for SIC-AAM was computed using the HoG detector described in Chapter 5. Figures 7.5, 7.6, and 7.7 show results of the
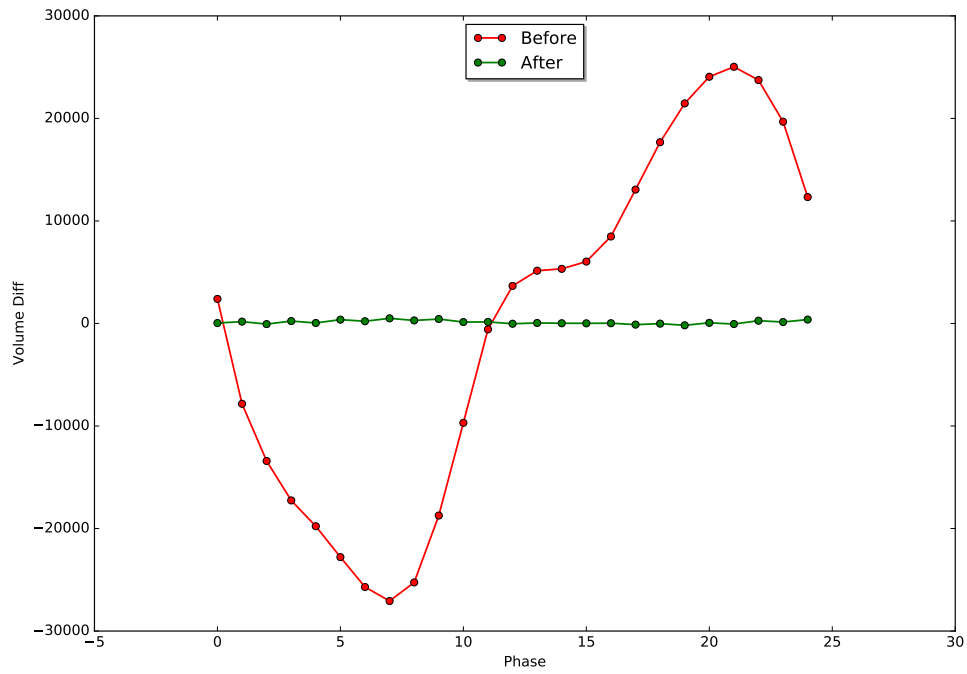
(a) Case 1
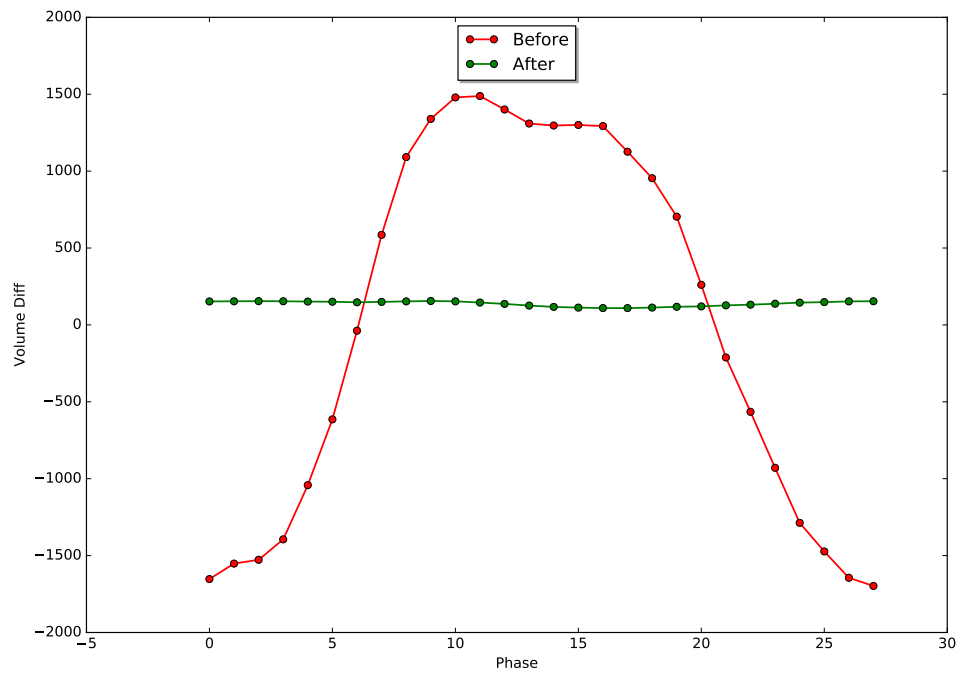
(b) Case 2

Figure 7.4: *Volume Constraint*

Figure 7.5: *Segmentation for Basal Slice*

Figure 7.6: *Segmentation for Mid Slice*
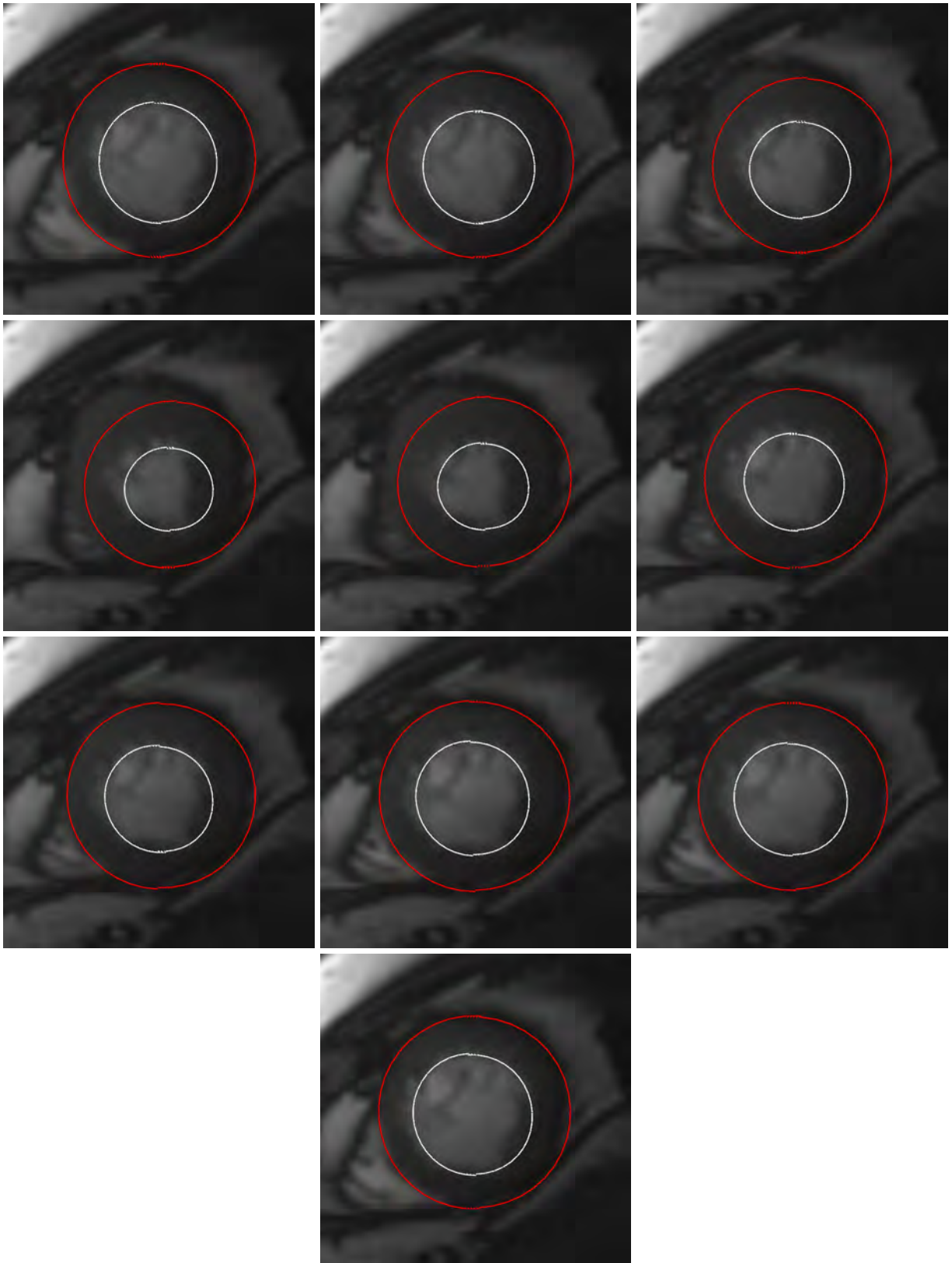
Figure 7.7: *Segmentation for Apical Slice*
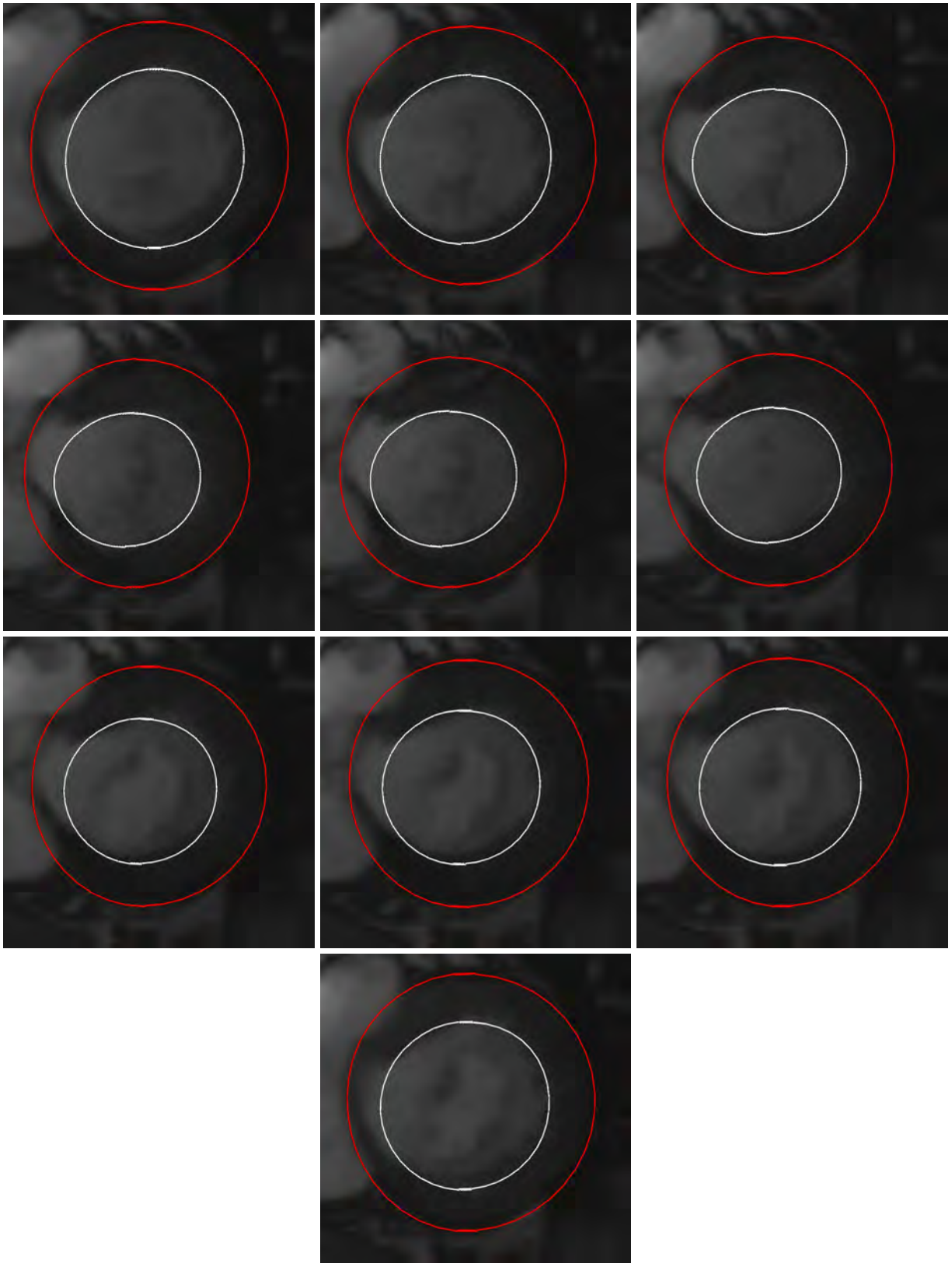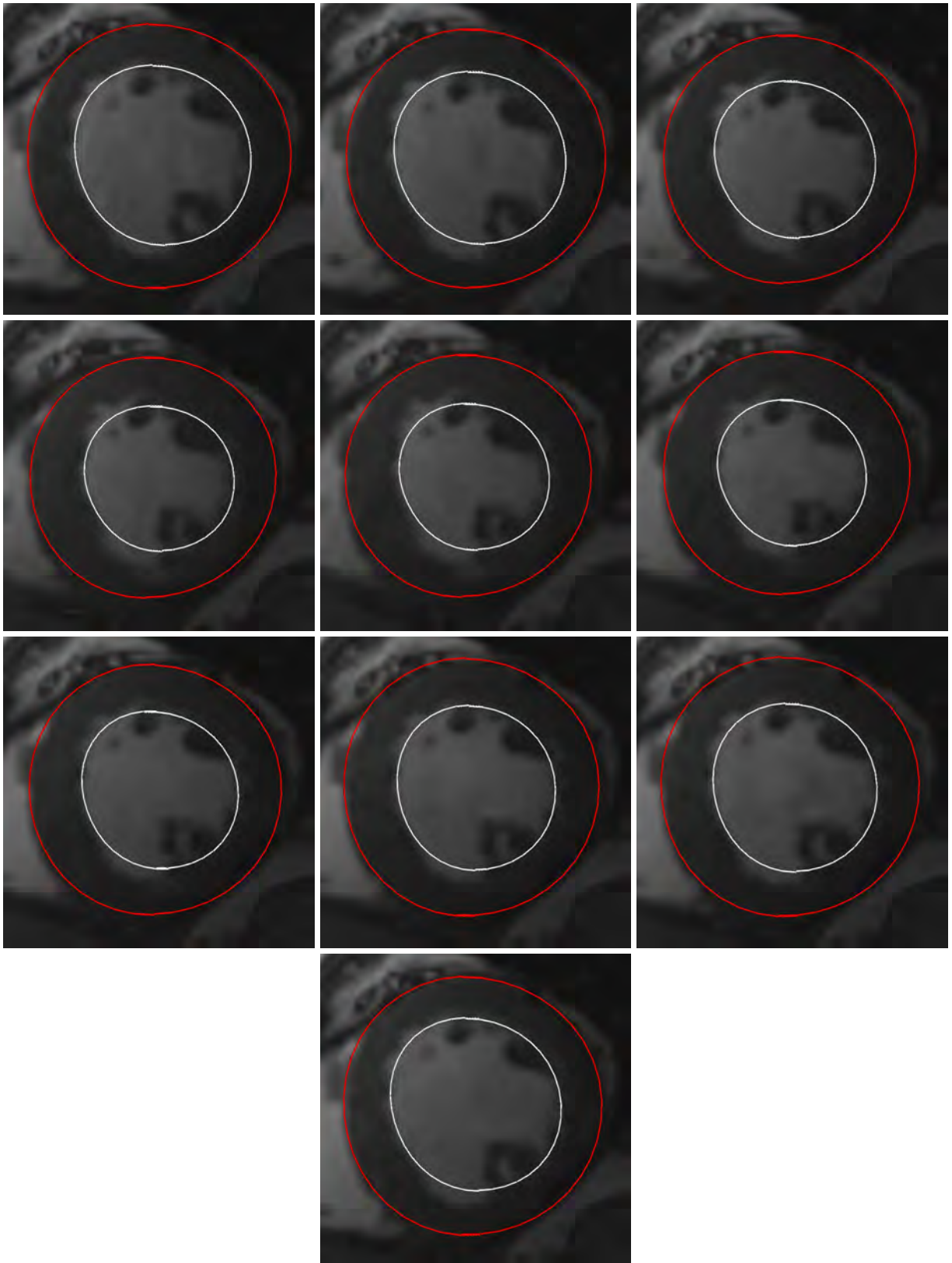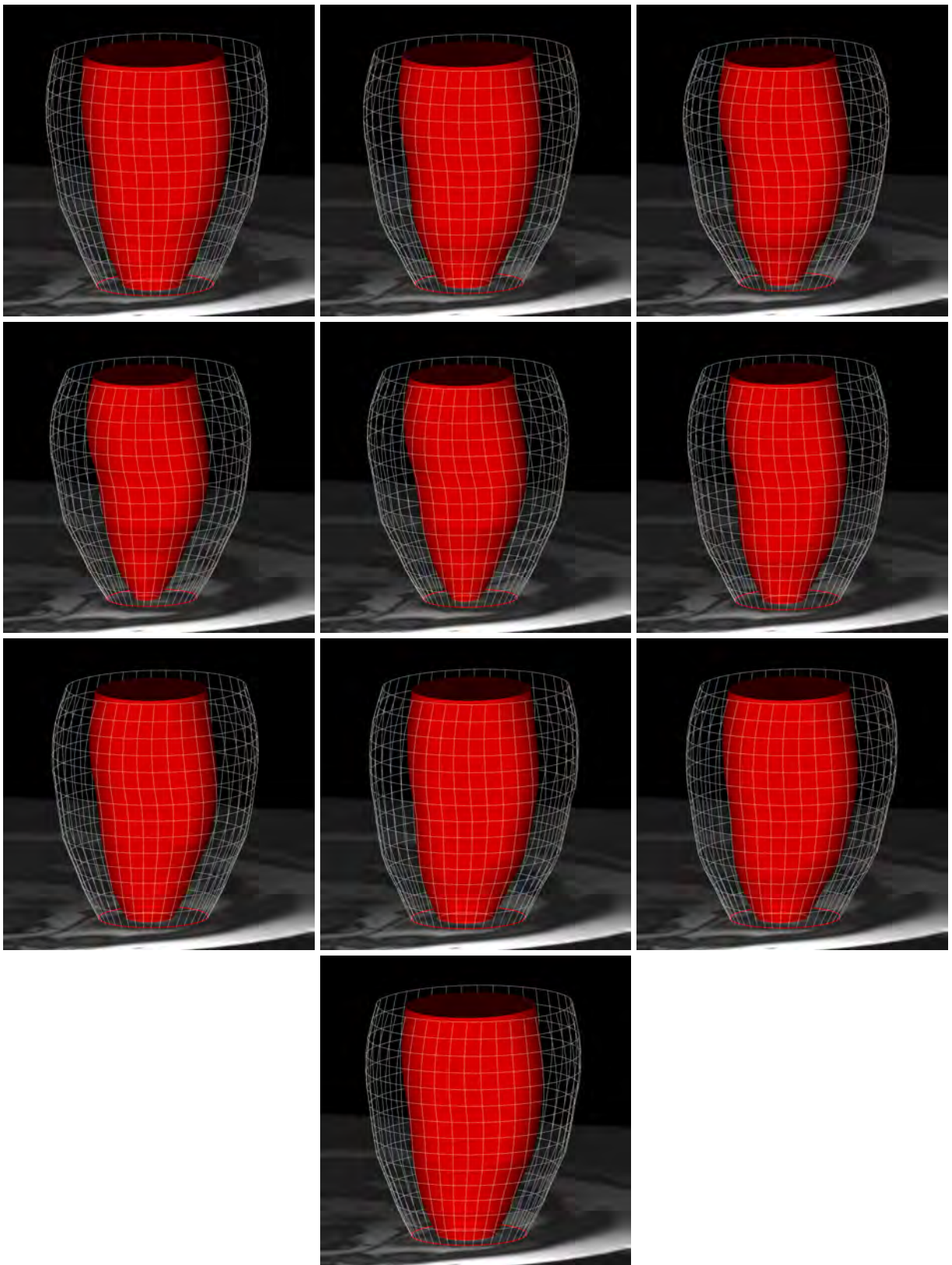
Figure 7.8: *4D LV Reconstruction*

CAP Consensus				
Fully Automated	Method	Precision	Recall	F-Score
no	AO	0.82(0.12)	0.88(0.15)	0.85
no	AU	0.91(0.13)	0.89(0.13)	0.89
no	DS	0.74(0.15)	0.80(0.17)	0.77
yes	SCR	0.87(0.16)	0.74(0.23)	0.79
yes	**Ours**	**0.64(0.15)**	**0.82(0.14)**	**0.72**
yes	INR	0.50(0.10)	0.89(0.17)	0.64

Table 7.5: *Segmentation overlap for CAP data*

method for basal, mid, and apical slices, respectively, from a test case. The figures contain every other frame for each slice. Figure 7.8 shows the same results in 3D.

To measure the accuracy of the method, we compute the precision (7.4), recall (7.5), and f-score (7.6) for the overlap between the set of ground truth M_{gt} and the estimated M_{est} myocardium masks. Table 7.5 contains scores of our method along with the scores of other methods that have been reported in [Suinesiaputra et al., 2014]. An f-score of 0.72 shows good agreement with the consensus ground truth, and it is comparable to those of the other two automated methods, INR [Margeta et al., 2012] and SCR [Jolly et al., 2012].

7.5   Summary

This chapter presented the experiments and performance analysis for our proposed detection and reconstruction techniques. The unified deformable model was applied to statically reconstruct the LV from cardiac cine MR images. The model was trained and tested on the MICCAI database and it achieves a mean Dice coefficient and mean MPD of 0.86 and 3.65 mm, respectively. Next, the proposed HoG based LV detector wasa trained and tested on both the MICCAI and CAP databases. The f-score averages are 0.89 and 0.91 respectively. Next, the automated static reconstruction technique using POIC-AAM and the unified model was trained and tested on the ED phase of the MICCAI database. The average f-score for endo, epi, and myocardium overlap were 0.80, 0.82, and 0.46, respectively. Finally, the automated variational technique for dynamic LV reconstruction

using SIC-AAM and temporal constraints was trained and tested on the CAP database. The average f-score for myocardium overlap across all the phases was 0.72.

CHAPTER 8

Conclusion

We now summarize the key contributions of the thesis and discuss promising directions for future research.

We have proposed a novel deformable model which combines a strong statistical prior learned from manually-segmented training data with a finite element deformable skin. Our model is unique in the sense that the statistical reference shape is embedded in its physical formulation and it evolves simultaneously along with the pose and local displacement parameters, all under the influence of external forces. The model is more generic in the sense that any kind of external forces (gradient, inertial, or optical flow) can be applied without any change to the model's formulation. Such abstraction of external forces facilitates the design of image potential functions that influence the shape of the model. For example, we can design a potential based on optical flow to change the shape of the model according to the LV motion across phases in the images. The finite element skin provides good local control over the smoothness of the surface, and by virtue of its ability to evolve independently, it is able to assume shapes that may not have been captured by the learned deformations of the statistical reference shape.

The survey by Petitjean and Dacher [2011] shows that either full automation has not been achieved or that it is not addressed/reported by a majority of the methods in literature. Many methods are semi-automated due to the unavailability of initial estimates for their models and, hence, they require some form of user iteration, such as indicating the LV position in the image by clicking in the middle of the blood pool or by drawing a circle around it. We addressed this crucial step by proposing a HoG based LV detector for the cardiac cine MR volume. The estimates of LV

position and scale obtained in this first step are used to initialize our model-based approaches to automatically reconstruct the LV with as few assumptions of the images as possible.

A method to automate static LV reconstruction using a combination of POIC-AAM and our unified model was proposed. While POIC-AAM fits to visual data using global appearance, the unified model fits to visual data using local image features. The complimentary combination of the two models provided the capability to capture the finer details present in visual data, while also paying attention to global visual cues.

Many of the reported methods in the research literature address the static reconstruction of the LV for just the ED and ES phases. It is hard for such methods to achieve temporally consistent reconstructions. We proposed a fully automated method for 4D LV reconstruction using the SIC-AAM statistical deformable model. The variational method incorporates temporal smoothness and volume constraints to recover LV shape and motion in a consistent manner.

Most methods reported in the research literature for LV reconstruction have been tested on private image databases [Suinesiaputra et al., 2014; Petitjean and Dacher, 2011] and, hence, an objective way to compare the performance of these methods is precluded. In this thesis, the results of our methods are reported for the publicly available MICCAI LV Segmentation Challenge and CAP Consensus image databases. The LV detection and the LV segmentation f-scores of our proposed methods show good agreement between the output of the automated system and the expert ground truth.

8.1 Future Work

Building a fully automated system to produce highly precise/accurate LV reconstructions is challenging. Although much effort has been invested in this work to tackle the many sub-problems of the system, there are some issues that we intend to address in future work.

Chapter 4 illustrated the benefits of adding local control to statistical deformable models. Both,

deterministic and statistical modes of deformation contribute to the changes in shape as external image based forces are applied. Currently, these forces are computed from local image features that provide very good border accuracy for segmentation problems. One avenue that requires more investigation is the incorporation of both local features and global appearance for the unified model (instead of having separate models), so as to improve the border accuracy of statistical deformable models, such as the AAM.

In our work, we did not investigate the cardiac cine MR acquisition and image formation processes themselves. Instead, the focus was more on the analysis of images that have already been acquired. The internal details of the acquisition process could reveal avenues for overcoming the challenges (as discussed in Section 1.1.4) faced by model based shape reconstruction methods. For example, if the contrast between the myocardium and outer organs (especially the lungs) can be enhanced to clearly distinguish the different tissue types, then models such as the AAM (that rely on appearance) will do a better job at delineating the epicardium boundary. Therefore, such changes to the acquisition process could prove to be conducive to improving the precision of model-based methods.

Reconstructing the shape and motion of the LV from cardiac short-axis cine MR images alone is a challenging problem. Due to the thickness of the slices and a gap of 8-10 mm between them, it is hard, or perhaps impossible, to estimate the motion along the LV longitudinal axis (z-direction). Tracking the valve plane and LV apex in long-axis images (4 chamber and 2 chamber) can help. Although since only two slices are imaged for the long axes, they do not yield a complete 3D volume that can be incorporated in appearance models. Moreover, the long-axis images can be mis-registered with respect to the short-axis images, and hence algorithms that consider the appearance from both short-axis and long-axis images could encounter misleading visual cues. Hence, a separate preprocessing phase to register the long-axis images with the 3D volume (defined by the short-axis images) is necessary. Having more long-axis slices along the cross-section of the LV should be beneficial in building a more accurate 3D volume.

Another way to improve LV motion estimates could be to integrate wall motion and deformation

information available in tagged MRI, with our current physics-based models. From our experience, adding more priors on the shape, motion, and appearance to the models will be beneficial.

Finally, the HoG detector and the statistical models of shape and appearance have to be trained on larger image sets so as to increase their generalization capabilities by incorporating more variations as part of prior knowledge. More image databases with 4D expert ground truth labels should be made available publicly. Currently, the CAP consensus database contains 100 patient volumes that have 4D consensus ground truth. More such public databases will help in building better models and more LV segmentation challenges will provide a better assessment of the state-of-the-art in fully-automated methods.

APPENDIX A

UniMR

The development of a fully automated shape reconstruction system requires a software infrastructure that facilitates easy access to some basic functions. UniMR is one such tool that we have developed. Some of the main features of this tool are described below.

1. **Visualization** — UniMR can be used to visualize cardiac cine MR volumes, LV geometry, LV bounding boxes, and myocardium boundary contours at multiple resolutions, as shown in Figure A.1. It also enables the real-time visualization of fitting iterations for an LV reconstruction algorithm. The slices for a patient's cine MR DICOM images are sequenced and positioned according to the coordinate system defined in the DICOM headers. The visualization window enables mouse interaction for zooming and rotating the volume view. Keyboard functionality enables the changing of slices/frames and the changing of image brightness/contrast. For debugging purposes (or just for curiosity), UniMR can display filtered MR volumes. As shown in Figure A.2, some of the filters are Canny edge, Gaussian smoothing, gradient magnitude, and gradient magnitude of an image-gradient-based potential function. Similarly, HoG cells (Figure A.3(a)) and optical flow vectors (Figure A.3(b)) can also be visualized.

2. **Prototyping** — UniMR employs an object-oriented segmentation framework that facilitates the easy plug-and-play of LV reconstruction method prototypes. Any new method can be prototyped and incorporated easily into UniMR without having to worry about the details of image volume and LV visualization. Several statistical and deterministic deformable-model-based methods were developed using this abstract segmentation interface. If a method does

68
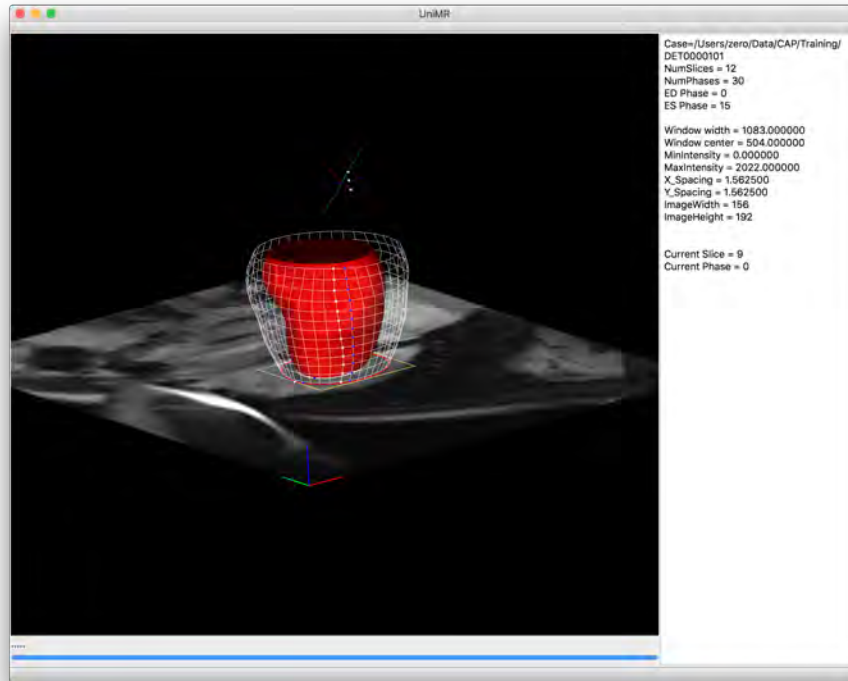
(a) Full Resolution
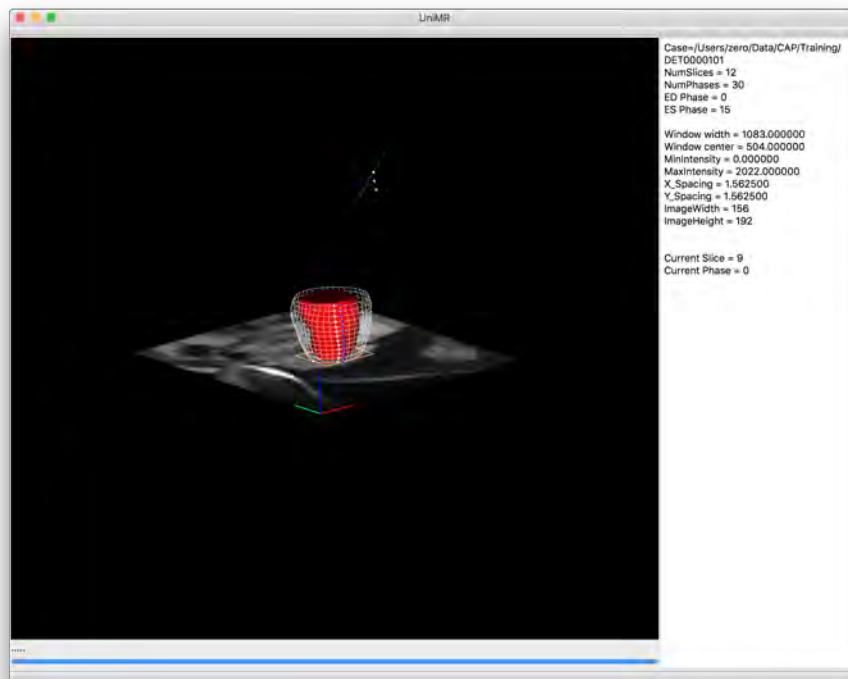
(b) Half Resolution

Figure A.1: *UniMR interface to visualize cardiac MRI volume multi-resolution pyramids*
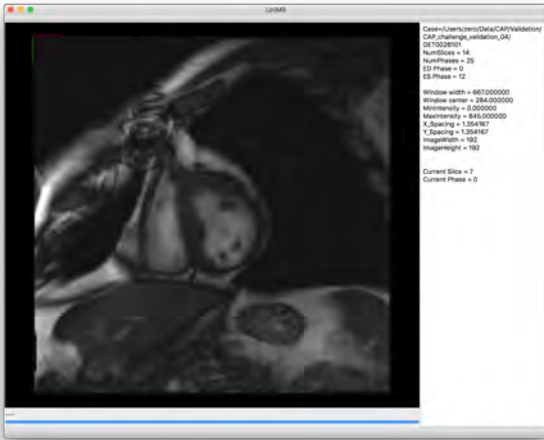
(a) Short-Axis MR image
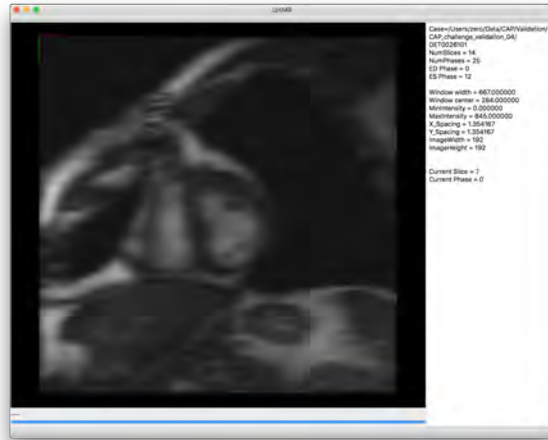
(b) Short-Axis MR image (blurred)
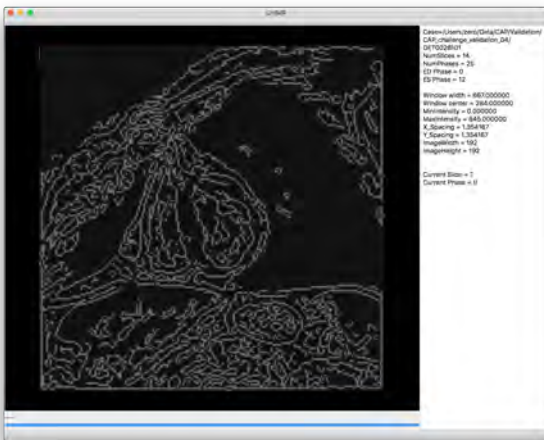
(c) Canny Edge
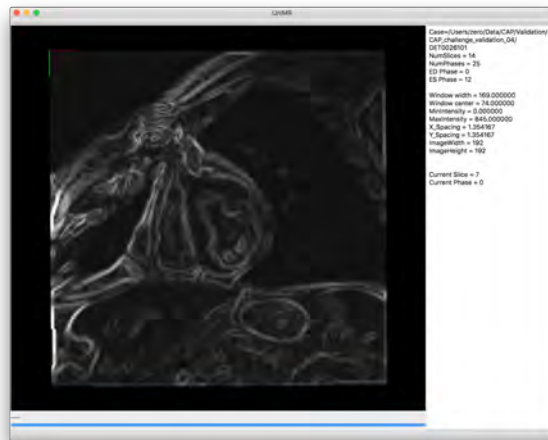
(d) Gradient Magnitude
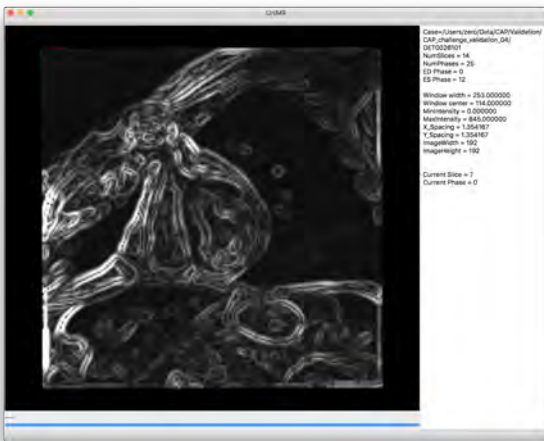
(e) Gradient Magnitude of a potential
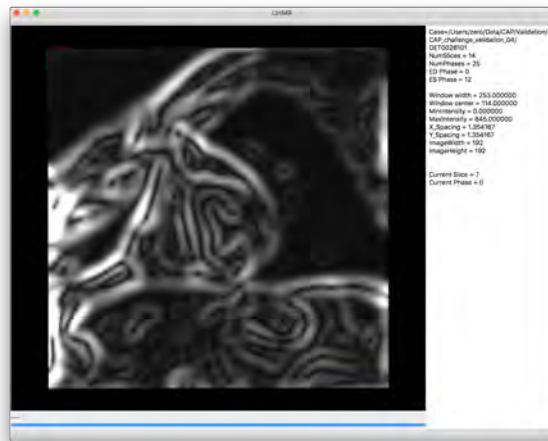
(f) Gradient Magnitude of a potential (blurred)

Figure A.2: *UniMR Features*
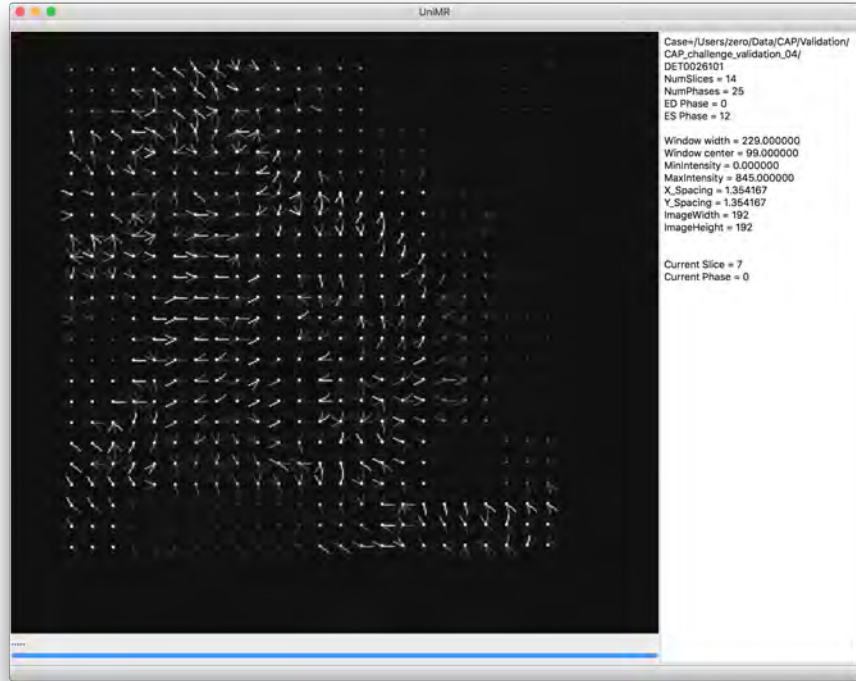
(a) Histogram Of Oriented Gradients (HoG)
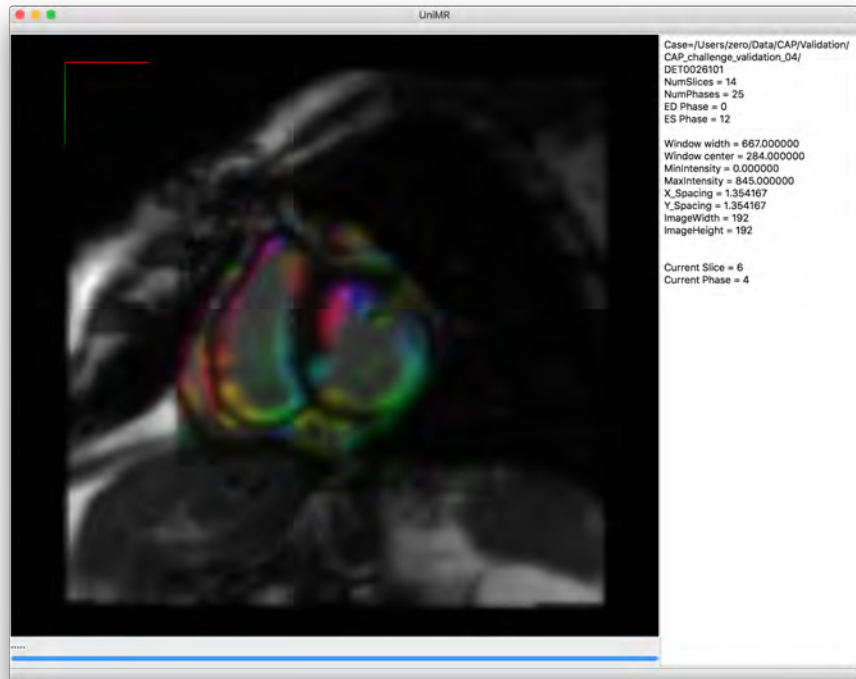
(b) Optical Flow

Figure A.3: *UniMR Features (continued)*

71

not work (and some early prototypes did fail), then one can easily de-link it from the main infrastructure.

3. **Experiments** — An LV reconstruction algorithm's performance is tested using overlap and surface positioning quantitative metrics such as precision, recall, dice, jaccard, and mean perpendicular distance. UniMR has functions that implement all these metrics. Training image volumes and validation image volumes can be explicitly specified in XML files. Leave-one-out validation can also be chosen. It supports the output of LV reconstruction algorithms (bounding boxes, myocardium boundaries, phase info, etc.) in a custom XML file format.

UniMR is written in the C/C++ language for Mac OSX 10.11. It uses the Insight Toolkit (ITK) [Johnson et al., 2013] for the low-level filtering/convolution operations and for loading DICOM images. The QT Graphical User Interface API was used to develop the front end OpenGL-based visualization windows. External libraries Armadillo [Sanderson and Curtin, 2016], LibSVM [Chang and Lin, 2011], and MLPack [Curtin et al., 2013] were incorporated for linear algebra, SVM, and machine learning algorithms, respectively. Whenever possible, OpenMP was used to parallelize tasks across multiple CPU cores.

In summary, UniMR enables the quick prototyping, performance analysis, and debugging of LV reconstruction algorithms, without having to worry about the details of volume loading, sequencing, and visualization.

APPENDIX B

Useful Mathematical Tools

The following sections briefly introduce some useful mathematical tools.

B.1 Singular Value Decomposition (SVD)

The Singular Value Decomposition (SVD) is a very useful matrix factorization that can provide insights into the structure of a matrix. Let $\mathbf{A} \in \mathcal{R}^{m \times n}, \mathbf{A} : \mathcal{R}^n \to \mathcal{R}^m$ be a map that takes vectors in $\mathcal{R}^n$ to vectors in $\mathcal{R}^m$. The SVD of $\mathbf{A}$ is given by $\mathbf{A} = \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^\mathsf{T} \\ \mathbf{V}_2^\mathsf{T} \end{bmatrix}$, where r is the rank of $\mathbf{A}$. Here,

- $\mathbf{U}_1 \in \mathcal{R}^{m \times r}$ forms an orthonormal basis for the range of $\mathbf{A}$;

- $\mathbf{V}_1 \in \mathcal{R}^{n \times r}$ forms an orthonormal basis for the range of $\mathbf{A}^\mathsf{T}$;

- $\mathbf{\Sigma}_r \in \mathcal{R}^{r \times r}$ is a diagonal matrix that contains the non-negative singular values sorted by descending magnitude;

- $\mathbf{U}_2 \in \mathcal{R}^{m \times (m-r)}$ forms an orthonormal basis for the null space of $\mathbf{A}^\mathsf{T}$;

- $\mathbf{V}_2 \in \mathcal{R}^{n \times (n-r)}$ forms an orthonormal basis for the null space of $\mathbf{A}$.

The SVD is widely used in least-squares fitting problems (pseudoinverse) and in computing the ranks, ranges, and null-spaces of matrices, and in computing low-rank matrix approximations.

B.2 Orthogonal Procrustes Analysis (OPA)

Orthogonal Procrustes Analysis (OPA) is a procedure that is used to align two given shapes. OPA finds an optimal rotation $\mathbf{R} \in \mathcal{R}^{3 \times 3}$ by performing an SVD of the inter-shape correlation matrix. If $\mathbf{x}_1 \in \mathcal{R}^{3 \times p}$ and $\mathbf{x}_2 \in \mathcal{R}^{3 \times p}$ are vectors for two origin-centered shapes, then the OPA procedure solves the following optimzation problem:

$$\underset{\mathbf{R}}{\arg \min} \, ||\mathbf{R}\mathbf{x}_2 - \mathbf{x}_1||_F^2, \tag{B.1}$$

whose solution is found by performing an SVD of the inter-shape correlation matrix to yield

$$\mathbf{x}_2 \mathbf{x}_1^\mathsf{T} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\mathsf{T}. \tag{B.2}$$

The optimal rotation matrix $\mathbf{R}$ is given by

$$\mathbf{R} = \mathbf{V}\boldsymbol{\Delta}\mathbf{U}^\mathsf{T}, \tag{B.3}$$

where $\boldsymbol{\Delta} = \mathrm{diag}([1, \ldots, 1, \det(\mathbf{V}\mathbf{U}^\mathsf{T})]) \in \mathcal{R}^{3 \times 3}$.

B.3 Principal Components Analysis (PCA)

In modeling scenarios where data are assumed to be generated from a low-dimensional basis, PCA is a technique that can be used for dimensionality reduction [Hastie et al., 2009]. It is an unsupervised technique for building a generative model of the given data by fitting a hyper-ellipsoid. The axes of the ellipsoid capture the uncorrelated (orthonormal) directions of maximum data variance, and they help in restricting attention to those "principal" dimensions. In addition to providing a low-rank basis, a reduced set of parameters for data models (built using PCA) can help in speeding up iterative optimization algorithms [Cootes et al., 1998; Gopal et al., 2013; Gopal and Terzopoulos, 2016].

Let the columns of the data matrix $\mathbf{D} \in \mathcal{R}^{m \times n}$ represent a set of n data points in m-dimensional space (assume $m >> n$). Computing the orthonormal maximum variance dimensions involves the following steps:

1. Compute the mean $\bar{\mathbf{x}} \in \mathcal{R}^m$ of the data points:

$$\bar{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{D}[i],$$

where $\mathbf{D}[i]$ is the ith column of $\mathbf{D}$.

2. Center the data:

$$\mathbf{X} = \mathbf{D} - \bar{\mathbf{x}} \mathbf{1}^\mathsf{T},$$

where $\mathbf{1} \in \mathcal{R}^n$ is a vector of ones.

3. Compute covariance matrix $\mathbf{A} \in \mathcal{R}^{n \times n}$:

$$\mathbf{A} = \mathbf{X}^\mathsf{T} \mathbf{X}.$$

4. Since $\mathbf{A}$ is a symmetric positive semidefinite matrix, the SVD can be used to factorize $\mathbf{A}$ as follows:

$$\mathbf{A} = \mathbf{Q} \Lambda \mathbf{Q}^\mathsf{T},$$

where the diagonal elements of $\Lambda = \mathrm{diag}([\lambda_1, \lambda_2, ..., \lambda_n])$ are the eigenvalues of $\mathbf{A}$ sorted in descending order, and the columns of $\mathbf{Q}$ are the associated eigenvectors of $\mathbf{A}$.

5. The principal components for the data in $\mathbf{D}$ are the eigenvectors of $\mathbf{X}\mathbf{X}^\mathsf{T}$. Therefore, columns of the matrix $\mathbf{X}\mathbf{Q} \in \mathcal{R}^{m \times n}$ form the n orthogonal principal components (although, these have to be normalized) with the corresponding variances $\lambda_1, \lambda_2, \ldots, \lambda_n$ respectively.

6. Select the top k principal components such that they account for $p\%$ of the variance; i.e.,

$$\frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{n} \lambda_i} \times 100 \geq p.$$

Usually, p is chosen in the range 90%–98%.

APPENDIX C

Unified Variational Method Formulation

This appendix provides more details about the formulation for the Unified Variational Method developed in Section 6.2.

C.1 Warping in 3D

We use a fixed tessellation of the LV shape using tetrahedral elements. Given the rigid $(\mathbf{q}_c, \mathbf{q}_\theta)$ and non-rigid $(\mathbf{q}_s)$ shape parameters, we map every point $\mathbf{x} \in \mathcal{R}^3$ in the base mesh volume $\bar{\mathbf{s}}$ to the image volume $I : \mathcal{R}^3 \to \mathcal{R}$ using a coordinate transform $\mathbf{N} : \mathcal{R}^3 \to \mathcal{R}^3$ defined as

$$\mathbf{N}(\mathbf{x}; \mathbf{q}_c, \mathbf{q}_\theta, \mathbf{q}_s) = \mathbf{R}(\mathbf{q}_\theta)\mathbf{W}(\mathbf{x}; \mathbf{q}_s) + \mathbf{T}(\mathbf{q}_c) \tag{C.1}$$

where $\mathbf{R}$ is rotation transform, $\mathbf{T}$ is translation transform, and $\mathbf{W}$ is a piecewise affine warp.

The piecewise affine warp $\mathbf{W}(\mathbf{x}; \mathbf{q}_s) : \mathcal{R}^3 \to \mathcal{R}^3$ maps $\mathbf{x}$ in the mean shape $\bar{\mathbf{s}}$ to a point $\mathbf{W}(\mathbf{x}; \mathbf{q}_s)$ inside the corresponding tetrahedron in the model shape $\mathbf{s}$ immersed in the test volume I. If $[\mathbf{p}_i^0, \mathbf{p}_j^0, \mathbf{p}_k^0, \mathbf{p}_l^0] \in \mathcal{R}^3$ are the 4 vertices of a tetrahedron in the mean shape and $[\mathbf{p}_i, \mathbf{p}_j, \mathbf{p}_k, \mathbf{p}_l] \in \mathcal{R}^3$ are the vertices in the corresponding tetrahedron in the model shape, then the warp takes the form,

$$\mathbf{W}(\mathbf{x}; \mathbf{q}_s) = \mathbf{p}_i + \alpha(\mathbf{p}_j - \mathbf{p}_i) + \beta(\mathbf{p}_k - \mathbf{p}_i) + \gamma(\mathbf{p}_l - \mathbf{p}_i), \tag{C.2}$$

77

where α, β, and γ are obtained from

$$
\begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \left[(\mathbf{p}_j^0 - \mathbf{p}_i^0) \quad (\mathbf{p}_k^0 - \mathbf{p}_i^0) \quad (\mathbf{p}_l^0 - \mathbf{p}_i^0) \right]^{-1} (\mathbf{x} - \mathbf{p}_i^0)
$$

Since the tetrahedron vertices $[\mathbf{p}_i, \mathbf{p}_j, \mathbf{p}_k, \mathbf{p}_l]$ in the model shape $\mathbf{s}$ are functions of the shape parameters $\mathbf{q}_s$, the warp Jacobian is given by

$$
\frac{\partial \mathbf{W}}{\partial \mathbf{q}_s} = \begin{bmatrix} \sum_v \frac{\partial W_x}{\partial p_{v_x}} \frac{\partial p_{v_x}}{\partial \mathbf{q}_s} \\ \sum_v \frac{\partial W_y}{\partial p_{v_y}} \frac{\partial p_{v_y}}{\partial \mathbf{q}_s} \\ \sum_v \frac{\partial W_z}{\partial p_{v_z}} \frac{\partial p_{v_z}}{\partial \mathbf{q}_s} \end{bmatrix} \in \mathcal{R}^{3 \times n},
\tag{C.3}
$$

where we are summing over the $v = i, j, k, l$ vertices, and where

$$
\frac{\partial W_x}{\partial p_{i_x}} = 1 - \alpha - \beta - \gamma, \quad \frac{\partial W_x}{\partial p_{j_x}} = \alpha, \quad \frac{\partial W_x}{\partial p_{k_x}} = \beta, \quad \frac{\partial W_x}{\partial p_{l_x}} = \gamma.
\tag{C.4}
$$

The derivatives $\frac{\partial p_{v_x}}{\partial \mathbf{q}_s}$, $\frac{\partial p_{v_y}}{\partial \mathbf{q}_s}$, and $\frac{\partial p_{v_z}}{\partial \mathbf{q}_s}$ are obtained from (6.6).

C.2 Rotation Parameterization

Due to its ease of use in iterative optimization schemes [Taylor and Kriegman, 1994], we choose the exponential map parameterization $\exp : so(3) \to SO(3)$ for the rotation $\mathbf{R}(\mathbf{q}_\theta) = \exp(\mathbf{q}_\theta^*))$, where $so(3) = \{\omega^* \in \mathcal{R}^{3 \times 3} | \omega \in \mathcal{R}^3\}$ is the special orthogonal Lie algebra denoting the space of skew-symmetric matrices. Due to the computation of parameter updates in an inverse compositional approach in SIC [Matthews and Baker, 2004], the derivatives of the transform $\mathbf{N}$ with respect to

$\mathbf{q}_\theta$, $\mathbf{q}_c$, and $\mathbf{q}_s$ are evaluated at $\mathbf{q} = \mathbf{0}$ and they take the forms

$$\left.\frac{\partial \mathbf{N}}{\mathbf{q}_\theta}\right|_{\mathbf{q}=0} = [\mathbf{W}(\mathbf{x}; \mathbf{q}_s)|_{\mathbf{q}_s=\mathbf{0}}]^* = [\mathbf{x}]^*; \tag{C.5}$$

$$\left.\frac{\partial \mathbf{N}}{\mathbf{q}_c}\right|_{\mathbf{q}=0} = \mathbf{I} \in \mathcal{R}^{3\times3} \quad \text{Identity matrix}; \tag{C.6}$$

$$\left.\frac{\partial \mathbf{N}}{\mathbf{q}_s}\right|_{\mathbf{q}=0} = \left.\frac{\partial \mathbf{N}}{\partial \mathbf{W}}\right|_{\mathbf{q}=0} \left.\frac{\partial \mathbf{W}}{\partial \mathbf{q}_s}\right|_{\mathbf{q}=0} = \left.\mathbf{R}(0)\frac{\partial \mathbf{W}}{\partial \mathbf{q}_s}\right|_{\mathbf{q}=0} = \left.\frac{\partial \mathbf{W}}{\partial \mathbf{q}_s}\right|_{\mathbf{q}=0}. \tag{C.7}$$

C.3 Incompressibility Constraint

If $\mathbf{p}_i$, $\mathbf{p}_j$, $\mathbf{p}_k$ and $\mathbf{p}_l$ are the 4 vertices of a tetrahedron in the model shape $\mathbf{s}$, then its volume is given by

$$V = \frac{1}{6}\det\left[(\mathbf{p}_j - \mathbf{p}_i) \quad (\mathbf{p}_k - \mathbf{p}_i) \quad (\mathbf{p}_l - \mathbf{p}_i),\right] \tag{C.8}$$

and since its vertices are functions of $\mathbf{q}_s$, the derivative $\frac{\partial V}{\mathbf{q}_s}$ can be found from the above expression.

C.4 Euler-Lagrange Equations

Assuming we have an initial estimate at $\mathbf{q}(t)$, we perform a first-order Taylor's expansion of each of the terms inside the energy functional

$$E(\mathbf{q}(t) + \mathbf{r}(t)) = \int_{\mathcal{T}} \alpha E_{data}(\mathbf{q}(t) + \mathbf{r}(t)) + E_{smooth}(\mathbf{q}(t) + \mathbf{r}(t)) + \gamma E_{vol}(\mathbf{q}(t) + \mathbf{r}(t)) \, dt, \tag{C.9}$$

where if $\mathbf{q} = [\mathbf{q}_c^T \mathbf{q}_\theta^T \mathbf{q}_s^T \mathbf{q}_a^T]^T$, then $\mathbf{r} = [\Delta\mathbf{q}_c^T \Delta\mathbf{q}_\theta^T \Delta\mathbf{q}_s^T \Delta\mathbf{q}_a^T]^T$. The Euler-Lagrange equation for the above functional with respect to $\mathbf{r}(t)$ is given by

$$\frac{\partial \mathcal{L}}{\partial \mathbf{r}} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \mathbf{r}_t} + \frac{d}{dt^2}\frac{\partial \mathcal{L}}{\partial \mathbf{r}_{tt}} = 0, \quad \text{where } \mathbf{r}_t = \frac{\partial \mathbf{r}(t)}{\partial t} \text{ and } \mathbf{r}_{tt} = \frac{\partial^2 \mathbf{r}(t)}{\partial t^2}. \tag{C.10}$$

This gives us an ODE of the following form:

$$\mathbf{A}\mathbf{r} = \mathbf{b}, \tag{C.11}$$

where

$$\mathbf{A} = \alpha \sum_{\mathbf{x}} \left(\mathbf{M}(\mathbf{x})\mathbf{M}(\mathbf{x})^T \right) + \gamma \left(\sum_i \nabla V_i(\mathbf{q}) \right) \left(\sum_i \nabla V_i(\mathbf{q}) \right)^T,$$

and

$$
\begin{aligned}
\mathbf{b} =\ & \mathbf{B}_1 \mathbf{r}_{tt} - \mathbf{B}_2 \mathbf{r}_{tttt} \\
& -\alpha \left(\sum_{\mathbf{x}} \left(\bar{\mathbf{a}}(\mathbf{x}) + \mathbf{P}_a(\mathbf{x})\mathbf{q}_a - I(\mathbf{N}(\mathbf{x}; \mathbf{q}_c, \mathbf{q}_\theta, \mathbf{q}_s)) \right) \mathbf{M}(\mathbf{x}) \right) \\
& -\gamma \left(\sum_i V_i(\mathbf{q}) - V_{tar} \right) \sum_i \nabla V_i(\mathbf{q})
\end{aligned}
$$

with

$$\mathbf{M}(\mathbf{x}) = \left[\left(\nabla \bar{\mathbf{a}}(\mathbf{x}) + \nabla \mathbf{P}_a(\mathbf{x})\mathbf{q}_a \right) \left(\frac{\partial \mathbf{N}}{\partial \mathbf{q}_c}, \frac{\partial \mathbf{N}}{\partial \mathbf{q}_\theta}, \frac{\partial \mathbf{W}}{\partial \mathbf{q}_s} \right), \mathbf{P}_a(1)(\mathbf{x}), ..., \mathbf{P}_a(m)(\mathbf{x}) \right]^T,$$

where $\mathbf{P}_a(i)$ refers to the ith column of $\mathbf{P}_a$.

We discretize the above ODE by using 2nd-order accurate finite differences for the 2nd-order and 4th-order terms:

$$\mathbf{r}_{tt} = \frac{\mathbf{r}_{(K-1)}^{(t+1)} - 2\mathbf{r}_{(K)}^{(t)} + \mathbf{r}_{(K-1)}^{(t-1)}}{1^2},$$

$$\mathbf{r}_{tttt} = \frac{\mathbf{r}_{(K-1)}^{(t+2)} - 4\mathbf{r}_{(K-1)}^{(t+1)} + 6\mathbf{r}_{(K)}^{(t)} - 4\mathbf{r}_{(K-1)}^{(t-1)} + \mathbf{r}_{(K-1)}^{(t-2)}}{1^4},$$

where K and $K - 1$ refer to iterations.

The coordinate transform $\mathbf{N}$ involves a piecewise affine warp that does not form a group. Hence, we employ the following approximation to update the parameters $\mathbf{q}(t) = [\mathbf{q}_c^T \mathbf{q}_\theta^T \mathbf{q}_s^T \mathbf{q}_a^T]^T$ with

$\mathbf{r}(t) = [\Delta\mathbf{q}_c^T \Delta\mathbf{q}_\theta^T \Delta\mathbf{q}_s^T \Delta\mathbf{q}_a^T]^T$ at every iteration:

$$\mathbf{R}(\mathbf{q}_\theta) \quad = \quad \mathbf{R}(\mathbf{q}_\theta)\mathbf{R}(\Delta\mathbf{q}_\theta);$$

$$\mathbf{T}(\mathbf{q}_c) = \mathbf{R}(\mathbf{q}_\theta)\mathbf{T}(\Delta\mathbf{q}_c) + \mathbf{T}(\mathbf{q}_c) \quad \Longrightarrow \quad \mathbf{q}_c = \mathbf{R}(\mathbf{q}_\theta)\Delta\mathbf{q}_c + \mathbf{q}_c;$$

$$\mathbf{W}(\mathbf{x};\mathbf{q}_s) \quad \leftarrow \quad \mathbf{W}(\mathbf{x};\mathbf{q}_s) \circ \mathbf{W}(\mathbf{x};\Delta\mathbf{q}_s)^{-1};$$

$$\mathbf{q}_a \quad = \quad \mathbf{q}_a + \Delta\mathbf{q}_a.$$

REFERENCES

Bistoquet, A., Oshinski, J., et al. (2007). Left ventricular deformation recovery from cine MRI using an incompressible model. *IEEE Transactions on Medical Imaging*, 26(9):1136–1153.

Caselles, V., Kimmel, R., and Sapiro, G. (1997). Geodesic active contours. *International Journal of Computer Vision*, 22(1):61–79.

Chan, T. F. and Vese, L. A. (2001). Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277.

Chang, C.-C. and Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

Cootes, T. F., Edwards, G. J., and Taylor, C. J. (1998). Active appearance models. In *Computer Vision – ECCV'98*, pages 484–498. Springer.

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

Curtin, R. R., Cline, J. R., Slagle, N. P., March, W. B., Ram, P., Mehta, N. A., and Gray, A. G. (2013). MLPACK: A scalable C++ machine learning library. *Journal of Machine Learning Research*, 14:801–805.

Dalal, N. and Triggs, B. (2005). Histograms of oriented gradients for human detection. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 886–893.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society; Series B*, pages 1–38.

Feng, C., Li, C., Zhao, D., Davatzikos, C., and Litt, H. (2013). Segmentation of the left ventri-

cle using distance regularized two-layer level set approach. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*, pages 477–484. Springer.

Frangi, A. F., Niessen, W. J., and Viergever, M. A. (2001). Three-dimensional modeling for functional analysis of cardiac images: A review. *IEEE Transactions on Medical Imaging*, 20(1):2–5.

Glass, L., Hunter, P., and McCulloch, A. (1991). *Theory of Heart: Biomechanics, Biophysics, and Nonlinear Dynamics of Cardiac Function*. Institute for Nonlinear Science. Springer New York.

Gopal, S., Otaki, Y., Arsanjani, R., Berman, D., Terzopoulos, D., and Slomka, P. (2013). Combining active appearance and deformable superquadric models for LV segmentation in cardiac MRI. In Karssemeijer, N. and Samei, E., editors, *Proc. SPIE 8869, Medical Imaging 2013: Image Processing*, pages 86690G:1–8.

Gopal, S. and Terzopoulos, D. (2014). A unified statistical/deterministic deformable model for LV segmentation in cardiac MRI. In Camara, O., Mansi, T., Pop, M., Rhode, K., Sermesant, M., and Young, A., editors, *Statistical Atlases and Computational Models of the Heart: Imaging and Modelling Challenges*, volume 8330 of *Lecture Notes in Computer Science*, pages 180–187. Springer, Berlin.

Gopal, S. and Terzopoulos, D. (2016). Automated model-based left ventricle segmentation in cardiac MR images. In Camara, O., Mansi, T., Pop, M., Rhode, K., Sermesant, M., and Young, A., editors, *Statistical Atlases and Computational Models of the Heart: Imaging and Modelling Challenges*, volume 9534 of *Lecture Notes in Computer Science*, pages 3–12. Springer, Berlin.

Hamarneh, G. and Gustavsson, T. (2004). Deformable spatio-temporal shape models: Extending active shape models to 2D + time. *Image and Vision Computing*, 22(6):461–470.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer Series in Statistics. Springer New York, second edition.

Horn, B. K. and Schunck, B. G. (1981). Determining optical flow. *Artificial Intelligence*, 17(1-3):319–331.

Huang, S., Liu, J., Lee, L., Venkatesh, S., Teo, L., Au, C., and Nowinski, W. (2009). Segmentation of the left ventricle from cine MR images using a comprehensive approach. *The MIDAS Journal – Cardiac MR Left Ventricle Segmentation Challenge*.

Isard, M. and Blake, A. (1998). Condensation: Conditional density propagation for visual tracking. *International Journal of Computer Vision*, 29(1):5–28.

Johnson, H. J., McCormick, M., Ibáñez, L., and The Insight Software Consortium (2013). *The ITK Software Guide*. Kitware, Inc., third edition.

Jolly, M. P. (2009). Fully automatic left ventricle segmentation in cardiac cine MR images using registration and minimum surfaces. *The MIDAS Journal – Cardiac MR Left Ventricle Segmentation Challenge*.

Jolly, M.-P., Guetter, C., Lu, X., Xue, H., and Guehring, J. (2012). Automatic segmentation of the myocardium in cine MR images using deformable registration. In *International Workshop on Statistical Atlases and Computational Models of the Heart*, pages 98–108. Springer.

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

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

Lelieveldt, B. P., van der Geest, R. J., Reiber, J. H., Bosch, J. G., Mitchell, S. C., and Sonka, M. (2001). Time-continuous segmentation of cardiac image sequences using active appearance motion models. In *Information Processing in Medical Imaging*, pages 446–452. Springer.

Lorenzo-Valdés, M., Sanchez-Ortiz, G. I., Elkington, A. G., Mohiaddin, R. H., and Rueckert,

D. (2004). Segmentation of 4D cardiac MR images using a probabilistic atlas and the EM algorithm. *Medical Image Analysis*, 8(3):255–265.

Lu, Y., Radau, P., Connelly, K., Dick, A., and Wright, G. (2009). Automatic image-driven segmentation of left ventricle in cardiac cine MRI. *The MIDAS Journal – Cardiac MR Left Ventricle Segmentation Challenge*.

Lu, Y.-L., Connelly, K. A., Dick, A. J., Wright, G. A., and Radau, P. E. (2013). Automatic functional analysis of left ventricle in cardiac cine MRI. *Quantitative Imaging in Medicine and Surgery*, 3(4):200.

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

Margeta, J., Geremia, E., Criminisi, A., and Ayache, N. (2012). Layered spatio-temporal forests for left ventricle segmentation from 4D cardiac MRI data. In *International Workshop on Statistical Atlases and Computational Models of the Heart*, pages 109–119. Springer.

Matthews, I. and Baker, S. (2004). Active appearance models revisited. *International Journal of Computer Vision*, 60(2):135–164.

McInerney, T. and Terzopoulos, D. (1996). Deformable models in medical image analysis: A survey. *Medical Image Analysis*, 1(2):91–108.

Metaxas, D. and Terzopoulos, D. (1993). Shape and nonrigid motion estimation through physics-based synthesis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(6):580–591.

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

Mitchell, S. C., Lelieveldt, B. P., van der Geest, R. J., Bosch, H. G., Reiver, J., and Sonka, M. (2001). 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.

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

Nocedal, J. and Wright, S. (2006). *Numerical optimization*. Springer Science & Business Media.

Papademetris, X., Sinusas, A. J., Dione, D. P., Constable, R. T., and Duncan, J. S. (2002). Estimation of 3-D left ventricular deformation from medical images using biomechanical models. *IEEE Transactions on Medical Imaging*, 21(7):786–800.

Paragios, N. (2002). A variational approach for the segmentation of the left ventricle in cardiac image analysis. *International Journal of Computer Vision*, 50(3):345–362.

Petitjean, C. and Dacher, J.-N. (2011). A review of segmentation methods in short axis cardiac MR images. *Medical Image Analysis*, 15(2):169–184.

Radau, P., Lu, Y., Connelly, K., Paul, G., Dick, A. J., and Wright, G. A. (2009). Evaluation framework for algorithms segmenting short axis cardiac MRI. *The MIDAS Journal – Cardiac MR Left Ventricle Segmentation Challenge*.

Sanderson, C. and Curtin, R. (2016). Armadillo: A template-based C++ library for linear algebra. *Journal of Open Source Software*, 1:22. http://arma.sourceforge.net/armadillo_joss_2016.pdf.

Shabana, A. A. (2013). *Dynamics of multibody systems*. Cambridge University Press.

Song, S. M. and Leahy, R. M. (1991). Computation of 3-D velocity fields from 3-D cine CT images of a human heart. *IEEE Transactions on Medical Imaging*, 10(3):295–306.

Stegmann, M. B. and Pedersen, D. (2005). Bi-temporal 3D active appearance models with ap-

plications to unsupervised ejection fraction estimation. In *Medical Imaging*, pages 336–350. International Society for Optics and Photonics.

Suinesiaputra, A., Cowan, B. R., Al-Agamy, A. O., Elattar, M. A., Ayache, N., Fahmy, A. S., Khalifa, A. M., Medrano-Gracia, P., Jolly, M.-P., Kadish, A. H., et al. (2014). A collaborative resource to build consensus for automated left ventricular segmentation of cardiac MR images. *Medical Image Analysis*, 18(1):50–62.

Taylor, C. J. and Kriegman, D. J. (1994). Minimization on the lie group SO(3) and related manifolds. Technical Report 9405, Yale University.

Terzopoulos, D. (1986). Regularization of inverse visual problems involving discontinuities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(4):413–424.

Terzopoulos, D. and Metaxas, D. (1991). Dynamic 3D models with local and global deformations: Deformable superquadrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(7):703–714.

Terzopoulos, D. and Szeliski, R. (1993). Tracking with Kalman snakes. In Blake, A. and Yuille, A., editors, *Active Vision*, pages 3–20. MIT press.

Terzopoulos, D. and Witkin, A. (1988). Physically based models with rigid and deformable components. *IEEE Computer Graphics and Applications*, 8(6):41–51.

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

Wang, H. and Amini, A. A. (2012). Cardiac motion and deformation recovery from MRI: A review. *IEEE Transactions on Medical Imaging*, 31(2):487–503.

World Health Organization (2016). Cardiovascular diseases. Fact sheet online: http://www.who.int/mediacentre/factsheets/fs317/en/.

Xu, C. and Prince, J. L. (1997). Gradient vector flow: A new external force for snakes. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 66–71.

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

Zhu, S.-C. and Yuille, A. (1996). Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(9):884–900.