Segmentation of Left Ventricle From 3D Cardiac MR Image Sequences Using A Subject-Specific Dynamical Model

Yun Zhu, Xenophon Papademetris, Albert Sinusas, James S. Duncan
Department of Biomedical Engineering and Diagnostic Radiology, Yale University
310 Cedar Street, New Haven, CT 06520
yun.zhu@yale.edu

Abstract

Statistical model-based segmentation of the left ventricle from cardiac images has received considerable attention in recent years. While a variety of statistical models have been shown to improve segmentation results, most of them are either static models (SM) which neglect the temporal coherence of a cardiac sequence or generic dynamical models (GDM) which neglect the inter-subject variability of cardiac shapes and deformations. In this paper, we use a subject-specific dynamical model (SSDM) that handles inter-subject variability and temporal dynamics (intra-subject variability) simultaneously. It can progressively identify the specific motion patterns of a new cardiac sequence based on the segmentations observed in the past frames. We formulate the integration of the SSDM into the segmentation process in a recursive Bayesian framework in order to segment each frame based on the intensity information from the current frame and the prediction from the past frames. We perform "Leave-one-out" test on 32 sequences to validate our approach. Quantitative analysis of experimental results shows that the segmentation with the SSDM outperforms those with the SM and GDM by having better global and local consistencies with the manual segmentation.

1. Introduction

Segmentation of the left ventricle (LV) from cardiac images plays an important role in quantitative functional analysis, such as ejection fraction estimation and myocardial motion analysis. In recent years, statistical shape models have been extensively used in LV segmentation (see [3] for an overview). For example, Mitchell *el al.* proposed a multistage Active Appearance Model (AAM) to segment left and right ventricles from cardiac MR images [7]. They later

extended their work to segmenting 3-D images [6]. These models, however, supply a prior just for the shape, but not for the motion of that shape. Hence, they are static in time and ignore the characterization of cardiac motion patterns.

The heart is a dynamic system, making time-independent segmentation inadequate. Perperidis *el al.* extended 3-D model to 4-D case by adding the temporal information [10]. They built two separate models to account for the inter-subject variability and cardiac dynamics (intra-subject variability) respectively. While these two models were successfully applied to the classification of cardiac images from normal volunteers and patients of hypertrophic cardiomyopathy, they are not related, making them inappropriate for LV segmentation.

Some work has been done to build dynamical models. Mitchell extended their AAM to 2-D Active Appearance Motion Model (AAMM) [6]. However, it is nontrivial to extend 2-D AAMM to segmenting a full 3-D cardiac sequence because of the high dimensionality. Sun proposed to learn the cardiac dynamics using a second-order dynamical model [11]. In the computer vision community, Blake proposed to use dynamical model to track objects in a filtering framework [1]. Cremers used a second-order autoregressive model to describe the silhouettes of a walking person [2]. While this approach shows improvement for gait tracking, a second-order system is insufficient to describe complex shape deformations such as cardiac dynamics. Also, they learn a uniform model for all sequences, thus ignoring the subject variability of the motion patterns, making them generic dynamical models (GDM).

Our objective is to build a subject-specific dynamical model (SSDM) to account for two factors that cause the cardiac shape variability simultaneously. One factor is the inter-subject variability, the other is the temporal dynamics caused by cardiac deformation during a cardiac cycle, as shown in Figure 1. These two factors are interactive and cannot be separated into two independent statistical models. Unfortunately, conventional Principal Components Analysis (PCA) and Independent Component Analysis (ICA) can

only account for one factor at one time, and therefore are inappropriate for our dynamical model.
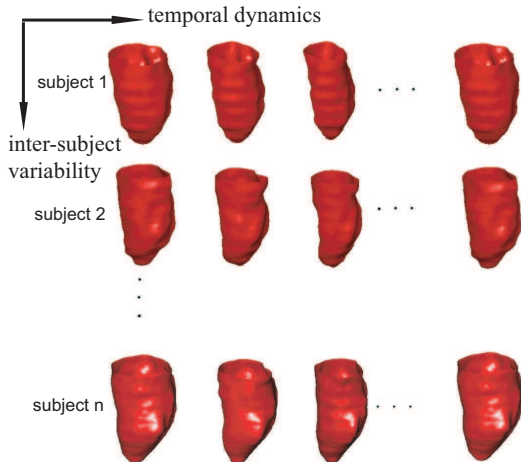

Figure 1. The interaction of cardiac temporal dynamics and inter-subject shape variability.

To address this problem, we extend conventional PCA and ICA to higher orders, and use Multilinear PCA (MPCA) and Multilinear ICA (MICA) [12, 13] to build a dynamical model that decomposes the training set in order to take care of the temporal dynamics (intra-subject variability) and inter-subject variability simultaneously. In addition, we design a dynamic prediction algorithm that can progressively identify the subject vector associated with a new cardiac sequence and use this subject vector to predict the subject-specific dynamics from segmentations observed in previous frames. We formulate the integration of such SSDM into the segmentation process in a recursive Bayesian framework. This framework models the evolution of endocardial (ENDO) and epicardial (EPI) surfaces driven both by intensity information from the current frame as well as the dynamical shape prior inferred from the past segmentations, with the knowledge learned from the training set.

This paper is structured as follows: In Section 2, we introduce the cardiac shape decomposition algorithm and dynamic prediction algorithm. In Section 3, we introduce a Bayesian formulation for the cardiac sequence segmentation that integrates the intensity information and SSDM. In Section 4, we provide experimental results and compare them with those obtained with the static model (SM) and GDM. Finally in Section 5, we conclude this paper. In Appendix A and B, we review the mathematical foundations of multilinear algebra, MPCA, and MICA.

2. The construction of SSDM

2.1. Shape Alignment

Training Data We acquired 32 sequences of breath-hold canine MR images from a GE signa 1.5 Tesla scanner. The image acquisition was triggered by electrocardiographic (ECG) signals recorded by a MR-compatible sensing device to suppress noises and artifacts induced by gradient magnetic fields. Axial images through the LV were obtained with gradient echo cine technique. The resulting 3-D image set consists of 16 temporal frames per cardiac cycle, with an in-plane resolution of 1.6mm and a slice thickness of 5mm. The cardiac sequences were temporally aligned using ECG signals. The ENDO- and EPI surfaces were manually outlined by cardiology collaborator using the BioImage Suite Software [8] and then sampled to 0.5 voxel resolution. In this work, we performed "Leave-one-out" test that alternately selects 31 sequences to build the SSDM and leaves one sequence to validate the automatic segmentation algorithm.

Alignment The second stage is to align cardiac shapes to a common coordinate system. In particular, we mapped all frames of each sequence to the first frame of the first sequence by performing inter-subject and intra-subject registrations as shown in Figure 2. We used an affine transform to account for the global shape difference, followed by a shape-based non-rigid transform that minimizes the difference in principal curvatures to accommodate detailed shape differences. The accuracy of the shape-based non-rigid transform was validated using implanted markers in [9]

Landmark Extraction and Propagation In the third stage, we first extracted 1095 landmarks on the ENDO surface and 1200 landmarks on the EPI surface in the first frame of the first sequence. We then propagate this set of landmarks to all frames of each sequence using the affine and non-rigid transforms obtained in the second stage. We illustrate this process in Figure 2. Finally, we obtained 2295 landmarks for each frame, and 16 frames in each sequence.

2.2. Cardiac Shape Decomposition

Appendix A reviews the basic mathematics of multilinear algebra, and appendix B overviews the fundamentals of MPCA and MICA. We use a third-order tensor to $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$ to represent the aligned cardiac shapes obtained in Section 2.1, where $I = 31$ is the number of subjects, $J = 16$ is the number of frames within a sequence, and $K = 2295 \times 3 = 6885$ is the dimension of the landmark vectors. By applying MPCA to tensor $\mathcal{X}$, we obtain

$$\mathcal{X} \simeq \mathcal{S} \times_1 \mathbf{U}^{\text{subject}} \times_2 \mathbf{U}^{\text{motion}} \times_3 \mathbf{U}^{\text{landmark}} \quad (1)$$

where $\mathcal{S} \in \mathbb{R}^{P \times Q \times R}$ is the core tensor which represents the interaction of the subject, motion, and landmark subspaces. Matrices $\mathbf{U}^{\text{subject}} \in \mathbb{R}^{I \times P}$, $\mathbf{U}^{\text{motion}} \in \mathbb{R}^{J \times Q}$ and $\mathbf{U}^{\text{landmark}} \in \mathbb{R}^{K \times R}$ are the subject subspace, motion subspace, and landmark subspace respectively. Matrix $\mathbf{U}^{\text{subject}}$ contains row vectors $\mathbf{u}_i^{\text{subject}} \in \mathbb{R}^P$ $(1 \leq i \leq I)$ of coeffi-

intra-subject registration
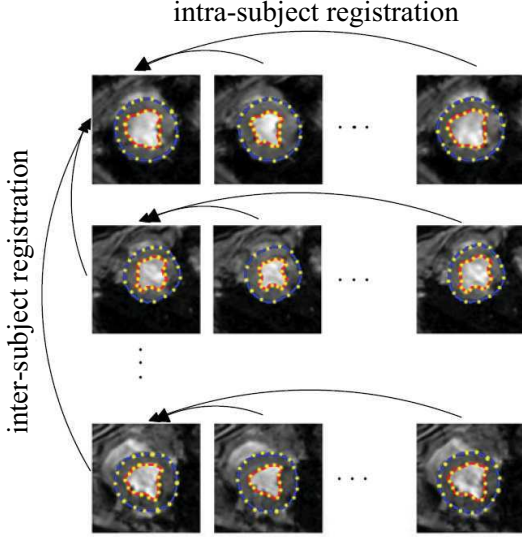
inter-subject registration

Figure 2. Overview of landmark propagation. All individual data sets are matched to an atlas, i.e. the first frame of the first sequence, via inter- and intra-subject registrations. The landmarks in the atlas are then copied to the individual subjects.

cients for each person i, and matrix $\mathbf{U}^{\mathrm{motion}}$ contains row vectors $\mathbf{u}_j^{\mathrm{motion}} \in \mathbb{R}^Q$ $(1 \le j \le J)$ for frame j.

While it is reasonable to apply PCA in the subject subspace, it is not appropriate to use it in the motion subspace because the deformation of cardiac shapes does not have a Gaussian distribution. Decomposition of the motion subspace using PCA would result in a set of mixed modes, as shown in Figure 3. To handle this problem, we adopt ICA in the motion subspace to obtain a set of independent modes in the motion subspace.

$$
\begin{aligned}
\mathcal{X} &\simeq \mathcal{S} \times_1 \mathbf{U}^{\mathrm{subject}} \times_2 \mathbf{U}^{\mathrm{motion}} \times_3 \mathbf{U}^{\mathrm{landmark}} \\
&= \mathcal{S} \times_1 \mathbf{U}^{\mathrm{subject}} \times_2 \mathbf{U}^{\mathrm{motion}} \mathbf{W}^T \mathbf{W}^{-T} \times_3 \mathbf{U}^{\mathrm{landmark}} \\
&= \tilde{\mathcal{S}} \times_1 \mathbf{U}^{\mathrm{subject}} \times_2 \tilde{\mathbf{U}}^{\mathrm{motion}} \times_3 \mathbf{U}^{\mathrm{landmark}} \qquad (2)
\end{aligned}
$$

where the core tensor $\tilde{\mathcal{S}} = \mathcal{S} \times_2 \mathbf{W}^{-T}$, the column vectors of $\tilde{\mathbf{U}}^{\mathrm{motion}}$ are the independent components of the motion subspace $\tilde{\mathbf{U}}^{\mathrm{motion}}$.

To reduce dimensions, 1) we select the complete eigenvectors in the motion subspace, i.e. $Q = J$, and perform MPCA in the subject and landmark subspaces to find the optimal P and R such that the approximation keeps more than 98% of the original energy; 2) we fix P and R, and perform MICA in the motion subspace and choose Q according to following criteria:

1. All selected modes correspond to significant shape variations;

2. After decomposition, more than 96% of the original energy is retained.

In practice, we reduced $I = 31$ to $P = 5$ and $K = 6885$ to $R = 15$ in the first step, and further reduced $J = 16$ to $Q = 3$ in the second step. We kept around 98.2% of the original energy in the first step and 96.6% of the original energy in the second step. In the second step, we observed three significant modes of shape variation: short-axial contraction, twisting of the heart, and long-axial contraction, as shown in Figure 4.
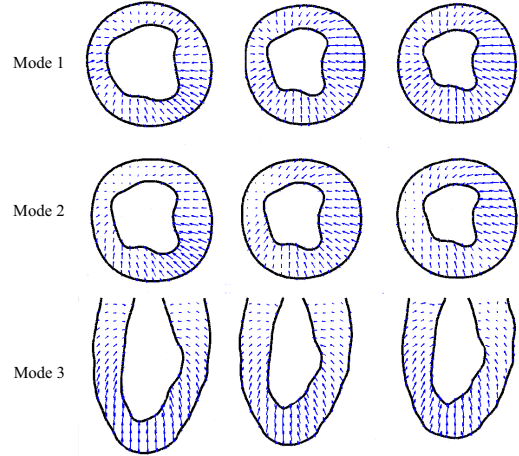

Figure 3. The mixed-mode decomposition in the motion subspace obtained with PCA. All three modes contain short-axial contraction. The dense displacement fields are obtained using the shape-based tracking algorithm described in [9].
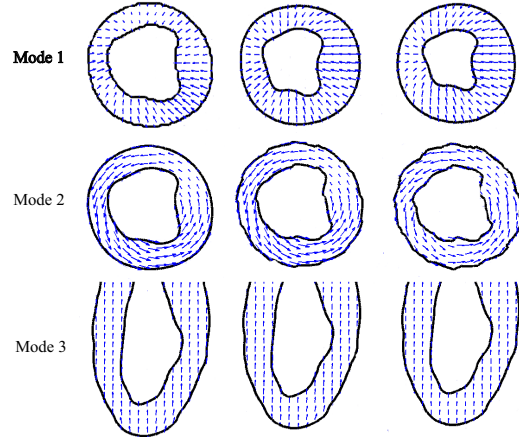

Figure 4. The significant modes of variation in the motion subspace obtained with ICA. Mode 1: short-axial contraction; Mode 2: twisting of myocardium; Mode 3: long-axial contraction.

2.3. Dynamic Prediction

In Section 2.2, we decomposed the cardiac shape training set into the subject subspace $\mathbf{U}^{\mathrm{subject}}$, motion subspace $\tilde{\mathbf{U}}^{\mathrm{motion}}$, and landmark subspace $\mathbf{U}^{\mathrm{landmark}}$. Thus, we can represent a new cardiac shape using

$$\mathbf{x} = \tilde{S} \times_1 \mathbf{u}^{\text{subject}} \times_2 \tilde{\mathbf{u}}^{\text{motion}} \times_3 \mathbf{U}^{\text{landmark}} \quad (3)$$

where $\mathbf{u}^{\text{subject}}$ is the subject vector and $\tilde{\mathbf{u}}^{\text{motion}}$ is the motion vector. Given the segmentations of a new cardiac sequence from frame 1 to frame $t-1$, we want to predict its segmentation in frame t. The idea is to first project the given segmentations to the subject subspace to identify the subject vector associated with this sequence, and then use this subject vector to construct the LV shape in frame t.

Let $\mathbf{x}_{1:t-1} = \{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_{t-1}\}$ denote the observed segmentation of a new cardiac sequence, we can predict the segmentation at frame t as follows:

- **Projection** The subject vector associated with this sequence can be represented as

$$\hat{\mathbf{u}}^{\text{subject}} = \mathbf{x}_{1:t-1} \cdot \mathbf{T}_{(1)}^{-1} \quad (4)$$

where $\mathbf{T}_{(1)}$ is the mode-1 unfolding of tensor $\mathcal{T} = \tilde{\mathcal{S}} \times_2 \tilde{\mathbf{u}}_{1:t-1}^{\text{motion}} \times_3 \mathbf{U}^{\text{landmark}}$.

- **Prediction** With this subject vector, we can predict the segmentation in frame t as

$$\mathbf{x}_t^* = \tilde{\mathcal{S}} \times_1 \hat{\mathbf{u}}^{\text{subject}} \times_2 \tilde{\mathbf{u}}_t^{\text{motion}} \times_3 \mathbf{U}^{\text{landmark}} \quad (5)$$

3. Cardiac Segmentation

Assume that we are given a cardiac sequence $I_{1:t} : \Omega \to \mathbb{R}$, where $I_{1:t}$ denotes the set of images $\{I_1, I_2, ..., I_t\}$ at different frames. Also, let $\mathbf{C}_t = \{\mathbf{C}_t^+, \mathbf{C}_t^-\}$ be the myocardial contour at frame t, where $\mathbf{C}_t^+$ is the ENDO surface and $\mathbf{C}_t^-$ is the EPI surface. Using the Bayesian formula, the problem of segmenting the current frame t can be addressed by maximizing the conditional probability

$$\begin{aligned}
\hat{\mathbf{C}}_t &= \arg\max_{\mathbf{C}_t} \mathcal{P}\left(\mathbf{C}_t | I_{1:t}\right) \\
&= \arg\max_{\mathbf{C}_t} \mathcal{P}\left(I_t | \mathbf{C}_t, I_{1:t-1}\right) \mathcal{P}\left(\mathbf{C}_t | I_{1:t-1}\right) \\
&= \arg\max_{\mathbf{C}_t} \mathcal{P}\left(I_t | \mathbf{C}_t, I_{1:t-1}\right) \int \mathcal{P}\left(\mathbf{C}_t | \mathbf{C}_{1:t-1}\right) \\
&\qquad \mathcal{P}\left(\mathbf{C}_{1:t-1} | I_{1:t-1}\right) d\mathbf{C}_{1:t-1} \\
&\overset{(a)}{=} \arg\max_{\mathbf{C}_t} \underbrace{\mathcal{P}\left(I_t | \mathbf{C}_t\right)}_{\text{data adherence}} \underbrace{\mathcal{P}\left(\mathbf{C}_t | \hat{\mathbf{C}}_{1:t-1}\right)}_{\text{dynamic prior}} \quad (6)
\end{aligned}$$

As explained in [2], we make two assumptions at step (a) of equation 6, which lead to a computationally more feasible problem:

- We assume that $I_{1:t}$ are mutually independent, i.e. $\mathcal{P}\left(I_t | \mathbf{C}_t, I_{1:t-1}\right) = \mathcal{P}\left(I_t | \mathbf{C}_t\right)$;

- We assume that the distributions of previous states to be strongly peaked around the maxima of the respective distributions, i.e. $\mathcal{P}\left(\mathbf{C}_{1:t-1} | I_{1:t-1}\right) = \delta\left(\mathbf{C}_{1:t-1} - \hat{\mathbf{C}}_{1:t-1}\right)$, where $\hat{\mathbf{C}}_i = \arg\max \mathcal{P}\left(\mathbf{C}_i | I_{1:i}\right)$ are the estimates of myocardial contours in the past, and $\delta(\cdot)$ denotes the Dirac Delta function [2].

Thus, equation 6 defines a recursive Bayesian formula to implement sequential cardiac segmentation, in which the segmentation of each particular time frame is based not only on the image data at that instant, but on the predictions from previous frames as well.

3.1. Data Adherence

Due to their robustness, region-based segmentation methods have been successfully applied to the segmentation of images with weak edges. Here, in our approach, we evolve a region-based deformable model based on the statistical intensity distribution from cardiac MR images.

To decide the intensity distribution of cardiac MR images, the entire image is partitioned by the ENDO- and EPI surfaces into three regions-LV blood pool, LV myocardium, and background, as shown in Figure 5. The LV blood pool and myocardium are homogeneous, and therefore can be modeled with a unimodal distribution. The most common distribution for MR images is Gaussian (Normal) distribution.

$$\mathcal{P}\left(I; \mu_l, \sigma_l\right) = \frac{1}{\sqrt{2\pi}\sigma_l} \exp\left\{-\frac{(I - \mu_l)^2}{2\sigma_l^2}\right\} \quad (7)$$

where μ_l is the mean of Gaussian distribution, and σ_l is its variation. For $l = 1$, it models the intensity distribution in LV blood pool. For $l = 2$, it models the intensity distribution in LV myocardium.
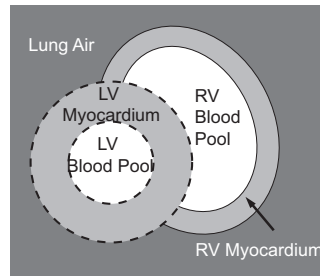

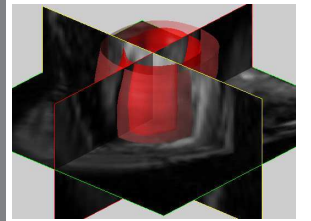Figure 5. The short-axis view of the heart. Two dotted circles are ENDO- and EPI contours.

Figure 6. An example of volumetric MR data set with automatically extracted ENDO- and EPI surfaces.

The background, however, is inhomogeneous (see Figure 5) because it contains more than one tissue (RV blood pool, RV myocardium, lung air, and etc). Therefore, modeling it with a single distribution would be insufficient. To handle this problem, we used a mixture model and invoke Expectation Maximization (EM) algorithm to fit the background histogram.

$$\mathcal{P}_3\left(I;\mu_3,\sigma_3\right)=\sum_{k=1}^{M}\alpha_k\mathcal{P}_{3,k}\left(I;\mu_{3,k},\sigma_{3,k}\right) \qquad (8)$$

where M is the number of components, α_k is the mixture proportion of component k that satisfies $\sum_{k=1}^{M}\alpha_k=1$, $\mu_{3,k}$ and $\sigma_{3,k}$ are the mean and deviation of its component Gaussian distribution. We show in Figure 7 the histogram of each region with fitted distribution function.
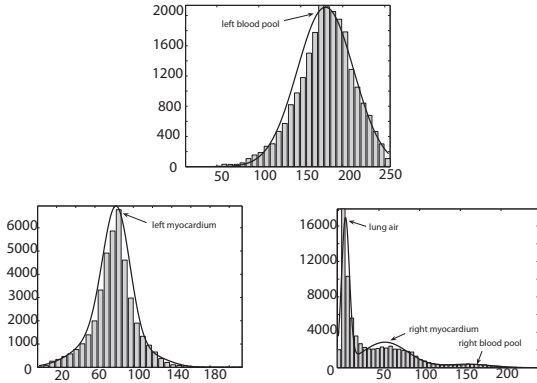

Figure 7. The histograms of LV blood pool, LV myocardium, and background with their fitted distributions.

Let $\Omega_{t,1}$, $\Omega_{t,2}$, and $\Omega_{t,3}$ denote the LV blood pool, LV myocardium and background in each frame t. Thus, the data adherence term can be defined as a log-likelihood function as follow:

$$\log\mathcal{P}\left(I_t|\mathbf{C}_t\right) = \log\mathcal{P}\left(I_t|\mathbf{C}_t^+,\mathbf{C}_t^-\right)$$
$$= \sum_{l=1}^{3}\int_{\Omega_{t,l}}\log\mathcal{P}\left(I_t;\mu_l,\sigma_l\right)d\mathbf{x} \quad (9)$$

The maximization of equation 9 can be interpreted as the propagation of $\mathbf{C}_t^+$ and $\mathbf{C}_t^-$ that maximizes the piecewise homogeneities.

3.2. Dynamic Prior

As shown in Section 2.3, we predict the ENDO- and EPI contours at frame t based on the previous segmentations using equation 4 and 5. Thus, we define the dynamic prior as follows:

$$\mathcal{P}\left(\mathbf{C}_t|\hat{\mathbf{C}}_{1:t-1}\right)\propto\exp\left\{-\alpha\int\|\mathbf{C}_t-\mathbf{C}_t^*\|^2\,d\mathbf{x}\Big/2\right\} \qquad (10)$$

where α is a weighting parameter, and $\mathbf{C}_t^*$ is the predicted contour using the dynamic prediction algorithm described in Section 2.3. We found in the experiment that $0.5\leq\alpha\leq2$ is applicable to most of the data sets.

3.3. Optimization

Combining equation 9 and 10, the maximization of the posterior probability defined in the Bayesian framework can be identified by the following Euler-Lagrange equations:

$$\frac{\partial\mathbf{C}_t^+}{\partial\tau} = \log\left(\frac{\mathcal{P}\left(I_t;\mu_2,\sigma_2\right)}{\mathcal{P}\left(I_t;\mu_1,\sigma_1\right)}\right)\mathbf{n}_t^+-\alpha\left(\mathbf{C}_t^+-\mathbf{C}_t^{*+}\right)$$
$$\frac{\partial\mathbf{C}_t^-}{\partial\tau} = \log\left(\frac{\mathcal{P}\left(I_t;\mu_3,\sigma_3\right)}{\mathcal{P}\left(I_t;\mu_2,\sigma_2\right)}\right)\mathbf{n}_t^--\alpha\left(\mathbf{C}_t^--\mathbf{C}_t^{*-}\right)$$

where $\mathbf{n}_t^+$ and $\mathbf{n}_t^-$ are the normals of $\mathbf{C}_t^+$ and $\mathbf{C}_t^-$ respectively, and τ is the time step used to numerically solve the above equations. Thus, we can recursively segment the ENDO- and EPI contours for each frame, based on the intensity information from the current frame and the prediction from previous frames.

4. Experiments

In this section, we show the experimental results from our algorithm and compare them with those obtained from the segmentation with the SM and GDM. As mentioned in Section 2.1, we adopted "Leave-one-out" test that alternately chooses one out of 32 sequences to validate our algorithm. Figure 8 shows the short-axis view of automatically segmented ENDO-and EPI surfaces on consecutive frames during ventricular systole. Figure 6 shows an example of volumetric MRI data set with automatically extracted ENDO-and EPI surfaces.
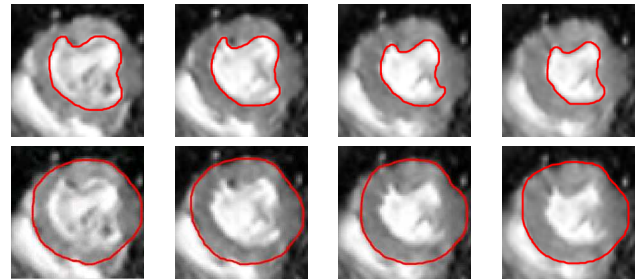

Figure 8. Short-axis view of ENDO-and EPI surfaces from MRI images during ventricular systole. Top: ENDO, Bottom: EPI.

To quantify the accuracy of our approach, we used two distance error metrics and two area error metrics: mean ab-

solute distance (MAD), Hausdorff distance (HD), percentage of true positives (PTP), and percentage of false positives (PFP).

Let A and B be two surfaces from the automatic segmentation and manual segmentation respectively. Suppose they are represented as point sets, i.e. $A = \{\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n\}$ and $B = \{\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_m\}$, we define MAD and HD as follows

$$
\begin{aligned}
\mathrm{MAD}\,(A, B) &= \frac{1}{2}\left\{\frac{1}{n}\sum_{i=1}^{n} d\,(\mathbf{a}_i, B) + \frac{1}{m}\sum_{j=1}^{m} d\,(\mathbf{b}_j, A)\right\} \\
\mathrm{HD}\,(A, B) &= \max\left\{\max_i d\,(\mathbf{a}_i, B), \max_j d\,(\mathbf{b}_j, A)\right\}
\end{aligned}
$$

where $d\,(\mathbf{a}_i, B) = \min_j |\mathbf{b}_j - \mathbf{a}_i|$. While MAD is a global measure of the match of two surfaces, HD reflects their local similarities.

Let Ω_A and Ω_B be two regions enclosed by the surfaces A and B respectively, we define PTP and PFP as follows

$$
\begin{aligned}
\mathrm{PTP} &= \frac{Volume\,(\Omega_A \cap \Omega_B)}{Volume\,(\Omega_B)} \\
\mathrm{PFP} &= \frac{Volume\,(\Omega_B) - Volume\,(\Omega_A \cap \Omega_B)}{Volume\,(\Omega_A)}
\end{aligned}
$$

4.1. Static versus Dynamical Models

Figure 9 compares the experimental results with the SM and SSDM. Table 1 and 2 use MAD, HD, PTP, and PFP to quantitatively analyze the segmentation results on ENDO- and EPI surfaces respectively. We observe that both MAD and HD decrease with the SSDM, which implies that SM is biased in the global sense because it is trapped easily in the local minima. Also, we observe the SSDM improves ENDO segmentation more than EPI segmentation because the epicardium does not move as significantly as the endocardium when the heart beats.
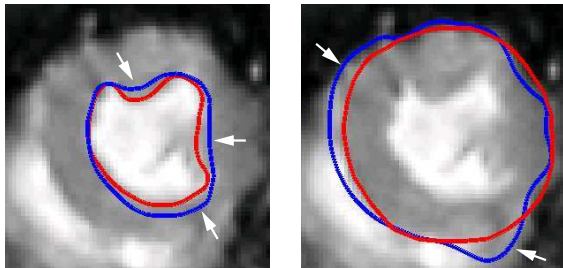

Figure 9. Comparing segmentation results using SM (blue) and SSDM (red). It can be seen that the segmentation with SM, when applied to noisy images, can easily be trapped in local minima.

Moreover, we calculated the MAD between the predicted contour and the contour from manual segmentation

Table 1. Comparison with the SM for segmentation of ENDO surfaces

	SM	SSDM
MAD(mm)	2.34 ± 0.12	0.60 ± 0.08
HD(mm)	3.54 ± 0.15	1.54 ± 0.12
PTP(%)	87.6 ± 2.44	98.7 ± 0.54
PFP(%)	12.4 ± 2.34	1.21 ± 0.49

Table 2. Comparison with the SM for segmentation of EPI surfaces

	SM	SSDM
MAD(mm)	3.23 ± 0.14	1.23 ± 0.09
HD(mm)	4.35 ± 0.21	1.67 ± 0.15
PTP(%)	84.2 ± 2.42	95.6 ± 0.89
PFP(%)	15.7 ± 2.43	4.5 ± 0.77

for each frame during ventricular systole (see Figure 10). We took the mean shape of SM as its predicted contour. We observe that SSDM progressively improves the prediction accuracy as additional segmentations from previous frames become available. In contrast, SM is most accurate in the middle of cardiac systole, at which time point the LV shape is closet to the mean shape of SM.
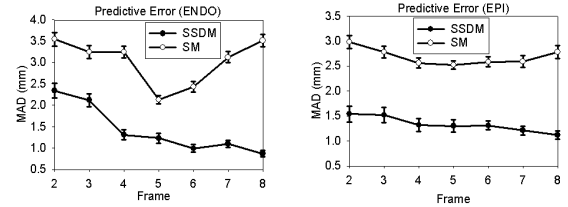

Figure 10. Comparing the prediction errors from the SM and SSDM.

4.2. Generic versus Subject-Specific Dynamical Models

In Figure 11, we compare the experimental results with GDM and SSDM. Table 3 and 4 use MAD, HD, PTP and PFP to present the quantitative analysis of the segmentation results on ENDO- and EPI surfaces respectively. We observe that while the MADs from GDM and SSDM are quite close, the HD from SSDM is significantly smaller than that from GDM. It implies that while GDM can capture the global deformation, it does not reflect the local shape variations.

Table 3. Comparison of GDM and SSDM for segmentation of ENDO surfaces

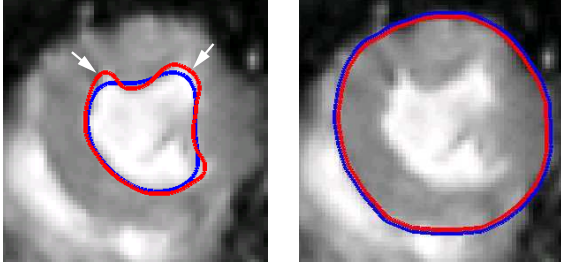|	GDM	SSDM
MAD (mm)	0.98 ± 0.09	0.60 ± 0.08
HD (mm)	3.51 ± 0.15	1.54 ± 0.12
PTP (%)	95.7 ± 0.65	98.7 ± 0.54
PFP (%)	4.5 ± 0.65	1.21 ± 0.49

Figure 11. Comparing segmentation results using GDM (blue) and SSDM (red). While the GDM captures the global shape deformation, it does not characterize the local shape variations.

Table 4. Comparison of GDM and SSDM for segmentation of EPI surfaces

	GDM	SSDM
MAD (mm)	1.34 ± 0.10	1.23 ± 0.09
HD (mm)	3.30 ± 0.15	1.67 ± 0.15
PTP (%)	93.3 ± 0.98	95.6 ± 0.89
PFP (%)	6.7 ± 0.80	4.5 ± 0.77

We also computed the HD between the predicted contours and manually segmented contours for each frame during ventricular systole, as shown in Figure 12. We found that the SSDM progressively improves the predictive accuracy while the GDM produces an almost constant error.
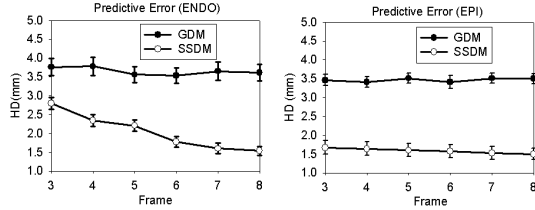

Figure 12. Comparing the prediction errors from GDM and SSDM.

5. Conclusion

In this paper, we have developed a subject-specific dynamical model (SSDM) that accounts for the fact that a particular cardiac shape in a given frame depends on the shapes observed in previous frames. In the training phase, we used MPCA and MICA to decompose the cardiac shape training set into subject subspace, motion subspace, and landmark subspace. In the prediction phase, we projected the segmentations in the past frames to the subject subspace to identify the subject vector associated with this cardiac sequence, followed by the prediction of cardiac shapes in the future frames with the estimated subject vector. The integration of the SSDM into the segmentation task is formulated on the basis of Bayesian framework, which combines the intensity information from the current frame and the dynamic prior from the past frames to estimate the LV bound-

aries. This process is repeated through a cardiac cycle.

Future work includes the development of a system that integrates segmentation and motion analysis into a single framework, where the motion module takes the segmented contours as input and estimates the spatially-dense myocardial displacement fields through a cardiac cycle based on the biomechanical model [9]. The estimated displacement fields can in turn be used to improve the segmentation results.

A. Multilinear Algebra

A tensor, also known as multi-dimensional array, is a higher order generalization of a vector (first-order tensor) and a matrix (second-order tensor). An element of tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times \ldots \times I_n}$ is denoted as $\mathcal{A}_{i_1 \ldots i_n}$ where $1 \leq i_n \leq I_n$. The mode-n fibers of tensor $\mathcal{A}$ are the I_n- dimensional vectors obtained from $\mathcal{A}$ by varying index i_n while keeping the other indices fixed. The mode-n unfolding of tensor $\mathcal{A}$ is defined by stacking its mode-n fibers in a particular order, as shown in Figure 13. We denote the n-mode unfolding of tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times \ldots \times I_n}$ as $\mathcal{A}_{(n)} \in \mathbb{R}^{I_n \times (I_{n+1} I_{n+2} \ldots I_N I_1 I_2 \ldots I_{n-1})}$. The n-mode product of tensor $\mathcal{A}$ by a matrix $\mathbf{U} \in \mathbb{R}^{J_n \times I_n}$, denotes by $\mathcal{A} \times_n \mathbf{U}$, is an $(I_1 \times I_2 \times \ldots \times I_{n-1} \times J_n \times I_{n+1} \times \ldots \times I_N)$-tensor of which the entries are given by

$$(\mathcal{A} \times_n \mathbf{U})_{i_1 i_2 \ldots i_{n-1} j_n i_{n+1} \ldots i_N} = \sum_{i_n} \mathcal{A}_{i_1 i_2 \ldots i_{n-1} i_n i_{n+1} \ldots i_N} \mathbf{U}_{j_n i_n}$$
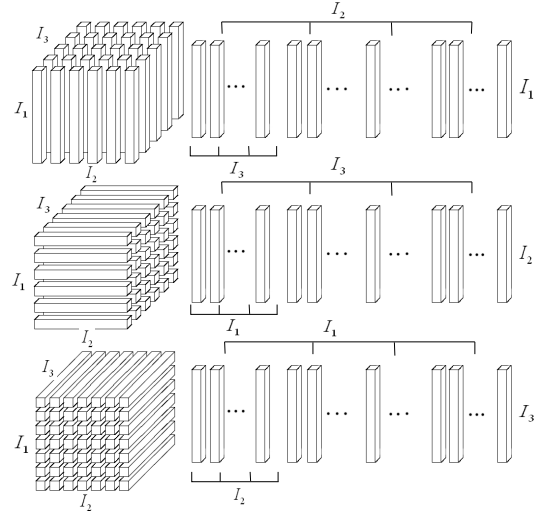
(11)

Figure 13. Unfolding of the $(I_1 \times I_2 \times I_3)$-tensor $\mathcal{A}$ to the $(I_1 \times I_2 I_3)$-matrix $\mathcal{A}_{(1)}$, the $(I_2 \times I_3 I_1)$-matrix $\mathcal{A}_{(2)}$ and the $(I_3 \times I_1 I_2)$-matrix $\mathcal{A}_{(3)}$; Left: Mode-1, 2, and 3 fibers; Right: Mode-1, 2, and 3 unfolded matrices.

B. Multilinear PCA and ICA

A matrix $\mathbf{A} \in \mathbb{R}^{I_1 \times I_2}$ is a second-order tensor that can be decomposed using singular value decomposition (SVD) as $\mathbf{A} = \mathbf{U}^{(1)}\mathbf{S}\mathbf{U}^{(2)T}$, where $\mathbf{U}^{(1)}$ and $\mathbf{U}^{(2)}$ are orthonormal matrices. In terms of the n-mode tensor products defined above, this matrix product can be rewritten as $\mathbf{A} = \mathbf{S} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)}$. By extension, an Nth-order tensor can be decomposed using multi-linear SVD (MSVD) [5] as follows

$$\mathcal{A} = \mathcal{S} \times_1 \mathbf{U}^{(1)}... \times_n \mathbf{U}^{(n)}... \times_N \mathbf{U}^{(N)} \qquad (12)$$

where tensor $\mathcal{S}$ is the core tensor governing the interaction between mode matrices $\mathbf{U}^{(n)}$ containing orthonormal vectors in the column space of $\mathbf{U}^{(n)}$.

In Principal Components Analysis (PCA), we can approximate matrix $\mathbf{A}$ by deleting the eigenvectors associated with the smaller singular values.

$$\mathbf{A} \simeq \mathbf{A}' = \mathbf{S}' \times_1 \mathbf{U}'^{(1)} \times_2 \mathbf{U}'^{(2)} \qquad (13)$$

Dimension reduction for Multilinear PCA (MPCA) is optimized in the least squares sense using Alternating Least Squares (ALS) method [5]. Analogously to equation 13, we can reconstruct tensor $\mathcal{A}$ as

$$\mathcal{A} \simeq \mathcal{A}' = \mathcal{S}' \times_1 \mathbf{U}'^{(1)}... \times_n \mathbf{U}'^{(n)}... \times_N \mathbf{U}'^{(N)} \qquad (14)$$

where $\mathcal{S}' \in R^{J_1 \times J_2 \times ... \times J_N}$ and $\mathbf{U}'^{(n)} \in \mathbb{R}^{I_n \times J_n}$ for all $n = 1, ..., N$.

Unlike PCA, Independent Component Analysis (ICA) looks for a linear combination of a set of statistically independent component whose non-Gaussianity is maximized. ICA can be computed from PCA solution through the rotation of the principal components such that they become independent [4]:

$$\mathbf{A}^T = \mathbf{U}^{(2)}\mathbf{S}\mathbf{U}^{(1)T} = \left(\mathbf{U}^{(2)}\mathbf{S}\mathbf{W}^{-1}\right)\left(\mathbf{W}\mathbf{U}^{(1)T}\right) = \mathbf{K}^T\mathbf{C}^T \qquad (15)$$

where $\mathbf{W}$ is an invertible transformation matrix computed by the ICA algorithm, $\mathbf{C} = \mathbf{U}^{(1)}\mathbf{W}^T$ are the independent components, and $\mathbf{K} = \mathbf{W}^{-T}\mathbf{S}\mathbf{U}^{(2)T}$ are the coefficients.

Analogously to equation 14, Multilinear ICA (MICA) is obtained using the relationship between SVD and ICA as shown above [13]:

$$
\begin{aligned}
\mathcal{A} &\simeq \mathcal{S} \times_1 \mathbf{U}^{(1)} \times ... \times \mathbf{U}^{(N)} \\
&= \mathcal{S} \times_1 \mathbf{U}^{(1)}\mathbf{W}_1^T\mathbf{W}_1^{-T}... \times_N \mathbf{U}^{(N)}\mathbf{W}_N^T\mathbf{W}_N^{-T} \\
&= \mathcal{S} \times_1 \tilde{\mathbf{U}}^{(1)}\mathbf{W}_1^{-T}... \times_N \tilde{\mathbf{U}}^{(N)}\mathbf{W}_N^{-T} \\
&= \left(\mathcal{S} \times_1 \mathbf{W}_1^{-T}... \times_N \mathbf{W}_N^{-T}\right) \times_1 \tilde{\mathbf{U}}^{(1)}... \times_N \tilde{\mathbf{U}}^{(N)} \\
&= \tilde{\mathcal{S}} \times_1 \tilde{\mathbf{U}}^{(1)}... \times_N \tilde{\mathbf{U}}^{(N)} \qquad (16)
\end{aligned}
$$

where the core tensor $\tilde{\mathcal{S}} = \mathcal{S} \times_1 \mathbf{W}_1^{-T}... \times_N \mathbf{W}_N^{-T}$.

References

[1] A. Blake and M. Isard. *Active Contours*. Springer, 1998. 1

[2] D. Cremers. Dynamical statistical shape priors for level set-based tracking. *IEEE TPAMI*, 28(8):1262–1273, 8 2006. 1, 4

[3] A. F. Frangi, W. J. Niessen, and M. A. Viergever. Three-dimensional modeling for functional analysis of cardiac images: A review. *IEEE TMI*, 20(1):2–25, 1 2001. 1

[4] J. Friedman, T. Hastie, and R. Tibshirani. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, New York, 2001. 8

[5] L. D. Lathauwer, B. D. Moor, and J. Vandewalle. A multilinear singular value decomposition. *SIAM J. Matrix Anal. Appl.*, 21(4):1253–1278, 2000. 8

[6] S. C. Mitchell, J. G. Bosch, B. P. F. Lelieveldt, R. J. van der Geest, J. H. C. Reiber, and M. Sonka. 3-d active appearance models: Segmentation of cardiac mr and ultrasound images. *IEEE TMI*, 21(9):1167–1178, 9 2002. 1

[7] S. C. Mitchell, B. P. F. Lelieveldt, R. J. van der Geest, H. G. Bosch, J. H. C. Reiber, and M. Sonka. Multistage hybrid active appearance models: Segmentation of cardiac mr and ultrasound images. *IEEE TMI*, 20(5):415–423, 5 2001. 1

[8] X. Papademetris, M. Jackowski, N. Rajeevan, R. T. Constable, and L. H. Staib. Bioimage suite: An integrated medical image analysis suite, section of bioimaging sciences, department of diagnostic radiology, yale school of medicine. http://www.bioimagesuite.org. 2

[9] X. Papademetris, A. J. Sinusas, D. P. Dione, R. T. Constable, and J. S. Duncan. Estimation of 3-d left ventricular deformation from medical images using biomechanical models. *IEEE TMI*, 21(7):786–800, 7 2002. 2, 3, 7

[10] D. Perperidis, R. Mohiaddin, and D. Rueckert. Construction of a 4d statistical atlas of the cardiac anatomy and its use in classification. In *MICCAI*, pages 402–410, 2005. 1

[11] W. Sun, M. Cetin, R. Chan, V. Reddy, G. Holmvang, V. Chandar, and A. Willsky. Segmenting and tracking of the left ventricle by learning the dynamics in cardiac images. In *IPMI*, pages 553–565, 2005. 1

[12] M. A. O. Vasilescu and D. Terzopoulos. Multilinear analysis of image ensembles: Tensorfaces. In *ECCV*, pages 447–460, 2002. 2

[13] M. A. O. Vasilescu and D. Terzopoulos. Multilinear independent component analysis. In *CVPR*, volume 1, pages 547–553, 2005. 2, 8