

ESTIMATION OF CARDIAC PHASES IN ECHOGRAPHIC IMAGES USING MULTIPLE MODELS*

Jacinto Nascimento

Jorge S. Marques

João Sanches

Instituto Superior Técnico/Instituto de Sistemas e Robótica
Av. Rovisco Pais, Torre Norte, 1049-001, Lisboa, Portugal

ABSTRACT

This paper presents an algorithm for tracking the left ventricle in echocardiographic sequences, using multiple models. The use of multiple dynamic models is appropriate since the heart motion presents two phases (diastole and systole) with different dynamics. The main difficulty concerns the low contrast and speckle noise present in ultrasound images. To overcome this problem a robust multiple model tracker is used, based on a bank of nonlinear filters, organized in a tree structure. This algorithm determines which model is active at each instant of time and updates its state by propagating the probability distribution, using robust estimation techniques. It is shown in the paper that the proposed algorithm simultaneously copes with several dynamic models and with outliers. Furthermore the proposed algorithm provides high level information that is not available when a single model is used.

1. INTRODUCTION

The evolution of the left ventricle during the cardiac cycle plays an important role in medical diagnosis. This information can be retrieved using ultrasound imaging. The estimation of left ventricle by image processing techniques is an important issue, since it helps medical doctors to interpret the data and allows a quantitative evaluation of the heart dynamic properties, e.g., the ejection fraction.

Heart segmentation is a difficult problem since ultrasound images have a low signal to noise ratio, being corrupted by multiplicative noise (speckle) and by artifacts. Furthermore, the boundaries of the heart cavities are not always visible, since there is a low contrast between the ventricle and the myocardium (see Fig. 1). This makes the segmentation of the heart contours specially difficult.

Temporal cues can be used to improve the segmentation results, since they provide additional information about the heart cavities. However, the motion of the left ventricle is not well described by a linear dynamic model since it presents abrupt changes corresponding to two phases of the cardiac cycle (systole and diastole). Therefore, poor dynamic models are usually used in this problem.

Several works have been published addressing the estimation of the left ventricle. Some of them use statistical methods to describe the heart contour and the tissue properties inside and outside the endocardium [1, 2]. Markov random fields are used to describe the endocardium imposing smoothing restrictions in space and time. The image pixels are considered as independent random

*THIS WORK WAS PARTIALLY SUPPORTED BY FCT AND POSI UNDER PROJECT HEART 3D.

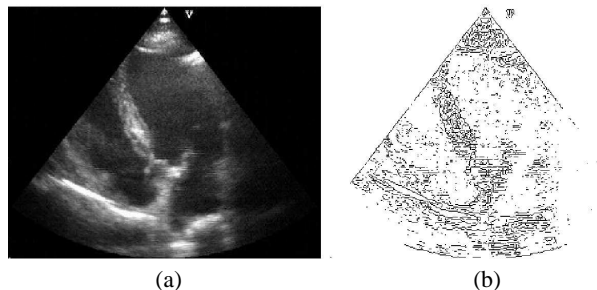


Fig. 1. Ultrasound image of the left ventricle (a), and the corresponding edge points (b).

variables with Rayleigh distribution. Good segmentation results have been recently achieved using a biomechanical model of the heart [3].

This paper proposes a different approach to overcome the difficulties associated with the segmentation of the left ventricle. Switched dynamic systems are used to represent the heart motion in both phases. This allows a more accurate representation of the heart motion, improving the tracking results. To deal with the speckle noise present in the ultrasound images, middle level features (strokes) are used instead of low level ones such as edge points. Robust filtering techniques are adopted to estimate the model parameters from noisy data with outliers [4]. The combination of these techniques (switched multiple models, middle level features and robust filtering) is explored in this paper.

The paper is organized as follows. Section 2 presents the problem formulation. Section 3 describes the robust tracker for switched multiple models, based on the propagation of Gaussian mixtures. Section 4 presents experiments results and section 5 concludes the paper.

2. PROBLEM FORMULATION

Let x_t be a vector describing the position and deformation of the left ventricle at time instant t . To account for changes in dynamic behaviour, it is assumed that x_t is described by two dynamic equations such that only one is active at each instant of time. Therefore

$$x_t = A_{k_{t-1}, k_t} x_{t-1} + w_t, \quad (1)$$

where $w_t \sim \mathcal{N}(0, Q_{k_{t-1}, k_t})$ is a white Gaussian noise, A_{k_{t-1}, k_t} is a matrix and $k_t \in \{1, \dots, m\}$ is the label of the active model at time instant t ; $m = 2$ is the number of the steady state models.

It is assumed that the label sequence k_1, \dots, k_t is a random sequence modeled by a first order Markov process with transition probability

$$T_{ij} = p(k_t = j \mid k_{t-1} = i), \quad (2)$$

where $i, j \in \{1, \dots, m\}$.

Many trackers are based on image features (edge points) detected on search lines, orthogonal to the predicted contour [5]. However, a large number of outliers is obtained when this approach is used in ultrasound images and the Kalman tracker fails to estimate the heart contour. To overcome this difficulty two techniques are used. First the edge points are organized in strokes using simple heuristic rules (see Fig. 2). Strokes are more reliable than edge points. Second robust filtering techniques are adopted to estimate the state of the switched dynamic model. Since we do not know which strokes are valid, a binary label is assigned to each one. Since the stroke labels are unknown, every possible combination of valid/invalid labels must be considered. Each sequence of stroke labels is called a *data interpretation*. Confidence degrees are assigned to each interpretation (association probabilities). Therefore, each interpretation has a different influence on the shape and motion estimates. The state estimate and the uncertainty measures are updated using all data interpretations.

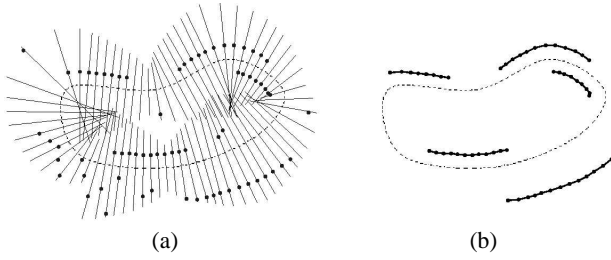


Fig. 2. Stroke detection: (a) edge detection and (b) linking (\bullet - denote the features obtained in the measurement lines orthogonal to the predicted contour).

Let $y(t)$ be a set of strokes detected in the image and let $y_{i,t}$ be a vector with the observations classified as valid according in the i -th interpretation. The sensor model associated with the i -th interpretation is given by

$$y_{i,t} = C_i x_t + \eta_{i,t} \quad (3)$$

where matrix C_i depends on the interpretation i and $\eta_{i,t} \sim \mathcal{N}(0, R_i)$ is a white Gaussian noise.

The switched dynamic system characterized by (1-3) has a hybrid state $z_t = (x_t, k_t)$ which includes the state vector x_t and the label of the active model k_t . The propagation of the hybrid state distribution is given by

$$p(z_t \mid z_{t-1}) = p(x_t \mid k_t, x_{t-1}, k_{t-1}) p(k_t \mid k_{t-1}, x_{t-1}). \quad (4)$$

The first term can be obtained from the dynamic equation (1) and the second term is T_{k_{t-1}, k_t} .

3. PROBABILITY DENSITY PROPAGATION

The inference problem can be stated as follows: *given a set observations $Y^t = \{y_1, y_2, \dots, y_t\}$ which contain outliers, what are the best estimates for the state vector, \hat{x}_t , and model label, \hat{k}_t .*

This is a nonlinear filtering problem. Assuming that the joint probability density function $p(x_t, k_t \mid Y^t)$ is known, the hybrid state can be estimated using the *maximum a posteriori* (MAP) method

$$(\hat{x}_t, \hat{k}_t) = \arg \max_{x_t, k_t} p(x_t, k_t \mid Y^t). \quad (5)$$

The *a posteriori* distribution $p(x_t, k_t \mid Y^t)$ can be obtained by a marginalization procedure

$$\begin{aligned} p(x_t, k_t \mid Y^t) &= \sum_{K^{t-1}} p(x_t, K^t \mid Y^t) \\ &= \sum_{K^{t-1}} c_{K^t} p(x_t \mid K^t, Y^t) \end{aligned} \quad (6)$$

where $c_{K^t} = p(K^t \mid Y^t)$ and $K^t = \{k_1, \dots, k_t\}$ is the sequence of model labels up to time instant t . If $p(x_t \mid K^t, Y^t)$ is a normal distribution $\mathcal{N}(\hat{x}_{K^t}, P_{K^t})$, the joint density $p(x_t, k_t \mid Y^t)$ defined in (5) is a mixture of Gaussians, each of them being associated to a different labeling sequence.

The propagation of the *a posteriori* density can be split into the following steps: multi model prediction and multi model filtering.

The *prediction step* aims to compute $p(x_t, k_t \mid Y^{t-1})$. This can be done as follows

$$p(x_t, k_t \mid Y^{t-1}) = \sum_{K^{t-1}} c_{K^t|t-1} p(x_t \mid K^t, Y^{t-1}) \quad (7)$$

where $c_{K^t|t-1} = P(K^t \mid Y^{t-1})$ is the predicted mixture coefficient computed by

$$c_{K^t|t-1} = T_{k_{t-1}, k_t} c_{K^{t-1}} \quad (8)$$

where $T_{k_{t-1}, k_t} = P(k_t \mid K^{t-1}, Y^{t-1})$. The computation of the mixture components for known K^t is straightforward: $p(x_t \mid K^t, Y^{t-1}) = \mathcal{N}(\hat{x}_{K^t|t-1}, P_{K^t|t-1})$ with mean and the covariance updated by

$$\hat{x}_{K^t|t-1} = A_{k_{t-1}, k_t} \hat{x}_{K^{t-1}} \quad (9)$$

$$P_{K^t|t-1} = A_{k_{t-1}, k_t} P_{K^{t-1}} A_{k_{t-1}, k_t}^T + Q_{k_{t-1}, k_t} \quad (10)$$

The *filtering step* aims to update the mixture coefficients as well as the state mean and covariance matrix. This can be accomplished in a robust way by using the S-PDAF method proposed in [6]

$$\hat{x}_{K^t} = \hat{x}_{K^t|t-1} + \sum_{i=1}^{m_i} \alpha_{i,t} K_{i,t} \nu_{i,t} \quad (11)$$

$$\begin{aligned} P_{K^t} &= \left[I - \sum_{i=1}^{m_i} \alpha_{i,t} K_{i,t} C_i \right] P_{K^t|t-1} \\ &+ \sum_{i=0}^{m_i} \alpha_{i,t} \hat{x}_{i,t} \hat{x}_{i,t}^T - \hat{x}_t \hat{x}_t^T \end{aligned} \quad (12)$$

where $\hat{x}_{i,t} = E\{x(t) \mid I_{i,t}, K^t, Y^t\}$, $\alpha_{i,t} \triangleq p(I_{i,t} \mid K^t, Y^t)$ is the association probability, $K_{i,t}$, $\nu_{i,t}$ are the Kalman gain and innovation, associated to the interpretation $I_{i,t}$.

After some manipulation the coefficients c_{K^t} can be obtained as follows [7]

$$c_{K^t} = \gamma c_{K^t|t-1} \sum_i^k \alpha_i(t) \prod_{j=1}^M \prod_{n=b^j}^{e^j} {}^k \mathcal{E}_i^j(s_n, t) \quad (13)$$

where γ is a normalization constant, $c_{K^t|t-1}$ is the predicted mixture coefficient, $\alpha_i(t)$ is the association probability assigned to the data interpretation $I_{i,t}$, M is the number of strokes, b^j, e^j are the indices of the j -th stroke, and \mathcal{E} is a normal or uniform distribution, depending on the stroke j being considered as valid or invalid in the interpretation $I_{i,t}$. The superscript k means that the quantities involved in (11-13) dependent on the label sequence.

The filter defined in (11-13) is denoted as *Robust Multi Model tracker*. The computation of (11), (12) and (13) is organized in a tree structure, each branch being characterized by c_{K^t} , x_{K^t} and P_{K^t} (K^t defines a tree path from the root to one of the leaves). The number of leaves (Gaussian modes) in this structure exponentially increases as time passes by. Assuming that we have m label values, the mixture will have m^t modes at time t . In practice, the number of modes must be limited. Several strategies can be used to achieve this goal, e.g., by using mode merging or component elimination [8]. The last method is used in this paper by discarding all mixture components with coefficients smaller than a threshold.

4. EXPERIMENTAL RESULTS

The proposed tracker was tested using sequences of echocardiographic images obtained with an ultrasound probe operating at 1.7 Mhz and sampled at 15 frames/sec. The goal is to track the walls of the left ventricle in ultrasound images and to estimate the corresponding phases of the cardiac cycle.

B-splines are used to represent the ventricle boundary. In this paper a B-spline with $N_c = 25$ control points, sampled at $N = 48$ samples is used. We have considered the group of Euclidean transformations [5] to model the heart motion assuming that we know the shape of the ventricle on the first frame. Four motion parameters are considered plus local deformations. The state is given by

$$x_t = [x_1, x_2, x_3, x_4, \theta_{x_1}, \dots, \theta_{x_{N_c}}, \theta_{y_1}, \dots, \theta_{y_{N_c}}]^T \quad (14)$$

with size $(4 + 2 \times N_c) \times 1$. The observation model is defined by

$$C = \begin{bmatrix} \mathbf{x} & \mathbf{1} & -\mathbf{y} & \mathbf{0} & \mathcal{B}_{N \times N_c} & \mathbf{O}_{N \times N_c} \\ \mathbf{y} & \mathbf{0} & \mathbf{x} & \mathbf{1} & \mathbf{O}_{N \times N_c} & \mathcal{B}_{N \times N_c} \end{bmatrix} \quad (15)$$

where C is a $2N \times (4 + 2N_c)$ matrix, \mathcal{B} is a $N \times N_c$ interpolation B-spline matrix, \mathbf{O} is a null matrix with the appropriate dimensions, \mathbf{x}, \mathbf{y} are $N \times 1$ vectors with the coordinates of the reference shape (see [5]), $\mathbf{1}, \mathbf{0}$ are $N \times 1$ vectors.

We have used two models to track the boundary of the left ventricle with dynamic matrices

$$A_m = \text{diag}(l_m), \quad m = 1, 2 \quad (16)$$

where $l_m = [\Delta_m \ \mathbf{1}]$ is a $(4 + 2N_c) \times 1$ vector and $\mathbf{1}$ is a $(3 + 2N_c) \times 1$ vector with $\Delta_1 = 0.9$, $\Delta_2 = 1.1$. These two models perform an expansion and a contraction of the reference

shape appropriate to represent the heart motion during the systole and diastole.

Figs. 3, 4 display tracking results obtained with the Robust Multi Model Tracker (RMMT) during a cardiac cycle.

The cardiac cycle comprises two different phases: diastole (relaxation) and systole (contraction). In the diastole the ventricle relaxes and the mitral valve is open. In the systole phase the ventricle contracts, and the mitral valve closes. The Robust Multi Model tracker must choose the dynamic model which describes the heart motion best at each instant of time. Fig. 3 shows the tracking results in the systole phase. During the systole the RMMT chooses the contraction model as expected (model 1). This phase is initiated by the peak of the ECG signal (QRS complex), which represents ventricular depolarization and ends after the occurrence of the T wave (see the right bottom figure in 3) which represents the ventricular repolarization. Fig. 4 corresponds to the relaxation phase (diastole). In the diastole the RMMT chooses the expansion model (model 2). This phase starts after the occurrence of the T wave (left top figure) until a peak of the ECG is observed (right bottom figure).

Fig. 5 shows the label estimate obtained with the RMMT during 3 cardiac cycles (bottom) and compares these estimates with ground truth (top). A perfect classification of the motion dynamics is obtained. During the systole phase, the contraction model is selected while in the diastole phase, the expansion model is chosen. These results were obtained with the probability transitions $p_{ii} = 0.8, i = 1, 2$ in matrix T .

The robust multi model tracker provides two levels of information: the output of the first level is an estimate of the heart boundary. The output of the second level is a binary signal which discriminates the two phases of the cardiac cycle. This type of information can not be obtained if a single model is used.

5. CONCLUSIONS

This paper presents a robust algorithm for tracking the left ventricle in a sequence of ultrasound images, based on switched dynamic models. The evolution of the state vector is described by a bank of two stochastic state equations, switched according to a Markov process. Furthermore, it is assumed that the visual features detected in the image contain outliers, i.e., invalid features which do not belong to the object boundary and prevent an accurate tracking of the ventricle contour. A robust filtering algorithm is proposed which is able to deal with multiple dynamics and invalid observations. This is accomplished by propagating the *a posteriori* density of the unknown parameters using Gaussian mixtures. Exact inference is possible if we allow the number of components to grow to infinity. To prevent this difficulty, pruning techniques are adopted. Experimental results are presented in the paper to assess the performance of the algorithm with real data. It is shown that the proposed method efficiently estimates the best active model, even in the presence of noisy measurements and outliers. The proposed algorithm also provides accurate estimates of the ventricle boundary. The proposed tracker outputs two levels of information: shape estimates and a binary classification of the contour motion, thus providing valuable information concerning anatomic and functional aspects. This is an important advantage, obtained by using the multiple models.

Acknowledgement: We thank Prof. Fausto Pinto from the Instituto Cardio Vascular de Lisboa for kindly providing the echographic data used in this work.

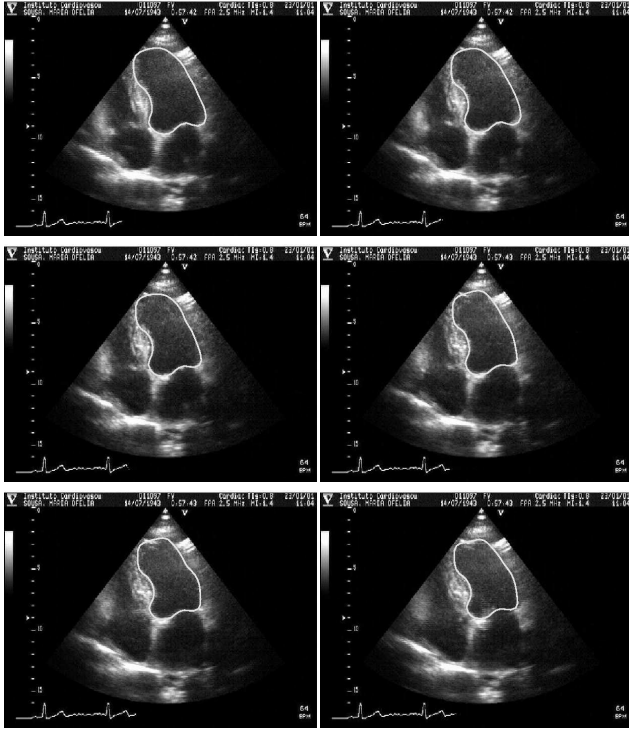


Fig. 3. Tracking in the second cardiac cycle (systolic phase) with Robust Multi Model: active contraction model, frames 35, 36, (first row) 37, 38 (second row), 39, 40 (third row).

6. REFERENCES

- [1] J. M. Dias and J. Leitão, "Wall position and thickness estimation from sequences of echocardiograms images," *IEEE Trans. on Medical Imaging*, vol. 15, no. 1, pp. 25–38, February 1996.
- [2] M. Figueiredo and J. Leitão, "Bayesian estimation of ventricular contours in angiographic images," *IEEE Trans. on Medical Imaging*, vol. 11, no. 3, September 1992.
- [3] Xenophon Papademetris, Albert J. Sinusas, R. Todd Constable Donald P. Dione, and James S. Duncan, "Estimation of 3-d left ventricular deformation from medical images using biomechanical models," *IEEE Trans. on Medical Imaging*, vol. 21, no. 7, pp. 786–799, July 2002.
- [4] Y. Bar-Shalom and T. Fortmann, *Tracking and Data Association*, Academic Press, 1988.
- [5] A. Blake and M. Isard, *Active Contours*, Springer, 1998.
- [6] J. Nascimento and J. S. Marques, "Robust shape tracking in the presence of cluttered background," *Proc. IEEE Int. Conf. on Image Processing*, vol. 3, pp. 82–85, September 2000.
- [7] J. Nascimento and J. S. Marques, "Improving the robustness of parametric shape tracking with switched multiple models," *Pattern Recognition*, vol. 35, no. 12, pp. 2711–2718, December 2002.
- [8] J. S. Marques and J. M. Lemos, "Optimal and suboptimal shape tracking based on switched dynamic models," *Image and Vision Computing*, pp. 539–550, June 2001.

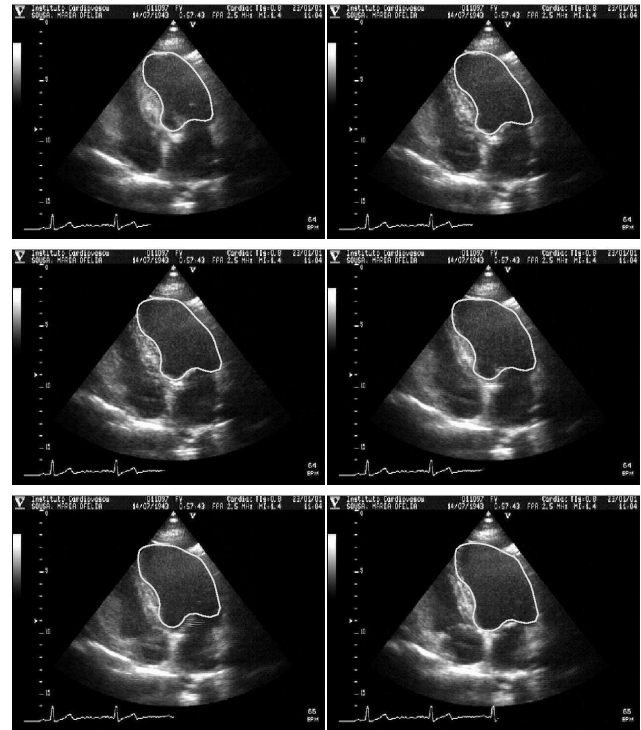


Fig. 4. Tracking in the second cardiac cycle (diastolic phase) with Robust Multi Model: active expansion model, frames 42, 44, (first row) 46, 48 (second row), 50, 52 (third row).

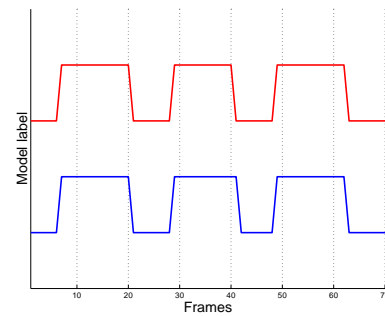


Fig. 5. Correct label sequence (top) and estimated label sequence obtained using the RMMT (bottom).