Published in final edited form as:

Comput Methods Biomech Biomed Eng Imaging Vis. 2019 ; 7(5-6): 690–706. doi:
10.1080/21681163.2019.1583607.

Integrated 3D Anatomical Model for Automatic Myocardial Segmentation in Cardiac CT Imagery

N. Dahiya^{a,*}, A. Yezzi^a, M. Piccinelli^b, E. Garcia^b

^aGeorgia Institute of Technology, North Ave NW, Atlanta, GA 30332, USA

^bEmory University School of Medicine, 101 Woodruff Circle, Atlanta, GA, 30322, USA

Abstract

Segmentation of epicardial and endocardial boundaries is a critical step in diagnosing cardiovascular function in heart patients. The manual tracing of organ contours in Computed Tomography Angiography (CTA) slices is subjective, time-consuming and impractical in clinical setting. We propose a novel multi-dimensional automatic edge detection algorithm based on shape priors and principal component analysis (PCA). We have developed a highly customized parametric model for implicit representations of segmenting curves (3D) for Left Ventricle (LV), Right Ventricle (RV), and Epicardium (Epi) used simultaneously to achieve myocardial segmentation. We have combined these representations in a region-based image modeling framework with high level constraints enabling the modeling of complex cardiac anatomical structures to automatically guide the segmentation of endo/epicardial boundaries. Test results on 30 short-axis CTA datasets show robust segmentation with error (mean $\pm$ std mm) of (1.46 ± 0.41) , (2.06 ± 0.65) , (2.88 ± 0.59) for LV, RV and Epi respectively.

Keywords

myocardial segmentation; shape analysis; principal component analysis; active contours; cardiac tomographic angiography (CTA)

1. Introduction

In 2015, the American Heart Association (AHA) reported (CDC 2015) that heart disease is the No. 1 cause of death in the United States, killing nearly 787,000 people in 2011 alone. Among various types of heart diseases, coronary artery disease is the most common, killing nearly 380,000 people annually. Further, the direct and indirect costs of heart disease total more than \$320 billion. Consequently, early and cost-effective diagnosis of heart diseases is an important requirement for clinical practitioners worldwide.

Cardiac Tomography Angiography (CTA), a non-invasive imaging technique, is becoming increasingly popular for cardiac examination, mainly due to its superior spatial resolution compared to MRI. This imaging modality is currently widely used for the diagnosis of Coronary Artery Disease (CAD) but it is not commonly used for the diagnosis of ventricular

*Corresponding author. ndahiya3@gatech.edu.

and atrial function. The primary reason for that is the lack of autonomous image processing techniques for cardiac medical imagery. Medical imagery, in general, suffers from poor contrast, low signal to noise ratio (SNR), weak or missing edges or smearing of edges due to patient movement. These among other reasons make it difficult to apply standard image processing techniques to medical image processing tasks and obviate the need to incorporate as much prior information about the objects of interest as possible.

In our work, we use shape prior model-based implicit parametric representations of each anatomical structure of interest (LV, RV, epicardium), and calculate parameters of each segmenting curve via gradient descent procedure to minimize an energy functional designed for the segmentation task.

1.1 Relationship to prior work

To overcome the shortcomings of the image acquisition process and inherent complexity of the heart anatomy, many different classes of techniques have been proposed over the years to aid diagnosis and prognosis of cardiac diseases in various imaging modalities and data acquisition protocols. Model based segmentation methods have become particularly prominent in literature and many different formulations have been proposed with many including some form of weak or strong shape prior. Active Shape and Appearance Models (Cootes and Taylor 1992; Cootes et al. 1995; Cootes and Taylor 1998; Cootes et al. 2001) based techniques have been very influential in the field of medical image processing, and have been used for both LV and RV segmentation (van Assen et al. 2006, 2008; Mitchell et al. 2001). A further enhancement of model based techniques is in using deformable models (Weese et al. 2001) which provides additional flexibility by allowing the models to freely deform in response to local image features while restraining the global shape. Ecabert et al. (2008) used this technique to present a model-based approach for fully automatic segmentation of the whole heart (four chambers, myocardium and great vessels) in cardiac CT imagery. Their model consisted of multi-compartment triangulated mesh of vertices and triangles. The authors used a 3D Global Hough Transform for whole heart initialization and estimated global pose using similarity transformations. This model was then allowed to adapt to local individual patient anatomy using shape-constrained deformable models.

Atlas based segmentation is another class of techniques that have become increasingly popular in recent years. An atlas describes the different structures present in a given type of image. It can be generated by manually segmenting an image or by integrating information from multiple segmented images from different individuals. Given an atlas, an image can be segmented by mapping its coordinate space to that of the atlas using a registration process (Rohlfing et al. 2003). Initially applied to brain segmentation, it is now being widely used for cardiac segmentation as well. Recently Shahzad et al. (2017) presented a fully automatic segmentation framework to segment the whole heart, aortic root, left atrium, right atrium, LV, and RV including both the blood-pool and the myocardial tissue in the ventricles from non-contrast-enhanced calcium scoring CT scan. Based on the work of Kiri li et al. (2010), this work used a multi-atlas-based segmentation process to register multiple atlas scans with corresponding manually annotated labels to an unseen subject's scan. The registration

process used a mutual-information (Mattes et al. 2003) based similarity measure as a cost function to be minimized with respect to a set of image transformation parameters.

While most methods referenced here require some form of training images or prior information, Zhu et al. (2014) presented a training free method for extraction of LV myocardium from CT images. The authors used the blood pool surface to approximate the heart surface and automatically detected the LV by examining the distribution of levelsets starting from the LV apex and further refining it by utilizing a geometric active contour model (Kichenessamy et al. 1995; Caselles et al. 1997) on the blood pool surface. After locating the endocardial surface, the authors used a variational region growing method (Gao et al. 2012) to locate the epicardial surface. These two surface estimates are then refined using an active contour model with shape constraints and myocardium extracted from the voxels between these two surfaces. While this method was used to obtain only LV myocardium, Zhu et al. (2013a) used a similar strategy to segment both the ventricles and myocardium. The RV is identified on heart surface constructed by using the LV as a seed region for the variational region growing method.

Active contour models have been widely used in medical image segmentation because of their flexibility and robustness. Prior information can be incorporated as well to restrict the optimization space. Chen et al. (2001) used the concept of an average shape model within the framework of geometric active contour models. Staib and Duncan (1992) introduced a parametric point model based on an elliptic Fourier decomposition of the land-mark points. Chakraborty et al. (1994) extended this approach to a hybrid segmentation model that incorporates both gradient and region-homogeneity information. Leventon et al. 2000 proposed a model-based segmenter that incorporated shape information as a prior model to restrict the flow of the geodesic active contour (Yezzi et al. 1997; Caselles et al. 1997). Their prior parametric shape model is derived by performing PCA on a collection of signed distance maps of the training shape. The segmenting curve then evolves according to two competing forces: 1) the gradient force of the image, and 2) the force exerted by the estimated shape where the parameters of the shape are calculated based on the image gradients and the current position of the curve. In contrast to these edge-based active contour models which utilize image gradient to stop the evolving contours on the object boundaries, region-based models (Chan and Vese 1999), which minimizes the Mumford-Shah (Mumford and Shah 1989; Chan and Vese 2001) energy functional have shown to be more robust to noise and initial placement of contours. These methods are more global in nature and avoid taking image intensity derivatives.

Tsai et al. (2003) developed a parametric model using pose and shape parameters for segmentation, where they describe the segmenting curve as a linear combination of eigenvectors that are obtained by performing PCA on the variations from the mean shape. Vikram et al. (2010) later applied a variation of this technique which incorporated region of confidence (ROC) labels indicating the reliability of edges in certain object regions, to the task of myocardial segmentation in pigs cardiac MR datasets.

1.2 Our contribution

Vikram et al. (2010) and Tsai et al. (2003) use the chan-vese image model which models the image as piece-wise smooth function and the evolution of the segmenting curve depends upon the pixel intensities within entire regions. That is, region-based models regard an image as the composition of a finite number of regions and rely on regional statistics for segmentation. The statistics of entire regions (such as sample mean and variance) are used to direct the movement of the curve toward the boundaries of the image.

More formally, the Chan-Vese energy is defined as:

$$E_{cv} = \int_{R^u} (I - \mu)^2 d\mathbf{x} + \int_{R^v} (I - \nu)^2 d\mathbf{x} \quad (1)$$

where R^u and R^v are the regions inside and outside the segmenting curve. It has been shown that the optimal choice for the constants μ and ν are the region means. This model can be used either on single contour with two regions, inside and outside, or multiple regions. It can also be easily extended from 2D to 3D. In case of 3D model, the image statistics are calculated on entire volumes rather than regions.

We have customized these frameworks for the task of 3D Cardiac CT imagery by introducing several high level constraints and making use of shape priors for three anatomical regions, LV, RV, and Epi simultaneously. We have developed a Chan-Vese like model for the interior region of the Epicardium i.e. RV, LV and Myocardium. To customize the model for myocardial segmentation, we have derived a mathematical formulation for enforcing a strict ordering of the image statistics of different regions in the appearance model which makes the segmentation even more robust to initial contour placement.

Additionally, the region surrounding the epicardium in cardiac imagery is very complex with different types of tissues exhibiting different intensity ranges. This makes it unreasonable to model the background as a coherent region with a single mean intensity. Consequently, we have developed a dynamic adaptive background model. We model the background as a binary cluster, one cluster exhibiting very low and the other very high image intensities. We have derived and present the mathematical formulation for doing so.

Instead of using the model separately for each region¹ which doesn't allow the shapes to be constrained to not overlap with each other, we use all three priors simultaneously with high level constraints. Our model includes a coupling energy functional with an overlap penalty between distinct regions which effectively couples the three priors. The shape models we have developed are defined directly in 3D space instead of operating in a slice-by-slice manner.

Note that our binary clustering strategy for background region together with strict ordering of the image statistics within the four regions imposes more prior information on the known aspects of the image appearance than other methods. These same adaptations together with

¹Since our models are defined directly in 3D the term "region" means a volume.

the overlap constraints between regions however rendered the problem non-convex and made the optimization much more challenging compared to other models. We have successfully addressed this complication by using a relaxation of the overlap constraint which permits small amounts of overlap during the optimization process but not in the final result. This relaxation together with momentum based gradient descent has resulted in a robust and reliable optimization process on our models.

Our formulation also overcomes several shortcomings of some of the latest methods outlined in the previous section. In particular, the multi-atlas based method fails to respond to local image variations. Consequently, it fails to differentiate between LV endo and epicardium. Zhu et. al's region growing method uses generic moment based priors and does not take into account the expected ordering of the average appearance intensities between various regions being segmented. In contrast, our formulation is based on anatomically specific PCA based shape priors and region based image appearance models highly customized to contrast enhanced cardiac CT images. This customized appearance model combined with customized PCA based shape priors is able to adequately capture both global and local image information and outperforms the other approaches as shown in later sections.

The rest of the paper is organized as follows. Section 2 provides a high level overview of our segmentation models and approach. Section 3 describes a variational gradient-based approach to align all the training shapes in the database to eliminate variations in pose. Based on this aligned training set, we show in Section 4 the development of an implicit parametric representation of the segmenting curve using principal component analysis. Section 5 describes the region-based models we have developed for image segmentation. In section 6 we show the experimental results obtained on human CTA datasets and experimentally illustrate the salient features of our algorithmic framework. Finally, in section 7 we conclude with a summary and some possible future research directions of this work.

2. Approach

In collaboration with Nuclear Cardiology R&D Laboratory at Emory University School of Medicine, we have applied our algorithm for myocardial segmentation in a 30 patient 3D CTA study. We use part of the studies for the training phase of our algorithm in which we extract mean shapes and principal modes of shape variation of each of the 3 anatomical structures, namely LV, RV and Epi. The training phase is based on manual expert segmentation of the anatomical structures by experts at Emory University Hospital. The result of the manual segmentation process is a binary mask for each region which we then convert to a signed distance function (SDF) (Osher and Sethian 1988) where the zero level set of the SDF represents shape boundary. If we define the image domain to be Ω , the boundary of the binary shape as $\partial\Omega$ and $\mathbf{x}$ as any point in the image domain, then the signed distance function $\text{SDF}(\mathbf{x})$ is defined as follows:

$$\text{SDF}(\mathbf{x}) = \begin{cases} +d(\mathbf{x}, \partial\Omega), & \mathbf{x} \in \Omega^+ \\ -d(\mathbf{x}, \partial\Omega), & \mathbf{x} \in \Omega^- \end{cases} \quad (2)$$

where $d(\mathbf{x}, \Omega)$ is the euclidean distance from a point $\mathbf{x}$ to the nearest point on the shape boundary and Ω^+ and Ω^- are, respectively, the parts of the domain outside and inside the shape boundary. Fast Marching methods (Sethian 1996) provide very efficient ways of computing the SDF from a binary shape.

Since shape is defined to be invariant to Euclidean similarity transformations (Dryden and Mardia 1998), we first perform image alignment on the set of training shapes before proceeding further. Once we have a set of aligned training SDF's, we perform PCA to extract mean shape representations and principal modes. In the actual segmentation phase, we allow optimization of a set of free parameters, pose and shape, for each region. The shape parameters correspond to a set of weights in a linear combination of mean shape and its principal modes. This linear combination is then allowed to deform by a set of pose parameters including translation, rotation and scale.

The parameter optimization process is based on the gradients of a region based energy functional designed to achieve the task of myocardial segmentation. This image modeling functional is based on highly successful and well understood image modeling frameworks developed over the years. We have added several high level modifications to these models leading to a highly customized and integrated 3D anatomical model for automatic myocardial segmentation in full 3D Cardiac CT Imagery.

3. Image Alignment

The various images in the training set vary in size and orientation. We need to employ an alignment technique as a preprocessing step to allow us to capture shape variations in our database without interference from pose variations. There are many works dealing with image alignment problem (Dryden and Mardia 1998; Goodall 1991; Cootes and Taylor 2001). For our purposes, we approach the task of image alignment in variational framework.

Let the training set $\mathbb{T}$ consist of n 3D binary images $\{I^1, I^2, \dots, I^n\}$, each with values of one inside and zero outside. The goal is to calculate a set of pose parameters $\{\mathbf{p}^1, \mathbf{p}^2, \dots, \mathbf{p}^n\}$ used to jointly align the binary training images, and hence remove any variations in shape due to pose differences. We focus on using similarity transformations to align these binary images to each other. That is, in three dimensions, each set of pose parameters, $\mathbf{p}^i$, comprises of translation in x, y, z directions, uniform scale h in 3 dimensions, and three rotations, *yaw*, *pitch* and *roll*. We represent the translation as a vector with three components $T = [x \ y \ z]^T$ and scale as a single parameter h .

Space rotations have three degrees of freedom, and admit several ways to represent and operate with them (Murray et al. 1994). Each representation has advantages and disadvantages. In our work we chose to use exponential or twist coordinates, also known as Euler-Rodrigues parameters (Cheng and Gupta 1989; Murray et al. 1994, p. 33), a compact representation requiring only three real numbers. The Euler-Rodrigues formula (Murray et al. 1994, p. 28) states that the rotation matrix represented by $v \in \mathbb{R}^3$ is given by:

$$R(v) = \cos(\|v\|)I + \sin(\|v\|)\left[\frac{v}{\|v\|}\right]_{\times} + (1 - \cos(\|v\|))\left[\frac{v}{\|v\|}\right]\left[\frac{v}{\|v\|}\right]^T \quad (3)$$

where v is the axis of rotation and its magnitude is the amount of rotation in radians and

$$[a]_{\times} := \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix} \in \text{Skew}_3 \quad (4)$$

is the cross-product (skew-symmetric) matrix such that $[a]_{\times}b = a \times b$, for all $a = (a_1 \ a_2 \ a_3)^T, b \in \mathbb{R}^3$.

The transformed image of I , based on pose parameters, $\mathbf{p}$, is defined as:

$$\hat{I}(\hat{\mathbf{x}}) = I(\mathbf{x}) \quad (5)$$

where $\hat{\mathbf{x}}$ and $\mathbf{x}$ represent any point in the transformed and original un-transformed image respectively. $\hat{\mathbf{x}}$ and $\mathbf{x}$ are related by the following relation:

$$\hat{\mathbf{x}} = hR(v)(\mathbf{x} - O) + (T + O) \quad (6)$$

where O is the center of rotation. The inverse transformation given by the relation:

$$\mathbf{x} = \frac{1}{h}R(v)^T(\hat{\mathbf{x}} - (T + O)) + O \quad (7)$$

While transforming any image (or its signed distance representation), we go over the transformed image and use the inverse relation from equation 7 to get location from original image where we can use tri-linear interpolation for sub-pixel locations.

In order to apply a variational framework for this process, and achieve sub-pixel accuracy in the alignment process we first convert the binary training images to signed distance functions. We then align the 3D SDF masks of the original binary images to a reference image by minimizing an energy functional w.r.t. the set of pose parameters. The objective of the alignment process is to maximize the overlap between a training image and a reference image. An intuitive way to formulate an energy functional representing this task could be as follows:

$$E_{align}(\text{Training Image}) = \int_{\Omega} |\chi_{train} - \chi_{ref}| d\hat{\mathbf{x}} \quad (8)$$

In equation 8, Ω is the training image domain and χ_{train} and χ_{ref} are the characteristic functions of the training and reference images respectively. In any image domain, the characteristic function is equal to 1 inside a shape and 0 outside it. The energy functional essentially measures the overlap between the reference image and image being aligned. If the boundary of the image being aligned matches exactly with that of the reference image

then the energy would be zero. Any deviation from that would lead to an increase in energy. Hence, taking the gradient of alignment energy w.r.t. the pose parameters and running a gradient descent procedure leads to the desired set of final optimized pose parameters which reduce the mismatch due to pose variations in the training images. The problem of aligning a group of images to each other is under determined. In order to get around this problem we can use one of the training images as a reference by keeping its pose parameters fixed and align all other images to this reference.

At each iteration of gradient descent process, we use the following pose parameter update equation:

$$\mathbf{p}_i^{n+1} = \mathbf{p}_i^n - \delta t \cdot \nabla_{\mathbf{p}_i} E_{align} \quad (9)$$

where $\nabla_{\mathbf{p}_i} E_{align}$ represents the gradient of the alignment energy w.r.t. i_{th} component of the pose parameters. Following the discussion in appendix A, the expressions for individual derivatives w.r.t. translation (T), scale (h) and rotation (v) are given by the following set of equations:

$$\nabla_T E_{align} = \int_{\hat{S}} (1 - 2\chi_{ref}) \left(\frac{\partial \hat{P}}{\partial T} \right)^T \hat{N} d\hat{S} \quad (10a)$$

$$\nabla_h E_{align} = \int_{\hat{S}} (1 - 2\chi_{ref}) \left(\frac{\partial \hat{P}}{\partial h} \right)^T \hat{N} d\hat{S} \quad (10b)$$

$$\nabla_v E_{align} = \int_{\hat{S}} (1 - 2\chi_{ref}) \left(\frac{\partial \hat{P}}{\partial v} \right)^T \hat{N} d\hat{S} \quad (10c)$$

In equations 10, $\hat{S}$ refers to the transformed zero levelset² of the image being aligned and $\hat{N}$ is the outward unit normal at each point $\hat{P}$ on the surface. The detailed derivations and final expressions that can be implemented for equation 10 are provided in equation 30 derived in section 5.5.

Based on the experience of clinical experts we chose a set of ten training images from a total of thirty datasets. Out of the ten training images we used one image as a reference image and aligned the rest using the gradient descent formulation of equation 9 and gradient expressions of equation 10 separately for each region LV, RV and Epi.

Figure 1, shows a subset of the training images for Epicardium. The binary black and white shape is the reference image. Each small image represents a 2D slice through the 3D volume which helps to better visualize the images. The red curve is the zero levelset of the corresponding signed distance function of the reference image. The other colored curves

²Since we are working in 3D images the zero levelset of their 3D SDF representations, $\hat{S}$, is a 3D surface and $\hat{P}$ refers to the collection of 3D zero crossing points on the surface.

represent the boundaries of five other training images before alignment. We can clearly see that these images are not exactly overlapping indicating the pose differences in the dataset. Figure 2 shows the same set of images after the alignment process is finished. Again we are showing 2D slices through the 3D volume. The final shapes are much closer to the reference image but clearly differences remain. These differences represent the shape variability in the dataset which we intend to capture through the PCA process and use as prior information in our segmentation models. This alignment process is run separately for each region namely, Epicardium, Left and Right Ventricles.

Using these aligned training image SDF's we use PCA to generate shape priors as described in the following section.

4. Shape Priors

After the Image Alignment phase, we use the aligned signed distance representation of training images for each region, LV, RV and Epi, for running a PCA algorithm (Leventon et al. 2000; Tsai et al. 2003; Vikram et al. 2010). The boundaries of each of the n aligned training shapes are embedded as the zero level set of n separate signed distance functions $\{\Psi_1, \Psi_2, \dots, \Psi_n\}$ with negative distances assigned to inside and positive distances assigned to the outside of the object. We compute the mean level set, Ψ_{mean} , as the average of the n aligned level set representations, $\Psi_{mean}(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n \Psi_i(\mathbf{x})$. To extract the shape variability, the mean shape, Ψ_{mean} , is subtracted from each of the aligned signed distance functions, Ψ_i , to create n mean-offset functions, $\Psi_1, \Psi_2, \dots, \Psi_n$. These mean-offset functions are then used to run PCA procedure to extract the shape priors which represent the shape variabilities.

We form n column vectors, $\vec{\Delta\Psi}_i$, consisting of N samples of each Ψ_i . The most common approach is to vectorize the $N_1 \times N_2 \times N_3$ grid of the training images into a large vector of $N = N_1 \times N_2 \times N_3$ samples in row major order by looping over all slices. Next we define the shape variability matrix as:

$$S = [\vec{\Delta\Psi}_1, \vec{\Delta\Psi}_2, \dots, \vec{\Delta\Psi}_n] \quad (11)$$

We then employ eigen-value decomposition as follows:

$$\frac{1}{n} S S^T = U \Sigma U^T \quad (12)$$

where $U = [\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n]$ is a $N \times n$ matrix whose columns represent the n modes of variations in the shape and Σ is a $n \times n$ diagonal matrix whose diagonal elements represent the corresponding non-zero eigen-values or singular values. The N values of the i_{th} column of U , $\vec{u}_i$, are arranged back into $N_1 \times N_2 \times N_3$ grid by undoing the earlier vectorization operation to yield $u_i(\mathbf{x})$, the i_{th} principal mode of variation which we also refer to as *eigen-shape*. Based on this PCA procedure we can extract maximum n *eigen-shapes*, $u_1(\mathbf{x}), u_2(\mathbf{x}), \dots, u_n(\mathbf{x})$.

Since we are dealing with 3D images, the dimensions of the matrix $\frac{1}{n}SS^T$, $N \times N$ are very large. Running singular value decomposition on such a large matrix is very computationally expensive and also requires large amount of computer resources. In our implementation we compute the eigen-vectors and eigen values of $\frac{1}{n}SS^T$ from a much smaller $n \times n$ matrix given by:

$$W = \frac{1}{n}S^TS \quad (13)$$

If $\mathbf{d}$ is an eigen-vector of W with eigen-value λ then $S\mathbf{d}$ is an eigen-vector of $\frac{1}{n}SS^T$ with eigen-value λ (Leveton 2000).

Using the mean level set function and selecting k n eigen-shapes we introduce a new level set function expressed as a linear combination of mean and k eigen shapes as follows:

$$\Psi(w, \mathbf{x}) = \Psi_{mean}(\mathbf{x}) + \sum_{i=1}^k w_i u_i(\mathbf{x}) = \Psi_{mean}(\mathbf{x}) + w \cdot U_k(\mathbf{x}) \quad (14)$$

where $w = [w_1, w_2, \dots, w_k]$ are the weights associated with the k eigen-shapes $U_k(\mathbf{x}) = [u_1(\mathbf{x}), u_2(\mathbf{x}), \dots, u_k(\mathbf{x})]$ with corresponding variances, $\{\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2\}$ given by the eigen-values calculated from PCA procedure. Equation 14 gives a finite representation of a level set function whose zero level set represents the boundary of the evolving shape. While the PCA analysis behind equation 14 was based on SDF's, it will not, in general, yield a strict SDF but only an approximation³.

As an illustration of how our shape prior model captures the shape variability from the aligned training images, Figure 3 shows how the mean RV shape changes by adding weighted combination of eigen-shapes to the mean. Each small image is a 2D slice through a 3D volume. The white pixels represent the inside of the mean RV shape while black pixels represent outside. The red curve which is the zero levelset of the mean, shows the boundary of the RV shape without any contribution from the eigen-shapes. The blue curve shows how the shape changes when we add σ_1 times the first principal mode of RV variation while magenta curve shows the shape with σ_1 times the first principal mode subtracted from the mean. We can clearly see how this model captures the shape variability. The RV shape generally varies a lot from patient to patient in terms of both size and orientation. The magenta curve particularly highlights the large variation in shape captured by the model.

Similarly, figure 4 shows the shape variability of Epicardium. In this blue and magenta curves show σ_2 times the second mode of variation being added and subtracted from the mean respectively. Again we can see the shape variability especially at the base slices where the shape of the epicardium changes dramatically.

³However this approximation has little consequence since we depend on the zero level set and not the values of the entire function itself.

The segmentation also need to accommodate the variations shape due to pose differences. To be able to handle pose variations $\mathbf{p}$ we modify our shape prior model in equation 14 as follows:

$$\hat{\Psi}(w, \hat{\mathbf{x}}) = \hat{\Psi}_{mean}(\hat{\mathbf{x}}) + \sum_{i=1}^k w_i \hat{U}_i(\hat{\mathbf{x}}) = \hat{\Psi}_{mean}(\hat{\mathbf{x}}) + w \cdot \hat{U}_k(\hat{\mathbf{x}}) \quad (15)$$

where $\hat{\mathbf{x}}$ is related to $\mathbf{x}$ by equation 6.

$$\hat{\mathbf{x}}(\mathbf{x}, \mathbf{p}) = hR(v)(\mathbf{x} - O) + (T + O)$$

We will use the zero level set of $\hat{\Psi}(w, \hat{\mathbf{x}})$ as the new representation of our shape. By varying the shape weights as well as pose parameters, we can capture a broad class of shape variations. Figure 5 shows an overview of the entire training process used to obtain the shape priors and summarizes the notation used at each step.

Since our ultimate goal is myocardial segmentation, we combine 3D shape prior models for LV, RV and Epi which implicitly divides the image into 4 regions as myocardium can now be defined as the region bounded by LV, RV, and Epi and the Background (BG) as the region outside Epi shape. We make use of the 3 shape prior models simultaneously instead of individually. This formulation allows us to impose high level constraints to customize the model for the challenging task of myocardial segmentation.

5. Image modeling

As described previously in section 1.1, there are two broad class of segmentation models used in variational computer vision, edge based and region based. We use region based models where the evolution of the segmenting curve depends upon the pixel intensities within entire regions. That is, region-based models regard an image as the composition of a finite number of regions and rely on regional statistics for segmentation. The statistics of entire regions (such as sample mean and variance) are used to direct the movement of the curve toward the boundaries of the image. This is in contrast to edge-based models where the evolution of the curve depends strictly on nearby pixel intensities (i.e., gradient information). As a result, region-based models are more global than edge-based models. Furthermore, because of the global nature of region-based models, these models do not require the use of inflationary terms commonly employed by edge-based techniques to drive the curve toward image boundaries. Region-based models are also more robust to noise since they do not employ gradient operators, which are inherently sensitive to noise, to explicitly detect the location of edges.

5.1 Region-Based model

In cardiac CT imagery, the interior of Epicardium consists of RV, LV and Myocardium which is the region of Epicardium excluding RV and LV. Additionally, everything excluding Epicardium i.e. the background consists of various anatomical structures. We have developed a region based model defined using an energy functional consisting of two parts:

$$E_{coupled} = \text{Interior Model} + \text{Background Model}$$

The **Interior Model** is a Chan-Vese like model that would appear as follows:

$$E_{coupled} = w_{LV} \int_{LV} (I - \mu_{LV})^2 d\hat{\mathbf{x}} + w_{RV} \int_{RV} (I - \mu_{RV})^2 d\hat{\mathbf{x}} + w_{Myo} \int_{Myo} (I - \mu_{Myo})^2 d\hat{\mathbf{x}} + \text{Background Model} \quad (16)$$

Due to the complexity of the background we treat the **Background Model** separately as described in the next section.

Note that in the above formulation, whenever the three regions plus the background (BG) are disjoint then these data fidelity terms function the same way as in the Chan-Vese formulation. This occurs when there is no overlap between any of the surface pairs. However, overlap could occur in which case these regions do not remain automatically disjoint. In such a scenario we have to more precisely define the domains of each potentially overlapping region. Accordingly, we define the region domains for LV and RV explicitly as the interior of their respective shape models. The background (BG) is defined explicitly as the exterior of Epicardium shape model. Myocardium (Myo) region domain, is defined using all three shape models implicitly as the region inside Epicardium shape and outside both LV and RV shapes. Based on these definitions whenever overlap occurs between the LV and RV the data fidelity penalty gets double counted along the overlap through the first two terms in equation 16. A similar double counting would occur between LV/BG, RV/BG pairs using the background model described in section 5.2 This acts as natural in-built penalty for overlap and thus discourages overlapping shapes. This definition of potentially overlapping region domains introduces an extra level of coupling between the regions and making equation 16 different from standard Chan-Vese energy. In section 5.3, we introduce an additional explicit overlap penalty to ensure the absence of overlap in the final segmented result.⁴

In equation 16, μ_{LV} , μ_{RV} , μ_{Myo} are the respective region mean intensities and w_{LV} , w_{RV} , and w_{Myo} are a set of weighting factors for each region. Using these weighting factors we can bias the evolving shapes to favor the expansion of a particular region. These weighing factors are real numbers and the values relative to each other lead to differing penalties on regions. For example, if w_{Myo} is more than w_{RV} and their respective means are same then a voxel being in Myocardium region would increase the energy more than being in RV region. So using these weights we can bias the expansion of one region compared to others. These weights could be subtly tuned to the specific parameters of a given CT system and contrast protocol. For the system and protocol used in our study we have found the set of weights as shown in section 6.2.

⁴In contrast to regular Chan-Vese, overlap may occur during evolution but the explicit overlap penalty introduced in section 5.3 ensures it doesn't occur in the final result.

In the standard Chan-Vese model there is no restriction on the ordering of the region means. Our model differs in this respect by introducing a strict ordering of the means of each region as shown in equation 17:

$$\mu_{Myo} \leq \mu_{RV} \leq \mu_{LV} \quad (17)$$

This enhancement allows added robustness against initial placement of contours. As an example, if the LV is initialized in the vicinity of the RV in the actual image the segmentation may drift towards RV boundaries. Enforcing this ordering prevents this drift and helps drive the contours towards the correct boundaries even if initialized poorly. While segmenting a dataset, we check and enforce this ordering at each gradient descent iteration. For example if the Myo mean exceeds the RV mean then we enforce the mean ordering constrain by setting both μ_{Myo} and μ_{RV} to be equal to the weighted average of their means.

5.2 Background model

In cardiac imaging the tissue surrounding the epicardium is very complex. There are different types of tissues, muscles, blood pools, calcified vessels and other organs. Each of these regions exhibit different intensity ranges. This leads to the presence of both bright and dark regions. Some regions can be even brighter than the LV. This makes it difficult to model the whole background as a single region with a mean intensity. The shape prior model tries to capture these variations and effectively bleeds the myocardium into the background region.

To overcome these difficulties, we model the background as a combination of two clusters, one darker than myocardium, and other brighter than LV. This leads to the following modification of energy functional:

$$E_{coupled} = \text{Interior Model} + w_{BG} \int_{BG} \min((I - c_{lo})^2, (I - c_{hi})^2) d\hat{\mathbf{x}} \quad (18)$$

The above formulation is a standard two level c-means clustering. By iterating over the background pixels, the background region is divided into two clusters where c_{lo} and c_{hi} are the cluster centers of the low and high intensity regions respectively.

The overall energy functional now becomes:

$$\begin{aligned} E_{coupled} = & w_{LV} \int_{LV} (I - \mu_{LV})^2 d\hat{\mathbf{x}} + w_{RV} \int_{RV} (I - \mu_{RV})^2 d\hat{\mathbf{x}} \\ & + w_{Myo} \int_{Myo} (I - \mu_{Myo})^2 d\hat{\mathbf{x}} + w_{BG} \int_{BG} \min((I - c_{lo})^2, (I - c_{hi})^2) d\hat{\mathbf{x}} \end{aligned} \quad (19)$$

Accordingly we modify the mean intensity ordering given in equation 17:

$$c_{lo} \leq \mu_{Myo} \leq \mu_{RV} \leq \mu_{LV} \leq c_{hi} \quad (20)$$

We start the c-means clustering algorithm by setting c_{lo} equal to μ_{Myo} and c_{hi} equal to μ_{LV} and then run the clustering iteration. In cases where the constraints in eq. 20 are violated at the end of clustering we again enforce the constrain by setting c_{lo} or c_{hi} equal to Myo or LV before running gradient descent iteration. This ensures segmentation better than one achieved using a single mean for background region without mean ordering constraints.

5.3 Overlap Penalty

As described in section 5.1, the basic energy term has an in-built overlap penalty which discourages overlap between regions. On top of that we introduce a separate overlap penalty term to further bias the different regions to not overlap. The overlap penalty is purely regularization term which induces a repulsive force whenever regions overlap. This extra overlap penalty term further couples the separate shape priors into a single higher level shape model.

$$E_{overlap} = \beta_{LR} \left(\int_{LV \cap RV} d\hat{\mathbf{x}} \right) + \beta_L \left(\int_{LV \cap BG} d\hat{\mathbf{x}} \right) + \beta_R \left(\int_{RV \cap BG} d\hat{\mathbf{x}} \right) \quad (21)$$

This overlap penalty energy term penalizes the amount of intersection between LV and RV, LV and BG, and RV and BG with tunable penalty factors. This formulation is made possible by our framework to use three shape priors simultaneously rather than individually.

5.4 Parameter Optimization

Based on the energy formulations described in the image modeling section, the task now is to find an optimal set of pose and shape parameters which minimize the energy functionals designed to segment the four regions. This goal is achieved using parameter optimization procedure which involves running a gradient descent algorithm and updating the pose($\mathbf{p}$) and shape($\mathbf{w}$) parameters for each region at every iteration which can be formulated as:

$$p_i^{(n+1)} = p_i^n - \delta t \cdot \nabla_{p_i} E_{total} \quad (22)$$

$$w_i^{(n+1)} = w_i^n - \delta t \cdot \nabla_{w_i} E_{total} \quad (23)$$

where we take the gradients of the total energy which includes the coupled energy with binary background ($E_{coupled}$) and overlap penalty energy ($E_{overlap}$), w.r.t. each pose and shape parameter. We choose a time step such that segmenting curves don't move by more than one pixel at every step. Additionally, we apply the rotation around the centroid of the evolving shape so as to orthogonalize the translation and scale as much as possible. Since we also update the shape weights every iteration, we update the centroid and recenter the pose parameters every iterations.

Since our entire formulation is in 3D and the energy integrals are volume integrals, the gradient expressions involve surface integrals for each shape respectively. At every iteration we recalculate the updated PCA model with the latest pose and shape parameters and extract the zero crossings of the shapes levelset surfaces. We then calculate the required energy

gradients as described in the next sections and update the pose and shape parameters using equation 22 and equation 23 respectively.

5.5 Energy gradients

In this section we derive the expressions for the gradients required for the gradient descent procedure with respect to both pose and shape parameters. In our approach we segment the cardiac images implicitly by varying the pose and shape parameters and represent the evolving shape boundary as $\hat{S}$ which is the zero levelset surface of shape model $\hat{\Psi}$ given by equation 15. Let $\hat{P}$ and P represent a point on the zero levelset surface of the pose transformed and un-transformed shape models given by equation 15 and equation 14 respectively. $\hat{P}$ and P are related to each other by the relation:

$$\hat{P}(w, \mathbf{p}) = hR(v)(P(w) - O) + (T + O)$$

Similar to the discussion in image alignment section 3 and based on the derivations in appendix A, the expressions for individual derivatives of total energy E w.r.t. translation (T), scale (h) and rotation (v) are given by the following equations:

$$\nabla_T E = \int_{\hat{P}} f \left(\frac{\partial \hat{P}}{\partial T} \right)^T \hat{N} d\hat{S}, \quad \nabla_h E = \int_{\hat{P}} f \left(\frac{\partial \hat{P}}{\partial h} \right)^T \hat{N} d\hat{S}, \quad \nabla_v E = \int_{\hat{P}} f \left(\frac{\partial \hat{P}}{\partial v} \right)^T \hat{N} d\hat{S} \quad (24)$$

The expressions for the scalar force term f for each region are derived in appendix B. Here we derive the gradient expressions $\left(\frac{\partial \hat{P}}{\partial \lambda} \right)^T \hat{N}$ where λ could be T , h or v .

5.5.1 Rotation Gradient—We have chosen to use the Euler-Rodrigues representation for rotation due to its compact nature. Gallego and Yezzi (2015) provide a number of derivations related to this particular representation of space rotations. In particular we make use of the following result to derive the expressions for gradient w.r.t. to rotation:

$$\frac{\partial R(v)u}{\partial v} = -R[u \times] \frac{vv^T + (R^T - I)[v \times]}{\|v\|^2} = -[Ru \times] \frac{vv^T + (I - R)[v \times]}{\|v\|^2} \quad (25)$$

where $u \in \mathbb{R}^3$ is a vector independent of v .

$$9 \left(\frac{\partial \hat{P}}{\partial v} \right)^T \hat{N} = \left(\frac{\partial}{\partial v} (hR(P - O) + (T + O)) \right)^T \hat{N} \quad (26a)$$

$$= \left(h \frac{\partial}{\partial v} R(P - O) \right)^T \hat{N} \quad (26b)$$

$$= \left(-h[R(P - O) \times] \frac{vv^T + (I - R)[v \times]}{\|v\|^2} \right)^T \hat{N} \quad (26c)$$

$$= \left(- \left[(\hat{P} - T - O) \times \right] \frac{vv^T + (I - R)[v \times]}{\|v\|^2} \right)^T \hat{N} \quad (26d)$$

$$= \left(\frac{vv^T + [v \times](R^T - I)}{\|v\|^2} \right) \underbrace{\left((\hat{P} - \hat{O}) \times \hat{N} \right)}_{\hat{z}} \quad (26e)$$

$$= \frac{vv^T \hat{z} + (v \times (R^T \hat{z}) - v \times \hat{z})}{\|v\|^2} = \frac{vv^T \hat{z} + (R^T R v) \times (R^T \hat{z}) - v \times \hat{z}}{\|v\|^2} \quad (26f)$$

$$= \frac{vv^T \hat{z} + R^T ((R v) \times \hat{z}) - v \times \hat{z}}{\|v\|^2} = \frac{vv^T \hat{z} + R^T (v \times \hat{z}) - v \times \hat{z}}{\|v\|^2} \quad (26g)$$

$$= \frac{vv^T \hat{z} + (R^T - I)(v \times \hat{z})}{\|v\|^2} = \frac{vv^T + (R^T - I)[v \times]}{\|v\|^2} (\hat{P} - \hat{O}) \times \hat{N} \quad (26h)$$

$$= \underbrace{\frac{vv^T + (R^T - I)[v \times]}{\|v\|^2}}_{M_v} \underbrace{(\hat{P} - \hat{O})}_{\hat{Q}} \times \hat{N} = M_v (\hat{Q} \times \hat{N}) \quad (26i)$$

Here in going from equation 26b to equation 26c we have made use of the derivative of rotation result provided in equation 25.

5.5.2 Translation gradient

$$\left(\frac{\partial \hat{P}}{\partial T} \right)^T \hat{N} = \left(\frac{\partial}{\partial T} (hR(P - O) + (T + O)) \right)^T \hat{N} \quad (27a)$$

$$= \left(\frac{\partial T}{\partial T} \right)^T \hat{N} = \hat{N} \quad (27b)$$

Going from equation 27a to equation 27b, the only term dependent on the translation is T itself.

5.5.3 Scale gradient

$$\left(\frac{\partial \hat{P}}{\partial h} \right)^T \hat{N} = \left(\frac{\partial}{\partial h} \left(hR(P - O) + \underbrace{(T + O)}_{\hat{O}} \right) \right)^T \hat{N} \quad (28a)$$

$$= (R(P - O))^T \hat{N} = \left(R \left(\frac{1}{h} R^T (\hat{P} - (T + O)) \right) \right)^T \hat{N} \quad (28b)$$

$$= \frac{1}{h} \underbrace{(\hat{P} - \hat{O})}_{\hat{Q}}^T \hat{N} \quad (28c)$$

$$= \frac{1}{h} \hat{Q} \cdot \hat{N} \quad (28d)$$

In the above set of equations we have used the inverse transform relation equation 7 and the fact that the rotation matrix R is orthogonal such that $R^T R = R R^T = I$.

5.5.4 Eigen-shape weights gradient—If $\hat{P}(w)$ is an evolving point on the evolving zero levelset of $\hat{\Psi}(w, \hat{\mathbf{x}}) = \hat{\Psi}_{mean}(\hat{\mathbf{x}}) + w \cdot \hat{U}_k(\hat{\mathbf{x}})$ (where $\hat{\Psi}_{mean}(\hat{\mathbf{x}}) \doteq \Psi_{mean}(\mathbf{x})$ and $\hat{U}_k(\hat{\mathbf{x}}) \doteq U_k(\mathbf{x})$), then we may compute the total derivative in w

$$\frac{\partial}{\partial w} (0 = \hat{\Psi}(w, \hat{P}(w))) = \hat{\Psi}_{mean}(\hat{P}(w)) + w \cdot \hat{U}_k(\hat{P}(w))$$

$$0 = \underbrace{\frac{\partial \hat{\Psi}(w, \hat{P})}{\partial \hat{\mathbf{x}}}}_{\nabla \hat{\Psi}} \frac{\partial \hat{P}}{\partial w} + \underbrace{\frac{\partial \hat{\Psi}}{\partial w}(w, \hat{P})}_{\text{direct partial}} = \left(\frac{\partial \hat{P}}{\partial w} \right)^T \underbrace{\hat{N}(\hat{P}) \|\nabla \hat{\Psi}(w, \hat{P})\|}_{\nabla \hat{\Psi}} + \underbrace{\hat{U}_k(\hat{P}(w))}_{\frac{\partial \hat{\Psi}}{\partial w}}$$

which consists of the direct partial derivative with respect to w together with a chain rule term which links the spatial movement of the zero level set with the partial derivatives in space.⁵ Finally from this equation we obtain:

$$\left(\frac{\partial \hat{P}}{\partial w} \right)^T \hat{N}(\hat{P}) = - \frac{\hat{U}_k(\hat{P})}{\|\nabla \hat{\Psi}(w, \hat{P})\|} = - \frac{U_k(P)}{\|\nabla \hat{\Psi}(w, \hat{P})\|} \quad (29)$$

At the end of equation 29, we use the last equality for computational efficiency such that we do not have to transform all the principal components.

5.5.5 Full gradient expression—In this section we just present the summary of the last sub-sections showing the consolidated gradient expressions that are eventually used in the gradient descent algorithm's update equations presented in section 5.4. If we denote the full vector of pose and shape parameters by λ then we may write more compactly:

⁵This kind of total derivative structure is generally referred to as the Material Derivative when the derivative is taken in time.

$$\lambda = \begin{bmatrix} v \\ h \\ T \\ w \end{bmatrix}, \quad \frac{\partial \hat{P}}{\partial \lambda} = \begin{bmatrix} \frac{\partial \hat{P}}{\partial v} & \frac{\partial \hat{P}}{\partial h} & \frac{\partial \hat{P}}{\partial T} & \frac{\partial \hat{P}}{\partial w} \end{bmatrix}, \quad \left(\frac{\partial \hat{P}}{\partial \lambda} \right)^T \hat{N} = \begin{bmatrix} M_v(\hat{Q} \times \hat{N}) \\ \frac{1}{h}(\hat{Q} \cdot \hat{N}) \\ \hat{N} \\ U_k(P) \\ -\frac{1}{\|\nabla \hat{\Psi}(w, \hat{P})\|} \end{bmatrix}$$

If we denote the total energy i.e. the sum of $E_{coupled}$ and $E_{overlap}$ as simply E then we may write:

$$\nabla_{\lambda} E = \begin{bmatrix} \frac{\partial E}{\partial v} \\ \frac{\partial E}{\partial h} \\ \frac{\partial E}{\partial T} \\ \frac{\partial E}{\partial w} \end{bmatrix} = \begin{bmatrix} \int_{\hat{P}} f \left(\frac{\partial \hat{P}}{\partial v} \right)^T \hat{N} d\hat{S} \\ \int_{\hat{P}} f \left(\frac{\partial \hat{P}}{\partial h} \right)^T \hat{N} d\hat{S} \\ \int_{\hat{P}} f \left(\frac{\partial \hat{P}}{\partial T} \right)^T \hat{N} d\hat{S} \\ \int_{\hat{P}} f \left(\frac{\partial \hat{P}}{\partial w} \right)^T \hat{N} d\hat{S} \end{bmatrix} = \begin{bmatrix} M_v \int_{\hat{P}} f(\hat{Q} \times \hat{N}) d\hat{S} \\ \frac{1}{h} \int_{\hat{P}} f(\hat{Q} \cdot \hat{N}) d\hat{S} \\ \int_{\hat{P}} f \hat{N} d\hat{S} \\ - \int_{\hat{P}} f \frac{U_k(P)}{\|\nabla \hat{\Psi}(w, \hat{P})\|} d\hat{S} \end{bmatrix} \quad (30)$$

Since we are optimizing a set of shape and pose parameters for three shapes, LV, RV, and Epi, we would have three sets of expressions given by equation 30, one for each shape. For example, for LV's scale derivative expression, we would have a set of zero crossing points $\hat{P}$ representing the LV zero levelset surface with corresponding unit normals $\hat{N}$ at each point. The force term f would be given by equation B5a. Scale h would be the current LV scale and $\hat{Q}$ would be calculated at each point for LV. Two similar instances of scale derivative would in turn apply to Epi and RV.

5.5.6 Momentum accelerated gradient descent—Standard gradient descent algorithms can generally suffer from slow convergence or can get stuck in a shallow valleys leading to sub-optimal results. There are several methods that help ameliorate these drawback (Sebastian 2017). Most common method employed to overcome slow convergence and oscillations in standard gradient descent is called momentum (Qian 1999) based accelerated gradient descent. In this scheme a fraction of previous update vector is added to the current update vector. This approach allows the descent procedure to pick up “speed” and helps overcome “shallow valleys” and approach better local minimums. Essentially, momentum increases for parameters whose gradient points in the same directions and reduces updates for parameters whose gradient changes directions. Our shape based approach can be combined with this method to further boost the gradient descent procedure in natural dynamic way and leads to better segmentation results.

6. Experimental results

We applied our integrated anatomical model for cardiac segmentation on 30 cardiac tomography angiography (CTA) patient studies obtained at Nuclear Cardiology R&D

Laboratory at Emory University School of Medicine. Each of the studies were hand segmented by clinical experts to obtain binary masks as gold standard segmentations for each anatomical region of interest. We used 10 out of the 30 studies as training images to derive the shape models.

6.1 Choosing number of principal components

As discussed in section 4, each of the training datasets produces a principal component (eigen shape) and during the automatic segmentation phase we can represent the model using k of those eigen shapes. As such, there is no automatic procedure for choosing an optimal k , instead it is a trial and error process. We applied our PCA model to the original binary hand segmented images. We conducted experiments by using different numbers of eigen shapes, and compared the accuracy of the segmentations obtained using a statistical technique of binary classification (FMe 2017) called F-Measure. Figure (6) shows the plot in increasing order of accuracy for the Epicardium region for different numbers of modes. We observed a prominent spike in performance going from 3 to 5 modes but a rather smaller increases for using additional eigen shapes. The left and right ventricle segmentations also exhibited a similar trend. Such an analysis would be beneficial for a future commercial application where hundreds of training images might be used with limited computational resources.

6.2 Automatic segmentation results

Figure 8 shows a sample CT image along with the shape prior model initialized with mean shape for each of the three regions. The red, yellow and blue curves are the zero level sets of Epi, RV and LV respectively. Each small image shows a 2D slice through the 3D volume where the curves are actually surfaces. The region enclosed by these curves is Myo and everything outside of the red (Epi) curve is BG.

As illustrated in figure 7, the segmentation process starts with this mean shape model where the eigen-shape weights are initially all set to zero. The pose parameters are also initialized to unity (translation and rotation are 0 but scale is 1) i.e. no transformations are applied yet to the model. The segmentation then proceeds with the momentum based gradient descent algorithm using the mathematical formulation derived in section 5. We run several iterations until convergence when the energy stops decreasing substantially. In each iteration, the shape and pose parameters are updated to yield progressively better segmentation.

Figure 9 shows the same test case which was initialized with the shape prior mean shape model after the gradient descent procedure is finished. During this process an optimized set of pose and shape parameters are obtained for each of the three regions. The difference in the starting and final shape for RV and Epi is especially stark. The RV region initially partially overlaps with myocardium and is at the edge of LV region. The gradient descent process pushes it to its own region leading to a much better final myocardial segmentation. We used $w_{LV} = 0.5$, $w_{RV} = 0.5$, $w_{Myo} = 1.5$ and $w_{BG} = 1$ for all experiments. As explained in section 5.1, tuning these weights can lead to more expansion of some regions compared to others. Having higher weights for Myo compared to LV and RV leads to expansion of these regions compared to Myo. This is in line with our knowledge that most of the region inside

Epi should be occupied by LV and RV compared to Myo. With these weights we can overcome the lack of contrast especially between RV and Myo regions.

Figure 10 shows the region-based model that corresponds to the final segmentation results in the previous example. The darkest and brightest region (visible in some 2D slices) outside the epicardium (red curve) correspond to the two clusters of the background region. Without this binary background formulation the Epicardium shape gets attracted towards these bright regions, generally calcified vessels, especially if they are located very close to epicardium. This problem is avoided due to our binary background and means ordering formulation.

Finally, figure 11 shows the 3D surfaces representing the manual segmentation generated by clinicians compared to the final result of the optimized PCA model generated by our algorithm. Compared to manual segmentation, our model is also much smoother.

After running our automatic segmentation algorithm, we convert the final segmenting curves of each shape back to binary masks, again one for each region. Clinical experts at Emory university calculated and compared the results with manual tracings and found an error (mean $\pm$ std) of 1.46 ± 0.41 mm for LV, 2.06 ± 0.65 mm for RV, and 2.88 ± 0.59 mm for Epi. The results for LV are most accurate as this region usually has the best contrast. The Epi and RV results are slightly lower than LV. Currently, there are very few, if any, tools that can perform fully automatic cardiac segmentation available in the market. Hence it is difficult to compare the improvements to any industry standard.

As mentioned in section 1.1, there have been many different segmentation methods reported in the literature. There are several issues that prevent a direct comparison of our algorithm with these other state-of-the-art methods. Firstly, all the methods reviewed in this paper reported segmentation accuracy results in trans-axial orientation of the data. Since our imaging experts are trained to recognize cardiac anatomy based on the hearts natural axes (short, vertical and horizontal long axes) they are more confident in their manual tracings done in short axis orientation compared to trans-axial orientation. Thus biventricular short axis segmentation leads to more accurate manual segmentation and reduced inter and intra observer variability of the boundaries. In practically all of cardiac imaging this natural axes approach has been adopted as the gold standard for manual segmentation of myocardial boundaries. Consequently, we have developed our algorithms based on short-axis orientation of all our datasets. Hence a direct comparison with techniques implemented in trans-axial orientation would generate a larger error due to the gold standard rather than the automatic segmentation techniques and thus not a fair comparison. Secondly, the method used to compare the accuracy of automated segmentation to manual tracings is not clear in some of the reviewed papers. This further makes it difficult to compare the results directly. Finally, comparison between techniques is further complicated by some techniques reporting their results only for the LV, some treat LV and RV myocardium separately while others treat them as a single component.

For these reasons, the best way to make a fair quantitative comparison between different myocardial segmentation algorithms is to use same datasets for testing as well as use the same segmentations as gold standard. To this end, we now use the same image datasets and

manual segmentations standards to compare the automatically segmented results from our new algorithm to those using the variational region growing method presented in Zhu et al. (2013b).

In order to compare Zhu's method and our new method, we used six out of the twelve pigs datasets as training data to generate our shape prior models for each shape. We then used our segmentation algorithm to generate automatic segmentation masks for all twelve datasets. Our clinical expert collaborators at Emory University then used the same error measurement algorithm to compare the results of both methods to the manual segmentations. Note that these manual segmentations were originally drawn from trans-axial slices since it was the only original data we had in common for comparisons between techniques.

Table 1, summarizes the mean errors and standard deviation for each shape. Note that our proposed algorithms shows a trend of smaller mean error compared to the variational region growing method for all three shapes. A one-tail paired t-test confirms the trend albeit not yet reaching statistical significance with $p = .055$ for the epicardial boundaries, $p = .20$ for the LV endocardium and $p = .41$ for the RV. Improvements in these results will require improvements in the reference gold standard by generating the manual segmentation from natural axis of the heart as well as averaging the results of multiple experts. In Zhu's paper the authors showed better results compared to the localized PCA based method presented in Vikram et al. (2010, 2011). The authors also showed their results to be competitive with the methods of Ecabert et al. (2008) and Zheng et al. (2008). Now we have shown a trend of improved results over Zhu's method.

Additionally, we ran our tests on a quad core 3.2 GHz Intel core i7 CPU with 16 GB RAM. Segmenting a new test case takes approximately 50 seconds on average. However, in our algorithms a lot of computation can be parallelized and the running time can be cut to about one third of current times.

7. Conclusion

We have presented a novel combination and extension of existing frameworks to automatically segment myocardial boundaries in cardiac CT images. Among their various applications, the proposed methodologies have been developed also in the context of a project for multimodality image fusion (Piccinelli et al. 2014). Function and anatomy are equally important in the assessment of Coronary Artery Disease. A powerful imaging strategy for a comprehensive evaluation of the patient status is to fuse a nuclear imaging test, i.e. single-photon emission computed tomography (SPECT) or positron emission tomography (PET), with an anatomical imaging test, i.e. a CTA. Currently, there are very few, if any, tools that can perform full automatic cardiac segmentation available in the market, which hampers the translation of a multimodality fusion approach to the clinical environment. The results obtained so far are particularly promising and a validation of a fully automated fusion procedure is underway to assess its feasibility and clinical reliability. The algorithmic framework can also be easily adapted to work with cardiac magnetic resonance imagery (MRI), which is increasingly becoming a prominent cardiac imaging modality.

Acknowledgments

This work was funded in part by a seed grant from the Coulter Foundation (Cou 2017), National Science Foundation (NSF) grant number CCF-1526848, National Institutes of Health (NIH) grant number R01 HL143350, and Army Research Office grant number ARO W911NF-18-1-0281.

Appendix A.: Region Based Active Contours and Active Surfaces

For its various advantages, in our work we have used region-based active surfaces for the task of 3D Cardiac CT segmentation. Here we provide some basic concepts related to active contours/surfaces frequently used in our mathematical formulations. Figure A1 shows a typical active contour in 2D where the curve C divides an image in two regions inside (R_{in}) and outside (R_{out}). In 3D it would be a surface (S) which divides a volume into inside and outside. Suppose we devise some energy functional for a particular task as follows:

$$E = \int_{R_{in}} f_{in} d\mathbf{x} \quad (A1)$$

We can modify equation A1 using the well known divergence theorem. Suppose $\vec{F}_{in}$ is a vector field such that $\nabla \cdot \vec{F} = f_{in}$, then equation A1 can be written as:

$$E = \int_{R_{in}} f_{in} d\mathbf{x} = \int_{R_{in}} (\nabla \cdot \vec{F}_{in}) d\mathbf{x} = \int_S \vec{F}_{in} \cdot \mathbf{N} dS \quad (A2)$$

where the last equality follows by using the divergence theorem which relates the volume integral of divergence of flux $\vec{F}$ to an equivalent surface integral. In equation A2, $\mathbf{N}$ is the outward unit normal to the surface S (curve C in 2D) and dS is a unit surface area element. As shown in Yezzi et al. (2003), the derivative of energy w.r.t. a parameter λ is given by the following expression:

$$\frac{\partial E}{\partial \lambda} = \int_S \nabla \cdot \vec{F}_{in} \left(\frac{\partial S}{\partial \lambda} \cdot \mathbf{N} \right) dS \quad (A3)$$

Figure A1.
Typical active contour

With our assumption that $\nabla \cdot \vec{F}_{in} = f_{in}$ the last equation becomes:

$$\frac{\partial E}{\partial \lambda} = \int_S f_{in} \left(\frac{\partial S}{\partial \lambda} \cdot N \right) dS \quad (A4)$$

Now if we are given an integral over an external region R_{out} as $E = \int_{R_{out}} f_{out} d\mathbf{x}$, we can rewrite it as follows:

$$E = \int_{\Omega} f_{out} d\mathbf{x} - \int_{R_{in}} f_{out} d\mathbf{x} \quad (A5)$$

where Ω is the complete image domain. Noting that the first term doesn't depend on λ and using equations A1 through A4, the derivative of this energy w.r.t. parameter λ is given by the following equation:

$$\frac{\partial E}{\partial \lambda} = 0 - \int_S f_{out} \left(\frac{\partial S}{\partial \lambda} \cdot N \right) dS = - \int_S f_{out} \left(\frac{\partial S}{\partial \lambda} \cdot N \right) dS \quad (A6)$$

In our work we have energy functionals with integrals of both inside and outside regions of the form:

$$E = \int_{R_{in}} f_{in} d\mathbf{x} + \int_{R_{out}} f_{out} d\mathbf{x} \quad (A7)$$

By using equation A4 and equation A6, the derivative w.r.t. λ of the previous energy is easily seen to be:

$$\frac{\partial E}{\partial \lambda} = \int_S f_{in} \left(\frac{\partial S}{\partial \lambda} \cdot N \right) dS - \int_S f_{out} \left(\frac{\partial S}{\partial \lambda} \cdot N \right) dS \quad (A8a)$$

$$= \int_S (f_{in} - f_{out}) \frac{\partial S}{\partial \lambda} \cdot N dS \quad (A8b)$$

In equation A8, λ could be any parameter such as translation (T), scale (h), rotation (ν) or shape weight (w). We have derived the partial derivatives of the surface $\frac{\partial S}{\partial \lambda}$ in the relevant places in the main text.

Although equation 8 in section 3 represents the alignment energy functional, it is more convenient to rewrite it as follows:

$$E_{align}(Training Image) = \int_{R_{in}} (1 - \chi_{ref}) d\hat{\mathbf{x}} + \int_{R_{out}} \chi_{ref} d\hat{\mathbf{x}} \quad (A9)$$

where R_{in} and R_{out} are the inside and outside regions of the training image shape being aligned with respect to the reference image. In this alignment energy, $f_{in} = 1 - \chi_{ref}$ and $f_{out} = \chi_{ref}$. Hence $f_{in} - f_{out} = 1 - 2\chi_{ref}$ which leads to the derivative expressions in equation 10.

Appendix B.: Outward Normal Force For Segmentation Energy

In section 5, we described our image segmentation model consisting of an interior model, a background model and an overlap penalty term. While these models are more complicated than the image alignment model, they are nevertheless of the same form as equation A7 consisting of integrals over an interior and exterior region. Consequently, the derivatives of the segmentation energy w.r.t. the pose and shape parameters of each shape, namely, LV, RV and Epi, are of the same form as equation A8. Here we derive the expressions for $f_{in} - f_{out}$ for each component of the segmentation energy for all three shapes.

Equation 16 and 18 defined the interior and background model respectively and the resulting region based energy functional, $E_{coupled}$, was defined as follows:

$$E_{coupled}(\hat{S}_{LV}, \hat{S}_{RV}, \hat{S}_{Epi}) = w_{LV} \int_{LV} (I - \mu_{LV})^2 d\hat{\mathbf{x}} + w_{RV} \int_{RV} (I - \mu_{RV})^2 d\hat{\mathbf{x}} + w_{Myo} \int_{Myo} (I - \mu_{Myo})^2 d\hat{\mathbf{x}} + w_{BG} \int_{BG} \min((I - c_{lo})^2, (I - c_{hi})^2) d\hat{\mathbf{x}} \quad (B1)$$

Following the same procedure as in appendix A, when taking the derivatives of $E_{coupled}$ w.r.t. LV parameters we note that the second and fourth integral terms do not depend on LV shape. Similarly, in case of RV the first and fourth terms do not depend on RV and for Epi the first and second terms are independent of Epi shape. Hence,

$$f_{LV}^{coupled} = \underbrace{w_{LV}(I - \mu_{LV})^2}_{f_{in}} - \underbrace{w_{Myo}(I - \mu_{Myo})^2 \chi_{epi}(1 - \chi_{RV})}_{f_{out}} \quad (B2a)$$

$$f_{RV}^{coupled} = \underbrace{w_{RV}(I - \mu_{RV})^2}_{f_{in}} - \underbrace{w_{Myo}(I - \mu_{Myo})^2 \chi_{epi}(1 - \chi_{LV})}_{f_{out}} \quad (B2b)$$

$$f_{epi}^{coupled} = \underbrace{w_{Myo}(I - \mu_{Myo})^2(1 - \chi_{LV})(1 - \chi_{RV})}_{f_{in}} - \underbrace{w_{BG}(I - c(I))^2 \chi_{epi}(1 - \chi_{RV})}_{f_{out}} \quad (B2c)$$

where χ_{epi} , χ_{RV} , and χ_{LV} are characteristic functions of Epi, RV and LV shapes respectively. The characteristic function terms can be explained by how we have defined the region domains due to the possibility of overlap between regions in section 5.1. Keeping with our modified background model described in 5.2, in equation B2c, we define the background parameter $c(I)$ as:

$$c(I) = \begin{cases} c_{lo}, & I \leq \frac{c_{lo} + c_{hi}}{2} \\ c_{hi}, & I > \frac{c_{lo} + c_{hi}}{2} \end{cases} \quad (B3)$$

The overlap penalty is defined as:

$$E_{overlap}(\hat{S}_{LV}, \hat{S}_{RV}, \hat{S}_{Epi}) = \beta_{LR} \left(\int_{LV \cap RV} d\text{magenta} \hat{\mathbf{x}} \right) + \beta_L \left(\int_{LV \cap BG} d\hat{\mathbf{x}} \right) + \beta_R \left(\int_{RV \cap BG} d\hat{\mathbf{x}} \right)$$

which may be rewritten as follows:

$$E_{overlap} = \beta_{LR} \int_{LV} \chi_{RV} d\hat{\mathbf{x}} + \beta_{LR} \int_{RV} \chi_{LV} d\hat{\mathbf{x}} + \beta_L \int_{LV} (1 - \chi_{epi}) d\hat{\mathbf{x}} + \beta_L \int_{BG} \chi_{LV} d\hat{\mathbf{x}} + \beta_R \int_{RV} (1 - \chi_{epi}) d\hat{\mathbf{x}} + \beta_R \int_{BG} \chi_{RV} d\hat{\mathbf{x}}$$

Following the same procedure as above and noting that for LV and RV $f_{out} = 0$ and for Epi $f_{in} = 0$, we get the following expressions:

$$f_{LV}^{overlap} = \beta_{LR} \chi_{RV} + \beta_L (1 - \chi_{epi}) \quad (B4a)$$

$$f_{RV}^{overlap} = \beta_{LR} \chi_{LV} + \beta_R (1 - \chi_{epi}) \quad (B4b)$$

$$f_{Epi}^{overlap} = -\beta_L \chi_{LV} + \beta_R \chi_{RV} \quad (B4c)$$

The overall force f that is used in the gradient expressions in equation 30, is the sum of the forces corresponding to $E_{coupled}$ and $E_{overlap}$ and is given by the following:

$$f_{LV} = f_{LV}^{coupled} + f_{LV}^{overlap} \quad (B5a)$$

$$f_{RV} = f_{RV}^{coupled} + f_{RV}^{overlap} \quad (B5b)$$

$$f_{Epi} = f_{Epi}^{coupled} + f_{Epi}^{overlap} \quad (B5c)$$

9. References

- Caselles V, Kimmel R, Sapiro G. 1997 Geodesic active contours. International Journal of Computer Vision. 22(61):61–79.
- CDC. 2015 Heart disease: Scope and impact. Available from: <http://www.theheartfoundation.org/heart-disease-facts/heart-disease-statistics/>.
- Chakraborty A, Staib L, Duncan J. 1994 An integrated approach to boundary finding in medical images. In: Proc. IEEE Workshop Biomedical Image Analysis p. 13–22.
- Chan T, Vese L. 1999 An active contour model without edges. In: Int. Conf. Scale-Space Theories in Computer Vision p. 141–151.
- Chan T, Vese L. 2001 A level set algorithm for minimizing the mumford-shah functional in image processing. In: IEEE Workshop on Variational and Level Set Methods in Computer Vision p. 161–168.

- Chen Y, Thiruvankadam S, Huang F, Wilson D, G MEA, Tagare H. 2001 On the incorporation of shape priors into geometric active contours. In: IEEE Workshop on Variational and Level Set Methods in Computer Vision p. 145–152.
- Cheng H, Gupta K. 1989 A historical note on finite rotations. *Journal of Applied Mechanics*. 56(1): 139–145.
- Cootes T, Edwards G, Taylor C. 2001 Active appearance models. *IEEE Trans on Pattern Recognition and Machine Intelligence*. 23(6):681–685.
- Cootes T, Taylor C. 1992 Smart snakes. In: *Proceedings of British Machine Vision Conference* p. 266–275.
- Cootes T, Taylor C. 1998 Active appearance models. In: *Proceedings of European Conference on Computer Vision* Springer; p. 484–498. vol. 2.
- Cootes T, Taylor C, Cooper C, Graham J. 1995 Active shape models - their training and application. *Computer Vision and Image Understanding*. 61(9):38–59.
- Cootes TF, Taylor CJ. 2001 Statistical models of appearance for computer vision. University of Manchester Report No.: MSU-CSE-06-2.
- Dryden I, Mardia K. 1998 Statistical shape analysis. John Wiley & Sons.
- Ecabert O, Peters J, Schramm H, Lorenz C, von Berg J, Walker M. 2008 Automatic model-based segmentation of the heart in ct images. *IEEE Trans Med Imag*. 27(9):1189–1201.
- Gallego G, Yezzi A. 2015 A compact formula for the derivative of a 3-d rotation in exponential coordinates. *Journal of Mathematical Imaging and Vision*. 51(3):378–384.
- Gao Y, Kikinis R, Bouix S, Shenton M, Tannenbaum A. 2012 A 3d interactive multi-object segmentation tool using local robust statistics driven active contours. *Med Image Anal*. 16(6): 1216–1217. [PubMed: 22831773]
- Goodall C 1991 Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society Series B*. 53(2):285–339.
- Kichenessamy S, Kumar A, Olver P, Tannenbaum A, Yezzi A. 1995 Gradient flows and geometric active contour models. In: *Proc. of Intl. Conf. Computer Vision* p. 810–815.
- Kiri li HA, Schaap M, Klein S, Papadopoulou SL, Bonardi M, Chen CH, Weustink AC, Mollet NR, Vonken EJ, van der Geest RJ, et al. 2010 Evaluation of a multi-atlas based method for segmentation of cardiac cta data: a large-scale, multicenter, and multivendor study. *Medical Physics*. 37(12):6279–6291. [PubMed: 21302784]
- Leventon M, Grimson E, Faugeras O. 2000 Statistical shape influence in geodesic active contours. In: *Proc. IEEE Conf. on Computer Vision and Pattern Recognition* p. 316–323. vol. 1.
- Leveton M 2000 Statistical models in medical image analysis [dissertation]. Boston (MA): Massachussets Institute of Technology.
- Mattes D, Haynor DR, Vesselle H, Lewellen TK, Eubank W. 2003 Pet-ct image registration in the chest using free-form deformations. *IEEE Transactions on Medical Imaging*. 22:120–128. [PubMed: 12703765]
- Mitchell SC, Lelieveldt B, van der Geest RJ, Bosch JG, Reiber JHC, Sonka M. 2001 Multistage hybrid active appearance model matching: Segmentation of left and right ventricles in cardiac mr images. *IEEE Trans Med Imag*. 20(5):415–423.
- Mumford D, Shah J. 1989 Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*. 42(6):577–685.
- Murray R, Li Z, Sastry S. 1994 A mathematical introduction to robotic manipulation. CRC Press.
- Osher S, Sethian J. 1988 Fronts propagating with curvature dependent speed: Algorithms based on hamilton-jacobi formulations. *Journal of Computational Physics*. 79(1):12–49.
- Piccinelli M, Faber T, Arepalli C, Appia V, Vinten-Johnsen J. 2014 Automatic detection of left and right ventricles from cta enables efficient alignment of anatomy with myocardial perfusion data. *Journal of Nuclear Cardiology*. 21:396–108.
- Qian N 1999 On the momentum term in gradient descent learning algorithms. *Neural Networks*. 12(1): 145–151. [PubMed: 12662723]

- Rohlfing T, Brandt R, Menzel R, Maurer CR. 2003 Segmentation of three-dimensional images using non-rigid registration: methods and validation with application to confocal microscopy images of bee brains. *Proc SPIE*. 5032:363–374. Available from: 10.1117/12.483558 .
- Sebastian R 2017 An overview of gradient descent optimization algorithms In: arXiv:1609.04747v2 [cs.LG]. p. 1–14.
- Sethian J 1996 A fast marching level set method for monotonically advancing fronts. *Proceedings of the National Academy of Sciences of the United States of America*. 93(4):1591–1595. [PubMed: 11607632]
- Shahzad R, Bos D, Budde RPJ, Pellikaan K, Niessen WJ, van der Lugt A, Walsum T. 2017 Automatic segmentation and quantification of the cardiac structures from non-contrast-enhanced cardiac ct scans. *Physics in Medicine & Biology*. 62(9):3798 Available from: <http://stacks.iop.org/0031-9155/62/i=9/a=3798> . [PubMed: 28248196]
- Staib L, Duncan J. 1992 Boundary finding with parametrically deformable contour models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 14:1061–1075.
- Statistical binary classification method. Available from: https://en.wikipedia.org/wiki/F1_score.
- Tsai A, Yezzi A, Wells W, Tempany C, Tucker D, Fan A, Grimson WE, Willsky A. 2003 A shape-based approach to the segmentation of medical imagery using level sets. *IEEE Transactions on Medical Imaging*. 22:137–154. [PubMed: 12715991]
- van Assen HC, Danilouchkine MG, Dirksen MS, Reiber JH, Lelieveldt BP. 2008 A 3-d active shape model driven by fuzzy inference: Application to cardiac ct and mr. *IEEE Transactions on Information Technology in Biomedicine*. 12(5):595–605. [PubMed: 18779074]
- van Assen HC, Danilouchkine MG, Frangi FF, Ordas S, Westenberg JJ, Reiber JH. 2006 Spasm: A 3d-asm for segmentation of sparse and arbitrarily oriented cardiac mri data. *Medical Image Analysis*. 10(2):286–303. [PubMed: 16439182]
- Vikram A, Ganapathy B, Abufadel A, Yezzi A, Faber T. 2010 A regions of confidence based approach to enhance segmentation with shape priors. In: *Proc. of SPIE-IS&T Electronic Imaging*, SPIE p. 7533–12.
- Vikram A, Ganapathy B, Yezzi A, Faber T. 2011 Localized principal component analysis based curve evolution: A divide and conquer approach. In: *Proc. of IEEE International Conference in Computer Vision* p. 1981–1986.
- Weese J, Kaus MR, Lorenz C, Lobregt S, Truyen R, Pekar V. 2001 Shape constrained deformable models for 3-d medical image segmentation In: *Information Processing in Medical Imaging*. Springer Berlin Heidelberg; p. 380–387.
- Yezzi A, Kichenassamy S, Kumar A, Olver P, Tannenbaum A. 1997 A geometric snake model for segmentation of medical imagery. *IEEE Transactions on Medical Imaging*. 16:199–209. [PubMed: 9101329]
- Yezzi A, Zollei L, Kapur T. 2003 A variational framework for integrating segmentation and registration through active contours. *Journal of Medical Image Analysis*. 7:171–185. [PubMed: 12868620]
- Zheng Y, Barbu A, Georgescu B, Scheuering M, Comaniciu D. 2008 Four-chamber heart modeling and automatic segmentation for 3d cardiac ct volumes using marginal space learning and steerable features. *IEEE Trans Med Imag*. 27(11):1668–1681.
- Zhu L, Gao Y, Appia V, Yezzi A, Arepalli C, Faber T, Stillman A, Tannenbaum A. 2013a . Automatic delineation of the myocardial wall from ct images via shape segmentation and variational region growing. *IEEE Trans Biomedical Imaging*. 60(10):2887–2895.
- Zhu L, Gao Y, Appia V, Yezzi A, Arepalli C, Faber T, Stillman A, Tannenbaum A. 2013b . Automatic delineation of the myocardial wall from ct images via shape segmentation and variational region growing. In: *IEEE Trans. on Biomedical Engineering* p. 2887–2895.
- Zhu L, Gao Y, Appia V, Yezzi A, Arepalli C, Faber T, Stillman A, Tannenbaum A. 2014 A complete system for automatic extraction of left ventricular myocardium from ct images using shape segmentation and contour evolution”. *IEEE Trans Image Processing*. 23(3):1340–1351.

Figure 1.

Set of Epicardium training images before alignment. The black and white image is the reference image with the red curve being the zero levelset of its corresponding signed distance function representation. Different color boundaries are other images that are aligned to this reference image. 2D slices of a 3D volume are shown for better visualization

Figure 2.

Set of Epicardium training images after alignment. The amount of overlap increases after the alignment process. The differences that still remain represent the shape variability we wish to capture through the PCA technique. 2D slices of a 3D volume are shown for better visualization

Figure 3.

Illustration of how the shape prior model captures the shape variability in case of RV. The red curve is the mean RV shape boundary and bright pixels show the interior for easier visualization. Blue curve shows $\Psi_{mean} + 1\sigma_1 u_1$ while magenta curve is $\Psi_{mean} - 1\sigma_1 u_1$. 2D slices through the 3D volume are shown here

Figure 4.

Illustration of how the shape prior model captures the shape variability in case of Epicardium. The red curve is the mean Epicardium shape boundary and bright pixels show the interior for easier visualization. Blue curve shows $\Psi_{mean} + 1\sigma_2u_2$ while magenta curve is $\Psi_{mean} - 1\sigma_2u_2$. 2D slices through the 3D volume are shown here

Figure 5.

Overview of the training process and notation used for elements of the shape prior model.

Figure 6.
Accuracy for the Epicardium region for different numbers of modes.

Figure 7.
Flowchart for the segmentation process.

Figure 8:

Shape prior model initialized with mean shape for each of the three regions. The red, yellow and blue curves are the zero level sets of Epi, RV and LV respectively. Each small image shows a 2D slice through the 3D volume

Figure 9:
Myocardial segmentation after gradient descent is finished. The red, yellow and blue curves are the zero level sets of Epi, RV and LV respectively. Each small image shows a 2D slice through the 3D volume

Figure 10:

The model corresponding to the final segmentation in example test case. The darkest and brightest regions are the two background clusters. In the interior of Epicardium (red curve), LV is the brightest region followed by RV and finally myocardium

Figure 11.
3D rendering of manual(left) vs automatic(right) segmentation

Table 1.

Comparison between proposed method and the variational region growing method (Zhu et al.) on same twelve pigs datasets (mean mm $\pm$ std mm)

Method	LV	RV	Epi
Proposed method	1.38 $\pm$ 0.70	1.88 $\pm$ 0.93	1.63 $\pm$ 0.74
Zhu et al.	1.64 $\pm$ 0.63	1.97 $\pm$ 0.63	3.21 $\pm$ 2.84