

# EPICARDIAL SEGMENTATION IN DYNAMIC CARDIAC MR SEQUENCES USING PRIORS ON SHAPE, INTENSITY, AND DYNAMICS, IN A LEVEL SET FRAMEWORK

Atiyeh Ghoreyshi      René Vidal

Center for Imaging Science, Department of Biomedical Engineering, Johns Hopkins University

## ABSTRACT

We propose a method for epicardial segmentation in dynamic MR sequences using priors on shape, intensity, and dynamics. The prior models are built from a training set of manually segmented sequences. The dynamics of the heart and background intensities are modeled with linear autoregressive models whose parameters are learnt from training data and then used as priors for segmentation purposes. The segmented training shapes are registered with respect to a common reference. Each registered shape is represented with a signed distance function, and a statistical shape model is learnt using PCA on these functions. Segmentation is achieved by minimizing a spatial-temporal generalization of the Mumford-Shah energy functional in a level set framework.

## 1. INTRODUCTION

Over the past few years, significant attention has been directed to the automatic segmentation of cardiac images using prior models on intensity and shape.

In the case of intensity-based image segmentation, one of the most common methods seeks to find a piecewise smooth approximation of the image by minimizing the Mumford-Shah energy functional [1]. In the case of a piecewise constant approximation, an image  $I(x, y)$  is segmented into two regions by finding a smooth contour,  $C$ , and the mean intensities inside and outside this contour. Explicit methods [2] represent  $C$  with a finite number of control points which are evolved to match the boundaries in the image. Implicit methods [3, 4, 5], known as level set methods, represent  $C$  as the zero level set of an implicit function  $\varphi$ , i.e.  $C = \{(x, y) : \varphi(x, y) = 0\}$ , and evolve this function to match the boundaries in the image.

Level set methods are nowadays widely used in cardiac image segmentation, because they result in coherent regions with smooth boundaries. While earlier work used edge-based features for segmentation [6, 7], more recent work combines edge and region-based features with prior knowledge [8]. However, epicardial segmentation in MR images can be very challenging due to lack of well defined edges, lack of contrast, existence of different textures, inhomogeneities, noise, etc.

Recent work has tried to overcome these problems by using priors on the shape and appearance of the region of interest [9, 10, 11, 8, 12, 13]. Although these methods significantly

improve the results, they are not accurate and reliable enough yet. In particular, shape variability among different slices, and also the displacement of organs with respect to each other and with respect to the scanner at different scanning times, pose significant challenges to the use of shape priors. Also, intensities inside and outside of the heart often have significant overlap in cardiac MR images, which poses challenges to the use of intensity priors. This has motivated the use of dynamics as an additional cue for segmentation [14].

In this paper, we propose a method to segment the epicardium in dynamic MR sequences, using priors on shape, intensity, and also dynamics. We model different regions with autoregressive (AR) models, and extract probability distributions for the AR parameters in different regions. We use this information as well as information about the shape and intensities in different regions to perform the segmentation task.

## 2. PRIORS ON INTENSITY AND DYNAMICS

In order to build prior models for the intensity and dynamics of the heart and the background, we describe the intensity of a pixel as the output of an AR model. Considering  $F$  consecutive frames in a video sequence, we assume that the intensity of every pixel in each frame is related to the intensities of the same pixel in previous frames by an AR model of order  $p$

$$I_f(\mathbf{x}) = \bar{I}_F(\mathbf{x}) + \sum_{i=1}^p a_i(\mathbf{x})(I_{f-i}(\mathbf{x}) - \bar{I}_F(\mathbf{x})), \quad (1)$$

where  $I_f(\mathbf{x})$  is the intensity of pixel  $\mathbf{x}$  in the  $f^{\text{th}}$  frame, and  $\bar{I}_F(\mathbf{x})$  is the mean intensity of pixel  $\mathbf{x}$  in the  $F$  consecutive frames. Rewriting equation (1) for all the frames in a temporal window of size  $F$ , and substituting  $I_i(\mathbf{x}) - \bar{I}_F(\mathbf{x})$  with  $I'_i(\mathbf{x})$ ,  $i = 1 \dots F$ , yields the matrix equation

$$\mathbf{a}(\mathbf{x})\Phi_F(\mathbf{x}) = U_F(\mathbf{x}), \quad (2)$$

where

$$\Phi_F(\mathbf{x}) = \begin{bmatrix} I'_1(\mathbf{x}) & I'_2(\mathbf{x}) & \cdots & I'_{F-p}(\mathbf{x}) \\ I'_2(\mathbf{x}) & I'_3(\mathbf{x}) & \cdots & I'_{F-p+1}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ I'_p(\mathbf{x}) & I'_{p+1}(\mathbf{x}) & \cdots & I'_{F-1}(\mathbf{x}) \end{bmatrix}, \quad (3)$$

$$\mathbf{a}(\mathbf{x}) = [ a_p(\mathbf{x}) \quad a_{p-1}(\mathbf{x}) \quad \cdots \quad a_1(\mathbf{x}) ], \quad \text{and} \quad (4)$$

$$U_F(\mathbf{x}) = [ I'_{p+1}(\mathbf{x}) \quad I'_{p+2}(\mathbf{x}) \quad \cdots \quad I'_F(\mathbf{x}) ]. \quad (5)$$

The least squares solution to equation (2) is given by

$$\mathbf{a}(\mathbf{x}) = U_F(\mathbf{x})\Phi_F(\mathbf{x})^\top (\Phi_F(\mathbf{x})\Phi_F(\mathbf{x})^\top)^{-1}. \quad (6)$$

However, singularities could occur if  $\Phi_F(\mathbf{x})$  is not full rank. This could happen, e.g., when a pixel has almost constant intensity in the  $F$  consecutive frames, or when  $F$  is too small to allow the intensities of a pixel to have enough variation. Since we are interested in choosing  $F$  as small as possible to avoid significant displacements of the boundary of the heart within the  $F$  frames, we add a regularization term to the least squares problem. Instead of minimizing  $\|\mathbf{a}(\mathbf{x})\Phi_F(\mathbf{x}) - U_F(\mathbf{x})\|^2$ , we minimize  $\|\mathbf{a}(\mathbf{x})\Phi_F(\mathbf{x}) - U_F(\mathbf{x})\|^2 + \eta\|\mathbf{a}(\mathbf{x})\|^2$ . The new solution for the AR parameters is given by

$$\mathbf{a}(\mathbf{x}) = U_F(\mathbf{x})\Phi_F(\mathbf{x})^\top (\Phi_F(\mathbf{x})\Phi_F(\mathbf{x})^\top + \eta\mathbf{I}_{p \times p})^{-1}. \quad (7)$$

Given a training set of manually segmented video sequences of a beating heart, we can calculate the AR parameters for every pixel in the image using a sliding temporal window of size  $F$ . With all the AR parameters in hand, we can now build two sets of histograms of mean intensity and AR parameters for the heart and the background, respectively. We use these histograms as the prior probability distribution functions (pdf's) for the intensity and dynamics of the regions.

### 3. PRIORS ON SHAPE

In order to capture the shape variability among all the shapes in our training set, we first align all the shapes in the training set using similarity transformations. For this step, we represent the shapes with binary images which are white on the heart and black on the background (true masks). We then extract the remaining variability using PCA on the space of shapes, and capture the dominant modes of shape variation by using the first few principal components.

For the first step, we register image  $\tilde{I}(\mathbf{x})$  to image  $I(\mathbf{x})$ , by minimizing the least squares energy functional

$$E_{align} = \frac{1}{2} \int (\tilde{I}(\tilde{\mathbf{x}}) - I(\mathbf{x}))^2 d\mathbf{x}. \quad (8)$$

In equation (8),  $\mathbf{x} = [x \quad y]^\top$  are the coordinates of  $I$ , and  $\tilde{\mathbf{x}} = [\tilde{x} \quad \tilde{y}]^\top$  are the transformed coordinates of  $\tilde{I}$  via a similarity transformation  $\tilde{\mathbf{x}} = hR\mathbf{x} + t$ , where  $h \in \mathbb{R}^+$ ,

$$R = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \in SO(2), \text{ and } t = \begin{bmatrix} t_x \\ t_y \end{bmatrix} \in \mathbb{R}^2$$

are the scale, rotation matrix, and translation vector, respectively. We minimize  $E_{align}$  using gradient descent with respect to the pose parameters  $\mathbf{p} = [t_x \quad t_y \quad h \quad \theta]^\top$ . The gradient of  $E_{align}$  with respect to  $\mathbf{p}$  is given by

$$\nabla_{\mathbf{p}} E_{align} = \int (\tilde{I}(\tilde{\mathbf{x}}) - I(\mathbf{x})) \nabla_{\mathbf{p}} \tilde{I}(\tilde{\mathbf{x}}), \quad (9)$$

where  $\nabla_{\mathbf{p}} \tilde{I}(\tilde{\mathbf{x}}) = [\mathbf{I}_{2 \times 2} \quad R\mathbf{x} \quad h \frac{\partial R}{\partial \theta} \mathbf{x}]^\top \nabla_{\tilde{\mathbf{x}}} \tilde{I}(\tilde{\mathbf{x}})$ . In our experiments, this simple least squares method is sufficient for the registration of the binary true masks.

After aligning the true masks, we build a signed distance function (sdf) for every registered true mask. An sdf is a function defined on the image plane, which gives the signed distance of any point on the image to a specific closed contour such that the points inside the contour have negative distance, and the points outside the contour have positive distance.

The so-computed sdf's are then used for building a statistical shape prior model using principal component analysis (PCA). That is, we write the sdf  $\phi_j(\mathbf{x})$  corresponding to the  $j^{\text{th}}$  shape,  $j = 1, \dots, N$ , as

$$\phi_j(\mathbf{x}) = \phi_0(\mathbf{x}) + \sum_{i=1}^n \alpha_{ij} \psi_i(\mathbf{x}), \quad (10)$$

where  $\phi_0(\mathbf{x})$  is the mean of all the sdf's,  $\psi_1, \dots, \psi_n$  are the  $n$  principal modes of shape variation, and  $\alpha_{ij}$  are the shape coefficients. In theory, sdf's are functions on a continuous domain. However, as the number of pixels is finite, we can approximate the sdf's by their values on a finite grid. More specifically, let  $\varphi_j \in \mathbb{R}^P$  be a vector containing the values of  $\phi_j(\mathbf{x})$  for all  $P$  pixels  $\mathbf{x}$ . Then, we can compute the functions  $\psi_i$ 's by applying PCA to the matrix

$$S = [ \varphi_1 \quad \varphi_2 \quad \dots \quad \varphi_N ] \in \mathbb{R}^{P \times N}. \quad (11)$$

We choose the number of principal components  $n$  using the model selection formula [15]

$$n = \underset{j=1, \dots, N}{\operatorname{argmin}} \frac{\sigma_{j+1}^2}{\sum_{i=1}^j \sigma_i^2} + \kappa \cdot j, \quad (12)$$

where  $\sigma_i$  is the  $i^{\text{th}}$  singular value of  $S$  and  $\kappa$  is a parameter.

Note that the space of sdf's is not a linear manifold. This can easily be seen by noting that the summation of two sdf's need not be an sdf. Therefore, by doing PCA on this manifold, we are using a linearization of the manifold around  $\phi_0(\mathbf{x})$ . However, since we previously registered the shapes, deviations of the shapes from  $\phi_0(\mathbf{x})$  are not significant, and the linearized manifold is a good approximation.

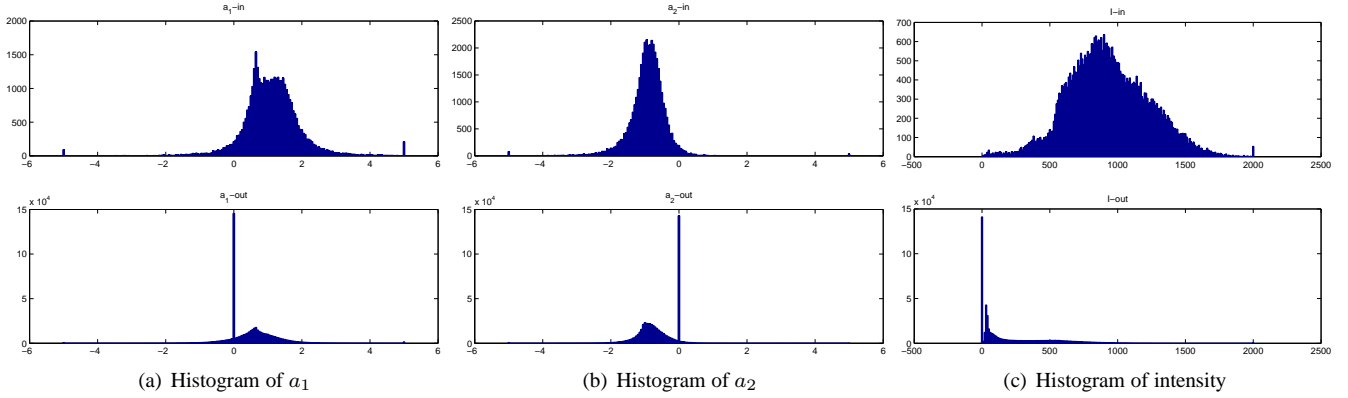
### 4. SEGMENTATION USING DYNAMICAL PRIORS

Given  $F$  new frames, we can use the prior shape model to parameterize the contour as the zero level set of the function

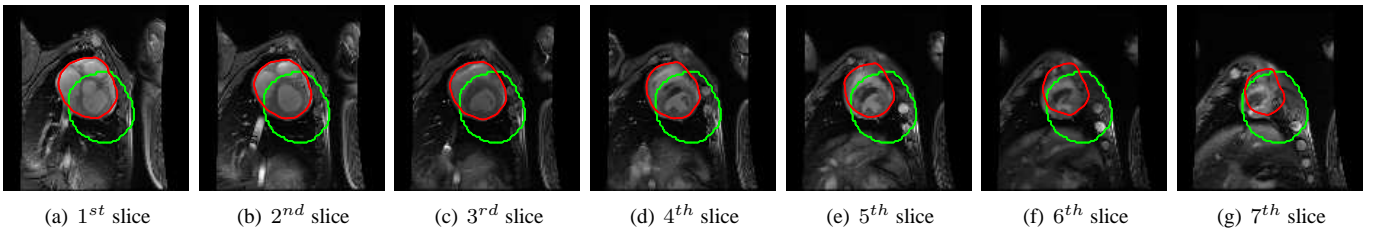
$$\phi(\mathbf{x}) = \phi_0(hR\mathbf{x} + t) + \sum_{i=1}^n \alpha_i \psi_i(hR\mathbf{x} + t), \quad (13)$$

where  $h$ ,  $\theta$ , and  $t$ , are the pose parameters, and  $\alpha_1 \cdots \alpha_n$  are the shape coefficients. The optimal contour for the  $F$  frames can be found by maximizing the log-likelihood function

$$f_{ML} = \log \operatorname{prob}(\text{intensity and AR parameters of the } F \text{ frames} \\ | h, \theta, t, \alpha_1 \cdots \alpha_n). \quad (14)$$



**Fig. 1.** Histograms of intensity and AR parameters for the regions corresponding to inside and outside of the heart



**Fig. 2.** Segmentation results. Green contour: initialization using  $\phi_0$ . Red contour: final segmentation.

Notice that maximizing  $f_{ML}$  is equivalent to minimizing

$$\begin{aligned}
 E_{ML} = -f_{ML} = & \\
 & - \int \log P_{in}(u_0(\mathbf{x}), a_1(\mathbf{x}), \dots, a_p(\mathbf{x})) (1 - H(\phi(\mathbf{x}))) d\mathbf{x} \\
 & - \int \log P_{out}(u_0(\mathbf{x}), a_1(\mathbf{x}), \dots, a_p(\mathbf{x})) H(\phi(\mathbf{x})) d\mathbf{x}.
 \end{aligned} \quad (15)$$

The functions  $\phi$  and  $H(\phi)$  are the sdf and the heaviside function respectively, corresponding to the current contour.  $H(\phi)$  is a binary function which is 1 where  $\phi \geq 0$ , and 0 where  $\phi < 0$ .  $P_{in}$  and  $P_{out}$  are the joint pdf's of the intensity and AR parameters for the heart and the background, respectively. In practice, we do not have access to these probability distributions. We only have histograms for the intensity and AR parameters computed from the training set. We use the histograms of these parameters as their individual pdf's. We then use the product of the individual pdf's as their joint pdf, assuming that they are independent. We also assume that the pixels are independent of each other, thus  $P_{in}$  and  $P_{out}$  are defined independently at each pixel.

We find the optimal pose parameters,  $\mathbf{p}$ , and the optimal shape coefficients  $\boldsymbol{\alpha} = [\alpha_1, \dots, \alpha_n]^\top$  by minimizing  $E_{ML}$  using gradient descent. The optimal segmentation is then given by substituting the optimal parameters in (13).  $\nabla_{\mathbf{p}} E_{ML}$ , the gradient of  $E_{ML}$  with respect to  $\mathbf{p}$ , is given by

$$\int \log \frac{P_{in}}{P_{out}}(u_0(\mathbf{x}), a_1(\mathbf{x}), \dots, a_p(\mathbf{x})) \delta(\phi(\mathbf{x})) \nabla_{\mathbf{p}} \phi(\mathbf{x}), \quad (16)$$

where  $\nabla_{\mathbf{p}} \phi(\mathbf{x}) = [\mathbf{I}_{2 \times 2} \quad R\mathbf{x} \quad h \frac{\partial R}{\partial \theta} \mathbf{x}]^\top \nabla_{\mathbf{x}} \phi(\mathbf{x})$ .  $\nabla_{\boldsymbol{\alpha}} E_{ML}$ , the gradient of  $E_{ML}$  with respect to  $\boldsymbol{\alpha}$ , is given by

$$\int \log \frac{P_{in}}{P_{out}}(u_0(\mathbf{x}), a_1(\mathbf{x}), \dots, a_p(\mathbf{x})) \delta(\phi(\mathbf{x})) \nabla_{\boldsymbol{\alpha}} \phi(\mathbf{x}), \quad (17)$$

where  $\nabla_{\boldsymbol{\alpha}} \phi(\mathbf{x}) = \boldsymbol{\psi}(hR\mathbf{x} + t)$ , and

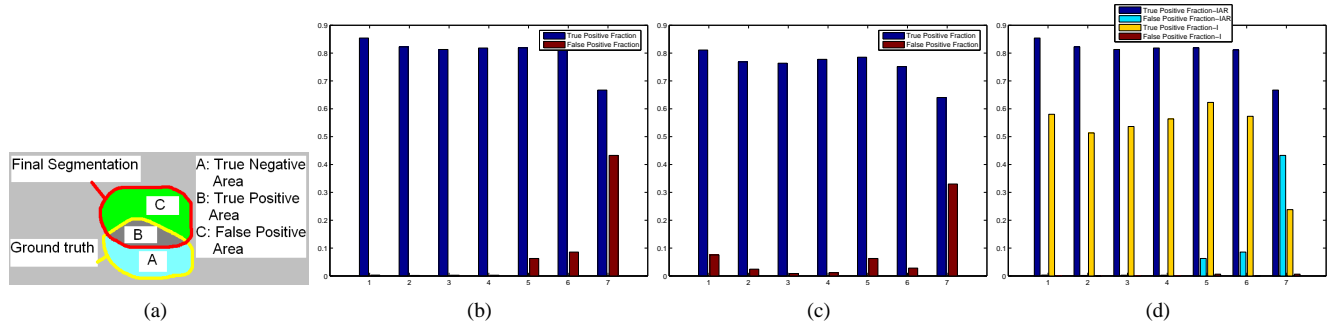
$$\boldsymbol{\psi} = [\psi_1(hR\mathbf{x} + t) \psi_2(hR\mathbf{x} + t) \dots \psi_n(hR\mathbf{x} + t)]^\top. \quad (18)$$

## 5. EXPERIMENTAL RESULTS

In our experiments, we use a dataset containing 7 video sequences corresponding to 7 axial slices of the heart. Each video sequence has 30 frames of size  $128 \times 128$  spanning over one complete beating cycle. For every slice, we use frames 1-5, 11-15 and 21-25 for training, and the rest for testing.

Figure 1 shows the histograms of intensity and AR parameters for the heart and the background. Figure 2 shows our segmentation results for various slices of the dataset. Notice that the method works almost perfectly on the basal slices. On the other hand, the algorithm loses accuracy on the apical slices, because on these slices, the variations of heart intensities are less significant than in the basal slices, and therefore comparable to intensity variations due to noise and artifacts in some neighboring regions. However, the apical slices are not particularly important during catheterization procedures, which are the main application of this work.

In order to give the reader a better way to evaluate our results, we provide numerical results for some intuitive and critical quantities that are illustrated in Figure 3(a). We define the



**Fig. 3.** (a) Demonstration of TP and FP. (b) Statistics of the results when initialized with  $\phi_0$ . (c) Statistics of the results with various initializations. (d) Comparing our method (IAR) versus the method without AR priors (I) when initialized with  $\phi_0$ .

True Positive Fraction for every image as  $\frac{B}{A+B}$ , and the False Positive Fraction as  $\frac{C}{A+B}$ . We perform two different types of experiments. First, for every slice in our testing set, we run our algorithm with the mean shape  $\phi_0$  as initialization. Figure 3(b) shows the mean statistics for the results on the 7 slices. Second, we segment the images initialized with a transformed version of the mean shape. We use similarity transformations using all the combinations given by  $t_x \in \{-10, 0, 10\}$  pixels,  $t_y \in \{-10, 0, 10\}$  pixels,  $\theta \in \{-90, 0, 90\}$  degrees, and  $h \in \{.8, 1, 1.2\}$ . Figure 3(c) shows the mean statistics of the results of this experiment for the 7 slices. Note that our method is highly robust to initialization.

We also tested a simplified version of our method in which we do not include AR parameter priors. This method is similar to the method in [13], except that we do not put any probability distribution on the shape priors. As can be seen in Figure 3(d), using AR parameters significantly improves the true positive fraction of the results, although it increases the false positive fraction in apical slices. However, note that we are more concerned about the true positive fraction. We do not want to miss any part of the heart at this stage, because it will result in further inaccuracy of the automatic segmentation of subdivisions of the heart that might follow our algorithm.

## 6. CONCLUSIONS AND FUTURE WORK

We have presented an algorithm for epicardial segmentation in dynamic cardiac MR sequences that uses priors on shape, intensity and dynamics. Our results showed that priors on dynamics indeed improve segmentation results. We also showed results on very challenging apical slices that are usually avoided. In the future, we plan to extend our algorithm to 3D volumes. This will help us get better results on apical slices, because it will enable us to use the information from adjacent slices.

## 7. ACKNOWLEDGEMENTS

Work supported by grant NIH-NHLBI 5-R01-HL-64795-01. We thank A. Genc, F. Miller, M. Zviman and H. Halperin.

## 8. REFERENCES

- [1] D. Mumford and J. Shah, "Boundary detection by minimizing functionals," in *IEEE CVPR*, 1985, pp. 22–26.
- [2] A. Witkin and D. Terzopoulos, "Snakes: Active contour models," *Int. Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1987.
- [3] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," in *IEEE CVPR*, 1995, pp. 694–699.
- [4] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.
- [5] A. Tsai, A. Yezzi Jr., and A.S. Willsky, "Curve evolution implementation of the Mumford-Shah functional for image segmentation, denoising, interpolation, and magnification," *IEEE Transactions on Image Processing*, vol. 10, no. 8, pp. 1169–1186, 2001.
- [6] K. Siddiqui, Y. B. Lauziere, A. Tannenbaum, and S. W. Zucker, "Area and length minimizing flows for shape segmentation," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 433–443, 1998.
- [7] M. Santarelli, V. Positano, C. Michelassi, M. Lombardi, L. Landini, "Automated cardiac MR image segmentation: theory and measurement evaluation," *Medical Eng. & Physics*, vol. 25, pp. 149–159, 2003.
- [8] C. Pluempitwiriwajew, J. Moura, Y. Wu, and C. Ho, "Stacs: New active contour scheme for cardiac MR image segmentation," *IEEE Transactions on Medical Imaging*, vol. 24, no. 5, pp. 593–603, 2005.
- [9] M. Kaus, J. von Berg, W. Niessen, and V. Pekar, "Automated segmentation of the left ventricle in cardiac MRI," *MICCAI*, 2003, pp. 432–439.
- [10] X. Huang, Z. Li, and D. Metaxas, "Learning coupled prior shape and appearance models for segmentation," in *MICCAI*, 2004, pp. 60–69.
- [11] R. Beichel, H. Bischof, F. Leberl, and M. Sonka, "Robust active appearance models and their application to medical image analysis," *IEEE Trans. on Medical Imaging*, vol. 24, no. 4, pp. 1151–1169, 2005.
- [12] Y. Shang, X. Yang, M. Zhu, B. Jin, and M. Liu, "Region and shape prior based geodesic active contour and application in cardiac," in *MICCAI*, 2005, vol. 3483, pp. 1102–1110.
- [13] M. Rousson and D. Cremers, "Efficient kernel density estimation of shape and intensity priors for level set segmentation," in *MICCAI*, 2005, vol. 3750, pp. 757–764.
- [14] A. Ravichandran, R. Vidal, and H. Halperin, "Segmenting a beating heart using polysegment and spatial GPCA," in *IEEE International Symposium on Biomedical Imaging*, 2005, pp. 634–637.
- [15] K. Kanatani, "Evaluation and selection of models for motion segmentation," in *Asian Conference on Computer Vision*, 2002, pp. 7–12.