3D active shape modeling for cardiac MR and CT image segmentation

Assen, Hans Christiaan van

'There is nothing worse than a sharp image of a fuzzy concept.'

Ansel Adams (1902–1984)

Cardiac LV Segmentation Using a 3D Active Shape Model Driven by Fuzzy Inference

Abstract

Manual quantitative analysis of cardiac left ventricular function using multi-slice CT is labor intensive because of the large datasets. In previous work, an intrinsically three-dimensional segmentation method for cardiac CT images was presented based on a 3D Active Shape Model (3D-ASM). This model systematically overestimated left ventricular volume and underestimated blood pool volume, due to inaccurate estimation of candidate points during the model update steps. In this paper, we propose a novel ASM candidate point generation method based on a Fuzzy Inference System (FIS), which uses image patches as an input. Visual and quantitative evaluation of the results for 7 out of 9 patients shows substantial improvement for endocardial contours, while the resulting volume errors decrease considerably (blood pool: -39 ± 29 cubic voxels in the previous model, -0.66 ± 6.2 cubic voxels in the current). Standard deviation of the epicardial volume decreases by approximately 50%.

3.1   Introduction

Deformable statistical models have proven to be highly useful in medical image analysis. Especially with respect to segmentation tasks, they can be very effective and therefore they form an active field of research. The knowledge in the model ranges from shape and shape variation to gray level information. Since 1992, Cootes *et al.* introduced several statistical modeling methods, the Point Distribution Model (PDM) [11], the Active Shape Model (ASM) [19], which is an extension to the PDM with a matching algorithm, and the Active Appearance Model (AAM) [27]: an extension of the ASM using a statistical intensity model of image patches. PDMs have been explored in 3D [16, 53], mainly for the analysis of complex 3D shapes. AAMs have also been developed in 3D, and applied to medical image segmentation problems [31]. Applications for ASMs, however, have mainly been limited to 2D or 2D plus time [52, 23].
In previous work, we developed a cross-modality 3D-ASM [60] for cardiac left ventricle (LV) segmentation of MR and CT images, while avoiding the necessity of intensity model retraining for each modality. We showed that the 3D-ASM is a promising tool for this purpose. However, due to the basic edge detection applied to define model updates, an underestimation of blood pool volume inside the LV and an overestimation of total LV-volume resulted. This was our primary motivation to design a method to more accurately localize the candidate boundary positions in the image slices. For this purpose, we chose to adopt fuzzy inference (FI) for two reasons:

- by applying a classification approach to image patches instead of pixel scan lines, sensitivity to local disturbances and noise can be greatly reduced,

- the model should be applicable to both MR and CT without retraining a statistical intensity model, because of the large training sets inherent to retraining. This independence is realized by selecting fuzzy clustering for classification, since it is based on relative intensity differences between tissues.

(a)   (b)
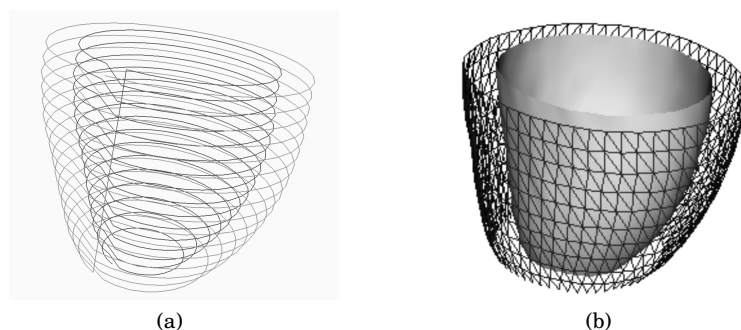
Figure 3.1: (a) Line parameterization, defining the point correspondence for the 3D cardiac left ventricle (b) Mesh used for matching the model to the image information.
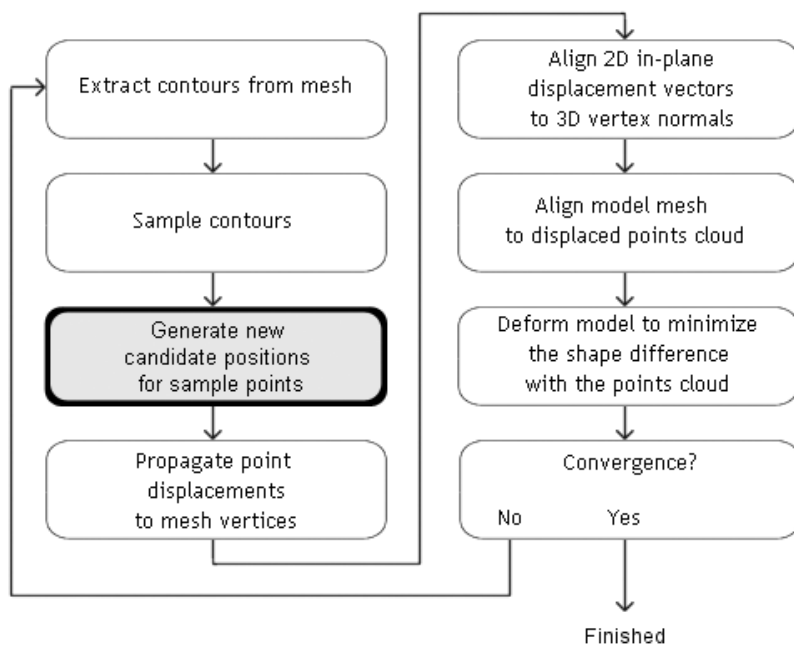

Figure 3.2: Schematic representation of the matching procedure used in this 3D-ASM. The focus of this paper is the generation of new candidate positions for the sample points that construct the mesh (indicated in gray).
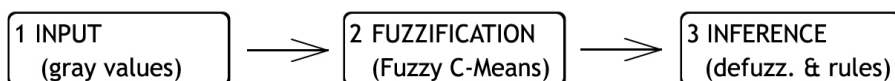

Figure 3.3: Takagi-Sugeno-based inference system.

3.2   Methodology

3.2.1   3D model generation

To perform statistical analysis on 3D-shapes, three important issues were addressed: point correspondence for 3-dimensional cardiac LV shapes, 3D shape alignment, and statistical modeling of 3D shape and shape variation.

The definition of point correspondences is an active field of research. Apart from a manual definition of point correspondence, different automatic methods have been explored. Styner *et al.* [17] presented an evaluation of four automatic point correspondence methods. This evaluation suggests the use of either DetCov or the MDL-based approach, however the DetCov method is not easily extendable to 3D. For our ASM however, we adopted an application specific solution and used a manually defined point correspondence [60] (see Fig. 3.1).

After line-parameterization of every shape in the training set, the resulting point set is resampled in the direction of the long axis. This is done to generate sets of sample points, that represent the same number of slices, and consequently the same number of sample points. Each shape sample can then be expressed as a 3n element vector x containing n concatenated 3-dimensional landmark points (xi, yi,zi) of a particular shape. Both the LV endo- and epicardial surface are represented in the same vector. Finally, the shape samples are aligned. The remaining differences are solely shape related, and thus the effect of trivial variations in pose and scale is eliminated. For alignment of the training samples, Procrustes analysis was applied. We adopted Besl and McKays iterative closest points (ICP) algorithm for registration of 3D points sets [59]. Shape modeling is analogous to the 2D case, except that 3D coordinates are used. To extract the modes of variation of the training set, the eigenvectors and covariance matrix are calculated using principal component analysis.

3.2.2   Model matching

The model was extended with a matching algorithm: an iterative procedure to determine update steps according to information from the target image set. During each update, the shape parameters are adjusted to best match the image evidence. For the 3D-ASM of the cardiac LV, a 3D triangular mesh was constructed from the sample points (see Fig. 3.1). During the matching, this mesh is intersected by the image planes, thus generating 2D contours spanned by the intersections of the mesh triangles. Model update information is represented by 2D point-displacement vectors, extracted from the 2D image slices. These 2D vectors are propagated to the 3D mesh representing the complete model update. This propagation is achieved by alignment of the 2D in-plane displacement vectors to the 3D surface normal vector. Similar to the training stage, scaling, rotation and translation differences between the current state of the model and the point cloud representing the model update step are eliminated by alignment. Both mesh states are aligned using the ICP method [59]. Successively, the parameter vector b controlling model deformation is adjusted:

$$\mathbf{b} = \Phi^T(\mathbf{x}_{proposed} - \mathbf{x}_{current}). \tag{3.1}$$

with $\mathbf{x}_{proposed}$ the vector representing the proposed shape, and $\mathbf{x}_{current}$ the vector representing the aligned current state of the mesh (see Fig. 3.2).

3.2.3 Edge detection using Fuzzy Inference

This paper addresses the generation of new point positions to define the model update steps in between iterations. In the classic ASM [19], model updates were generated using a statistical gray level model in each sample point, implying modality dependence. Since we require our ASM to be modality independent, we developed an alternative decision scheme based on the Takagi-Sugeno Fuzzy Inference System [61], because this is based on relative intensity differences. This system works as follows (see Fig. 3.3):

1. *input*
   For each intersection point between the mesh and the image, our system considers an image patch, centered on this point. Thus, a large number of patches from the stack of images are sampled (see Fig. 3.4). Patch size was selected such that myocardium pixels, pixels from the blood pool and/or air in the lung are included in the patch.

2. *fuzzification*
   To determine the location of the transition from blood to myocardium or from myocardium to air, gray values are first classified. To ensure modality independence, only relative gray value differences between blood, myocardium and air are used. This makes Fuzzy C-Means (FCM) clustering a suitable algorithm to distinguish between three classes, and use the class transitions as borders. The classes used are bright, dark, and medium bright, which represent blood pool, air and myocardium respectively. To obtain a balanced class distribution, FCM was applied to the intersection rings from all the images simultaneously.

3. *inference (see Fig. 3.5)*
   For each pixel, three fuzzy membership degrees (FMDs) result from the fuzzification, above. Based on these FMDs, the inference step looks like:

   - *for each pixel*
     if (gray value is bright) then pixel is blood pool if (gray value is medium) then pixel is myocardium if (gray value is dark) then pixel is air

   - *for each line*
     if (majority of pixels is class i) then line is class i else line is unclassified

   - *transition*

     endocardial border *from outside to inside, for all spatial vertex locations, find first transition from myocardium to blood pool*

     epicardial border *from inside to outside, at the septum, find first transition from myocardium to blood pool, at the lateral wall, find first transition from myocardium to air, at the inferior wall, interpolate septum and lung candidates*

The minimum membership degree for a pixel to be effectively classified is 0.5 for blood and air and 0.7 for myocardium. If a pixel does not have the required minimum membership value it is not classified (fuzzy border) and thus not considered for inference (see Fig. 3.5). The new candidate point resulting from this decision scheme is located at the transition found. If no position can be generated because of the lack of a proper transition in the image patch, the particular model point is updated using the result from neighboring patches in the model for which a decision could be made.
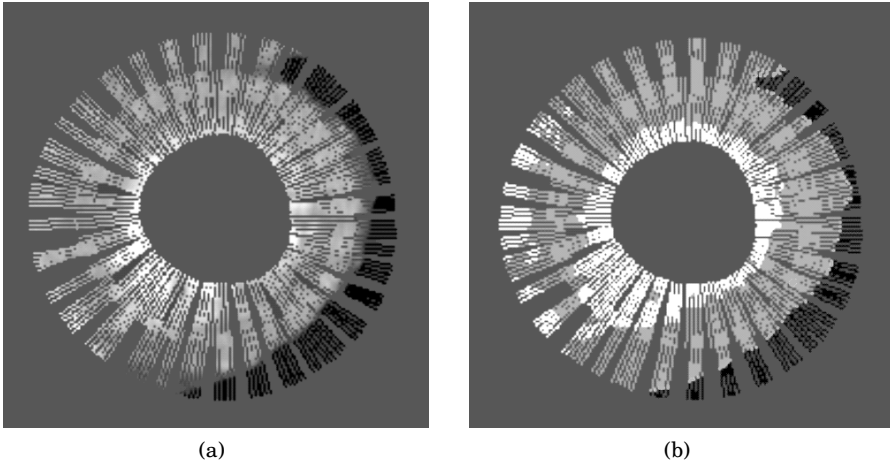
Figure 3.4: Ring of image patches from a cardiac short axis image slice. (a) The true pixel gray values. (b) Ring of image patches after classification by the FI decision scheme.
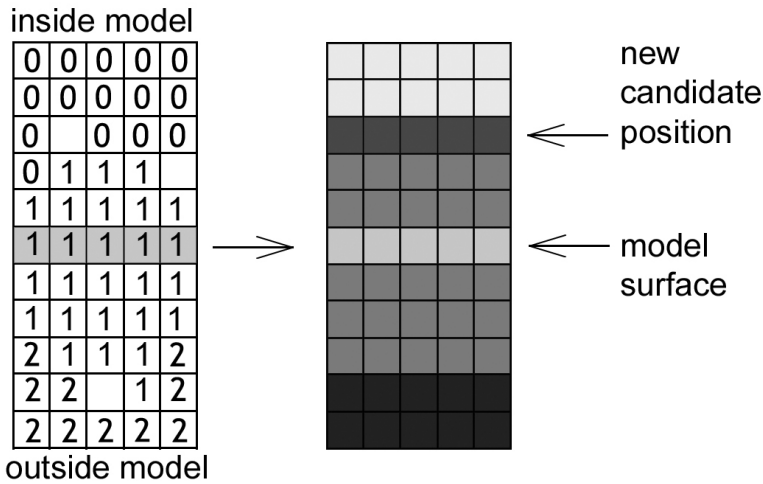

Figure 3.5: Left, an image patch with classified pixels is shown. The pixels are classified to blood (0), myocardium (1), and air (2) or none (empty) (part 1 of the inference step). The right patch shows classified lines according to part 2 of the inference step. The darker lines (below) are air, the medium bright lines are myocardium, and the bright lines are blood pool. In the patch, which is centered on the model surface (indicated with arrow), a new candidate position is found (also indicated).

3.3   Experimental Setup

To evaluate the proposed Fuzzy Inference method for updating the candidate points, we tested the 3D-ASM on cardiac CT data from 9 patients comparing both the simple convolution-based edge detection and the newly implemented FI-based method. For this, a statistical shape model was generated using expert drawn contours of a group of 53 patients and normals, from 3D MR data [60]. The shape parameterization presented in Section 3.2.1 was applied, where each sample was divided in 16 slices, each containing 32 points for the epicardial contour and 32 points for the endocardial contour. To reduce model dimensionality, the model was restricted to represent 99% of the shape variation present in the training data, resulting in 33 modes of variation.

The 3D ASM was applied to 9 short axis CT acquisitions of cardiac LVs. Scans were acquired with CT scanners from two different vendors, and had an axial slice thickness of approximately 1 mm and an in-plane resolution of 0.5 mm/pixel. All data sets were reformatted to yield short-axis image slices.

Prior to matching, the model pose was initialized manually. The initial model scale was equal to the average model scale of the training set. The model shape was initialized to the mean training shape, whereas position was manually initialized inside the cardiac LV. The class centers of the three tissue classes used by FCM were initialized identically for each iteration and for each patient. During model matching, deformation was limited by constraining each component of the model deformation parameter vector between -3σ and $+3\sigma$.

The model search ran for a fixed number of iterations, the same for both the FI-based model and the convolution-based model. For the FI-based model, a two-stage matching was employed: initially, the convolution method was used until the update step size between iterations substantially decreased. From there, the final adjustments, small scale and pose changes and deformation of the model were realized using the FI-based point generation. The model states from the last iteration for both models were used for comparing the two candidate point generation methods.

The method was visually evaluated to assess whether the new candidate point generation method is an improvement with respect to the convolution-based technique, by comparing results from the same iteration in the matching process. In case of matching failure, the match was reported as failure and excluded from further quantitative evaluation. In addition, the error in LV endo- and epicardial volume with respect to the volume from expert drawn contours was compared: these parameters are required for the clinical evaluation of cardiac function. The expert contours for evaluation were generated in the same manner as the contours used for model training, however, different experts were involved.

3.4   Results

In all 9 cases, the matching procedure converged to a stable solution. However, 7 out of 9 matches resulted in visually plausible contours, while 2 matches showed faulty candidate points, as a result of either data truncation or poor image quality. Reconstruction artifacts in the images, like the well-known stepping artifact, also caused faulty candidate point generation in these two cases. This occurred for both candidate point generation methods.

In the 7 accepted cases, the Fuzzy Inference method clearly outperformed the convolution-based technique: results from both candidate point generation methods for five pa-
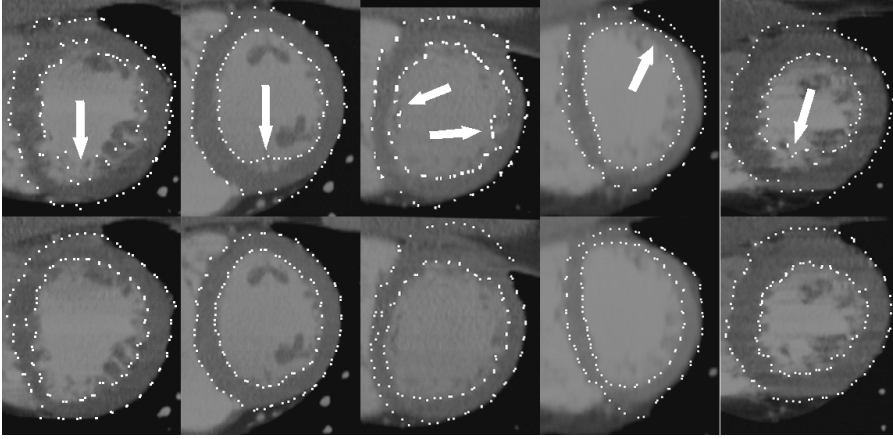
Figure 3.6: Image slices with candidate update points for five patients. The top row shows candidate positions generated with the convolution-based technique, the bottom row shows the same slices processed with the FI method. The columns show images from five different patients.

tients are shown in Figure 3.6. The final contours in all 7 successful cases were closer to the manually defined contours for the fuzzy inference method than for the convolution-based method (see Fig. 3.7). For the convolution-based method, the average error in the epicardial volume was 33 ± 41 cubic voxels, and -39 ± 29 cubic voxels for the blood pool volume, whereas for the proposed fuzzy inference method, these errors were 33 ± 21 cubic voxels and -0.66 ± 6.2 cubic voxels respectively.
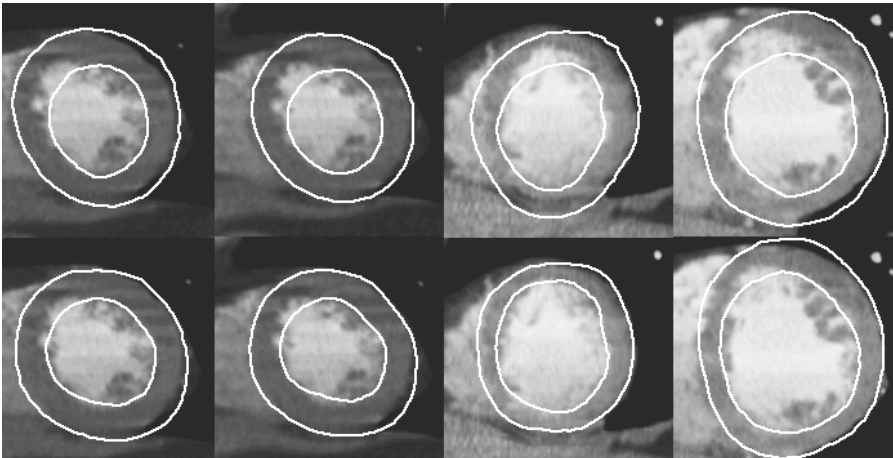

Figure 3.7: Image slices with automatically detected contours. The top row shows contours defined with the convolution-based point generation, while the bottom row shows those from the FI-based method in the same image slices.

3.5   Discussion and conclusions

In the 3D-ASM presented earlier [60], edge information was derived by a convolution of a simple edge template with various pixel strips extracted from image slices intersecting the model. This resulted in points located at the strongest edge in the neighborhood, while often the expert identified edge was characterized by a weaker transition. If the candidate edge point repetitively defines spurious model update positions, the model as whole may be misplaced, adversely influencing accuracy.

The main contribution of this work was the incorporation of a fuzzy-inference-based method for the generation of model update positions. We expected the combination of the FCM with knowledge-based rules to provide a more robust method for model update generation. The presented experiments show a considerably better performance of the FI-based method compared to the convolution-based technique. This applies to many structures in the images, like papillary muscle and trabeculations (see arrows in Fig. 3.6), and for multiple slices in the image stacks.

Quantitatively, the results of the presented pilot validation study show a substantial reduction in blood pool volume error (from -39 ± 29 to -0.66 ± 6.2 cubic voxels). The error in epicardial volume, however, did not decrease much compared to the convolution-based results, although its standard deviation decreased by approximately 50%. Further reduction of this systematic overestimation of epicardial volume may be achieved by decoupling the endo- and epicardial shape models, which is a topic of ongoing research.

In general, we conclude that the novel fuzzy-inference-based candidate point generation method is a definite improvement over convolution-based methods. Though further research is still required, the experiments presented here demonstrate a decrease in blood pool volume error, and visually better contour definition.