

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/303179696>

A new ASM framework for left ventricle segmentation exploring slice variability in cardiac MRI volumes

Article in *Neural Computing and Applications* · May 2016

DOI: 10.1007/s00521-016-2337-1

CITATIONS

9

READS

70

3 authors, including:



Carlos Santiago

Technical University of Lisbon

22 PUBLICATIONS 88 CITATIONS

[SEE PROFILE](#)



Jorge S Marques

University of Lisbon

268 PUBLICATIONS 3,324 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Automated detection and characterization of impact craters [View project](#)



Medical Imaging Multimodality Breast Cancer Diagnosis User Interface [View project](#)

A new ASM framework for left ventricle segmentation exploring slice variability in cardiac MRI volumes

Carlos Santiago¹ · Jacinto C. Nascimento¹ · Jorge S. Marques¹

Received: 13 October 2015 / Accepted: 3 May 2016
© The Natural Computing Applications Forum 2016

Abstract Three-dimensional active shape models use a set of annotated volumes to learn a shape model. Using unique landmarks to define the surface models in the training set, the shape model is able to learn the expected shape and variation modes of the segmentation. This information is then used during the segmentation process to impose shape constraints. A relevant problem in which these models are 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 (thus, 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 by proposing a framework to deal with the variable number of slices in the training set and a resampling strategy for the test phase to segment the left ventricle in cardiac MRI volumes with any number of slices. The proposed method was evaluated on a public database with 660 volumes of both healthy and diseased patients, with promising results.

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

1 Introduction

The 3D segmentation of the left ventricle (LV) in cardiac MRI is part of the routine procedure of evaluating cardiac function [20]. This task is typically performed by cardiologists by manually tracing the LV border in each slice of the MRI volume, at two different time instances: the end-systolic and the end-diastolic phases. This procedure consumes valuable clinical time, which is why there has been a lot of research in automatizing the LV segmentation. However, several problems associated with cardiac MRI image modality hamper the performance of these methods, such as distinguishing the endocardium (inner border of the LV) from papillary muscles and trabeculations [35].

A popular approach is to use shape model methods [35] that learn the expected shape of the LV from a training set of annotated images. This information is then used to constrain the segmentation process and prevent unexpected shape estimates. Among the shape model methods, the most common is the Active Shape Model (ASM) [10], which is able to capture the expected shape and its main modes of deformation in a simple linear model. The ASM methodology is based in a two-phase strategy: (1) a training phase, in which statistical information about the shape is extracted from a training set to learn the shape model, and (2) a test phase, in which the learned shape model is applied to an unseen image by fitting the model to observation points extracted from that image.

The downside of ASM is that this method assumes the border of the target object is described by a set of unique points, typically called landmarks, such that each landmark

✉ Carlos Santiago
carlos.santiago@ist.utl.pt

¹ Instituto Superior Técnico, Avenida Rovisco Pais 1,
Torre Norte, 1049-001 Lisbon, Portugal

is located at a specific 3D position on that object. Under this assumption, it is possible to establish correspondences between the landmarks in two different representations of the object—a necessary condition to learn the shape model in the training phase. In the context of cardiac MRI segmentation, the assumption that the LV can be described by unique landmarks is not valid for two reasons. First, there are not enough anatomical landmarks within each slice of the MRI volume to define the LV contour with specific contour points [38]. Previous methods have all dealt with this problem by simply resampling the slice contours in arc-length starting at a specific point [2, 32]. Second, the number of slices in the MRI volume varies from patient to patient and along the cardiac cycle [40], which means that the number of contours used to define the LV (one for each slice) also varies. This leads to the following question: *how can we learn a shape model from volumes with a variable number of slices?* One of the focuses of this work is to provide an answer to this question, by proposing a framework to resample the LV models in the training set to a predefined number of contours.

Once the training phase has been completed, and shape model has been learned, another issue arises in the test phase. The test volume may have a different number of slices than the learned shape model, which means the slice contours will not be located at the same positions as the volume slices. One approach would be to interpolate the test volume intensities to obtain new slices. However, motion artifacts and the large distance between slices make this approach a bad choice, meaning that another alternative is required. This leads to a second question: *how can we segment a test volume without interpolating voxel intensities between slices?* This work also addresses this problem by proposing an alternative approach based on resampling the learned shape model.

In summary, the contribution of this work is twofold: (1) a framework to learn the shape model is proposed, which is able to deal with the variable number of slices in the training phase, and (2) in the test phase, a resampling scheme is also proposed that allows the mean shape and the main modes of deformation of the shape model to be sampled at different positions along the LV axis. A preliminary version of this work was published at IbPRIA 2015 [40], with the difference that, in this new version, a comprehensive review of the state of the art is provided and the experimental evaluation of the method has been extended to include the segmentation of more than 600 additional cardiac MRI volumes.

The remaining of the paper is as follows. Section 2 presents a review of the state of the art in automatic cardiac MRI segmentation strategies. Section 3 describes the proposed methodology. Section 4 describes the experimental setup used to evaluate the proposed method, and Sect. 5

shows the results obtained. Finally, Sect. 6 concludes the paper with final remarks and future work.

2 State of the art

This section provides an overview of the main approaches in medical image segmentation problems, with particular focus on cardiac MRI segmentation.

The most common approaches can be subdivided into the following categories: (1) bottom-up methods [11, 21, 41, 46], (2) active contours or deformable models [3, 16, 17, 22, 31, 39, 44], (3) active shape/appearance models (ASM/AAM) [2, 10, 23, 30, 32, 37], database-guided methods [6, 7, 15, 45, 47, 48], and atlas-based methods [24, 25, 28, 49, 50]. Each one of these approaches is different in: (1) the prior knowledge used to constrain the problem, (2) how they search for the target organ in the image, and (3) the amount of a priori training required. This makes them have intrinsic advantages and drawbacks, depending on the application. A brief review of each approach is presented next.

Bottom-up approaches [11, 21, 41, 46] mainly rely on standard image processing techniques to detect the border of the target organ. Among those techniques are edge detection methods, morphological operators, and thresholding. The main advantage of these methods is their low computational complexity. However, they are sensitive to initial conditions and to image quality.

Deformable models have been popularized by two distinct approaches: active contours [22] and level sets [27], both of which were able to increase the robustness of the segmentations while keeping the computational complexity low. Nonetheless, the former approach was still sensitive to the initialization, and the latter was sensitive to image conditions. Later developments aimed at further improving the segmentation results under difficult image conditions by using more robust estimation methods [8, 31, 37, 44]. Alternatively, other methods proposed using shape and texture priors [12, 13, 26, 33, 34]. However, the prior knowledge embedded in the optimization function of these approaches was often designed by hand or learned using a small training set. This means that they were most likely unable to capture all the information about the shape or texture required to perform well in more diverse datasets, particularly if both diseased and healthy patients were analyzed.

The above issues led to the development of more sophisticated supervised learning models, in which training sets of annotated data were used to fully learn the shape and appearance of the target object. One of the most popular methods developed using this approach is the Active Shape Model (ASM) [10]. The main idea behind

ASMs is to describe the shape of an organ using a set of unique landmarks and the corresponding statistics, i.e., mean shape and main modes of deformation. The shape statistics are learned in a training phase, using an annotated training set. Then, in the test phase, the learned shape model is applied to a test image. The segmentation is obtained by, first, extracting edge points using a boundary-driven method, and then fitting the shape model to those edge points. Since the performance of the boundary-driven method greatly depends on image quality, later approaches either started combining ASMs with models the appearance of the target organ, namely using Active Appearance Models (AAM) [5, 9, 29], or proposing robust estimation methods [1, 36, 37]. Recently, 3D ASM/AAMs have been developed to address volumetric segmentation of 3D medical images [2, 18, 32]. However, the higher dimensionality increases the difficulty of describing the organ shape using unique landmarks and requires a larger dataset to fully capture the 3D shape and texture variations. Another main issue with ASM and AAM methods is that they need the model to be initialized close to the desired boundary.

On the other hand, database-guided methods are completely independent of an initial guess [6, 7, 15, 45, 47, 48]. Instead, these methods search the whole parameters space, which guarantees reproducible results, but with the disadvantage of greatly increasing the computational complexity. Furthermore, these methods also require a large number of training data to estimate the model parameters and are not robust to image conditions that are not represented in the training set.

Finally, atlas-based approaches rely on a labeled atlas of a cardiac MRI to obtain the segmentation of the LV. The atlas is typically generated from one manually segmented volume or by combining information from multiple segmented volumes [35]. Then, a new volume is segmented by mapping its coordinate space to the atlas through a non-rigid registration process [42]. This type of methods have been widely used in brain-related applications, and there are several software tools available for this purpose (e.g., SPM¹ and FSL²). For cardiac MRI, there has also been some research [24, 25, 28, 49, 50], because these methods allow more flexible segmentations for shapes that are not represented in the training set, compared to shape model approaches. However, they do not impose anatomical constraints on the transformation that maps the test volume to the atlas, and the intensity values of the transformed volume have to be computed through interpolation, which should be avoided as is explained in Sect. 3.

The focus of this work is to propose a framework for using an ASM to obtain the 3D segmentation of cardiac MRI volumes. In particular, this framework is able to deal with the problem of finding the unique landmarks in the LV models in the training data and proposes a method of resampling the learned shape model, i.e., the mean shape and the main modes of deformation.

3 Proposed method

Consider the diagram shown in Fig. 1. During the training phase, the annotations of the MRI volumes in the training set consist of manually drawn contours of the LV border in each slice of the volume. These contours are located at equally spaced positions along the LV axis (perpendicularly to the volume slices) and define the LV surface. However, the number of slices in the volumes is different, which means the number of contours that define the LV surface also varies (first stage in Fig. 1, under training Phase). Consequently, it is not possible to establish correspondences between the landmarks in different LV surfaces; thus, the shape model cannot be learned using the standard ASM approach. To circumvent this problem, a polynomial interpolation scheme is used to describe the position of a surface landmark along the LV axis (second stage in Fig. 1). This allows each training shape to be resampled in a predefined number of slices and guarantees that all of them have the same landmarks. The outcome of the training phase is, thus, a shape model with a previously specified number of slices (third stage in Fig. 1).

When we proceed to the test phase, the variable number of slices may cause another issue to arise. If the test volume has a different number of slices than the specified value used in the training phase (fifth stage in Fig. 1), then the position of the shape model contours will not match the position of the test volume slices. One possible approach would be to interpolate the test volume to determine the intensity values at the same positions as the shape model contours. However, the low spatial resolution of MRI along the LV axis and motion artifacts can cause significant displacements in the location of the LV contour in consecutive slices. Therefore, the LV border in the resulting interpolated image will often be blurred, as shown by the example in Fig. 2. In these cases, it is very difficult to accurately determine the location of the LV border, which means interpolating voxel intensities is not a good approach. An alternative approach is proposed that consists in resampling the shape model instead (going from the third to the forth stage in Fig. 1). In this way, the new shape model contours will be located at the same positions as the volume slices, and the

¹ <http://www.fil.ion.ucl.ac.uk/spm/>.

² <http://fsl.fmrib.ox.ac.uk/fsl/>.

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

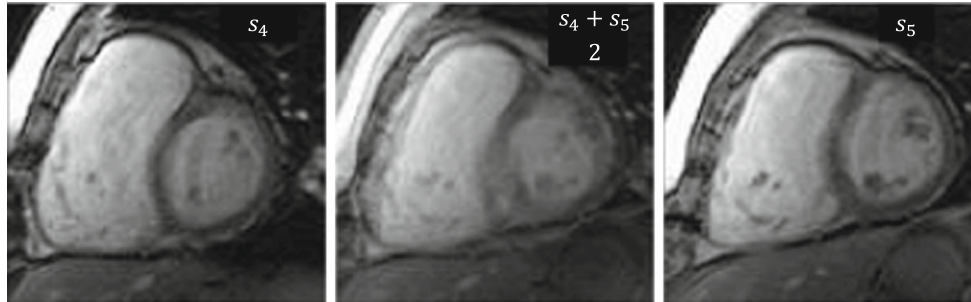
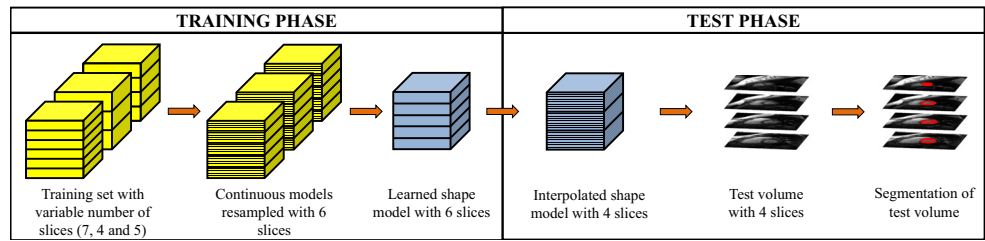


Fig. 2 Example of an interpolated image, at $s = \frac{s_4 + s_5}{2}$, obtained by linear interpolation between two consecutive slices, s_4 and s_5

segmentation can proceed as normal (last two stages in Fig. 1).

3.1 Training phase

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

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

where $m = 1, \dots, S^v$, and S^v is the number of slices in volume v . Assuming that the LV contours have all been previously sampled, in arc-length, with N points, we define $\mathbf{x}_v(s_m) \in \mathbb{R}^{2N \times 1}$ as the LV contour on the m th slice, such that

$$\mathbf{x}_v(s_m) = [\mathbf{x}^1{}^\top(s_m), \mathbf{x}^2{}^\top(s_m), \dots, \mathbf{x}^N{}^\top(s_m)]^\top, \quad (2)$$

where $\mathbf{x}^i(s_m) = [x_1^i, x_2^i]^\top \in \mathbb{R}^{2 \times 1}$ is the position of the i th point. We also assume that there is a correspondence between the i th point, $\mathbf{x}^i(s_m)$, and the i th point of another contour, $\mathbf{x}^i(s_{m'})$, $m' \neq m$, i.e., they represent the same landmark but in different slices. We wish to model the slice contour as a function of the axial position, $\hat{\mathbf{x}}_v(s)$, for any $s \in [0, 1]$ in order to resample the surface models in the training set. The next section describes the proposed methodology.

3.1.1 Interpolation methodology

In this work, the position of the contour points in a specific volume v , $\hat{\mathbf{x}}_v(s)$, is interpolated by 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 $\mathbf{C}_v \in \mathbb{R}^{2N \times K}$ is the coefficient matrix associated to volume v , defined by

$$\mathbf{C}_v = \begin{bmatrix} \mathbf{c}_1^1 \\ \mathbf{c}_1^2 \\ \vdots \\ \mathbf{c}_1^N \\ \mathbf{c}_2^1 \\ \mathbf{c}_2^2 \end{bmatrix}, \quad (4)$$

where $\mathbf{c}_j^i \in \mathbb{R}^{1 \times K}$ is the line of \mathbf{C}_v associated with the j th coordinate of the i th contour point. The coefficient matrix, \mathbf{C}_v , is specific of volume v and does not depend on the slice position s . On the other hand, the polynomial basis functions, $\boldsymbol{\psi}(s) = [1, s, \dots, s^{K-1}]^\top$, depend only on the slice position, s .

This representation provides an estimate of the LV contour for any position $s \in [0, 1]$, i.e., it is a continuous representation of the contour points along the LV axis, as shown in Fig. 3, that allows the surface models in the training set to be resampled with a predefined number of slice contours. However, first, the coefficient matrices \mathbf{C}_v have to be estimated for each volume using the corresponding annotations.

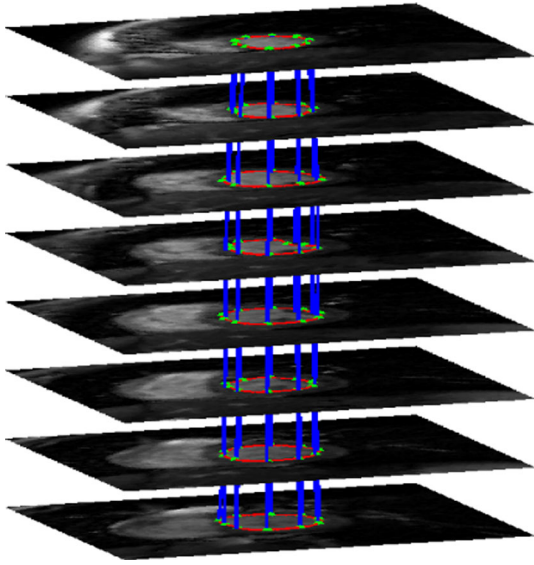


Fig. 3 Continuous representation of the surface model using a polynomial interpolation of the position of the contour points along the LV axis. The red lines represent the slice contours, the green dots are the contour points, and the blue lines are the interpolated position along the LV axis (trajectories) (color figure online)

Let $X_v^i \in \mathbb{R}^{2 \times S^v}$ denote the known position (trajectory) of the i th point in the contour along the axial position s ,

$$X_v^i = \begin{bmatrix} X_{1v}^i \\ X_{2v}^i \end{bmatrix} = \begin{bmatrix} x_1^i(s_1), & \dots, & x_1^i(s_{S^v}) \\ x_2^i(s_1), & \dots, & x_2^i(s_{S^v}) \end{bmatrix} \quad (5)$$

$$= [x^i(s_1), \dots, x^i(s_{S^v})].$$

The coefficient line $c_j^i \in \mathbb{R}^{1 \times K}$ is associated with the trajectory points X_{jv}^i and is computed by finding

$$c_j^i = \arg \min_c \|X_{jv}^i - \Psi c\|^2 + \gamma \|c\|^2, \quad (6)$$

where $\Psi = [\psi(s_1), \dots, \psi(s_{S^v})]^\top \in \mathbb{R}^{S^v \times K}$ is the concatenation of the polynomial basis functions $\psi(s_m)$ for $m = 1, \dots, S^v$, and γ is a regularization constant. This is a ridge regression formulation [19], which has the following solution

$$c_j^{i\top} = (\Psi^\top \Psi + \gamma I)^{-1} \Psi^\top X_{jv}^{i\top}, \quad (7)$$

where I is the $K \times K$ identity matrix. The regularization term constrains the solution and allows the estimation of c_j^i for any value of K , which differs from the ordinary least squares approach which only allows solutions for $K \leq S^v$.

The solution (7) can be computed for all the lines in C_v simultaneously, leading to

$$C_v^\top = (\Psi^\top \Psi + \gamma I)^{-1} \Psi^\top X_v^\top, \quad (8)$$

where $X_v = [x_v(s_1), \dots, x_v(s_{S^v})] \in \mathbb{R}^{2N \times S^v}$.

After computing C_v for each volume v in the training set, the corresponding surface models are sampled using (3) at the positions $s_m = \frac{m-1}{S^v-1}$, $m = 1, \dots, S^r$, where S^r is the desired number of slices. This guarantees that all the surface models have the same number of landmarks.

3.1.2 Learning the shape statistics

Once all the surface models in the training set have been resampled, it is possible to learn a shape model (from the second to the third stage in Fig. 1). Following the standard ASM methodology [10], we assume any surface model results from deforming the mean shape and applying a transformation associated with 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

$$E(\theta) = \sum_{m=1}^{S^r} \sum_{i=1}^N \left\| T_\theta \left(\hat{x}^i \left(\frac{m-1}{S^r-1} \right) \right) - x^{i \text{ref}} \left(\frac{m-1}{S^r-1} \right) \right\|^2, \quad (9)$$

where $\hat{x}^i \left(\frac{m-1}{S^r-1} \right)$ is the i th point in the slice contour located at the position $\frac{m-1}{S^r-1}$, $x^{i \text{ref}} \left(\frac{m-1}{S^r-1} \right)$ is the corresponding point in a reference surface model (for instance, one of the training surface models randomly selected) located at the same axial position, and $T_\theta(\cdot)$ is a 2D similarity transformation with parameters $\theta = \{a, t\}$, applied to all slices, such that

$$T_\theta(\hat{x}^i(s)) = \hat{X}^i(s)a + t, \quad (10)$$

where

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

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 (9) leads to a standard least squares solution similar to [10].

After the training surfaces have been aligned, the mean shape of each slice, $\bar{x}(s)$, is computed as the average slice contour over all the volumes in the training set. The first L main modes of deformation, $D(s) = [d_1(s), \dots, d_L(s)] \in \mathbb{R}^{2N \times L}$, and the corresponding eigenvalues, $\lambda_l(s)$, are obtained by principal component analysis (PCA), where $d_l(s) \in \mathbb{R}^{2N \times 1}$ 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.

3.2 Test phase

The shape model learned using the method described in the previous sections is used to segment other cardiac MR volumes (test volumes). As mentioned above, the number of slices in a test volume, which we denote as S^t , may not be the same as the learned shape model, $S^r \neq S^t$.

Since resampling volume intensities is not a good approach (recall Sect. 3), 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 (from the third to the fourth stage in 1). The following subsections describe: (1) the interpolation of the shape model and (2) the estimation of the model parameters that segment the LV in a test volume.

3.2.1 Resampling the learned shape model

The mean shape can be easily interpolated using the methodology described in Sect. 3.1. We compute the corresponding coefficient matrix, $\bar{\mathbf{C}}$, using

$$\bar{\mathbf{C}}^\top = (\Psi^\top \Psi + \gamma \mathbf{I})^{-1} \Psi^\top \bar{\mathbf{X}}^\top, \quad (11)$$

which is similar to (8), except the trajectories used in this case correspond to the trajectories of the points in the mean contour, $\bar{\mathbf{X}} = [\bar{\mathbf{x}}(s_1), \dots, \bar{\mathbf{x}}(s_{S^r})]$. Then, the mean shape is resampled using (3) at S^t slices, located at $s = \frac{m-1}{S^t-1}$, with $m = 1, \dots, S^t$. Thus, a mean shape with S^t equally spaced slice contours is obtained.

Regarding the interpolation of the main modes of deformation, applying the previous methodology will likely lead to unwanted results. For instance, if we want to determine the first mode of deformation in a specific slice position, s , that does not coincide with the positions of the learned shape model, we cannot simply interpolate using the known first modes in the other positions. The reason is that the deformation modes are learned independently for each slice, and then, they are sorted according to the value of the corresponding eigenvalues. Therefore, it is not possible to guarantee that the first mode describes the same deformation in all the slices.

In this work, we use a simpler approach that consists in finding the correspondences between deformation modes in consecutive slices and use the resulting pairs of deformation modes 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 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 $F(\cdot)$ maps the correspondences between the deformation modes in s_m to s_{m+1} . This mapping is achieved by searching for the deformation mode in the $(m+1)$ th slice that is most similar to $\mathbf{d}_l(s_m)$, according to the Euclidean distance,

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

The 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

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

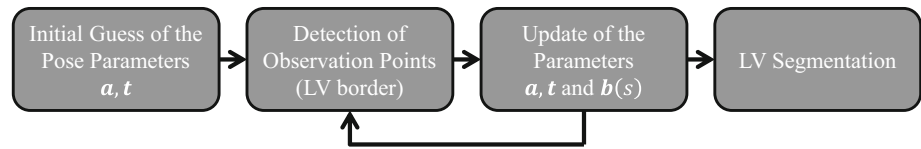
where $T_\theta(\cdot)$ is defined in (10). This means that the segmentation of the test volume is obtained by finding the parameters for the pose transformation, $\theta = \{\mathbf{a}, \mathbf{t}\}$, and the deformation coefficients, $\mathbf{b}(s)$, that better defines the LV border. This final step is addressed in the following section.

3.2.2 Segmentation of a test volume

Given the test volume, the segmentation of the LV is obtained by estimating the pose, defined by transformation T_θ 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 (outer border of the LV), papillary muscles and trabeculations [35], that should be considered as noise or outliers.

In this work, the estimation of the shape model parameters in the test phase is achieved by using a robust estimation method called EM-RASM [37]. The algorithm iteratively updates the model parameters as shown in Fig. 4. First, an initial guess of the pose parameters is provided—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, ideally located at LV border, are detected in the vicinity of the model (second block in Fig. 4). These observation points are searched in each slice of the volume, along lines orthogonal to the contour model, by applying an edge detector (see [4] Sect. 5.2 for details).

Fig. 4 EM-RASM: estimation of the model parameters that segment the test volume



The approach used to extract the observation points often leads to the detection of observation points that do not belong to the LV border (outliers), e.g., on intensity transitions associated with the epicardium or papillary muscles. The EM-RASM is able to handle these outliers by assuming that each observation point may be an outlier or a valid point. It assigns each observation point a weight based on the probability that it belongs to the LV border. The weights determine their influence in the estimation of the model parameters, θ 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 [37] for an in-depth description), computed over all the slice contours simultaneously (third block in Fig. 4).

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

4 Experimental setup

The proposed method was evaluated on the publicly available dataset [2], containing 33 sequences of 3D cardiac MRI short axis volumes. Each sequence covers one cardiac cycle with 20 volumes of a different patient, whose age ranges between 8 and 15 years old. Out of the 33 patients, only two are healthy and two have unknown diagnosis, which means most of them are diseased patients. However, not all the diseases are related to the LV (e.g., enlarged right ventricle). Manually drawn endocardial contour are also provided and will be used to evaluate the proposed algorithm.

For each test sequence, the shape model was trained using the remaining 32 sequences, i.e., the results were obtained in a leave-one-sequence-out scheme. The slice contours in the training set were resampled, in arc-length, with $N = 40$ points, and the shape model was learned with $S^r = 10$ slices. Empirical experiments (see Sect. 5.1) showed that the best parameters for the polynomial interpolation were $K = 8$ and $\gamma = 10^{-4}$. The number of

deformation modes used was $L = 10$, which corresponded to over 95 % of the variability in the training set.

In the test phase, the segmentations obtained were quantitatively evaluated using the average Dice similarity coefficient [14], d_{Dice} , and the average minimum distance between the shape model points and the ground truth, d_{AV} , measured in mm. These metrics are computed as follows. Let $\tilde{\mathbf{x}} = \{\tilde{\mathbf{x}}(s_1), \dots, \tilde{\mathbf{x}}(s_{S'})\}$ and $\mathbf{g} = \{\mathbf{g}(s_1), \dots, \mathbf{g}(s_{S'})\}$ denote the set of contours that define the segmentation obtained using the proposed method and the ground truth, respectively, of a test volume with S' slices. Also, let $R(\tilde{\mathbf{x}})$ be the 3D region delimited by the contours in $\tilde{\mathbf{x}}$ and $R(\mathbf{g})$ be the 3D region delimited by the ground truth segmentation. The Dice similarity coefficient is given by

$$d_{\text{Dice}}(R(\tilde{\mathbf{x}}), R(\mathbf{g})) = 2 \frac{V(R(\tilde{\mathbf{x}}) \cap R(\mathbf{g}))}{V(R(\tilde{\mathbf{x}})) + V(R(\mathbf{g}))}, \quad (17)$$

where $V(\cdot)$ denotes the volume of a region, and \cap denotes the intersection. A Dice coefficient of 1 means there is a perfect match between the two segmentations, and a value of 0 means the corresponding regions do not even overlap. Regarding the average minimum distance metric, the smallest distance between the ground truth and the i th point of the shape model on the m th slice, $\tilde{\mathbf{x}}^i(s_m)$, is defined as

$$d(\tilde{\mathbf{x}}^i(s_m), \mathbf{g}(s_m)) = \min_j \|\mathbf{g}^j(s_m) - \tilde{\mathbf{x}}^i\|_2, \quad (18)$$

where $\mathbf{g}(s_m)^j$ is the j th point of the ground truth segmentation of the m th slice. Thus, the average minimum distance between the two segmentations is

$$d_{\text{AV}}(\tilde{\mathbf{x}}, \mathbf{g}) = \frac{1}{NS'} \sum_{m=1}^{S'} \sum_{i=1}^N d(\tilde{\mathbf{x}}^i(s_m), \mathbf{g}(s_m)). \quad (19)$$

5 Results

In this section, statistical results are presented and discussed for three scenarios: (1) sensitivity of the proposed method to parameter variation, namely the parameters associated with the resampling scheme used in the training set to normalize the number of slices in each volume; (2) sensitivity of the EM-RASM to the model initialization in a synthetic case; and (3) the segmentation of the left ventricle in cardiac MRI.

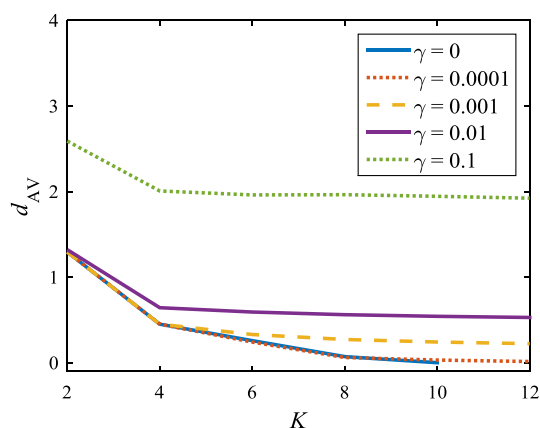


Fig. 5 Sensitivity to parameter variation. Each line shows the d_{AV} metric between the original contours and the interpolated ones for different values of K and γ

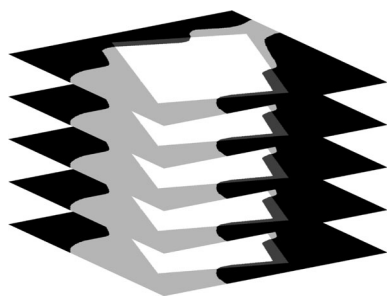
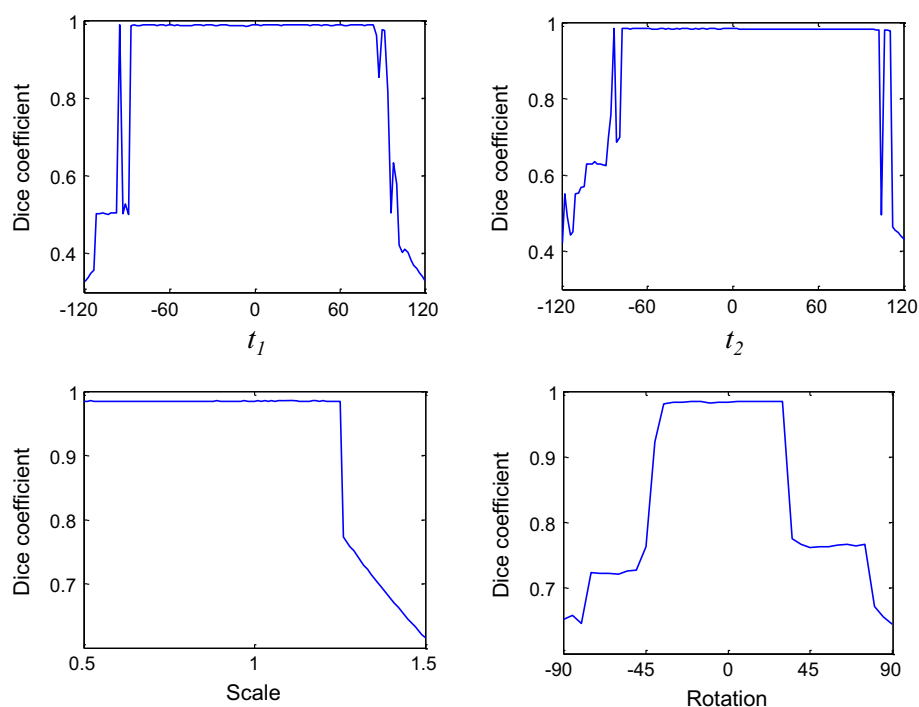


Fig. 6 Synthetic example used to evaluate the sensitivity of the algorithm to initialization

Fig. 7 Sensitivity to the model initialization using Dice coefficient. Each plot shows the performance of the algorithm (ML) for different initial values of the transformation parameter: t_1 , t_2 , scale, and rotation

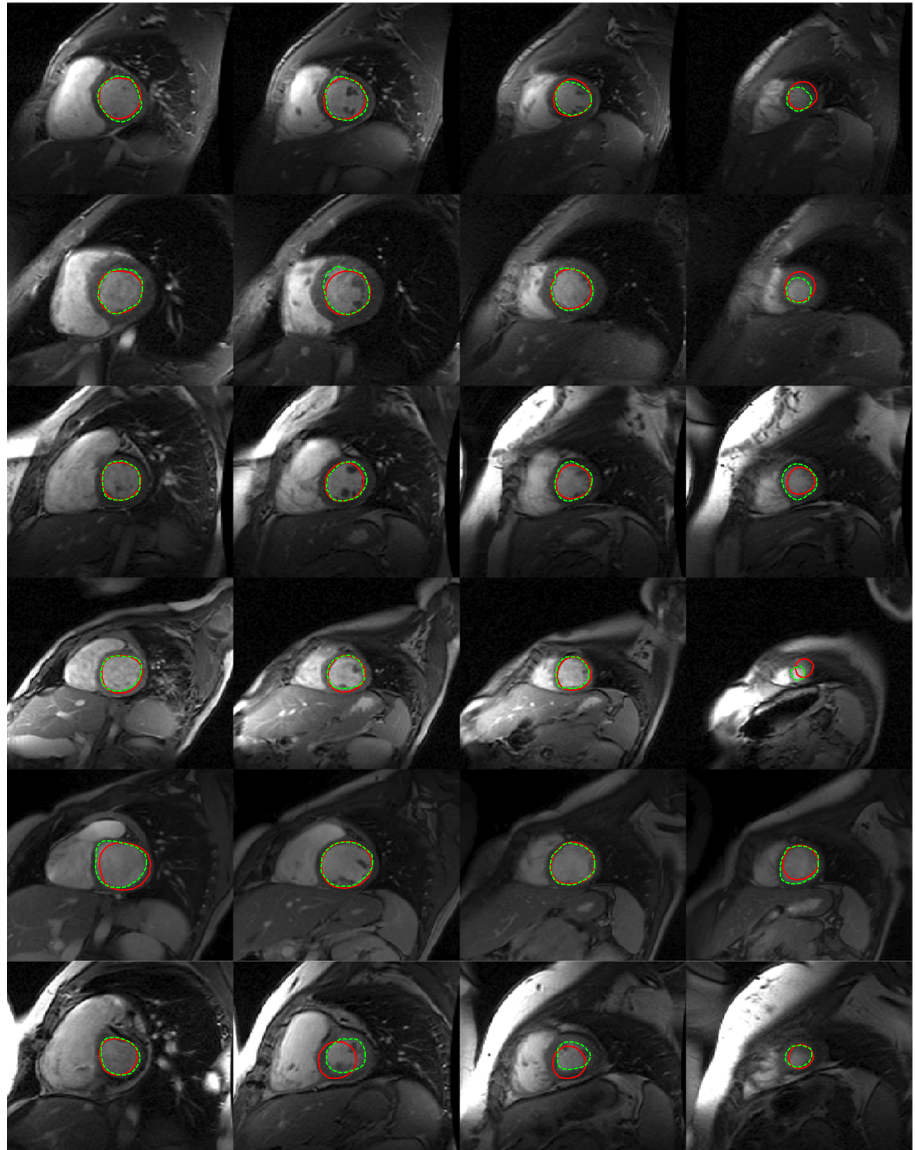


5.1 Sensitivity to parameter variation

The performance of the resampling scheme described in Sect. 3.1 depends on two main parameters: (1) the number of polynomial basis functions, K , and (2) the regularization constant, γ , used to estimate the coefficient matrix in (8). The first parameter is related to the degree of the polynomial used to model the trajectory of a contour point along the slices. There are two possible scenarios: (1) K is smaller than the number of slices in the volume, S^v , in which case the polynomial is not able to accurately describe the trajectories of the contour points, i.e., the interpolated model is an approximation of those trajectories, and (2) $K \geq S^v$, in which case the accuracy of the interpolated model is only limited by the regularization term. On the other hand, the regularization parameter, γ , is responsible for keeping the entries in the coefficient matrix small, which keeps the trajectories estimates smooth.

Figure 5 shows the d_{AV} metric between the original contours in the training set and the corresponding interpolated contours (using the same number of slices), for different values of $K = \{2, 4, 6, 8, 10, 12\}$ and $\gamma = \{0, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}\}$. The figure shows that higher values of K lead to better results. It is possible to see that for each curve, there is a value of K beyond which the accuracy improvement is not significant. As for the regularization constant, it is possible to see that smaller values lead to better results. However, it is important to note that for $\gamma = 0$, the computation of the coefficient matrix has

Fig. 8 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)



numerical issues for $K > S^v$, because $\Psi^T \Psi$ becomes non-invertible. In such cases, the d_{AV} was not computed and the volume was removed from the results. Since $S^v \leq 10$ for all the volumes in the dataset, it was not possible to compute any coefficient matrix for $K = 12$. Based on these results, the choice of parameter values to use in the following sections was $K = 10$ and $\gamma = 10^{-4}$.

5.2 Sensitivity to initialization

The synthetic example shown in Fig. 6 was used to evaluate the sensitivity of the algorithm to the initialization. The example consists of a stack of 5 slices with the same gray-scale image depicting a white rectangle in a

gray and black background. The several initial values for the transformation parameters θ (which define the rotation, scaling, and translation of the shape model) were tested and, for each case, the Dice coefficient of the final segmentation was computed. Figure 7 shows the results obtained of varying each transformation parameter separately. In this example, the parameters that lead to the correct segmentation are $t_1 = t_2 = 0$, a scale of 1 and a rotation of 0 degrees. It is possible to see that, for each plot, there is a wide plateau within which the algorithm is able to converge to the correct solution. This shows that the algorithm is robust to the initialization conditions and that it is able to obtain the expected segmentation unless the initial parameter estimates are very far from the correct values.

5.3 Left ventricle segmentation

This subsection presents the results of the segmentations obtained using the proposed algorithm in the LV segmentation problem. Comparative results are also presented, to evaluate the benefits and disadvantages of the proposed method against its two state-of-the-art counterparts: the standard ASM approach and the Active Appearance Model (AAM) approach.

Figure 8 shows examples of the segmentations obtained using the proposed method. In most cases, the algorithm is able to accurately determine the location of the LV and disregards papillary muscles and other misleading structures. This is due to the robust parameter estimation provided by the EM-RASM algorithm, which assigns low weights to these outliers. However, the model has difficulties in dealing with the LV border whenever the volume has misaligned slices (e.g., see bottom row of Fig. 8), as these random variations are not captured by shape model (i.e., they cannot be represented using the available deformation modes). Furthermore, the proposed method does not always provide accurate segmentations in the apical slices, due to: (1) the lack of contrast between the intensity inside and outside the LV (first and second rows in Fig. 8) and (2) the variability in both size and position of the LV contour in these slices (fourth row in Fig. 8). Since in this work, the detection of the LV border is based on edge detection, obtaining accurate segmentation of apical slices is a challenging task.

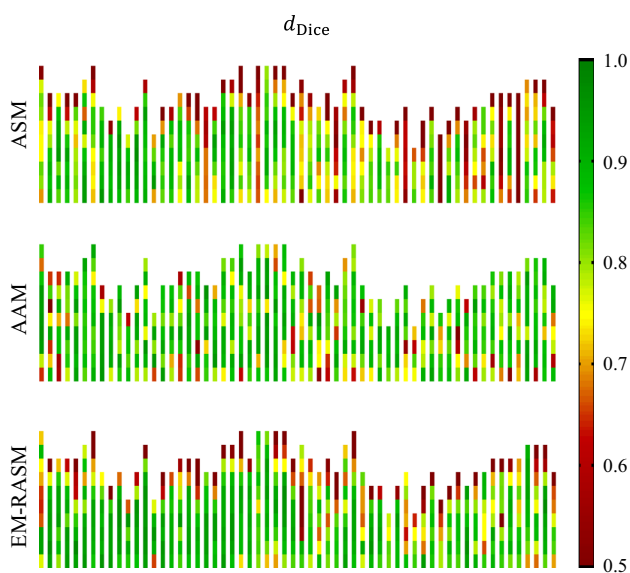


Fig. 9 Quantitative evaluation of 60 volumes of the dataset using the Dice coefficient. Each column corresponds to one volume, and each pixel in the columns is associated with a volume slice. The colors are defined by the colorbar (greener is better) (color figure online)

Table 1 Statistical results of the overall performance, for the 660 volumes in the dataset [2], using the proposed method and comparison with: (1) the standard ASM [10] and (2) the AAM methodology proposed in [43]

	ASM [10]	AAM [43]	EM-RASM
Dice (%)	73.1 (13.1)	81.7 (3.6)	79.4 (8.1)
AV (mm)	4.7 (3.0)	2.9 (0.8)	3.5 (1.4)

Each table entry shows the mean (above) and standard deviation (below)

Figure 9 shows a quantitative evaluation of the proposed algorithm (bottom row) for 60 volumes with discriminative results for each slice. It is possible to see that the algorithm achieves good segmentations for the majority of the volume slices. However, some of them are not quite accurate, particularly in the apical slices. This is due to the already mentioned problems: (1) inability to fit misaligned slices and (2) difficulty in finding evidence of the LV border in the apical slices.

These results are compared with two other approaches: (1) the standard ASM, proposed in [10], and (2) the AAM methodology proposed in [43]. Figure 9 shows that the former is not always able to problem the correct segmentations, mainly due to its inability to distinguish the LV border from other anatomical structures in the image, such as papillary muscles. Furthermore, it also struggles with the correct segmentation of the LV in the apical slices, for the same reasons as the proposed method. On the other hand, the AAM methodology provides more reliable results for the apical slices, since it uses texture information instead of an edge detector. However, the accuracy of the AAM in the remaining slices is more variable.

Statistical results for the overall performance are depicted in Table 1. It is possible to see that the proposed method outperforms the standard ASM and achieves similar result to the AAM methodology. A slightly lower performance is expected since the AAM approach uses texture information to guide the segmentation, which is a significant advantage over the edge detector used in this work.

6 Conclusions

This paper proposes a 3D Active Shape Model (ASM) for the segmentation of the left ventricle in cardiac MRI. Although ASM-based approaches are common in medical image segmentation problems, this approach requires the number of slices in each training surface models to be normalized first.

We propose to deal with this issue by using a continuous representation for the surface model, which allows the slice contour in any position to be obtained. This approach is used to resample all the surface models in the training set, thus establishing a correspondence between the landmarks (surface points) in different models.

The same problem arises in the test phase, where the learned shape model may have to be applied to a test volume with 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. Then, the segmentation is obtained by estimating the shape model parameters using the EM-RASM technique [37], which increases the robustness in the presence of outliers.

The results show that the proposed method achieves good segmentation results, outperforming the standard ASM approach and achieving similar results to the AAM methodology proposed in [43].

Future work should focus on improving the proposed methodology by: (1) allowing local translations of each slice contour, to account for possible misalignments between consecutive slices, and (2) using a more robust method to extract observation points from the volume slices, such as using texture information (similarly to the AAM), in order to improve the accuracy in the slices where the LV border is harder to detect.

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

References

1. Abi-Nahed J, Jolly MP, Yang GZ (2006) Robust active shape models: a robust, generic and simple automatic segmentation tool. In: Larsen R, Nielsen M, Sparring J (eds) Medical image computing and computer-assisted intervention–MICCAI 2006. Springer, Berlin, Heidelberg, pp 1–8
2. Andreopoulos A, Tsotsos JK (2008) Efficient and generalizable statistical models of shape and appearance for analysis of cardiac MRI. *Med Image Anal* 12(3):335–357
3. Billet F, Sermesant M, Delingette H, Ayache N (2009) Cardiac motion recovery and boundary conditions estimation by coupling an electromechanical model and cine-MRI data. In: Ayache N, Delingette H, Sermesant M (eds) Functional imaging and modeling of the heart. Springer, Berlin, Heidelberg, pp 376–385
4. Blake A, Isard M (1998) Image processing techniques for feature location. In: Active contours. Springer, London, pp 97–113
5. Bosch JG, Mitchell SC, Lelieveldt BPF, Nijland F, Kamp O, Sonka M, Reiber JHC (2002) Automatic segmentation of echocardiographic sequences by active appearance motion models. *IEEE Trans Med Imaging* 21(11):1374–1383
6. Carneiro G, Georgescu B, Good S, Comaniciu D (2008) Detection and measurement of fetal anatomies from ultrasound images using a constrained probabilistic boosting tree. *IEEE Trans Med Imaging* 27(9):1342–1355
7. Carneiro G, Nascimento JC (2010) Multiple dynamic models for tracking the left ventricle of the heart from ultrasound data using particle filters and deep learning architectures. In: Conference computer vision and pattern recognition (CVPR)
8. Chen T, Babb J, Kellman P, Axel L, Kim D (2008) Semiautomated segmentation of myocardial contours for fast strain analysis in cine displacement-encoded MRI. *IEEE Trans Med Imaging* 27(8):1084–1094
9. Cootes T, Beeston C, Edwards G, Taylor C (1999) A unified framework for atlas matching using active appearance models. In: Kuba A, Sámal M, Todd-Pokropek A (eds) Information processing in medical imaging. Springer, Berlin, Heidelberg, pp 322–333
10. Cootes TF, Taylor CJ, Cooper DH, Graham J (1995) Active shape models—their training and application. *Comput Vis Image Underst* 61(1):38–59
11. Cousty J, Najman L, Couprie M, Clément-Guinaudeau S, Goissen T, Garot J (2007) Automated, accurate and fast segmentation of 4D cardiac MR images. In: Sachse FB, Seemann G (eds) Functional imaging and modeling of the heart. Springer, Berlin, Heidelberg, pp 474–483
12. Cremers D (2006) Dynamical statistical shape priors for level set-based tracking. *IEEE Trans Pattern Anal Mach Intell* 28(8):1262–1273
13. Cremers D, Osher S, Soatto S (2006) Kernel density estimation and intrinsic alignment for shape priors in level set segmentation. *Int J Comput Vis* 69(3):335–351
14. Dice LR (1945) Measures of the amount of ecologic association between species. *Ecology* 26(3):297–302
15. Georgescu B, Zhou XS, Comaniciu D, Gupta A (2005) Database-guided segmentation of anatomical structures with complex appearance. In: Conference computer vision and pattern recognition (CVPR)
16. Gopal S, Terzopoulos D (2014) A unified statistical/deterministic deformable model for LV segmentation ins cardiac MRI. In: Camara O, Mansi T, Pop M, Rhode K, Sermesant M, Young A (eds) Statistical atlases and computational models of the heart. Imaging and modelling challenges. Springer, Berlin, Heidelberg, pp 180–187
17. Grosgeorge D, Petitjean C, Caudron J, Fares J, Dacher JN (2011) Automatic cardiac ventricle segmentation in MR images: a validation study. *Int J Comput Assist Radiol Surg* 6(5):573–581
18. Heimann T, Meinzer HP (2009) Statistical shape models for 3D medical image segmentation: a review. *Med Image Anal* 13(4):543–563
19. Hoerl AE, Kennard RW (1970) Ridge regression: biased estimation for nonorthogonal problems. *Technometrics* 12(1):55–67
20. Hundley WG, Bluemke DA, Finn JP, Flamm SD, Fogel MA, Friedrich MG, Ho VB, Jerosch-Herold M, Kramer CM, Manning WJ et al (2010) ACCF/ACR/AHA/NASCI/SCMR 2010 expert consensus document on cardiovascular magnetic resonance: a report of the American College of Cardiology Foundation Task Force on Expert Consensus Documents. *J Am Coll Cardiol* 55(23):2614–2662
21. Jolly M (2009) Fully automatic left ventricle segmentation in cardiac cine MR images using registration and minimum surfaces. *MIDAS J Cardiac MR Left Ventricle Segm Chall* 4
22. Kass M, Witkin A, Terzopoulos D (1988) Snakes: active contour models. *Int J Comput Vis* 1(4):321–331
23. Kaus MR, Jv Berg, Weese J, Niessen W, Pekar V (2004) Automated segmentation of the left ventricle in cardiac MRI. *Med Image Anal* 8(3):245–254
24. Lorenzo-Valdés M, Sanchez-Ortiz GI, Elkington AG, Mohiaddin RH, Rueckert D (2004) Segmentation of 4D cardiac MR images using a probabilistic atlas and the EM algorithm. *Med Image Anal* 8(3):255–265
25. Lötjönen J, Kivistö S, Koikkalainen J, Smutek D, Lauerma K (2004) Statistical shape model of atria, ventricles and epicardium

- from short-and long-axis MR images. *Med Image Anal* 8(3): 371–386
26. Lynch M, Ghita O, Whelan PF (2008) Segmentation of the left ventricle of the heart in 3-D+t MRI data using an optimized non-rigid temporal model. *IEEE Trans Med Imaging* 27(2):195–203
 27. Malladi R, Sethian J, Vemuri B (1995) Shape modeling with front propagation: a level set approach. *IEEE Trans Pattern Anal Mach Intell* 17:158–175
 28. Medrano-Gracia P, Cowan BR, Bluemke DA, Finn JP, Lima JA, Suinesiaputra A, Young AA (2013) Large scale left ventricular shape atlas using automated model fitting to contours. In: Ourselin S, Rueckert D, Smith N (eds) *Functional imaging and modeling of the Heart*, vol 7945., *Lecture Notes in Computer Science* Springer, Berlin Heidelberg, pp 433–441
 29. Mitchell S, Lelieveldt B, van der Geest R, Bosch H, Reiber J, Sonka M (2001) Multistage hybrid active appearance model matching: segmentation of left and right ventricles in cardiac MR images. *IEEE Trans Med Imaging* 20(5):415–423
 30. Mitchell SC, Bosch JG, Lelieveldt BP, van der Geest RJ, Reiber JH, Sonka M (2002) 3-D active appearance models: segmentation of cardiac MR and ultrasound images. *IEEE Trans Med Imaging* 21(9):1167–1178
 31. Nascimento JC, Marques JS (2008) Robust shape tracking with multiple models in ultrasound images. *IEEE Trans Image Process* 17(3):392–406
 32. O'Brien SP, Ghita O, Whelan PF (2011) A novel model-based 3D time left ventricular segmentation technique. *IEEE Trans Med Imaging* 30(2):461–474
 33. Paragios N (2003) A level set approach for shape-driven segmentation and tracking of the left ventricle. *IEEE Trans Med Imaging* 22(6):773–776
 34. Paragios N, Deriche R (2002) Geodesic active regions and level set methods for supervised texture segmentation. *Int J Comput Vis* 46(3):223–247
 35. Petitjean C, Dacher J (2011) A review of segmentation methods in short axis cardiac MR images. *Med Image Anal* 15(2):169–184
 36. Rogers M, Graham J (2006) Robust active shape model search. In: Heyden A, Sparr G, Nielsen M, Johansen P (eds) *Computer vision—ECCV 2002*. Springer, Berlin, Heidelberg, pp 517–530
 37. Santiago C, Nascimento J, Marques J (2015) 2D Segmentation using a robust active shape model with the EM algorithm. *IEEE Trans Image Process* 24(8):2592–2601. doi:[10.1109/TIP.2015.2424311](https://doi.org/10.1109/TIP.2015.2424311)
 38. Santiago C, Nascimento JC, Marques JS (2013) Performance evaluation of point matching algorithms for left ventricle motion analysis in MRI. In: *Engineering in medicine and biology society (EMBC), 2013 35th annual international conference of the IEEE*. IEEE, pp 4398–4401
 39. Santiago C, Nascimento JC, Marques JS (2015) Automatic 3-D segmentation of endocardial border of the left ventricle from ultrasound images. *IEEE J Biomed Health Inform* 19(1):339–348. doi:[10.1109/JBHI.2014.2308424](https://doi.org/10.1109/JBHI.2014.2308424)
 40. Santiago C, Nascimento JC, Marques JS (2015) Robust 3D active shape model for the segmentation of the left ventricle in MRI. In: Paredes R, Cardoso JS, Pardo XM (eds) *Pattern recognition and image analysis—IbPRIA'15*. Springer, Switzerland, pp 283–290
 41. Sonka M, Zhang X, Siebes M, Bissing M, Dejong S, Collins S, McKay C (1995) Segmentation of intravascular ultrasound images: a knowledge-based approach. *IEEE Trans Med Imaging* 14:719–732
 42. Studholme C, Hill DL, Hawkes DJ (1997) Automated three-dimensional registration of magnetic resonance and positron emission tomography brain images by multiresolution optimization of voxel similarity measures. *Med Phys* 24(1):25–35
 43. Tzimiropoulos G, Pantic M (2013) Optimization problems for fast aam fitting in-the-wild. In: *Proceedings of the IEEE international conference on computer vision*. pp 593–600
 44. Uzunbas MG, Zhang S, Pohl KM, Metaxas D, Axel L (2012) Segmentation of myocardium using deformable regions and graph cuts. In: *2012 9th IEEE international symposium on biomedical imaging (ISBI)*. IEEE, pp 254–257
 45. Weng J, Singh A, Chiu M (1997) Learning-based ventricle detection from cardiac mr and ct images. *IEEE Trans Med Imaging* 16(4):378–391
 46. Zhang L, Geiser E (1984) An effective algorithm for extracting serial endocardial borders from 2-D echocardiograms. *IEEE Trans Biomed Eng BME* 31:441–447
 47. Zheng Y, Barbu A, Georgescu B, Scheuering M, Comaniciu D (2008) Four-chamber heart modeling and automatic segmentation for 3-D cardiac CT volumes using marginal space learning and steerable features. *IEEE Trans Med Imaging* 27(11):1668–1681
 48. Zhou XS, Comaniciu D, Gupta A (2005) An information fusion framework for robust shape tracking. *IEEE Trans Pattern Anal Mach Intell* 27(1):115–129
 49. Zhuang X, Hawkes D, Crum W, Boubertakh R, Uribe S, Atkinson D, Batchelor P, Schaeffter T, Razavi R, Hill D (2008) Robust registration between cardiac MRI images and atlas for segmentation propagation. In: Reinhardt JM, Pluim JPW (eds) *Medical imaging*. International Society for Optics and Photonics, SPIE, pp 691408
 50. Zhuang X, Rhode KS, Razavi RS, Hawkes DJ, Ourselin S (2010) A registration-based propagation framework for automatic whole heart segmentation of cardiac MRI. *IEEE Trans Med Imaging* 29(9):1612–1625