3D active shape modeling for cardiac MR and CT image segmentation
Assen, Hans Christiaan van

Citation

Version: Corrected Publisher’s Version
Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden

Downloaded from: https://hdl.handle.net/1887/4460

Note: To cite this publication please use the final published version (if applicable).
Hamlet: ‘Do you see yonder cloud that’s almost in shape of a camel?’
Lord Polonius: ‘By the mass, and ’tis like a camel, indeed.’
Hamlet: ‘Methinks it is like a weasel.’
Lord Polonius: ‘It is backed like a weasel.’
Hamlet: ‘Or like a whale?’
Lord Polonius: ‘Very like a whale.’

William Shakespeare (1564–1616), in Hamlet

3D-Active Shape Model Matching for Left Ventricle Segmentation in Cardiac CT
Abstract

Manual quantitative analysis of cardiac left ventricular function using multi-slice CT is labor intensive because of the large data sets. We present an automatic, robust and intrinsically three-dimensional segmentation method for cardiac CT images, based on 3D-Active Shape Models (3D-ASMs).

ASMs describe shape and shape variations over a population as a mean shape and a number of eigenvariations, which can be extracted by e.g. Principal Component Analysis (PCA). During the iterative ASM matching process, the shape deformation is restricted within statistically plausible constraints ($3\sigma$). Our approach has two novel aspects: the 3D-ASM application to volume data of arbitrary planar orientation, and the application to image data from another modality than which was used to train the model, without the necessity of retraining it.

The 3D-ASM was trained on MR data and quantitatively evaluated on 17 multi-slice cardiac CT data sets, with respect to calculated LV volume (blood pool plus myocardium) and endocardial volume. In all cases, model matching was convergent and final results showed a good model performance. Bland-Altman analysis however, showed that blood pool volume was slightly underestimated and LV volume was slightly overestimated by the model. Nevertheless, these errors remain within clinically acceptable margins.

Based on this evaluation, we conclude that our 3D-ASM combines robustness with clinically acceptable accuracy. Without retraining for cardiac CT, we could adapt a model trained on cardiac MR data sets for application in cardiac CT volumes, demonstrating the flexibility and feasibility of our matching approach. Causes for the systematic errors are edge detection, model constraints, or image data reconstruction. For all these categories, solutions are discussed.

2.1 Introduction

In medical image analysis and segmentation, deformable statistical models have proven to be highly successful and form an important field of research \[50, 51\]. Their success mainly stems from the a-priori knowledge contained in these models, ranging from shape and shape variation knowledge to gray level information.

In 1992, Cootes et al. introduced the Point Distribution Model as a description of the characteristic shape and shape variations of a set of example shapes \[11\]. In 1995 Cootes et al. further developed the PDM into a versatile trainable method for shape modeling and matching in the form of the Active Shape Model \[19\]. Active Shape Models model the characteristic shape and shape variations over a population of example shapes, and can be applied to image segmentation to limit the segmentation result to "statistically plausible" shape instances. In 1998, Cootes presented an extension of the ASM in the form of an Active Appearance Model, which is an extension with a statistical model of an intensity patch \[27\]. Both Active Shape Models and Active Appearance Models have found widespread application in 2D \[11, 21, 22, 27, 29\] and 2D + time \[52\] medical image segmentation problems.

However, applications so far for Active Shape Models (modeling + matching), have been mostly limited to those two categories, although Point Distribution Models (PDMs),
have been explored in 3D [15, 16, 53, 54], mainly for shape analysis of complex 3D shapes. Nevertheless, application of 3D Point Distribution Models to image segmentation is still largely unexplored territory.

The primary motivation of this work was to develop a segmentation method for the 3D cardiac ventricle that:

- treats segmentation in an intrinsically three-dimensional manner and exploits 3D continuity of LV shape,
- is applicable to various types of 3D cardiac image data, either from different modalities or acquisition protocols without retraining of the statistical model,
- can segment an LV data set using only a few images with different orientations, e.g., a four-chamber view, a two-chamber view and a short-axis view (see Fig. 2.1). This sparse data matching enables LV function analysis without the necessity of acquiring a large number of image slices and is a major reason for us to choose the ASM approach for this segmentation task.

In this paper, we present a 3D extension of the Active Shape Models that meets the requirements mentioned above. We focus on the development of a novel matching mechanism for 3D Point Distribution Models, with two important properties:

- the left ventricle shape is represented in 3D by mapping the 3D PDM to a mesh structure consisting of two surfaces, which can be intersected with planes of arbitrary orientation. These planes mimic image slices from which edge information is extracted to drive the model matching. By mapping these 2D position changes from the image slices to the 3D mesh points, the 3D shape and pose parameters are updated,
- the mesh used to represent the 3D model can be intersected by planes of any orientation, as mentioned in the first property above. The statistical model itself contains only shape information and is therefore modality independent. The edge detection technique applied to extract edge information from the 2D image slices is also independent of image modality. Therefore, the 3D-ASM is applicable to image data from different modalities and from different acquisition protocols.

The 3D-ASM was quantitatively evaluated with stacks of parallel short-axis images. In the same evaluation, we illustrate the model's independence of image modality with the direct application of the 3D-ASM, which was trained on cardiac MR data, to segmentation of multi-slice cardiac CT images. Moreover, the model was qualitatively tested on axial CT image stacks to demonstrate the validity of the concept of segmenting left-ventricular image data of arbitrary planar orientation.

The outline of the paper is as follows. In the next section, 3D model generation is treated, and the matching algorithm is discussed in detail. In Section 3, the cardiac 3D CT image data used for evaluation will be mentioned briefly, while Section 4 contains the results of the validation study.
2.2 Methodology

2.2.1 Model generation

Model generation in 3D involves three important issues, which are discussed in this section. First, point correspondence for 3-dimensional cardiac left-ventricular (LV) shapes is discussed. Second, shape alignment in 3D is addressed, and finally, statistical modeling of 3D shape and shape variation is discussed.

Point correspondence in the 3D cardiac left ventricle

An important condition for generating an Active Shape Model is the definition of point correspondence between shapes. Together with alignment, point correspondence is one of the most important problems of the extension of the ASM from 2D to 3D. Methods for defining generic point-correspondences between shapes currently form an active field of research [53, 55, 56]. However, because of the relatively simple shape of the 3D cardiac left ventricle, we adopted an application specific definition for point correspondence in the left ventricle.

Figure 2.2(a) shows the shape of a left ventricular endo- and epicardial surfaces and the parameterization that was applied. To define the parameterization representing a particular shape instance, each contour is sampled at equidistant angles with respect to the LV long axis. The sampling in each slice starts at the posterior junction of the left and right ventricle, which had been indicated by an expert. The coordinates of the samples of the endo- and epicardial contours in the stack of image slices together form one shape vector. The endocardial contours are sampled counter clockwise and from the apical to the basal slice, filling half the shape vector. The other half of the vector is filled by sampling the epicardial contours counter clockwise, from the basal slice towards the apical slice. This yields a line-parameterization with a specific point-to-point ordering.

After this line-parameterization is constructed for each left ventricle shape in the data set, the resulting point set is resampled in the direction of the long axis. Thus, each left ventricle shape in the training data is resampled to a fixed number of slices, and consequently to the same number of sample points. The latter is of critical importance because of the required point correspondence. Each shape sample can then be ex-
Figure 2.2: (a) Line parameterization, defining the point correspondence for the 3D cardiac left ventricle. (b) 3D triangular mesh constructed from the PDM. Each point is connected to its corresponding point in the next slice and to the previous point in the next slice, thus creating triangles.

pressed as a 3n element vector $\mathbf{x}$ containing $n$ concatenated 3-dimensional landmark points $\{(x_i, y_i, z_i)\}$ of a particular shape.

**Statistical shape modeling**

In order to minimize the effect of trivial variations in pose and scale in the statistical shape model, shape alignment is required. Shape samples are aligned using Procrustes analysis [57, 58], where all shapes in the training set are aligned using an iterative least-squares distance minimization. For the registration of corresponding point sets, we adopted Besl and McKay’s iterative closest points (ICP) algorithm for registration of 3D points sets [59]. This algorithm uses quaternions to represent scaling and rotation, assuming that 3D translation is represented in a separate pose vector.

After Procrustes alignment, the residual variation in shape in the training data set is modeled. The aligned set of shapes is transformed onto a basis of eigenvectors describing the shape variation in the set. The eigenvectors are calculated by applying a Principal Component Analysis (PCA) on the covariance matrix of the training set. A point in the eigenspace is a linear combination of eigenvectors representing a particular shape $\mathbf{x}$ in 3D, while the origin in the eigenspace corresponds with the average shape from the training set. Any shape synthesized from the model can be written as

$$\mathbf{x} = \overline{\mathbf{x}} + \Phi \mathbf{b}.$$  \hspace{1cm} (2.1)

with $\Phi$ a matrix containing the $m$ eigenvectors and $\mathbf{b}$ a $m$-dimensional ($m \leq N$) parameter vector controlling the deformation of the model.

Each eigenvector represents a certain proportion of shape variation, so that the total variance in the training data is the sum of all the eigenvalues. Modes corresponding with small eigenvalues contribute only little to the shape variation. To reduce model dimensionality, usually the number of axes in the eigenspace is limited. The number of allowed modes is determined by the proportion of variance in the training data that
2.2 Methodology

the model should represent. If $p_v$ is the required proportion of total variance $V_{total}$ present in the training set, the number of modes $k$ can be determined by

$$
\sum_{i=1}^{k} \lambda_i \geq p_v V_{total}
$$

(2.2)

2.2.2 Matching Algorithm

To match the 3D PDM to image data, a 3D triangular mesh from the points in the PDM was constructed by connecting neighboring points within one slice. To form triangles, each point is connected by an edge to the corresponding point in the next slice and to the point previous to the corresponding point in the next slice (see Fig. 2.2(b)).

The proposed matching procedure is based on distance minimization between two point sets, one representing the current mesh state and the other being a set of candidate boundaries resulting from image template matching in planar intersections of

![Diagram of the matching algorithm](image)

**Figure 2.3:** Schematic representation of the matching algorithm used in the 3D-ASM.

![Diagram of different intersection directions](image)

**Figure 2.4:** Different possible intersection directions for the LV model. Left, a short-axis intersection and right, an axial intersection.
2.2 Methodology

the triangle mesh. The second point cloud represents the target shape, to which the model is fitted within the statistical deformation constraints to guarantee plausible left-ventricular shapes. The matching consists of the following steps (see Fig. 2.3):

1. **Mesh intersection.** From the current state of the mesh, model sample points in the individual image planes are derived by dividing the mesh into two parts by the image plane. Each point (vertex) in the mesh is labeled to either side of the image plane. The in-plane contour points are generated by intersecting the mesh edges with differently labeled end points with the image plane. Endo- and epicardial sample points are generated simultaneously. Figure 2.4 shows two possible directions for intersection, a short-axis intersection and an axial intersection.

2. From each 2D intersection contour, a set of evenly distributed sample points is generated. For this, the center of gravity of the contour coordinates is calculated. Using this point as the center of rotation, the 2D contour is sampled with equal angle intervals, i.e. a fixed number of points per contour is generated.

3. For each sample point, generate a candidate boundary position. The candidate boundary position is estimated by maximizing cross correlation between a filter template (step filter) and a sampled line profile perpendicular to the contour. This yields a 2D displacement vector for each of the sample points (see Fig. 2.5(a)).

4. Projection of candidate boundary points on the mesh vertices. The candidate boundary points on the contours in the image slices do not coincide with the mesh points. Therefore, the point displacements in the image slices are propagated to the mesh vertices spanning the intersected triangle edge. Typically, displacements of more than one sample point contribute to the displacement of a mesh point. Multiple edges above and below any vertex can be intersected by image slices.

5. Align the in-plane displacement vector of each mesh point to the average 3D normal vector of the neighboring triangles in the mesh.

6. Align the current state to the proposed state of the mesh, using the ICP alignment used during model generation [59].

7. Calculate shape differences between the aligned states of the mesh, and successively adjust the parameter vector $b$ controlling the model deformation:

$$b = \Phi^T(x_{\text{proposed}} - x_{\text{current}})$$

with $x_{\text{proposed}}$ the vector representing the proposed shape, and $x_{\text{current}}$ the vector representing the aligned current state of the mesh. Thus, the mesh in the current state is deformed to optimally match the new shape. This deformation is constrained within the bounds of the statistical description, i.e. within a fixed number of standard deviations from the average (non-deformed) model.

Repeat until convergence.

During the matching phase, in step 3, template matching between a sampled in-plane profile perpendicular to the intersection contour and a filter template was used to
generate a candidate position for each point in the mesh. A simple step filter profile of nine pixels served as a filter template. To find the epicardial wall at the septal wall, a rising edge filter template was used, whereas all other regions, a descending edge filter was employed. The length of the sampled in-plane profiles was fixed to 21 pixels. Matching was performed in two steps: a Euclidian transformation only in the first few iterations, followed by simultaneous shape and pose adaptation. This was motivated by the observation that the model is not yet close to its final position if it still undergoes large pose changes. This way, the risk of convergence to neighboring structures can be decreased. In case no clearly defined edge could be detected along a sampled in-plane profile, the new position of the sample point is estimated by interpolation between reliable neighboring edge results.

2.3 Experimental setup

2.3.1 Training data

To assess the clinical potential of the presented ASM matching, an ASM was generated using expert drawn epicardial and endocardial contours of the cardiac left ventricle of a group of 53 patients and normals, from 3D MR data. The basal and the apical slices of the ventricle in the model were selected as the last slice in which still both an endo- and epicardial contours were drawn by the expert. Images were 256x256 pixels; field of view: 400-450 mm; pixel sizes: 1.56-1.75 mm.

To define point correspondence between the training samples, the parameterization presented in Section 2. 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. This resulted in 1024 points for the model.

In order to reduce model dimensionality, the model was restricted to represent 99% of the shape variation present in the training data, resulting in 33 modes for statistical shape deformation description.

2.3.2 Evaluation data

The 3D-ASM performance was tested on 17 CT acquisitions of left ventricles. The number of slices in these acquisitions ranges from 15 to 39 of which between 12 and 24 contain left ventricular data. Of the 17 data sets that were used, six were acquired using a Toshiba Acquilion 4-slice CT-scanner, and had an axial slice thickness of approximately 1 mm and an in-plane resolution of 0.5 mm/pixel. Eleven data sets were acquired using GE Lightspeed 4-slice CT scanners with a slice thickness of 1.25 mm and an in-plane resolution of approximately 0.5 mm/pixel. All data sets were reformatted to yield short-axis image slices. No normals were included for evaluation; all data sets included in the evaluation population were acquired from patients.

2.3.3 Model matching parameters

Prior to matching, the model pose was initialized manually. For each patient, an initial model scale and the mid-ventricular slice were indicated manually. The model shape was initialized to the mean training shape. In case of matching failure, initialization was interactively repeated by application of an offset in the starting position or orientation. In case of repeated failure, the match was excluded from further quantitative
evaluation.
Throughout most of the model surface areas, the step edge filters for the endocardial surface were defined the same as those for the epicardial surface: a descending edge when traversing the surface from inside to outside. Only at the septal wall, the epicardial edge filters were defined with opposite edge transitions. To avoid epicardial edges in the image attracting the endocardial surface of the model and vice versa, a-priori knowledge about the gray levels of the several tissues was utilized to reject implausible edge candidates: the average gray value of the blood pool was required to be higher than a certain Houndsfield threshold. The average gray value of the lung was required to be lower than another certain Houndsfield unit.
During the model matching to the patient data, model deformation was limited by constraining each component of the model deformation parameter vector between -3s and +3s. The ASM as described here did not have a stop criterion indicating that convergence had been achieved. During evaluation, the model search ran for a fixed number of iterations. The model state in the last iteration was included in a quantitative comparison with manually drawn expert contours.

2.3.4 Quantitative evaluation

To quantitatively evaluate the model, volumes from the endocardial and epicardial contours were calculated using Simpson's rule. Areas were calculated enclosed by the contours that resulted from the model in mm$^2$ and summed these over the slices included in the evaluation. Slices were included in the evaluation if the model was able to produce a full contour for that slice. Slices for which the model produced only partial contours (see Fig. 2.5(a), bottom-center) were excluded from the quantitative evaluation procedure. From the calculated automatic and manual epicardial volumes (LV volume) and endocardial volumes (blood pool volume) regression formulas were calculated. Bland-Altman analysis of the LV (i.e. blood pool plus myocardium) volume and the blood pool volume was performed to determine whether segmentation results of the model show a systematic error.

Figure 2.5: (a) All contours for this patient generated by the 3D-ASM shown in the 3D cardiac image data set (short-axis). (b) Contours generated by the 3D-ASM for an axial stack of 3D cardiac image CT data.
2.4 Results

In all 17 cases, the matching procedure resulted in visually plausible contours. No match was excluded from quantitative evaluation, according to the exclusion criterion mentioned above. Figure 2.5 shows representative examples of plausible contours resulting from the 3D-ASM in short-axis and axial data sets respectively. From the calculated automatic and manual epicardial volumes (LV volume) and endocardial volumes (blood pool volume) regression formulas were calculated. For epicardial volume an excellent correlation was found: $y = 1.02x + 16.4 (R = 0.99)$, with $y$ denoting the LV volume calculated from automatic contours and $x$ the LV volume calculated from expert contours. For the endocardial volume, we found the same excellent correlation factor: $y = 0.88x + 1.4 (R = 0.99)$, with $y$ the endocardial volume calculated from automatically generated contours, and $x$ the blood pool volume resulting from manually drawn contours. Bland-Altman plots for endo- and epicardial volume are shown in Figure 2.6. The regression formulas and the Bland-Altman plots reveal that there is a slight systematic underestimation of the blood pool volume and a slight systematic overestimation of the epicardial volume by the model.

2.5 Discussion and conclusions

In this paper, the feasibility of a 3D-ASM for automatic segmentation of the cardiac left ventricle from MSCT image data was investigated. For all subjects in the quantitative evaluation, the 3D-ASM converged. Visually plausible contours for the endo- and epicardial transitions in all image data sets were generated (see Fig. 2.5). The model produced the visually plausible contours for short-axis data sets as well as for
2.5 Discussion and conclusions

Figure 2.7: (a) An example where model limitations prohibit the model to deform according to the edge-information in the image (arrow). Edge-detection resulted in good candidate points, but the model surface was unable to reproduce local myocardial thinning. (b) Contours that the same model generated for a mid-ventricular slice of the same study are shown. This shows that this model shape limitation can be a local issue.

Axial data sets.
Quantitative evaluation from the short-axis data sets however, shows that ED LV volume (blood pool plus myocardium) is systematically overestimated by the model, whereas ED blood pool is systematically underestimated (see Fig. 2.6). This can also be seen from the regression formulas presented for both volumes. One of the reasons for this may be the lack of apex data in the model. The apex is a well recognizable landmark of the cardiac left ventricle, and therefore including it in the statistical LV shape description may result in a more representative model. Missing the apex in the model is a cause for uncertainty about the position of the model in the direction of the LV long axis.

In some cases however, generation of an accurate segmentation failed. This can have a number of causes, which can be divided in three categories:

- **image data**
  In three data sets, image data was truncated close to the LV frontal side due to reformatting. The truncation itself showed to be a very strong edge, attracting and thus misleading the model epicardial surface. In the short-axis test set of another patient, expected 3D continuity of the data set was not present, because of an in-plane shift of one slice with respect to the rest of the stack. If a stack of images has slices that are shifted in an in-plane direction the model may not be able to follow the surfaces. The 3D model requires 3D continuity between slices in the image data set and thus provides 3D continuity in the contours that result from the 3D segmentation process. To resolve these issues, image reformatting should be performed with care.

- **model limitations**
In one patient, the myocardium showed drastic local thinning, that the model was unable to reproduce (see Fig. 2.7). The deformation proposed as a result of the edge-detection could not be accommodated by model. This may be caused by a lack of representative sample shapes in the training data set. Secondly, because of missing apical data in the model, uncertainty in the position of the model (closer to the base, or more towards the apex) arises. If the model is positioned too close to the base or apex, local deformations of the model will not match local shape characteristics any more. Incorporation of apical data in the model and extending the training data set with additional patient data should solve these issues.

- edge-detection

Edge-detection yielding candidate edge positions at transitions between fat and air instead of transitions between myocardium and epicardial fat, causes a systematic overestimation of LV volume. Moreover, uncertainty about the best values to choose for the average blood pool gray value and the lung gray value influences edge-detection accuracy. This is a common tuning problem. And thirdly, in 3 patients, an edge between epicardial fat and lung (air) attracted the model epicardial surface more than the edge between myocardium and epicardial fat. The model deformed to fit the epicardial fat, thus yielding overestimation of the total LV volume. Therefore, for the edge-detection algorithm, we suggest adding a preprocessing step that distinguishes between tissue types before the 3D-ASM is applied. This would also support the objective of a modality independent model.

The ultimate goal of this work is the application of the model to a sparse data set, i.e., a small number of image slices that are not necessarily parallel, as shown in Figure 2.1. That however, requires a weighing scheme while updating the model’s pose and shape parameters, since the sparse data set will cause a large amount of model points not to be updated. Only sample points close to model intersections by the image slices are updated with the help of edge information. Other sample points are not updated in between iterations, which means that they tend to retain their position. These points will have to be discarded while updating the model parameters, i.e. they have to be assigned a weight equal to zero.

Once this weighing scheme is implemented, it can also be used to assign weights to sample points according to the strength of the edges that attract these points. Thus, a strong edge has more influence than a weak edge, and points that have no edge in their vicinity at all can be assigned a weight zero, as with the sample points missing image data in the sparse data application.

In the current evaluation, the model was initialized manually. In the future, we want to further automate the model matching as much as possible, so initialization should also be completely automatic. Due to this automation, the risk of initialization further from the final segmentation result will be increased. In order to reduce model dependency of the initialization, the model can be allowed to broaden its view in the initial matching process, and reducing its view when the matching process converges towards its final result.

In conclusion, in this paper we demonstrate the feasibility of a modality independent 3D-ASM. Based on the quantitative validation presented here, the 3D-ASM combines robustness with clinically acceptable accuracy. Without labor intensive retraining on cardiac CT data, we could adapt a model trained on MR data sets for application in cardiac CT data volumes.