NATIONAL UNIVERSITY OF SINGAPORE

Left Ventricle Segmentation Using Data-driven Priors and Temporal Correlations

by
Jia Xiao

A thesis submitted in partial fulfillment for
the degree of Masters of Engineering

in the
Faculty of Engineering
Department of Electrical and Computer Engineering

December 2010

Abstract

Cardiac MRI has been widely used in the study of heart diseases and transplant rejections using small animal models. However, due to low image quality, quantitative analysis of the MRI data has to be performed through tedious manual segmentation. In this thesis, a novel approach based on data-driven priors and temporal correlations is proposed for the segmentation of left ventricle myocardium in cardiac MR images of native and transplanted rat hearts. To incorporate data-driven constraints into the segmentation, probabilistic maps generated based on prominent image features, i.e., corner points and scale-invariant edges, are used as priors for endocardium and epicardium segmentation, respectively. Non-rigid registration is performed to obtain the deformation fields, which are then used to compute the averaged probabilistic priors and feature spaces. Integrating data-driven priors and temporal correlations with intensity, texture, and edge information, a level set formulation is adopted for segmentation. The proposed algorithm was applied to 3D+t cardiac MR images from eight rat studies. Left ventricle endocardium and epicardium segmentation results obtained by the proposed method respectively achieve $87.1 \pm 2.61\%$ and $87.79 \pm 3.51\%$ average area similarity and $83.16 \pm 8.14\%$ and $91.19 \pm 2.78\%$ average shape similarity with respect to manual segmentations done by experts. With minimal user input, myocardium contours obtained by the proposed method exhibit excellent agreement with the gold stan-

dard and good temporal consistency. More importantly, it avoids inter- and intra-observer variations and makes accurate quantitative analysis of low-quality cardiac MR images possible.

Acknowledgements

This thesis would not have been successfully completed without the kind assistance and help of the following individuals.

First and foremost, I would like to express my deepest appreciation to my supervisors Associate Professor Ashraf Kassim and Assistant Professor Sun Ying for their unwavering guidance and support throughout the course of this research. I am grateful for their continual encouragement and advice that have made this project possible.

I would like to thank Dr. Yi-Jen L. Wu, the researcher from Pittsburgh NMR Center for Biomedical Research, USA, for her help and effort in providing the manual segmentation ground truth.

I would also like to thank Mr. Francis Hoon, the Laboratory Technologist of the Vision and Image Processing Laboratory, for his technical support and assistance.

Last but not least, I would like to extend my gratitude to my fellow lab mates for their help and enlightenment.

Contents

List of Publications

Xiao Jia, Chao Li, Ying Sun, Ashraf A. Kassim, Yijen L. Wu, T. Kevin Hitchens, and Chien Ho, "*A Data-driven Approach to Prior Extraction for Segmentation of Left Ventricle in Cardiac MR Images*", IEEE International Symposium on Biomedical Imaging(ISBI)'09, Boston, USA, June 2009.

Chao Li, **Xiao Jia**, and Ying Sun, "*Improved Semi-automated Segmentation of Cardiac CT and MR Images*", IEEE ISBI'09, Boston, USA, June 2009.

Xiao Jia, Ying Sun, Ashraf A. Kassim, Yijen L. Wu, T. Kevin Hitchens, and Chien Ho, "*Left Ventricle Segmentation in Cardiac MRI Using Data-driven Priors and Temporal Correlations*" [abstract], 13th Annual Society for Cardiovascular Magnetic Resonance (SCMR) Scientific Sessions, Phoenix, USA, January, 2010.

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Problem Statement

Small rodent animal models are widely used in evaluating pharmacological and surgical therapies for cardiovascular diseases. With the help of noninvasive imaging tools, like cardiac magnetic resonance imaging (MRI), in-vivo quantitative analysis of heart function of small animal models becomes possible in cardiac pathological studies and therapy evaluations [1, 2].

Reliable quantitative analysis of cardiac MRI data requires accurate segmentation of the left ventricle (LV) myocardium, which is tedious and time-consuming when performed manually. In addition to its high labor cost, manual segmentation also suffers from inter- and intra-observer variations. Therefore, it is desirable to

design an automated segmentation system which produces accurate and consistent segmentation results.

Automated segmentation of small animal MRI data is very challenging, and existing algorithms lack accuracy as well as robustness in solving such segmentation problems. Different from human hearts, rat hearts are small in size, therefore cardiac MR images acquired from rats normally have very limited spatial resolution and low signal-to-noise ratio (SNR). In allograft rejection studies [2], the transplanted rat heart is placed in recipient's abdomen and edges are not as well defined as is found when the native heart is surrounded by the lung. Moreover, turbulent blood flow often causes confusing edges in the LV cavity.

Although many approaches have been reported for the automated segmentation of human hearts [3], few methods have been proposed to segment small animal hearts. The STACS method proposed in [4] has been shown to produce relatively accurate segmentation results on short-axis cardiac MR images of a rat by combining region-based and edge-based information with an elliptical shape prior and contour smoothness constraint. Proposed in [5], a deformable elastic template has been utilized to segment left and right ventricles of mouse heart simultaneously in 3D cine MR images.

The above mentioned methods achieved acceptable segmentation in MR images of native rat heart, but they perform poorly on MRI data used in the study of

animal heart transplantation. To realize accurate automatic segmentation of the LV myocardium in MR images of both native and transplanted rat hearts, new approaches have to be explored.

1.2 Contributions

In this thesis, a novel method is proposed for the segmentation of LV myocardium in cardiac MRI for both native and transplanted rat hearts, incorporating data-driven priors as well as temporal correlations.

The extraction of prominent features and the generation of data-driven priors were originally introduced in our previous publication [6]. Derived from prominent features on individual images, the prior maps are representative of corresponding image data yet embedded with anatomical prior knowledge that is complementary to pixel-wise information, e.g., image intensity. Combining the prior maps and pixel-wise information, the proposed method achieves accurate and robust segmentation.

In addition to the data-driven priors, the segmentation results are further refined through the incorporation of temporal correlations. Though some research works have been done on temporally constrained segmentation, misleading point-to-point correspondence caused by inaccurate registration is still the major challenge yet

to be overcome. In the proposed approach, point-to-point correspondences for epicardium and endocardium segmentations are constructed separately through non-rigid registration. Utilizing the previously extracted features as prior knowledge, registration accuracy is enhanced significantly. With reliable frame-to-frame registration, not only image data of neighboring frames are incorporated into the segmentation, prior maps of neighboring frames are also utilized to provide complementary information that is absent in the image to be segmented.

Through accurate automatic segmentation, the proposed method enables efficient quantitative analysis of low quality rat MRI data and avoids inter- and intra-observer variations.

1.3 Thesis Organization

This thesis is organized as follows. A review of related works is presented in Chapter 2. Chapter 3 provides a detailed introduction on the proposed approach. Experimental results and performance evaluations are given in Chapter 4. In Chapter 5, the thesis is summarized and possible future work is discussed.

Chapter 2

Background and Previous Work

2.1 Segmentation

Medical image segmentation attracted enormous attention from the research community in the past few decades. Several approaches have been widely adopted in solving different segmentation problems, and some of the popular methods have been extensively developed recently. It is common that different approaches are used in conjunction for solving specific segmentation problems.

One important group of segmentation methods can be considered as pixel classification methods, including thresholding, classifiers, supervised or unsupervised clustering methods, and Markov random field (MRF) models [3].

Other techniques have also been developed, including artificial neural networks, atlas-based approaches, and deformable models. In the application of cardiac MRI segmentation, methods based on deformable models have been widely studied and adopted. A review of different approaches using deformable models is provided in this section.

In segmentation applications where the most discriminant features are intensity distribution patterns instead of pure intensity values, texture features are often extracted and utilized predominantly. Proposed by Rousson et al. in [7], an effective segmentation method based on texture information is introduced in Section 2.1.2.

Due to high noise level and complex anatomic structures, prior knowledge is often used in segmenting medical images. As a new type of prior, the data-driven prior, will be introduced in this thesis. A brief summary of previous work on incorporating prior knowledge into segmentation is provided at the end of this section.

2.1.1 Deformable Models

According to [3], deformable model based methods are defined as physically motivated, model-based techniques for delineating region boundaries by using closed parametric/non-parametric curves or surfaces that deform under the influence of internal and external forces. Internal forces are determined from the curve or

surface to make it smooth or close to a predefined appearance. External forces are normally computed from the image to deform the contour so that the object boundaries can be correctly delineated.

Deformable model based methods have the following advantages: 1) object boundaries are defined as closed parametric or non-parametric curves, and the final segmentation results can be deformed from an initial contour according to internal and external forces; and 2) by introducing the internal force, boundaries of segmented objects are smooth and can be biased towards different appearances, and this is particularly important because desired object boundaries do not have arbitrary appearances in most medical segmentation applications. There are also limitations of deformable based approaches [3]: an initial contour should be placed before the deformation, and in some cases, final outcomes are very sensitive to the initialization; and choosing appropriate parameters can also be time consuming.

2.1.1.1 Parametric Active Contours

Snakes

Initially introduced as "snakes" in [8], this classical active contour approach was effective in solving a wide range of segmentation problems. Through energy minimization, snakes evolve a deformable model based on image features.

Let us define a contour C parameterized by arc length s as

$$C(s) = \{(x(s), y(s)) : 0 \leq s \leq L\} \ \Re \longrightarrow \Omega, \tag{2.1}$$

where L denotes the length of the contour C and Ω denotes the entire domain of an image $I(x, y)$. An energy function $E(C)$ can be defined on the contour such as:

$$E(C) = E_{\text{int}} + E_{\text{ext}}, \tag{2.2}$$

where E_{int} and E_{ext} denote the internal and external energies, respectively. The internal energy function determines the regularity (or the smoothness) of the contour. A common definition of the internal energy is a quadratic function:

$$E_{\text{int}} = \int_0^1 \alpha |C'(s)|^2 + \beta |C''(s)|^2 ds, \tag{2.3}$$

where α controls the tension of the contour, and β controls the rigidity of the contour. The external energy term that determines the criteria of contour evolution depending on the image $I(x, y)$ can be defined as

$$E_{\text{ext}} = \int_0^1 E_{\text{img}}\left(C(s)\right) ds, \tag{2.4}$$

where $E_{\text{img}}(x, y)$ denotes a scalar function defined on the image plane, so that local minimum of E_{img} attracts the snakes to edges. A common example of the edge attraction function is a function of the image gradient given by

$$E_{\text{img}}(x, y) = \frac{1}{\lambda \left|\nabla G_\sigma * I(x, y)\right|}, \tag{2.5}$$

where G denotes a Gaussian smoothing filter with standard deviation σ, λ is the suitable constant chosen and $*$ is the convolution operator. Solving the problem of snakes is to find the contour C that minimizes the total energy term E with the given set of weights α and β.

Classic snakes suffer from two major limitations: 1) initial contours have to be sufficiently close to the correct object boundaries to provide accurate segmentation. However, without prior knowledge, it is impossible for most segmentation applications to initialize contours close to object boundaries; 2) classic snakes are not capable of detecting more than one objects simultaneously, as it maintains the same topology during contour evolution.

Gradient Vector Flow

To overcome the problem that classic snakes encounter in segmenting objects with concave boundary regions [9], gradient vector flow (GVF) was introduced in [10] as an external force. It is a 2D vector field $V(s) = [u(s), v(s)]$ that minimizes the following objective function

$$E = \int\int \mu(u_x^2 + u_y^2 + v_x^2 + v_y^2) + |\nabla f|^2 |V - \nabla f|^2 dxdy, \tag{2.6}$$

where u_x, u_y, v_x and v_y are the spatial derivatives of the field, μ is the blending parameter, and ∇f is the gradient of the edge map which is defined as the negative external force, i.e., $f = -E_{\text{ext}}$. The objective function is composed of two terms:

the regularization term and the data driven term. The data-driven term dominates this function in the object boundaries (i.e., $|\nabla f|$ is large), while the regularization term dictates the function in areas where the intensity is constant (i.e., $|\nabla f|$ tends to zero). The GVF is obtained by solving the following Euler equations by using calculus of variations, and the normalized GVF is used as the static external force of the snake:

$$\mu\nabla^2 u - (u - f_x)(f_x^2 + f_y^2) = 0 \tag{2.7}$$

$$\mu\nabla^2 v - (v - f_y)(f_x^2 + f_y^2) = 0 \tag{2.8}$$

where ∇^2 is the Laplacian operator.

Although GVF solves the problem associated with concave boundaries, it has its own limitation caused by the diffusion of flow information: GVF creates similar flow for strong and weak edges, which can be considered a drawback in some applications.

2.1.1.2 Geometric Active Contours

Extended from classic snakes, the geometric active contour (GAC) model introduced in [11] overcomes some snakes' limitations. The model is given by

$$E(C) = \int_0^1 f(|\nabla I(C(s))|)|C_s|ds \tag{2.9}$$

$$= \int_o^{L(C)} f(|\nabla I(C(s))|)ds, \tag{2.10}$$

where the function f is the edge detecting function defined in (2.5), ds is the Euclidean element of length and $L(C)$ is the Euclidean length of the curve C defined by

$$L(C) = \int_0^1 |C_s| ds = \int_0^{L(C)} ds. \tag{2.11}$$

Though some short comings of classic snakes are overcome, the GAC model still suffers from one major limitation: the curve can only be evolved towards one direction (inwards or outwards). As a result, the initial curve has to be placed completely inside or outside of the object of interest.

Level Sets

Level sets are a class of deformable models that have been studied most intensively in the area of medical image segmentation. Initially proposed in [12], Osher and Sethian represent a contour implicitly via 2D Lipchitz continuous function $\phi(x, y) : \Omega \rightarrow \Re$, defined in the image plane. The function $\phi(x, y)$ is called level set function, and a particular level, usually the zero level of $\phi(x, y)$ is defined as the contour, such as

$$C = \{(x, y) : \phi(x, y) = 0\}, \ \forall (x, y), \in \Omega \tag{2.12}$$

where Ω denotes the entire image plane.

As the level set function $\phi(x, y)$ evolves from its initial stage, the corresponding set of contours C, i.e., the red contours in Fig. 2.1, propagate. With this definition,

the evolution of the contour is equivalent to the evolution of the level set function, i.e., $\partial C/\partial t = \partial\phi(x,y)/\partial t$. The advantage of using the zero level is that a contour can be defined as the border between a positive area and a negative area, so the contours can be identified by just checking the sign of $\phi(x,y)$. The initial level set function $\phi(x,y)$: $\Omega \rightarrow \Re$ may be given by the signed distance from the initial contour such as,

$$\begin{aligned} \phi_0(x,y) &\equiv \{\phi(x,y) : t = 0\} \\ &= \pm D\left((x,y), N_{x,y}(C_0)\right) \end{aligned} \qquad \forall (x,y) \in \Omega, \tag{2.13}$$

where $\pm D(a,b)$ denotes the signed distance between a and b, and $N_{x,y}(C_0)$ denotes the nearest neighboring pixel on initial contours $C_0 \equiv C(t=0)$ from (x,y). The initial level set function is zero at the initial contour points given by

$$\phi_0(x,y) = 0,\ \forall (x,y) \in C_0. \tag{2.14}$$

The deformation of the contour is generally represented in a numerical form as a partial differential equation (PDE). A formulation of contour evolution using the magnitude of the gradient of $\phi(x,y)$ was initially proposed by Osher and Sethian as

$$\frac{\partial\phi(x,y)}{\partial t} = |\nabla\phi(x,y)|(\nu + \epsilon\kappa(\phi(x,y))), \tag{2.15}$$

where ν denotes a constant speed term to push or pull the contour, $\kappa(\cdot) : \Omega \rightarrow \Re$ denotes the mean curvature of the level set function $\phi(x,y)$ given by

$$\kappa(\phi(x,y)) = \mathrm{div}\left(\frac{\nabla\phi}{\|\nabla\phi\|}\right) = \frac{\phi_{xx}\phi_y^2 - 2\phi_x\phi_y\phi_{xy} + \phi_{yy}\phi_x^2}{(\phi_x^2 + \phi_y^2)^{\frac{3}{2}}}, \tag{2.16}$$

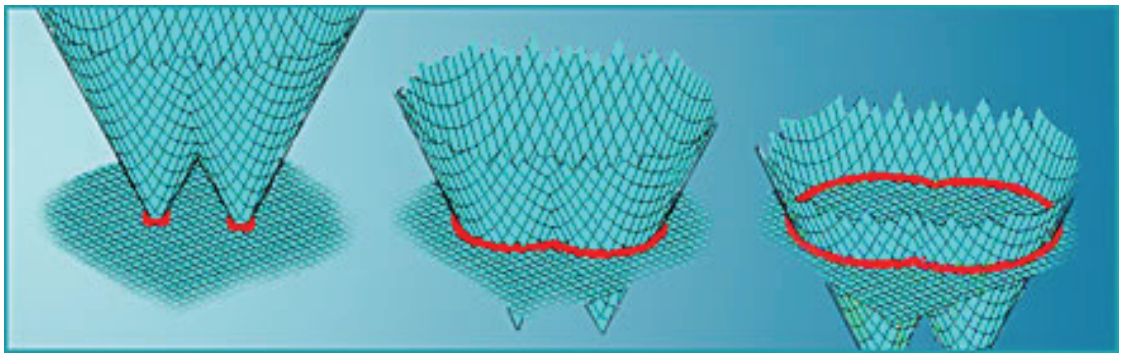

Figure 2.1: Evolution of level set function

where ϕ_x and ϕ_{xx} denote the first and second order partial derivatives of $\phi(x, y)$ with respect to x respectively, and ϕ_y and ϕ_{yy} denote the same with respect to y. The role of the curvature term is to control the regularity of the contours as the internal energy term E_{int} does in the classical snake model, and ϵ controls the balance between the regularity and robustness of the contour evolution.

An outstanding characteristic of the level set method is that contours can split or merge as the topology of the level set function changes. Therefore, level set methods can detect more than one object simultaneously, and multiple initial contours can be placed. Figure 2.1 shows how the initial separated contours merge as the topology of level set function varies. This flexibility and convenience provide a means for an automated segmentation by using a predefined set of initial contours.

Another advantage of the level set method is the possibility of curve evolution in dimensions higher than two. The mean curvature of the level set function (see (2.16)) can be easily extended to deal with higher dimensions. This is very useful

in propagating a surface to segment volume data.

The computational cost of level set methods is high because the computation should be done on the same dimension as image plane Ω. Thus, the convergence speed is slower than other segmentation methods, particularly local filtering based methods. The high computational cost can be compensated by using multiple initial contours. The use of multiple initial contours increases the convergence speed by cooperating with neighbor contours quickly. Level set methods with faster convergence, called fast marching methods, have been studied intensively for the last decade [13].

However, in traditional level set methods, the level set function ϕ can develop shocks, very sharp and/or flat shape during the evolution, which makes further computation highly inaccurate. To avoid these problems, a common numerical scheme is to initialize the function ϕ as a signed distance function before the evolution, and then "reshape" (or "re-initialize") the function ϕ to be a signed distance function periodically during the evolution. Indeed, the re-initialization process is crucial and cannot be avoided in using traditional level set methods.

Variational Level Set

To realize the level set method without re-initialization, a novel way of level set formulation which is easily implemented by simple finite difference scheme has been proposed by Li et al. [14].

Re-initialization in traditional level set methods has been extensively used as a numerical remedy for maintaining stable curve evolution and ensuring desirable results. From the practical viewpoint, the re-initialization process can be quite complicated, expensive, and have subtle side effects. It is crucial to keep the evolving level set function as an approximate signed distance function during the evolution, especially in a neighborhood around the zero level set. It is well known that a signed distance function must satisfy a desirable property of $|\nabla\phi| = 1$. Conversely, any function ϕ satisfying $|\nabla\phi| = 1$ is the signed distance function plus a constant [15]. A metric to characterize how close a function ϕ is to a signed distance function in $\Omega \subset \Re^2$ is defined by:

$$P(\phi) = \int_\Omega \frac{1}{2} \left(|\nabla\phi| - 1\right)^2 dxdy. \tag{2.17}$$

This metric plays a key role in the variational formulation. The variational formulation is defined as:

$$E(\phi) = \mu P(\phi) + E_m(\phi), \tag{2.18}$$

where $\mu > 0$ is the parameter controlling the effect of penalizing the deviation of ϕ from a signed distance function, and $E_m(\phi)$ is a certain energy that would drive the motion of the zero level curve of ϕ.

The gradient flow that minimizes the functional E is defined as:

$$\frac{\partial\phi}{\partial t} = -\frac{\partial E}{\partial\phi}. \tag{2.19}$$

For a particular functional $E(\phi)$ defined explicitly in terms of ϕ, the Gateaux derivative can be computed and expressed in terms of the function ϕ and its derivatives [16].

The variational formulation described in (2.18) is applied to active contours for image segmentation, and the zero level set curve of ϕ can evolve to the desired features in the image. The energy E_m is defined as a functional that depends on image data, and it is named as external energy. Accordingly, the energy $P(\phi)$ is called the internal energy of the function ϕ.

During the evolution of ϕ according to the gradient flow in (2.19) that minimizes the functional (2.18), the zero level curve is moved by the external energy term E_m. Meanwhile, due to the penalizing effect of the internal energy, the evolving function ϕ is automatically maintained as an approximate signed distance function during the evolution according to the evolution in (2.19). As a result, the re-initialization procedure is completely eliminated in the above formulation.

2.1.2 Texture Segmentation

In an attempt to extract texture features to assist segmentation, structure tensor based methods were first introduced by Bigun et al. [17]. To overcome the problem of dislocated edges in feature channels caused by Gaussian smoothing, Rousson et al. [7] combine the nonlinear structure tensor proposed in [18] and vector-valued

diffusion introduced in [19] to obtain a diffusion based feature space. Applying the variational framework proposed in their earlier publication [20] on the extracted feature space, Rousson et al. implement maximum a posteriori segmentation by energy minimization.

Diffused Feature Space

For a given image I, the structure tensor matrix is defined as:

$$u = \nabla I \nabla I^T = \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}, \tag{2.20}$$

where I_x and I_y are gradients of image I along the x and y direction respectively. To reduce noise while preserving edges, nonlinear diffusion (based on Perona and Malik [21]) is applied. The diffusion equation is

$$\partial_t u = \operatorname{div}\left(g\left(|\nabla u|\right)\nabla u\right), \tag{2.21}$$

where g is a decreasing function. For vector-valued data:

$$\partial_t u_i = \operatorname{div}\left(g\left(\sum_{k=1}^{N} |\nabla u_k|^2\right)\nabla u_i\right), \ \forall i, \tag{2.22}$$

where u_i is an evolving vector channel and N the number of channels. All channels are coupled by a joint diffusivity, so an edge in one channel inhibits smoothing in the others. The diffusivity function g is defined as:

$$g\left(|\nabla u|\right) = \frac{1}{|\nabla u| + \epsilon}, \tag{2.23}$$

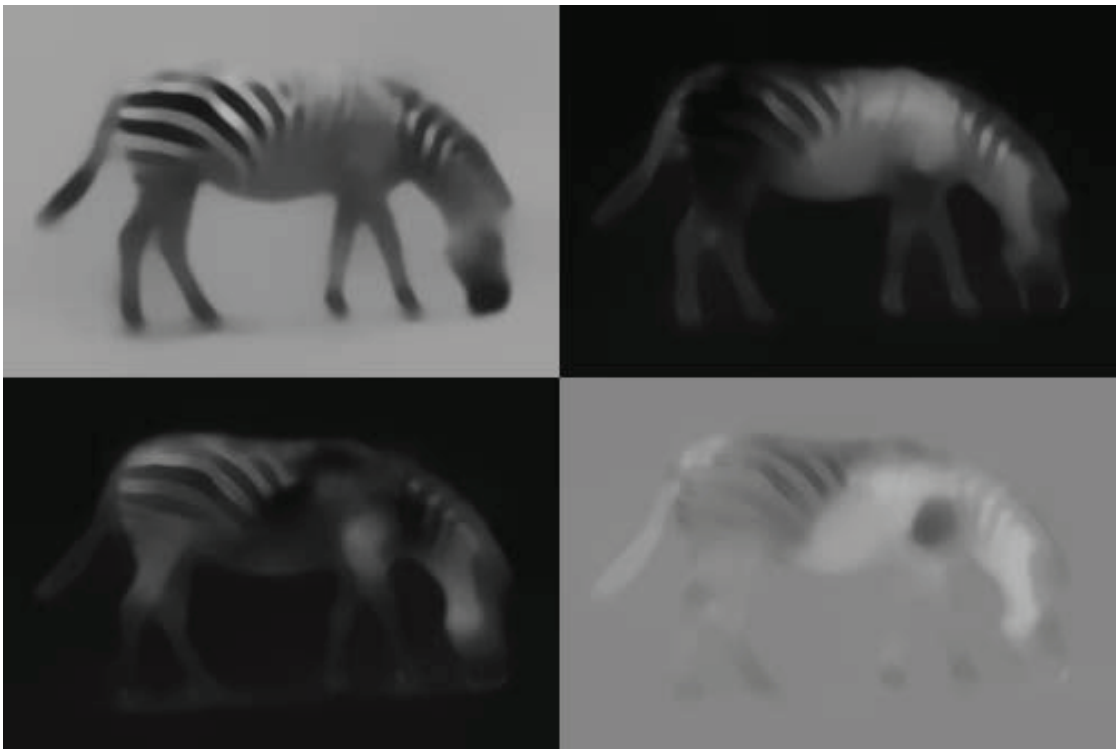

Figure 2.2: Feature channels $(u_1, \ldots, u_4)$ obtained by smoothing $I, I_x^2, I_y^2, I_x I_y$ from left to right and top to bottom.[1]

where ϵ is a small positive constant added to avoid numerical problems.

By applying (2.22) with initial conditions $u_1 = I$, $u_2 = I_x^2$, $u_3 = I_y^2$,$u_4 = I_x I_y$ and the diffusivity function $g(s) = 1/s$, features can be extracted as illustrated in Fig. 2.2.

Adaptive Segmentation

The image segmentation can be found by maximizing a posteriori partitioning probability $p\left(\mathcal{P}(\Omega)|I\right)$ where $\mathcal{P}(\Omega) = \{\Omega_1, \Omega_2\}$ is a partition of the image domain

[1]Figure taken from "Active unsupervised texture segmentation on a diffusion based feature space" [7]

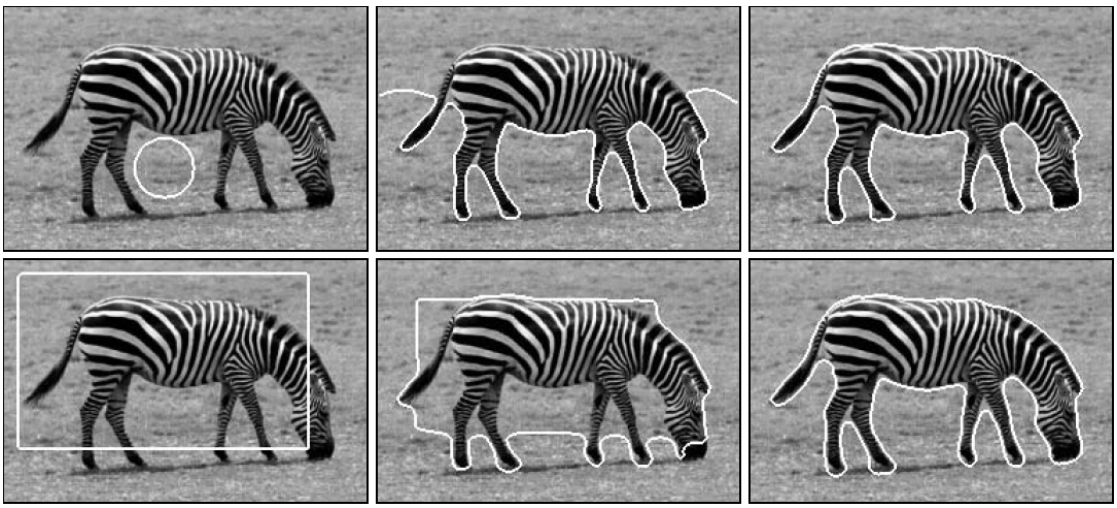

Figure 2.3: Texture segmentation results[1]

Ω. Instead of using original image I, the segmentation energy functional is defined based on the vector-valued image $u = (u_1, \ldots, u_4)$.

Let $p_1(u(x))$ and $p_2(u(x))$ be the probability density function for the value $u(x)$ to be in Ω_1 and Ω_2, respectively. With $\partial\Omega$ being the boundary between Ω_1 and Ω_2, the segmentation is found by minimizing the energy

$$E(\Omega_1, \Omega_2) = -\int_{\Omega_1} \log p_1(u(x))\, dx - \int_{\Omega_2} \log p_2(u(x))\, dx. \tag{2.24}$$

To model the statistics of each region, a general Gaussian approximation is used for all four channels. Let $\{\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1\}$ and $\{\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2\}$ be the vector's means and covariance matrices of the Gaussian approximation in Ω_1 and Ω_2. The probability of $u(x)$ to be in Ω_i is:

$$p_i(u(x)) = \frac{1}{(2\pi)^2 |\boldsymbol{\Sigma}_i|^{1/2}} e^{-\frac{1}{2}(u(x)-\boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1}(u(x)-\boldsymbol{\mu}_i)}. \tag{2.25}$$

Here information in each channel is assumed to be uncorrelated, and the probability

density function (pdf) $p_i\left(u(x)\right)$ can be estimated using the joint density probability of each component:

$$p_i\left(u(x)\right) = \prod_{k=1}^{4} p_{k,i}\left(u_k(x)\right). \tag{2.26}$$

Let $H_\epsilon(z)$ and $\delta_\epsilon(z)$ be regularized versions of the Heaviside and Dirac functions. Adding a regularization constraint on the length of $\partial\Omega$, the energy (2.24) can be minimized with respect to the whole set of parameters $\{\partial\Omega, \boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2\}$ using the following evolution equation (see [20] for details):

$$\phi_t(x) = \delta_\epsilon\left(\phi(x)\right)\left(\nu \mathrm{div}\left(\frac{\nabla\phi}{|\nabla\phi|}\right) + \log\frac{p_1\left(u(x)\right)}{p_2\left(u(x)\right)}\right), \tag{2.27}$$

while the Gaussian parameters are updated at each iteration as:

$$\begin{cases} \mu_i(\phi) = \int_\Omega u(x)\chi_i[\phi(x)]dx / \int_\Omega \chi_i[\phi(x)]dx \\ \Sigma_i(\phi) = \int_\Omega \left(\mu_i - u(x)\right)\left(\mu_i - u(x)\right)^T \chi_i[\phi(x)]dx / \int_\Omega \chi_i[\phi(x)]dx \end{cases} \tag{2.28}$$

where $\chi_1(z) = H_\epsilon(z)$ and $\chi_2(z) = 1 - H_\epsilon(z)$.

2.1.3 Incorporating Priors

For the integration of prior knowledge, some segmentation methods [5] based on deformable templates were proposed. In [5], obtained from manual segmentation of cardiac ventricles in a reference data set, a topological and geometric model of the ventricles is used as a deformable elastic template to simultaneously segment both left and right ventricles in 3D cine MR images of rat hearts. Although this method

utilizes both topological and geometric characteristics of ventricles, the deformable template obtained from one reference dataset (or even several reference datasets) is not representative enough, and it only captures very limited topological variations.

Besides deformable templates, many existing segmentation methods [4, 22, 23] use shape priors. In [4] and [22], elliptical shape priors were used for both endocardium and epicardium segmentation by including a shape prior term in the energy functional. Learned from training samples, probabilistic shape priors proposed in [23] constrain the segmentation by optimizing a statistical metric between the evolving contour and the prior model.

Incorporating shape priors, the above mentioned methods significantly enhanced the segmentation robustness. However, due to the fact that these shape priors are not representative of any particular image, the effect of incorporating such shape priors is just adding another regulatory force that prevents the contours from having very unlikely shapes. Moreover, obtaining a large number of manually processed training samples can be very time consuming. For these reasons, it is desirable to extract representative priors from the image itself without the training process. Ideally, the extracted priors should carry useful information about the image structure which can be utilized as a piece of reliable prior knowledge to guide the contour deformation towards correct segmentation.

2.2 Registration

Similar to segmentation, image registration is also a fundamental image processing problem that has been extensively studied [24, 25, 26]. The registration process can be simply interpreted as a process of aligning or matching two or more images having similar contents. Under the context of multi-frame segmentation problem, registration provides correspondence from one image to another, which is useful when complementary image information appears on different frames.

In general, available registration approaches can be grouped into two classes: feature-based and intensity-based methods. To register images based on features, a preprocessing step is required to extract appropriate features, such as salient points or edges. By matching corresponding features, the deformation field can be calculated by interpolation. The intensity-based methods measure similarity using pixel intensity values directly. In this thesis, we only focus on intensity-based registration methods, as automated accurate detection of unique features is too difficult to achieve on rat cardiac MR images.

Depending on the application, similarity measures can be different. Sum-of-absolute-difference (SAD) and sum-of-square-differences (SSD) similarity measures have been compared in [27] for the registration of cardiac positron emission tomography (PET) images. The SAD and SSD similarity measures assume constant brightness for corresponding pixels, and therefore, are mostly used in intra-modality image

registration. Another group of similarity measures calculate cross-correlation (CC). CC is an optimal measure for registration in the case of linear relationship between the intensity values in the image to be registered as it can compensate difference in gain and bias. For different imaging modalities, similarity measures based on joint entropy, mutual information, and normalized mutual information generally result in better registration [28, 29].

Within the intensity-based class of registration approaches, one of the well-known methods uses the concept of diffusion to perform image-to-image registration based on the optical flow [30]. Another class of methods named free form deformation (FFD) [26] calculate the transformation using a set of sparse spaced control points, which are not linked to any specific image features, and finding the extreme of the similarity measure defined in the neighborhood at the control points. After interpolated from the displacement of sparse control points according to certain smooth constraints, the deformation field is calculated. PDEs are also used to model the deformation by physical analogies [24]. Some Markov random field (MRF) based registration methods are also proposed in recent publications [25]. The B-spline based FFD is discussed in detail in the following sub-section.

2.2.1 B-spline Based Free Form Deformation

Initially proposed in [26], B-spline based FFD has been used in the application of 3D breast MR image registration. The deformation model consists of two parts: global and local transformation.

Global Motion Model

A rigid transformation which is parameterized by 6 degrees of freedom (describing rotations and translations) has been used to model the global motion. In 3-D, an affine transformation can be used to describe the rigid transformation:

$$T_{\text{global}}(x, y, z) = \begin{pmatrix} \theta_{11} & \theta_{12} & \theta_{13} \\ \theta_{21} & \theta_{22} & \theta_{23} \\ \theta_{31} & \theta_{32} & \theta_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} \theta_{14} \\ \theta_{24} \\ \theta_{34} \end{pmatrix}, \tag{2.29}$$

where the coefficients Θ parameterize the 12 degrees of freedom of the transformation.

Local Motion Model

Affine transformation only captures the global motion, therefore an additional transformation is required to model the local deformation. In medical images acquired at different time instances, local transformation can vary significantly. Therefore, parameterized transformation is not capable of modeling it. Different from parameterized models, B-spline based FFD models deform an object by

manipulating an underlying mesh of control points. The resulting deformation controls the shape of the 3-D object and produces a smooth and C^2 continuous transformation.

To define a spline-based FFD, we denote the domain of the image volume as $\Omega = \{(x, y, z)|0 \leq x < X, 0 \leq y < Y, 0 \leq z < Z\}$. Let Φ denote a $n_x \times n_y \times n_z$ mesh of control points $\phi_{i,j,k}$ with uniform spacing δ. Then the FFD can be written as the 3-D tensor product of the familiar 1-D cubic B-splines:

$$T_{\text{local}}(x, y, z) = \sum_{l=0}^{3} \sum_{m=0}^{3} \sum_{n=0}^{3} B_t(u) B_m(v) B_n(w) \phi_{i+l,j+m,k+n}, \tag{2.30}$$

where $i = \lfloor x/n_x \rfloor - 1$,$j = \lfloor y/n_y \rfloor - 1$, $k = \lfloor z/n_z \rfloor - 1$, $u = x/n_x - \lfloor x/n_x \rfloor - 1$,$v = y/n_y - \lfloor y/n_y \rfloor - 1$, $w = z/n_z - \lfloor z/n_z \rfloor - 1$, and B_l represents the l^{th} basis function of the B-spline

$$\begin{aligned} B_0(u) &= (1-u)^3/6, \\ B_1(u) &= (3u^3 - 6u^2 + 4)/6, \\ B_2(u) &= (-3u^3 + 3u^2 + 3u + 1)/6, \\ B_3(u) &= u^3/6. \end{aligned} \tag{2.31}$$

To achieve the best compromise between the degree of nonrigid deformation and associated computational cost, a hierarchical multi-resolution approach is adopted. Let $\Phi^1, \ldots, \Phi^L$ denote a hierarchy of control point meshes at different resolutions in a coarse to fine fashion. Each control mesh Φ^l and the associated spline-based FFD defines a local transformation T^l_{local} at each level of resolution and their sum

defines the local transformation T_{local}:

$$T_{\text{local}}(x, y, z) = \sum_{l=1}^{L} T_{\text{local}}^{l}(x, y, z). \tag{2.32}$$

Although the local transformation can vary significantly from region to region, the deformation should be characterized by a smooth transformation locally. The penalty term introduced to regularize the smoothness of the deformation is defined as:

$$C_{\text{smooth}} = \frac{1}{V} \int_0^X \int_0^Y \int_0^Z \left[\left(\frac{\partial^2 T}{\partial x^2} \right)^2 + \left(\frac{\partial^2 T}{\partial y^2} \right)^2 + \left(\frac{\partial^2 T}{\partial z^2} \right)^2 + 2 \left(\frac{\partial^2 T}{\partial xy} \right)^2 + 2 \left(\frac{\partial^2 T}{\partial xz} \right)^2 + 2 \left(\frac{\partial^2 T}{\partial yz} \right)^2 \right] dxdydz, \tag{2.33}$$

where V denotes the volume of the image domain.

Normalized Mutual Information

Mutual information is based on the concept of information theory and expresses the amount of information that one image A contains about a second image B

$$C_{\text{similarity}}(A, B) = H(A) + H(B) - H(A, B), \tag{2.34}$$

where $H(A)$, $H(B)$ denote the marginal entropies of A, B and $H(A, B)$ denotes their joint entropy, which is calculated from the joint histogram of A and B. If both images are aligned, the mutual information is maximized. To avoid any dependency on the amount of image overlap, normalized mutual information (NMI) can be used as a measure of image alignment. According to [26], the image similarity is defined

based on one form of NMI:

$$C_{\text{similarity}}(A, B) = \frac{H(A) + H(B)}{H(A, B)}. \tag{2.35}$$

Optimization

To obtain the optimal transformation, a cost function associated with the global transformation parameters Θ, as well as the local transformation parameters Φ is minimized. Combining the smoothness penalty described in (2.33) and the similarity measure in (2.35), the function that needs to be minimized is:

$$C(\Theta, \Phi) = -C_{\text{similarity}}\left(I(t_0), T\left(I(t)\right)\right) + \lambda C_{\text{smooth}}(T), \tag{2.36}$$

where λ is the weighting parameter which defines the tradeoff between the alignment of the two image volumes and the smoothness of the transformation.

2.3 Joint Registration & Segmentation

As two of the most important fundamental image processing problems, image segmentation and registration have been studied separately for decades. Recent development in the image processing research community shows a trend of integrating segmentation and registration, and proposed methods are sometimes known as "Segistration" [31]. In general, joint segmentation and registration methods can be divided into two groups depending on the availability of initial segmentation.

Proposed in [31], using the reference with initial segmentation as an atlas image, target image can be segmented by the "Registration+Segmentation" model. Let I_1 be the atlas image containing the atlas shape $\bar{C}$, I_2 the target image that needs to be segmented, and v be the deformation field from I_2 to I_1, i.e., the transformation is centered in I_2, defining the non-rigid deformation between the two images. Final segmentation in I_2 can be obtained by minimizing the following energy:

$$E(v, \tilde{C}) = Seg(I_2, \tilde{C}) + \text{dist}(v(\bar{C}), \tilde{C}) + Reg(I_1, I_2, v), \tag{2.37}$$

where $\tilde{C}$ is the boundary contour of the desired anatomical shape in I_2. The first and last term denote the segmentation and registration functional, respectively. The second term measures the distance between the transformed atlas $v(\bar{C})$ and the current segmentation $\tilde{C}$ in the target image.

Another class of approaches requires no initial segmentation (or atlas image). The aim of this case is to find a displacement field for the registration and a segmentation of objects in both images. Actually a segmentation of the template is computed and then carried out over to the reference by the computed displacement field. In this process, registration and segmentation interact sequentially: in each iteration, the segmentation uses feedback from the last registration step and vice versa. One representative approach of this class is proposed by Unal and Slabaugh in [32].

Chapter 3

Proposed Method

3.1 The Cine MRI

The cine MRI data that primarily used in the experiments consists of short axis (SA) images, which are cross-sectional images of rat hearts transversal to its major axis. All MRI datasets are in 4D; for each time instant, one set of volumetric images is acquired.

Image acquisition is done by scanning multiple successive slices at different times within the cardiac cycle. The acquisition is triggered by an electrocardiography (ECG) signal, and several cycles of the heart must be acquired. Several cardiac images corresponding to the same cardiac phase can be averaged to improve the PSNR. Fig. 3.1 illustrates the image acquisition process.

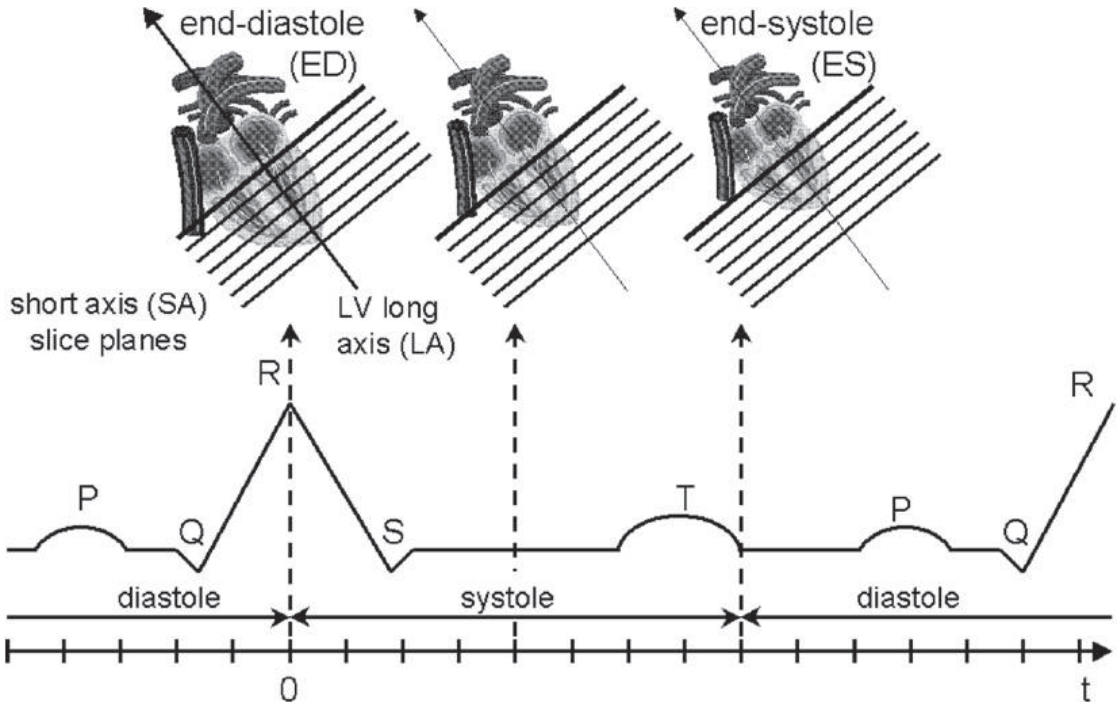

Figure 3.1: Illustration of MRI acquisition[2]

In Figs. 3.2 and 3.3, down-sampled version of datasets acquired from native and transplanted rats are illustrated. The actual dataset normally consists of 10 slices (each column shown in Figs. 3.2 and 3.3 illustrates a set of volumetric images scanned at a particular instance), and for each slice of images, about 10 frame of images were acquired (each row shown in Figs. 3.2 and 3.3 illustrates a set of images scanned from the same short axis plane at different phases of a cardiac cycle).

Cine images of native and heterotopic transplanted hearts are shown in Fig. 3.4, A and B are short-axis images of a native rat heart with a bright blood pulse sequence

[2]Figure taken from "A review of cardiac image registration methods" [33]

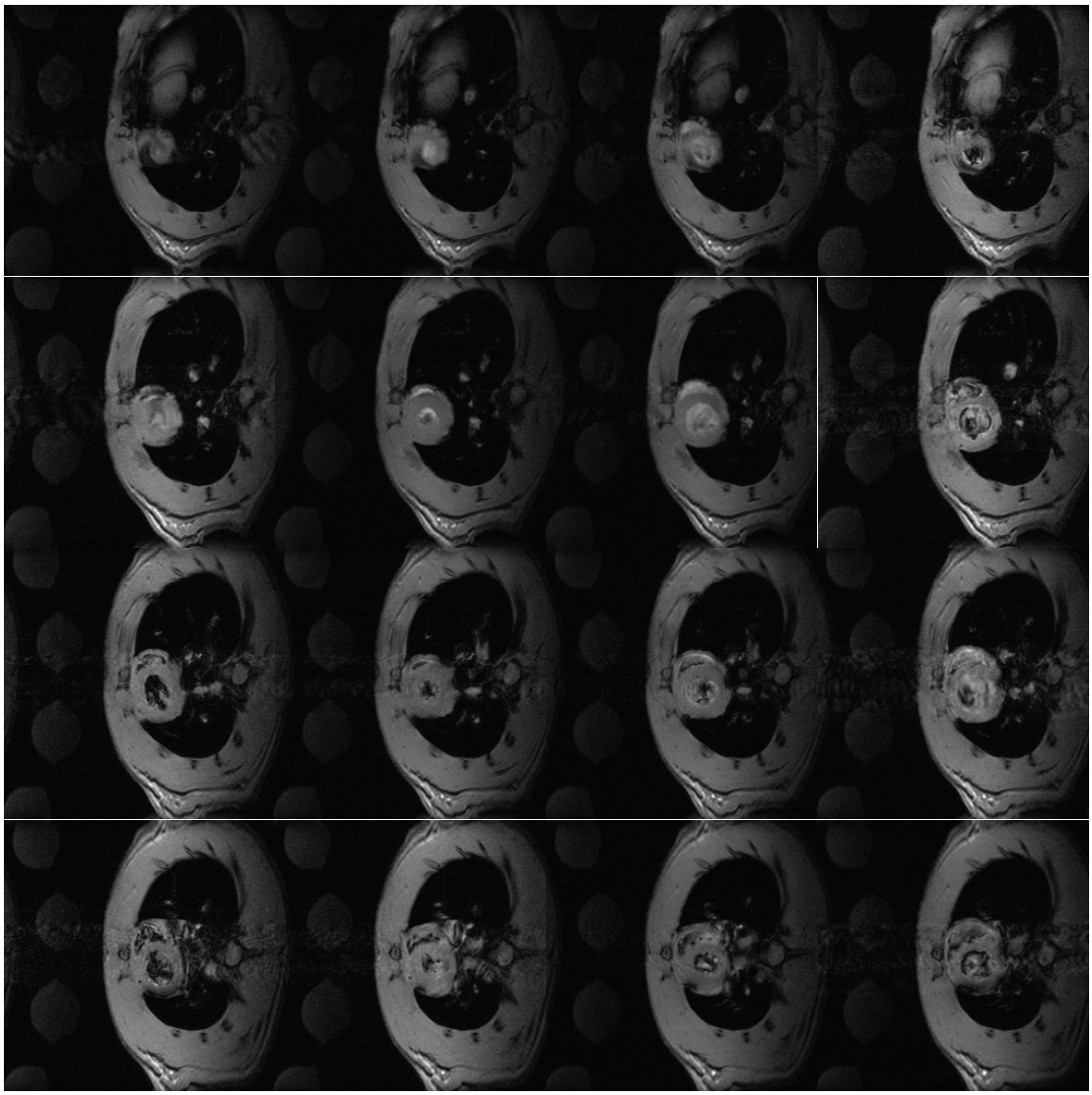

Figure 3.2: Illustration of MRI data of native hearts

(FLASH), C and D are short-axis images of a native rat heart with a black-blood spin-echo pulse sequence, E and F are short-axis images of a transplanted rat heart in the abdomen with a black-blood spin-echo pulse sequence. The left column (A, C, E) are images acquired at the end-diastole (ED), whereas the right column (B, D, F) are acquired at the end-systole (ES).

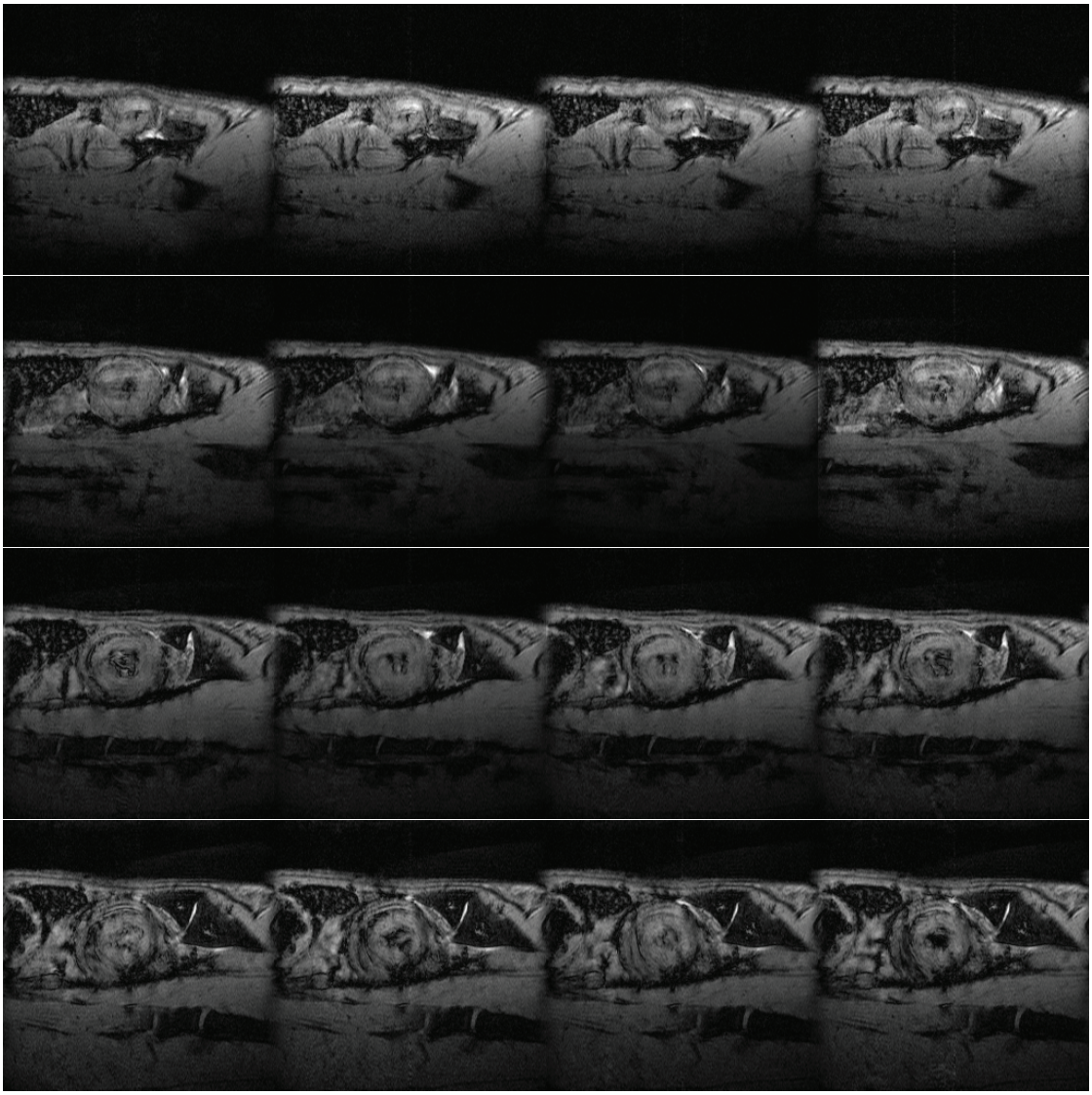

Figure 3.3: Illustration of MRI data of transplanted hearts

3.2 Slice-by-slice Segmentation

For a given 3D+t MRI dataset, frames in the same slice are acquired at different time instances, capturing the myocardial motion in a cardiac cycle from ED to ES and back to ED. Due to the myocardium wall motion, some prominent image fea-

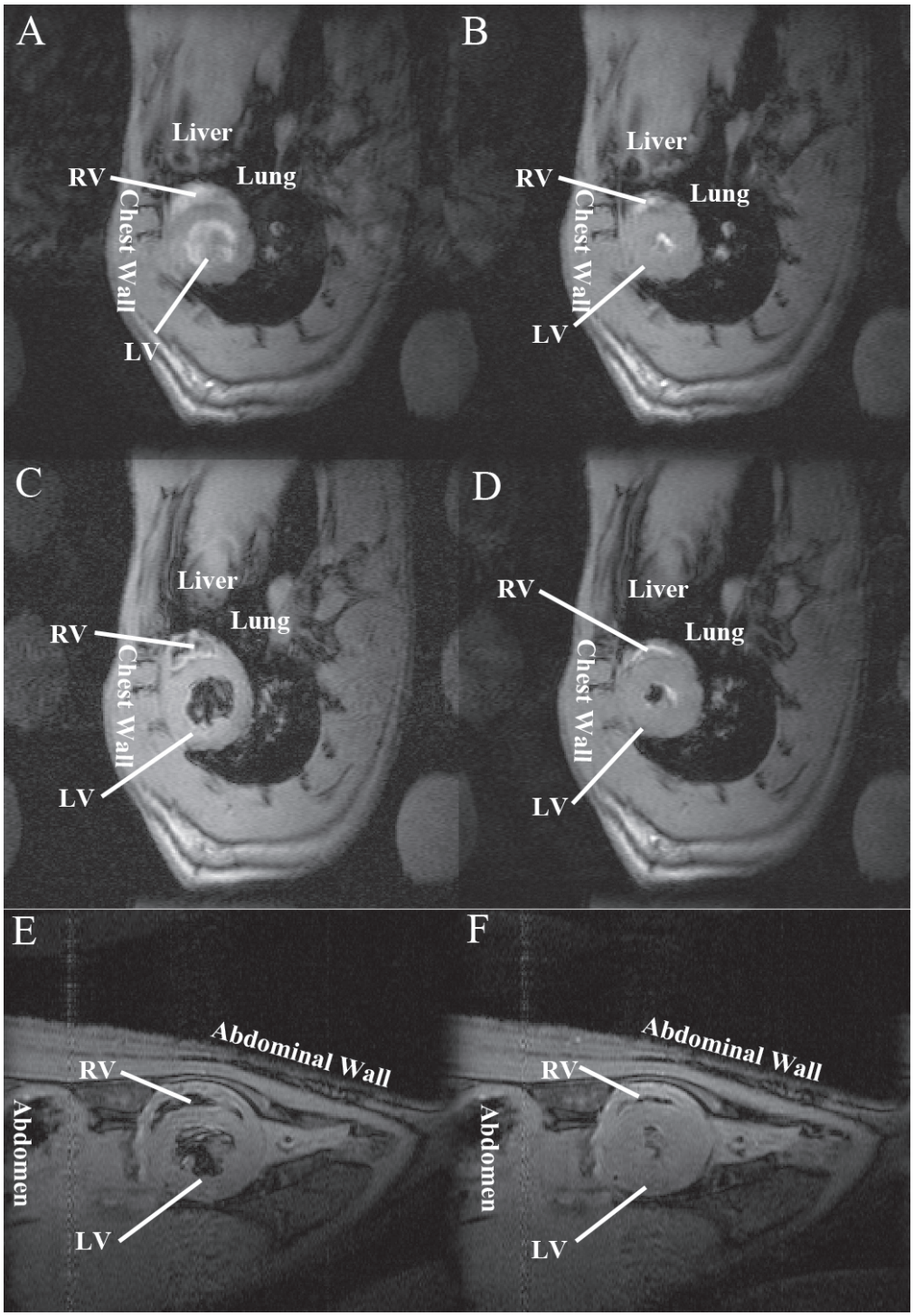

Figure 3.4: Cine imaging for native and heterotopic transplanted hearts

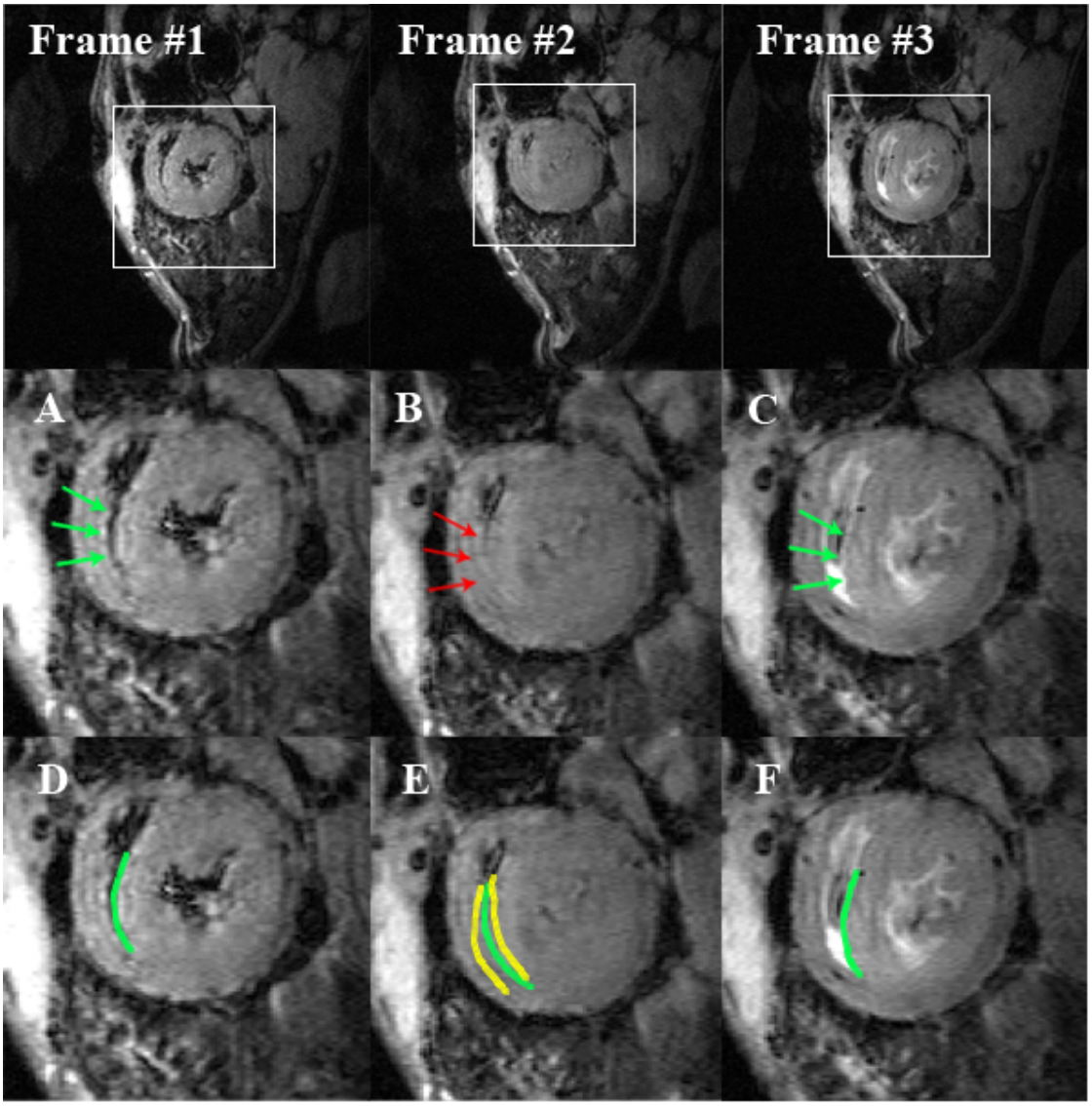

Figure 3.5: Illustration of segmentation ambiguity caused by the lack of prominent image feature

tures are not visible on one frame, but they can be observed on neighboring frames. As illustrated in Fig. 3.5, within the same slice, edges pointed by green arrows are observed on frames #1 and #3 but not on frame #2. It shows that although most image features on frames in the same slice are consistent, some information present on different frames can also be complementary. For the segmentation problem par-

ticularly, given one single 2D image, sometimes the "correct" segmentation is not unique and there exist multiple acceptable solutions (illustrated in yellow segments in Fig. 3.5, however only green segments are correct according to expert's manual segmentation). Interestingly, the ambiguity is largely eliminated when a sequence of images are provided, and that is exactly why experts go back and forth between frames to refine their drawings when they perform manual segmentation. As a result, to resolve the segmentation ambiguity, even if just one image in the 3D+t dataset needs to be segmented, the proposed method not only utilizes the frame of interest, but also considers all the remaining frames in the same slice.

There is always a tradeoff between system performance and the amount of user interaction required. We managed to find an efficient solution that significantly enhances the robustness of the proposed method while requiring only minimum user input. For each slice of images, one frame is manually selected as the reference frame, and a rough center of the epicardium in the reference frame is provided by the user, then the system correctly initializes the region of interest (ROI), according to which the images are cropped. Without affecting the segmentation accuracy, image cropping largely reduces the computational cost.

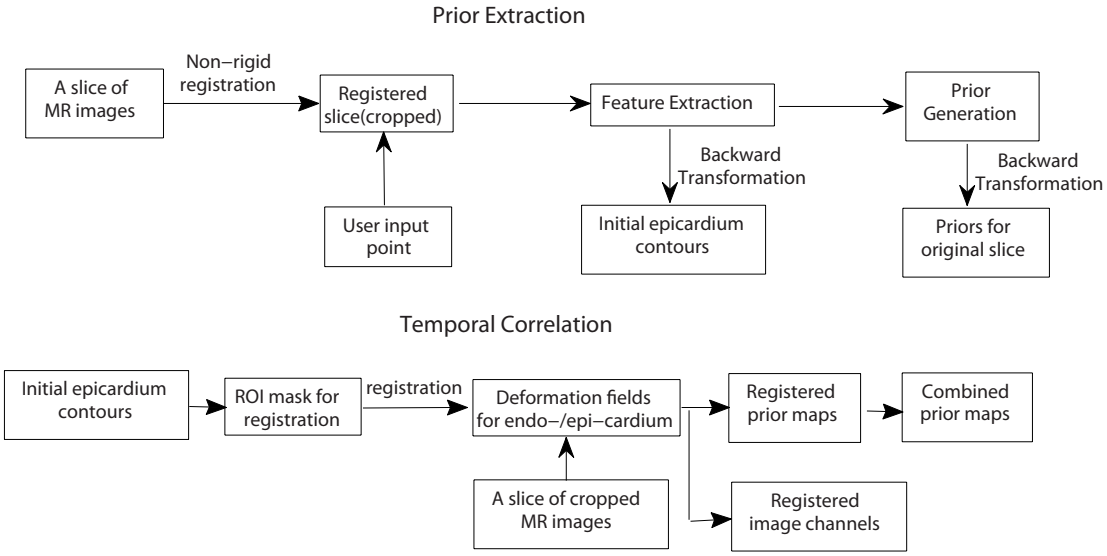

Figure 3.6: Steps of the acquisition of data-driven priors and the establishment of temporal correlations

3.3 Algorithm Overview

The proposed method realizes the segmentation of the LV myocardium in rat MRI in five steps: 1) preprocessing, 2) generate diffused structure tensor space, 3) extract the data-driven priors, 4) establish temporal correspondences, and 5) energy formulation and optimization. Details on the extraction of data-driven priors and the construction of temporal point correspondences are illustrated in Fig. 3.6.

To enhance image quality and suppress systematic intensity inhomogeneity, image preprocessing is performed before further operations. Due to the fact that intensity values in the LV cavity are inconsistent and fluctuate dramatically during blood flow into and out of the cavity while intensity values in the myocardium region are

consistent and homogeneous, diffused structure tensor space is constructed based on the image to capture texture information.

Given an user input point, all frames in a slice are cropped, and non-reference frames are registered to the selected reference frame. Prominent features, i.e., corner points inside LV cavity and scale-invariant edges along the LV epicardium, are extracted from the registered images. Probabilistic priors maps are then generated based on the distributions of extracted corner points and edges. According to the detected scale-invariant edges along LV epicardium, an initial epicardium contour can be estimated. As all the above operations are performed on the registered image sequence, probabilistic priors and initial epicardium contours need to be deformed backward to match the unregistered original slice.

After the data-driven priors are obtained, correspondences among pixel points on frames in the same slice are formed through non-rigid registration. Minimizing the registration error according to the ROI masks defined by initial epicardium contours generated previously, cropped original frames are registered to the reference slice, and different deformation fields are then interpolated for the purpose of endocardium segmentation and epicardium segmentation separately. With the deformation fields, frames in the same slice are connected: any point on any frame can be matched to another point on any other frame within that slice.

Combining the data-driven priors, temporal correlation and the diffused structure

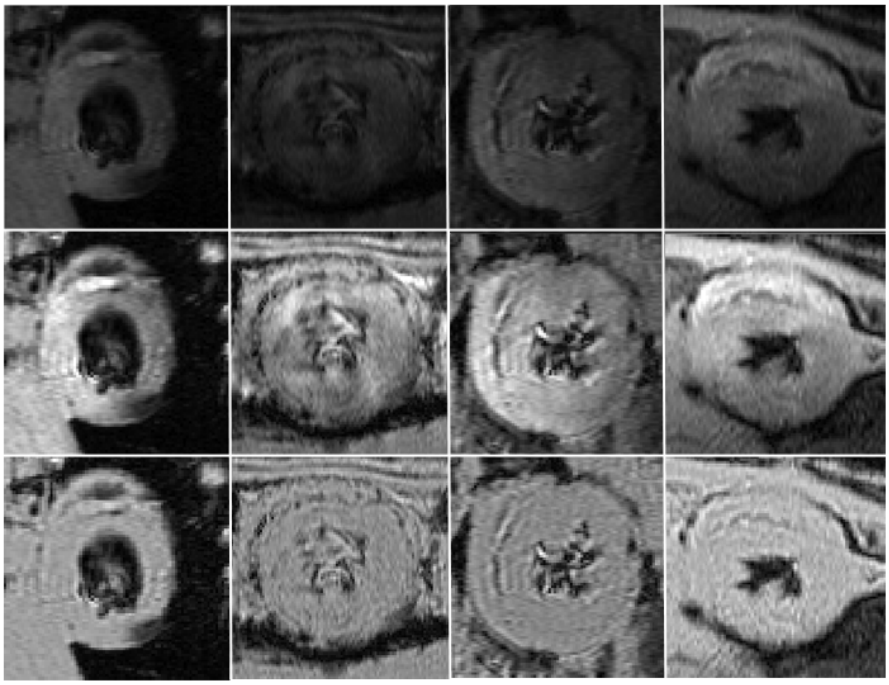

Figure 3.7: Preprocessing. First row: original images. Second row: images after contrast enhancement. Third row: images after contrast enhancement and inhomogeneity correction.

tensor space, an energy functional is formulated. By minimizing the energy functional in a level set framework, the LV myocardium is automatically segmented.

3.4 Preprocessing

Due to low contrast and inhomogeneity of raw cardiac MR images of rat hearts, preprocessing including contrast enhancement as well as inhomogeneity correction are performed prior to feature extraction. Image contrast is enhanced by histogram equalization, and inhomogeneity is corrected by the algorithm proposed by Axel et al. in [34]. In [34], the bias field of the MR image is estimated and corrected

using an approximation of the image of a uniform phantom, which is obtained by blurring the original image.

In Fig. 3.7, cropped original images are displayed in the first row, images after contrast enhancement are shown in the second row, and images in the last row are obtained after contrast enhancement and inhomogeneity correction. One can observe that through preprocessing, edges in the original MR images are preserved, and systematic intensity inhomogeneity is removed.

3.5 Diffused Structure Tensor Space

As shown in the top rows in Figs. 3.8 and 3.9, intensity values in the LV cavity are inconsistent and fluctuate dramatically during blood flow into and out of the cavity while intensity values in the myocardium region are normally consistent and homogeneous. This implies that texture can be used as a discriminant feature.

Presented by Rousson et al. in [7], unsupervised texture segmentation can be realized in the diffused structure tensor feature space, which is also referred to as the "feature space" for short in this thesis. Details are described in Section 2.1.2. Applying Rousson's method, the feature space can be constructed from the cropped rat MR images. As shown in Figs. 3.8 and 3.9, the first channel is obtained by diffusing original image, i.e., I, and the second, third and fourth

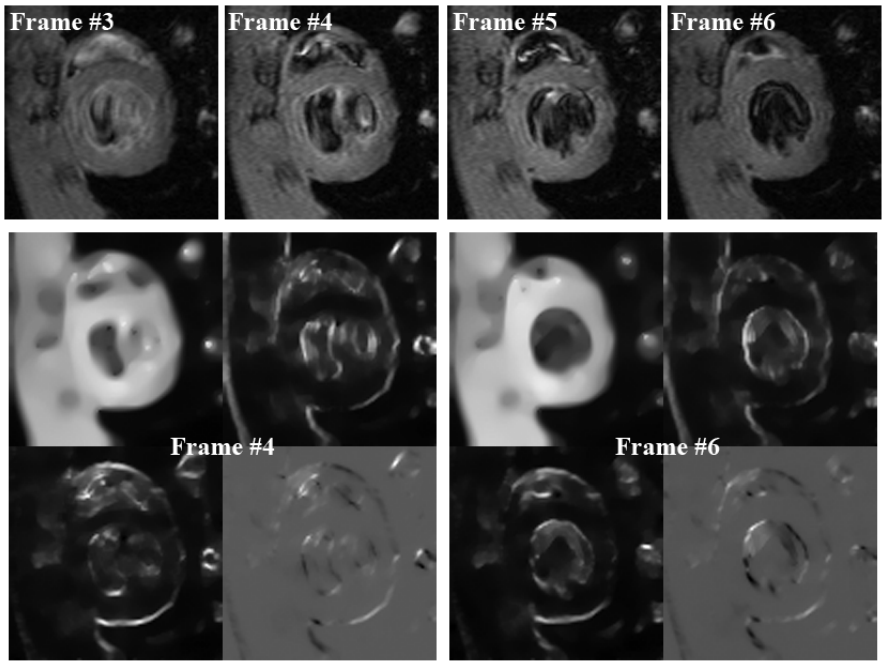

Figure 3.8: Diffused structure tensor space of a native rat heart

channels are diffused from I_x^2, I_y^2 and $I_x I_y$, respectively. I_x is the gradient of I in the x direction, and I_y is the gradient of I in the y direction.

One can observe that especially in the second and third channel, myocardium regions are homogeneous and with low intensity, whereas the LV cavity regions are generally bright. In some images, if we know the epicardium boundary, the classification of points within LV cavity and points on LV myocardium can be easily done by common stochastic classifiers. The feature space is shown to be superior to the original image in discriminating textures, therefore the proposed method uses the feature space predominately.

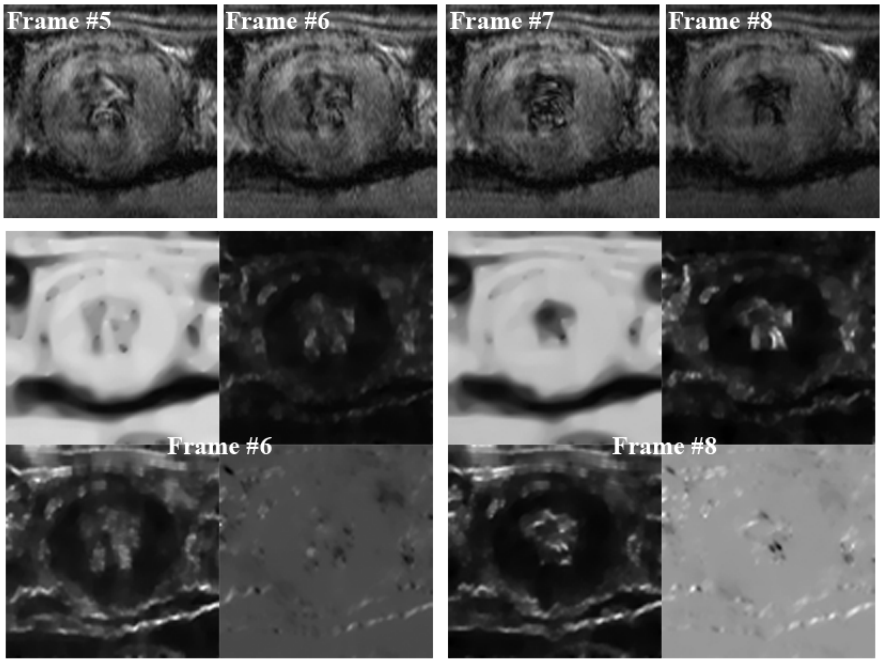

Figure 3.9: Diffused structure tensor space of a transplanted rat heart

3.6 Acquisition of Data-driven Priors

Deformable models without prior knowledge cannot provide accurate segmentation and may result in leaking due to noise and complexity of organ structures. To overcome this problem, probabilistic priors are automatically generated based on extracted features. Prominent feature points, i.e., corner points and scale-invariant edges, are detected respectively as good indicators of the LV cavity and the epicardium. To remove undesired feature points, we assume that the point provided by user is close to the true center of epicardium and the shape of the epicardium is approximately circular.

3.6.1 Registration

Different from the method previously proposed in [6], we now first register all remaining frames in the selected slice to the reference frame through non-rigid registration, so the registered frames and the reference frame form a "registered slice".

The reason of introducing the registered slice is to compensate heart motion, therefore minimize the error in prior extraction step. As aforementioned, to segment any frame, the proposed method utilizes all frames within that slice. Specifically in the acquisition of data-driven prior for epicardium segmentation, scale-invariant edges on current frame and neighboring frames are detected, and those edges are combined to generate a data-driven prior map for current frame. Therefore, it is desirable that features detected in neighboring frames can be correctly mapped onto the current frame, which means that the heart motion has to be compensated.

Since motion compensation is particularly important to the generation of data-driven prior for epicardium segmentation, it is crucial that registration error especially along the epicardium boundary is minimized. As shown in Figs. 3.10 and 3.11, the first row displays cropped original images within the same slice, and the corresponding registered images are in the second row. The middle column shows the reference image with a set of green landmarks along epicardium, and the same set of landmark points are shown in red on neighboring frames. One can

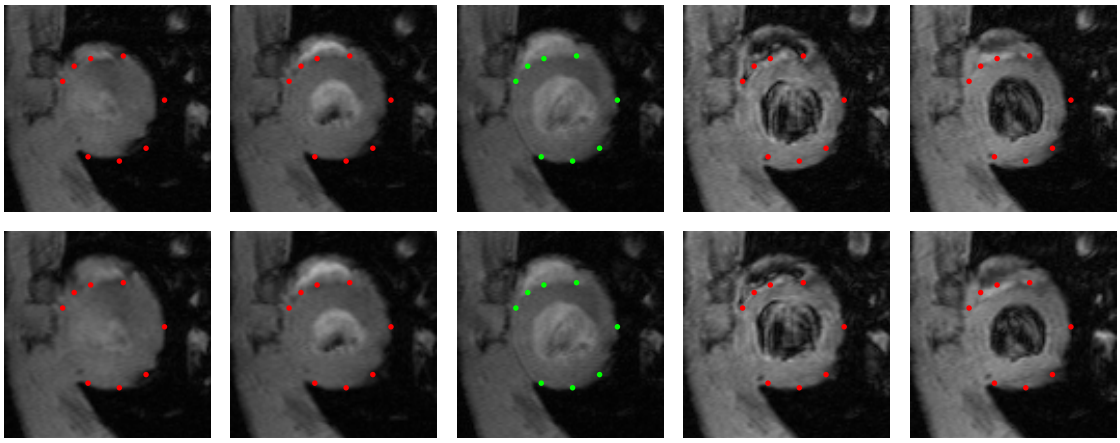

Figure 3.10: Illustration of registration accuracy along epicardium (native rat heart). First row: original images. Second row: registered images.

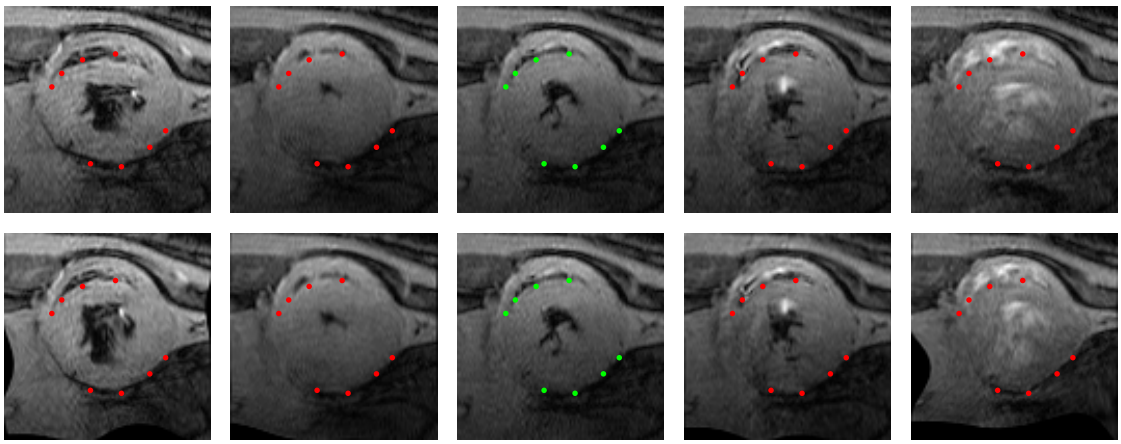

Figure 3.11: Illustration of registration accuracy along epicardium (transplanted rat heart). First row: original images. Second row: registered images.

observe that through registration, original images are deformed such that points along epicardium in neighboring frames are mapped to the ones in the reference frame. As a result, epicardium wall motion is compensated.

3.6.2 Priors for Endocardium

As previously illustrated, the intensity distribution in the LV cavity is unpredictable and inhomogeneous. To extract features that are only available in the LV cavity but not in the myocardium and at the same time complementary to low level image features, we apply corner point detection.

Compared to blob detectors, corner detectors are more robust against intensity inhomogeneities inside the LV cavity region caused by turbulent blood flow. Here we use the algorithm proposed by Rosten and Drummond in [35, 36] to detect corner points. These corner points are mostly located either inside the LV cavity or outside the epicardium, with very few in the myocardium. Therefore, it is possible to filter out corner points outside the LV cavity. Taking the center point provided by the user as the origin, we convert all corner points to polar coordinates and obtain the distribution of the corner points with respect to radius R. A threshold radius R^* is then calculated using Gaussian mixture models to extract the corner points inside the LV cavity (see Fig. 3.12).

Using the extracted corner points as sample points, for the LV cavity, a relative probability density function having values in the range $[0, 1]$, $p_{\text{prior}}(x, y)$ (see Fig. 3.12(d)), is obtained by kernel density estimation. To avoid the domination of priors in the energy functional, we define a conservative prior probability map

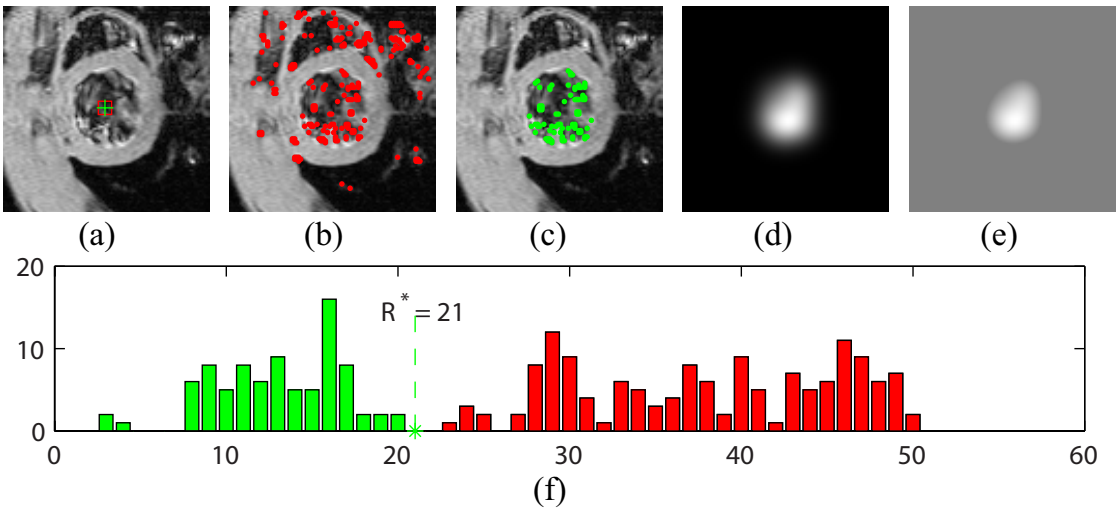

Figure 3.12: Extraction of endocardium prior. (a) User provided point; (b) All corner points detected; (c) Corner points within the LV cavity; (d) Relative probability density map; (e) Prior map for endocardium segmentation; (f) Distribution of corner points in polar coordinates.

as:

$$P_{\text{prior}}(x,y) = \begin{cases} p_{\text{prior}}(x,y) & \text{if } p_{\text{prior}}(x,y) > 0.5 \\ 0.5 & \text{otherwise} \end{cases} . \tag{3.1}$$

As shown in Fig. 3.12(e), a high prior probability indicates that the point is more likely to be inside the LV cavity than outside. On the other hand, a prior probability of 0.5 indicates no preference between the inside and outside of the LV cavity.

As the probabilistic maps are obtained from the registered slice, we deform the maps backward to get prior maps that match the original slice.

3.6.3 Priors for Epicardium

Similar to the feature selection for endocardium segmentation, prominent feature that represents the epicardium boundary has to be complementary to low level image features, and at the same time, it can only be uniquely detected along the epicardium, or detected features that do not lie on the epicardium boundary can be effectively filtered out. The scale-invariant edge is one of the most discriminative features that fits the requirements.

We detect scale-invariant edges according to the method described in [37, 38]. In order to select edges along the epicardium, we first filter out undesired ones by examining edge directions: desired edges should be tangent to the epicardium, which is approximately circular. If a detected edge and the corresponding radial direction are nearly perpendicular, the edge is preserved; otherwise, the edge is discarded.

In polar coordinates, we remove edges inside the LV cavity using the threshold R^* obtained previously. After that, edges along the epicardium can be approximately extracted by selecting edges with radius in the range from $\tilde{R}$ to $\tilde{R}+ \Delta r$, where $\tilde{R}$ is the minimum radius value of edges not inside the cavity and Δr is a tolerance distance that assures the true boundary is inside its capture range. In our experiments, Δr is set to be 10 pixels.

Let $S_{i,j}$ denote the set of extracted edges in the j^{th} frame of the i^{th} slice. Since

the epicardium is generally approximately circular, we estimate the distribution of edge points along radial directions. Let $N(\mu_{i,\theta}, \sigma^2_{i,\theta})$ denote a global Gaussian distribution of the epicardial radius, estimated from edge points from all frames of i^{th} slice along angle θ and its neighboring directions. To estimate the distribution $N(\mu_{ij,\theta}, \sigma^2_{ij,\theta})$ for the j^{th} frame, we first remove outliers from set $S_{i,j}$ based on $N(\mu_{i,\theta}, \sigma^2_{i,\theta})$. Next, we estimate the distribution of the epicardial radius in different directions using the remaining edge points. In case the edge points are not detected along certain segments in the j^{th} frame, we use the global distribution $N(\mu_{i,\theta}, \sigma^2_{i,\theta})$ to approximate $N(\mu_{ij,\theta}, \sigma^2_{ij,\theta})$. By mapping the estimated distributions $N(\mu_{ij,\theta}, \sigma^2_{ij,\theta})$ back to Cartesian coordinates, the probabilistic map (normalized to [0, 1]) of the epicardium for a particular frame, as shown in Fig. 3.13, can be obtained.

To estimat an initial epicardium boundary for the j^{th} frame in the i^{th} slice, we convert points in polar coordinates with radius $\mu_{ij,\theta}$ for each angle θ back to Cartesian coordinates. As shown in Fig. 3.13, the estimated initial epicardium boundary is plotted in blue.

Similarly, as the probabilistic maps are obtained from the registered slice, we deform the maps backward to get prior maps for the original slice.

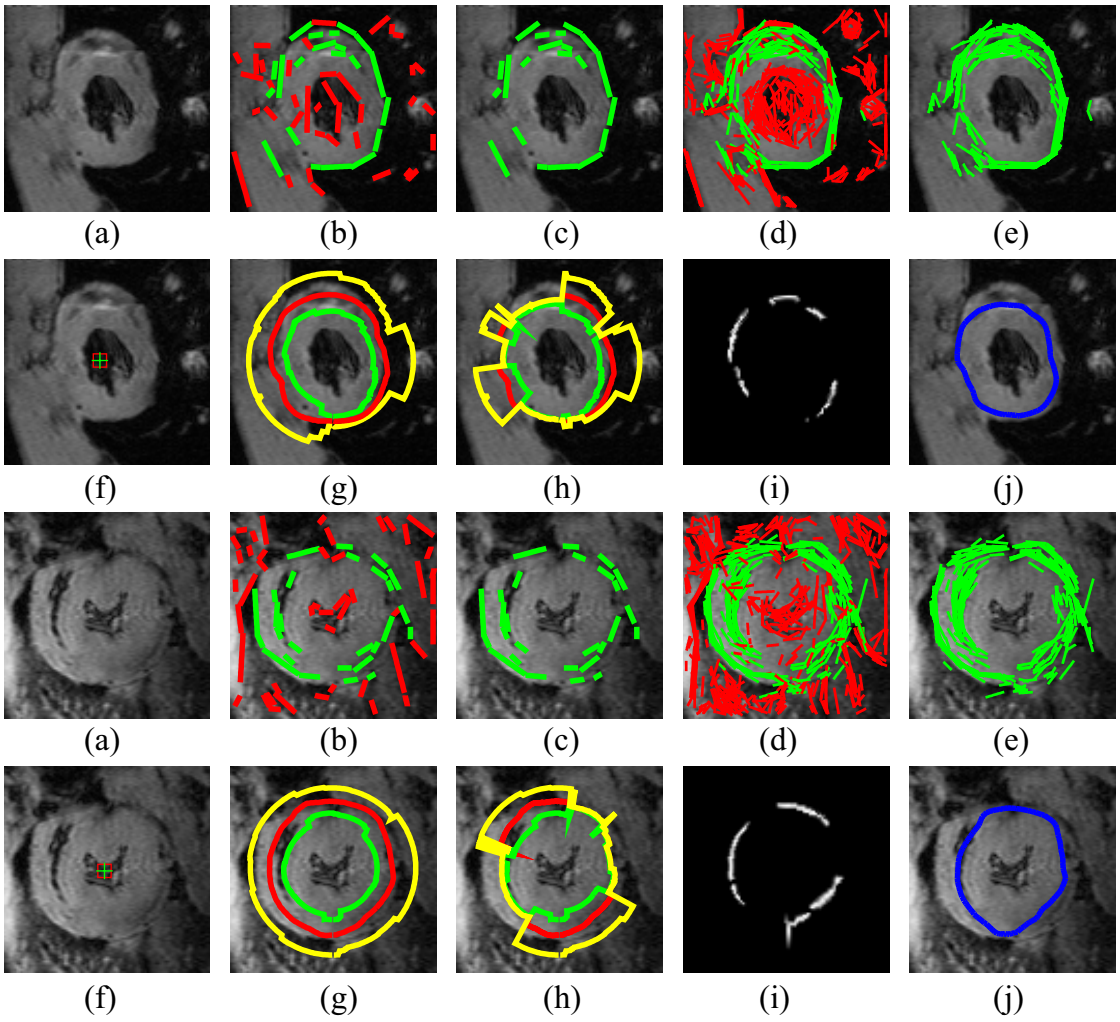

Figure 3.13: Extraction of epicardium prior. (a) Original image. (b) Edges detected from the current image. (c) Edges in the current frame after filtering. (d) Edges detected from all image in the slice. (e) All edges in the slice after filtering. (f) User provided point. (g) Illustration of $N(\mu_{i,\theta}, \sigma^2_{i,\theta})$. (h) Illustration of $N(\mu_{ij,\theta}, \sigma^2_{ij,\theta})$. (i) Prior map for epicardium segmentation. (j) Estimated initial epicardium boundary.

3.7 Establishment of Temporal Correlations

Temporal correspondences between different frames in the same slice are constructed through image registration. Here, the purpose of performing image regis-

tration is to get the deformation fields, which explicitly show heart motion. Using the deformation fields obtained from registration, point correspondences are established. Then complementary information that is present on neighboring frames are correctly transformed onto current frame, therefore segmentation ambiguity can be effectively resolved.

Different from normal human MRI, MR images of rat hearts have very limited resolution, low SNR and very irregular LV cavity shape. Most importantly, intensity of the blood in the LV cavity changes dramatically among different frames within the same slice, as shown in Figs. 3.8 and 3.9. All the above mentioned differences make the frame-to-frame registration a very difficult task.

3.7.1 Registration

Most methods achieve registration by minimizing a cost function that consists of two terms: a data term and a smoothness term. The data term is defined in such a way that by minimizing it, the pixel-wise similarity between target image and the registered floating image is maximized. The smoothness term is sometimes referred as the regularization term, which constrains discontinuous deformation by penalizing local deviations on the deformation fields. The energy minimization process is actually a compromise between the pixel-wise difference defined by the data term and spatial constraint specified in the smoothness term.

Due to the fact that the endocardium deforms more dramatically than the epicardium does, it is necessary to define different smoothness costs in the LV cavity, myocardium, and regions outside epicardium, such that a weak smoothness penalty allows large deformations in the LV cavity and myocardium while a strong smoothness penalty ensures local continuity of the deformation fields in other regions. However, the proper definition of such a smoothness cost is difficult in actual implementation. Without a spatially varying smoothness constraint, registration accuracy will be compromised. Therefore in this thesis, to avoid registration error caused by improper definition of the smoothness cost, we perform registration using two spatially uniform smoothness constraints with different ROI masks to obtain deformation fields for endocardium and epicardium segmentations separately. Here we adopt the B-spline based non-rigid registration proposed by Rueckert et al. in [26], and the method is implemented by Dirk-Jan Kroon from the University of Twente.

To overcome problems caused by inconsistent intensity in the region of LV cavity, we register feature maps instead of original images. For each frame in a particular slice, its feature map is obtained from the constructed feature space. As we already have the initial epicardium boundary in every frame, we define a ring shape region (shown in the second column of Fig. 3.14) in which most of the points belong to the myocardium. Then average intensity of the myocardium can be estimated easily by calculating mean intensity value within the ring region. Note that the

estimation here does not have to be very accurate. With the estimated average myocardium intensity, we define the feature map as follows:

$$I_{\mathrm{FM}} = |I - \bar{i}| + |u_1 - \bar{i}| + \sum_{k=2,3} u_k, \tag{3.2}$$

where $\bar{i}$ is the estimated average myocardium intensity, I is the original image, and u_k is the k^{th} channel in the feature space. u_1, u_2, and u_3 are respectively shown in the third, fourth, and fifth column of Fig. 3.14, and I_{FM} is shown in the last column. We provide more examples in Fig. 3.15: images in the first row are cropped original MR images, and the corresponding feature maps are shown in the second row.

As shown in Figs. 3.14 and 3.15, instead of a mix of bright and dark pixels, the LV cavity region is filled with pixels having higher intensities compared to pixels in the myocardium region. As a result, to register the feature maps, there is no need to define a similarity measure as what normally used in inter-modality image registration, and here we choose SSD as the similarity measure.

3.7.1.1 Endocardium

In the process of generating deformation fields for endocardium segmentation, we focus on enhancing the registration accuracy within the LV cavity. Therefore, we define a ROI using the initial epicardium contours as a mask (as shown in white in Fig. 3.16(c) and (f)) to calculate registration error.

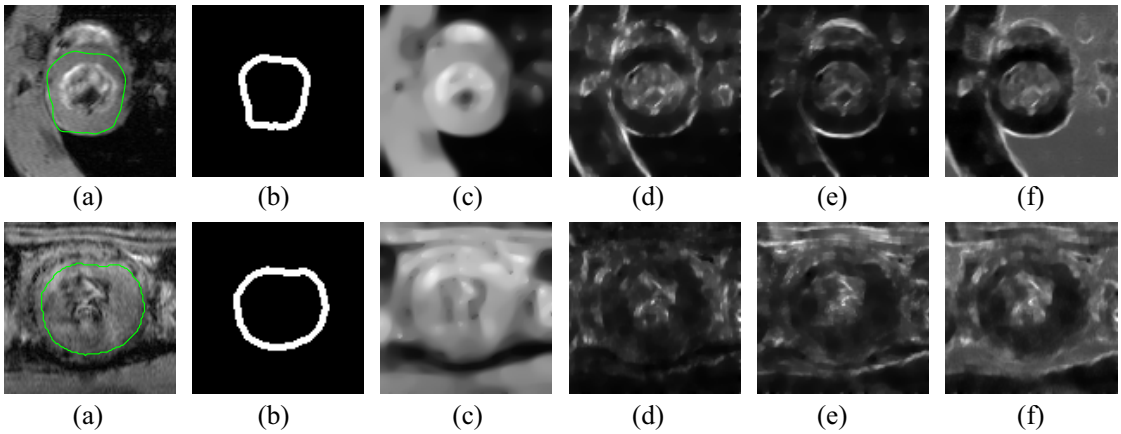

Figure 3.14: Feature maps for MR images of native and transplanted rat hearts.(a) Estimated initial epicardium boundary. (b) Ring shape mask. (c)-(e) Feature channels u_1, u_2, and u_3. (f) Feature map.

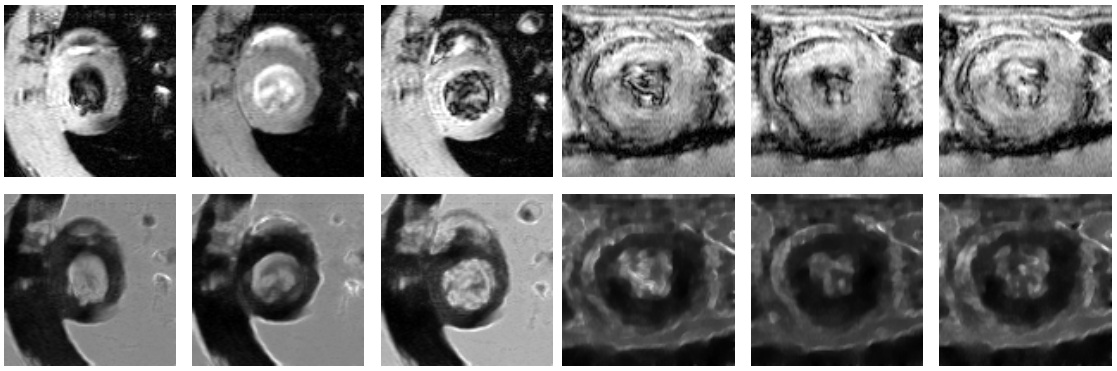

Figure 3.15: More feature maps. First row: original images. Second row: corresponding feature maps.

The registration results are shown in Fig. 3.17. The first and third row respectively displays cropped original MR images of a native and transplanted rat heart, and the second and last row shows corresponding registered images. Similar to Section 3.6.1, the middle column shows the reference image with a set of green landmarks along endocardium, and the same set of landmark points are shown in red on neighboring frames. One can observe that through registration, original images are

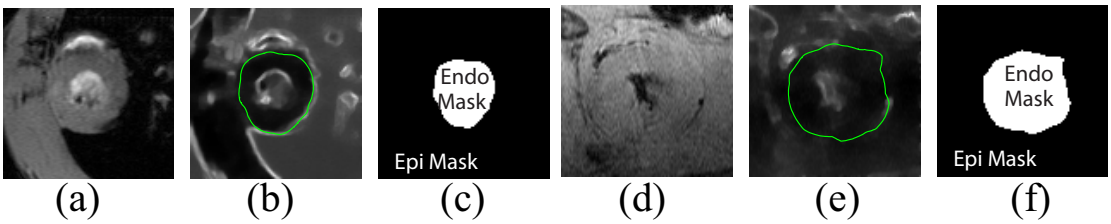

Figure 3.16: Registration Masks. (a,d) Original image. (b,e) Estimated initial epicardium boundary plotted on the feature map. (c,f) Registration mask.

deformed such that points along endocardium in neighboring frames are mapped to the ones in the reference frame.

Here we use $T_{\text{endo}}^{(i,j)}$ to represent the deformation field for endocardium segmentation from the i^{th} frame to the j^{th} frame. For a slice with 10 frames, deformation fields are stored in a 10 by 10 array, which has all diagonal elements equal to zero.

Ideally, for any pair of frames A and B within a slice, the deformation field for endocardium segmentation has to be obtained by registering frame A and B. However, the computational cost would be very high if such mechanism was adopted in the proposed method. In our implementation, we only register the non-reference frames to the selected reference frame once. We use the following equations to get the deformation field $T_{\text{endo}}^{(i,j)}$ for any pair of frames:

$$T_{\text{endo}}^{(r,n)} = B\left(T_{\text{endo}}^{(n,r)}\right), \tag{3.3}$$

$$T_{\text{endo}}^{(n_1,n_2)} = T_{\text{endo}}^{(n_1,r)} + B\left(T_{\text{endo}}^{(n_2,r)}\right) \tag{3.4}$$

where frame r is the reference frame, frames n, n_1, and n_2 are non-reference frames in the same slice, and $B\left(T_{\text{endo}}^{(i,j)}\right)$ is the backward transformation function that

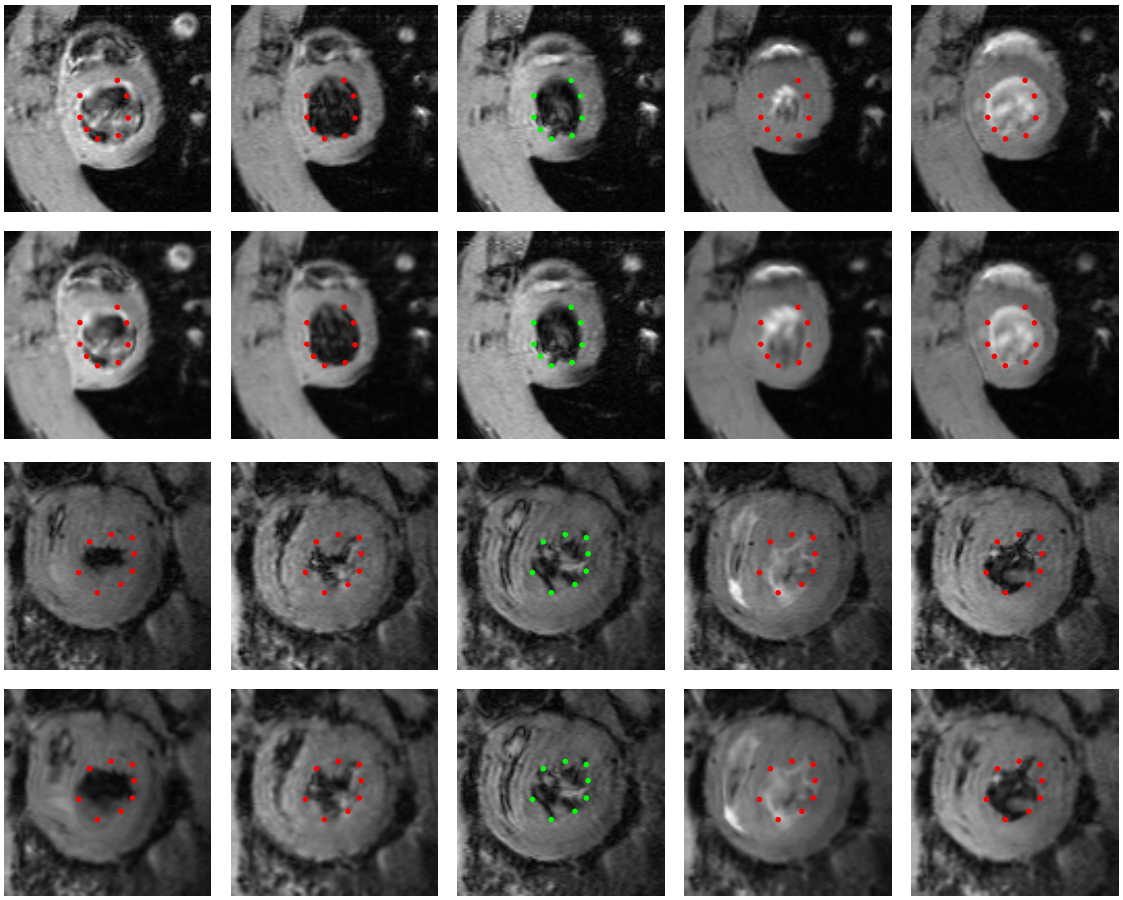

Figure 3.17: Registration results for endocardium segmentation

calculates the deformation field $T_{\text{endo}}^{(j,i)}$ based on $T_{\text{endo}}^{(i,j)}$.

3.7.1.2 Epicardium

Similar to Section 3.7.1.1, we only minimize the registration error within the mask (shown in black in Fig. 3.16(c) and (f)). Registration results are shown in Fig. 3.18. One can observe that through registration, original images are deformed such that points along epicardium in neighboring frames are mapped to the ones in the

reference frame.

Here we define $T_{\text{epi}}^{(i,j)}$ to represent the deformation field from the i^{th} frame to the j^{th} frame. In the implementation, we also only register the non-reference frames to the selected reference frame once. To get the deformation field $T_{\text{epi}}^{(i,j)}$ for any pair of frames, we use similar equations:

$$T_{\text{epi}}^{(r,n)} = B\left(T_{\text{epi}}^{(n,r)}\right), \tag{3.5}$$

$$T_{\text{epi}}^{(n_1,n_2)} = T_{\text{epi}}^{(n_1,r)} + B\left(T_{\text{epi}}^{(n_2,r)}\right) \tag{3.6}$$

3.7.2 Combined Feature Spaces

To resolve the segmentation ambiguity caused by the lack of prominent image features on the frame of interest, we propose to use the image information in the remaining frames within the same slice as complementary information to constrain the segmentation.

To find a combined feature space utilizing all frames within the same slice, we use the weighted average:

$$\widetilde{u}_{\text{endo/epi},k}^{i} = \alpha_c u_k^i + \alpha_n \sum_{j \neq i} T_{\text{endo/epi}}^{(j,i)}(u_k^j), \tag{3.7}$$

where $\widetilde{u}_{\text{endo/epi},k}^{i}$ is the combined k^{th} channel in the feature space of the i^{th} frame for endocardium or epicardium segmentation, u_k^i is the k^{th} channel in the feature

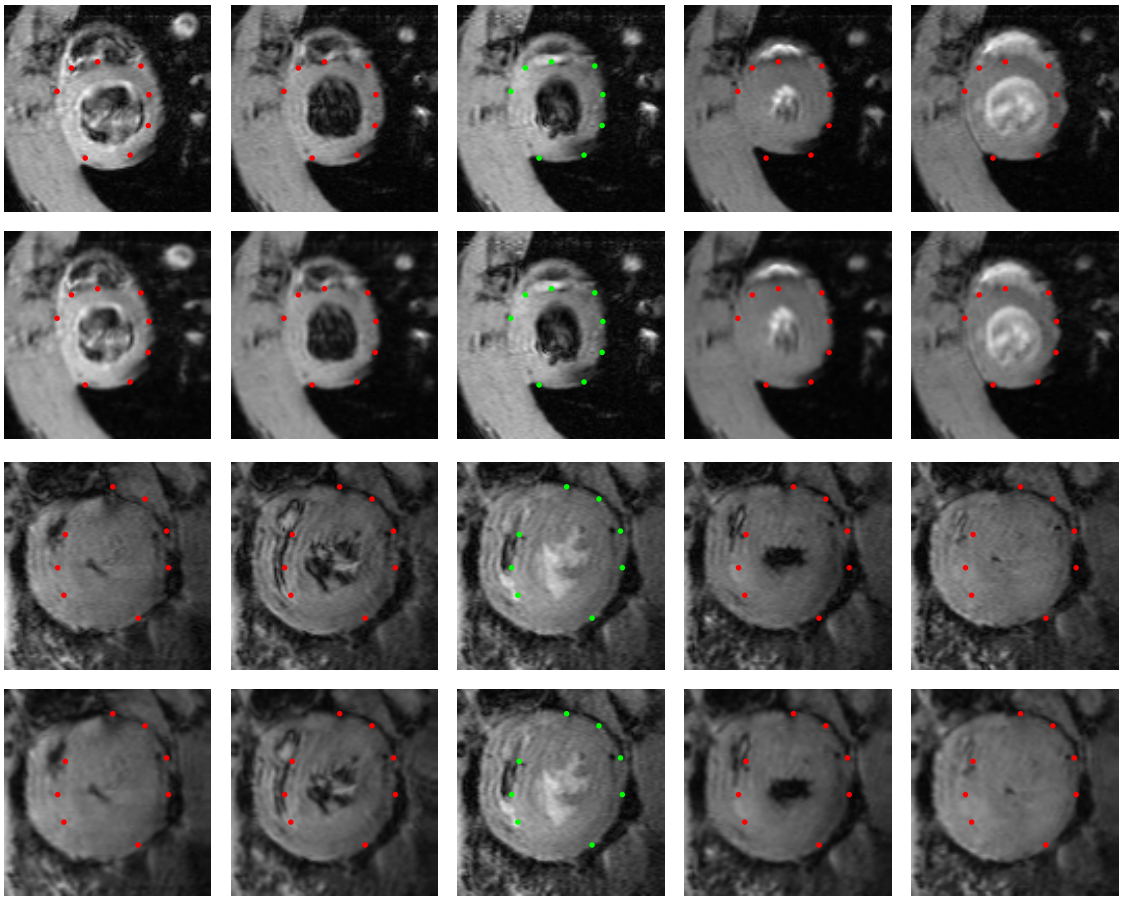

Figure 3.18: Registration results for epicardium segmentation

space of the i^{th} frame obtained in Section 3.5, and α_c and α_n are weights for the current and neighboring frames, respectively.

As shown in Figs. 3.19 and 3.20, current frame and neighboring frames are combined according to (3.7). It is easy to observe that the combined feature space provides more information at regions where current frame displays indiscriminative image feature. Incorporating information present on all frames in the same slice, the combined feature space has enhanced discriminability, therefore the segmentation ambiguity problem is effectively resolved.

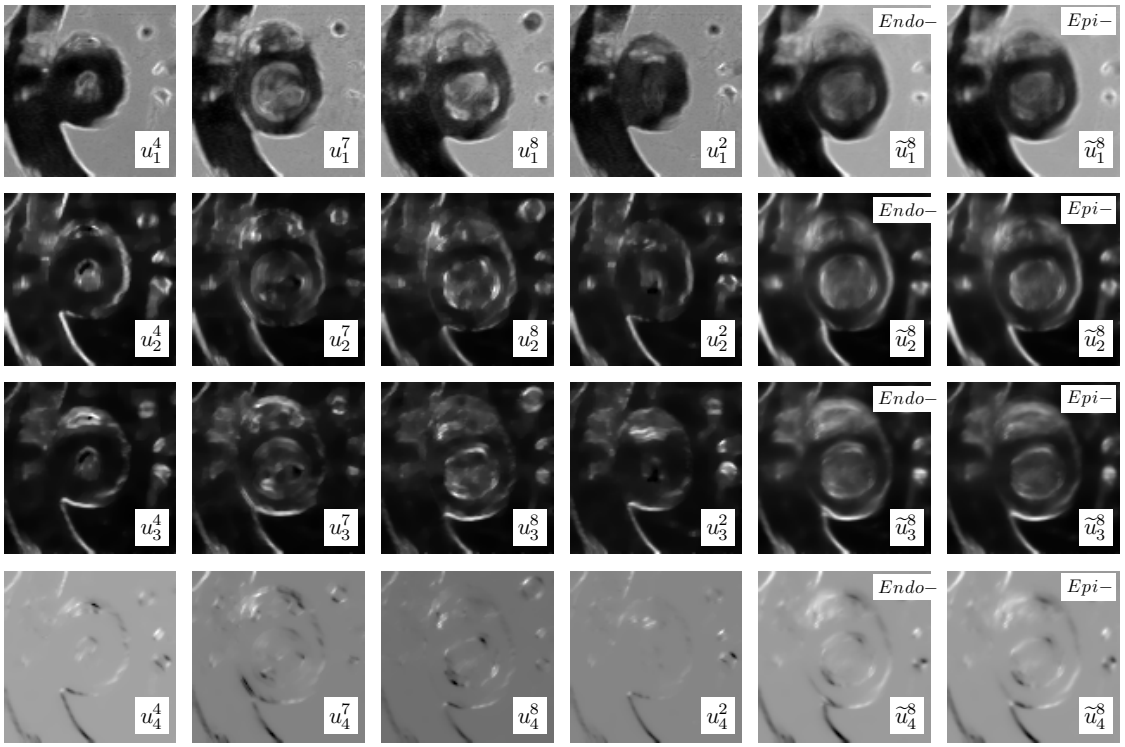

Figure 3.19: Combination of feature spaces (native rat heart). First four columns: feature space of individual frames. Fifth column: combined feature space for endocardium segmentation. Last column: combined feature space for epicardium segmentation.

3.7.3 Combined Probabilistic Prior Maps

In our previous publication [6], although features extracted from all frames within one slice were utilized in generating the prior map for a particular frame, features detected on the frame of interest is assumed to most reflect the image structure and used primarily in calculating the prior map. As a result, the generated prior maps are conservative and preserve only information with high probability. In this thesis, to utilize the reliable information in prior maps to the most possible extent, we also take prior maps of neighboring frames into consideration when one frame

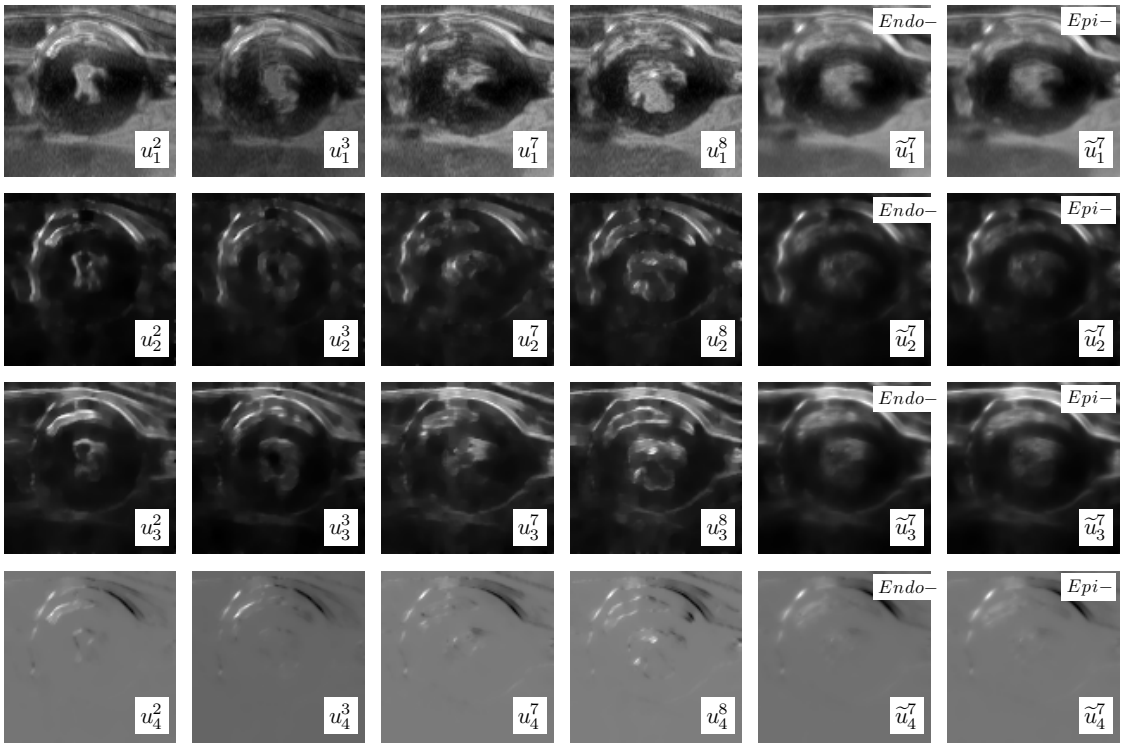

Figure 3.20: Combination of feature spaces (transplanted rat heart). First four columns: feature space of individual frames. Fifth column: combined feature space for endocardium segmentation. Last column: combined feature space for epicardium segmentation.

is segmented.

Here we also use weighted average:

$$\widetilde{P}^i_{\text{endo/epi}} = \beta_c P^i_{\text{endo/epi}} + \beta_n \sum_{j \neq i} T^{(j,i)}_{\text{endo/epi}}(P^j_{\text{endo/epi}}), \tag{3.8}$$

where $\widetilde{P}^i_{\text{endo/epi}}$ is the combined prior map of the i^{th} frame for endocardium or epicardium segmentation, $P^i_{\text{endo/epi}}$ is the prior map of the i^{th} frame for endocardium or epicardium segmentation obtained in Section 3.6, and β_c and β_n are weights for the current and neighboring frames, respectively.

As shown in Fig. 3.21, the current frame prior map and complementary informa-

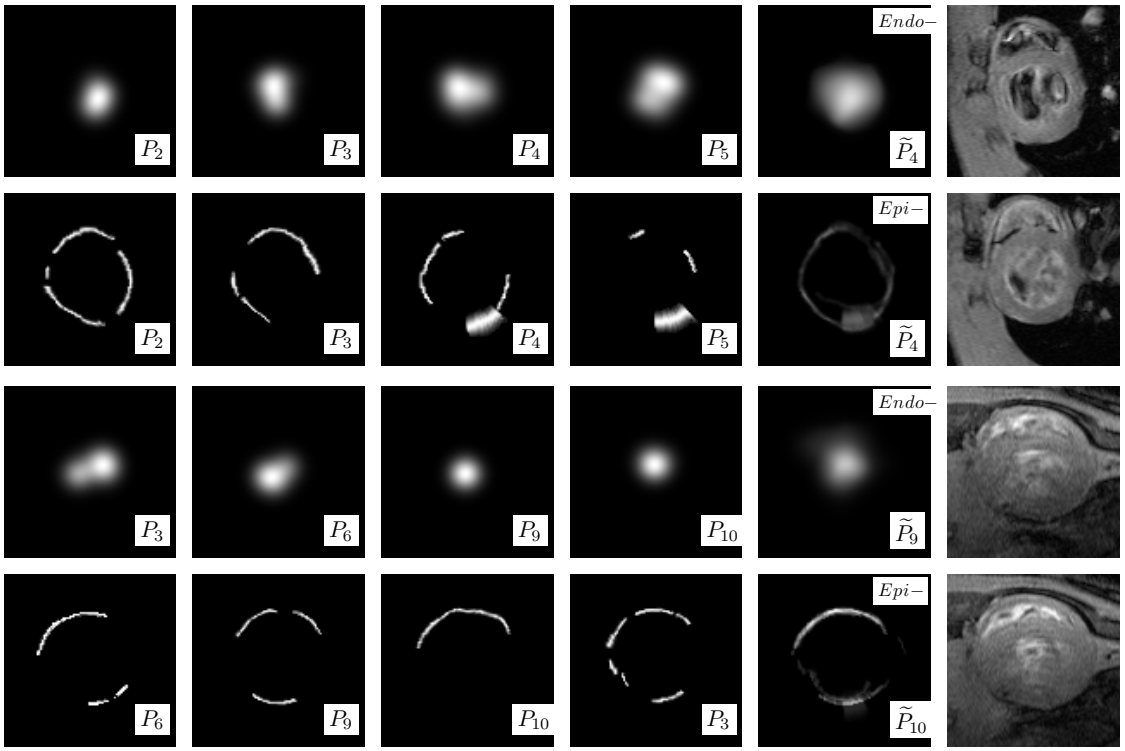

Figure 3.21: Combination of prior maps. First four columns: prior maps of individual frames. Fifth column: combined prior maps. Last column: corresponding original images.

tion in prior maps of the rest frames are integrated to get a combined prior map. For endocardium prior maps, one can observe that the combined prior map ($\widetilde{P}^4_{\text{endo}}$ and $\widetilde{P}^9_{\text{endo}}$) represents the LV cavity better than the individual prior map (P^4_{endo} and P^9_{endo}). Similarly, for epicardium prior maps, information in the combined prior map ($\widetilde{P}^{10}_{\text{epi}}$ and $\widetilde{P}^{10}_{\text{epi}}$) is augmented by incorporating complementary information in prior maps of other frames, and the combined prior map is more representative to the epicardium than the individual prior map (P^4_{epi} and P^{10}_{epi}).

3.8 Energy Formulation

For a contour C in the i^{th} frame, which is embedded as the zero level set of function ϕ_i, the energy functional is defined as:

$$J(\phi_i) = \lambda_r J_r(\phi_i) + \lambda_t J_t(\phi_i) + \lambda_e J_e(\phi_i), \tag{3.9}$$

$$J_t(\phi_i) = \sum_{j \neq i} J_r \left(T^{(i,j)}_{\text{endo/epi}}(\phi_i) \right), \tag{3.10}$$

where $J_r(\phi_i)$ is the region-based term incorporating extracted priors and images in feature channels; $J_t(\phi_i)$ is the temporal constraint term; $J_e(\phi_i)$ is the edge-based term moving the contour towards the object boundaries; and λ_r, λ_t and λ_e are wights that regulate relative strength of above mentioned three terms, respectively. In our implementation, we set $\lambda_r = 1$, $\lambda_t = 0.1$, and $\lambda_e = 0.5$, for both endocardium and epicardium segmentation.

The edge-based term $J_e(\phi_i)$ is common for both endocardium and epicardium segmentation. Let g_i be an inverse edge indicator function:

$$g_i = \frac{1}{1 + |\nabla G_\sigma * I_i|^2}, \tag{3.11}$$

where G_σ is the Gaussian kernel with standard deviation σ and I_i is the preprocessed image in the i^{th} frame. Thus, J_e is given by

$$J_e(\phi_i) = \int_\Omega g_i \delta_\epsilon(\phi_i(x,y)) |\nabla \phi_i(x,y)| dxdy, \tag{3.12}$$

where δ_ϵ is the regularized Dirac function.

According to Section 3.7.2, we have constructed the feature space $\widetilde{u}^i = (\widetilde{u}^i_1, ..., \widetilde{u}^i_4)$ for the i^{th} frame. The probability density function for $\widetilde{u}^i(x, y)$ to be in the foreground Ω_1 and the background Ω_2 can be estimated by

$$p_j\left(\widetilde{u}^i(x,y)\right) = \prod_{k=1}^{4} p_{k,j}(\widetilde{u}^i_k(x,y)) \qquad j \in \{1,2\}, \tag{3.13}$$

where $p_{k,j}\left(\widetilde{u}^i_k(x,y)\right)$ represents the likelihood of a point (x, y) belonging to Ω_j, based on the k^{th} channel (I, I_x^2, I_y^2 or $I_x I_y$) of the combined feature space, $\widetilde{u}^i$.

Here we adopt Gaussian approximation for all channels. Since the image $\widetilde{u}^i$ is vector-valued, we have to deal with covariance matrices. Let $\{\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1\}$ and $\{\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2\}$ be the vector's means and covariance matrices of the Gaussian approximation in Ω_1 and Ω_2. The probability of $\widetilde{u}^i$ to be in Ω_j, $j \in \{1, 2\}$, is:

$$p_j\left(\widetilde{u}^i(x,y)\right) = \frac{1}{(2\pi)^2|\boldsymbol{\Sigma}_j|^{1/2}} e^{-\frac{1}{2}\left(\widetilde{u}^i(x,y)-\boldsymbol{\mu}_j\right)^T \boldsymbol{\Sigma}_j^{-1}\left(\widetilde{u}^i(x,y)-\boldsymbol{\mu}_j\right)}. \tag{3.14}$$

3.8.1 Endocardium Segmentation

The region-based term is then defined as

$$\begin{aligned} J_r(\phi_i) = &- \int_{\Omega_1} \ln\left[p_1\left(\widetilde{u}^i_{\text{endo}}\right) \cdot \widetilde{P}^i_{\text{endo}}\right] dxdy \\ &- \int_{\Omega_2} \ln\left[p_2(\widetilde{u}^i_{\text{endo}}) \cdot \left(1 - \widetilde{P}^i_{\text{endo}}\right)\right] dxdy, \end{aligned} \tag{3.15}$$

where $\widetilde{P}^i_{\text{endo}}$ is the combined probabilistic map for the LV cavity shown in Fig. 3.21.

The temporal constraint term is defined as:

$$\begin{aligned} J_t(\phi_i) = \sum_{j\neq i} \Big\{ &- \int_{\Omega_1} \ln \left[T^{(j,i)}_{\text{endo}} \left(p_1\left(\widetilde{u}^j_{\text{endo}}\right) \cdot \widetilde{P}^j_{\text{endo}} \right) \right] dxdy \\ &- \int_{\Omega_2} \ln \left[T^{(j,i)}_{\text{endo}} \left(p_2\left(\widetilde{u}^j_{\text{endo}}\right) \cdot \left(1 - \widetilde{P}^j_{\text{endo}} \right) \right) \right] dxdy \Big\}. \end{aligned} \tag{3.16}$$

Finally, the level set evolution equation can be derived as:

$$\begin{aligned} \frac{d\phi_i}{dt} = \delta_\epsilon(\phi_i) \Big[&\lambda_r \ln \frac{p_1(\widetilde{u}^i_{\text{endo}}) \cdot \widetilde{P}^i_{\text{endo}}}{p_2(\widetilde{u}^i_{\text{endo}}) \cdot (1 - \widetilde{P}^i_{\text{endo}})} + \lambda_t \sum_{j\neq i} T^{(j,i)}_{\text{endo}} \left(\ln \frac{p_1(\widetilde{u}^j_{\text{endo}}) \cdot \widetilde{P}^j_{\text{endo}}}{p_2(\widetilde{u}^j_{\text{endo}}) \cdot (1 - \widetilde{P}^j_{\text{endo}})} \right) \\ &+ \lambda_e \text{div} \left(g_i \frac{\nabla \phi_i}{|\nabla \phi_i|} \right) \Big]. \end{aligned} \tag{3.17}$$

3.8.2 Epicardium Segmentation

Unlike the segmentation of the endocardium, the region-based term in epicardium segmentation incorporates the combined priors by applying a spatially varying weight $\omega^i(x, y) = 1 - \widetilde{P}^i_{\text{epi}}(x, y)$, where $\widetilde{P}^i_{\text{epi}}$ is the combined probabilistic map shown in Fig. 3.21. Therefore, the level set evolution equation is defined as follows:

$$\begin{aligned} \frac{d\phi_i}{dt} = \delta_\epsilon(\phi_i) \Big[&\lambda_r \omega^i \ln \frac{p_1(\widetilde{u}^i_{\text{epi}})}{p_2(\widetilde{u}^i_{\text{epi}})} + \lambda_t \sum_{j\neq i} T^{(j,i)}_{\text{epi}} \left(\omega^j \ln \frac{p_1(\widetilde{u}^j_{\text{epi}})}{p_2(\widetilde{u}^j_{\text{epi}})} \right) \\ &+ \lambda_e \text{div} \left(g_i \frac{\nabla \phi_i}{|\nabla \phi_i|} \right) \Big]. \end{aligned} \tag{3.18}$$

Chapter 4

Results & Discussion

4.1 Material

4.1.1 Study Population

Altogether 8 sets of 3D+t MRI data provided by the Pittsburgh NMR Center for Biomedical Research (Pittsburgh, PA, USA) were used in our experiments. These datasets were obtained by taking MR scans on a group of rats consisting of 4 rats with native heart and the other 4 with transplanted heart. As larger rats have more fatty tissue that can cause difficulties in surgical dissection, all rats used were 8-10 weeks of age and weighed between 250 and 300 g.

4.1.2 Transplantation Model

Unlike normal clinical practice, heterotopic heart and lung transplantation models were chosen for MRI studies. In the heterotopic heart and lung transplantation model, in addition to keeping in place the native heart and lung, the recipient rat receives another heart and lung located outside the chest. The reasons for the heterotopic transplantation models are twofold. First, this enables studies of the entire rejection process without many physiologic alterations of a transplanted animal, because the heart and lung grafts do not have a life-supporting function. Second, a cardiopulmonary bypass system is not available for rodents, so orthotopic heart transplantation is not feasible at the present time. The total ischemic time for the transplant surgery is about 30 min.

4.1.3 Image Acquisition

All eight MRI datasets used in our experiments were acquired by a Brucker AVANCE DRX 4.7 Tesla system. The MRI protocol has the following parameters: TR = one cardiac cycle (about 180 ms); TE = 8-10 ms; NEX = 4; flip angle = 90°; field-of-view = 3-4 cm; slice thickness = 1-1.5 mm; in-plane resolution = 117-156 μm; 4D image data resolution = 256×256×10×10 pixel.

4.1.4 The Gold Standard

Manual segmentation results provided by an experienced research scientist from the Pittsburgh NMR Center for Biomedical Research were used as the "Gold Standard" in evaluating the performance of the proposed method. There are in total 101 images with manual segmentation results, of which 76 are images of native rat hearts and 25 are images of transplanted rat hearts.

4.2 Qualitative Analysis

4.2.1 Agreement With Image Features

In Fig. 4.1, we qualitatively compare the segmentation results obtained by various automated methods and by manual segmentation. Cropped original images are shown in the top row, in which the first three images are samples of native rat heart MRI and the remaining three are MR images of transplanted rat hearts. From the second to the fifth row, each row shows, respectively the segmentation results automatically generated by methods using level sets without prior information, with elliptical shape prior, with data-driven priors introduced in our previous publication [6], and with both data-driven priors and temporal correlations as proposed in this thesis. The last row shows the "gold standard" obtained through manual segmentation by experts.

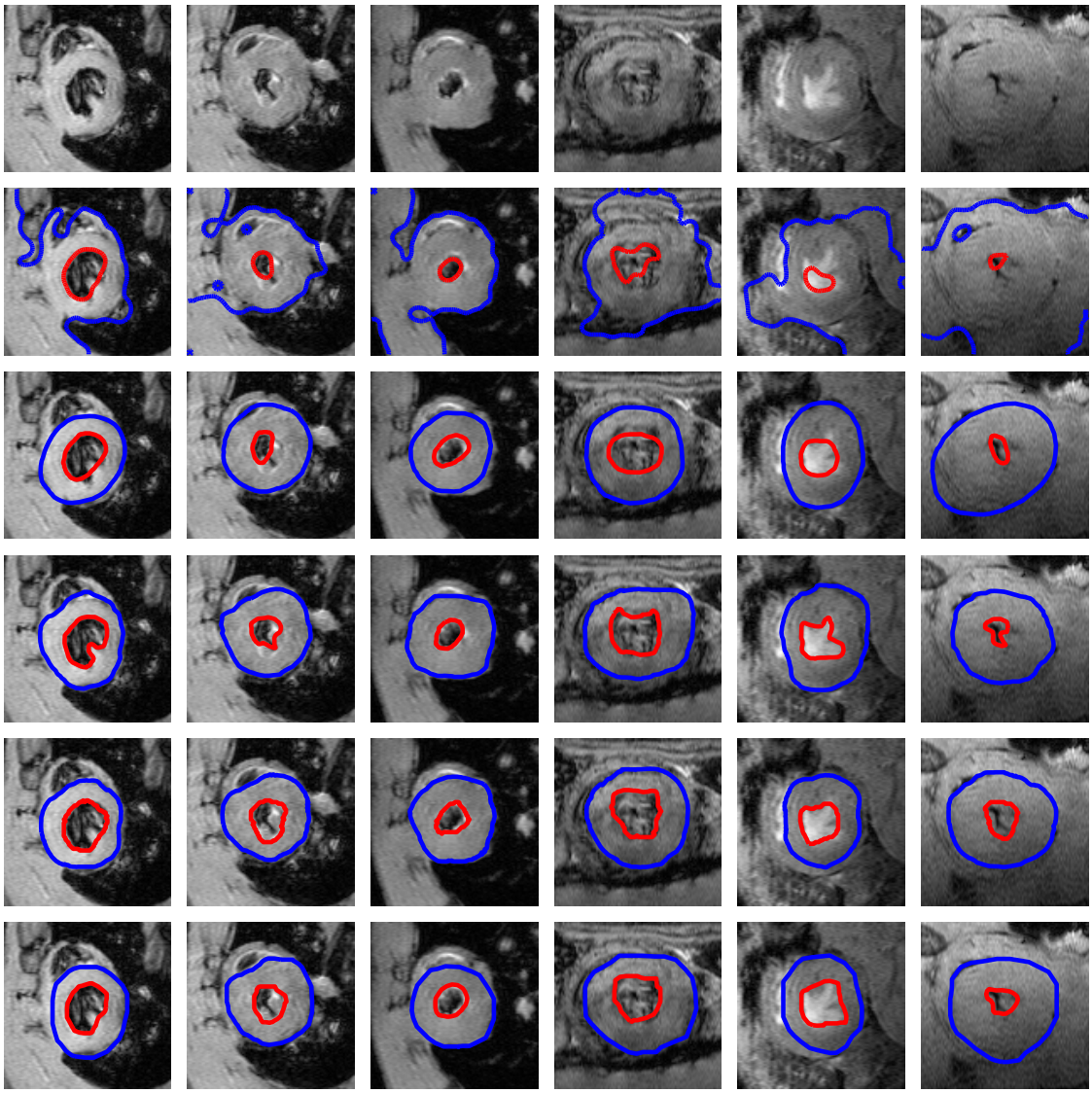

Figure 4.1: Agreement with image features of segmentation results

One can observe that when no prior information is adopted, leaking occurs. It is resulted from the fact that in rat MR images, "object" and "background" sometimes carry similar image information and cannot be successfully discriminated by low level image features. Therefore, spatial constraints need to be applied. Results in the third row show that elliptical shape prior successfully avoids leaking and

prevents the contour from having a random shape. Though shape prior enhances robustness, it does compromise the segmentation accuracy in cases when the endocardium or epicardium has a very different shape compared to the prior. Examples can be found in the last two images of the third row.

Different from normal shape priors, the data-driven priors regulate the level set evolution according to representative image features extracted from the image itself, therefore the segmentation accuracy is not compromised by inappropriate assumptions about the image. Resulting segmentations obtained by exerting data-driven priors (see the fourth row in Fig. 4.1) show better agreement with image features compared to the ones in the third row. As observed from the segmentation results, one limitation of utilizing only data-driven priors is that sometimes the segmented endocardium does not include the papillary muscles, which should be enclosed by the endocardium contour according to expert's explanation. Our proposed method incorporating both data-driven priors and temporal correlations successfully overcomes this limitation (see the fifth row in Fig. 4.1). As shown in the last two rows, the endocardial and epicardial boundaries detected by the propose method in the fifth row are very close to their manual counterparts in the last row.

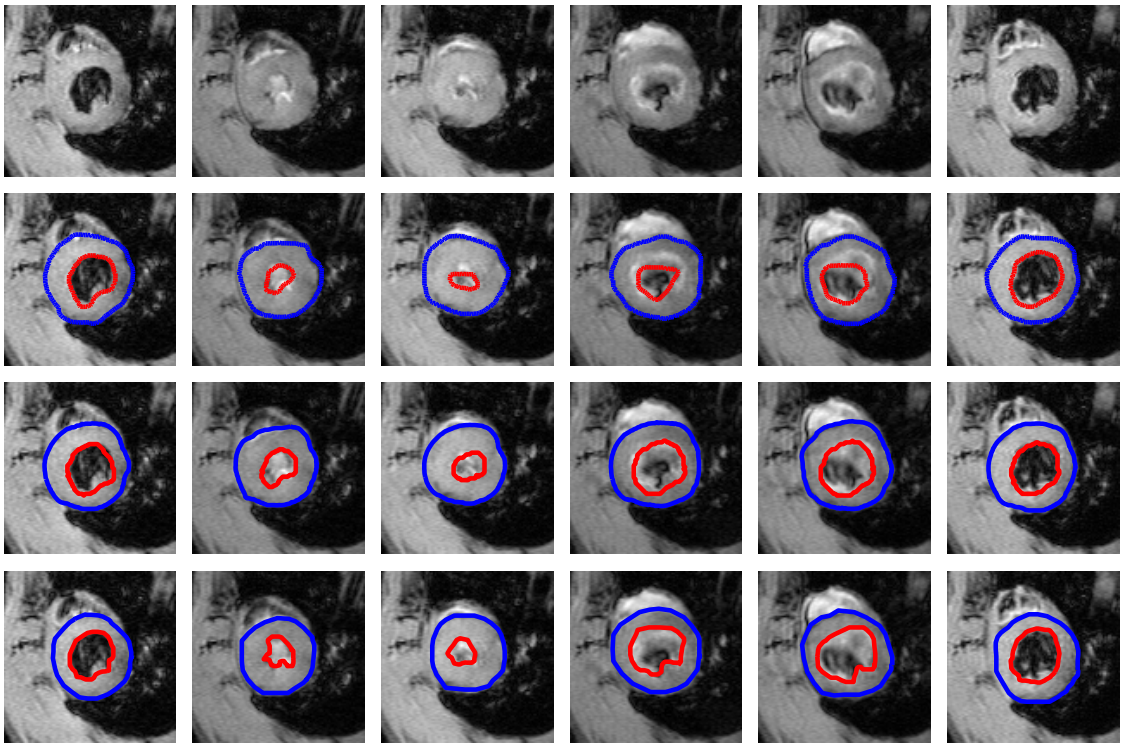

Figure 4.2: Comparison of temporal consistency of segmentation results

4.2.2 Temporal Consistency

We compare the temporal consistency of segmentation results obtained by different methods in Fig. 4.2. The first row displays a slice of native rat heart MR images acquired at different time instances in a cardiac cycle; and we show corresponding myocardium boundaries detected using the algorithm described in [6], the proposed method in this thesis, and expert's manual segmentation in the second, third and fourth row, respectively.

Generally, the myocardium motion is smooth. As a result, the myocardium boundaries in adjacent frames should have similar appearance but with different scales

to reflect the contraction or expansion of the myocardial wall. Observed from the second row of Fig. 4.2, epicardium and endocardium boundaries in neighboring frames have quite different shapes. This reveals that without considering point correspondence between adjacent frames, the method introduced in [6] fails in maintaining temporal consistency of the segmentation.

On the contrary, the myocardium boundaries obtained by the proposed method in this thesis reflect a smooth motion of the myocardial wall. By taking account of the complementary information present in neighboring frames, the proposed method effectively resolves the segmentation ambiguity discussed in the beginning of Section 3.2. Compared with the manual segmentations, the myocardium boundaries detected by the proposed method actually exhibit better consistency as far as temporal smoothness is concerned. Indeed, even for the experts, it is very hard to manually delineate contours by considering both image feature from the current frame and contour point correspondences in neighboring frames at the same time. This also explains the existence of the intra-observer variation, which refers to the segmentation variation encountered by the same expert in an attempt to segment the same image more than once. Therefore, the proposed method outperforms manual segmentation as it maintains temporal consistency of the segmentation and avoids intra- and inter-observer variations.

4.3 Quantitative Analysis

To quantitatively evaluate the accuracy of the proposed segmentation method against manual segmentation, we measure area similarities and shape similarities.

Table 4.1: Area similarity

		Endocardium		Epicardium		Myocardium	
Dataset	# of images	Mean	STD	Mean	STD	Mean	STD
1	34	0.8609	0.0302	0.8824	0.0201	0.8134	0.0628
2	26	0.8812	0.0248	0.8603	0.0212	0.8245	0.0560
3	10	0.8787	0.0180	0.8387	0.0207	0.8641	0.0486
4	6	0.8757	0.0225	0.8644	0.0100	0.8474	0.0313
5	20	0.8702	0.0221	0.9280	0.0220	0.8414	0.0314
6	3	0.8614	0.0110	0.8348	0.0109	0.8762	0.0603
7	1	0.8927	0	0.8350	0	0.8963	0
8	1	0.8665	0	0.8318	0	0.8843	0
Total	101	0.8710	0.0261	0.8779	0.0351	0.8322	0.0548

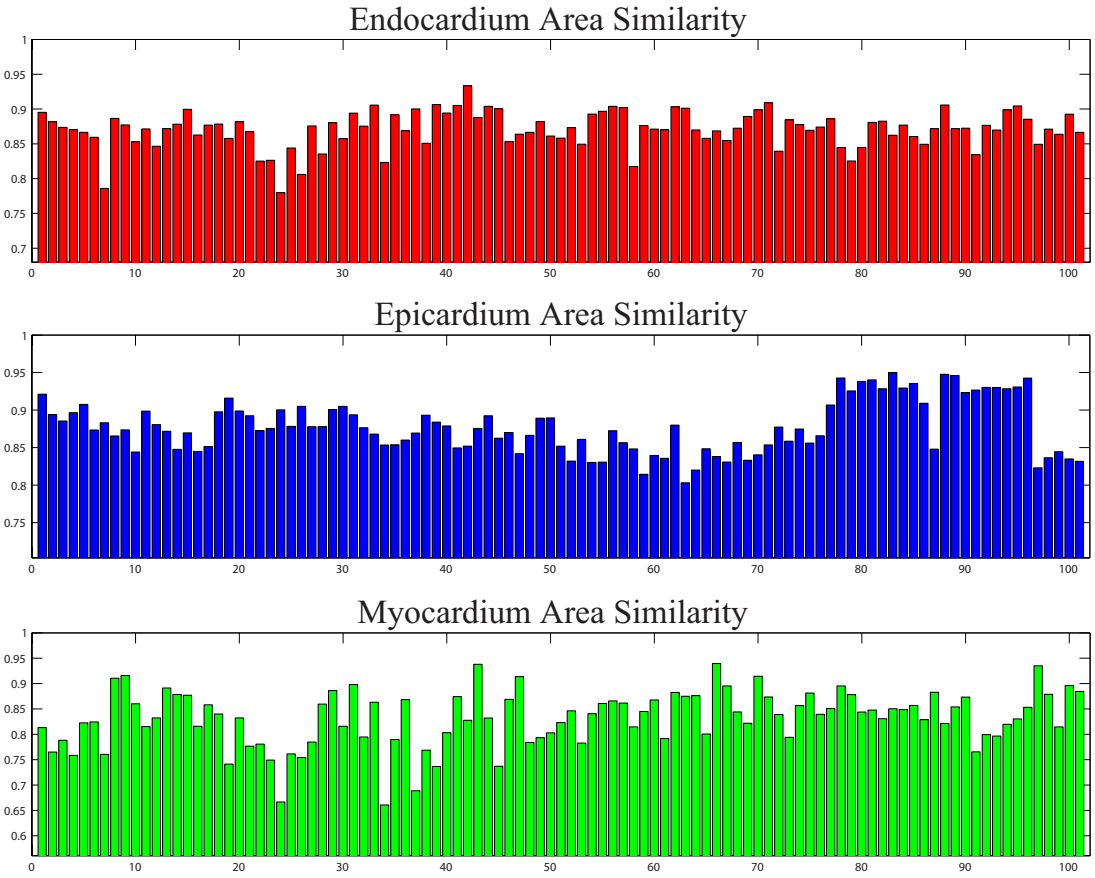

Figure 4.3: Area Similarity

4.3.1 Area Similarity

We first measure the area similarities, S_{area}, between the ROI masks (generated from endocardium, epicardium and myocardium boundaries) obtained by the proposed method and the corresponding masks by manual segmentation.

Here S_{area} follows the same definition as in [4]. The area similarity of myocardium is defined as follows:

$$S_{\text{area}} = \frac{2n(A_1 \bigwedge A_2)}{n(A_1) + n(A_2)}, \tag{4.1}$$

where A_1 and A_2 are binary images whose "on" pixels represent the regions of the segmented object, $\bigwedge$ is the element-wise "and" operator, and n represents the cardinality of A or the number of "on" pixels in the binary image A.

For the eight studies tested in our experiments, the distributions of S_{area} for endocardium, epicardium, and myocardium have mean values of 0.871, 0.8779, and 0.8322 with standard deviations 0.0261, 0.0351, and 0.0548, respectively. Details are provided in Table 4.1 and Fig. 4.3. These results show that the proposed method has similar performances compared to those reported in [4] and [22]. Although area similarity values are not significantly improved, considering the fact that the methods presented in [4] and [22] only work on images of native rat hearts, the proposed method has expanded the capability for reliable segmentation of transplanted rat heart images, which are much more difficult to segment than the native ones.

4.3.2 Shape Similarity

Shape similarity measures the difference in local orientation between two different segmentations. Different from area similarity, shape similarity is more sensitive to local variations in the object shape. By measuring the shape similarity, we further evaluate the performance of our segmentation method.

According to [4], steps in calculating the shape similarity measure is illustrated in

Fig. 4.4. Let C_1 and C_2 be two contours. The contour C_1 is the set of coordinates of the reference contour, or the gold standard contour, and the contour C_2 is the set of coordinates of the contour obtained automatically by the proposed method, which we will call the automatic contour. The goal is to find a similarity measure $S_{\text{shape}} \in [0, 1]$ that quantitatively assesses how similar the shape of the two contours C_1 and C_2 are. To determine the shape similarity measure, the first step is to generate the binary edge maps E_1 and E_2, where the "on" pixels represent the pixels on each of the two contours being compared. Then the shape of the contours in each binary edge map is propagated by applying the signed Euclidean distance transform

$$\mathcal{D}(x,y) = \begin{cases} -\min\limits_{(i,j)\in C} \sqrt{(x-i)^2 + (y-j)^2}, & \text{if}(x,y) \in \Omega_1 \\ \min\limits_{(i,j)\in C} \sqrt{(x-i)^2 + (y-j)^2}, & \text{if}(x,y) \in \Omega_2 \end{cases}, \tag{4.2}$$

where (x, y) represents the pixels in the image domain, $(i, j) \in C$ represents the pixels on the contour C, and Ω_1 and Ω_2 are sets of pixels inside and outside contour C, respectively. Applying the signed Euclidean distance transform in (4.2) to the binary edge maps E_1 and E_2, the corresponding distance maps, $\mathcal{D}_2$ and $\mathcal{D}_2$, can be obtained respectively. These distance maps simply contain the scaling replicas of the contour shapes, represented in different level sets, throughout the image domain.

In the third step, we calculate the corresponding phase maps by taking the inverse

tangent of the ratio of the gradient components in each distance map, i.e.,

$$\Phi_i(x,y) = \tan^{-1}\frac{\nabla_y \mathcal{D}_i(x,y)}{\nabla_x \mathcal{D}_i(x,y)} \quad \text{for} \quad i = 1, 2, \tag{4.3}$$

where $\nabla_x \mathcal{D}_i$ and $\nabla_y \mathcal{D}_i$ represent the x and y components of the gradient of the distance map $\mathcal{D}_i$, respectively.

In the fourth step, the normalized phase similarity between the two contours are computed according to

$$S_{\text{phase}} = \frac{\Phi_1 - \Phi_2 - \pi}{\pi}. \tag{4.4}$$

The index S_{phase} takes values in $[0, 1]$. A value of 1 for S_{phase} indicates that the contours have the same phase and a value of 0 refers to the maximum phase difference of π.

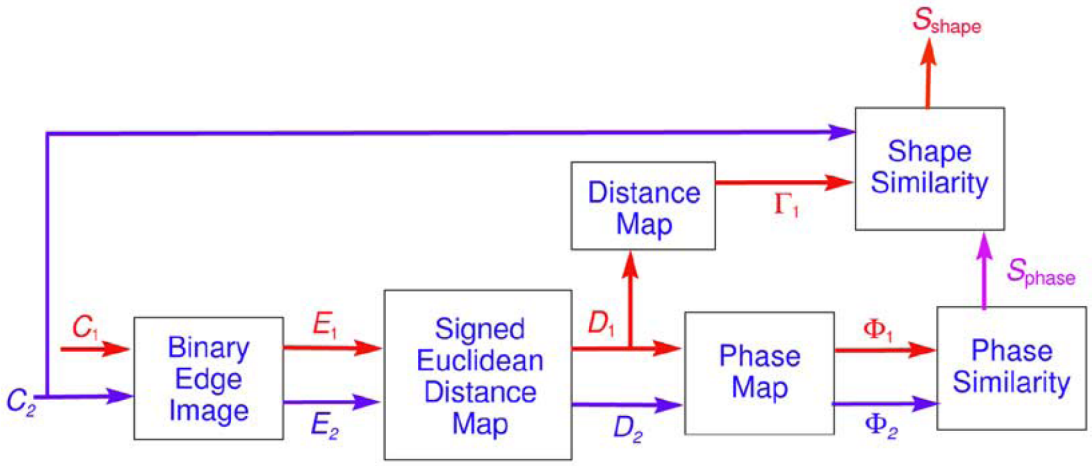

Figure 4.4: Flowchart for calculating shape similarity measure[3]

[3]Figure taken from "STACS: New active contour scheme for cardiac MR image segmentation" [4]

In the final step, the shape similarity is measured by taking the weighted sum of the phase similarity measure along C_2 (the automatic contour) against C_1 (the reference contour), i.e.,

$$S_{\text{shape}} = \frac{1}{n(C_2)} \sum_{(x,y)\in C_2} \Gamma_1(x,y) S_{\text{phase}}(x,y), \tag{4.5}$$

where C_2 is the set of pixels on the automatic contour, $n(C_2)$ denotes the cardinality of C_2, or the number of pixels on the contour C_2, and $\Gamma_1(x,y) \in [0,1]$ is derived from $\mathcal{D}_1$, the distance map of the reference contour, as

$$\Gamma_1(x,y) = \exp\left\{-\frac{\mathcal{D}_1^2(x,y)}{\sigma^2}\right\}, \tag{4.6}$$

where σ^2 is a positive constant.

For the eight studies tested in our experiments, distribution of S_{shape} for endocardium and epicardium has a mean value of 0.8316 and 0.9119 with standard deviation 0.0814 and 0.0278, respectively. Details are provided in Table 4.2 and Fig. 4.5. These results show a significant improvement compared to those have been reported in [4] and [22], despite the fact that images used in [4] and [22] are all MRI of native rat hearts, which are much easier to segment.

4.4 Discussion

As segmentation results obtained by the proposed method achieve 0.8 to 0.9 average area similarity and shape similarity with very small standard deviations, we

Table 4.2: Shape similarity

		Endocardium		Epicardium	
Dataset	# of images	Mean	STD	Mean	STD
1	34	0.8253	0.0822	0.9065	0.0273
2	26	0.8621	0.0784	0.9089	0.0302
3	10	0.8548	0.0557	0.9355	0.0247
4	6	0.8399	0.0884	0.9202	0.0241
5	20	0.8032	0.0698	0.9063	0.0244
6	3	0.7585	0.1700	0.9291	0.0209
7	1	0.7627	0	0.9255	0
8	1	0.8303	0	0.9326	0
Total	101	0.8316	0.0814	0.9119	0.0278

conclude that our algorithm consistently produces accurate segmentation results.

Regarding the area similarity measure, we observe that epicardium area similarity is normally greater than endocardium and myocardium area similarities, and endocardium similarity is generally greater than myocardium similarity. The reason is quite straightforward: the epicardium contour has the largest area, therefore without normalization, epicardium area similarity is generally the greatest; the endocardium contour has a smaller area compared to the epicardium contour,

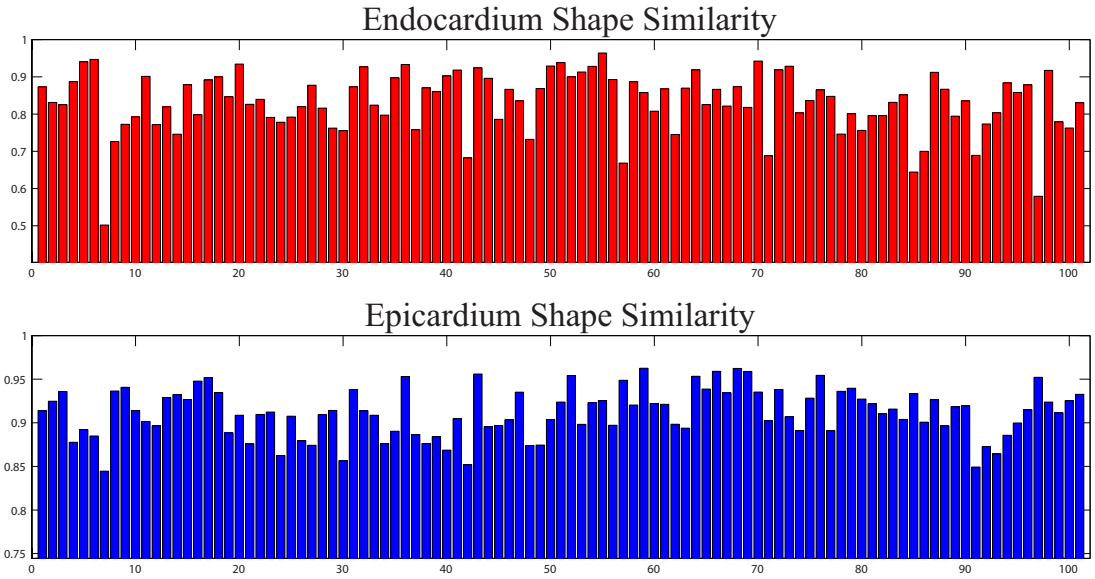

Figure 4.5: Shape Similarity

therefore the endocardium area similarity has smaller values as it is more sensitive to segmentation error; due to the fact that myocardium is determined by the combination of endocardium and epicardium, myocardium has accumulated segmentation error, therefore lower area similarity.

Interestingly, although the area similarity values and the shape similarity values are not directly comparable, we do observe that for endocardium, area similarity (0.871 $\pm$ 0.0261) is greater than shape similarity (0.8316 $\pm$ 0.0814), but for epicardium, area similarity (0.8779 $\pm$ 0.0351) is smaller than shape similarity (0.9119 $\pm$ 0.0278). One reasonable explanation is that compared to area similarity, the shape similarity is more sensitive to the contour size. As contour shape is extremely sensitive to segmentation error when the contour size is small, we observe epicardium shape

similarity is always greater than endocardium shape similarity.

Our experiments are performed in the MATLAB environment on a workstation with a quad core CPU running at 3.0GHz. The average processing time is about 4 minutes per slice, which consists of 10 two-dimensional MR images. As current experiments are primarily designed for the purpose of feasibility test, our method was not implemented to optimize the computational efficiency. The processing time can be significantly reduced if the proposed method is implemented in a more efficient programming language, and at the same time, multi-threading is enabled to take advantage of the multi-core processor system.

Chapter 5

Conclusion & Future Work

5.1 Conclusion

In this thesis, we introduced a novel method for the segmentation of LV myocardium in transplanted rat cardiac MRI utilizing data-driven priors and temporal correlations.

Different from normal shape priors which are used to enhance segmentation robustness by penalizing unlikely contour shapes, the data-driven priors introduced in this thesis improves both accuracy and robustness of the segmentation by providing reliable information that is extracted from the the image itself and complementary to low level image features. The essential difference between the proposed method and other automated segmentation methods utilizing prior knowledge is how the

prior knowledge is interpreted as a cost in the segmentation energy formulation: most methods build certain type of model according to some prior knowledge, descriptive or statistical, and define a cost based on the degree of geometrical or topological agreement between the current contour and the model; however, the proposed method extracts prominent image features from the image and uses the features that conform with descriptive prior knowledge to generate prior maps, which are applied as confidence maps to spatially bias the segmentation cost function.

Compared with traditional ways of incorporating prior knowledge into segmentation, the data-driven prior shows its superiority in twofold: first, based on very general descriptive prior knowledge, the proposed method automatically generates prior maps for any image, therefore it avoids the tedious training or modeling process; second, the proposed method only incorporates prior knowledge that coincides with prominent image features into the segmentation, therefore the segmentation results are not compromised by the fact that prior knowledge is only a piece of very general and observational information, and some of it can be wrong on a particular image.

To resolve the problem of segmentation ambiguity caused by the lack of discriminative image feature on a particular frame of image, complementary information from neighboring frames is incorporated into the segmentation. Point-to-point correspondence between pixels on neighboring frames is constructed through image

registration. To reduce registration error due to the inconsistent intensity distribution in original MR images, feature maps are calculated and registered to obtain deformation fields, which are applied not only on channels in the feature space, but also on data-driven priors to form point correspondence between the frame of interest and its neighboring frames.

Experimental results show that myocardium contours obtained by the proposed method exhibit excellent agreement with image features and the gold standard. At the same time, smooth myocardial wall motion reflected by the automatically segmented contours reveals that the proposed method largely resolves the problem of segmentation ambiguity and effectively preserves the temporal consistency of the segmentation. Undoubtedly, the method proposed in this thesis has improved performance over the one we previously presented in [6].

Being able to produce reliable and accurate myocardium segmentation results, the proposed method not only significantly reduces the cost and processing time compared to manual segmentation, but also successfully circumvents intra- and inter- observer variations.

5.2 Future Work

Segmentation accuracy can be further enhanced by incorporating spatial smoothness constraints. As myocardium in the apical or basal slice normally has diminishing size and irregular shape, it is true that MR images in apical and basal slices are more difficult to segment than the ones in middle slices. Under the assumption that the heart has a spatially smooth anatomic structure, it is reasonable to constrain the smoothness of the LV myocardium surface in the segmentation. Formulated in the level set framework, the proposed method can be easily extended for 3D volume segmentation. In our future work, we aim to refine the segmentation accuracy, especially in the apical and basal regions, by incorporating myocardium surface smoothness constraints.

Bibliography

[1] D. J. Stuckey, C. A. Carr, D. J. Tyler, E. Aasum, and K. Clarke, "A novel MRI method to detect altered left ventricular ejection and filling patterns in rodent models of disease," *Magnetic Resonance in Medicine*, vol. 60, pp. 582–587, Aug 2008.

[2] Y.-J. L. Wu, K. Sato, Q. Ye, and C. Ho, "MRI investigations of graft rejection following organ transplantation using rodent models," *Methods Enzymol*, vol. 386, pp. 73–105, 2004.

[3] D. L. Pham, C. Xu, and J. L. Prince, "A survey of current methods in medical image segmentation," *Annual Review of Biomedical Engineering*, vol. 2, pp. 315–338, 2000.

[4] C. Pluempitiwiriyawej, J. M. F. Moura, Y.-J. L. Wu, and C. Ho, "STACS: New active contour scheme for cardiac MR image segmentation," *IEEE Trans. Med. Imag.*, vol. 24, pp. 593–603, May 2005.

[5] J. Schaerer, Y. R. P. Clarysse, B. Hiba, P. Croisille, J. Pousin, and I. E. Magnin, "Simultaneous segmentation of the left and right heart ventricles in 3D cine MR images of small animals," *Computers in Cardiology*, vol. 32, pp. 231–234, 2005.

[6] X. Jia, C. Li, Y. Sun, A. A. Kassim, Y. L. Wu, T. K. Hitchens, and C. Ho, "A data-driven approach to prior extraction for segmentation of left ventricle in cardiac MR images," in *Proc. IEEE International Symposium on Biomedical Imaging(ISBI)*, pp. 831–834, Jul 2009.

[7] M. Rousson, T. Brox, and R. Deriche, "Active unsupervised texture segmentation on a diffusion based feature space," in *Proc. IEEE Conf. on Comp. Vis. Patt. Recog. (CVPR)*, vol. 2, pp. 699–704, Jun 2003.

[8] M. Kass, A. Witikin, and D. Terzopoulos, "Snakes: active contour models," *Int. Journal of Computer vision*, vol. 1, pp. 321–331, Jan 1988.

[9] C. Davatzikos and J. L. Prince, "An active contour model for mapping the cortex," *IEEE Trans. Med. Imag.*, vol. 14, pp. 65–80, Mar 1995.

[10] C. Xu and J. Prince, "Gradient vector flow: a new external force for snakes," in *Proc. IEEE Conf. on Comp. Vis. Patt. Recog. (CVPR)*, pp. 66–71, Jun 1997.

[11] S. Kichenassamy, A. Kumar, P. Oliver, A. Tannenbaum, and A. Yezzi, "Gradient flows and geometric active contours models," in *Proc. International Conference on Computer Vision (ICCV)*, pp. 810–815, 1995.

[12] S. Osher and J. A. Sethian, "Fronts propagating with curvature dependent speed: algorithms based on Hamiton-Jacobi formulations," *Journal of Computer physics*, vol. 79, no. 1, pp. 12–49, 1988.

[13] J. Sethian, "A fast marching level set method for monotonically advancing fronts," in *Proc. National Academy of Sciences*, vol. 93, pp. 1591–1595, Sep 1996.

[14] C. Li, C. Xu, C. Gui, and D. Fox, "Level set evolution without re-initialization: A new variational formulation," in *Proc. IEEE Conf. on Comp. Vis. Patt. Recog. (CVPR)*, vol. 1, pp. 430–436, 2005.

[15] V. I. Arnold, *Geometrical methods in the theory of ordinary differential equations*. Springer, second ed., 1988.

[16] L. Evans, *Partial differential equations*. American Mathematical Society, 1998.

[17] J. Bigun, G. H. Granlund, and J. Wiklund, "Multidimensional orientation estimation with applications to texture analysis and optical flow," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, pp. 775–790, Aug 1991.

[18] T. Brox and J. Weickert, "Nonlinear matrix diffusion for optic flow estimation," in *Proc. 24th DAGM Symposium, Lecture Notes in Computer Science*, vol. 2449, pp. 446–453, Sept 2002.

[19] D. Tschumperle and R. Deriche, "Diffusion tensor regularization with constraints preservation," in *Proc. IEEE Conf. on Comp. Vis. Patt. Recog. (CVPR)*, vol. 1, pp. 948–953, Dec 2001.

[20] M. Rousson and R. Deriche, "A variational framework for active and adaptative segmentation of vector valued images," in *Proc. IEEE Workshop on Motion and Video Computing*, pp. 56–61, Dec 2002.

[21] P. Perona and J. Malik, "Scale space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 629–639, 1990.

[22] T. Chen, J. Babb, P. Kellman, L. Axel, and D. Kim, "Semi-automated segmentation of myocardial contours for fast strain analysis in cine displacement-encoded MRI," *IEEE Trans. Med. Imag.*, vol. 27, pp. 1084–1094, Aug 2008.

[23] M. Rousson and N. Paragios, "Prior knowledge, level set representations and visual grouping," *International Journal of Computer Vision*, vol. 76, pp. 231–243, Mar 2008.

[24] G. E. Christensen, R. D. Rabbitt, and M. I. Miller, "Deformable templates using large deformation kinematics," *IEEE Transactions on Image Processing*, vol. 5, pp. 1435–1447, 1996.

[25] B. Glocker, N. Komodakis, G. Tziritas, N. Navab, and N. Paragios, "Dense image registration through MRFs and efficient linear programming," *Medical Image Analysis*, vol. 12, pp. 731–741, Dec 2008.

[26] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: application to breast MR Images," *IEEE Transactions on Medical Imaging*, vol. 18, pp. 712–721, Aug 1999.

[27] C. K. Hoh, M. Dahlbom, G. Harris, Y. Choi, R. A. Hawkins, M. E.Philps, and J. Maddahi, "Automated iterative three-dimensional registration of positron emission tomography images," *J. Nucl. Med.*, vol. 34, no. 11, pp. 2009–2018, 1993.

[28] R. Kim, T. Aw, S. Bacharach, and R. Bonow, "Correlation of cardiac MRI and PET images using lung cavities as landmarks," in *Proc. IEEE Conf. Computers in Cardiology*, pp. 49–52, 1991.

[29] P. Slomka, D. Dey, C. Przetak, and R. Baum, "Automated 3-D spatial integration of 18-F FDG wholebody PET with CT," *Journal of Nuclear Medicine*, vol. 41, no. 6, p. 59, 2000.

[30] J. P. Thirion, "Image matching as a diffusion process: an analogy with maxwell's demons," *Medical Image Analysis*, vol. 2, no. 3, pp. 243–260, 1998.

[31] F. Wang and B. C. Vemuri, "Simultaneous registration and segmentation of anatomical structures from brain MRI," in *Proc. Medical Image Computing and Computer-Assisted Intervention - MICCAI 2005*, vol. 3749, pp. 17–25, 2005.

[32] G. Unal and G. Slabaugh, "Coupled PDEs for non-rigid registration and segmentation," in *Proc. IEEE Conf. on Comp. Vis. Patt. Recog. (CVPR)*, vol. 1, pp. 168–175, Jul 2005.

[33] T. Makela, P. Clarysse, O. Sipila, N. Pauna, Q. Pham, T. Katila, and I. Magnin, "A review of cardiac image registration methods," *IEEE Trans. Med. Imag.*, vol. 21, pp. 1011–1021, 2002.

[34] L. Axel, J. Costantini, and J. Listerud, "Intensity correction in surface coil MR imaging," *American Journal of Radiology*, vol. 148, pp. 418–420, Feb 1987.

[35] E. Rosten and T.Drummond, "Fusing points and lines for high performance tracking," in *Proc. International Conference on Computer Vision (ICCV)*, vol. 2, pp. 1508–1515, Nov 2005.

[36] E. Rosten and T.Drummond, "Machine learning for high speed corner detection," in *Proc. European Conference on Computer Vision (ECCV)*, vol. 1, pp. 430–443, May 2006.

[37] I. Kokkinos, P.Maragos, and A. Yuille, "Bottom-up and top-down object detection using primal sketch features and graphical models," in *Proc. IEEE Conf. on Comp. Vis. Patt. Recog. (CVPR)*, vol. 2, pp. 1893–1900, Jun 2006.

[38] I. Kokkinos, P.Maragos, and A. Yuille, "Unsupervised learning of object deformation models," in *Proc. International Conference on Computer Vision (ICCV)*, pp. 1–8, Oct 2007.