Robust 3D Active Shape Model for the Segmentation of the Left Ventricle in MRI

Carlos Santiago(✉), Jacinto C. Nascimento, and Jorge S. Marques

Institute for Systems and Robotics, Instituto Superior Técnico, Lisboa, Portugal
carlos.santiago@ist.utl.pt

Abstract. 3D Active shape models use a set of annotated volumes to learn a shape model. The shape model is defined by a fixed number of landmarks at specific locations and takes shape constraints into account in the segmentation process. A relevant problem in which these models can be used is the segmentation of the left ventricle in 3D MRI volumes. In this problem, the annotations correspond to a set of contours that define the LV border at each volume slice. However, each volume has a different number of slices (i.e., a different number of landmarks), which makes model learning difficult. Furthermore, motion artifacts and the large distance between slices make interpolation of voxel intensities a bad choice when applying the learned model to a test volume. These two problems raise the following questions: (1) *how can we learn a shape model from volumes with a variable number of slices?* and (2) *how can we segment a test volume without interpolating voxel intensities between slices?* This paper provides an answer to these questions and proposes a 3D active shape model that can be used to segment the left ventricle in cardiac MRI.

Keywords: Active shape model · 3D segmentation · Cardiac MRI · Interpolation

1 Introduction

Active shape models (ASMs) are commonly used to segment medical images because they lead to robust shape estimates [1]. The ASM approach uses the concept of Point Distribution Model (PDM) to learn the shape statistics from a set of annotated volumes (training set), which defines the surface of the object using a set of labeled landmarks. However, this strategy assumes that all the surface models have the same landmarks.

This assumption is not always true. For example, in the segmentation of the left ventricle (LV), in MRI volumes, the number of slices associated to the LV is subject dependent. This means that the number of landmarks used to define the

C. Santiago—This work was supported by FCT [UID/EEA/5009/2013] and [SFRH/BD/87347/2012].

R. Paredes et al. (Eds.): IbPRIA 2015, LNCS 9117, pp. 283–290, 2015.
DOI: 10.1007/978-3-319-19390-8_32

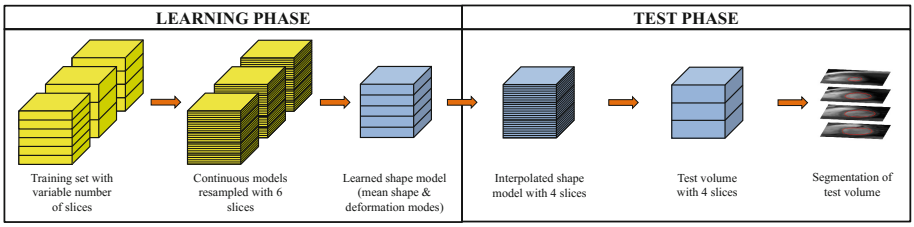

Fig. 1. Diagram of the proposed approach: (1) Learning phase - learning an ASM from volumes with a variable number of slices; and (2) Test phase - applying the learned model to a test volume.

LV surface is different. 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 paper answers this question. Before learning the shape statistics, the surface models in the training set are normalized with respect to the number of slices. The normalization is done by modeling the position of each landmark (surface point) along the LV axis through interpolation. This allows the LV surfaces to be resampled at different positions (i.e., to determine the number of slices in the surface model), which allows the shape model for each slice to be learned. A schematic illustration of the proposed approach is depicted in Fig. 1 (learning phase).

After learning the shape model, we use it to segment new volumes (test phase), which may have a different number of slices. This means that the 2D contours of the learned model may not match the slices of the test volumes. Interpolation is a feasible solution [2–4], but causes the loss of contrast along the LV border (see Sect. 3). This leads to the second question: *how can we segment a test volume without interpolating voxel intensities between slices?*. We propose an alternative approach that consists in interpolating the learned model statistics, i.e., mean shape and modes of deformation.

The remainder of this paper is organized as follows. Section 2 describes the proposed shape representation and how it is used to resample the training surface models. Section 3 explains how the learned model is applied to segment a test volume. Section 4 describes the experimental setup and the results obtained using the proposed approach, and final conclusions are presented in Sect. 5.

2 Learning Phase - Interpolating the Training Surface Models

In order to learn the shape model, we resample the surface models in the training set using an interpolated model of landmarks' position along the LV axis. Under the assumption that, for any training volume v, the first (basal) slice is located at the $s_1 = 0$ and the last (apical) slice at $s_{S^v} = 1$, the axial position of the slices is given by

$$s_m = \frac{m-1}{S^v - 1}, \tag{1}$$

where $m = 1, \ldots, S^v$, and S^v is the number of slices in volume v. Let $\boldsymbol{x}_v(s_m) \in \mathbb{R}^{2N\times 1}$ be the LV contour on the m-th slice, defined by N points,

$$\boldsymbol{x}_v(s_m) = \left[\boldsymbol{x}^{1\top}(s_m), \boldsymbol{x}^{2\top}(s_m), \ldots, \boldsymbol{x}^{N\top}(s_m)\right]^\top, \tag{2}$$

where $\boldsymbol{x}^i(s_m) = \left[x_1^i, x_2^i\right]^\top \in \mathbb{R}^{2\times 1}$ is the position of the i-th point. We assume that all the contours are sampled with N points and 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. We wish to model the slice contour as a function of the axial position, $\widehat{\boldsymbol{x}}_v(s)$, for any $s \in [0, 1]$.

2.1 Interpolated Surface Model

The proposed approach aims to describe the position of the contour points of a specific volume v, $\widehat{\boldsymbol{x}}_v(s)$, along the LV axis by using a combination of K polynomial basis functions, $\boldsymbol{\psi}(s) \in \mathbb{R}^{K\times 1}$,

$$\widehat{\boldsymbol{x}}_v(s) = \boldsymbol{C}_v\boldsymbol{\psi}(s), \tag{3}$$

where $\boldsymbol{C}_v \in \mathbb{R}^{2N\times K}$ is the coefficient matrix associated to volume v. The coefficient matrix is specific of volume v and each line in the matrix is associated to a coordinate of a specific contour point. On the other hand, the polynomial basis, $\boldsymbol{\psi}(s) = \left[1, s, \ldots, s^{K-1}\right]^\top$, depend only on the slice position, s.

This representation provides an estimate of the LV contour for any position $s \in [0, 1]$. This will be used to resample the surface models in the training set. However, first, the coefficient matrix $\boldsymbol{C}_v$ associated to the surface model, v, has to be estimated from the corresponding annotations. This is addressed in the following subsection.

2.2 Resampling the Surface Models in the Training Set

In order to resample a specific surface model of volume v in the training set, the corresponding coefficient matrix $\boldsymbol{C}_v$ has to be computed, based on the available annotations. First, let us denote $\boldsymbol{X}_v^i \in \mathbb{R}^{2\times S^v}$ as the position (trajectory) of the i-th point in the contour along the axial position s,

$$\boldsymbol{X}_v^i = \begin{bmatrix}\boldsymbol{X}_{1v}^i \\ \boldsymbol{X}_{2v}^i\end{bmatrix} = \begin{bmatrix}x_1^i(s_1), \ldots, x_1^i(s_{S^v}) \\ x_2^i(s_1), \ldots, x_2^i(s_{S^v})\end{bmatrix} = \left[\boldsymbol{x}^i(s_1), \ldots, \boldsymbol{x}^i(s_{S^v})\right]. \tag{4}$$

The coefficient $\boldsymbol{c}_j^i \in \mathbb{R}^{1\times K}$, which is the line from matrix $\boldsymbol{C}_v$ associated with the trajectory points $\boldsymbol{X}_{jv}^i$, is computed by finding

$$\boldsymbol{c}_j^i = \arg\min_{\boldsymbol{c}} \|\boldsymbol{X}_{jv}^{i\top} - \boldsymbol{\Psi}\boldsymbol{c}^\top\|^2 + \gamma\|\boldsymbol{c}\|^2, \tag{5}$$

where $\boldsymbol{\Psi} = [\boldsymbol{\psi}(s_1), \ldots, \boldsymbol{\psi}(s_{S^v})]^\top \in \mathbb{R}^{S^v \times K}$ is the concatenation of the polynomial basis $\boldsymbol{\psi}(s_m)$ for $m = 1 \ldots, S^v$, and γ is a regularization constant. This is a ridge regression formulation [5] that has the following solution

$$\boldsymbol{c}_j^i = \boldsymbol{X}_{jv}^i \boldsymbol{\Psi} \left(\boldsymbol{\Psi}^\top \boldsymbol{\Psi} + \gamma \boldsymbol{I} \right)^{-1}, \tag{6}$$

where $\boldsymbol{I}$ is the $K \times K$ identity matrix. This approach differs from the ordinary least squares due to the regularization term, which constrains the solution and allows the estimation of $\boldsymbol{c}_j^i$ not only for $K \leq S^v$, but also when $K > S^v$.

The solution (6) can be computed for all the lines in $\boldsymbol{C}_v$, leading to

$$\boldsymbol{C}_v = \boldsymbol{X}_v \boldsymbol{\Psi} \left(\boldsymbol{\Psi}^\top \boldsymbol{\Psi} + \gamma \boldsymbol{I} \right)^{-1}, \tag{7}$$

where $\boldsymbol{X}_v = [\boldsymbol{x}_v(s_1), \ldots, \boldsymbol{x}_v(s_{S^v})] \in \mathbb{R}^{2N \times S^v}$. Now, the contour, $\widehat{\boldsymbol{x}}_v(s)$, can be obtained for any position $s \in [0, 1]$ using (3).

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

2.3 Learning the Shape Statistics

Once all the surface models in the training set have been resampled, 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. Therefore, before computing the shape statistics, all the surface models have to be aligned. This is done by finding, for each surface, a global (pose) transformation T_θ that minimizes the following sum of squared errors

$$E(\theta) = \sum_{m=1}^{S^r} \sum_{i=1}^{N} \left\| T_\theta \left(\widehat{\boldsymbol{x}}^i \left(\frac{m-1}{S^r - 1} \right) \right) - \boldsymbol{x}_{\text{ref}}^i \left(\frac{m-1}{S^r - 1} \right) \right\|^2, \tag{8}$$

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

$$T_\theta \left(\widehat{\boldsymbol{x}}^i(s) \right) = \widehat{\boldsymbol{X}}^i(s) \boldsymbol{a} + \boldsymbol{t}, \tag{9}$$

where

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

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 (8) leads to a standard least squares solution similar to [6].

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

3 Robust 3D ASM - Segmenting a Test Volume

The shape model learned by the previous methods is used to segment other cardiac MR volumes. As before, the number of slices in a test 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 volume to determine the intensity values at the same axial positions as the shape model contours. However, the spatial resolution of MRI between axial slices is very low (approximately 10 mm spacing between slices), and motion artifacts can cause significant displacements in the location of the LV contour in consecutive slices. Consequently, the location of the LV border may become hard to determine in interpolated images, as shown in Fig. 2

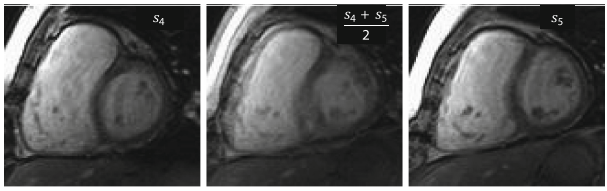

Fig. 2. Example of an interpolated image, at $s = \frac{s_4+s_5}{2}$, obtained by linear interpolation between two consecutive slices, s_4 and s_5.

We propose a different approach that consists in resampling the shape model (mean shape and deformation modes) to have the same number of slices as the test volume. The mean shape can be easily interpolated using the strategy in Sect. 2. We compute the corresponding coefficient matrix, $\overline{\boldsymbol{C}}$, using (7), and resample the mean shape at S^t slices, $s = \frac{m-1}{S^t-1}$, with $m = 1, \ldots, S^t$, using (3).

On the other hand, computing the main modes of deformation for intermediate slices is not straightforward. The reason is that 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 possible to find corresponding deformation modes in different slices. Therefore, we use a simpler approach, that consists in finding the correspondences between deformation modes in consecutive slices and use them to perform a linear interpolation, as follows. 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, $\boldsymbol{D}(s) = [\boldsymbol{d}_1(s), \ldots, \boldsymbol{d}_L(s)]$, 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,

$$\alpha = \frac{s - s_m}{s_{m+1} - s_m}. \tag{10}$$

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

$$\boldsymbol{d}_l(s) = (1 - \alpha)\boldsymbol{d}_l(s_m) + \alpha \boldsymbol{d}_{F(l)}(s_{m+1}) \tag{11}$$

$$\lambda_l(s) = (1 - \alpha)\lambda_l(s_m) + \alpha \lambda_{F(l)}(s_{m+1}), \tag{12}$$

where $F(\cdot)$ maps the correspondences between the deformation modes in s_m to s_{m+1},

$$F(l) = \arg\min_n \left\| \boldsymbol{d}_l(s_m) - \boldsymbol{d}_n(s_{m+1}) \right\|. \tag{13}$$

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

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

$$\boldsymbol{x}(s) = T_\theta \left(\overline{\boldsymbol{x}}(s) + \boldsymbol{D}(s)\boldsymbol{b}(s) \right). \tag{14}$$

This means that the segmentation of the test volume is obtained by finding the parameters for the pose transformation, $\theta = \{\boldsymbol{a}, \boldsymbol{t}\}$, and the deformation coefficients, $\boldsymbol{b}(s)$. In this work, this is achieved using the robust estimation method called EM-RASM, which is able to compute the shape model parameters in the presence of outliers. See [7] for an in-depth description of this methodology.

4 Results

The proposed method was evaluated on a set of 20 volumes extracted from the publicly available dataset provided by Andreopoulos and Tsotsos [4]. This database provides the endocardial contour of the LV that will be considered as ground truth.

The results were obtained using a leave-one-out scheme, where the shape model was trained using 19 volumes and then applied to the remaining (test) volume. In all the tests, each slice contour was resampled, in arc-length, at $N = 40$ points, and the surface models were resampled at $S^r = 8$ slices, using $K = 6$ and $\gamma = 10^{-5}$. Therefore, the total number of points in the surface models was $N \times S^r = 320$. In the test phase, the total number of points in the resampled shape model was $N \times S^t$ (it depended on the test volume). The segmentation was quantitatively evaluated using the average Dice similarity coefficient [8], d_{Dice}, and the average minimum distance between the surface model points and the ground truth, d_{AV}, measured in mm.

Some examples of the segmentations are shown in Fig. 3. It is possible to see that the obtained segmentations are similar to the ground truth, although the

algorithm performs better in the basal slices than in the apical slices. This is due to the fact that the LV chamber is very small in apical slices, and its borders are often irregular. The overall results, according to the Dice coefficient and the average minimum distance, were $d_{\text{Dice}} = 0.88 \pm 0.07$ and $d_{\text{AV}} = 1.3 \pm 0.7$ mm.

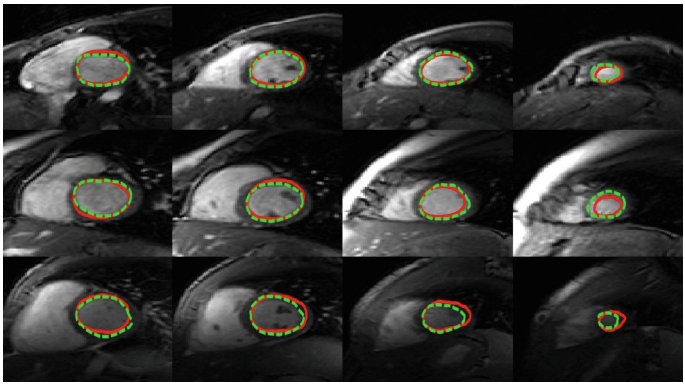

Fig. 3. Examples of the obtained segmentations. Each line shows a different volume and each row a different slice, starting at the basal slice (left) and ending at the apex (right). The red and dashed green lines are the output of the proposed algorithm and the ground truth, respectively (Color figure online).

5 Conclusion

This paper proposes a 3D Active Shape Model (ASM) for the segmentation of the left ventricle in cardiac MRI. Although ASMs based approaches are commonly used to segment medical images, they cannot be directly used in cardiac MRI the volumes have the variable number of slices.

We propose to deal with this issue by using a continuous representation for the surface model, which allows the surface model to be resampled to a predefined number of slices. By resampling all the surface models in the training set, we establish a correspondence for the landmarks (surface points) between all the surface models. Furthermore, the same problem arises in the test phase, because the learned model may have a different number of slices. The proposed approach interpolates the learned model, i.e., the mean shape and the main modes of deformation, to avoid interpolating intensity values between the volume slices.

The results show that the proposed method is able to accurately segment the LV.

References

1. Heimann, T., Meinzer, H.: Statistical shape models for 3d medical image segmentation: a review. Med. Image Anal. **13**(4), 543–563 (2009)

2. Mitchell, S.C., Bosch, J.G., Lelieveldt, B.P., van der Geest, R.J., Reiber, J.H., Sonka, M.: 3-d active appearance models: segmentation of cardiac mr and ultrasound images. IEEE Trans. Med. Imaging **21**(9), 1167–1178 (2002)
3. Kaus, M.R., von Berg, J., Weese, J., Niessen, W., Pekar, V.: Automated segmentation of the left ventricle in cardiac mri. Med. Image Anal. **8**(3), 245–254 (2004)
4. Andreopoulos, A., Tsotsos, J.K.: Efficient and generalizable statistical models of shape and appearance for analysis of cardiac mri. Med. Image Anal. **12**(3), 335–357 (2008)
5. Hoerl, A.E., Kennard, R.W.: Ridge regression: biased estimation for nonorthogonal problems. Technometrics **12**(1), 55–67 (1970)
6. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models-their training and application. Comput. Vis. Image Unders. **61**(1), 38–59 (1995)
7. Santiago, C., Nascimento, J.C., Marques, J.S.: A robust active shape model using an expectation-maximization framework, In: 2014 21th IEEE International Conference on Image Processing (ICIP), pp. 6076–6080. IEEE (2014)
8. Dice, L.R.: Measures of the amount of ecologic association between species. Ecology **26**(3), 297–302 (1945)