Segmentation of the Left Ventricle in Cardiac MRI Using a Probabilistic Data Association Active Shape Model

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

Abstract—The 3D segmentation of endocardium of the left ventricle (LV) in cardiac MRI volumes is a challenging problem due to the intrinsic properties of this image modality. Typically, the object shape and position are estimated to fit the observed features collected from the images. The difficulty inherent to the LV segmentation in MRI is that the images contain outliers (i.e., observations not belonging to the LV border) due to the presence of other structures. This paper proposes a robust approach based on the Active Shape Model (ASM) that is able to circumvent the above problem. More specifically, the ASM will be guided by probabilistic data association filtering (PDAF) of strokes (i.e. line segments) computed in the neighborhood of the shape model. Thus, the proposed approach, termed herein as ASM-PDAF, will perform the following main steps: 1) edge detection (low-level features) in the vicinity of the shape model; 2) edge grouping (mid-level features) to obtain potential LV strokes; and 3) filtering using a PDAF framework (highlevel features) to update the ASM. Experimental results on a public cardiac MRI database show that the proposed approach outperforms previous literature research.

I. INTRODUCTION

The 3D segmentation of the endocardium of the left ventricle (LV) in cardiac MRI is a challenging problem [1]. A popular approach relies on the use of a shape model, namely, the Active Shape Model (ASM). This model is guided by image observation (or feature) points, computed in each slice of the MRI volume, in the attempt to provide a fairly good LV segmentation. However, if the observations are not located at the LV boundary (i.e., they are outliers), the ASM model is not able to fit the true boundary, providing misleading results. This paper proposes a robust ASM algorithm that is able to deal with outliers, inspired in the probabilistic data association filter (PDAF) [2]. This methodology has proved to be competitive in several contexts, e.g., in lip and vehicle tracking [3], LV tracking in ultrasound (US) images [4], or other medical applications [5]. The advantage of the proposed approach over these methods is the use of a shape model guided by PDAF, instead of a generic deformable model with no shape prior that may provide unpredictable shape estimates.

The ASM-PDAF formulation is based on two key concepts. First, it does not rely on single observation points (low-level features). Instead, mid-level features (strokes) are computed by grouping observation points into strokes. Second, two labels (valid/invalid) are assigned for each stroke. Since the stroke labels are unknown *a priori*, we have to consider all possible combinations of valid/invalid labels a termed as a *data interpretation*. A confidence degree is assigned to each data interpretation, called *data association probabilities*, that determines its influence in the estimation of the segmentation, thus all the data interpretations contribute to segmenting the LV with different weights. This allows for a robust segmentation when the image contains outliers.

Previous methods have been proposed to deal with outliers, e.g., by using random sampling methods [6], or applying the Robust Point Matching algorithm [7]. Similar to our work, a recently new method was proposed in [8], based on the Expectation-Maximization algorithm, which assumes that each observation point may be an outlier. Although this approach shows improvements in the segmentation accuracy over the previous methods, it does not take advantage of all the information available about the points extracted from the volume slices. The method proposed in this paper addresses the segmentation in a similar way to [8], but explores a different approach to define the observation model that is based on PDAF.

II. ROBUST ACTIVE SHAPE MODEL

The Active Shape Model (ASM) method is based on the assumption that the shape of an object can be learned from a dataset of labeled volumes. Therefore, the segmentation of a test volume can be defined by the mean shape and a linear combination of the main modes of deformation contained in the dataset. In the LV segmentation problem, the MRI volumes are characterized by having a low resolution along the third dimension and with a variable number of slices. Consequently, it is not easy to learn a 3D shape model. We use the approach proposed in [9] to define the 3D shape model as a function of the slice position and to learn the mean shape and the main modes of deformation.

Let $\mathbf{x}(s_m) \in \mathbb{R}^{2N \times 1}$ be the model at the *m*-th slice, defined by N points. The position of the *i*-th model point is given by

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

where $\overline{\boldsymbol{x}}^{i}(s_{m})$ is the *i*-th point in the mean shape, $\boldsymbol{D}^{i}(s_{m})$ is a $2 \times K$ matrix with the K main modes of deformation of the *i*-th model point, $\boldsymbol{b}(s_{m})$ is a vector with the deformation coefficients, and

$$\boldsymbol{A} = \left[\begin{array}{cc} a_1 & -a_2 \\ a_2 & a_1 \end{array} \right], \quad \boldsymbol{t} = \left[\begin{array}{c} t_1 \\ t_2 \end{array} \right]$$

C. Santiago, J. Nascimento and J. Marques are with Instituto de Sistemas e Robótica, Instituto Superior Técnico, Lisboa, Portugal.

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



Fig. 1. Extraction of strokes from a volume slice, in this particular case the second slice s_2 is considered. (Left) detection of edge points (in white) along lines (in yellow) orthogonal to the contour model (in red); (middle) edge points linked to form strokes (each having a different color); (right) zoom in of the 7th stroke (yellow).

are parameters of a global transformation that scales, rotates and translates the shape model.

Given a test volume, the segmentation of the LV is obtained by determining the global transformation parameters, $\boldsymbol{\theta} = \{\boldsymbol{a}, \boldsymbol{t}\}, \text{ and deformation coefficients, } \boldsymbol{b}(s_m), \text{ such that}$ $\boldsymbol{x}(s_m)$ fits the LV border in all slices, $m = 1, \ldots, S$. Consequently, the first step of the segmentation is to locate LV border in the volume slices. Several approaches may be used for this purpose. In this work, we will assume that the LV border is located along edges on the volume slices. Thus, for each slice, we perform the edge detection method described in [10], which consists of detecting edge points along search lines orthogonal to the contour model (see Fig. 1 left). Then, we group the edge points to form strokes (see Fig. 1 middle). Unfortunately, some of the detected strokes are outliers (i.e., they do not belong to the LV border) and should not be taken into account when computing the model parameters. This means that segmentation method must be able to simultaneously estimate the segmentation and identify which strokes are reliable. The proposed formulation is able to tackle these two challenges, as will be detailed in the following sections.

A. Observation Model

Suppose that an initial guess for the location of the LV border is available, i.e., the model parameters at t = 0, $\Theta_{(t_0)} = \{a, t, b(s_1), \dots, b(s_m)\}_{(t_0)}$, are known.

The observation model is based on strokes, obtained using the Mutual Favorite Paring algorithm [11], which groups edge points into line segments based on a distance criterion. An example is shown in Fig. 1 (left to middle), where, from a set of 40 edge points (left), nine strokes are formed (right), significantly reducing the complexity of the observation model.

Formally, for a particular slice s, let $\mathbf{Y}(s) = \{\mathbf{Y}^1(s), \dots, \mathbf{Y}^L(s)\}$ denote a set of L strokes, where $\mathbf{Y}^i(s) = \left[\mathbf{y}^{i1}(s)^{\top}, \dots, \mathbf{y}^{iM^i}(s)^{\top}\right]^{\top} \in \mathbb{R}^{2M^i \times 1}$ is the *i*-th stroke, and $\mathbf{y}^{ij}(s)$ is the *j*-th edge point in that stroke.

Also, let $\boldsymbol{X}^i(s) = \left[\boldsymbol{x}^{i1}(s)^{\top}, \dots, \boldsymbol{x}^{iM^i}(s)^{\top} \right]$ be the model associated to stroke $\dot{\mathbf{Y}}^{i}(s)$, such that $\mathbf{y}^{ij}(s)$ is detected along a search line passing through, $x^{ij}(s)$ (see Fig. 1 right). The goal of the segmentation algorithm is to fit the shape model to the valid strokes, i.e., to the strokes located on the LV border. Finally, let $I^i(s) = \{0,1\}$ be a binary label assigned to the *i*-th stroke $Y^{i}(s)$ at slice s, where $I^{i}(s) = 1$ means it is valid, and $I^{i}(s) = 0$ means it is invalid (outlier). We do not know which strokes are valid, i.e., the labels $I^1(s), \ldots, I^L(s)$ are unknown. Consequently, a total of 2^L possible combinations of valid/invalid strokes (henceforth denote as data interpretations) have to be taken into consideration. However, not all the strokes can be simultaneously considered as valid. If two strokes, say $Y^{1}(s)$ and $Y^{2}(s)$, overlap with respect to the shape model, the data interpretations considering $I^{1}(s) = I^{2}(s) = 1$ have to be disregarded, because they cannot be both located at the LV border. For instance, in Fig. 1 right, the yellow and green strokes cannot be simultaneously considered valid. This reduces the number of possible data interpretations.

The LV border is given by a particular interpretation, say $I^{\star}(s) = \{I^{1^{\star}}(s), \dots, I^{L^{\star}}(s)\}$. We define the probability of Y(s) as

$$p(\mathbf{Y}(s) \mid \boldsymbol{\Theta}, I^{\star}(s)) = \prod_{i=1}^{L} \begin{bmatrix} I^{i^{\star}}(s) & 1 - I^{i^{\star}}(s) \end{bmatrix} \begin{bmatrix} \mathcal{N}\left(Y^{i}(s); \mathbf{X}^{i}(s), \mathbf{\Sigma}^{i}(s)\right) \\ \mathcal{U}(V_{\mathbf{X}^{i}(s)}) \end{bmatrix},$$
(2)

where $\mathcal{N}(\cdot; \mu, \Sigma)$ is a Gaussian distribution with mean μ and covariance Σ , and $\mathcal{U}(V_{C^i(s)\boldsymbol{x}(s)})$ is a uniform distribution in the vicinity of $\boldsymbol{X}^i(s)$. This means that the valid strokes are generated by adding a Gaussian perturbation to the LV border, and that the outliers are generated with equal probability within a validation gate (i.e. the length of the orthogonal lines) around the LV border. However, we do not know $I^*(s)$, which means we have to simultaneously determine



Fig. 2. Example of data interpretations on a particular slice of the MRI volume: (left) the detected strokes, each with a different color (the white line corresponds to the current model estimate), and each of the following images show the most likely interpretations (higher probability, $w_k(s_m)$).

the model parameters and the interpretation $I^*(s)$. In this work, this is achieved using the Expectation-Maximization (EM) algorithm [12] that we detail next.

B. Expectation-Maximization

Let $\mathcal{Y} = \{\mathbf{Y}(s_1), \dots, \mathbf{Y}(s_S)\}$ be the set of strokes detected in all the slices in the volume. Also, let $\mathcal{I} = \{I(s_1), \dots, I(s_S)\}$ be the corresponding interpretations. Ideally, the model parameters would be estimated by solving the following problem

$$\arg\max_{\boldsymbol{\Theta}\mathcal{I}} \mathcal{P}\left(\boldsymbol{\mathcal{Y}}, \mathcal{I}, \boldsymbol{\Theta}\right), \tag{3}$$

where $\mathcal{P}(\mathcal{Y}, \mathcal{I}, \Theta)$ is the log-posterior probability, which is defined by

$$\mathcal{P}(\boldsymbol{\mathcal{Y}}, \mathcal{I}, \boldsymbol{\Theta}) = \sum_{m=1}^{S} \log p\left(\boldsymbol{Y}(s_m), \boldsymbol{\Theta}, I(s_m)\right), \quad (4)$$

assuming that the label combinations for each volume slice are statistically independent. The problem formulated in (4) is infeasible, because it involves a marginalization over a large number of possible interpretations. The EM algorithm allows the computation of a suboptimal solution to this problem, in which the labels, $I^1(s_m), \ldots, I^L(s_m), m =$ $1, \ldots, S$, are hidden variables of the model (i.e., they are not observed).

The EM algorithm finds the model parameters by iterating between two steps: 1) the *E*-step, in which the expectation of $\mathcal{P}(\mathcal{Y}, \mathcal{I}, \Theta)$ is computed; and 2) the *M*-step, in which the model parameters are updated by maximizing the expectation obtained in the E-step. These steps are explained below.

1) *E-step:* Let $\widehat{\Theta}_{(t)}$ denote the current estimate of the model parameters, for iteration *t*, where $\widehat{\Theta}_{(0)}$ is the initial guess for these parameters. At each iteration, a new set of strokes, \mathcal{Y} , is extracted from the volume. We define the auxiliary function

$$Q\left(\boldsymbol{\Theta}; \widehat{\boldsymbol{\Theta}}_{(t)}\right) = \mathbb{E}_{\mathcal{I}}\left[\mathcal{P}(\boldsymbol{\mathcal{Y}}, \mathcal{I}, \boldsymbol{\Theta}) \middle| \boldsymbol{\mathcal{Y}}, \widehat{\boldsymbol{\Theta}}_{(t)}\right], \qquad (5)$$

where $\mathbb{E}_{\mathcal{I}}[\cdot]$ denotes the expectation over all possible combinations of labels, i.e. the data interpretations. This leads

to

$$Q\left(\boldsymbol{\Theta}; \widehat{\boldsymbol{\Theta}}_{(t)}\right) = \sum_{m=1}^{S} \sum_{k} w_k(s_m) \log p\left(\boldsymbol{Y}(s_m), \boldsymbol{\Theta}, I_k(s_m)\right),\tag{6}$$

where, for each slice m, $I_k(s_m) = [I_k^1(s_m), \ldots, I_k^L(s_m)]$ is the k-th possible interpretation, $w_k(s_m)$ is the probability of $I_k(s_m)$, given by

$$w_{k}(s_{m}) = p\left(I_{k}(s_{m}) \middle| \mathbf{Y}(s_{m}), \widehat{\mathbf{\Theta}}_{(t)}\right)$$

$$\propto p\left(\mathbf{Y}(s_{m}) \middle| \widehat{\mathbf{\Theta}}_{(t)}, I_{k}(s_{m})\right) p\left(I_{k}(s_{m}) \middle| \widehat{\mathbf{\Theta}}_{(t)}\right) (7)$$

such that $\sum_{k} w_{I_k(s)} = 1$, and

$$p(\mathbf{Y}(s_m), \mathbf{\Theta}, I_k(s_m)) = p(\mathbf{Y}(s_m)|\mathbf{\Theta}, I_k(s_m)) p(I_k(s_m)|\mathbf{\Theta}) p(\mathbf{\Theta}).$$
(8)

The first term in (8) is given by (2) and $p(\Theta)$ is a prior on the model parameters. In this work, we assume that all the interpretations are equally probable *a priori*, i.e., $p(I_k(s_m)|\Theta)$ is a constant. Figure 2 shows an example slice, depicting the detected strokes (left) and the three interpretations with higher probability.

2) *M-step:* In this step, the model parameters are updated by solving the following optimization problem

$$\widehat{\boldsymbol{\Theta}}_{(t+1)} = \arg \max_{\boldsymbol{\Theta}} Q\left(\boldsymbol{\Theta}; \widehat{\boldsymbol{\Theta}}_{(t)}\right) \,. \tag{9}$$

We simplify this step by maximizing first with respect to the transformation parameters, a, t, and only then for $b(s_1), \ldots, b(s_S)$. The update equations are obtained by straightforward derivatives operations, similar to ones in [8].

The method described above will be denoted as Active Shape Model with Probabilistic Data Association Filtering (ASM-PDAF).

III. RESULTS

The proposed method was evaluated on a public cardiac MRI dataset [13], which contains 33 sequences of healthy and diseased patients. Each sequence has 20 volumes at different phases covering one complete cardiac cycle. The volumes have a variable number of slices, ranging from 8-15, and the spacing between slices range from 6 to 13 mm.

Each volume slice is an image with 256×256 pixels and with a resolution of 0.93-1.64 mm. The dataset also provides a manual segmentation of the LV that will be used as ground truth.

The shape model was learned using a leave-one-sequenceout scheme, which means that in order to segment the volumes of one particular sequence, the shape model was learned using all the remaining 32 sequences. The segmentations were evaluated using the Dice coefficient [14], d_{Dice} , and the average minimum distance between the segmentation and the ground truth, d_{AV} .

The shape model was initialized using $\mathbf{b}_{(0)}(s_m) = \mathbf{0}$, $m = 1, \ldots, S$ and $\mathbf{a}_{(0)} = \begin{bmatrix} 1 & 0 \end{bmatrix}^{\top}$, i.e., using the mean shape (undeformed) with no scaling and no rotation. The translation parameters, $\mathbf{t}_{(0)}$, were obtained by manually providing a point in the center of the LV. Figure 3 shows examples of the obtained segmentations.

In order to show the benefits of the proposed approach, the results were compared with a previous approach, called EM-RASM [8], that considers individual edge points as observation points. Statistical results are shown in Table III for this method and for ASM-PDAF. It is possible to see that the proposed method outperforms the previous one.

 TABLE I

 Comparison of the results with a previous method.

	EM-RASM [8]	ASM-PDAF
d_{Dice}	78.9 (21.4)	81.1 (15.9)
$d_{\rm AV}$	2.05 (2.56)	1.6 (1.5)

IV. CONCLUSIONS

This paper proposes a robust approach to determine the parameters of a shape model based on the Active Shape Model (ASM) and the *probabilistic data association filter* (PDAF) for segmenting the LV, called ASM-PDAF. The method relies on strokes extracted from the volume slices, by performing: 1) an edge detection step (low-level features) in the vicinity of the shape model; and 2) an edge grouping step (mid-level features) to obtain potential LV strokes. Then, these strokes are filtered by using a PDAF framework (high-level features). The results show that the proposed approach outperforms a previous state-of-the-art approach (EM-RASM).

References

- 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] Y. Bar-Shalom, *Tracking and data association*. Academic Press Professional, Inc., 1988.
- [3] J. C. Nascimento and J. S. Marques, "Robust shape tracking in the presence of cluttered background," *Multimedia, IEEE Transactions on*, vol. 6, no. 6, pp. 852–861, 2004.
- [4] —, "Robust shape tracking with multiple models in ultrasound images," *Image Processing, IEEE Transactions on*, vol. 17, no. 3, pp. 392–406, 2008.
- [5] P. Abolmaesumi and M. R. Sirouspour, "An interacting multiple model probabilistic data association filter for cavity boundary extraction from ultrasound images," *Medical Imaging, IEEE Transactions on*, vol. 23, no. 6, pp. 772–784, 2004.



Fig. 3. Segmentation results. Each column shows a different volume (from a different sequence) and each row shows a different slice of the volume. The red line is the segmentation using the proposed method and the green dashed line is the ground truth.

- [6] M. Rogers and J. Graham, "Robust active shape model search," in *Computer Vision–ECCV 2002*. Springer, 2006, pp. 517–530.
- [7] J. Abi-Nahed, M. P. Jolly, and G. Z. Yang, "Robust active shape models: A robust, generic and simple automatic segmentation tool," in *Medical Image Computing and Computer-Assisted Intervention– MICCAI 2006.* Springer, 2006, pp. 1–8.
- [8] C. Santiago, J. C. Nascimento, and J. S. Marques, "Segmentation of 2D Objects Using a Robust Active Shape Model with the EM Algorithm," to appear in IEEE Transactions on Image Processing, 2015.
- [9] —, "Robust 3d active shape model for the segmentation of the left ventricle in mri," in *to be presented at IbPRIA*, 2015.
- [10] A. Blake and M. Isard, Active shape models. Springer, 1998.
- [11] 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.
- [12] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 1–38, 1977.
- [13] 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.
- [14] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.