

Bidirectional Segmentation of Three-Dimensional Cardiac MR Images Using a Subject-Specific Dynamical Model*

Yun Zhu, Xenophon Papademetris,
Albert J. Sinusas, and James S. Duncan

Departments of Biomedical Engineering and Diagnostic Radiology,
Yale University, USA
yun.zhu@yale.edu

Abstract. Statistical model-based segmentation of the left ventricles has received considerable attention these years. While many statistical models have been shown to improve segmentation results, most of them either belong to (1) static models (SM) that neglect the temporal coherence of a cardiac sequence, or (2) generic dynamical models (GDM) that neglect the individual differences of cardiac motion. In this paper, we propose a subject-specific dynamical model (SSDM) that can simultaneously handle inter-subject variability and temporal cardiac dynamics (intra-subject variability). We also design a dynamic prediction algorithm that can progressively predict the shape of a new cardiac sequence at a given frame based on the shapes observed in earlier frames. Furthermore, to reduce the accumulation of the segmentation errors throughout the entire sequence, we take into account the periodic nature of cardiac motion and perform bidirectional segmentation from a certain frame in a cardiac sequence. “Leave-one-out” validation on 32 sequences show that our algorithm can capture local shape variations and suppress the propagation of segmentation errors.

1 Introduction

The segmentation of the left ventricle (LV) from cardiac magnetic resonance (MR) images is an important prerequisite for quantitative analysis of cardiac function, such as the measurement of ejection fraction and myocardial motion. To determine the myocardial boundaries at weak edges, many researchers have proposed the use of statistical models to regularize the segmentation process (see [1] for an overview). However, most of these models are static models (SM), which supply a prior just for shape, but not the motion of that shape. To take into account the information in the time domain, Perperidis *et al.* extended 3-D models to 4-D case and built two separate models to account for the inter-subject variability and cardiac temporal dynamics (intra-subject variability) respectively [2]. While these two models are able to differentiate cardiac sequences from normal

* This work is supported by the grant 5R01HL082640-02.

subjects from patients with hypertrophic cardiomyopathy, they are inadequate for the segmentation purposes because these two models are unrelated.

At present, there remains limited research concerning the the development of dynamical shape models. For example, Senegas *et al.* proposed a linear dynamical model to approximate cardiac dynamics [3]. Sun *et al.* proposed learning cardiac dynamics using a second-order nonlinear model [4]. While these dynamical models are superior to SM, they are homogeneous in time¹ and therefore insufficient in describing complex shape deformations, such as cardiac dynamics. Also, by supplying a uniform model to all sequences, they ignore the subject variability of motion patterns. Therefore, they are generic dynamical models (GDM).

This paper offers two major contributions. First, we propose a subject-specific dynamical model (SSDM) to simultaneously account for the subject-specific variations of cardiac shapes and inhomogeneous motion. To build this SSDM, we need to differentiate two factors that cause cardiac shape variability. One is inter-subject variability, and the other is temporal dynamics caused by cardiac deformation during a cardiac cycle, as shown in Figure 1. Because conventional Principal Components Analysis (PCA) and Independent Component Analysis (ICA) can only focus on one factor, we extend them to higher orders by utilizing Multilinear PCA (MPCA) [5] and Multilinear ICA (MICA) [6] to decompose the training set and describe the interaction of inter-subject variability and temporal dynamics. We also design a dynamic prediction algorithm that can progressively predict the subject-specific dynamics at the current frame while we gather more information from earlier frames.

Second, in order to further decrease the accumulation of segmentation errors throughout the entire cycle, we propose a forward-backward segmentation

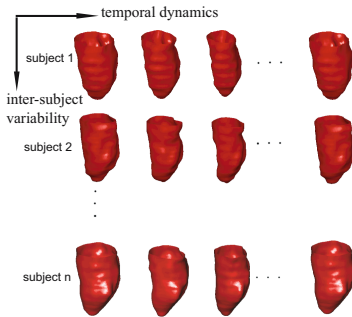


Fig. 1. The interaction of cardiac temporal dynamics and inter-subject variability

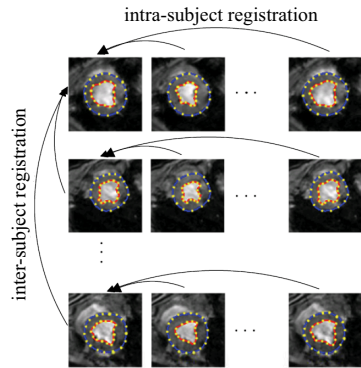


Fig. 2. Overview of landmark propagation

¹ A dynamical model is time-homogeneous if the conditional probability of state t given its previous states only depends on the time difference between those states, i.e. $\mathcal{P}(\mathbf{S}_t | \mathbf{S}_{t-1}, \mathbf{S}_{t-2}, \dots, \mathbf{S}_{t-m}) = \mathcal{P}(\mathbf{S}_{t+n} | \mathbf{S}_{t+n-1}, \mathbf{S}_{t+n-2}, \dots, \mathbf{S}_{t+n-m})$ for all n and m .

strategy that utilizes the periodic nature of cardiac motion and simultaneously performs bidirectional segmentation from a certain frame in a cardiac cycle.

2 Method

2.1 The Construction of SSDM

Data Set. We acquired 32 sequences of canine MR images from a GE signa 1.5 Tesla scanner triggered by electrocardiographic (ECG) signals that were recorded by a MR-compatible ECG sensing device to suppress noises and artifacts induced by gradient magnetic fields. The resulting 3-D image set consists of 16 temporal 3-D frames per cycle, with an in-plane resolution of 1.6mm and a slice thickness of 5mm. The ENDO- and EPI surfaces were manually outlined by our cardiology collaborator using the Bioimage Suite Software [7]. To align all cardiac shapes to a common coordinate system, we first extracted 1095 landmarks on the ENDO surface and 1200 landmarks on the EPI surface in the first frame of the the first sequence. Then we propagated this set of landmarks to all frames in each sequence by mapping those frames to the first frame of the first sequence using inter- and intra-subject registration, as shown in Figure 2. We used both an affine transform to account for the global shape difference and a shape-based non-rigid transform in [8] to accommodate the detailed shape differences. Thus, we obtained 2295 landmarks for each frame. To maximize the effective size of the training set, we adopted the “Leave-one-out” approach, which alternately selects one sequence for validation and uses the rest as the training set.

Shape Decomposition. In this paper, we use MPCA and MICA to decompose cardiac shapes (see [5,6] for an overview of MPCA and MICA). Here, we denote the aligned cardiac shapes as a third-order tensor $\mathcal{S} \in \mathbb{R}^{I \times J \times K}$, where $I = 31$ is the number of subjects, $J = 16$ is the number of frames within a cardiac sequence, and $K = 2295 \times 3 = 6885$ is the dimension of landmark vectors. By applying MPCA to tensor \mathcal{S} , we can decompose it as follows:

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

where $\mathcal{Z} \in \mathbb{R}^{P \times Q \times R}$ is the core tensor which represents the interaction of 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 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 coefficients for each subject i , and matrix $\mathbf{U}^{\text{motion}}$ contains row vectors $\mathbf{u}_j^{\text{motion}} \in \mathbb{R}^Q$ ($1 \leq j \leq J$) of coefficients for frame j .

While it is reasonable to assume that the subject subspace has a Gaussian distribution, the motion subspace does not have a Gaussian distribution. Therefore, we propose utilizing ICA to decompose temporal dynamics to a set of independent

motion patterns in the motion subspace while using PCA in the subject subspace and landmark subspace.

$$\begin{aligned}
 \mathcal{S} &\simeq \mathcal{Z} \times_1 \mathbf{U}^{\text{subject}} \times_2 \mathbf{U}^{\text{motion}} \times_3 \mathbf{U}^{\text{landmark}} \\
 &= \mathcal{Z} \times_1 \mathbf{U}^{\text{subject}} \times_2 \mathbf{U}^{\text{motion}} \mathbf{W}^T \mathbf{W}^{-T} \times_3 \mathbf{U}^{\text{landmark}} \\
 &= \tilde{\mathcal{Z}} \times_1 \mathbf{U}^{\text{subject}} \times_2 \tilde{\mathbf{U}}^{\text{motion}} \times_3 \mathbf{U}^{\text{landmark}}
 \end{aligned} \tag{2}$$

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

To reduce dimensions, we first fixed $Q = J$ and performed MPCSA in the subject subspace and landmark subspace to reduce $I = 31$ to $P = 5$ and $K = 6885$ to $R = 15$, such that the remaining energy was around 98.2% of the original energy. Second, we performed MICA in the motion subspace to find independent variables correspondent to significant shape variations. We selected three most significant modes: short-axial contraction, twisting of the heart, and long-axial contraction. We retained around 96.6% of the original energy after decomposition. Therefore, $Q = 3$ is identified for our model.

Dynamic Prediction Algorithm. Given the segmentation of a cardiac sequence from frame 1 to $t - 1$, we want to predict the segmentation at frame t . The idea is to project the given segmentation to the subject subspace in order to identify the subject vector associated with this sequence, then followed by the prediction of the segmentation at time t with this subject vector.

Let $\mathbf{S}_{1:t-1} = \{\mathbf{S}_1, \mathbf{S}_2, \dots, \mathbf{S}_{t-1}\}$ denote the observed segmentation of a new cardiac sequence, the subject vector associated with this sequence can be represented as

$$\hat{\mathbf{u}}^{\text{subject}} = \mathcal{B} \times_1 \mathbf{S}_{1:t-1} \tag{3}$$

where the ‘‘projection tensor’’ \mathcal{B} is obtained by retensorizing matrix $\mathbf{B}_{(1)} = \mathbf{T}_{(1)}^{-T}$ (the matrix $\mathbf{T}_{(1)}$ is the 1-mode unfolding of tensor $\mathcal{T} = \tilde{\mathcal{Z}} \times_2 \tilde{\mathbf{u}}_{1:t-1}^{\text{motion}} \times_3 \mathbf{U}^{\text{landmark}}$).

With this subject vector, we can predict the segmentation at frame t as

$$\mathbf{S}_t^* = \tilde{\mathcal{Z}} \times_1 \hat{\mathbf{u}}^{\text{subject}} \times_2 \tilde{\mathbf{u}}_t^{\text{motion}} \times_3 \mathbf{U}^{\text{landmark}} \tag{4}$$

2.2 Forward-Backward Segmentation

Assume that we are given consecutive cardiac images $I_{1:t} = \{I_1, I_2, \dots, I_t\}$. Also, we denote \mathbf{S}_t as the segmentation at frame t . Thus, the forward segmentation of frame t can be addressed by maximizing the conditional probability: $\mathcal{P}(\mathbf{S}_t | I_{1:t}) \propto \mathcal{P}(I_t, \mathbf{S}_t | I_{1:t-1}) = \mathcal{P}(I_t | \mathbf{S}_t, I_{1:t-1}) \int \mathcal{P}(\mathbf{S}_t | \mathbf{S}_{1:t-1}) \mathcal{P}(\mathbf{S}_{1:t-1} | I_{1:t-1}) d\mathbf{S}_{1:t-1} \stackrel{(a)}{\approx} \mathcal{P}(I_t | \mathbf{S}_t) \mathcal{P}(\mathbf{S}_t | \hat{\mathbf{S}}_{1:t-1}^+)$.

As explained in [9], we make two assumptions at step (a): (1) $I_{1:t}$ are mutually independent, i.e. $\mathcal{P}(I_t | \mathbf{S}_t, I_{1:t-1}) = \mathcal{P}(I_t | \mathbf{S}_t)$; and (2) the distribution of previous states to be strongly peaked around the maxima of the respective distributions:

$\mathcal{P}(\mathbf{S}_{1:t-1}|I_{1:t-1}) \approx \delta(\mathbf{S}_{1:t-1} - \hat{\mathbf{S}}_{1:t-1}^+)$, where $\hat{\mathbf{S}}_{1:t-1}^+ = \arg \max \mathcal{P}(\mathbf{S}_{t-1}|I_{1:t-1})$ is the result from forward segmentation, and $\delta(\cdot)$ denotes the Dirac delta function.

Similarly, we have $\mathcal{P}(\mathbf{S}_t|I_{t+1:N}) = \int \mathcal{P}(\mathbf{S}_t|\mathbf{S}_{t+1:N}) \mathcal{P}(\mathbf{S}_{t+1:N}|I_{t+1:N}) d\mathbf{S}_{t+1:N} = \mathcal{P}(\mathbf{S}_t|\hat{\mathbf{S}}_{t+1:N}^-)$, where $\hat{\mathbf{S}}_{t+1:N}^- = \arg \max \mathcal{P}(\mathbf{S}_{t+1}|I_{t+1:N})$ is the result from backward segmentation. With all image data $I_{1:N}$ in the entire sequence, we now combine the forward and backward segmentation by maximizing the conditional probability $\mathcal{P}(S_t|I_{1:N})$. As explained in [10], this conditional probability can be simplified as

$$\mathcal{P}(\mathbf{S}_t|I_{1:N}) \propto \mathcal{P}(\mathbf{S}_t|I_{1:t}) \mathcal{P}(\mathbf{S}_t|I_{t+1:N}) \propto \underbrace{\mathcal{P}(I_t|\mathbf{S}_t)}_{\text{data adherence}} \underbrace{\mathcal{P}(\mathbf{S}_t|\hat{\mathbf{S}}_{1:t-1}^+)}_{\text{forward dynamics}} \underbrace{\mathcal{P}(\mathbf{S}_t|\hat{\mathbf{S}}_{t+1:N}^-)}_{\text{backward dynamics}} \quad (5)$$

2.3 Data Adherence

To decide the intensity distribution of cardiac MR images, we partition the entire the image by the ENDO- and EPI surfaces into three regions - LV blood pool, LV myocardium, and background. The LV blood pool and myocardium have homogenous intensities, and can therefore be modeled with a unimodal distribution. The most common distribution for MR images is Gaussian (Normal) distribution $\mathcal{P}(I; \mu_l, \sigma_l) = \frac{1}{\sqrt{2\pi}\sigma_l} \exp\left\{-\frac{(I-\mu_l)^2}{2\sigma_l^2}\right\}$, where μ_l is the mean of Gaussian distribution, and σ_l is its deviation. For $l = 1$, it models the intensity distribution in LV blood pool. For $l = 2$, it models the intensity distribution in LV myocardium.

The background, however, is inhomogeneous because it contains more than one tissues (RV blood pool, RV myocardium, and lung air). Therefore, we use a mixture model $\mathcal{P}_3(I; \mu_3, \sigma_3) = \sum_{k=1}^M \alpha_k \mathcal{P}_{3,k}(I; \mu_{3,k}, \sigma_{3,k})$, 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 standard deviation of its component Gaussian distribution. Here, we use $M = 3$ for the background of the cardiac MR images.

Let $\Omega_{t,1}$, $\Omega_{t,2}$, and $\Omega_{t,3}$ denote the LV blood pool, LV myocardium, and background in each frame t respectively. Thus, the data adherence term can be defined as a log-likelihood function $\log \mathcal{P}(I_t|\mathbf{S}_t) = \sum_{l=1}^3 \int_{\Omega_{t,l}} \log \mathcal{P}(I_t; \mu_l, \sigma_l) dx$.

2.4 Forward and Backward Dynamics

As shown in Section 2.1, we can predict the ENDO- and EPI surfaces at frame t from previous frames using Equations 3 and 4. Thus, the forward and backward dynamics can be defined as $\mathcal{P}(\mathbf{S}_t|\hat{\mathbf{S}}_{1:t-1}^+) \propto \exp\left\{-\frac{\alpha}{2} \int \|\mathbf{S}_t - \mathbf{S}_t^{+*}\|^2\right\}$ and $\mathcal{P}(\mathbf{S}_t|\hat{\mathbf{S}}_{t+1:N}^-) \propto \exp\left\{-\frac{\alpha}{2} \int \|\mathbf{S}_t - \mathbf{S}_t^{-*}\|^2\right\}$, where α is a weighting parameter,

and S_t^{+*} and S_t^{-*} are the predicted contours in forward and backward directions from the dynamic prediction algorithm described in Section 2.1. We found in the experiments that $0.5 \leq \alpha \leq 2$ is applicable for most of the data sets. In the experiments, the first frame of the sequence was manually segmented, and the segmentation of the current frame was initialized by the segmentation of the previous frame.

3 Experiments

Let A and B be two surfaces from automatic and manual segmentation respectively. Suppose A and B are represented by point sets, i.e. $A = \{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_N\}$ and $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_M\}$. The difference between these two surfaces can be described by mean absolute distance as $MAD(A, B) = \frac{1}{2} \left\{ \frac{1}{N} \sum_{i=1}^N d(\mathbf{a}_i, B) + \frac{1}{M} \sum_{j=1}^M d(\mathbf{b}_j, A) \right\}$, and Hausdorff distance as $HD(A, B) = \frac{1}{2} \left\{ \max_i d(\mathbf{a}_i, B) + \max_j d(\mathbf{b}_j, A) \right\}$, 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. We also computed the percentage of correctly segmented voxels as $PTP = \frac{\text{Volume}(\Omega_A \cap \Omega_B)}{\text{Volume}(\Omega_A)}$, where Ω_A and Ω_B are the two volumes enclosed by the surfaces A and B respectively.

3.1 SSDM Versus SM and GDM

Figure 3 provides a visual comparison of the results obtained with SM and SSDM, while Figure 4 provides a visual comparison of the results obtained with GDM and SSDM. It can be seen that SM is biased in the global sense because it is trapped easily in the local minima. Figure 5 (black and red lines) provides a quantitative evaluation of segmentation results over the entire cardiac cycle. It shows that for ENDO boundaries, when SSDM is used, MAD improves by 0.38mm and HD by 1.97mm. For EPI boundaries, when SSDM is used, MAD improves by 0.11mm and HD by 1.63mm. This implies that while GDM can capture global shape deformations, it misses out on the local shape variations.

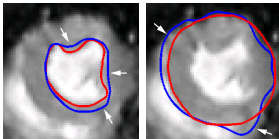


Fig. 3. Comparing the segmentation results using SM (blue) and SSDM (red)

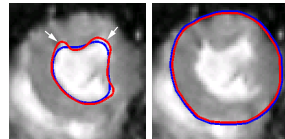


Fig. 4. Comparing segmentation results using GDM (blue) and SSDM (red)

3.2 Forward Versus Forward-Backward Segmentation

Figure 6 shows the qualitative results over an entire cycle with and without the backward strategy, while Figure 5 (red and green lines) illustrates the quantitative results. At the end of a cycle, when the backward strategy is added, MAD and HD improve by 0.90mm and 1.52mm, respectively, for the ENDO boundaries. Also, MAD and HD improve by 1.01mm and 2.04mm, respectively, for the EPI boundaries. This shows that forward-backward segmentation produces more consistent and robust delineation of LV boundaries over the full cardiac

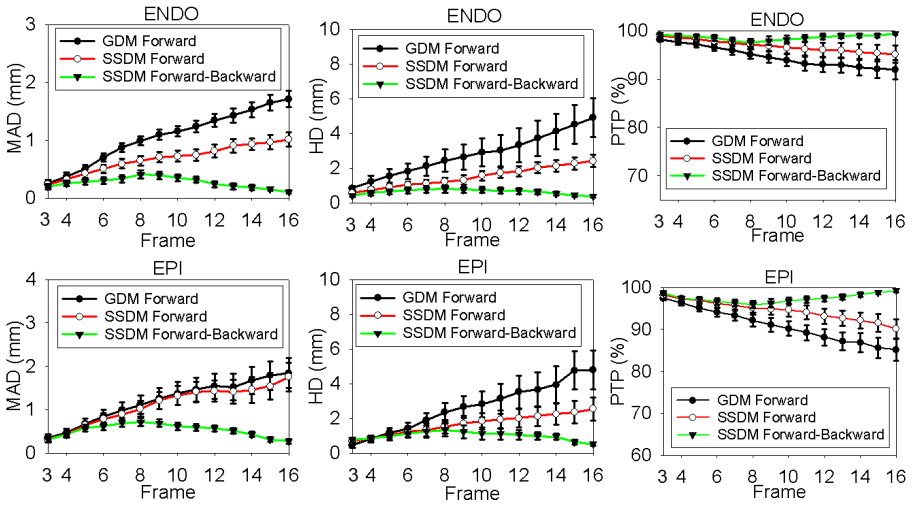


Fig. 5. Comparison among three algorithms over 32 sequences

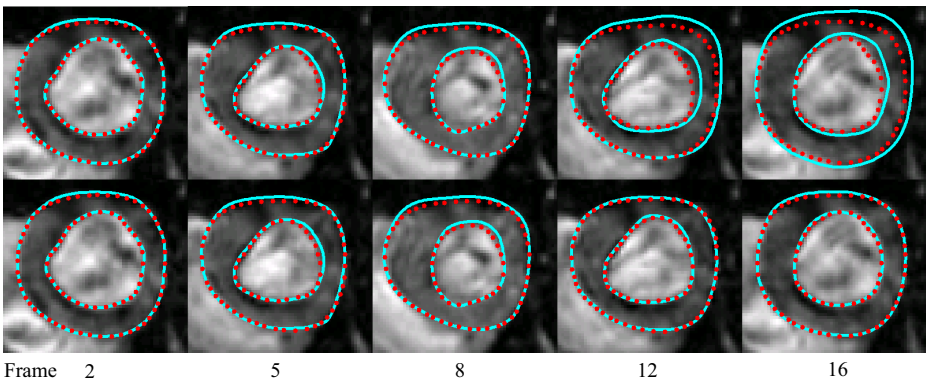


Fig. 6. Qualitative validation over an entire cardiac cycle (shown at frame 2, 5, 8, 12, and 16): automatically segmented contours (cyan solid lines), manual contours (red dotted lines). Upper: forward segmentation; Lower: forward-backward segmentation.

cycle despite the slight increase of error in the middle frames of a cycle. Also, the EPI boundaries are more difficult to segment. The segmentation errors of EPI boundaries in one frame quickly accumulate when the backward strategy is not added.

4 Conclusion

In this paper, we have proposed a subject-specific dynamical model (SSDM) to decompose training samples into subject subspace, motion subspace, and landmark subspace using MPCA and MICA. Given a new cardiac sequence, it can first identify the subject vector associated with this new sequence, and then use this subject vector to predict the LV contours in the future frames. To reduce the propagation of segmentation errors throughout the entire cardiac sequence, we proposed to perform forward-backward segmentation simultaneously from one frame in a cardiac sequence. We integrated the SSDM and the forward-backward strategy into the LV segmentation process on the basis of recursive Bayesian framework. Future work includes the extension of this approach to other modalities, such as CT and ultrasound images.

References

1. Frangi, A.F., Niessen, W.J., Viergever, M.A.: Three-dimensional modeling for functional analysis of cardiac images: A review. *IEEE TMI* 20(1), 2–25 (2001)
2. Perperidis, D., Mohiaddin, R., Rueckert, D.: Construction of a 4-d statistical atlas of the cardiac anatomy and its use in classification. In: Duncan, J.S., Gerig, G. (eds.) *MICCAI 2005*. LNCS, vol. 3749, pp. 402–410. Springer, Heidelberg (2005)
3. Senegas, J., Netsch, T., Cocosco, C.A., Lund, G., Stork, A.: Segmentation of medical images with a shape and motion model: A bayesian perspective. In: *MMBIA*, pp. 157–168 (2004)
4. Sun, W., Cetin, M., Chan, R., Reddy, V., Holmvang, G., Chandar, V., Willsky, A.: Segmenting and tracking of the left ventricle by learning the dynamics in cardiac images. In: Christensen, G.E., Sonka, M. (eds.) *IPMI 2005*. LNCS, vol. 3565, pp. 553–565. Springer, Heidelberg (2005)
5. Lathauwer, L.D., Moor, B.D., Vandewalle, J.: A multilinear singular value decomposition. *SIAM J. Matrix Anal. Appl.* 21(4), 1253–1278 (2000)
6. Alex, M., Vasilescu, O., Terzopoulos, D.: Multilinear independent component analysis. *CVPR* 1, 547–553 (2005)
7. Papademetris, X., Jackowski, M., Rajeevan, N., Constable, R.T., Staib, L.H.: Bioimage suite: An integrated medical image analysis suite, section of bioimaging sciences, department of diagnostic radiology, yale school of medicine, <http://www.bioimagesuite.org>
8. Papademetris, X., Sinusas, A., Dione, D.P., Constable, R.T., Duncan, J.S.: Estimation of 3-d left ventricular deformation from medical images using biomechanical models. *IEEE TMI* 21(7), 786–800
9. Cremers, D.: Dynamical statistical shape priors for level set-based tracking. *IEEE TPAMI* 28(8), 1262–1273 (2006)
10. Forsyth, D.A., Ponce, J.: *Computer Vision: A Modern Approach*. Prentice Hall, Englewood Cliffs (2002)