

# DEFORMABLE STRUCTURES LOCALIZATION AND RECONSTRUCTION IN 3D IMAGES

Davide Moroni, Sara Colantonio, Ovidio Salvetti

*Institute of Information Science and Technologies (ISTI), Italian National Research Council (CNR), Pisa, Italy*

Mario Salvetti

*Department of Mathematics, University of Pisa, Pisa, Italy*

**Keywords:** Deformable structures, Image segmentation, Artificial neural networks, Cardiac MRI.

**Abstract:** Accurate reconstruction of deformable structures in image sequences is a fundamental task in many applications ranging from forecasting by remote sensing to sophisticated medical imaging applications. In this paper we report a novel automatic two-stage method for deformable structure reconstruction from 3D image sequences. The first stage of the proposed method is focused on the automatic identification and localization of the deformable structures of interest, by means of fuzzy clustering and temporal regions tracking. The final segmentation is accomplished by a second processing stage, devoted to identify finer details using a Multi-level Artificial Neural Network. Application to the segmentation of heart left ventricle from MRI sequences are discussed.

## 1 INTRODUCTION

In the field of computer vision, analyzing the deformation pattern of non-rigid structures may convey useful information in a variety of settings. For example satellite image sequences display temporal evolution of complex structures like clouds and vortices, whose analysis is essential for meteorological forecast (Papin et al., 2000); reinforcement of speech recognition by visual data may also be based on the analysis of lips deformation (Bregler and Konig, 1994). More interestingly for our purposes, deformable structures show up in human anatomy (lungs and heart being key examples) and their deformation modes are of key importance in understanding the functional properties of the related organs and assessing their health-state (Moroni et al., 2006).

Imaging modalities provide an invaluable aid in analyzing such complex structures. However image sequences contain a huge amount of high dimensional data (2 or 3 spatial dimensions plus time) which cannot be fully exploited unless with the help of suitable tools for image processing and pattern recognition. The main goal is to compactly but faithfully describe deformable structure in such a way to allow for deformation pattern characterization and assessment. Such

an *encoding* would be useful to build up a reference database for similarity searches or data mining procedures.

Of course, an essential step in characterizing deformable structures is first of all their localization and reconstruction from an image sequence.

In this paper we address this preliminary problem, assuming that structures are periodically deformable 3D structures. The developed method consists in a two-stage procedure, based on fuzzy clustering and Artificial Neural Networks (ANN), for the identification and reconstruction of the deformable structures of interest in an image sequence. As an application of this procedure we describe the segmentation of heart Left Ventricle (LV) from Magnetic Resonance (MR) image sequences, extending and refining (Colantonio et al., 2005). Actually the LV, pumping oxygenated blood to the body, is of great importance in assessing the health or pathological states of the heart since it has been demonstrated that different pathologies of the heart are deeply correlated to its dynamics.

The paper is organized as follows. In Section 2 we define the class of structures we are interested in, making explicit the necessary assumptions. Then in Section 3 the proposed approach is outlined and its basic modules are described in detail. More pre-

cisely in Section 3.1 we address the problem of automatic localization of periodically deforming structures, while in Section 3.2 we propose a method for the actual reconstruction. In Section 4 we discuss LV segmentation and show some experimental results, whereas conclusions and directions for further work are briefly collected in Section 5.

## 2 PERIODICALLY DEFORMABLE STRUCTURES

A structure  $O$  embedded in the background space  $\Omega \subset \mathbb{R}^3$  is a collection

$$O = \{V^\alpha\}_{\alpha=1,2,\dots,k}$$

of smooth surfaces  $V^\alpha$  embedded in  $\Omega$ . The smoothness assumption is a quite common hypothesis in computational anatomy (see e.g. (Grenander and Miller, 1998)) and it is satisfied in practice to a large extent. We use, moreover, collection of surfaces - instead of a single one- to be able to describe and characterize also subparts of the structures.

A deformable structure  $C = (O_t)_{t=1,2,\dots}$  is a temporal sequence of structures satisfying some smoothness constraint. Each  $O_t = \{V_t^\alpha\}_{1 \leq \alpha \leq k}$  should be regarded as the *snapshot* of the deformable structure at time  $t$ .

We require that each surface  $V_t^\alpha$  appearing in the snapshot at time  $t$  can be smoothly deformed into  $V_{t+1}^\alpha$  in the subsequent snapshot. Tears or crack of any structure subpart are, therefore, ruled out; moreover, in such a way, we avoid dealing with changes in topology, that would require to model shape transitions. Such a task would be essential for example in meteorological applications, but is far beyond our present scopes in biomedical problems.

Finally, a periodically deformable structure is a deformable structure for which there exists an integer  $T$  such that  $\forall t : O_t = O_{t+T}$ . In other words, the deformable structure depicts a periodic motion; thus, a periodically deformable structure is characterized by a finite list of snapshots  $(O_0, O_2, \dots, O_{T-1})$ , which will be referred to as its deformation cycle.

It is assumed that a 3D image sequence  $(S_t)_{0 \leq t \leq T-1}$  has been acquired from which morphology of the structure can be inferred so as to represent faithfully a physical body or phenomenon of interest. Without loss of generality, considering biomedical applications, we assume that each scan  $S_t$  consists of a set of  $h$  parallel 2D slices. The pixels of each slice are identified by their position  $(x, y)$  in the slice plane and by a third coordinate,  $z$  ( $z = 1, \dots, h$ ), which refers

to the index of the slice itself in the stack. In the following, the three coordinates  $(x, y, z)$  will be referred to as a voxel and  $S_t(x, y, z)$  will denote its intensity value.

## 3 TWO-STAGE DEFORMABLE STRUCTURE SEGMENTATION

We address the problem of deformable structures segmentation with a two-stage method, which, firstly, automatically localize the deformable structure and then extracts its finer details, looking for precise contours of the whole structure and of its subparts.

To each scan  $S_t$ , the following two-stage procedure is applied:

1. *Structure automatic localization*: a cluster analysis, based on the fuzzy  $c$ -means algorithm, is applied to identify and label homogeneous regions in each scan. Through a region tracking procedure, the behavior of these regions is analyzed over an entire cycle, in order to extract a rough approximation  $C' = \{O'_t\}_{0 \leq t \leq T-1}$  of the deformable structure  $C$ .
2. *Segmentation refinement*:  $C'$  is used to compute the approximate orientation of the real structure  $C$ , which, in turn, is used to extract three-dimensional features processed by a dedicated ANN, in order to complete the segmentation, by identifying accurate contours of  $C$ .

### 3.1 Automatic Localization of Deformable Structures

We assume that shape descriptors of the deformable structure tracked on time exhibit a periodical behavior, with main frequency concentrated in the motion frequency. Further we assume that the subparts of the deformable structures appear as homogeneous regions at some scale. However the latter assumption is dictated by our implementations and can be substituted without altering the spirit of this contribution.

#### 3.1.1 Clustering

Homogeneous image regions are first labelled using an unsupervised clustering method, based on the fuzzy  $c$ -means algorithm (FCM) (Bezdek, 1981). This algorithm groups a set of data in a predefined number of regions so as to iteratively minimize a criterion function, namely the sum-of-squared-distance from region centroids, weighted by a cluster membership function. A membership grade  $p \in [0, 1]$  is asso-

ciated to each element of the data set, describing its probability to be in a particular cluster.

The FCM algorithm is applied to each 3D scan  $S_t$  to produce a number of clusters: for any voxel  $\mathbf{x}$ , a features vector  $(I_0(\mathbf{x}), I_1(\mathbf{x}), I_2(\mathbf{x}), \dots, I_r(\mathbf{x}))$  is computed so that  $I_0(\mathbf{x}) = S_t(\mathbf{x})$ , and for  $d = 1, \dots, r$ , we set  $I_d(\mathbf{x}) = \mathcal{G}_d * S_t(\mathbf{x})$ , where  $\mathcal{G}_d$  is a Gaussian kernel with standard deviation  $\sigma \propto d$ .

This, in turn, induces a partition of the image domain into a set  $P_t = \{R_t^1, R_t^2, \dots\}$  of disjoint connected regions, where the upper indices  $1, 2, \dots$  are region labels. In the following,  $\rho_t$  will denote the generic region in  $P_t$ .

### 3.1.2 Region Tracking

Once eliminated regions of negligible volume (island removal), an intra-cycle tracking procedure is performed. A simple centroid-based tracking algorithm associates, to any region  $\rho_t \in P_t$  in the phase  $t$ , its correspondent region  $T(\rho_t) \in P_{t+1}$  in the subsequent phase  $t + 1$ . The procedure can be iterated, thus producing a region sequence

$$\rho_t = T^0(\rho_t) \rightarrow T^1(\rho_t) \rightarrow T^2(\rho_t) \rightarrow \dots$$

which may be thought as the evolution of the starting region  $\rho_t$  in the different phases. Considering  $t = 0$  as reference phase, for each  $\rho_t \in P_t$  the regions appearing in its evolution are collected in a list  $\text{Ev}(\rho_0) = (T^t(\rho_0))_{0 \leq t \leq T-1}$ .

### 3.1.3 Features Extraction

For any region  $\rho_0 \in P_0$ , the behavior in time of a shape descriptor  $G$  (such as elementary geometric properties: volume, inertia moments etc. ) can be estimated by evaluating  $G$  for every element in the list  $\text{Ev}(\rho_0)$ , thus obtaining a vector  $f_G(\rho_0) = (G.T_t(\rho_0))_{0 \leq t \leq T-1}$ .

To detect the oscillatory behavior of  $f_G(\rho_0)$ , it is effective and convenient to switch to frequency domain and consider its power spectrum density (PSD) function.

In more detail, first the vector  $f_G(\rho_0)$  is normalized:

$$\mu_G(\rho_0) = \sum_{t=0}^{T-1} G.T_t(\rho_0)$$

$$f'_G(\rho_0) = \frac{1}{\mu_G(\rho_0)} f_G(\rho_0)$$

to obtain a scale invariant vector and, thus, getting rid of the relative size of  $G$ . Then the Fourier transform in non negative frequencies is computed:

$$F'_G(\rho_0) = \mathcal{F}(f'_G(\rho_0))$$

and the PSD is estimated using the periodogram method:

$$\text{PSD}_G(\rho_0) = |F'_G(\rho_0)|^2$$

The first harmonic coefficient  $v_G(\rho_0)$  in  $\text{PSD}_G(\rho_0)$  is then selected as a salient feature. Indeed, for fixed regions the variations in  $G$  during time are essentially due to noise; instead for regions in periodic motion the spectrum power is concentrated in the motion frequency.

Finally, for a predetermined list  $\{G, H, \dots\}$  of shape descriptors, a features vector  $I(\rho_0) = (\mu_G(\rho_0), v_G(\rho_0), \mu_H(\rho_0), v_H(\rho_0), \dots)$  is associated to each  $\rho_0 \in P_0$ .

### 3.1.4 Region Classification

Let  $O'_t$  denote the region corresponding to the deformable structure  $C$  at the phase  $t$ . At first, the reference phase is considered and  $O'_0$  is searched among regions  $\rho_0 \in P_0$ , taking into account their features vectors  $I(\rho_0)$ . More precisely, a set of learning examples is used to introduce a Mahalanobis distance in the feature space. Let  $I_1, I_2, \dots, I_s$  be a set of observed feature vectors relative to a training set of regions  $C'$  with mean  $m$  and covariance matrix  $\Sigma$ . The associated Mahalanobis distance, defined by

$$D(I) = \left( (I - m)^t \Sigma^{-1} (I - m) \right)^{1/2},$$

measures the dissimilarity of a feature vector w.r.t. to the expected region feature vector. Thus, for any new case,  $O'_0$  is selected among candidate regions  $\rho_0 \in P_0$  according to the criterion:

$$O'_0 = \arg \min_{\rho_0 \in P_0} D(I(\rho_0))$$

In subsequent phases, the region  $O'_t$  is singled out by means of the tracking algorithm, namely  $O'_t$  is defined as  $T^t(O'_0)$ .

## 3.2 Segmentation Refinement

The localization of the deformable structure in the previous section supplies as a byproduct a rough approximation of its boundary surface, which may suffer from poor intensity contrast or the presence of spurious structures.

The aim of this section is to refine the segmentation found in the previous section and to identify as well the contours of the structure subparts.

The set up is as follows. Let  $\Omega \subset \mathbb{R}^3$  be the image domain of the scan  $S_t$ . First we define in 3.2.1 a features function  $\mathcal{F}_t : \Omega \rightarrow \mathbb{R}^s$ , that assigns to each point  $\mathbf{x} \in \Omega$  a vector  $\mathcal{F}(\mathbf{x})$  of local features extracted from the image data  $S_t$ . Then we use an approach based

on Multi-Level Artificial neural networks (MANN) to find functions  $\Phi_\alpha : \Omega \times \mathbb{R}^s \rightarrow \mathbb{R}$  ( $\alpha = 1, 2, \dots, k$ ) s.t. the level sets:

$$V_\alpha = \{\mathbf{x} \in \Omega \mid \Phi_\alpha(\mathbf{x}, \mathcal{F}(\mathbf{x})) = 0\} \quad \alpha = 1, 2, \dots, k$$

correspond to the surface  $V^\alpha$  respectively.

The functions  $\Phi_\alpha$  are learned using a training set of segmented images and they can be used subsequently to segment new instances.

### 3.2.1 Features Extraction

Given a scan  $S_t : \Omega \rightarrow \mathbb{R}$  we briefly describe how a features function  $\mathcal{F}_t : \Omega \rightarrow \mathbb{R}^s$  may be constructed. Since the neural network will eventually use this function for the identification of image edges, it is clear that the function  $\mathcal{F}_t$  should include “edge detector”-like clues.

The involved features can be divided into two classes. First, low-level features are considered: they are context-independent and do not require any knowledge and/or pre-processing. Some examples are voxel position, gray level value, gradients and other differentials, texture, and so forth. Middle-level features are also selected, since voxel classification can benefit from more accurate clues, specific of the problem at hand. In particular, the knowledge of the deformable structure orientation, obtained as a byproduct in Section 3.1, can be used to individuate an Intrinsic Reference System (IRS) suitable to describe the structure shape. If, in addition, a priori information about the structure shape is available, a reliable clue for detecting edges in the images is given by the gradient along the normal direction to the expected edge orientation. Moreover, a multiscale approach is adopted: the features are computed on blurred images, supplying information about the behavior of the voxel neighborhood, which results in a more robust classification.

### 3.2.2 MANN-based Voxel Classification

The set of selected features are processed to accomplish the voxel classification by means of a Multilevel Artificial Neural Network (MANN), which assures several computational advantages (Di Bona et al., 2003).

For each voxel  $\mathbf{x}$ , its computed features vector  $\mathcal{F}_t(\mathbf{x})$  is divided into vectors  $\mathcal{F}_t^i(\mathbf{x})$ , each one containing features of the same typology and/or correlated. Then each  $\mathcal{F}_t^i(\mathbf{x})$  is processed by a dedicated classifier based on an unsupervised Self Organizing Maps (SOM) architecture. The set of parallel SOM modules constitutes the first level of the MANN which aims at clustering each portion of the feature vector into crisp classes, thus reducing the computational complexity.

Cluster indexes, in turn, are the input of the final decisional level, operated by a single EBP network. The output of this last module consists in a vector of membership grade of the voxel  $\mathbf{x}$  to the various surfaces  $V^\alpha$  ( $1 \leq \alpha \leq k$ ). The SOM modules are trained according to Kohonen algorithm (Kohonen, 1997). For the EBP module, a set of 3D scans should be pre-classified by an expert observer and used for supervised training, performed according to the Resilient Back-Propagation algorithm (Riedmiller and Braun, 1993).

## 4 STUDY CASE: LEFT VENTRICLE SEGMENTATION

Accurate segmentation of cardiac cavities is fundamental in assessing cardiac function and determining quantitative parameters. Magnetic Resonance Imaging (MRI) is a high quality and well-established imaging modality in analyzing heart diseases and has proved to be more reliable than other techniques, both in supplying accurate and reproducible morphological information and in assessing heart functions (ACC\AHA Task Force, 2005). However, due to noise or acquisition artifacts, visual information can be corrupted or ill defined: in a usual edge map of a cardiac MR slice, boundaries belonging to the LV appear broken or, even worse, close to stronger edges of other structures. In such cases, only expert knowledge may help: the exact location of the contours cannot be based only on image evidence, but should be learned from examples provided by expert observers. Usually, researchers have tried to design *ad hoc* algorithms able to incorporate a priori information about the LV shape. Model based surface detector have been widely used: for example, (Declercq et al., 1997) employed a Canny-Deriche edge detector in a 3D polar map to segment endocardial and epicardial surfaces, while (Faber et al., 1991) defined a hybrid spherical-cylindrical coordinate system. Snakes, since their introduction in the seminal paper by (Kass et al., 1988), have been a powerful tool in cardiac images analysis for segmentation and motion tracking. Recent improvements in this field include works by (Jolly et al., 2001), who reduced sensitivity to initial contour through Dijkstra algorithm, and by (Paragios, 2002) and (Huang et al., 2004) who introduced deformable models influenced by forces derived from image region information. (Mitchell et al., 2002; Üzümcü et al., 2003) used the concept of active appearance model (AAM). An AAM is a technique of analysis by synthesis, which, in principle, could describe any heart through a set

of learned 3D shapes and a set of allowed variations. Although AAMs mitigate the segmentation problem exploiting a shape prior, they are however very sensitive to initialization and intensity variations occurring at random locations. Aiming at solving this two issues, in (de Bruijne and Nielsen, 2004) a particle filtering scheme is introduced in the AAM framework to get relatively initialization-independent results, while a shape model inference on the basis of pixel classification is used to cope with local intensity variations. Finally, also neural networks approaches have been proposed. In (Stalidis et al., 1999) a Generating-Shrinking Neural Classifier is used to distinguish among lung, blood and myocardium points. This classification allows to extract a set of points on myocardial surfaces and, then, to assess parameters for a wavelets-based model. Two dedicated neural networks are presented instead in (Coppini et al., 1995). The first is used to select from an edge map boundaries belonging to the ventricle. To get a meaningful and unbroken surface, these edges are further processed: a thin plate model for the left ventricle is introduced and a stable configuration of minimum potential is found by means of an analog neural network implementation. The methods reported above depend tightly from the choice of model parameters and initial conditions. The last problem has traditionally been solved with a manual intervention by an expert observer, but this contrasts the need of a fully automatic segmentation.

#### 4.1 Data

The two-stage method described in the previous section has been applied to short axis gradient echo MR images, acquired with the FIESTA, GENESIS SIGNA MRI device (GE medical system), 1.5 Tesla, TR = 4.9 ms, TE = 2.1 ms, flip angle 45° and resolution  $r_x \times r_y \times r_z = (1.48 \times 1.48 \times 8)$  mm. Sets of  $T = 30$  3D scans, consisting of  $h = 11$  2D slices, were acquired at the rate of 30 ms for cardiac cycles [diastole-systole-diastole]. Various clinical cases were considered, for a total of 360 scans, corresponding to 12 cardiac cycles.

#### 4.2 Experimental Results

The LV segmentation can be viewed as a bi-modal problem (Paragios et al., 2002): the structure of interest is the myocardium, which can be identified and extracted seeking the separation among the endocardium and the epicardium surfaces.

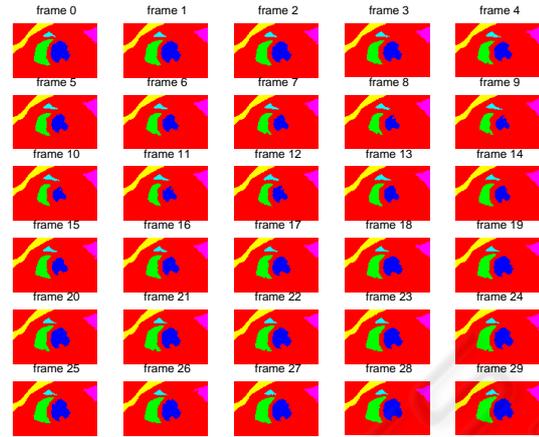


Figure 1: Visualization of the results of clusterization and tracking algorithm (slice number 6).

##### 4.2.1 LV Localization

FCM was applied separately to each scan to produce two cluster using 2 as fuzziness parameter; we considered as a feature vector  $(I_0, I_1, \dots, I_r)$  where  $I_d = \mathcal{G}_d * S_t$  and  $\mathcal{G}_d$  is a gaussian kernel of standard deviation  $d$  times the inslice resolution  $r_x$ :

$$\mathcal{G}_d(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}(d \cdot r_x)^{1/2}} \exp\left(\frac{-\|\mathbf{x}\|^2}{(d \cdot r_x)^2}\right)$$

Experimental testing showed that setting  $r = 2$  is sufficient to get a good partition of the image domain. The result of the tracking procedure on a middle slice is shown in Figure 1. The convex-hull volume and the inertia moments were considered as geometric properties. The use of convex volume (instead of the simpler volume) reduces the effect of papillary muscles that sometimes move towards the boundary of the region corresponding to the LV. Processing was performed only on middle slices, thus eliminating the apical cap and the basal segments of the LV. Analysis of various clinical cases has been used to introduce the Mahalanobis distance  $D$ ; for simplicity, the covariance matrix  $\Sigma$  has been assumed to be diagonal.

##### 4.2.2 LV Boundary Extraction

The previously found region corresponds roughly to the LV cavity (LVC) and may be used to introduce an IRS. Since LV is essentially bullet shaped, a hybrid spherical/cylindrical reference system is suitable to characterize its geometry and extract salient edge information. To describe more in detail the IRS, suppose, without loss of generality, that the  $z$ -axis of

$\Omega \subset \mathbb{R}^3$  coincides with the long axis of the LV computed in the previous section and that it is oriented from the apex to the base of the LV. A point  $\mathbf{O} = (0, 0, z_0)$  on the long axis is selected as the switching point between cylindrical and spherical coordinates. Cylindrical coordinates  $(r, \theta, h)$  are assigned to points  $\mathbf{x} = (x, y, z) \in \mathbb{R}^3$  satisfying  $z - z_0 \geq 0$ :

$$r(x, y, z) = \sqrt{x^2 + y^2} \quad (1)$$

$$\theta(x, y, z) = \tan^{-1}\left(\frac{y}{x}\right) \quad (2)$$

$$h(x, y, z) = z - z_0 \quad (3)$$

whereas spherical coordinates  $(r, \theta, \phi)$  are given to points satisfying  $z - z_0 \leq 0$ :

$$r(x, y, z) = \sqrt{x^2 + y^2 + (z - z_0)^2} \quad (4)$$

$$\theta(x, y, z) = \tan^{-1}\left(\frac{y}{x}\right) \quad (5)$$

$$\phi(x, y, z) = \cos^{-1}\left(\frac{z - z_0}{r(x, y, z)}\right) \quad (6)$$

(Inverse tangent must be suitably defined to take the correct quadrant into account)

Note that  $r$  and  $\theta$  are defined consistently everywhere. The unit vector field  $\hat{r}(\mathbf{x}) = \frac{\partial}{\partial r} \mathbf{x} / \|\frac{\partial}{\partial r} \mathbf{x}\|$  (pointing in direction of increasing radial coordinate  $r$ ) is then given in cartesian coordinates by:

$$\hat{r}(\mathbf{x}) = \begin{cases} (\cos \theta, \sin \theta, 0) & \text{if } z - z_0 \geq 0 \\ (\cos \theta \sin \phi, \sin \theta \sin \phi, \cos \phi) & \text{if } z - z_0 \leq 0 \end{cases} \quad (7)$$

Note that the field  $\hat{r}$  is almost orthogonal to cardiac surfaces and, therefore, the derivative  $\frac{\partial S_t}{\partial r}$  along the radial direction may be used as a clue for edge detection. Indeed for a point on a cardiac surface, the modulus of radial derivative is likely to be a high fraction of total gradient magnitude (see Figure 2). Moreover the degree of freedom in the choice of the switching point  $\mathbf{O}$  may be used to tune the IRS to the peculiar cardiac geometry under exam. An automatic procedure for the selection of the switching point is described in the Appendix. The hybrid reference system is used to associate to each point  $\mathbf{x} = (x, y, z) \in \Omega$ , a vector consisting of the following features extracted from the data :

- **Position:** The position of a point  $\mathbf{x}$  w.r.t. the IRS is expressed as a quadruple  $(r, \theta, \phi, h)$ . If  $z - z_0 \leq 0$  the entries  $r, \theta, \phi$  represent its spherical coordinates, whereas  $h$  is set to 0. Similarly, for  $z - z_0 \geq 0$ , the entries  $r, \theta, h$  represent its cylindrical coordinates whereas  $\phi$  is set to  $\pi/2$ . Notice that with this choice both definitions agree for points in the plane  $z = z_0$ .
- **Intensity:** The intensity value  $S_t(\mathbf{x})$  as well as its smoothed analogues  $\mathcal{G}_d * S_t(\mathbf{x})$ .

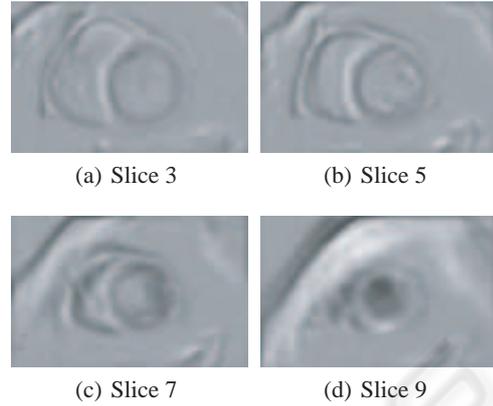


Figure 2: Example of computed features: radial derivative.

- **Gradient norm:** Gradient norm  $\|\nabla(\mathcal{G}_d * S_t(\mathbf{x}))(\mathbf{x})\|$  of the smoothed images  $\mathcal{G}_d * S_t$ .
- **Radial derivative:** The radial derivative of the smoothed images  $\frac{\partial \mathcal{G}_d * S_t(\mathbf{x})}{\partial r}(\mathbf{x}) = \nabla(\mathcal{G}_d * S_t(\mathbf{x})) \cdot \hat{r}$ .

Using the 2-level ANN, voxels are classified on the basis of their features vector as belonging or not to epi- and endocardial surfaces. More in detail, the set of extracted features is divided into two vectors  $\mathcal{F}_1, \mathcal{F}_2$  containing respectively 1) position and intensity and 2) position, gradient norm and radial derivative. The position w.r.t. IRS is replicated in both vectors because it reveals salient for clustering both features subsets. Then, the first level of the MANN consists of two SOM modules, which have been defined as 2D lattice of neurons and dimensioned experimentally, controlling the asymptotic behavior of the number of excited neurons versus the non-excited ones, when increasing the number of total neurons (Di Bono et al., 2004).

A  $8 \times 8$  lattice SOM was then trained for clustering the features vector  $\mathcal{F}_1$ , while  $\mathcal{F}_2$  was processed by a  $10 \times 10$  lattice SOM.

A single EBP module has been trained to combine the results of the first level and supply the final response of the MANN. The output layer of this final module consists in two nodes, which are used separately for reconstructing the epicardium and the endocardium. Since each cardiac surface divides the space into two connected regions (one of which is bounded), each output node can be trained using the signed distance function with respect to the relative cardiac surface. In this way, points inside the surface are given negative values, whereas positive values are given to points in the outside. Henceforth the surface of interest correspond to the zero-level set of the output

function.

Different architectures have been tested, finding the best performance for a network with only one hidden layer of 15 units. A manual segmentation was performed with expert assistance on the available data. A set of 240 scans was used for network training, while the remaining ones were used for network performance test.

The voxel classification, supplied by the MANN, may be directly used for visualization purposes by using an isosurface extraction method, as shown in Figure 3. Figure 4 shows the intersection of the two cardiac surfaces with a slice plane.



Figure 3: Different views of the rendered left ventricle at end diastole. The surfaces are obtained applying marching cubes on the two output functions of the network. To eliminate satellites, a standard island removing procedure is applied.

## 5 CONCLUSIONS

In this paper we presented a general approach to the localization and reconstruction of periodically deformable structures, based on fuzzy-clustering and multilevel artificial neural networks.

The elective case studies are represented by the analysis of heart deformable anatomical structures. Actually, for demonstrating the effectiveness of the proposed framework, we have shown the preliminary results in the reconstruction of the heart left ventricle.

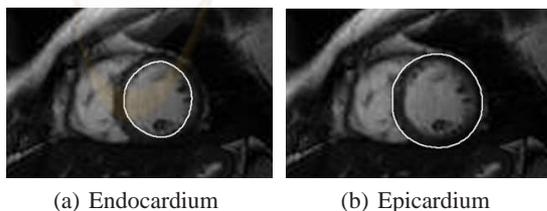


Figure 4: Intersection of cardiac surfaces with a slice plane.

The next step will be to employ the obtained results for defining a method to characterize the state of the deformable structure (a task already started in (Colantonio et al., 2006; Moroni et al., 2006)) with the goal of building up a reference database for similarity searches or data mining procedures.

## ACKNOWLEDGMENTS

This work was partially supported by European Project Network of Excellence MUSCLE - FP6-507752 (Multimedia Understanding through Semantics, Computation and Learning) and by European Project HEARTFAID “A knowledge based platform of services for supporting medical-clinical management of the heart failure within the elderly population”(IST-2005-027107).

## REFERENCES

- ACC\AHA Task Force (2005). ACC\AHA 2005 guideline update for the diagnosis and management of chronic heart failure in the adult. Technical report, American College of Cardiology Foundation and the American Heart Association.
- Bezdek, L. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithm*. Plenum Press, New York.
- Bregler, C. and Konig, Y. (1994). “Eigenlip” for robust speech recognition. In *Acoustics, Speech, and Signal Processing ICASSP-94.*, volume II, pages 669–672.
- Colantonio, S., Moroni, D., and Salvetti, O. (2005). MRI left ventricle segmentation and reconstruction for the study of the heart dynamics. In *IEEE ISSPIT*, pages 213–218, Athens, Greece.
- Colantonio, S., Moroni, D., and Salvetti, O. (2006). Shape comparison and deformation analysis in biomedical applications. In *Eurographics Italian Chapter Conference*, pages 37–43.
- Coppini, G., Poli, R., and Valli, G. (1995). Recovery of the 3-D shape of the left ventricle from echocardiographic images. *IEEE TMI*, 14(2):301–317.
- de Bruijne, M. and Nielsen, M. (2004). Image segmentation by shape particle filtering. *ICPR*, 03:722–725.
- Declerck, J., Feldmar, J., Goris, M., and Betting, F. (1997). Automatic registration and alignment on a template of cardiac stress and rest reoriented SPECT images. *IEEE Trans. Med. Imag.*, 16:727–737.
- Di Bona, S., Niemann, H., Pieri, G., and Salvetti, O. (2003). Brain volumes characterisation using hierarchical neural networks. *Artificial Intelligence in Medicine*, 28(3):307–322.
- Di Bono, M., Pieri, G., and Salvetti, O. (2004). A tool for system monitoring based on artificial neural networks. *WSEAS Transactions on Systems*, 3(2):746–751.

- Faber, T., Stokely, E., Peshock, R., and Corbett, J. (1991). A model-based four-dimensional left ventricular surface detector. *IEEE Tran. Med. Imag.*, 10:321–329.
- Grenander, U. and Miller, M. I. (1998). Computational anatomy: an emerging discipline. *Q. Appl. Math.*, LVI(4):617–694.
- Huang, X., Metaxas, D., and Chen, T. (2004). Meta-Morphs: Deformable shape and texture models. *CVPR*, 01:496–503.
- Jolly, M., Duta, N., and Funka-Lea, G. (2001). Segmentation of the left ventricle in cardiac MR images. In *ICCV*, pages 501–508.
- Kass, M., Witkin, A., and Terzopoulos, D. (1988). Snakes: Active contour models. *IJCV*, 1:321–331.
- Kohonen, T. (1997). *Self-Organizing Maps*, volume 30 of *Springer Series in Information Sciences*. 2nd edition.
- Mitchell, S., Bosch, J., Lelieveldt, B., van der Geest, R., Reiber, J., and Sonka, M. (2002). 3D active appearance models: Segmentation of cardiac MR and ultrasound images. *IEEE TMI*, 21(9):1167–1178.
- Moroni, D., Perner, P., and Salvetti, O. (2006). A general approach to shape characterization for biomedical problems. In Perner, P., editor, *Industrial Conference on Data Mining ICDM - Workshop on Mass-Data Analysis of Images and Signals*, IBAI CD-Report, pages 56–65, Leipzig.
- Papin, C., Boutheymy, P., Mèmin, E., and Rochard, G. (2000). Tracking and characterization of highly deformable cloud structures. In *Computer Vision - ECCV 2000*, volume 1843 of *LNCS*, pages 428–442. Springer Verlag.
- Paragios, N. (2002). A variational approach for the segmentation of left ventricle in cardiac image analysis. *IJCV*, 50(3):345–362.
- Paragios, N., Rousson, M., and Ramesh, V. (2002). Knowledge-based registration segmentation of the left ventricle: A level set approach. In *Proc. Sixth IEEE Workshop on Application of Computer Vision*, pages 37–42.
- Riedmiller, M. and Braun, H. (1993). A direct adaptive method for faster backpropagation learning: The RPROP algorithm. In *Proc. of the IEEE Intl. Conf. on Neural Networks*, pages 586–591, San Francisco, CA.
- Stalidis, G., Maglaveras, N., Dimitriadis, A., and Pappas, C. (1999). Using learning classification and four-dimensional parametric modeling for the analysis of myocardial thickening. In *Computers in Cardiology*.
- Üzümcü, M., Frangi, A., Sonka, M., Reiber, J., and Lelieveldt, B. (2003). ICA vs. PCA active appearance models: Application to cardiac MR segmentation. In *LNCS*, volume 2878, pages 451–458.

## APPENDIX: IRS SWITCHING POINT SELECTION

An automatic procedure that tune the IRS to the peculiar cardiac geometry under exams may be devised

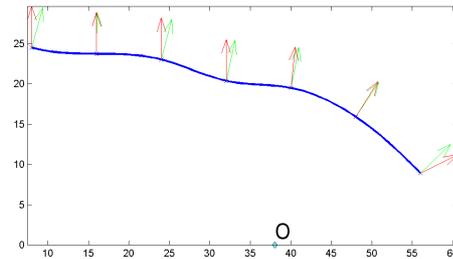


Figure 5: The mean  $z$ -profile of the LVC is plotted, together with outer normals to the curve (green) and radial vector  $\hat{r}_O$  (red) in selected points.

exploiting our knowledge of LVC.

Let  $O = (0, 0, z_0)$  be a point on the LVC long axis and let  $\hat{r}_O$  be the unit vector field given by eq. 7. A natural objective function that estimates orthogonality of the field  $\hat{r}_O$  w.r.t. endocardial surface is given by:

$$J(O) = \int_{\text{Endocardium}} (1 - \hat{r}_O \cdot \hat{n})^2 \quad (8)$$

where  $\hat{n}$  is the outer normal to the endocardial surface. In the previous equation, it is possible to approximate the (unknown) endocardial surface with the boundary of the LVC:

$$J(O) = \int_{\partial(\text{LVC})} (1 - \hat{r}_O \cdot \hat{n})^2 \quad (9)$$

Note that, in the spirit of Section 3.2.1, suitability of IRS to the cardiac geometry means in particular that  $\hat{r}_O$  should be orthogonal to cardiac surfaces. Therefore selection of the switching point may be translated into the optimization problem:

$$\hat{O} = \arg \min_O J(O) \quad (10)$$

However, for our purposes, it is not necessary to solve the optimization problem in eq. 10; indeed it is not meaningful to compute a so fine estimation of the switching point on the base of the rough data provided by LVC boundary. Instead, we prefer to convert this 2D fitting problem (integration on a surface) to a 1D problem (integration on a curve), by taking into account the  $z$ -axis mean profile of the LVC. This consists in the curve obtained considering the mean radius of the sections of LVC with a pencil of parallel planes  $\{z = \text{const}\}$  (see figure 5).

In this new setting the orthogonality criterion 10 may be restated with obvious modification.