A NEW ROBUST ACTIVE SHAPE MODEL FORMULATION FOR CARDIAC MRI SEGMENTATION

Carlos Santiago, Jacinto C. Nascimento, Jorge S. Marques

Institute for Systems and Robotics, LARSyS,
Instituto Superior Técnico, Univ. Lisboa, Portugal

ABSTRACT

The 3D segmentation of the left ventricle (LV) in cardiac MRI is a challenging problem, due to the presence of other anatomical structures and artifacts (outliers) around the LV. In this paper, a new formulation of a *Robust Active Shape Model* (RASM) is presented that is able to deal with those outliers. Instead of using the traditional one-to-one mapping of edge points and model points to compute the shape model parameters, the proposed approach uses a one-to-many mapping strategy and groups these edge points into edge segments (strokes). Then, a probabilistic framework provides a robust estimation of the model parameters, in which the influence in the segmentation of the unreliable outliers is reduced. The proposed method was tested on a public dataset comprising 660 volumes. The results indicate that this methodology provides accurate segmentations that are competitive with other state-of-the art methods.

Index Terms— 3D Segmentation, Cardiac MRI, Active Shape Model, Expectation-Maximization

1. INTRODUCTION

Several left ventricle (LV) segmentation algorithms for cardiac MRI have been proposed in the last decade [1], with the goal of relieving cardiologists of the tedious task of manually tracing the LV border to assess cardiac function of a patient. However, many of these algorithms struggle with distinguishing the LV border from other anatomical structures, such as the papillary muscles, leading to incorrect segmentations.

Shape models, such as the Active Shape Model (ASM) [2], have been a popular approach to prevent unexpected shape estimates in the segmentation [3, 4, 5, 6]. An ASM based approach comprises two phases: 1) a training phase, in which the expected shape and variation of the LV are learned from training data; and 2) a test phase, in which the learned shape model is used to guide the segmentation of new data. In the latter, after initializing the model, candidate border points are extracted from the image and used to estimate the ASM parameters that fit the model to the desired border. Despite being powerful segmentation tools, the accuracy of ASMs depends on the image quality and is usually low in the presence of outliers [7]. Alternative approaches have proposed robust active shape models (RASM) to deal with this issue [7, 8], that are able to disregard the outliers in the image either by using only a subset of the candidate border points or by weighting them according to some criteria.

In this work, a new Bayesian formulation for a RASM is proposed for 3D segmentation of the LV, inspired in the 2D robust estimation method proposed in [8]. In this new formulation, the candidate border points in each MRI slice are first grouped into edge segments, denoted as strokes. Then, the strokes from the whole MRI volume are used to guide the segmentation using an Expectation-Maximization framework. This means that more reliable information is provided to the RASM, thus improving its accuracy.

The remainder of this paper is organized as follows. Section 2 describes the proposed formulation. Section 3 provides the experimental setup to evaluate the proposed method and shows the obtained results. Finally, Section 4 concludes the paper with final remarks.

2. ROBUST ACTIVE SHAPE MODEL

As mentioned above, an ASM based approach requires a training phase to learn the expected shape and the main modes of deformation. The procedure used in this work for the training phase is based on the method proposed in [6], which allows a shape model to be learned from MRI volumes with a variable number of slices. Using this approach, the LV segmentation in a particular slice, m, is given by the contour model $\widetilde{\boldsymbol{x}}(s_m) = \left[\widetilde{\boldsymbol{x}}^1(s_m)^\top, \widetilde{\boldsymbol{x}}^2(s_m)^\top, \ldots, \widetilde{\boldsymbol{x}}^N(s_m)^\top\right]^\top \in \mathbb{R}^{2N\times 1}$, where s_m is the position of the m-th slice along the LV axis, N is the number of points in the model. The 2D position of the i-th model points $\widetilde{\boldsymbol{x}}^i(s_m) = \left[\widetilde{x}_1^1(s_m), \quad \widetilde{x}_2^1(s_m)\right]^\top \in \mathbb{R}^2$ is given by

$$\widetilde{\boldsymbol{x}}^i(s_m) = \boldsymbol{A}(\overline{\boldsymbol{x}}^i(s_m) + \boldsymbol{D}^i(s_m)\boldsymbol{b}(s_m)) + \boldsymbol{t}, \qquad (1)$$

where $\overline{\boldsymbol{x}}^i(s_m) \in \mathbb{R}^{2\times 1}$ is the expected position of the i-th model point, $\boldsymbol{D}^i(s_m) \in \mathbb{R}^{2\times K}$ is the matrix of the K main modes of deformation of the i-th model point, $\boldsymbol{b}(s_m) \in \mathbb{R}^{K\times 1}$

This work was supported by FCT [SFRH/BD/87347/2012] and [UID/EEA/50009/2013].

are the deformation coefficients, $\boldsymbol{A}$ is a rotation and scaling matrix with parameters $\boldsymbol{a} = [a_1, a_2]^\top$, and $\boldsymbol{t} = [t_1, t_2]^\top$ are the translation parameters.

During the test phase, the goal is to determine the best model parameters $\boldsymbol{\theta} = \{\boldsymbol{a}, \boldsymbol{t}, \boldsymbol{b}(s_1), \ldots, \boldsymbol{b}(s_S)\}$, such that $\widetilde{\boldsymbol{x}}(s_m), m = 1, \ldots, S$, fits the LV border in all of the S slices of the test volume. Assuming an initial guess of the model parameters $\boldsymbol{\theta}$ has been provided, the segmentation process starts by extracting candidate border points from the volume slices. In this work, these candidate points, denoted as observations, are obtained by detecting edge points along search lines orthogonal to the contour model, using the same edge detector as [8]. Fig. 1 (left) shows an example of the detected edge points (red dots) in a particular slice of the MRI volume.

Since most of these edge points belong to a common edge in the image, using them separately to estimate the model parameters is not taking advantage of all the information available. Therefore, the detected edge points are grouped into edge segments, denoted as strokes, using the Mutual Favorite Pairing algorithm [9]. This algorithm determines if two individual edge points should belong to the same group if they mutually the best match according to a distance criteria, which in this work is defined by the difference between the signed distance of the edge point to the corresponding model point (i.e., the model point located along the same search line). Fig. 1 (middle) shows the strokes obtained from the edge points extracted from the images (left). Note that not all the detected strokes correspond to the LV border. Some strokes are outliers associated with papillary muscles (within the LV) and with the epicardium, that should not be considered for the segmentation. The next section describes the observation models used to explain the relationship between the model points and both types of strokes.

2.1. Observation Model

Let $\boldsymbol{\mathcal{Y}}(s_m) = \{\boldsymbol{Y}^1(s_m), \ldots, \boldsymbol{Y}^{P^m}(s_m)\}$ denote the set of all the detected strokes on the m-th slice. The i-th stroke is given by M^i edge points, $\boldsymbol{Y}^i(s_m) = \left[\boldsymbol{y}^{i1}(s_m)^\top, \ldots, \boldsymbol{y}^{iM^i}(s_m)^\top\right] \in \mathbb{R}^{2M^i \times 1}$, where $\boldsymbol{y}^{ij}(s_m) \in \mathbb{R}^{2\times 1}$ is the position of the j-th edge point in the i-th stroke. Also, let $\boldsymbol{X}^i(s_m) \in \mathbb{R}^{2M^i\times 1}$ denote the corresponding model segment comprising the model points, $\widetilde{\boldsymbol{x}}^{ij}(s_m), j = 1, \ldots, M^i$. Fig. 1 (right) shows an example for the third stroke, $\boldsymbol{Y}^3(s_5)$, located on the slice $m = 5$. If all the detected strokes were used to determine the shape model parameters, the resulting segmentation would try to fit the valid strokes (located along the LV border) and the invalid ones (outliers) simultaneously, leading to incorrect estimates of the LV border.

Since we do not know which strokes are valid and which are outliers, two observation models are considered. Let $k^i(s_m)$ be an unknown binary label associated with the i-th stroke, such that $k^i(s_m) = 1$ if $\boldsymbol{Y}^i(s_m)$ is valid, and $k^i(s_m) = 0$ otherwise. For valid strokes, we assume that each edge point in that stroke is given by

$$\boldsymbol{y}^{ij}(s_m) = \widetilde{\boldsymbol{x}}^{ij}(s_m) + \boldsymbol{v}^{ij}(s_m), \tag{2}$$

where $\boldsymbol{v}^{ij}(s_m) \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}^{ij}(s_m))$ is a zero mean white Gaussian noise with diagonal covariance matrix $\boldsymbol{\Sigma}^{ij}(s_m)$. Assuming statistical independence between the edge points in a stroke, it is possible to conclude that the probability of the i-th stroke, $\boldsymbol{Y}^i(s_m)$, knowing $k^i(s_m) = 1$ is

$$\begin{aligned} p\left(\boldsymbol{Y}^i(s_m)\middle|k^i(s_m)\text{=}1, \boldsymbol{\Theta}\right) &= \prod_{j=1}^{M^i} p\left(\boldsymbol{y}^{ij}(s_m)\middle|k^i(s_m)\text{=}1, \boldsymbol{\Theta}\right) \\ &= \prod_{j=1}^{M^i} \mathcal{N}\left(\boldsymbol{y}^{ij}(s_m); \widetilde{\boldsymbol{x}}^{ij}(s_m), \boldsymbol{\Sigma}^{ij}(s_m)\right). \end{aligned} \tag{3}$$

On the other hand, outlier strokes are assumed to follow a uniform distribution

$$p\left(\boldsymbol{Y}^i(s_m)\middle|k^i(s_m)\text{=}0, \boldsymbol{\Theta}\right) = \prod_{j=1}^{M^i} \mathcal{U}(V_{\widetilde{\boldsymbol{x}}^{ij}(s_m)}), \tag{4}$$

within a validation gate $V_{\widetilde{\boldsymbol{x}}^{ij}(s_m)}$ in the vicinity of $\widetilde{\boldsymbol{x}}^{ij}(s_m)$. Finally, we assume that the two labels, $k^i(s_m) = 1$ and $k^i(s_m) = 0$, occur with probabilities $P(k^i(s_m) = 1) = p_1$ and $P(k^i(s_m) = 0) = p_0$, respectively, with $p_0 + p_1 = 1$, which are also unknown.

Using the above formulation, the segmentation of the LV is obtained by finding the parameters $\boldsymbol{\Theta} = \{\boldsymbol{\theta}, p_0, p_1\}$ that maximize the joint probability

$$\begin{aligned} \mathcal{P}(\boldsymbol{\mathcal{Y}}, \boldsymbol{\mathcal{K}}, \boldsymbol{\Theta}) &= \log p(\boldsymbol{\mathcal{Y}}, \boldsymbol{\mathcal{K}}, \boldsymbol{\Theta}) \\ &= \log p\left(\boldsymbol{\mathcal{Y}}|\boldsymbol{\mathcal{K}}, \boldsymbol{\Theta}\right) + \log p(\boldsymbol{\mathcal{K}}) + \log p(\boldsymbol{\Theta}), \end{aligned} \tag{5}$$

where $\boldsymbol{\mathcal{Y}} = \{\boldsymbol{\mathcal{Y}}(s_1), \ldots, \boldsymbol{\mathcal{Y}}(s_S)\}$ is the set of all the detected strokes in all the slices of the volume, and $\boldsymbol{\mathcal{K}} = \{\boldsymbol{\mathcal{K}}(s_1), \ldots, \boldsymbol{\mathcal{K}}(s_S)\}$ is the set of the corresponding labels, where $\boldsymbol{\mathcal{K}}(s_m) = \{k^1(s_m), \ldots, k^{P^m}(s_m)\}$. Assuming that the strokes are conditionally independent, (5) can be rewritten as

$$\begin{aligned} \mathcal{P}(\boldsymbol{\mathcal{Y}}, \boldsymbol{\mathcal{K}}, \boldsymbol{\Theta}) = \sum_{m=1}^{S}\sum_{i=1}^{P^m}\sum_{j=1}^{M^i} &\log p\left(\boldsymbol{y}^{ij}(s_m)\middle|k^i(s_m), \boldsymbol{\Theta}\right) \\ &+ \log p(k^i(s_m)) + \log p(\boldsymbol{\Theta}). \end{aligned} \tag{6}$$

The first term corresponds to the likelihood of the strokes, the second term is the probability of the observation model, $k^i(s_m)$, and the last term is related to a prior probability for $\boldsymbol{\Theta}$, which in this work is defined as in [8].

Since the analytical maximization of (6) is infeasible, the Expectation-Maximization (EM) algorithm is used. The following section describes EM framework for this problem.

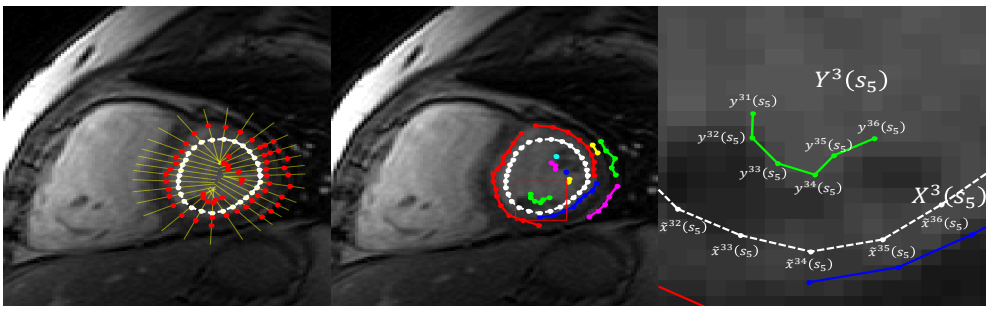

Fig. 1. Strokes detected in a volume slice: (left) detection of multiple edge points along lines orthogonal to the contour model; (middle) edge points grouped into strokes; (right) zoom on a particular stroke.

2.2. Expectation Maximization algorithm

The EM algorithm obtains the maximization of (6) by iterating between two steps: 1) *E-step*, in which the expectation of the (6) is updated by computing the probability of each stroke based on the current parameters; and 2) *M-step*, in which the parameters are updated by maximizing the expectation obtained in the E-step.

Let $\widehat{\boldsymbol{\Theta}}_{(t)}$ denote the current estimate of the model parameters, where $\widehat{\boldsymbol{\Theta}}_{(0)}$ is the initial guess provided by the initialization. At each iteration, a new set of strokes, $\boldsymbol{\mathcal{Y}}$, is extracted from the volume. In the E-step, an auxiliary function is defined as

$$Q\left(\boldsymbol{\Theta};\widehat{\boldsymbol{\Theta}}_{(t)}\right)=\mathbb{E}_{\boldsymbol{\mathcal{K}}}\left[\mathcal{P}(\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{K}},\boldsymbol{\Theta})\middle|\boldsymbol{\mathcal{Y}},\widehat{\boldsymbol{\Theta}}_{(t)}\right], \tag{7}$$

where $\mathbb{E}_{\boldsymbol{\mathcal{K}}}[\cdot]$ denotes the expectation over the two observation models. Combining (7) with (3) and (4) leads to

$$\begin{aligned}Q\left(\boldsymbol{\Theta};\widehat{\boldsymbol{\Theta}}_{(t)}\right) &\propto c+\log p(\boldsymbol{\Theta})\\ &+\sum_{m=1}^{S}\sum_{i=1}^{P^m}\sum_{j=1}^{M^i} w_0^i(s_m)\left[\log\mathcal{U}(V_{\widetilde{\boldsymbol{x}}^{ij}(s_m)})+\log p_0\right]\\ &+w_1^i(s_m)\left[\left(\boldsymbol{y}^{ij}(s_m)-\boldsymbol{A}\left(\overline{\boldsymbol{x}}(s_m)+\boldsymbol{D}(s_m)\boldsymbol{b}(s_m)\right)-\boldsymbol{t}\right)^{\top}\right.\\ &\left.\boldsymbol{\Sigma}^{ij\,-1}(s_m)\left(\boldsymbol{y}^{ij}(s_m)-\boldsymbol{A}\left(\overline{\boldsymbol{x}}(s_m)+\boldsymbol{D}(s_m)\boldsymbol{b}(s_m)\right)-\boldsymbol{t}\right)\right],\end{aligned} \tag{8}$$

where w_0^i and w_1^i denote the confidence degree of each stroke, such that $w_0^i(s_m)+w_1^i(s_m)=1$ and

$$w_1^i(s_m)\propto\widehat{p}_{1(t)}\,\mathcal{N}\left(\boldsymbol{y}^{ij}(s_m);\widetilde{\boldsymbol{x}}^{ij}(s_m),\Sigma^{ij}(s_m)\right) \tag{9}$$

$$w_0^i(s_m)\propto\widehat{p}_{0(t)}\,\mathcal{U}\left(V_{\widetilde{\boldsymbol{x}}^{ij}(s_m)}\right). \tag{10}$$

These weights correspond to the probability of $k_1^i(s_m)$ and $k_0^i(s_m)$ being the correct label for stroke $\boldsymbol{Y}^i(s_m)$, given the current model estimate.

In the M-step, the model parameters are updated by solving the optimization problem

$$\widehat{\boldsymbol{\Theta}}_{(t+1)}=\arg\max_{\boldsymbol{\Theta}} Q\left(\boldsymbol{\Theta};\widehat{\boldsymbol{\Theta}}_{(t)}\right). \tag{11}$$

We simplify this optimization by maximizing first with respect to the transformation parameters, $\boldsymbol{a},\boldsymbol{t}$, and only then for $\boldsymbol{b}(s_1),\ldots,\boldsymbol{b}(s_S)$ and p_0 and p_1. In all the cases, the update equations are obtained by taking the straightforward derivative of (8) with respect to each parameter, leading to linear systems of equations (as in [8]). However, the difference in this case is that each edge point in a stroke has the same weight, $w_1^i(s_m)$, computed in (9). In the case of the transformation parameters, $\boldsymbol{a},\boldsymbol{t}$, and the model probabilities, p_0,p_1, all the slices in the volume contribute to their estimation, while the deformation parameters in the m-th slice, $\boldsymbol{b}(s_m)$ are computed using only the strokes detected in that slice. In practice, the update equations correspond to a weighted least squares solution to the problem of determining a one-to-many mapping between strokes extracted from the volume and the corresponding model points. The weight of each stroke, $w_1^i(s_m)$, determines its contribution to the estimation of the parameters based on the probability of being a valid stroke. Consequently, the model iteratively updates towards fitting the strokes with higher probabilities. The segmentation process stops once the changes in parameters from one iteration to the next is below a specific threshold.

3. RESULTS

To evaluate the proposed method, a public dataset of cardiac MRI [3] is used. This dataset consists of MRI sequences of 20 volumes for 33 patients (most of them diseased), with a total of 660 volumes to be segmented. Annotations of the LV segmentation are provided and are used as ground truth (GT). Each volume has 4-10 slices in which the LV is visible (the remaining slices are disregarded) and each volume slice is a 256×256 image with $0.9-1.6$ mm/pixel of resolution. The spacing between slices is 6-13mm.

The training and test phases are performed in a cross validation fashion, i.e., for each test sequence, the remaining 32×20 annotated volumes are used to learn the shape model. Each volume in the test sequence is segmented independently, which means that no tracking information is used. The initialization of the transformation parameters, $\boldsymbol{t}$ is obtained by

human input of the rough location of the LV center in the basal slice, and the other parameters are set to $\widehat{a}_{2(0)} = 0$, $\widehat{\boldsymbol{b}}_{(0)}(s_m) = \mathbf{0}$ and $\widehat{p}_{0(0)} = \widehat{p}_{1(0)} = 0.5$. The quantitative evaluation is based on two metrics: 1) the Dice similarity coefficient; and 2) the average perpendicular distance (AV).

Fig. 2 shows examples of the segmentations obtained and the corresponding manual annotations for 4 slices of 5 different volumes. It is possible to see that the automatic segmentation is similar to the GT in most cases. Table 1 shows the overall statistical results of the proposed method and presents a comparison with other state of the art methods. The table shows that the proposed approach outperforms the standard 3D ASM [2], as well as the approach used in [6]. Furthermore, it achieves competitive results against the LV endocardium segmentation method proposed in [10] (see performance for *"all"* in Table 1, p. 605)[1] and the method proposed in [11] (see Table 1, p. 6). Note that the results from [10] and [11] were obtained on other datasets.

Table 1. Comparison of the statistical performance. The last column shows the percentage of volumes used to obtain the results (volumes with AV < 5 mm were excluded). Dashed entries mean all volumes were included.

	Dice (%)	**AV (mm)**	**% Good**
Cootes et al. [2]	73 ± 13	4.7 ± 3.0	-
Santiago et al. [8]	79 ± 8	3.5 ± 1.4	-
Huang et al. [10]	89 ± 4	2.2 ± 0.5	81.5 ± 18.0
Gopal et al. [11]	84 ± 4	3.7 ± 0.6	-
Proposed	80 ± 7	3.3 ± 1.2	-
	83 ± 6	2.6 ± 1.1	84.8

4. CONCLUSION

A new formulation for a robust active shape model is proposed for 3D segmentation of the LV in cardiac MRI. The proposed approach extends a previous approach [8] to 3D data and uses edge segments (strokes) extracted from the volume slices to guide the segmentation, instead of the traditional edge points. The results show that the proposed algorithm provides accurate segmentations that are competitive with other state-of-the art methods.

5. REFERENCES

[1] C. Petitjean and J. Dacher, "A review of segmentation methods in short axis cardiac MR images," *Medical image analysis*, vol. 15, no. 2, pp. 169–184, 2011.

[2] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models-their training and application," *Computer vision and image understanding*, vol. 61, no. 1, pp. 38–59, 1995.

[1]The *all* performance is chosen, since it is the comparison modality that contains only the endocardium of the LV with healthy and diseased patients, such as in our database.

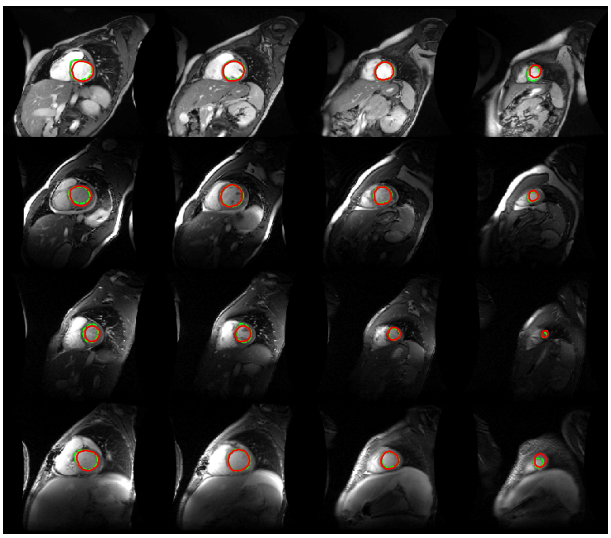

Fig. 2. Examples of the segmentations obtained with the proposed algorithm (red) and the GT (green). Each row shows a different volume and each column shows different slices from the base (left) to the apex (right).

[3] A. Andreopoulos and J.K. Tsotsos, "Efficient and generalizable statistical models of shape and appearance for analysis of cardiac MRI," *Medical Image Analysis*, vol. 12, no. 3, pp. 335–357, 2008.

[4] T. Heimann and H. Meinzer, "Statistical shape models for 3D medical image segmentation: A review," *Medical image analysis*, vol. 13, no. 4, pp. 543–563, 2009.

[5] H. Zhang et al., "4-D cardiac MR image analysis: left and right ventricular morphology and function," *Medical Imaging, IEEE Transactions on*, vol. 29, no. 2, pp. 350–364, 2010.

[6] C. Santiago, J.C. Nascimento, and J.S. Marques, "Robust 3D Active Shape Model for the Segmentation of the Left Ventricle in MRI," in *Pattern Recognition and Image Analysis*, vol. 9117 of *Lecture Notes in Computer Science*, pp. 283–290. Springer International Publishing, 2015.

[7] M. Rogers and J. Graham, "Robust active shape model search," in *Computer Vision–ECCV 2002*, pp. 517–530. Springer, 2006.

[8] C. Santiago, J.C. Nascimento, and J.S. Marques, "2D Segmentation Using a Robust Active Shape Model With the EM Algorithm," *Image Processing, IEEE Transactions on*, vol. 24, no. 8, pp. 2592–2601, Aug 2015.

[9] D.P. Huttenlocher and S. Ullman, "Recognizing solid objects by alignment with an image," *International Journal of Computer Vision*, vol. 5, no. 2, pp. 195–212, 1990.

[10] S. Huang et al., "An image-based comprehensive approach for automatic segmentation of left ventricle from cardiac short axis cine MR images," *Journal of digital imaging*, vol. 24, no. 4, pp. 598–608, 2011.

[11] S. Gopal and D. Terzopoulos, "A Unified Statistical/Deterministic Deformable Model for LV Segmentation in Cardiac MRI," in *Statistical Atlases and Computational Models of the Heart. Imaging and Modelling Challenges*, pp. 180–187. Springer, 2014.