Segmenting and Tracking the Left Ventricle by Learning the Dynamics in Cardiac Images[*]

W. Sun[1], M. Çetin[1], R. Chan[2], V. Reddy[2], G. Holmvang[2], V. Chandar[1], and A. Willsky[1]

[1] Laboratory for Information and Decision Systems,
Massachusetts Institute of Technology, Cambridge, MA USA
[2] Cardiovascular MR-CT Program, Massachusetts General Hospital,
Harvard Medical School, Boston, MA USA
waltsun@mit.edu

Abstract. Having accurate left ventricle (LV) segmentations across a cardiac cycle provides useful quantitative (e.g. ejection fraction) and qualitative information for diagnosis of certain heart conditions. Existing LV segmentation techniques are founded mostly upon algorithms for segmenting static images. In order to exploit the dynamic structure of the heart in a principled manner, we approach the problem of LV segmentation as a recursive estimation problem. In our framework, LV boundaries constitute the dynamic system state to be estimated, and a sequence of observed cardiac images constitute the data. By formulating the problem as one of state estimation, the segmentation at each particular time is based not only on the data observed at that instant, but also on predictions based on past segmentations. This requires a dynamical system model of the LV, which we propose to learn from training data through an information-theoretic approach. To incorporate the learned dynamic model into our segmentation framework and obtain predictions, we use ideas from particle filtering. Our framework uses a curve evolution method to combine such predictions with the observed images to estimate the LV boundaries at each time. We demonstrate the effectiveness of the proposed approach on a large set of cardiac images. We observe that our approach provides more accurate segmentations than those from static image segmentation techniques, especially when the observed data are of limited quality.

1 Introduction

Of the cardiac chambers in the heart, the left ventricle (LV) is quite frequently analyzed because its proper function, pumping oxygenated blood to the entire body, is vital for normal activity. One quantitative measure of the health of the LV is ejection fraction (EF). This statistic measures the percentage volume of blood transmitted out of the LV in a given cardiac cycle. To compute EF, we need to have segmentations of the LV at multiple points in a cardiac cycle; namely, at end diastole (ED) and end systole (ES). In addition, observing how the LV

evolves throughout an entire cardiac cycle allows physicians to determine the health of the myocardial muscles. Segmented LV boundaries can also be useful for further quantitative analysis. For example, past work [7, 18] on extracting the flow fields of the myocardial wall assumes the availability of LV segmentations throughout the cardiac cycle.

Automatic segmentation of the left ventricle (LV) in bright blood breath-hold cardiac magnetic resonance (MR) images is non-trivial because the image intensities of the cardiac chambers vary due to differences in blood velocity [24]. In particular, blood that flows into the ventricles produces higher intensities in the acquired image than blood which remains in the ventricles [9]. Locating the LV endocardium is further complicated by the fact that the right ventricle and aorta often appear jointly with the LV in many images of the heart. Similarly, automatic segmentation of low signal-to-noise ratio (SNR) cardiac images (e.g. body coil MR or ultrasound) is difficult because intensity variations can often obscure the LV boundary.

Several approaches exist for LV segmentation. Goshtasby and Turner [9], as well as Weng *et al.* [27] and Geiger *et al.* [8], apply intensity thresholding and then a local maximum gradient search to determine the final segmentation. Such gradient-based methods rely primarily on local information. When the image statistics inside an object's boundary are distinctly different from those outside, the use of region statistics may be more appropriate, especially if the discontinuity at the boundary is weak or non-uniform. Tsai *et al.* [26] consider region-based segmentations of the LV. Chakraborty *et al.* [3] consider combining both gradient and region techniques in the segmentation of cardiac structures. Similarly, Paragios [20] uses gradient and region techniques to segment two cardiac contours, the LV endocardium and epicardium. In all three papers, active contours (or curve evolution) [2, 4, 13, 15, 16, 19], a technique which involves evolving a curve to minimize (or maximize) a related objective functional, are used to determine the segmentation. In our work, we also take advantage of region-based information and curve evolution.

Static segmentation methods are limited by the data available in an individual frame. During a single cardiac cycle, which lasts approximately 1 second, the heart contracts from end diastole (ED) to end systole (ES) and expands back to ED. Over this time, MR systems can acquire approximately 20 images of the heart. Because adjacent frames are imaged over a short time period (approximately 50 ms), the LV boundaries exhibit strong temporal correlation. Thus, previous LV boundaries may provide information regarding the location of the current LV boundary. Using such information is particularly useful for low SNR images, where the observation from a single frame alone may not provide enough information for a good segmentation. There exists some past work which simply uses the previous frame's LV boundary as the prediction for the boundary in the current frame [8, 12]. Meanwhile, Zhou *et al.* [28] consider LV shape tracking by combining predictions, obtained through linear system dynamics assumed known, with observations. Their technique uses landmark points to represent the LV boundaries, introducing the issue of correspondence. All uncertainties are as-
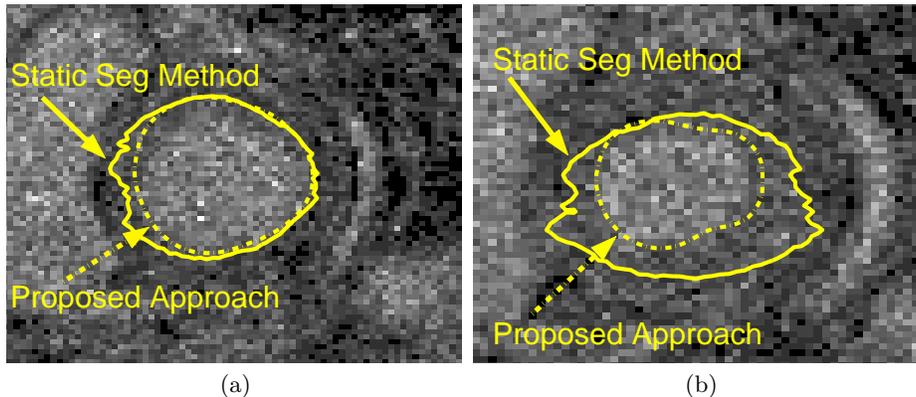
Fig. 1. (a) Segmentation of the fourth frame in the cardiac cycle. (b) Segmentation of the eighth frame (near end systole) in the cardiac cycle.

sumed to be Gaussian. Senegas *et al.* [21] use a Bayesian framework for tracking using a sample-based approach to estimate the densities. They use spherical harmonics for the shape model, and a simple linear model to approximate the cardiac dynamics.[3] In our work, we use non-linear dynamics in the recursive estimation of the LV boundary. We represent the LV by level sets to avoid issues inherent with marker points [23] and apply principal components analysis on the level sets to determine a basis to represent the shapes. Furthermore, we propose a method for learning a non-trivial dynamic model of the LV boundaries and apply this model to obtain predictions. Finally, we compute the maximum *a posteriori* (MAP) estimate using curve evolution.

In particular, we propose a principled Bayesian approach for recursively estimating the LV boundaries across a cardiac cycle. In our framework, LV boundaries constitute the dynamic system state we estimate, and a cardiac cycle of mid-ventricular images constitutes the data. From a training set of data, we learn the dynamics using an information-theoretic criterion [11]. More specifically, this involves finding a non-parametric density estimate of the current boundary conditioned on previous boundaries. The densities are approximately represented by using sample-based (i.e. particle filtering [1]) methods.

For the test data, we apply a particle filter to recursively estimate the LV boundary. Starting with the segmentations at the initial frames, we use the non-linear dynamic model learned from the training data to predict the boundary at the next frame. We then incorporate the image observation of this frame to produce a posterior density estimate of the LV boundary at each frame, which involves computing the MAP estimate at each frame using curve evolution. This procedure is then repeated for each subsequent frame. We apply the proposed algorithm to high and low SNR cardiac data to illustrate that our technique works in both regimes. We also demonstrate the improvements provided by the proposed method over results obtained from static LV segmentation methods, as shown in Figure 1.

---

[3] We thank the reviewer who brought Senegas *et al.*'s related work to our attention.
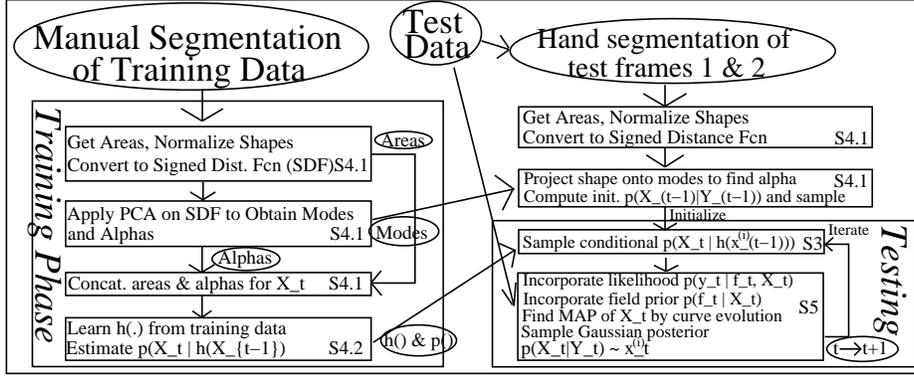
Fig. 2. Block diagram of our technique illustrating both the training and testing phases. Section data inside each block indicate where we describe the specific actions in text.

2 Framework and Methodology

We formulate the LV segmentation and tracking problem as an estimation of the posterior distribution at each time t_0 based on data from $t = 1$ to $t = t_0$. First, let y_t be the image data which are measurements of the blood and tissue intensity field f_t. Then, define X_t as the dynamic system state which contains information about the LV boundary at t. The segmentation problem involves finding the (f_t, X_t) pair which maximizes $p(f_t, X_t|Y_t)$, where $Y_t = [y_1, y_2, \ldots, y_t]$. We then recursively compute X_t to track the LV boundary across the entire cardiac cycle.

We apply Bayes' Theorem to the posterior $p(f_t, X_t|Y_t)$. Assuming that X_t is a Markov process and observing that $p(Y_{t-1})$ and $p(Y_t)$ do not depend on X_t,

$$p(f_t, X_t|Y_t) \propto p(y_t|f_t, X_t)p(f_t|X_t)p(X_t|Y_{t-1}) \tag{1}$$

$$= p(y_t|f_t, X_t)p(f_t|X_t)\int_{X_{t-1}} p(X_t|X_{t-1})\int_{f_{t-1}} p(f_{t-1}, X_{t-1}|Y_{t-1})df_{t-1}dX_{t-1},$$

where $p(y_t|f_t, X_t)$ is the likelihood term, $p(f_t|X_t)$ is the field prior, and $p(X_t|Y_{t-1})$ is the prediction density. From Eqn. (1), we observe the recursive nature of the problem (i.e. $p(f_t, X_t|Y_t)$ is written as a function of $p(f_{t-1}, X_{t-1}|Y_{t-1})$).

Given this framework, applying it to the LV tracking problem is not straightforward. One of the challenges involves the presence of arbitrary, non-Gaussian probability densities. In Section 3, we discuss the use of a sample-based approach to non-parametrically represent the densities in Eqn. (1). In addition, the dynamic model of the LV boundaries, hence the forward density $p(X_t|X_{t-1})$, needs to be learned using statistics from the training data. We discuss the procedure for learning in Section 4. Finally, we explain in Section 5 how we practically compute the MAP estimate of X_t and use this information to produce a segmentation as well as an estimate of the posterior $p(f_t, X_t|Y_t)$. Experimental results are shown in Section 6, and we summarize the work in Section 7. Figure 2 shows a block diagram representation of the algorithmic framework we propose.

3 Sample-Based Methods

Because many of the densities in Eqn. (1) have no simple closed-form, we use sample-based methods, such as particle filters [1, 5, 6, 14], to approximate these densities. Such methods represent a probability density using a set of weighted samples drawn from that density. Suppose we have an equally-weighted set of N samples $x_{t-1}^{(i)}$ that represent $p(X_{t-1}|Y_{t-1})$, a term which appears as part of $p(X_t|Y_{t-1})$ in the formulation according to

$$p(X_t|Y_{t-1}) = \int_{X_{t-1}} p(X_t|X_{t-1})p(X_{t-1}|Y_{t-1})dX_{t-1}. \tag{2}$$

From the conditional distribution $p(X_t|X_{t-1})$ (assume known for now), we next obtain M samples $x_{t|t-1}^{(i,j)}$ from $p(X_t|X_{t-1} = x_{t-1}^{(i)})$ for each i. Since the sample points for $p(X_{t-1}|Y_{t-1})$ are equally-weighted, $p(X_t|Y_{t-1})$ can similarly be approximated by the $N \times M$ equally-weighted samples $x_{t|t-1}^{(i,j)}$.

To complete the recursion as shown in Eqn. (1), we make an approximation for the marginalization of f_{t-1}. In particular, we choose the f_{t-1} which maximizes the posterior rather than marginalizing over f_{t-1}. Thus, we have

$$p(X_{t-1}|Y_{t-1}) = \int_{f_{t-1}} p(f_{t-1}, X_{t-1}|Y_{t-1})df_{t-1} \approx \max_{f_{t-1}} p(f_{t-1}, X_{t-1}|Y_{t-1}). \tag{3}$$

In the above discussion, we have described how, given $p(f_{t-1}, X_{t-1}|Y_{t-1})$, we can obtain $p(X_t|Y_{t-1})$ assuming $p(X_t|X_{t-1})$ is known. In the next section, we explain how we estimate $p(X_t|X_{t-1})$ through learning the system dynamics.

4 Learning the Dynamics

A number of approaches can be taken to learn the dynamics of an evolving system. We can consider purely physics-based models to constrain and explain the dynamics of a given problem [17, 22]. The drawback is that systems that accurately model physics may require high dimensional states and/or a complex set of differential equations that model the interaction between adjacent masses in the system. Alternatively, we may assume a statistical model that can either be parametric or non-parametric. For the former, the challenge is to find a parametric model that is well-matched to the problem structure and captures the statistical variability inherent in the problem. For richer modeling capacity, one can turn to non-parametric models, which can be computationally difficult. In Section 4.2, we explain our non-parametric, yet computationally tractable approach to learning the dynamics of LV boundaries. Before discussing this method, we first provide a description of the system state X_t.

4.1 Implicit Parametric Shape Model and State Representation

The set of LV boundaries have different internal areas and different shapes across a cardiac cycle and between patients. We want to represent these boundaries in a simple, low-dimensional, yet accurate, manner. To accomplish this, we use principal components analysis (PCA) on the shape variability to obtain a basis for the shapes [16]. We then represent each shape by a linear combination of the
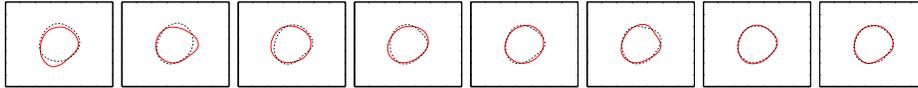
Fig. 3. Illustration of LV shape variability. $\pm 1\sigma$ for the first eight primary modes of variability (left to right). Solid curve represents $+1\sigma$ while dashed represents -1σ.

basis elements. The tracking problem reduces to learning the time evolution of the coefficients of the basis elements.

Starting with a training set of manually-segmented and registered data, we determine the area of each LV. Normalizing with respect to area, we create signed distance functions whose zero level sets are the shapes [23]. Leveraging on Leventon's PCA on shapes [16], we obtain a mean shape $\bar{\psi}$ and the primary modes of variability ψ_i (for i=1,2, …, K, where K is the number of shapes in the dataset) across the entire training set. In effect, we use a single basis to represent the shapes across the entire cardiac cycle. Figure 3 shows the eight primary modes of variability from the training set used in the experimental results presented in Section 6. For a given signed distance function ψ in the training set, $\psi = \bar{\psi} + \sum_{i=1}^{K} \alpha_i \psi_i$, where α_i's are a set of constants. It is known that for shapes which do not vary greatly, the primary few modes of variability can explain the majority of the variability of the data. In our training set, the first eight modes explain 97% of the variability in our specific training set of data. Thus, we approximately represent each ψ by the eight element vector $\boldsymbol{\alpha} = [\alpha_1; \alpha_2; \ldots; \alpha_8]^T$. By using PCA, a given curve (LV segmentation) can be approximately represented by a vector containing its area A and $\boldsymbol{\alpha}$.

Given this representation, we define the state X_t with the notion that the dynamics are a second-order system. This choice is made because higher-order systems require a larger state, while first-order systems do not adequately capture whether we are in the diastolic or systolic phase. Thus, we represent our state X_t as an eighteen-dimensional vector containing the area of the LV and the shape variabilities at frames t and $t-1$, namely $X_t = [A_t; \boldsymbol{\alpha}_t^T; A_{t-1}; \boldsymbol{\alpha}_{t-1}^T]^T$.

4.2 A Maximally-Informative Statistic

We propose learning the dynamics from a training set of data based on a technique [11] which produces a non-parametric density estimate of $p(X_t|X_{t-1})$. This estimate is obtained by using an information-theoretic criterion to maximize the predictive power of the observations.

Since the dimensionality of the conditional density may be large, we consider only the portion of the state X_{t-1} that is statistically pertinent to the prediction of X_t. Thus, we introduce a function $h(X_{t-1})$ which seeks to reduce dimensionality yet capture all information in X_{t-1} that relates to X_t (achieved exactly only when $I(X_t; X_{t-1}) = I(X_t; h(X_{t-1}))$, where $I(X_t, X_{t-1})$ is the mutual information between X_t and X_{t-1}). We can then create an estimate of $p(X_t|h(X_{t-1}))$ as an equally-informative yet simpler representation of $p(X_t|X_{t-1})$.

In practice, we constrain h to be linear, which likely precludes it from being a sufficient statistic. However, we choose the parameters of h such that

$I(X_t; h(X_{t-1}))$ is maximized, thus making h maximally-informative within this class. Details regarding h and the maximization are in [25]. Once the parameters of h are determined, we obtain a kernel density estimate of $p(X_t|h(X_{t-1}))$, where for kernel size we use the plug-in method of Hall *et al.* [10].

5 Finding the MAP Estimate by Curve Evolution

Now, we incorporate the data at time t to obtain the posterior $p(f_t, X_t|Y_t)$. Given equally-weighted samples $x_{t|t-1}^{(i,j)}$ for $p(X_t|Y_{t-1})$ as described in Section 3, one could in principle weight the particles by the likelihood and field priors to obtain a representation of $p(f_t, X_t|Y_t)$. Such an approach may work if the training data are rich. However, when we have a limited amount of training data, we make the assumption that the posterior distribution of X_t is Gaussian and determine this distribution by first computing its MAP estimate to determine the mean parameter (since we do not have a method in place to compute the posterior covariance, we approximate it to be a diagonal matrix with individual variances determined empirically from the shape variability in the training data). Maximizing $p(f_t, X_t|Y_t)$ to obtain the MAP estimate is equivalent to minimizing

$$E(f_t, X_t) = -\log p(y_t|f_t, X_t) - \log p(f_t|X_t) - \log p(X_t|Y_{t-1}), \qquad (4)$$

which involves a likelihood term $p(y_t|f_t, X_t)$, the prior on the field $p(f_t|X_t)$, and a prediction term $p(X_t|Y_{t-1})$. We discuss each term in Eqn. (4) individually.

5.1 Likelihood Term

Because we are interested in locating the boundary, we apply a simple observation model which assumes that the intensities are piecewise constant, with a bright intensity representing blood and a darker one representing the myocardium. Intensity variations in the observation, such as those due to differences in blood velocity [9], are modeled through a multiplicative random field (other choices of noise models can be handled in our framework, with the result being a different observation model). Mathematically, the observation model is

$$y_t(z) = \begin{cases} f_t^{R_{in}(X_t)} \cdot n(z) & , z \in R_{in}(X_t) \\ f_t^{R_{out}(X_t)} \cdot n(z) & , z \in R_{out}(X_t), \end{cases} \qquad (5)$$

where $f_t^{R_{in}(X_t)}$ and $f_t^{R_{out}(X_t)}$ are the constant, but unknown, field intensities for the blood pool region inside, R_{in}, and the myocardial region immediately outside (within five pixels), R_{out}, of the LV boundary, respectively, and $n(z)$ is spatially independent, identically distributed lognormal random field with $\log n(z)$ a Gaussian random variable having zero mean and variance σ_n^2. Note that we explicitly indicate the dependence of the regions on X_t. Given the field intensity $f_t^{R(X_t)}$ and the observation model of Eqn. (5), $\log y_t(z)$ is normally distributed with mean $\log f_t^{R(X_t)}$ and variance σ_n^2. Thus,

$$p(y_t|f_t, X_t) \propto \qquad (6)$$

$$exp\left(-\int_{z \in R^{in}(X_t)} \frac{(\log y_t(z) - \log f_t^{R_{in}(X_t)})^2}{2\sigma_n^2} \, dz - \int_{z \in R^{out}(X_t)} \frac{(\log y_t(z) - \log f_t^{R_{out}(X_t)})^2}{2\sigma_n^2} \, dz \right).$$

Since we have a second order model, X_t contains LV boundary information at both t and $t-1$. For the likelihood term, the regions R_{in} and R_{out} are determined by the boundary information from time t.

5.2 Field Priors

In applications where it is possible to extract prior field information, we incorporate a field prior into the problem. The mean log intensity inside is approximately stationary across a cardiac cycle. We compute the mean and variance of the log intensity inside (u and σ_u^2, resp.) and that immediately outside the curve (v and σ_v^2, resp.) from the training data and use this as a field prior to obtain

$$p(f_t|X_t) \propto exp(-\frac{(\log f_t^{R_{in}} - u)^2}{2\sigma_u^2})exp(-\frac{(\log f_t^{R_{out}} - v)^2}{2\sigma_v^2}). \qquad (7)$$

5.3 Prediction Term

Next, we want to provide a model for the prediction term. In Section 3, we described having equally-weighted samples $x_{t|t-1}^{(i,j)}$ to approximately represent our prediction term $p(X_t|Y_{t-1})$. We model this prediction density with a Parzen density estimate using these sample points. Mathematically,

$$p(X_t|Y_{t-1}) = \frac{1}{MN} \sum_{(i,j)} k(X_t; x_{t|t-1}^{(i,j)}, \sigma^2) = \frac{1}{MN} \sum_{(i,j)} \frac{1}{\sqrt{2\pi}\sigma} exp(\frac{-d^2(X_t, x_{t|t-1}^{(i,j)})}{2\sigma^2}),$$
$$(8)$$

where $k(X; \mu, \sigma^2)$ represents a Gaussian kernel with mean μ and variance σ^2 as determined from the bandwidth [10], MN is the number of samples, and $d(X_t, x)$ is a *distance* measure [25] between X_t and sample x.

5.4 Curve Evolution

Having the likelihood, prediction, and prior as above, and defining $F_t^i(X_t) = \log f_t^{R_{in}(X_t)}$ and $F_t^o(X_t) = \log f_t^{R_{out}(X_t)}$, Eqn. (4) becomes

$$E(f_t, X_t) \; = \alpha(\int_{z \in R_{in}(X_t)} \frac{(\log y_t(z) - F_t^i(X_t))^2}{2\sigma_n^2} \; dz + \int_{z \in R_{out}(X_t)} \frac{(\log y_t(z) - F_t^o(X_t))^2}{2\sigma_n^2} dz)$$

$$+\beta(\frac{(F_t^i(X_t) - u)^2}{2\sigma_u^2} + \frac{(F_t^o(X_t) - v)^2}{2\sigma_v^2}) + \gamma \log(\frac{1}{MN} \sum_{(i,j)} \frac{1}{\sqrt{2\pi}\sigma} exp(\frac{-d^2(X_t, x_{t|t-1}^{(i,j)})}{2\sigma^2}))),$$
$$(9)$$

where α, β, γ are weighting parameters specified based on the quality of data. For instance, in low SNR images, α is less heavily-weighted relative to β and γ. Details of the minimization process, which involves coordinate descent and curve evolution, may be found in [25].

6 Experimental Results

We apply the proposed technique on 2-D mid-ventricular slices of data, although it is also applicable to 3-D with a corresponding increase in computational complexity. The dataset we use contains twenty frame time sequences of breath-hold
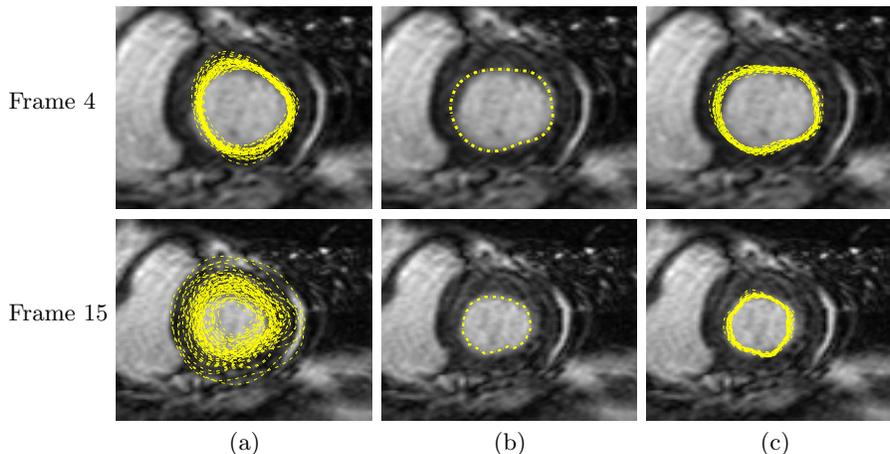
Fig. 4. (a) Curves representing predictions of the LV segmentation (observed MR image in background). (b) Segmentation of an MR image by obtaining the MAP estimate of X_t. (c) Curves representing samples of the posterior density $p(f_t, X_t|Y_t)$.

cardiac MR images, each representing a single cardiac cycle. We do not consider arrhythmia because only patients having sustained and hemodynamically-stable arrhythmia can be practically imaged and analyzed. Such a condition is very rare. Anonymized data sets of were obtained from the Cardiovascular MR-CT Program at Massachusetts General Hospital.

6.1 Training

As discussed in Section 4.1, we represent each manually segmented LV from the training set (a total of 840 frames) by a shape variability vector $\boldsymbol{\alpha}$ and an area A. We obtain the state X_t for each frame t in the cardiac cycle. Then, we learn the dynamics of our system by maximizing $I(X_t; h(X_{t-1}))$, where we approximate h by a linear function, and use gradient ascent on the parameters of h to find the maximum. We obtain a density estimate of $p(X_t|h(X_{t-1}))$ for use in test data.

6.2 Testing

We take sequences of twenty frames (ones not included in the training set), each a single cardiac cycle, as input for testing. For initialization, we assume that a user provides a segmentation of the first two frames in the sequence. The segmentations can be approximate segmentations using some automated method, an expert hand-segmentation, or predicted using a segmentation from a neighboring 2-D slice of the same patient at the same time. From these segmentations, we obtain the initial posterior $p(f_2, X_2|Y_2)$. Using particle filters and curve evolution as described, we recursively estimate the posterior for each frame.

6.3 Results

In Figure 4, we show the segmentation and tracking of the LV based on a test image sequence. Figure 4(a) shows LV boundaries extracted from samples of
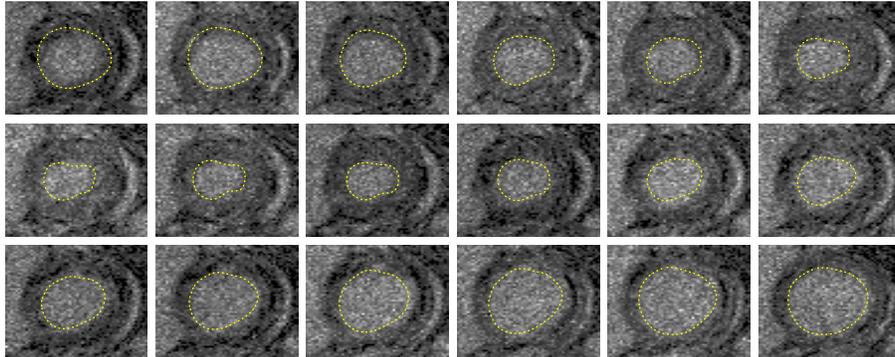
Fig. 5. MAP estimate of segmentations from frame 3 to 20 of a full cardiac cycle.

$p(X_t|h(X_{t-1}))$. Due to space constraints, we show two representative frames from the cardiac cycle. Note that these predictions are obtained based on segmentations from previous frames and on the learned dynamic model, but before incorporating the data shown in Figure 4(a). Figure 4(b) shows the MAP estimate of X_t, which involves incorporating the observed data. This estimate is obtained by minimizing Eqn. (9) and provides what qualitatively appears to be a reasonable segmentation of the LV boundary. Quantitatively, we measure accuracy by computing the symmetric difference between the segmentation and the manually-segmented truth normalized by the area of the truth. Here, the average value across the cardiac cycle of test data is 0.04. Finally, Figure 4(c) shows equally-weighted samples of the posterior density $p(X_t|Y_t)$. This example shows good results, but since the quality of the images are very good, static segmentation methods yield results similar to those shown in Figure 4(b).

We now consider low SNR images where static segmentation may not produce reasonable results. To simulate low SNR conditions, we add independent, lognormal multiplicative noise to MR images to produce a noisy dataset. Using dynamics trained from the MR image training set and initializing again using hand-segmentations on the first two frames, we estimate the LV boundaries. Figure 5 shows segmentations for a full cardiac cycle by taking the MAP estimate of X_t overlaid on the corresponding noisy MR data. The segmentations appear to provide accurate localizations of the LV boundaries despite low quality data.

Figure 1 provides a visual comparison between our approach and one using static segmentation. Only two frames are shown due to space limitations, but they are representative of the results obtained throughout the cardiac cycle. Quantitatively across the entire cardiac cycle, the normalized symmetric difference from our approach is 0.08, while that for static segmentation is 0.17. The static segmentation method is obtained by replacing the $p(X_t|Y_{t-1})$ term in our formulation with a curve length prior and is similar to the region-based segmentations described in the introduction [3, 20, 26]. In both illustrations, incorporating dynamics into the segmentation process using the approach we propose results in better estimates than those using a static segmentation method.

7   Conclusion

We have proposed a principled method to recursively estimate the LV boundary across a cardiac cycle. In the training phase, we learn the dynamics of the LV by obtaining a non-parametric density estimate for the system dynamics. From this, we produce predictions which, used in conjunction with the observations from a new frame, estimate the LV boundary in this frame. The process is repeated through a cardiac cycle. This approach uses information from temporal neighbors to produce better segmentations than using observations at the current frame alone. We have illustrated this method on high and low SNR images. Our formulation produces reasonable estimates using either set of measurements.

A number of extensions to this work may be considered. For instance, our ongoing work considers the generalization to general non-parametric densities for the posterior when a rich enough training set is available. Also, in the learning phase, one might be interested in explicitly incorporating physical constraints to the dynamic system. Adding such constraints may help to eliminate boundary estimates which are known to be physically impossible. In addition, other forms of the function h may be considered. More general non-linear functions may yield a more informative statistic at the cost of greater computational complexity, while a time-varying one may be more informative if sufficient training data is available. In this work, we have posed the problem as a forward recursive filter. Our current work considers improving the estimates by the use of smoothing. Finally, we note that although we track only 2-D slices of the LV in this paper, a natural experimental extension involves applying the technique to 3-D LV data.

References

1. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for on-line non-linear/non-Gaussian Bayesian tracking," *IEEE Transactions on Signal Processing,* 50:2 (2002) 174:188.
2. V. Caselles, F. Catte, T. Coll, and F. Dibos, "A geometric model for active contours in image processing," *Numerische Mathematik,* 66 (1993) 1-31.
3. A. Chakraborty, L. Staib, and J. Duncan, "Deformable boundary finding in medical images by integrating gradient and region information," *IEEE Transactions on Medical Imaging,* 15 (1996) 859-870.
4. Y. Chen, H. D. Tagare, S. Thiruvenkadam, F. Huang, D. Wilson, K. S. Gopinath, and R. W. Briggs, "Using prior shapes in geometric active contours in a variational framework," *Int'l Journal of Computer Vision,* 50:3 (2002) 315-328.
5. A. Doucet, S. J. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statistics and Computing,* 10:3 (2000) 197-208.
6. P. M. Djuric, J. H. Kotecha, J. Zhang, Y. Huang, T. Ghirmi, M. F. Bugallo, and J. Miguez, "Particle filtering," *IEEE Signal Processing Magazine,* 20:5 (2003) 19-38.
7. J. S. Duncan, A. Smeulders, F. Lee, and B. Zaret, "Measurement of end diastolic shape deformity using bending energy," *Computers in Cardiology,* (1988) 277-280.
8. D. Geiger, A. Gupta, L. A. Costa, and J. Vlontzos, "Dynamic programming for detecting, tracking and matching deformable contours," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 17:3 (1995) 294-302.

9. A. Goshtasby and D. A. Turner, "Segmentation of cardiac cine MR images for extraction of right and left ventricular chambers," *IEEE Transactions on Medical Imaging,* 14:1 (1995) 56-64.

10. P. Hall, S. J. Sheather, M. C. Jones, and J. S. Marron, "On optimal data-based bandwidth selection in kernel density estimation," *Biometrika,* 78:2 (1991) 263-269.

11. A. Ihler, "Maximally Informative Subspaces," MS Thesis, MIT (2000).

12. M. Jolly, N. Duta, and G. Funka-Lea, "Segmentation of the left ventricle in cardiac MR images," *Proc of the IEEE Int'l Conf on Computer Vision,* 1 (2001) 501-508.

13. M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int'l Journal of Computer Vision,* (1987) 321-331.

14. J. H. Kotecha and P. M. Djuric, "Gaussian particle filtering," *IEEE Transactions on Signal Processing,* 51:10 (2003) 2592-2601.

15. G. Kuhne, J. Weickert, O. Schuster, and S. Richter, "A tensor-driven active contour model for moving object segmentation," *Proceedings of the 2001 IEEE Int'l Conference on Image Processing,* 2 (2001) 73-76.

16. M. Leventon, E. Grimson, and O. Faugeras, "Statistical shape influence in geodesic active contours," *Proc IEEE Conf on Comp Vision & Patt Rec,* 1 (2000) 316-323.

17. A. McCulloch, J. B. Bassingthwaighte, P. J. Hunter, D. Noble, T. L. Blundell, and T. Pawson, "Computational biology of the heart: From structure to function," *Progress in Biophysics and Molecular Biology,* 69:2-3 (1998) 153-155.

18. J. C. McEachen II and J. S. Duncan, "Shape-based tracking of left ventricular wall motion," *IEEE Transactions on Medical Imaging,* 16:3 (1997) 270-283.

19. N. Paragios and R. Deriche, "Geodesic Active Contours and Level Sets for the Detection and Tracking of Moving Objects," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 22 (2000) 266-280.

20. N. Paragios, "A variational approach for the segmentation of the left ventricle in cardiac image analysis," *Int'l Journal of Computer Vision,* 50:3 (2002) 345-362.

21. J. Senegas, T. Netsch, C. A. Cocosco, G. Lund, and A. Stork, "Segmentation of Medical Images with a Shape and Motion Model: A Bayesian Perspective," *Computer Vision Approaches to Medical Image Analysis (CVAMIA) and Mathematical Methods in Biomedical Image Analysis (MMBIA) Workshop,* (2004) pp. 157-168.

22. M. Sermesant, C. Forest, X. Pennec, H. Delingette, and N. Ayache, "Deformable biomechanical models: Applications to 4D cardiac image analysis," *Medical Image Analysis,* 7:4 (2003) 475-488.

23. J. A. Sethian, "Level Set Methods: Evolving Interfaces in Geometry, Fluid Mechanics, Computer Vision, and Material Science," *Cambridge Univ. Press,* 1996.

24. G. K. von Schutthess, "The effects of motion and flow on magnetic resonance imaging," In *Morphology and Function in MRI,* Ch. 3 (1989) 43-62.

25. W. Sun, M. Çetin, R. Chan, V. Reddy, G. Holmvang, V. Chandar, A. Willsky, "Segmenting and Tracking the Left Ventricle by Learning the Dynamics in Cardiac Images," *MIT LIDS Technical Report 2642,* (Feb 2005).

26. A. Tsai, A. Yezzi, W. Wells, C. Tempany, D. Tucker, A. Fan, W. E. Grimson, and A. Willsky, "A shape-based approach to the segmentation of medical imagery using level sets," *IEEE Transactions on Medical Imaging,* 22:2 (2003) 137-154.

27. J. Weng, A. Singh, and M. Y. Chiu, "Learning-based ventricle detection from cardiac MR & CT images," *IEEE Trans on Medical Imaging,* 16:4 (1997) 378-391.

28. X. S. Zhou, D. Comaniciu, and A. Gupta, "An information fusion framework for robust shape tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 27:1 (2005) 115-129.