Active Shape Models and Segmentation of the Left Ventricle in Echocardiography

Nikos Paragios[1,⋆], Marie-Pierre Jolly[2], Maxime Taron[1], and Rama Ramaraj[2]

[1] Atlantis Research Group, Certis,
Ecole Nationale des Ponts et Chaussees, Champs sur Marne, France
[2] Imaging and Visualization Department,
Siemens Corporate Research, Princeton, NJ, USA

Abstract. Segmentation of the left ventricle in echocardiographic images is a task with important diagnostic power. We propose a model-based approach that aims at extracting the left ventricle for each frame of the cardiac cycle. Our approach exhibits several novel elements. Modelling consists of two separate components, one for the systolic and one for the diastolic moment. Segmentation is considered in two steps. During the first step a linear combination of the systolic and the diastolic model is to be recovered - that dictates the new model - along with a similarity transformation that projects this model to the desired image features. During the second step, a linear combination of the modes of variation for the systolic and diastolic models is recovered for precise extraction of the endocardium boundaries. The process is considered in the temporal domain where constraints are introduced to couple information across frames and to lead to a smooth solution. Promising results demonstrate the potentials of the presented framework.

1 Introduction

Cardiovascular diseases are a major health concern world-wide. The left ventricle and in particular the endocardium is a structure of particular interest since it performs the task of pumping oxygenated blood to the entire body. Echocardiographic apical views when processed can determine the ejection fraction, a critical measure of the heart cycle. While segmenting the ventricle in the systolic and diastolic frame could be sufficient to provide such a measure, continuous tracking of the endocardium could further improve diagnosis.

Portability and low acquisition cost are the most attractive elements of echocardiographic imaging [14] while the presence of low signal-to-noise (SNR) ratio is an important limitation. Model-free segmentation techniques aim at separating the intensity properties of the image entities and fail to cope with noise and speckle in echocardiography. The use of prior knowledge that encodes the geometric form of the structure of interest is a reasonable way to deal with corrupted data.

Prior art in echocardiography consists of data-driven [6] and model-based segmentation approaches [9, 1]. One can also separate the techniques that perform segmentation

⋆ Research was carried out during the affiliation of the author with Siemens Corporate Research from November 1999 to March 2003.

R. Kimmel, N. Sochen, J. Weickert (Eds.): Scale-Space 2005, LNCS 3459, pp. 131–142, 2005.

in the polar [6] or in the raw space. Statistical/Bayesian formulations [6], active shape and appearance models [4, 1, 17], snakes and active contours [9], deformable models and templates [8] and level set techniques [2, 11] are well established techniques considered to address the segmentation of the left ventricle in echocardiographic images.

In this paper, we propose an active shape-inspired variational framework for fast, reliable and automatic segmentation of the endocardium for ultrasonic images. Our approach involves modelling, extraction of primitives, rough segmentation and border detection. We consider two separate model spaces, one for the diastolic and one for the systolic case. We recover the average shape and the modes of variations for each model through a Principal Component Analysis using a set of registered training examples.

Extraction of important primitives (ventricular walls, valve plane) that are used to initiate the segmentation process is the first step towards automatic 2D+time segmentation. Then, a linear combination of the two average models (systolic & diastolic) and the parameters of a similarity transformation between this new model and the image are incrementally recovered through a robust minimization. One should note that such a model space is dynamic. The parameters of this transformation are constrained to be smooth in the temporal domain. Precise endocardium segmentation is determined through a linear combination of the moments of variation that describe training sets, the systolic and the diastolic one. Such combination is constrained over time.

The remainder of this paper is organized as follows. In the next section, we address shape registration and modelling of the left ventricle. Global segmentation that involves a global transformation between the model-space and the image is presented in section 3 while local refinements are considered in section 4. Discussion and perspectives are addressed in section 5.

2 Modelling the Geometric Structure of the Endocardium

Building compact representations from a set of examples is a well studied problem in imaging and vision. The selection of appropriate models to represent all examples of the training set within a common pose is a critical component of such a process. Once such selection has been established, one would like to align all training examples to the same pose. Then modelling can then be performed using well known statistical techniques.

2.1 Global Registration, Mutual Information and Implicit Representations

Registration of shapes [15] is an open, interesting and challenging problem in imaging, vision and in particular in medical image analysis. Such application is not within the scope of the report, and therefore the prior art will be omitted and the adopted technique to address the problem will be briefly presented. Overviews of shape and image registration techniques are available at [10, 15]. Details on the considered approach to align the training examples can be found at [7]. Modelling requires global registration between the samples in the training set and establishment of local correspondences between them. Let us assume that a set of ground truth that consist of n components is available $[s_1, s_2, ..., s_n]$. Global alignment is equivalent with finding parametric transformations $\mathcal{A}_i$ between the training set examples and a target shape s such that

$$i \in [1, ...n] : \mathcal{A}_i(s) = s_i$$

where s is the common pose to be recovered. An emerging way to represent shapes is through the use of implicit representations. Such approaches are quite popular when the task involves tracking moving interfaces [11]. Inspired by the work proposed in [12] we represent shapes using distance transforms and implicit representations;

$$\phi_i(\omega) = \begin{cases} 0, & \omega \in s_i \\ d(\omega, s_i), & otherwise \end{cases}$$

where ω is the pixel location and $d(\omega, s_i)$ the minimum Euclidean distance between this pixel and the shape s_i.

The selected representation is translation/rotation invariant. Scale variations can be considered to be global illumination changes in the space of distance transforms. Therefore, registration under scale variations is equivalent with matching different modalities that refer to the same structure of interest. Mutual information [3, 16] is an invariant technique according to a monotonic transformation of the two input random variables. The use of such criterion to perform shape registration within the space of distance transforms was proposed in [7]. Such criterion is based on the global characteristics of the structures of interest. In order to facilitate the notation let us denote: (**i**) the source representation ϕ_i as f, and (**ii**) the target representation ϕ as g.

In the most general case, registration is equivalent with recovering the parameters $\mathbf{\Theta} = (\theta_1, \theta_2, ..., \theta_N)$ of a parametric transformation $\mathcal{A}$ such that the mutual information between $f_\Omega = f(\Omega)$ and $g^{\mathcal{A}}_\Omega = g\big(\mathcal{A}(\mathbf{\Theta}; \Omega)\big)$ is maximized for a given sample domain Ω;

$$MI(X^{f_\Omega}, X^{g^{\mathcal{A}}_\Omega}) = \mathcal{H}\left[X^{f_\Omega}\right] + \mathcal{H}\left[X^{g^{\mathcal{A}}_\Omega}\right] - \mathcal{H}\left[X^{f_\Omega, g^{A}_\Omega}\right]$$

where $\mathcal{H}$ represents the differential entropy. Such quantity represents a measure of uncertainty, variability or complexity and consists of three components: (i) the entropy of the model, (ii) the entropy of the projection of the model given the transformation, and (iii) the joint entropy between the model and the projection that encourages transformations where f explains g. One can use the above criterion and an arbitrary transformation (rigid, affine, homographic, quadratic) to perform global registration that is equivalent with minimizing:

$$E(A(\mathbf{\Theta})) = -MI(X^{f_\Omega}, X^{g^{\mathcal{A}}_\Omega}) = -\int\int_{\mathcal{R}^2} p^{f_\Omega, g^{\mathcal{A}}_\Omega}(l_1, l_2) log \frac{p^{f_\Omega, g^{\mathcal{A}}_\Omega}(l_1, l_2)}{p^{f_\Omega}(l_1) p^{g^{\mathcal{A}}_\Omega}(l_2)} dl_1 dl_2$$

where (i) p^{f_Ω} corresponds to the probability density in $f_\Omega \ \big(\ [\Phi_{\mathcal{D}}(\Omega)]\big)$, (ii) $p^{g^{\mathcal{A}}_\Omega}$ corresponds to density in $g^{\mathcal{A}}_\Omega \ \big(\ [\Phi_{\mathcal{S}}(\mathcal{A}(\mathbf{\Theta}; \Omega))] \ \big)$, and (iii) $p^{f_\Omega, g^{\mathcal{A}}_\Omega}$ is the joint density. Such framework can account for various global motion models. We consider similarity registration between the training examples for the endocardium shapes.

Registration examples for the particular class of endocardium shapes are shown in [FIG. (1)]. Once training examples have been aligned, one should address the problem of recovering point(element)-wise correspondences. Such a deformation field $L(\mathbf{\Theta}; \mathbf{x})$ can be recovered either using standard optical flow constraints or through the use of warping techniques like the free form deformations method [13], which is a popular approach in graphics, animation and rendering [5].

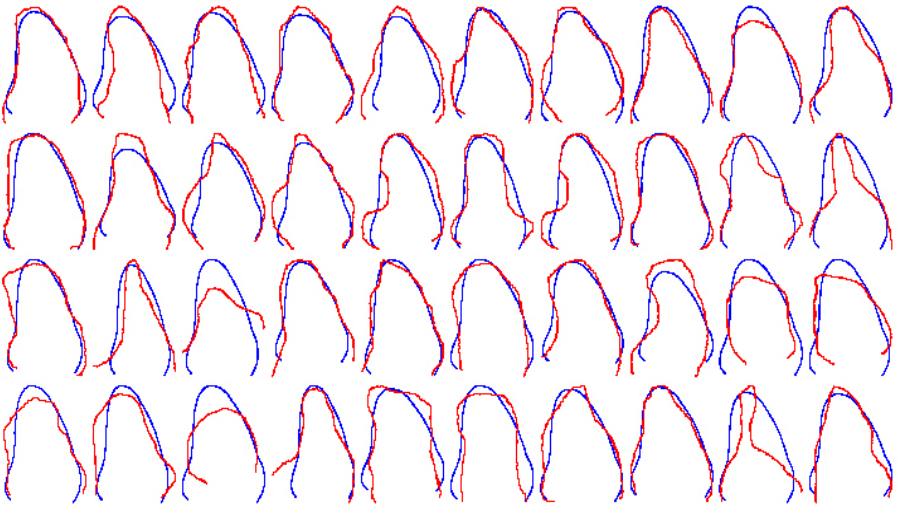

Fig. 1. Global Registration on the Space of Implicit Representations Using Mutual Information

2.2 Local Registration, Free Form Deformations and Implicit Representations

The essence of FFD is to deform an object by manipulating a regular control lattice P overlaid on its volumetric embedding space. Opposite to optical flow techniques, FFD techniques support smoothness constraints, exhibit robustness to noise and are suitable for modelling large and small non-rigid deformations. Furthermore, under certain conditions, it can support a dense registration paradigm that is continuous and guarantees a one-to-one mapping.

We consider an Incremental Cubic B-spline Free Form Deformation (FFD) to model the local transformation L. To this end, dense registration is achieved by evolving a control lattice P according to a deformation improvement $[\delta P]$. The inference problem is solved with respect to - the parameters of FFD - the control lattice coordinates.

Let us consider a regular lattice of control points

$$P_{m,n} = (P^x_{m,n}, P^y_{m,n});\ m = 1, ..., M, \quad n = 1, ..., N$$

overlaid to a structure

$$\Gamma_c = \{\mathbf{x}\} = \{(x, y) | 1 \leq x \leq X, 1 \leq y \leq Y\}$$

in the embedding space that encloses the source structure. Let us denote the initial configuration of the control lattice as P^0, and the deforming control lattice as $P = P^0 + \delta P$. Under these assumptions, the incremental FFD parameters are the deformations of the control points in both directions (x, y);

$$\mathbf{\Theta} = \{(\delta P^x_{m,n}, \delta P^y_{m,n})\};\ (m, n) \in [1, M] \times [1, N]$$

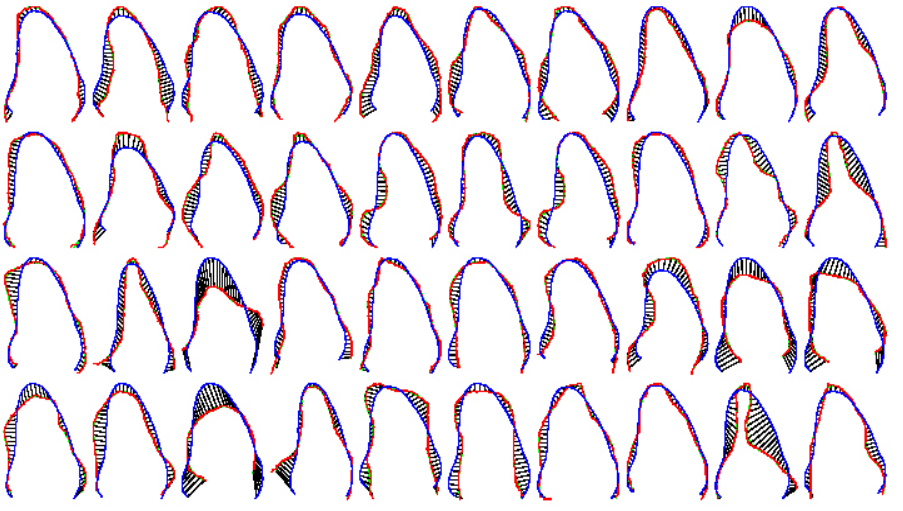

Fig. 2. Local Registration on the Space of Implicit Representations Using Free Form Deformations

The motion of a pixel $\mathbf{x} = (x, y)$ given the deformation of the control lattice from P^0 to P, is defined in terms of a tensor product of Cubic B-spline:

$$L(\mathbf{\Theta}; \mathbf{x}) = \mathbf{x} + \delta L(\mathbf{\Theta}; \mathbf{x}) = \sum_{k=0}^{3} \sum_{l=0}^{3} B_k(u) B_l(v) (P^0_{i+k,j+l} + \delta P_{i+k,j+l})$$

where $i = \lfloor \frac{x}{X} \cdot M \rfloor + 1$, $j = \lfloor \frac{y}{Y} \cdot (N \rfloor + 1$, $u = \frac{x}{X} M - \lfloor \frac{x}{X} \cdot M \rfloor$ and $v = \frac{y}{Y} N - \lfloor \frac{y}{Y} \cdot N \rfloor$.

The terms of the deformation component refer to (i) $\delta P_{i+l,j+l}$, $(k, l) \in [0, 3] \times [0, 3]$ consists of the deformations of pixel $\mathbf{x}$'s (sixteen) adjacent control points, (ii) $\delta L(\mathbf{x})$ is the incremental deformation at pixel $\mathbf{x}$, and (iii) $B_k(u)$ is the k^{th} basis function of a Cubic B-spline ($B_l(v)$ is similarly defined).

Local registration now is equivalent with finding the best lattice P configuration such that the overlaid structures coincide. Since structures correspond to distance transforms of globally aligned shapes, the Sum of Squared Differences (SSD) can be considered as the data-driven term to recover the deformation field $L(\mathbf{\Theta}; \mathbf{x}))$;

$$E_{data}(\mathbf{\Theta}) = \iint_{\Omega} \left(\Phi_{\hat{\mathcal{D}}}(\mathbf{x}) - \Phi_{\mathcal{S}}(L(\mathbf{\Theta}; \mathbf{x})) \right)^2 d\mathbf{x}$$

The use of such technique to model the local deformation registration component introduces in an implicit form some smoothness constraint that can deal with a limited level of deformation. In order to further preserve the regularity of the recovered registration flow, one can consider an additional smoothness term on the deformation field δL. We consider a computationally efficient smoothness term:

$$E_{smoothness}(\mathbf{\Theta}) = \iint_{\Omega} \left(\left|\left| \frac{\partial \delta L(\mathbf{\Theta}; \mathbf{x})}{\partial x} \right|\right|^2 + \left|\left| \frac{\partial \delta L(\mathbf{\Theta}; \mathbf{x})}{\partial y} \right|\right|^2 \right) d\mathbf{x}$$

Such smoothness term is based on a classic error norm that has certain known limitations. One can replace this smoothness component with more elaborated norms. Within the proposed framework, an implicit smoothness constraint is also imposed by the Spline FFD. Therefore there is not need for introducing complex and computationally expensive regularization components.

The Data-driven term and the smoothness constraints term can now be integrated to recover the local deformation component of the registration and solving the correspondence problem: $E(\mathbf{\Theta}) = E_{data}(\mathbf{\Theta}) + \alpha E_{smoothness}(\mathbf{\Theta})$, where α is the constant balancing the contribution of the two terms. The calculus of variations and a gradient descent method can be used to optimize such objective function [7]. The performance of the proposed framework on the Systolic Left Ventricle dataset is demonstrated in [Fig. (2)].

2.3 Composite Model Building

Let us assume that two sets of ground truth that consist of n components are available, one for the diastolic $[\mathbf{d}_1, \mathbf{d}_2, ..., \mathbf{d}_n]$ and one for the systolic case $[\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_n]$. Without loss of generality, one can assume that the elements of each set consists of m points defined on the Euclidean plane $\left(\mathbf{d}_i = (\mathbf{x}_1^i, \mathbf{x}_2^i, ..., \mathbf{x}_m^i)\right)$ and are registered to a common pose.

Principle Component Analysis (PCA) can be applied to capture the statistics of the corresponding elements across the training examples. PCA refers to a linear transformation of variables that retains - for a given number o_1, o_2 of operators - the largest amount of variation within the training data, according to:

$$\mathbf{d} = \overline{\mathbf{d}} + \sum_{k=1}^{o_1} \lambda_k^d \, (\mathbf{u}_k^d, \mathbf{v}_k^d), \quad \mathbf{s} = \overline{\mathbf{s}} + \sum_{k=1}^{o_2} \lambda_k^s \, (\mathbf{u}_k^s, \mathbf{v}_k^s)$$

where $\overline{\mathbf{d}}$ (resp. $\overline{\mathbf{s}}$) is the mean diastolic shape, o_1 (resp. o_2) is the number of retained modes of variation, $(\mathbf{u}_k^d, \mathbf{v}_k^d)$ (resp. $(\mathbf{u}_k^s, \mathbf{v}_k^s)$) are these modes (eigenvectors), and λ_j^d (resp. λ_j^s) are linear factors within the allowable range defined by the eigenvalues.

Once average models for the systolic and diastolic cases are considered, one can further assume that these models are registered, therefore there is an one-to-one correspondence between the points that define these shapes. Let $\left(\overline{\mathbf{d}} = (\mathbf{x}_1^d, \mathbf{x}_2^d, ..., \mathbf{x}_m^d)\right)$ be the diastolic average model and $(\overline{\mathbf{s}} = (\mathbf{x}_1^s, \mathbf{x}_2^s, ..., \mathbf{x}_m^s))$ the systolic one. Then one can define a linear space of shapes as follows:

$$\overline{\mathbf{c}}(\alpha) = \alpha \, \overline{\mathbf{s}} + (1 - \alpha) \, \overline{\mathbf{d}}, \quad 0 \leq \alpha \leq 1$$

One then can define a linear space of deformations that can account for the systolic, the diastolic frame as well as the frames in between:

$$\mathbf{c}(\alpha, \lambda_k^d, \lambda_s^d) = \overline{\mathbf{c}}(\alpha) + \sum_{k=1}^{o_1} \lambda_k^d \, (\mathbf{u}_k^d, \mathbf{v}_k^d) + \sum_{k=1}^{o_2} \lambda_k^s \, (\mathbf{u}_k^s, \mathbf{v}_k^s)$$

The most critical issue to be addressed within this process is the registration of the training examples as well as the registration of the systolic and diastolic average shapes. The approach proposed in [7] that performs registration in the implicit space of distance

functions using a combination between mutual information criterion and a free-form deformation principle is used. Such an approach can provide one-to-one correspondences between shapes for any given number of sampling elements. The resulting composite model is of limited complexity, can account for the systolic and the diastolic form of the endocardium as well as for the frames between the two extrema.

2.4 Composite Active Shape Models

Active shapes assume an average model, a certain number of modes of variation and the existence of corresponding image features. Without loss of generality one can assume that for each point j on the model space $\mathbf{c}(\alpha, \lambda_k^d, \lambda_s^d)$ the corresponding image point has been recovered $\mathbf{y}_j$. Then, the objective is to recover a set of parameters that will move each point in the model space $\mathbf{c}_j$ to the corresponding location in the image space $\mathbf{y}_j$. Such a task is performed in two stages where first a global transformation $\mathcal{T}$ between the model and the image is recovered that minimizes:

$$E_{data}(\alpha, \mathcal{T}) = \sum_{j=0}^{m} \rho \left(||\mathcal{T}(\overline{\mathbf{c}}_j(\alpha)) - \mathbf{y}_j|| \right)$$

according to some metric function ρ where $\mathcal{T}$ is a global transformation, similarity in our case

$$\mathcal{T}(x, y) = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} c \\ d \end{bmatrix}$$

that consists of a translation, a rotation and a scaling component and α defines the model space. The selection of the transformation should be consistent with the one adopted during the learning stage. It is important to point out that the model is not static since refers to a linear combination of the systolic and the diastolic model. Therefore, the process aims to recover simultaneously the combination of these two models that better accounts for the shape of the true data points and the optimal transformation between the model and the image space.

One can recover these parameters through an incremental update of the transformation. The corresponding location of the model points in the image plane could be used to improve the segmentation be seeking an incremental update on the transformation $\mathcal{T}(;)$ such that the projection of the $\overline{\mathbf{c}}_j$ moves closer to its true position $\mathbf{y}_j$ in the image.

3 Rough Segmentation of the Endocardium

The left ventricle is bounded on each side by the walls which tend to appear brighter in the ultrasound clip due to the various reflections from the tissue. In apical (both 2 chamber and 4 chamber views), the left ventricle is bounded on the bottom side by the mitral valve which connects it to the left atrium. The mitral valve is constantly moving (opening and closing) and its reflections are well recovered by the acquisition process.

We consider two parabolic equations to recover a rough approximation/detection of these walls which are the areas with the highest brightness. The parabola model the walls of the left ventricle but also outline the left atrium. The next step is to extract and track

the position of the mitral valve that separates the left ventricle and the left atrium. The approach relies on the observation that if the valve is closed, the two heart chambers are clearly separated while if the valve is open, the two chambers are connected. Two ellipses are used to model the ventricle and the atrium and the plane that best separates these ellipses and is consistent over time is considered to be the valve plane.

3.1 Recovering Correspondences

The most critical part within the presented framework is solving the correspondence problem, between the actual projection of the model and the optimal position. Such task within the active shape model is solved using a normalized intensity profile in the normal direction. We consider a probabilistic formulation of the problem. One would like to recover a density $p_{border}(;)$ that can provide the probability of a given pixel ω being at the boundaries of the endocardium. Within the considered framework, one can constrain the search in the direction normal to the model projection. The ventricular area consists of blood pool and heart walls. Endocardium border detection is equivalent with finding the boundaries between these two classes.

A description on the statistical properties of the blood pool as well as cardiac wall can be recovered. Let $p_{wall}(;)$ being the probability of a given intensity being part of the endocardium walls and $p_{blood}(;)$ the density the describes the visual properties of the blood pool. Then or correspondences between the model and the image are meaningful in places where there is a transition (wall to blood pool) between the two classes. Given a local partition one can define a transition probability between these two classes. Such partition consists of two line segments $[\mathcal{L}(\mathcal{T}(\mathbf{x}_j)), \mathcal{R}(\mathcal{T}(\mathbf{x}_j))]$ that live in the normal direction $[\mathcal{T}(\mathcal{N}_j)]$ of the model curve at element $\mathcal{T}(\mathbf{x}_j)$. The origins of these line segments is the point of interest $\mathcal{T}(\mathbf{x}_j)$, the have the same slope and opposite directions. One can assume that this point is a projection of the model point $\mathbf{x}_j$:

$$p_{border}(\mathcal{T}(\mathbf{x}_j)) = p\left([wall|\omega \in \mathcal{L}(T(\mathbf{x}_j))] \cap [\text{blood}|\omega \in \mathcal{R}(T(\mathbf{x}_j))]\right)$$

These conditions can be considered independent, leading to the following form for the border density:

$$\begin{aligned} p_{border}(\mathcal{T}(\mathbf{x}_j)) &= p\left(wall|\omega \in \mathcal{L}(T(\mathbf{x}_j))\right)\ p\left(\text{blood}|\omega \in \mathcal{R}(T(\mathbf{x}_j))\right) \\ &= \textstyle\prod_{\omega\in\mathcal{L}} p_{wall}(I(\omega))\ \prod_{\omega\in\mathcal{R}} p_{blood}(I(\omega)) \end{aligned}$$

One can evaluate this probability under the condition that the blood pool and wall density functions are known. The use of -log function can be considered to overcome numerical constraints, that is equivalent with finding the minimum of:

$$E(\phi) = \sum_{\omega\in\mathcal{L}(\phi)} \lambda|I(\omega)| + \sum_{\omega\in\mathcal{R}(\phi)} \frac{(I(\omega)-\mu)^2}{2\sigma^2}$$

after dropping out the constant terms where blood pool is modelled using an exponential distribution (λ) and tissue/walls using a Gaussian distribution (μ, σ). Thus, the most probable correspondence is recovered through the evaluation of $E(\phi)$ where ϕ is a point in the line defined by the projected normal. The search space for ϕ is considered to be all image locations respecting two conditions; (i) live in the normal $\mathcal{T}(\mathcal{N}_j)$, and (ii)

their distance from the current projection $\mathcal{T}(\overline{\mathbf{c}}_j(\alpha))$ is within a given search window. Once such correspondences were established the mechanism presented in [**Sec. 2.4**] is considered to determine the optimal solution through the estimation of the parameters of the transformation $(\alpha_t, \mathcal{T}_t)$.

3.2 Constraints on the Motion and the Position of the End-Valve Points

The motion of the valve plane is very critical to the operation of the endocardium. Such motion is consistent over time, and quite often exhibits a symmetric form. Without loss of generality, one can assume that the first $\overline{\mathbf{c}}_0(\alpha)$ and the last point $\overline{\mathbf{c}}_m(\alpha)$ of the model correspond to the valve end points. The displacement of these points from one frame to the next can be recovered in an implicit form.

Let $(\alpha_{t-1}, \mathcal{T}_{t-1})$ be the model and its transformation to the image plane towards the desired image features in the previous frame. Then, given some estimates on the current solution $(\alpha_t, \mathcal{T}_t)$ one can constrain the implicit motion of the valve points as follows:

$$E_{valve\ motion}(\alpha_t, \mathcal{T}_t) = \\ \psi\left(|\mathcal{T}_{t-1}(\overline{\mathbf{c}}_0(\alpha_{t-1})) - \mathcal{T}_t(\overline{\mathbf{c}}_0(\alpha_t))|\right) + \psi\left(|\mathcal{T}_{t-1}(\overline{\mathbf{c}}_m(\alpha_{t-1})) - \mathcal{T}_t(\overline{\mathbf{c}}_m(\alpha_t))|\right)$$

where ψ is an error metric - the Euclidean in our case - $\mathcal{T}_{t-1}(\overline{\mathbf{c}}_m(\alpha_{t-1}))$ is the position of the valve point at frame $t-1$, $\mathcal{T}_t(\overline{\mathbf{c}}_m(\alpha_t))$ the corresponding projection at frame t and $\mathcal{T}_{t-1}(\overline{\mathbf{c}}_m(\alpha_{t-1})) - \mathcal{T}_t(\overline{\mathbf{c}}_m(\alpha_t))$ the displacement of this point from one frame to the next. Such term will constrain the motion of the valve plane to be smooth over time.

Such a term accounts for the relative motion of the valve points but not for their actual position. To this end, one can introduce constraints forcing the model projections of the valve points to be close to the valve-plane earlier recovered ($\alpha_{valve}\ x + \beta_{valve}\ y + \gamma_{valve} = 0$). The distance between the current positions of the model valve points $(\overline{\mathbf{c}}_0(\alpha), \overline{\mathbf{c}}_m(\alpha))$ and their projections to the valve-plane $(\mathbf{p}_0(t), \mathbf{p}_m(t))$ is a term to be minimized;

$$E_{valve\ projection}(\alpha_t, \mathcal{T}_t) = \psi\left(|\mathbf{p}_0(t) - \mathcal{T}_t(\overline{\mathbf{c}}_0(\alpha_t))|\right) + \psi\left(|\mathbf{p}_m(t) - \mathcal{T}_t(\overline{\mathbf{c}}_m(\alpha_t))|\right)$$

One can consider a step further by recovering the exact position of the valve points in the image and then use these positions during the segmentation process. To this end, a model is built on the image profile for the left and the right end-valve points using an image patch cantered at the ground truth position of the valve. Then, these patches are normalized and an average model is recovered. Standard matching techniques are considered within a search area in the vicinity of the projected valve position to recover the most prominent valve points.

3.3 Smoothness Constraints on the Transformation Parameters

The motion of the ventricle also should fulfil certain constraints. It has to be periodic, exhibit a shrinking between the diastolic and the systolic frame and an expansion for the last part of the cardiac cycle. Such conditions can be imposed in various forms. Direct motion constraints (like the one earlier considered) focus on the distance of a model point in two consecutive frames. Such constraints though do not encode the continuity of the

model. We consider an implicit form, where continuity is imposed on the parameters of the model ($\alpha(t)$) and the transformation ($\mathcal{T}(t)$);

$$E_{smoothness}(\alpha_t, \mathcal{T}_t) = \sum_{k=-\tau}^{\tau} \left(\omega\left(|\alpha(t) - \alpha(t+k)|\right) + w \sum_{p \in \mathcal{T}} \omega\left(|p(t) - p(t+k)|\right) \right)$$

where $p \in \mathcal{T}$ is the set of the similarity transformation parameters (a, b, c, d), ω a monotonically decreasing function and $[-\tau, \tau]$ is the interval where continuity on the rough segmentation parameters is imposed. Such a term will keep distance small between the registration parameters from the model space to the image within a couple of frames that is equivalent with constraining the motion of the endocardium from one frame to the next.

The objective function is minimized using a two-stage robust incremental estimate technique. The calculus of Euler-Lagrange equations with respect to the transformation parameters leads to a 4×4 linear system that has a closed form solution. Once such an estimate is available, the optimal model space α is recovered through an exhaustive search within the $[0, 1]$ integral according to some quantization step.

4 Refine Segmentation

Once, appropriate models and similarity transformations were recovered for all frames of the cardiac clip, the next step is precise extraction of the endocardium walls. Such a task is equivalent with finding a linear combination of the modes of variation that deforms globally the model projection towards the desired image features. The space of variations consists of the diastolic and the systolic models. Opposite to the rough segmentation case where the scale of the model is fixed, the need of a blending parameter does not exist between systolic and diastolic models of variation is not present. Under the assumption of existing correspondences $\mathbf{y}_j$ and the global transformation $(\alpha, \mathcal{T})$ for a given frame t - that is omitted from the notation -, these linear coefficients are recovered through:

$$E_{data}(\lambda_0^d, ..., \lambda_0^s, ...) = \sum_{j=0}^{m} \rho\left(||\mathcal{T}\left(\overline{\mathbf{c}}_j(\alpha)\right) + \sum_{k=1}^{o_1} \lambda_k^d \left(\mathbf{u}_k^d, \mathbf{v}_k^d\right) + \sum_{k=1}^{o_2} \lambda_k^s \left(\mathbf{u}_k^s, \mathbf{v}_k^s\right) - \mathbf{y}_j||\right)$$

Similar to the case of global transformation, one can assume now that the form of the ventricle changes gradually during the cardiac cycle. The geometry of the recovered solution is determined according to the set of coefficients $(\lambda_0^d, ..., \lambda_0^s, ...)$. Therefore, imposing constraints of smoothing deformation from one frame-to-the next is equivalent with seeking the lowest potential of

$$E_{smoothness}(\lambda_0^d, ..., \lambda_0^s, ...) =$$
$$\sum_{k=-\tau}^{\tau} \left(\sum_{l=1}^{o_1} \omega(\lambda_l^d(t) - \lambda_l^d(t+k)) + \sum_{l=1}^{o_2} \omega(\lambda_l^s(t) - \lambda_l^s(t+k)) \right)$$

Last, but not least additional constraints using the position of the valve points could be considered, that aims at moving the projections of the model valve points to the their

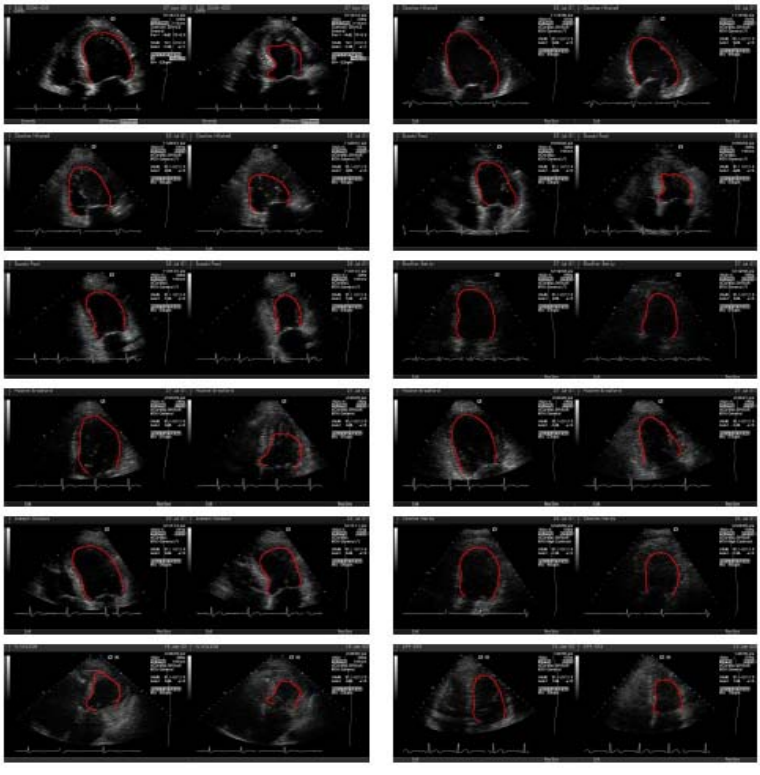

Fig. 3. Endocardium Segmentation for Apical Views for the diastolic frame and the systolic frame

true positions. The objective function is minimized using a robust incremental estimate technique. The calculus of Euler-Lagrange equations with respect to the unknown variables $(\lambda_0^d, ..., \lambda_0^s, ...)$ leads to a $[o_1 + o_2] \times [o_1 + o_2]$ linear system that has a closed form solution. Such step is repeated until convergence.

5 Conclusions

In this paper we have proposed a composite time-consistent 2D+time active shape model for the segmentation of the left ventricle in echocardiography. The approach exhibits certain novel elements, notably in the modelling and the segmentation phase.

Validation of the method was performed using a representative set of fifty patients for 2 and 4 champers views [**Fig. 3**] where the output of the proposed technique is superimposed to the ground truth. The objective was precise delineation of the ventricle, a much harder task than estimation of the ejection fraction. 50% of the time sonographers have accepted the result as it was while for the 25% of the remaining validation set, minor adjustments, notably in the valve position were sufficient to make the solution, same as the one pointed out from the clinical experts.

Future directions of our method involve epicardium segmentation and tracking. Such an objective is a natural extension that will improve results and the diagnostic power of the method since one could derive volume curves, EF radial strain, etc.

References

1. J. Bosch, S. Mitchell, B. Lelieveldt, F. Nijland, O. Kamp, M. Sonka, and J. Reiber. Automatocic Segmentation of Echocardiographic Sequences by Active Appearance Motion Models. *IEEE Transactions on Medical Imaging*, 21:1374–1383, 2002.
2. Y. Chen, H. Thiruvenkadam, H. Tagare, F. Huang, and D. Wilson. On the Incorporation of Shape Priors int Geometric Active Contours. In *IEEE Workshop in Variational and Level Set Methods*, pages 145–152, 2001.
3. A. Collignon, F. Maes, D. Vandermeulen, P. Suetens, and G. Marchal. Automated multi-modality image registration using information theory. In *Information Processing in Medical Imaging*, pages 263–274, 1995.
4. T. Cootes, A. Hill, C. Taylor, and J. Haslam. Use of Active Shape Models for Locating Structures in Medical Imaging. *Image Vision and Computing*, 12:355–366, 1994.
5. P. Faloutsos, M. van de Panne, and D. Terzopoulos. Dynamic Free-Form Deformations for Animation Synthesis. *IEEE Transactions on Visualization and Computer Graphics*, 3:201–214, 1997.
6. I. Herlin and N. Ayache. Feature extraction and analysis methods for sequences of ultrasound images. In *European Conference in Computer Vision*, pages 43–55, 1992.
7. X. Huang, N. Paragios, and D. Metaxas. Establishing Local Correspondences towards Compact Representations of Anatomical Structures. In *Medical Imaging Copmuting and Computer-Assisted Intervention*, 2003.
8. M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active Contour Models. *International Journal of Computer Vision*, 1:321–332, 1988.
9. D. Linker and V. Chalana. A Multiple Active Contour Model for Cardiac Boundary Detection on Echocardiographic Sequences. *IEEE Transactions on Medical Imaging*, 15:290–298, 1996.
10. J. Maintz and M. Viergever. A Survey for Medical Image Registration. *Medical Image Analysis*, 2:1–36, 1998.
11. S. Osher and N. Paragios. *Geometric Level Set Methods in Imaging, Vision and Graphics*. Springer Verlag, 2003.
12. N. Paragios, M. Rousson, and V. Ramesh. Non-Rigid Registration Using Distance Functions. *Computer Vision and Image Understanding*, 2003. to appear.
13. D. Rueckert, L.I. Sonda, C. Hayes, D. Hill, M. Leach, and D. Hawkes. Nonrigid registration using free-form deformations: Application to breast MR images. *IEEE Transactions on Medical Imaging*, 18:712–721, 1999.
14. C. Rumack, S. Wilson, and W. Charboneau. *Diagnostic Ultrasound*. Mosby, 1998.
15. R. Veltkamp and M. Hagedoorn. State-of-the-art in Shape Matching. Technical Report UU-CS-1999-27, Utrecht University, 1999. http://webdoc.gwdg.de/ebook/ah/2000/techrep/CS-1997/1997-27.pdf.
16. P. Viola and W. Wells. Aligment by Maximization of Mutual Information. In *IEEE International Conference in Computer Vision*, pages 16–23, 1995.
17. S. Zhou, D. Comaniciu, and A. Krishnan. Coupled-Contour Tracking through Non-orthogonal Projections and Fusion for Echocardiography. In *European Conference on Computer Vision*, volume I, pages 336–349, 2004.