

# A 3D Active Shape Model for Left Ventricle Segmentation in MRI

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

## Abstract

3D active shape models perform the estimation of deformable surfaces in 3D volumes using landmark statistics. First, the average shape and deformation modes are obtained from a set of annotated landmarks (training set). Then, these statistics are used to segment new volumes, involving the estimation of the alignment parameters and deformation coefficients. This approach is well suited to segment the shape of the left ventricle in 3D MRI volumes. However, there are several challenging issues since each MRI volume has a different number of slices, which means that it is difficult to establish correspondences between landmarks of different volumes. This leads to the main question: *how can we learn a shape model from volumes with a variable number of slices?* Motion artifacts and the large distance between slices make interpolation of voxel intensities a bad choice. The question can, thus, be reformulated: *how can we use active shape models without interpolating voxel intensities between slices?* This chapter provides an answer to these questions. We propose an interpolated model that allows the landmarks of each training volumes to be resampled. Then, we propose a resampling method for the learned shape model (mean shape and main modes of deformation), in order to ensure that it matches the slices of the test volume without requiring voxel interpolation. The proposed algorithm was tested in 20 MRI volumes using a leave-one-out scheme. The results show that it achieves good segmentation accuracy, with an average Dice coefficient of  $0.88 \pm 0.06$  and an average minimum distance to the ground truth of  $1.2 \pm 0.7$  mm.

## I. INTRODUCTION

Active shape models (ASMs) have been widely used in medical image analysis for their ability to include shape information in the segmentation process. Over the past decade, several ASM based methods have been proposed to address the 3D segmentation of medical images [1]. Most of them use the Point Distribution Model (PDM) [2] to learn the shape statistics from a set of annotated volumes (training set). The PDM defines the surface of an object using a set of labeled landmarks. These landmarks correspond to specific locations such that there is a correspondence between the  $i$ -th landmark on one surface and the  $i$ -th landmark on another surface. This allows the computation of shape statistics (mean shape and modes of deformation) to be computed. However, it requires that all the surfaces are described by the same landmarks.

The previous assumption is not always true. For instance, consider the left ventricle (LV) in 3D cardiac magnetic resonance (MR) volumes. An MR volume consists of a set of 2D images orthogonal to the LV axis and equally spaced. The surface of the LV is often defined by the 2D contours on each volume slice. However, the dimensions of the heart depend on the person and on the cardiac phase, which means the number of slices required to cover the heart varies. This leads to the following question: *how can we learn the shape statistics of the LV surface from volumes with a variable number of slices?*

This chapter addresses this problem by normalizing the surfaces of the volumes in the training phase, with respect to the number of slices. The normalization is achieved by modeling the position of each landmark (surface point) along the LV axis through interpolation. This allows the surface to be resampled at a set of predefined slices, which guarantees that the surface models in the training set have the same number of landmarks. A schematic illustration of this approach is shown in Fig. 1 (learning phase). Once the landmark correspondences have been established, we learn a shape model for each volume slice.

After computing the shape statistics (mean shape and main modes of deformation), we use them to segment a new volume, as shown in Fig. 1 (test phase). However, as before, this new volume may have

a different number of slices, which means that the 2D contours of the learned model may not match the volume slices. In most 3D segmentation problems this is not an issue because the interpolation of voxel intensities provides reliable data. This is not the case in cardiac MR volumes due to the low resolution along the LV axis (e.g., the spacing between slices is typically 1cm) and to motion artifacts [3]. Interpolating the voxel intensities in this scenario often leads to a loss of contrast between the blood pool (inside the endocardium of the LV) and the myocardium (outside the endocardium of the LV), making the segmentation task very difficult (see Section V for an example). This leads to a second question: *how can we segment a test volume without interpolating voxel intensities between slices?* We propose to interpolate the shape model instead, which means resampling the shape statistics (mean shape and the main modes of deformation). This way, the contours of the shape model are located at the same axial position as the volume slices, which means there is no need to interpolate voxel intensities.

The next section provides an overview of the state of the art for 3D segmentation using shape models. Sections III-IV describe the proposed shape representation and how it is used to resample the training surfaces. Section V explains the interpolation of the learned shape model and describes the estimation of the shape model parameters that segment the LV in a test volume. Section VI describes the experimental setup used to evaluate the proposed method and Section VII shows the results obtained. Final conclusions are presented in Section VIII.

## II. STATE OF THE ART

Learning 3D shape models, defined by landmarks is not a simple task [1]. In some problems, such as in cardiac MR volumes, there are no salient features in the volume and it is difficult to obtain unique landmark positions. Consequently, establishing correspondences between landmarks of two objects is not trivial. A popular approach [4]–[6] to overcome this issue is based on the Iterative Closest Point (ICP) algorithm [7]. This algorithm registers two shapes by iteratively establishing landmark correspondences between the two shapes using the best pose estimate and aligning the shapes based on those correspondences. Several variants of the ICP algorithm have been proposed to improve its results and to allow the use of different numbers of landmarks [8]–[10].

In the case of the LV segmentation problem, where the LV surface is defined by a set of contours (one per slice), the ICP would be unable to effectively align two LV surfaces with a different number of slices. That is why many works in the 3D LV segmentation problem use other approaches. Mitchell et al. [11], for example, propose the use of a normalized cylindrical coordinate system to define the position of the landmarks. First, they resample each contour at predefined angle intervals around the LV axis, which guarantees that all the 2D contours have the same number of landmarks. It also guarantees that there is a correspondence between the landmarks in consecutive slices. Then, they interpolate the position of the landmarks along the LV axis using linear interpolation. This allows them to resample the number of slices in a volume. Finally, by assuming a fixed position for the basal and axial slices in all the volumes, landmark correspondences between different volumes are easily established. A similar approach is also used by Andreopoulos et al. [3].

Another work [12] performs a preliminary step to establish the landmarks' position and correspondences, based on volumetric registration. Frangi et al. [13] also proposed an automatic method to determine the landmarks' position using this approach. They use a volumetric mesh, in which voxels are labeled based on the type of structure they belong to. These allow the alignment of the volume to a reference one using a non-rigid registration algorithm [14]. After the two volumes have been aligned, the reference landmarks are used to establish the position of the landmarks in the other volume.

As an alternative, the level-set method [15] has also been used. For instance, Tsai et al. [16] learn the shape statistics using the signed distance maps in the training set, instead of the landmark's positions. This approach has the advantage of not requiring landmark correspondences between shapes, as well as the advantage of being able to change the topology of the segmentation. The latter is not of particular interest for the problem of segmentation of the LV in MRI, since the LV does not change topology, but may be useful in other applications such as the segmentation of brain structures [17].

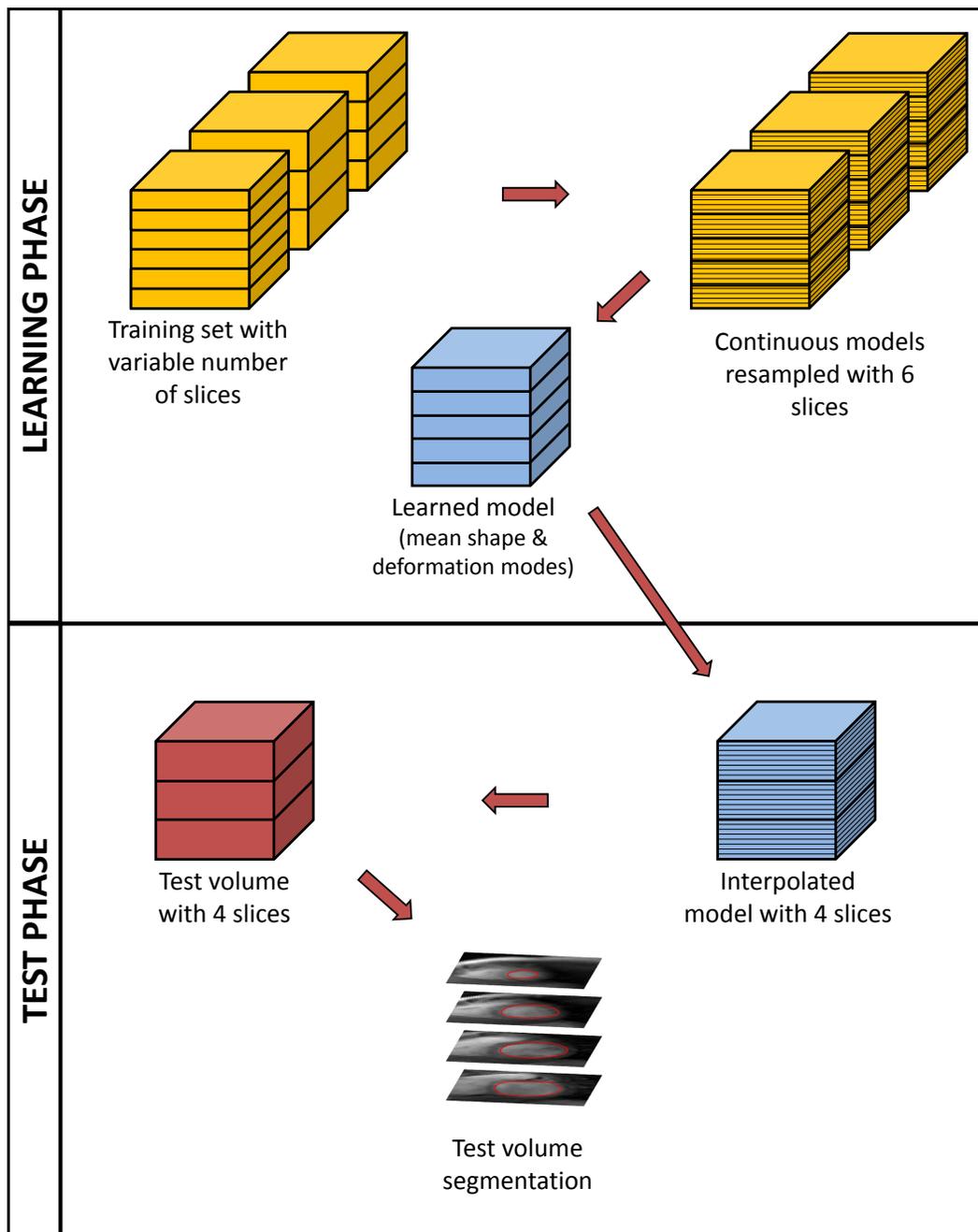


Fig. 1. Learning and test phases of ASM for volumes with a variable number of slices.

As previously mentioned, there is also an additional problem when applying a learned shape model to 3D cardiac MRI: the interpolation of voxel intensities located between two slices may not provide reliable data due to the large distance between slices and to motion artifacts. To overcome this problem, Andreopoulos et al [3] proposed using a preprocessing step that corrects misalignments between consecutive slices using an image registration procedure (translation only). Correcting the misalignments increases the reliability of the interpolated voxel intensities. A different approach was also introduced by Van Assen et al. [12], [18]. They start by building a triangular mesh using the landmarks of the learned model. This mesh has edges that intersect the volume slices. In an initial step, the set of points that intersect the volume slices are the ones used to obtain the necessary displacement vectors that segment the volume. Then, the displacement vectors are propagated to the corresponding landmarks in the shape model and used

to update the model parameters. In this way, they are able to use the available intensity information in the volume without having to interpolated voxel intensities. Nonetheless, some works have used trilinear interpolation to obtain the intensity values at any position within the limits of the volume [11], [19].

In the proposed method, landmarks correspondences are obtained in a way that is similar to the approach used in [3], [11]. However, instead of using a linear interpolation to resample the number of slices, which may create sharp edges in the surface of the LV (due to the large distance between the volume slices), we use a polynomial interpolation of the landmarks' position along the LV axis (see Fig. 2 for an illustration). In addition, our method is different in the test phase. First, we do not interpolate voxel intensities between slices; instead, we resample the shape model so that it fits the test volume. Second, we use a robust estimation method to obtain the shape parameters that segment that volume, based on the EM-RASM algorithm proposed in [20].

### III. INTERPOLATION OF SURFACE MODELS WITH A VARIABLE NUMBER OF SLICES

The PDM approach requires that all surfaces in the training set have the same number of landmarks. As discussed before, in cardiac MRI, this poses a problem because the volumes have a variable number of slices. For a specific volume,  $v$ , the slices are located at equally spaced axial positions

$$s_m = \frac{m - 1}{S^v - 1}, \quad (1)$$

where  $m = 1, \dots, S^v$  and  $S^v$  is the total number of slices in volume  $v$ . We assume that the basal slice is located at  $s_1 = 0$  and that the apical slice is located at  $s_{S^v} = 1$ . The shape model is learned based on medical segmentations, which correspond to the 2D contour of the LV on each slice. This means that the LV surface of volume  $v$  in the training set is defined by a set of  $S^v$  contours located at the axial positions defined by (1).

In order to learn the shape model, the surface models in the training set must have the landmarks at corresponding positions. This can be achieved by: 1) using specific landmarks to define each slice contour; and 2) using a fix number of slices in all volumes.

Regarding the first step, we resample the contours in arc-length at  $N$  points, starting at a specific anatomical landmark. Let  $\mathbf{x}_v(s_m) \in \mathbb{R}^{2N \times 1}$  be the left ventricle contour on the  $m$ -th slice of volume  $v$

$$\mathbf{x}_v(s_m) = \begin{bmatrix} \mathbf{x}^1(s_m) \\ \mathbf{x}^2(s_m) \\ \vdots \\ \mathbf{x}^N(s_m) \end{bmatrix} = \begin{bmatrix} x_1^1(s_m) \\ x_2^1(s_m) \\ x_1^2(s_m) \\ x_2^2(s_m) \\ \vdots \\ x_1^N(s_m) \\ x_2^N(s_m) \end{bmatrix}, \quad (2)$$

where  $\mathbf{x}^i(s_m) = [x_1^i(s_m), x_2^i(s_m)]^\top \in \mathbb{R}^{2 \times 1}$  is the position of the  $i$ -th point. This guarantees that there is a correspondence between the  $i$ -th point of one contour and the  $i$ -th point of another contour, i.e., they represent the same landmark. Concerning the second step, we use an interpolated/approximate model of the landmark positions along the LV axis, defined by  $\hat{\mathbf{x}}_v(s)$ . We wish to model the slice contour as a function of the axial position  $s \in [0, 1]$ . This is done using a combination of  $K$  polynomial basis functions,  $\boldsymbol{\psi}(s) \in \mathbb{R}^{K \times 1}$ ,

$$\hat{\mathbf{x}}_v(s) = \mathbf{C}_v \boldsymbol{\psi}(s), \quad (3)$$

where  $C_v \in \mathbb{R}^{2N \times K}$  is the coefficient matrix associated to volume  $v$ , defined by

$$C_v = \begin{bmatrix} c_1^1 \\ c_2^1 \\ c_1^2 \\ c_2^2 \\ \vdots \\ c_1^N \\ c_2^N \end{bmatrix}, \quad (4)$$

and where the line vector,  $c_j^i \in \mathbb{R}^{1 \times K}$ , contains the  $K$  coefficients associated to the  $j$ -th coordinate of the  $i$ -th contour point. This coefficient matrix is specific of volume  $v$ , *i.e.*, each surface is interpolated using a different coefficient matrix. On the other hand, the polynomial basis,  $\psi(s) = [1, s, \dots, s^{K-1}]^\top$ , depend only on the slice position,  $s$ .

This representation provides an estimate of the LV contour for any position  $s \in [0, 1]$  along the LV axis. Ultimately, this means that we are able to redefine the surface of any volume using a predefined number of slices, as shown in Fig. 2. This approach is used to resample the surface models in the training set. However, the coefficient matrix,  $C_v$ , associated to the surface in volume,  $v$ , has to be estimated from the corresponding annotations. This problem is addressed in the following section.

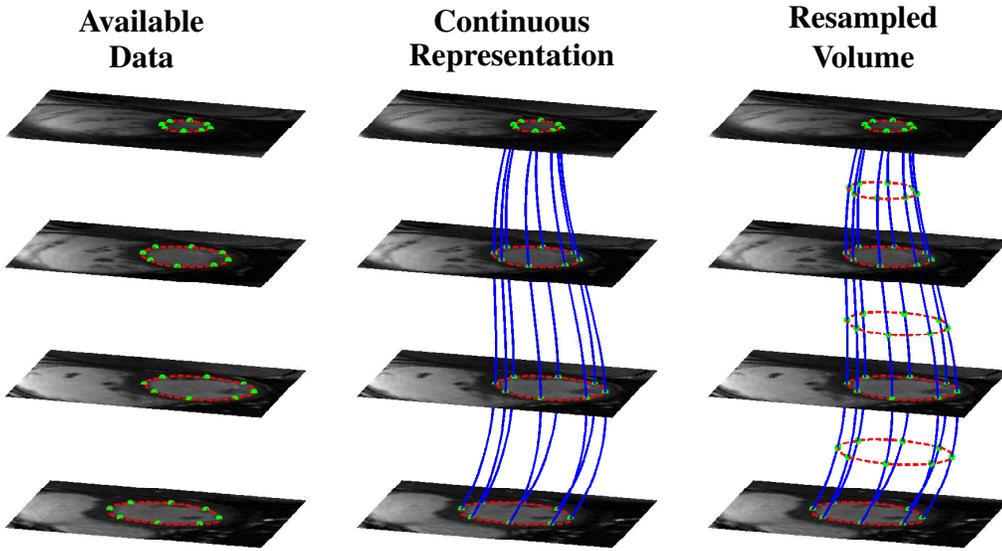


Fig. 2. Illustration of the resampling process. In this example, the available data (left) consists of an MR volume with 4 slices and respective contours in red. The blue lines show the interpolated model for a subset of the contour points (green dots), obtained using the 4 contours (middle), which allows us to obtain the contour at any axial location between the basal and apical slices (right).

#### IV. LEARNING THE SHAPE MODEL

We use the interpolation method described in the previous section to resample the surfaces of the training set. This section explains the computation of the coefficient matrix,  $C_v$ , for each volume  $v$  in the training set and the computation of the shape model.

### A. Resampling the surface models in the training set

The surface of the LV of a particular volume  $v$  is represented by a matrix  $\mathbf{X}_v \in \mathbb{R}^{2N \times S^v}$ , given by the concatenation of the slice contours,

$$\mathbf{X}_v = [\mathbf{x}_v(s_1), \mathbf{x}_v(s_2), \dots, \mathbf{x}_v(s_{S^v})]. \quad (5)$$

Each pair of lines in  $\mathbf{X}_v$ , denoted by  $\mathbf{X}_v^i \in \mathbb{R}^{2 \times S^v}$ , can be regarded as samples of the trajectory of the  $i$ -th contour point as a function of the slice position,  $s_m$  (see the green dots along each blue line in Fig. 2). Specifically, the trajectory samples are given by

$$\mathbf{X}_v^i = \begin{bmatrix} x_1^i(s_1), \dots, x_1^i(s_{S^v}) \\ x_2^i(s_1), \dots, x_2^i(s_{S^v}) \end{bmatrix} = \begin{bmatrix} \mathbf{X}_{1v}^i \\ \mathbf{X}_{2v}^i \end{bmatrix}, \quad (6)$$

where,  $\mathbf{X}_{jv}^i$ ,  $j = 1, 2$ , corresponds to a coefficient line vector  $\mathbf{c}_j^i \in \mathbb{R}^{1 \times K}$ , which is a line from matrix  $\mathbf{C}_v$  (recall (4)).

The trajectory sample points are used to estimate  $\mathbf{c}_j^i$  by

$$\mathbf{c}_j^i = \arg \min_{\mathbf{c}} \|\mathbf{X}_{jv}^i{}^\top - \Psi \mathbf{c}^\top\|^2 + \gamma \|\mathbf{c}\|^2, \quad (7)$$

where  $\Psi = [\boldsymbol{\psi}(s_1), \dots, \boldsymbol{\psi}(s_{S^v})]^\top \in \mathbb{R}^{S^v \times K}$  is the concatenation of the polynomial basis  $\boldsymbol{\psi}(s_m)$  for  $m = 1 \dots, S^v$ , and  $\gamma$  is a regularization constant. This is a ridge regression problem [21] that can be solved by

$$\mathbf{c}_j^i = \mathbf{X}_{jv}^i \Psi (\Psi^\top \Psi + \gamma \mathbf{I})^{-1}, \quad (8)$$

where  $\mathbf{I}$  is the  $K \times K$  identity matrix. Setting  $\gamma = 0$  would lead to the Ordinary Least Squares (OLS) solution. A regularization term is used ( $\gamma > 0$ ) because the OLS solution can only be computed for  $K \leq S^v$ , which means it requires at least the same number of sample points,  $S^v$ , as the number of basis functions,  $K$ . Since our goal is to use a sufficiently large  $K$  ( $K = 6$  was the value used in our tests) and to use the same  $K$  for all volumes, the OLS would not be suitable. The solution (8) can be computed for all the lines in  $\mathbf{C}_v$ , leading to

$$\mathbf{C}_v = \mathbf{X}_v \Psi (\Psi^\top \Psi + \gamma \mathbf{I})^{-1}, \quad (9)$$

Now, the contour,  $\widehat{\mathbf{x}}_v(s)$ , can be obtained for any position  $s \in [0, 1]$  using (3).

This approach is used to resample all the surface models included in the training set at  $s_m = \frac{m-1}{S^r-1}$ ,  $m = 1, \dots, S^r$ , where  $S^r$  is the desired number of slices. This guarantees that all volumes have the same number of landmarks.

### B. Learning the shape statistics

Once all the surface models in the training volumes have been resampled, then it is possible to learn a shape model. We assume a surface model results from deforming the mean shape and applying a transformation associated to the pose of the LV [1]. Therefore, in order to compute the shape statistics, all the surface models have to be aligned. This is done by finding, for each surface, a global (pose) transformation  $\mathbf{T}_\theta$  that minimizes the following sum of squared errors

$$E(\boldsymbol{\theta}) = \sum_{m=1}^{S^r} \sum_{i=1}^N \|\mathbf{T}_\theta(\widehat{\mathbf{x}}^i(s_m)) - \mathbf{x}_{\text{ref}}^i(s_m)\|^2, \quad (10)$$

where  $\mathbf{x}_{\text{ref}}$  is a reference shape (for instance, one of the training shapes randomly selected), and  $\mathbf{T}_\theta(\cdot)$  is a 2D similarity transformation with parameters  $\boldsymbol{\theta} = \{\mathbf{a}, \mathbf{t}\}$ , applied to all slices, such that

$$\mathbf{T}_\theta(\widehat{\mathbf{x}}^i(s_m)) = \widehat{\mathbf{X}}^i(s_m) \mathbf{a} + \mathbf{t}, \quad (11)$$

where

$$\widehat{\mathbf{X}}^i(s_m) = \begin{bmatrix} \widehat{x}_1^i(s_m) & -\widehat{x}_2^i(s_m) \\ \widehat{x}_2^i(s_m) & \widehat{x}_1^i(s_m) \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \quad \mathbf{t} = \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}.$$

We are only interested in the translation, rotation and scaling within the axial (slice) plane to guarantee that the slice contours remain orthogonal to the LV axis. The minimization of (10) leads to a standard least squares solution similar to the alignment algorithm presented in [2].

After the training surfaces have been aligned, the mean shape of each slice,  $\bar{\mathbf{x}}(s_m)$ , is computed as the average slice contour over all the volumes in the training set. The main modes of deformation,  $\mathbf{D}(s_m) = [\mathbf{d}_1(s_m), \dots, \mathbf{d}_L(s_m)] \in \mathbb{R}^{2N \times L}$ , and the corresponding eigenvalues,  $\boldsymbol{\lambda}(s_m) = [\lambda_1(s_m), \dots, \lambda_L(s_m)]^\top \in \mathbb{R}^{L \times 1}$ , are obtained by applying Principal Component Analysis (PCA) [1], where  $\mathbf{d}_l(s_m) \in \mathbb{R}^{2N \times 1}$  and  $\lambda_l(s_m) \in \mathbb{R}$  are the  $l$ -th main mode of deformation at the  $m$ -th slice and corresponding eigenvalue, respectively, and  $L \leq 2N$  is the number of main deformation modes that are used.

## V. ASM FOR 3D DATA

After the training phase, the learned shape model can then be used to segment a new MRI volume - the test phase. However, the number of slices in the new volume, which we denote as  $S^t$ , may not be the same as the learned shape model,  $S^r$ . In case  $S^t \neq S^r$ , one possible approach would be to interpolate the test volume to determine the intensity values at the axial positions of the shape model contours. This would require computing interpolated images. However, the spatial resolution of MRI between axial slices is very low, i.e., the distance between two consecutive slices is very large, and significantly larger than the distance between two consecutive pixels in a slice. Typical values for the distance between slices is 10mm, whereas the distance between two pixels in a slice is approximately 1mm. Furthermore, motion artifacts can cause significant displacement between the location of the LV contour in consecutive slices. Therefore, interpolating images often leads to the loss of contrast between the blood pool and the myocardium, which determines the location of the LV boundary. The images in Fig. 3 show the result of computing an interpolated image on a slice located between two consecutive slices, using trilinear interpolation. The edges in the new image are blurred and, therefore, it is difficult to accurately determine the location of the LV contour. This means that this approach is a bad choice for 3D segmentation of cardiac MRI.

We use a different approach that consists in resampling the learned shape model, *i.e.*, the mean shape and the main modes of deformation. This guarantees that the shape model contours are located at the same axial positions as the volume slices, and thus voxel interpolation is no longer required. The following sections address: *A.* the interpolation of the learned shape model, and *B.* the estimation of the model parameters that segment the LV in a test volume.

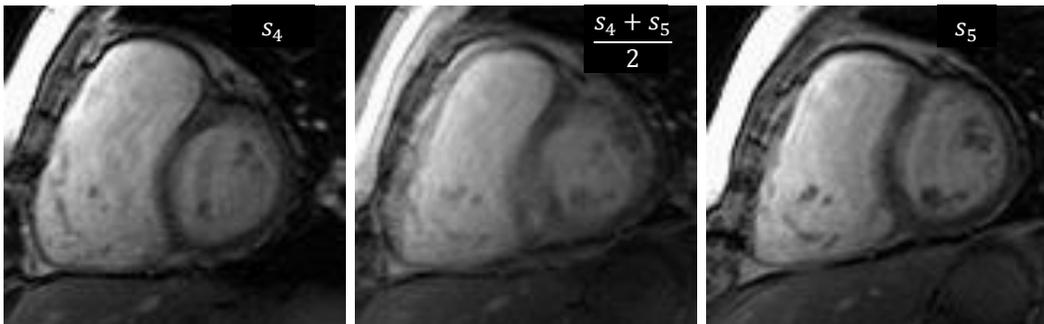


Fig. 3. Slice interpolation. Example of an interpolated image (middle), located at  $s = \frac{s_4 + s_5}{2}$ , obtained by linear interpolation between slices  $s_4$  and  $s_5$ .

### A. Interpolation of shape statistics

In the learning phase, we obtained the mean shape,  $\bar{\mathbf{x}}(s_m)$ , and the main modes of deformation,  $\mathbf{D}(s_m)$ , and their corresponding eigenvalues,  $\lambda(s_m)$ . These shape statistics were computed for the axial positions  $s_m = \frac{m-1}{S^r-1}$ ,  $m = 1, \dots, S^r$ . Now, given a test volume with  $S^t$  slices, we wish to obtain the shape statistics for new slice positions  $s_m = \frac{m-1}{S^t-1}$ ,  $m = 1, \dots, S^t$ , where  $S^t \neq S^r$ .

The mean shape in the new slice positions is computed using the same strategy described in the previous sections, i.e., by computing the corresponding coefficient matrix,  $\bar{\mathbf{C}}$ , and resampling at the new slice positions. Formally, let  $\bar{\mathbf{X}} = [\bar{\mathbf{x}}(0), \dots, \bar{\mathbf{x}}(1)] \in \mathbb{R}^{2N \times S^t}$  be the concatenation of all the  $S^t$  slices of the mean shape. The corresponding coefficient matrix,  $\bar{\mathbf{C}}$ , is computed using (9), where the trajectory samples are now given by  $\bar{\mathbf{X}}$ . Then, the mean shape is resampled at  $S^t$  slices using (3).

On the other hand, resampling the main modes of deformation at intermediate slices is not straightforward, since we need to match deformation modes of different slices. In fact, the modes of deformation are sorted according to the value of the corresponding eigenvalues. Since eigenvalues are learned independently for each slice, it is not easy to find corresponding deformation modes in different slices. In this chapter, we adopt a simple approach that consists of finding the nearest correspondence between deformation modes in consecutive slices and use them to perform a linear interpolation.

Consider a slice position,  $s \in [s_m, s_{m+1}]$ , located between slices  $s_m$  and  $s_{m+1}$ . The deformation modes at this slice,  $\mathbf{D}(s) = [\mathbf{d}_1(s), \dots, \mathbf{d}_L(s)]$ , are determined using linear interpolation between corresponding deformation modes in  $s_m$  and  $s_{m+1}$ . Let  $\alpha \in [0, 1]$  be the relative distance of slice  $s \in [s_m, s_{m+1}]$  to  $s_m$ ,

$$\alpha = \frac{s - s_m}{s_{m+1} - s_m}. \quad (12)$$

Without loss of generality, we assume that  $s_m$  is the closest slice (i.e.,  $\alpha \leq 0.5$ ). The  $l$ -th deformation mode and corresponding eigenvalue are given by

$$\mathbf{d}_l(s) = (1 - \alpha)\mathbf{d}_l(s_m) + \alpha\mathbf{d}_{F(l)}(s_{m+1}) \quad (13)$$

$$\lambda_l(s) = (1 - \alpha)\lambda_l(s_m) + \alpha\lambda_{F(l)}(s_{m+1}), \quad (14)$$

where  $F(\cdot)$  defines the correspondence between the deformation modes in  $s_m$  and  $s_{m+1}$ ,

$$F(l) = \arg \min_n \|\mathbf{d}_l(s_m) - \mathbf{d}_n(s_{m+1})\|. \quad (15)$$

This interpolation process is repeated for all the deformation modes at all the required slices, i.e., for  $l = 1, \dots, L$  and for  $s = \frac{m-1}{S^t-1}$ , with  $m = 1, \dots, S^t$ .

Once all the deformation modes and eigenvalues have been computed, we define the LV surface as

$$\mathbf{x}(s) = \mathbf{T}_\theta(\bar{\mathbf{x}}(s) + \mathbf{D}(s)\mathbf{b}(s)), \quad (16)$$

where  $\mathbf{b}(s)$  are the deformation coefficients. This means that the segmentation of the test volume is obtained by finding the parameters for the pose transformation,  $\theta = \{\mathbf{a}, \mathbf{t}\}$ , and the deformation coefficients,  $\mathbf{b}(s)$ .

The following section describes the estimation of the pose and shape parameters for a new test volume using a robust estimation method.

### B. Automatic surface estimation

Given the test volume, the segmentation of the LV is obtained by estimating the pose, defined by a similarity transformation  $\mathbf{T}_\theta$  with parameters  $\theta = \{\mathbf{a}, \mathbf{t}\}$ , and deformation coefficients of the shape model,  $\mathbf{b}(s)$ . However, automatically obtaining these parameters is difficult due to the presence of other structures in the images, such as the epicardium, papillary muscles and trabeculations [22], that should be considered as noise or outliers.

In this work, the automatic segmentation of the LV is achieved by using the EM-RASM estimation method [20], which is robust in the presence of outliers. An overview of the estimation method is shown in Fig. 4.

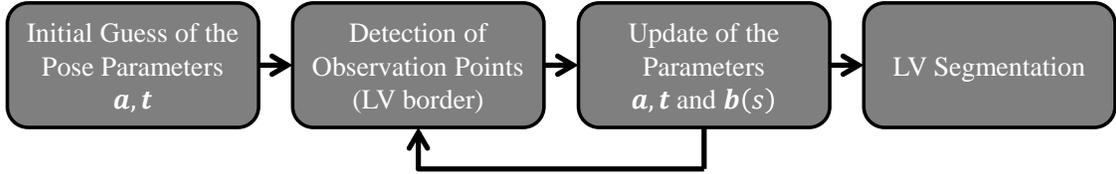


Fig. 4. Block diagram of the automatic segmentation process using the EM-RASM algorithm [20]. The observation points include outliers that should be rejected by the update block.

First, an initial guess of the pose parameters is required - a rough location of the LV center in the basal slice (first block in Fig. 4). We assume that the initial values for the deformation coefficients are  $\mathbf{b}(s) = \mathbf{0}$ , *i.e.*, that the mean shape is a good initialization. With these parameters, we can determine the location of the slice contours. Then, observation points, corresponding to the LV border, are detected in the vicinity of the model (second block in Fig. 4). These observation points are searched in each slice, along lines orthogonal to the contour model, as shown in Fig. 5. The LV border is detected along the search lines by applying an edge detector (see [23] Section 5.2 for details). This approach often leads to the detection of outliers, *i.e.*, observation points that do not belong to the LV border (see Fig. 5). These outliers should not be taken into account in the update of the pose and deformation parameters. The EM-RASM is able to handle the outliers by assuming that each observation point may be either an outlier or a valid point. It assigns a weight to each observation point proportional to the probability that the point belongs to the LV border. The weights determine their influence in the estimation of the model parameters,  $\theta$  and  $\mathbf{b}(s)$ . Since outliers typically get lower weights, their influence in the estimation procedure is reduced and the results are more robust. The final update equations correspond to the weighted least squares solution to the problem of minimizing the distance between each observation point and the corresponding model point (see [20] for further details), computed over all the slice contours simultaneously (third block in Fig. 4).

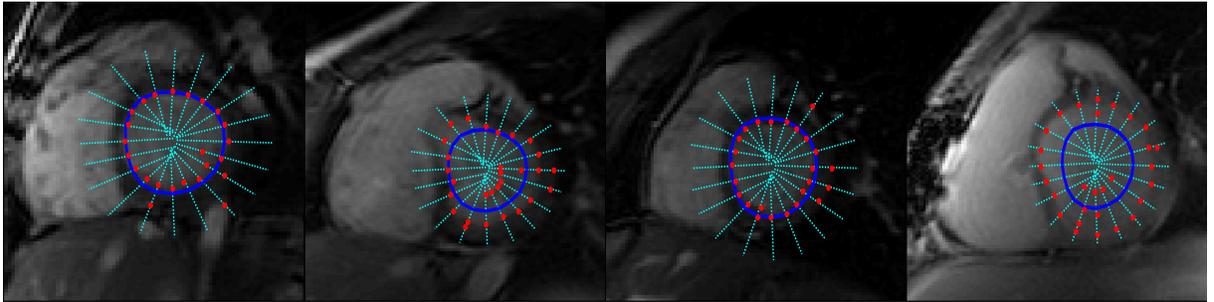


Fig. 5. Detection of observation points. In each example, the slice contour is shown in blue, the search lines in dashed cyan, and the red dots represent the detected observation points.

Once the parameters have been updated, the slice contours are updated and new observation points are extracted from the volume. This process iterates until no significant changes in the parameters occur. The final position of the slice contours determine the segmentation of the LV in the MRI volume (fourth block in Fig. 4).

## VI. EXPERIMENTAL SETUP

The proposed method was evaluated on a set of 20 volumes extracted from a publicly available dataset provided by Andreopoulos and Tsotsos [3]. These volumes consist of end-diastolic short axis cardiac MR volumes, acquired from 20 different subjects at the Department of Diagnostic Imaging of the Hospital for Sick Children in Toronto, Canada. They were acquired using the FIESTA scan protocol and a GE Genesis Signa MR scanner. The age of the subjects ranged between 8 and 15 and they displayed not only healthy

hearts but also heart abnormalities, such as enlarged right ventricles and ischemia. The image slices were acquired with  $256 \times 256$  pixels, with a resolution of  $0.93 - 1.64$  mm. The number of slices in the volumes ranged from 5 to 10 (recall that we are only interested in the slices depicting the endocardial border of the LV) and the spacing between consecutive slices range from 6 to 13 mm. The dataset also provided the endocardial contour of the LV, which was considered as ground truth (GT). The segmentations obtained using the proposed method were evaluated by comparison with the GT.

The segmentation obtained with the proposed approach was quantitatively evaluated using two metrics: 1) the average Dice similarity coefficient [24], and 2) the average minimum distance between the surface points and the GT. These metrics were computed as follows. Let  $z(s) \in \mathbb{R}^{2M \times 1}$  be the GT contour at slice  $s$ , defined by  $M$  points (the GT and the contour model may have a different number of points). Also, let  $R_{x(s)}$  and  $R_{z(s)}$  be the regions delimited by the obtained slice contour,  $x(s)$ , and by the corresponding GT contour,  $z(s)$ , respectively. The average Dice similarity coefficient is given by

$$d_{\text{Dice}} = \frac{1}{S^t} \sum_{m=1}^{S^t} 2 \frac{A(R_{x(s_m)} \cap R_{z(s_m)})}{A(R_{x(s_m)}) + A(R_{z(s_m)})}, \quad (17)$$

where  $A(\cdot)$  denotes the area of a region. The average minimum distance is given by

$$d_{\text{AV}} = \frac{1}{NS^t} \sum_{m=1}^{S^t} \sum_{i=1}^N \min_j \|\mathbf{x}^i(s_m) - \mathbf{z}^j(s_m)\|, \quad (18)$$

and was measured in mm.

The results were obtained using a leave-one-out scheme, where the shape model was trained using 19 volumes and then the model was applied to segment the remaining volume. This process was repeated for each test volume. In all the tests, the contours in the training surface models were resampled in arc-length to have  $N = 40$  points, and the surface model was resampled to have  $S^r = 8$  slices (regardless of the number of slices of the test volume). This means that the shape model was learned using a total number of points of  $N \times S^r = 320$ . Then, the shape model was resampled to have the same number of slices as the test volume,  $S^t$ , which means the total number of points in the test phase was  $N \times S^t$  (it depends on the test volume).

The interpolation models were computed using the parameters  $K = 6$  and  $\gamma = 10^{-4}$ . These values were empirically chosen by comparing the interpolated model with the corresponding training surface models for different values of  $K$  and  $\gamma$ . Fig. 6 (left) shows the error obtained for different values of  $K$  (using  $\gamma = 10^{-4}$ ), and Fig. 6 (right) shows the error for different values of  $\gamma$  (using  $K = 6$ ). It is possible to see that for values of  $K > 6$ , the average error does not justify the increase in computational complexity of using larger values of  $K$ . Regarding the regularization parameter  $\gamma$ , it is concluded that, for values of  $\gamma < 10^{-4}$ , the average error does not significantly change. This parameter can be interpreted as a confidence degree of a prior over the coefficients in matrix  $C_v$ . By decreasing the value of  $\gamma$ , the influence of the prior over the estimation of  $C_v$  is reduced, and the estimation is primarily influenced by the observed data (the trajectories  $\mathbf{X}_v$ ). On the other hand, a higher value of  $\gamma$  helps the estimation of  $C_v$  when the number of slices is smaller than  $K$ . For this reason, we chose  $\gamma = 10^{-4}$  for the following tests.

The obtained shape statistics are exemplified in Fig. 7. The figure shows the mean shape and the variation introduced by the two first modes of deformation. It is possible to see that, besides local (2D) deformation, these modes also capture the 3D shape variation, caused by misalignments between consecutive slices. In order to determine the influence of the number of deformation modes,  $L$ , used in the shape model (recall that we use  $L \leq 2N$  modes of the deformation), we compute the error of approximating the contours of the training surfaces by (16), i.e., a linear combination of the mean shape and the main modes of deformation. The results are shown in Fig. 8. The plot shows that using more modes of deformation than  $L = 10$  does not significantly improve the accuracy of the approximation. Again, choosing a larger

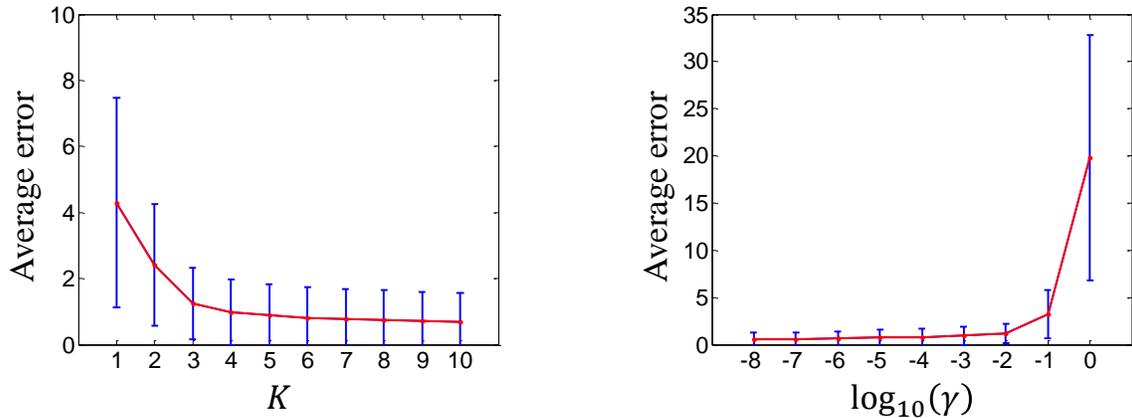


Fig. 6. Average error (in pixels) of the interpolated model for: (left) different number of polynomial basis functions,  $K$  (using  $\gamma = 10^{-4}$ ), and (right) different values of the regularization coefficient,  $\gamma$  (using  $K = 6$ ).

value of  $L$  would only lead to an increase of the computational complexity of the algorithm. Furthermore,  $L = 10$  corresponds to approximately 90% of the variation shown in the training set, which is also a common criterion to select the number of deformation modes used in the shape model [1].

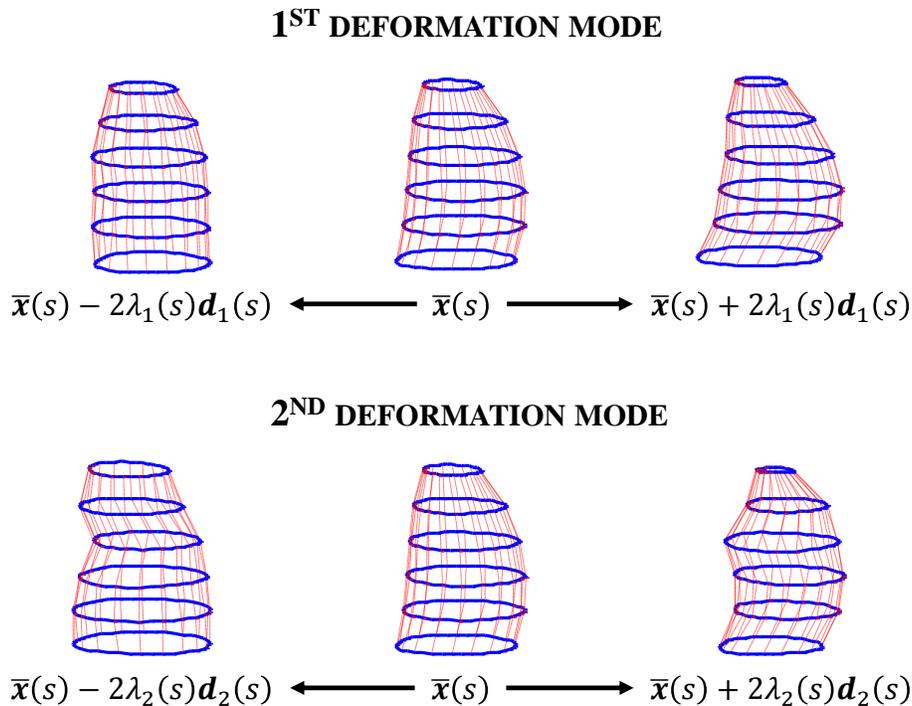


Fig. 7. Shape statistics for the LV. The top row shows the shape variation along the first mode of deformation and the bottom row the same for the second mode of deformation. The shape in the middle column corresponds to the mean shape  $\bar{x}(s)$ . These contours were computed for all the slice positions,  $s = s_1, s_2, \dots, s_6$ .

## VII. RESULTS

This section shows the evaluation of the segmentations obtained using the proposed method. Some examples of the segmentations are shown in Fig. 11. It is possible to see that the obtained segmentations are close to the ground truth. However, in some cases (e.g., the second row), the algorithm is not able to

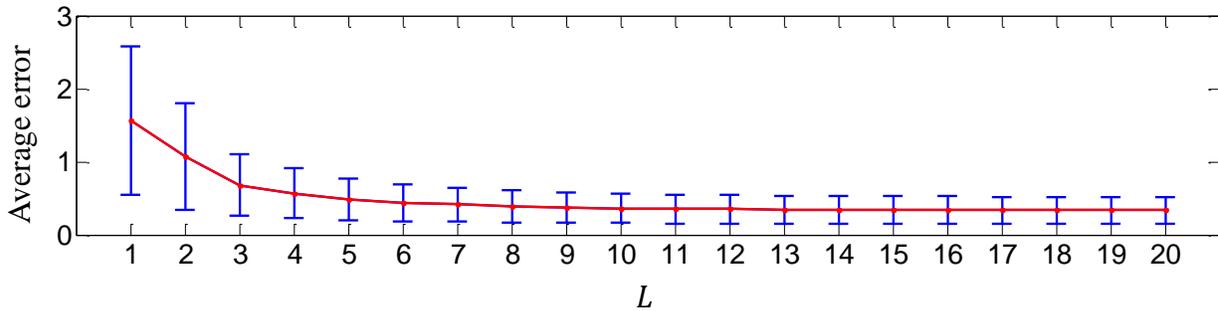


Fig. 8. Average error (in pixels) of the shape model approximation for different values of the number of deformation modes,  $L$ .

accuracy segment both the apical and basal slices. This is because it is not always possible to find a proper pose transformation that fits all slices, particularly in volumes where there is significant misalignment between slices. The statistics for the two metrics are shown in Fig. 9 for each test volume. The overall results achieved in these tests were  $d_{\text{Dice}} = 0.88 \pm 0.06$  and  $d_{\text{AV}} = 1.2 \pm 0.7$  mm.

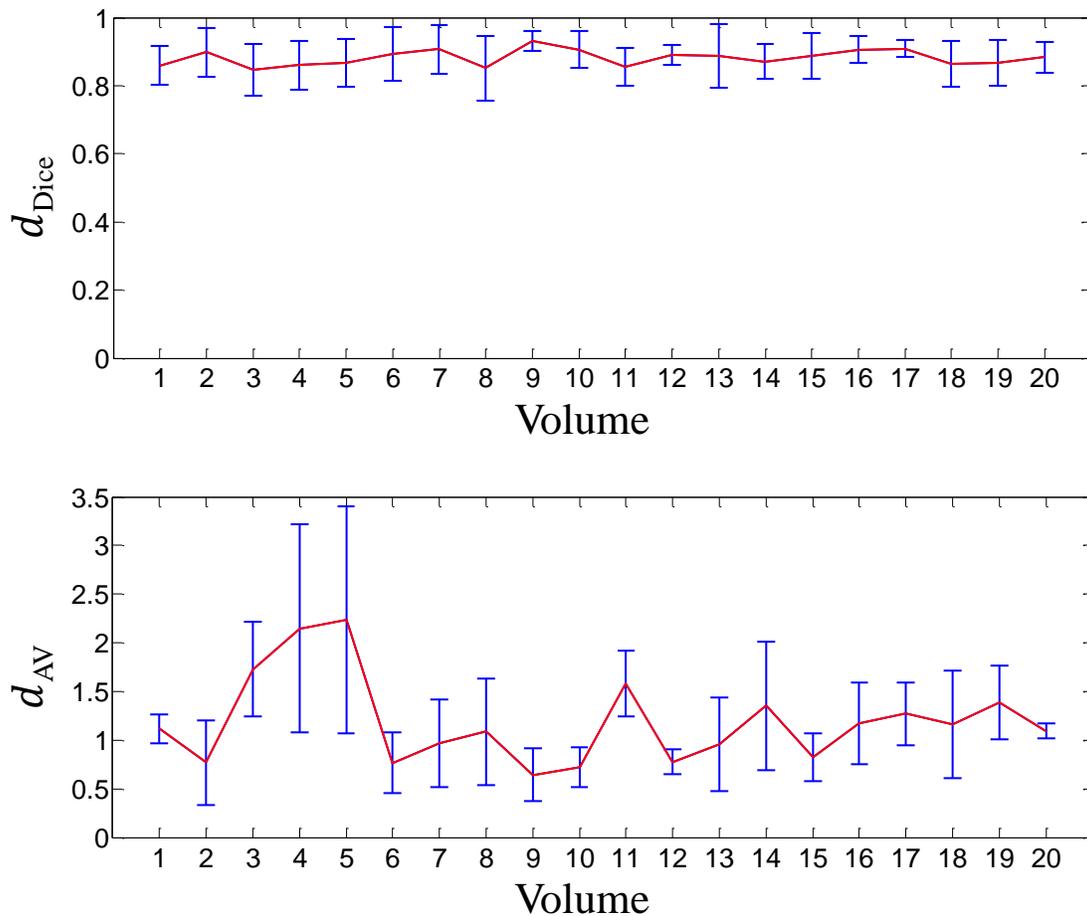


Fig. 9. Statistical results for all the test volumes using a leave-one-out scheme. The top row shows the Dice coefficient and the bottom row shows the average minimum distance metric (in mm).

Although the proposed method is able to accurately segment the LV, the apical slice remains the most difficult part of the volume to segment. This is due to the fact that the LV chamber is very small in these

slices, its borders are often irregular and this makes it difficult to detect the endocardium. This can be verified by the results in Fig. 10, which shows a 3D representation of the LV surface, as well as a color coded representation of each slice contours, where the color depends on the corresponding Dice coefficient (green corresponds to a good segmentation and red corresponds to a poor segmentation). In this figure, it is possible to see that most apical slices have poorer accuracy than the basal slices. Nonetheless, the segmentation accuracy is still high, with Dice coefficients of approximately  $d_{\text{Dice}} \approx 0.7$ .

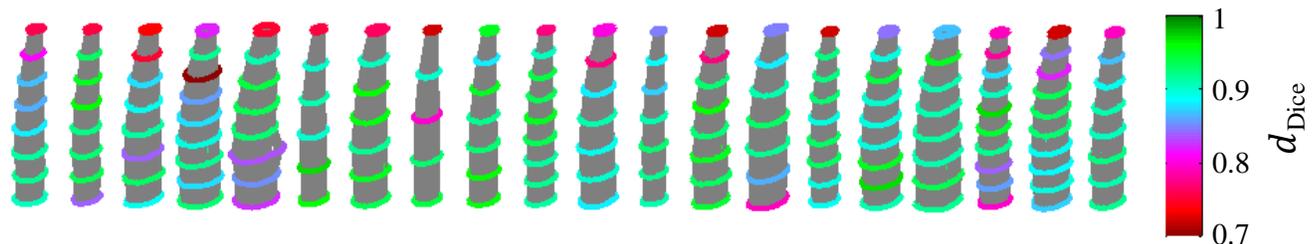


Fig. 10. Segmentations obtained using the proposed method. The color code shows the accuracy of the segmentation in each slice and for each volume (according to the Dice coefficient).

## VIII. CONCLUSION

Many medical image problems use an Active Shape Model (ASM) based approach to include shape constraints into the segmentation process. However, obtaining the 3D segmentation of cardiac MR volumes has additional difficulties due to the variable number of slices.

Learning a 3D shape model based on a training set of surfaces with a variable number of slices is not easy. We propose to deal with this issue by using a continuous (interpolated) representation for the surface. With this representation, we are able to obtain a smooth continuous surface that can be resampled to have a predefined number of slices. By resampling all the surfaces in the training set, we establish a 1-1 correspondence between the landmarks (surface points) of all the training surfaces. Then, the shape model can be easily learned by aligning the surfaces and applying PCA to determine the shape statistics.

On the other hand, in order to use the learned shape model in a test volume, one has to be careful in case the number of slices of the shape model is different from the number of slices of the test volume. In this case, carelessly applying trilinear interpolation to obtain the intensities of locations in between slices may result in the appearance of dubious edges, due to the large distance between consecutive slices and by the misalignment caused by motion artifacts. We address this issue by resampling the learned model so that it has the same number of slices as the test volume, noting that this involves resampling the mean shape as well as the main modes of deformation. Only then we apply the model to the test volume and estimate the parameters that best segment the volume. This means finding the pose and deformation of the LV. We restrict the possible transformations to rotation, scaling and translation in 2D, to ensure the surface points remain within the volume slices. The model parameters are obtained using a robust estimation technique, based on the EM algorithm [20].

The proposed approach was tested using 20 volumes from the Andreopoulos and Tsotsos dataset [3]. The shape model was learned using a leave-one-out scheme and the segmentations were evaluated using the Dice similarity coefficient and the average minimum distance metric. The results show that the proposed method is able to accurately segment the LV. However, further improvements may still be achieved. For instance, when the volume slices are misaligned, finding a proper similarity transformation to fit the learned model to the volume is nearly impossible. In these cases, a preprocessing step may be required to correct the misalignment as proposed in [3]. Alternatively, one could allow the possibility of having minor independent translations associated to each slice, allowing the model to correct the misalignment of the slices.

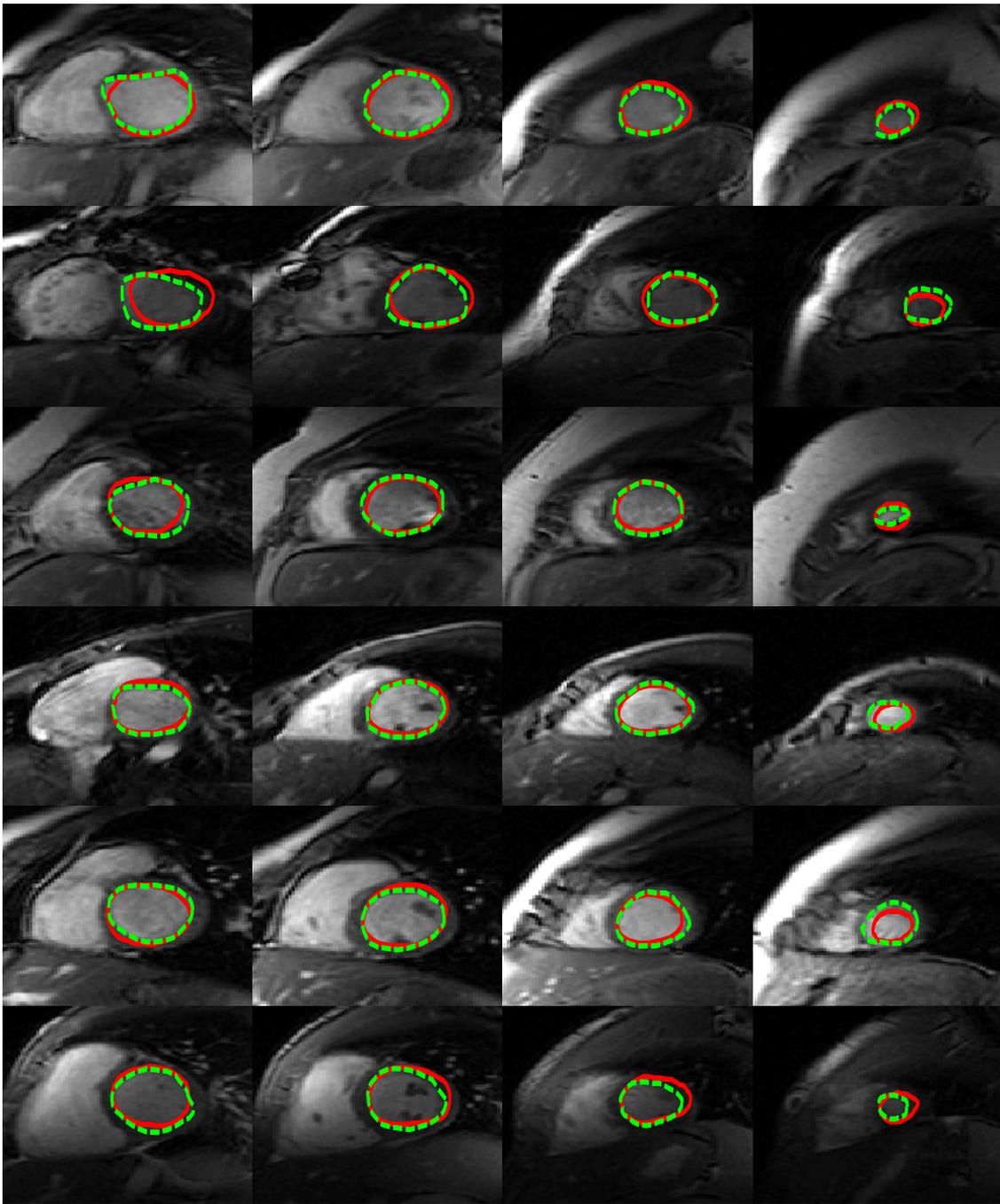


Fig. 11. Examples of the obtained segmentations. Each line shows a different volume and each row a different slice: the left column corresponds to the basal slice and the right column to the apex slice. The red contour is the obtained segmentation and the dashed green is the ground truth.

## REFERENCES

- [1] 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.
- [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.
- [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] A. Cauce and C. J. Taylor, "Building 3d sulcal models using local geometry," *Medical Image Analysis*, vol. 5, no. 1, pp. 69–80, 2001.
- [5] H. Chen and B. Bhanu, "Shape model-based 3d ear detection from side face range images," in *Computer Vision and Pattern Recognition-Workshops, 2005. CVPR Workshops. IEEE Computer Society Conference on*. IEEE, 2005, pp. 122–122.

- [6] D. Fritz, D. Rinck, R. Dillmann, and M. Scheuering, "Segmentation of the left and right cardiac ventricle using a combined bi-temporal statistical model," in *Medical Imaging*. International Society for Optics and Photonics, 2006, pp. 614 121–614 121.
- [7] P. J. Besl and N. D. McKay, "Method for registration of 3-d shapes," in *Robotics-DL tentative*. International Society for Optics and Photonics, 1992, pp. 586–606.
- [8] D. Chetverikov, D. Stepanov, and P. Krsek, "Robust euclidean alignment of 3d point sets: the trimmed iterative closest point algorithm," *Image and Vision Computing*, vol. 23, no. 3, pp. 299–309, 2005.
- [9] B. Jian and B. C. Vemuri, "Robust point set registration using gaussian mixture models," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 33, no. 8, pp. 1633–1645, 2011.
- [10] F. Pomerleau, F. Colas, R. Siegwart, and S. Magnenat, "Comparing icp variants on real-world data sets," *Autonomous Robots*, vol. 34, no. 3, pp. 133–148, 2013.
- [11] S. C. Mitchell, J. G. Bosch, B. P. Lelieveldt, R. J. van der Geest, J. H. Reiber, and M. Sonka, "3-d active appearance models: segmentation of cardiac mr and ultrasound images," *Medical Imaging, IEEE Transactions on*, vol. 21, no. 9, pp. 1167–1178, 2002.
- [12] H. C. Van Assen, M. G. Danilouchkine, A. F. Frangi, S. Ordás, J. J. Westenberg, J. H. Reiber, and B. P. Lelieveldt, "Spasm: a 3d-asm for segmentation of sparse and arbitrarily oriented cardiac mri data," *Medical Image Analysis*, vol. 10, no. 2, pp. 286–303, 2006.
- [13] A. F. Frangi, D. Rueckert, J. A. Schnabel, and W. J. Niessen, "Automatic construction of multiple-object three-dimensional statistical shape models: Application to cardiac modeling," *Medical Imaging, IEEE Transactions on*, vol. 21, no. 9, pp. 1151–1166, 2002.
- [14] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: application to breast mr images," *Medical Imaging, IEEE Transactions on*, vol. 18, no. 8, pp. 712–721, 1999.
- [15] S. Osher and J. A. Sethian, "Fronts propagating with curvature-dependent speed: algorithms based on hamilton-jacobi formulations," *Journal of computational physics*, vol. 79, no. 1, pp. 12–49, 1988.
- [16] A. Tsai, A. Yezzi Jr, 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," *Medical Imaging, IEEE Transactions on*, vol. 22, no. 2, pp. 137–154, 2003.
- [17] A. Tsai, W. Wells, C. Tempany, E. Grimson, and A. Willsky, "Mutual information in coupled multi-shape model for medical image segmentation," *Medical Image Analysis*, vol. 8, no. 4, pp. 429–445, 2004.
- [18] H. C. Van Assen, M. G. Danilouchkine, M. S. Dirksen, J. Reiber, and B. P. Lelieveldt, "A 3-d active shape model driven by fuzzy inference: application to cardiac ct and mr," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 12, no. 5, pp. 595–605, 2008.
- [19] M. R. Kaus, J. v. Berg, J. Weese, W. Niessen, and V. Pekar, "Automated segmentation of the left ventricle in cardiac mri," *Medical Image Analysis*, vol. 8, no. 3, pp. 245–254, 2004.
- [20] C. Santiago, J. C. Nascimento, and J. S. Marques, "A robust active shape model using an expectation-maximization framework," in *Image Processing (ICIP), 2014 21th IEEE International Conference on*. IEEE, 2014, pp. 6076–6080.
- [21] A. E. Hoerl and R. W. Kennard, "Ridge regression: Biased estimation for nonorthogonal problems," *Technometrics*, vol. 12, no. 1, pp. 55–67, 1970.
- [22] 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.
- [23] A. Blake and M. Isard, *Active shape models*. Springer, 1998.
- [24] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.