Automating 3D Echocardiographic Image Analysis

Gerardo I. Sanchez-Ortiz and Jérôme Declerck, Miguel Mulet-Parada, and J. Alison Noble

Medical Vision Laboratory, Department of Engineering Science, University of Oxford, Parks Road, Oxford OX1 3PJ, UK.
giso@robots.ox.ac.uk
http://www.robots.ox.ac.uk/∼mvl

Abstract. 3D echocardiography is a recent cardiac imaging method actively developed for quantitative analysis of heart function. A major barrier for its use as a quantitative tool in routine clinical practice is the absence of accurate and robust segmentation and tracking methods necessary to make the analysis fully automatic. In this article we present a fully-automated 3D echocardiographic image processing protocol for assessment of left ventricular (LV) function. We combine global image information provided by a novel multi-scale fuzzy-clustering segmentation algorithm, with local boundaries obtained with phase-based acoustic feature detection. We fit and track the LV surface using a 4D continuous transformation. To our knowledge this is the first report of a completely automated method. The protocol is viable for clinical practice. We exhibit and compare qualitative and quantitative results on three 3D image sequences that have been processed manually, in semi-automatic manner, and in fully automated fashion. Volume curves are derived and the ejection fractions errors with respect to manual segmentation are shown to be below 5%.

1 Introduction

The last few years have seen the emergence of 3D echocardiography acquisition systems in the market. Methods of acquisition are improving (in terms of spatial and temporal resolution), moving now towards real-time volumetric acquisition. However, interpretation and analysis of the data is more complex and time consuming than for conventional 2D echocardiography. As recent research studies have shown [1,2,3], the use of three-dimensional data provides more precise information on the pathophysiology of the heart than conventional analysis of 2D views ([4,5,6] and references therein), especially for volume and ejection fraction calculation.

Previous work has shown the feasibility of reconstructing a three-dimensional surface of the heart from sparse views [2]. However, in that work, the amount of interaction required to obtain a reconstruction of the endocardium was very large. In [1], a 3D finite element mesh of the left ventricular (LV) myocardium is computed and used to perform strain analysis. The approach is interesting,

S.L. Delp, A.M. DiGioia, and B. Jaramaz (Eds.): MICCAI 2000, LNCS 1935, pp. 687–696, 2000.
© Springer-Verlag Berlin Heidelberg 2000

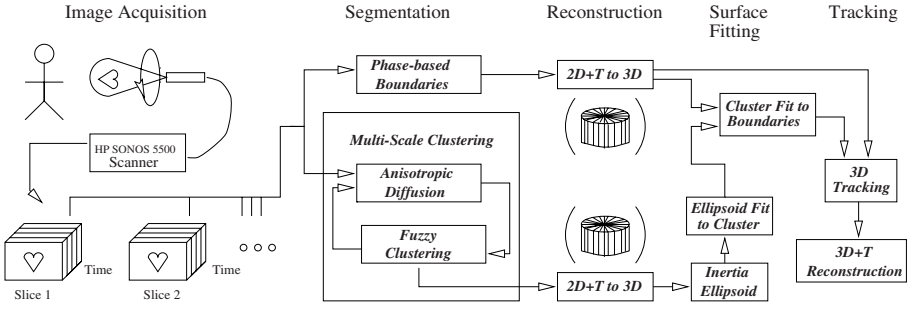


Fig. 1. Diagram shows the process for automatic analysis of 3D echocardiography.

however the analysis depended on the high quality images acquired from open-chest and is obviously not viable for routine protocol.

In this paper we present a framework for fully-automated 3D transthoracic echocardiographic image analysis, viable for routine clinical practice. The image data is acquired using a transthoracic 3D probe, the final result of the processing is a segmented left ventricular cavity over the whole cardiac cycle. No user interaction is necessary for this process. Previous work [3] from our laboratory has shown that it is possible to perform automatic tracking of the LV cavity, given a good segmentation of the image sequence and a correct initialisation for the tracker. This paper addresses especially these two issues. First, we initialise a 4D surface tracker [7] using a surface estimate of the LV cavity. The cavity surface estimate is obtained automatically with a novel method for multi-scale fuzzy-clustering based on combined ideas from [8,9]. We then use the tracker to follow LV features detected with an acoustic-based edge detector for 2D cardiac sequences [6]. In the following sections we present the details of the method, and the experiments which have been conducted on three human healthy subjects. Left ventricular cavity volume curves are computed and compared to those of a manual segmentation performed by an expert.

2 Materials and Methods

The method consists of the following steps: first, feature points are extracted from the image to get a precise localisation of the endocardial boundary. Secondly, an approximate LV cavity is estimated using an iterative process of anisotropic diffusion and fuzzy clustering. Finally, a series of matching steps are required to a) generate a model of the initial endocardial surface and b) map this model through the cardiac sequence. Successful matching over the sequence requires a good initialisation shape, which is provided by the first matching. The procedure is schematised in Figure 1 and detailed in the following.

2.1 Image Acquisition

Digital 3D echocardiographic data was acquired on a HP SONOS 5500 ultrasound machine using a 3-5MHz rotating transducer, and stored as a sequence of

2D echograms (2D+T image), one for each probe angle (see Figure 2 for examples of 2D+T images). For normal studies we used 12 coaxial planes, one every 15 degrees (see Figure 3(a)); for one "dense" dataset we acquired 60 slices, one every 3 degrees. Data was acquired at a frame rate of 25 frames per second, the pixel size being 0.5mm x 0.33mm. Scanning was performed using ECG and respiration gating, on an apical view (*i.e.* the probe was located at the apex and roughly aligned with the LV long axis).

2.2 Phase-Based Image Features Extraction

A phase-based spatio-temporal feature detection method is used to find candidate endocardial border points. An early implementation of the method used is described in [6] and the performance of the technique is tested in detail on clinical data in [10] within the context of 2D echocardiographic image tracking. Briefly, the idea is to detect endocardial border points according to their phase signature (edge shape) rather than intensity gradient information. The reason for this choice is that the acoustic reflection from the endocardial border varies according to the relative angle between the boundary and the transducer. The net effect is to produce a border with variable intensity contrast around its length. This makes it difficult to detect the endocardial border using an intensity-gradient based approach. This problem is even more of an issue in 3D echocardiography, since some 2D acquisition planes might be non-optimal because they are determined automatically once the first plane position has been chosen. In [6], feature asymmetry is proposed as an alternative measure for endocardial border detection and a spatio-temporal (2D+T) version of this idea is developed.

2.3 Scale-Space Clustering for LV Region Extraction

In order to have an automated surface fitter that reliably finds and tracks the LV wall without human intervention we need boundary points to be present at reasonable intervals all over the LV surface. Because ultrasound images have a very low signal-to-noise ratio in regions of the LV wall parallel to the insonification beam, obtaining boundary points all over the endocardium is difficult even with the phase-based method. However, this problem can be overcome if the fitter is initialized with a surface that lies reasonably close from the target LV surface.

In order to extract an estimate of the wall in every region of the LV cavity, we use a multi-scale fuzzy clustering algorithm as a complementary segmentation method. This algorithm does not rely exclusively on the local differential structure of the data but takes into account the global characteristics of the image. In this way, a continuous approximation of the LV cavity boundary is provided even in regions of the images with low contrast and low signal-to-noise ratio. This also diminishes the effect of outliers detected by the phase-based method that correspond to noise and other anatomical structures. The results of the clustering method are then used to initialize the surface fitter in the manner described in Section 2.4.

Fuzzy Clustering We depart from Bezdek's fuzzy c-means clustering algorithm [11] which uses information about image attributes (like intensity) to divide the image domain into a pre-determined number of regions (clusters) and assigns every pixel in the image a degree of membership to the clusters, *i.e.* a probability of belonging to each of the regions. The algorithm groups the attributes by iteratively approaching a minimum of an energy function that measures the dissimilarity between the pixel attributes and those attributes of the cluster centres of each region (the energy function is typically a distance in the attributes' vector space).

The image attributes we use are intensity and position in an elliptic-cylindrical coordinate system, which is a natural choice for the 2D+T LV long-axis images (in [8] we investigated a similar approach for 3D MR images, with an elliptic coordinate system). The attributes are rescaled to homogenize their value ranges in the attribute space.

The origin of the elliptic-cylindrical coordinate system is first placed in the centre of the image (placing the left ventricle near the centre of the imaging window is therefore an image acquisition requirement, although anyway this is normally the case given that the left ventricle is the imaging subject), and then the position is refined by computing the centre of mass of the LV cluster. The cluster corresponding to the LV cavity is automatically identified as the one with the lowest intensity and position. In case some pixels belonging to this cluster are scattered around in the image (which rarely happens after the image has been smoothed as described in the next section), the largest connected component is computed to select only the points of the cluster belonging to the LV cavity.

None of the only two parameters to which clustering could be sensitive proved to be crucial to get a correct estimation of the LV cavity. These two parameters are, the number of clusters, and the geometry-intensity weight, which determines how to weight the geometric position with respect to the image intensity in the attributes space. For all the datasets (around 100 2D+T images corresponding to 3 studies), using 6 clusters and a weight of 1 gave satisfactory results. Small variations to these values only modify the quality of the segmentation results (*i.e.* how close the cluster approximates the LV cavity, but they do not, for instance, fuse the cavity and the background clusters). If precise identification of the LV is required, these values can be useful to weight geometry v. intensity in images with poor contrast or boundaries definition. However, a cluster that approximates the LV cavity is sufficient for our protocol.

Anisotropic Diffusion for Scale-Space Generation In order to overcome the problematic effect of intensity fluctuations of the noisy ultrasound images, the clustering process is performed at different levels of resolution in a scale-space of the image [12]. The scale-space is generated using the knowledge-based anisotropic diffusion (KBAD) algorithm [9]. Anisotropic diffusion algorithms smooth the image intensity (I) while preserving sharp edges by using the heat diffusion equation $\frac{\partial I(x,y)}{\partial \tau} = \nabla \cdot (c(x,y) \nabla I(x,y))$, where τ is the diffusion time (related to the scale of resolution) and the conductance c is normally a monotonically decreasing function of the magnitude of the intensity gradient

($c = c(\|\nabla I\|)$). In the KBAD scheme the conductance term is a tensor and an explicit function of the the position $\mathbf{p}$ (in 2D+T in our image domain), the image intensity and its gradient, *i.e.* $\mathbf{C} = \mathbf{C}(\mathbf{p}, I, \nabla I)$. It can therefore incorporate *a priori* and *a posteriori* information of the geometric and dynamic characteristics of the image. It can also be used to introduce a probabilistic measure of the image intensity distribution [8,13], as explained in the next section.

Multi-scale Fuzzy-Clustering The KBAD scheme used for generating the scale-space gets feedback from the clustering in progress. The fuzzy classification of the image domain provides a measure of the *a posteriori* probability that neighbouring pixels belong to the same tissue type, and is therefore incorporated into the diffusion process by means of the conductance function, penalizing or encouraging diffusion between pixels depending on the probability of them to belong to the same cluster. *A priori* knowledge about the system is also introduced by choosing the elliptic-cylindrical coordinate system for the clustering attribute space.

The clustering is updated at regular intervals during the diffusion process (*i.e.* at different levels of the scale-space), and the initially coarse segmentation of the image is gradually improved until it converges to a meaningful region partition, as the smoothing action of the diffusion process clears the image from noise. The first clustering is done after some iterations of the diffusion, then repeated at regular intervals. The computational expense of repeating the clustering at different scales of resolution is not high because energy minimisation is faster in the lower dynamic range of the smoothed image. Since the segmentation does not have to be precise, the process is performed on subsampled images (reduced by a factor 8) making processing time shorter (a 2D+T image can, for instance, be processed in under half a minute in a O2 SGI work station).

The combined diffusion-clustering algorithm penalizes inter-cluster diffusion and encourages intra-cluster diffusion, resulting in homogeneous intensity clusters with high contrast between them. The two driving mechanisms of the diffusion process are, on the one hand, the gradient based function governed by the differential structure of the image intensity function, and on the other hand, the intensity and spatial based clustering through which knowledge has been introduced by the elliptic-cylindrical symmetry of the image. Since the clustering scheme uses non-local information in order to perform the classification, the diffusion process is enriched with information about the global characteristics of the image.

2.4 Surface Fitting and Motion Tracking

The phase-based feature points obtained for each of the 2D+T images are reconstructed in 3D (*i.e.* slices placed coaxially). The LV cluster is reconstructed in 3D in the same manner, and then used to create an ellipsoid of inertia by computing the eigenvectors of its mass (points) distribution. The surface of the LV cavity at end diastole is obtained in three steps: i) reconstruction of the ellipsoid of inertia; ii) matching of this ellipsoid surface to the cluster points in order to

get a first approximation of the shape of the cavity; and iii), matching of the latter to the feature points detected with the feature asymmetry algorithm.

This LV cavity shape is then sequentially fitted to the feature points at all time frames. A 3D matching is initialised with the result obtained at the previous time frame. Each of these matching steps is processed using a variant of the method described in [7]: a surface S_1 in an Image1 (either the ellipsoid, the manually segmented shape or the cluster shape) is deformed using B-spline tensor products to some feature points extracted from an Image2 (these being the boundary of the cluster or the points obtained as described in the segmentation section). The geometric transformation $\mathbf{f}$ that maps the points in Image1 to Image2 is such that the cost function defined by $C(\mathbf{f}) = \sum_{\mathbf{M_i} \in S_1} d[\mathbf{f}(\mathbf{M_i}) - CP_2(\mathbf{f}(\mathbf{M_i}))]^2$ is minimized, where $\mathbf{M_i}$ is a feature point from S_1, and the function CP_2 finds the closest feature point in Image2. d is a distance measure between two points. Computations of the local B-spline transformation is iterated until convergence is achieved. Details of the subset selection and minimisation process are outlined in [7]. As a final refinement, the control points of the 3D B-spline tensor products that define the deformations at each time frame are interpolated over time using a periodic temporal B-spline. The final deformation of the surface over time is therefore a continuous 4D (3D+T) tensor product of B-splines from which valuable dynamic information can be computed.

3 Results and Discussion

3.1 Phase-Based Feature Points Extraction

An example of feature detection on a slice at one time frame is shown in Figure 2(b), and a representation of the detected points over a spatio-temporal sequence is shown in Figure 2(c). Points are generally well defined over most of the endocardium, but a significant amount of spurious edges belonging to other anatomical structures or produced by noise and image acquisition artifacts are also detected and might perturb the motion tracking process. For instance, in Figure 2(g) we can see some boundary outliers in an image with a strong intensity inhomogeneity at the centre of the LV cavity (the inhomogeneity is due to a bad gain setting during image acquisition). For these reasons a careful initialisation of the surface fitting is essential to avoid tracking boundary outliers.

3.2 LV Cavity Cluster

The 2D+T sequence is partitioned into clusters, as shown in Figure 2(d). The boundary of the central cluster (which corresponds to the LV cavity) is shown in Figure 2(e) and (f). Only the end-diastolic cluster is used for initialisation, but the clustering has been performed in 2D+T for more robustness. In Figure 2(h) we can see the clustering results on the image with the inhomogeneity artifact shown in Figure 2(g), and in Figure 2(i) we can see volume rendering results using maximum intensity projection of the same LV cluster in 2D+T space. Notice

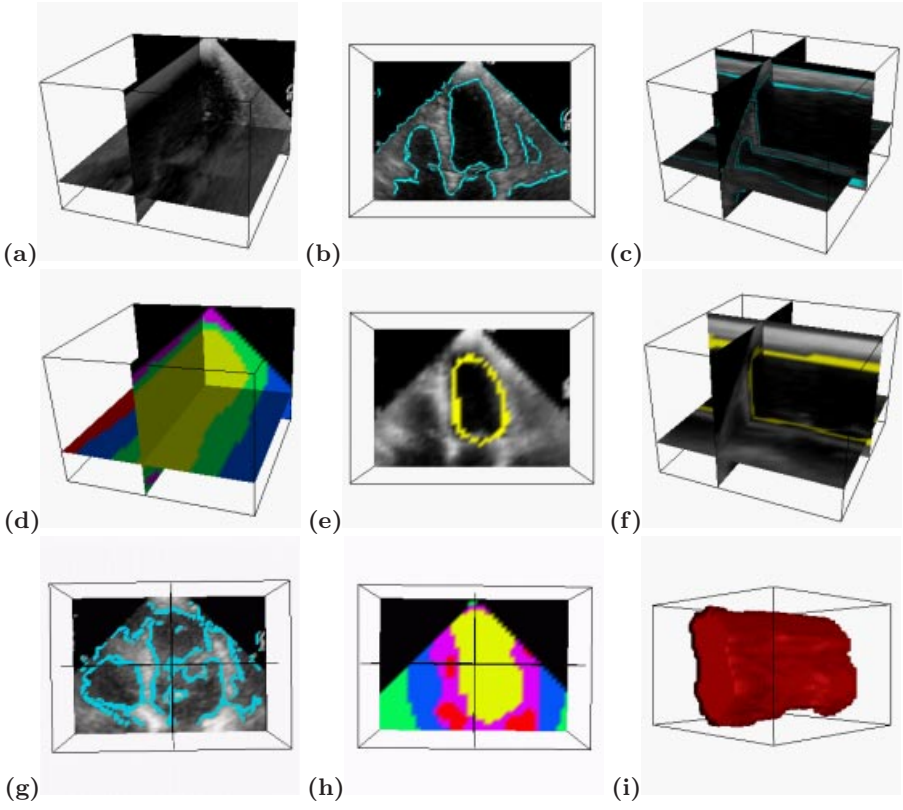
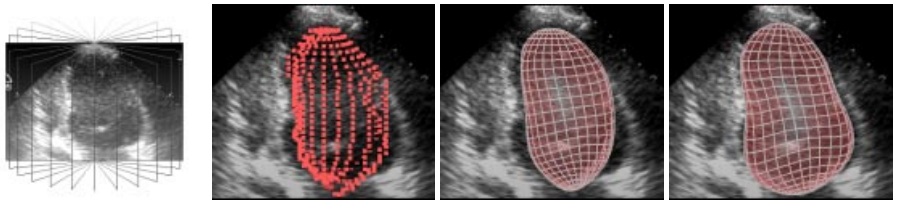


Fig. 2. **(a)** Example of original 2D+T image (depth is time). Feature detection overlaid on the original 2D+T image, at a particular time frame **(b)** and over the sequence **(c)**. **(d)** Clustering results (depth is time), showing the LV cavity cluster in yellow (light grey). The boundary of the LV cluster is overlaid on the subsampled image at a particular time frame **(e)** and over the sequence **(f)**. **(g)** Example of low quality ultrasound image with phase-based edges overlaid. **(h)** Clustering results of the same image (showing that the LV cluster is a good approximation to the cavity), and **(i)** volume rendering (maximum intensity projection) of the LV cluster in 2D+T space.

that the clustering algorithm works fine in spite of the intensity inhomogeneity (because it uses geometric as well as intensity information).

Once the 2D+T clusters have been computed for all planes of the image (all probe angles as shown in Figure 3(a)), the clusters are reconstructed in 3D (Figure 3(b) shows the boundaries of the reconstructed 3D cluster). The ellipsoid of inertia is computed from these points. This computed ellipsoid is then deformed onto the cluster points as shown in Figure 3(c). Finally, this deformed shape is deformed again onto the feature asymmetry points (Figure 3(d)). This last surface defines the shape of the LV cavity which is used for motion tracking.



(a) **(b)** **(c)** **(d)**

Fig. 3. **(a)** A 3D dataset with 12 planes. **(b)** 3D reconstructed boundary of the LV cavity cluster. **(c)** Cluster's ellipsoid of inertia deformed onto the cluster. **(d)** Surface from **(c)** deformed onto phase-based edges.

The reconstruction was robust on the three datasets processed. It provided a good initialisation for the tracking in all cases: the maximum distance between the surface and a manually segmented surface was around 8mm for all data.

3.3 Motion Tracking Using Dense vs. Sparse Data Acquisition

The effect of using sparse data on our protocol was investigated (previous results from our laboratory [3] had indicated that volume estimation from sparse data is feasible provided a good initialization for the tracker). We used the dense dataset (see Section 2.1) and selected 60 (full resolution), 30, 12 or 6 of the original planes. Tracking was performed using the automated method on each of these 4 sets of planes, and LV volumes over the 3D sequence were computed. The maximum volume variation between the sets at any given time was below 6% of the volume. In order to be consistent with the other acquisitions, all other results using this dataset (patient (1)) were obtained using only 12 planes.

3.4 Comparing User Guided and Fully Automatic Processing

Four tracking processes are compared on each of the three datasets: a) Manual segmentation performed by an expert for the complete 3D sequence; b) automatic tracking using manual segmentation of the end-diastolic surface as initialisation; c) automatic tracking using as an initial surface an ellipsoid manually placed by an expert with the help of a 3D graphic interface; and d) the fully automatic method, *i.e.* initialised using clustering as described in Sections 2.3 and 3.2. Figure 4 shows the volume curves obtained with these four methods on the three datasets. If the initialisation is not sufficiently close to the target LV cavity (method (c)), the tracking fails and the volumes are meaningless. On the other hand, the trackings (b) and (d) give similar results, showing that clustering based automatic initialisation gives as good results as initialisation from manual segmentation.

While initialising the tracker with a surface close to the actual LV reduces errors due to the sparsity of the data (see [3]), our results show a systematic

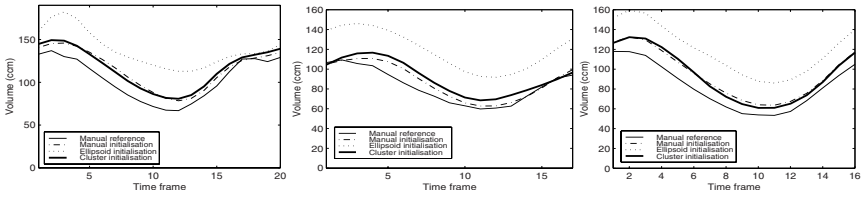


Fig. 4. The plots show volume curves for three different patients. In each of them, the full line shows results of manual segmentation. Using the automatic tracking, three initialisation shapes were used: a manual segmentation (dashed lines), a manually selected ellipsoid (dotted lines), and the automatic cluster's inertia-ellipsoid deformed to the cluster (thick lines).

Fig. 5. The table shows the ejection fractions computed for the three patients using all four reconstruction methods described in relation to Figure 4.

	Ejection Fraction (%)		
Method	Patient (1)	Patient (2)	Patient (3)
Manual segmentation	51	45	54
Manual initialisation	46	43	52
Ellipsoid initialisation	38	37	46
Cluster initialisation	46	41	54

overestimation of the volume. The main reason is that endocardial boundaries are not clearly defined in some regions of the lateral wall, often shadowed by the rib cage. In these regions the tracker followed spurious edges located outside the cavity. However, the computed ejection fractions are, in the case of the fully automatic method, less than 5% over the value obtained from the manual segmentation (see table in Figure 5).

4 Conclusions and Future Work

We demonstrated the feasibility of an automatic method to track endocardial boundary from 3D transthoracic ultrasound data. We first use *multi-scale clustering* to provide an approximate segmentation of the LV cavity. This estimated shape is used to initialise the *3D fitter and tracker*, which then follows the precise boundary candidates obtained with the *phase-based* method. Using local and global information from the 3D image sequences, the combination of these three techniques overcomes the problem of tracking sensitivity to the initial shape: we obtained very similar tracking results using a manually defined surface and the automatically defined one. Ejection fractions computed with the fully automatic method were less than 5% over the value obtained using manual segmentation by an expert. No user interaction is necessary for this process and the technique is viable for routine clinical practice.

In order to improve the LV surface tracking (and volume estimate) we are currently developing a filter based on the clustering results, to remove outliers from the endocardial boundary candidates. Also, normal vectors to the edges provided by the feature detector are being used as additional information for tracking. We have also investigated segmenting the reconstructed image in 3D space (instead of 2D+T) using an elliptic coordinate system for clustering and a 3D version of the feature detector. Results show that the method could also be used on 3D images obtained with a different acquisition protocol (*e.g.* using a real time 3D transducer-array).

Results from a clinical study comparing volumes obtained with our protocol and those from SPECT Multi-gated Acquisition (MUGA) will be available by the time of the meeting. This research is funded by grants from the EPSRC (grant GR/L52444) and the MRC (grant G9802587).

References

[1] X. Papadimetris, A. Sinusas, D. Dione, and J. Duncan. 3D cardiac deformation from ultrasound images. In *MICCAI*, pages 420–429, Cambridge, UK, Oct 1999.

[2] F. Sheehan, E. Bolson, R. Martin, G. Bashein, and J. McDonald. Quantitative three-dimensional echocardiography: methodology, validation and clinical applications. In *MICCAI*, pages 102–109, Boston, MA, USA, Oct. 1998.

[3] G.I. Sanchez-Ortiz, J.A. Noble, G.J.T. Wright, J. Feldmar, and M. Mulet-Parada. Automated LV motion analysis from 3D echocardiography. In *MIUA*, Oxford, UK, July 19-20 1999.

[4] A. Giachetti. On-line analysis of echocardiographic image sequences. *Medical Image Analysis*, 2(3):261–284, 1998.

[5] G. Jacob, A. Noble, M. Mulet-Parada, and A. Blake. Evaluating a robust contour tracker on echocardiographic sequences. *Medical Image Analysis*, 3(1):63–76, 1999.

[6] M. Mulet-Parada and J.A. Noble. 2D+T Boundary Detection in Echocardiography. In *MICCAI*, pages 186–196, Cambridge, Mass. U.S.A, Oct. 1998.

[7] J. Declerck, J. Feldmar, M.L. Goris, and F. Betting. Automatic registration and alignment on a template of cardiac stress and rest reoriented SPECT images. *IEEE Transactions on Medical Imaging*, 16(1):1–11, 1997.

[8] G.I. Sanchez-Ortiz. Fuzzy Clustering Driven Anisotropic Diffusion: Enhancement and Segmentation of Cardiac MR Images. In *IEEE Nuclear Science Symposium and Medical Imaging Conference*, volume 3, pages 1873–1875, Toronto, Canada, Nov. 1998.

[9] G.I. Sanchez-Ortiz, D. Rueckert, and P. Burger. Knowledge-based tensor anisotropic diffusion of cardiac MR images. *Medical Image Analysis*, 3(1):77–101, 1999.

[10] M. Mulet-Parada. *Intensity independent feature extraction and tracking in echocardiographic sequences*. PhD thesis, University of Oxford, 2000.

[11] J.C. Bezdek and P.F.Castelaz. "prototype classification and feature selection with fuzzy sets". *IEEE Trans. on systems, man and cybernetics*, SMC-7:87–92, 1977.

[12] B.M. ter Haar Romeny (Ed.). *Geometry-Driven Diffusion in Computer Vision*. Computational Imaging and Vision. Kluwer Academic Publishers, 1994.

[13] S.R. Arridge and A. Simmons. Application of multi-spectral probabilistic diffusion to dual echo mri. In *MIUA*, Leeds, UK, 1998.